AGRICULTURAL DECISION ANALYSIS

JOCK R. ANDERSON, JOHN L.

THE IOWA STATE UNIVERSITY

AGRICULTURAL DECISION ANALYSIS

ILLON, and J. BRIAN HARDAKER

RESS / AMES, IOWA, U.S.A.

Jock R. Anderson is Associate Professor of Agricultural Economics and Business Management, University of New England, Armidale, Australia and holds the M.Agr.Sci. degree from the University of Queensland and the Ph.D. degree from the University of New England.

John L. Dillon is Professor of Farm Management, University of New England, Armidale, Australia and holds the Ph.D. degree in Agricultural Economics from Iowa State University.

Brian Hardaker is Senior Lecturer in Agricultural Economics and Business Management, University of New England, Armidale, Australia and holds the M.Sc. degree from the University of London and the Ph.D. degree from the University of New England.

Composed and printed by Science Press, 300 W. Chestnut Street, Ephrata, Pa. 17522

First edition, 1977

Library of Congress Cataloging in Publication Data

Anderson, Jock R 1941–
Agricultural decision analysis.

Includes bibliographies and index.
1. Farm management—Decision making—Statistical methods. 2. Bayesian statistical decision theory.
I. Dillon, John L., joint author. II. Hardaker, J. B., joint author. III. Title.
S561.A559 658.4'03 76-46456
ISBN 0-8138-0400-4

DEDICATED TO THE MEMORY OF

GERALD W. DEAN

CONTENTS

Sections of the text marked with Skippy are relatively difficult and can be skipped without loss of continuity.

PREFACE

Our concern in this book is with the making of decisions—it is to be hoped good ones—in a managerial context. We are all continually confronted by decision making. Decisions lie at the heart of society in all its aspects. Whether in politics or government, in farming or banking, in making cars or catching whales, in scientific research or medical diagnosis, or in our personal lives, decisions are unavoidable. Like it or not, decision making is part of the human burden.

The concepts and procedures we discuss are relevant to all decision making. As yet, though, they have not reached high operational pitch in some decision areas, largely because decision making has only come under theoretical scrutiny by economists, psychologists, statisticians, and management scientists in the last few decades. However, the development of decision analysis for agricultural management (particularly farm and business) has been relatively rapid and successful. In a variety of ways the authors have played some part in these developments, so it seemed appropriate that we should produce a text oriented to managerial applications in agriculture.

We will be concerned exclusively with risky choice, i.e., the problem that prevails when a decision maker must choose between alternatives, some or all of which have consequences that are not certain. We believe the study of risky decision making is most important since most managerial decisions involve elements of uncertainty—a fact that is not yet adequately reflected in economics training generally and management training in particular.

The approach to risky choice that we follow is a conditionally normative and logical one, often termed Bernoullian or Bayesian decision theory but more usually just decision theory or decision analysis. Given the decision maker's goal, this approach indicates which alternative he ought to take. It is based on the decision maker's *personal* strengths of belief about the occurrence of uncertain events and his *personal* evaluation of potential consequences. The emphasis in this definition on personal aspects should be noted. It is this which some critics of the approach have bemoaned, since they presumably desire either universal agreement on beliefs or universal answers to decision problems. We leave it to you to make your own judgment on the value and validity of the personal approach. Our evangelism attests to our conviction.

We do not believe that all the theoretical niceties discussed are relevant to all real-world agricultural decisions. How much effort should be put into any particular decision depends on the time available, the cost of analysis, and the importance of the decision. We do strongly believe that familiarity with the decision analysis approach—even at its simplest level of merely asking, What choices? What consequences? What chances?—will certainly lead to better agricultural management. This will be true regardless of whether the decision context is the farm itself; agribusiness or commerce; or government policy, research, or extension.

We have discussed the material in this book with many people. In particular, for their stimulus we must thank our students at the University of New England and the seminar groups at the Indian Agricultural Research Institute, New Delhi, and at Gödöllö University, Budapest, with whom we tested much of the presentation. Ross Drynan, John Kennedy, Chad Perry, and Javier Troncoso were kind enough to give us comments on an earlier draft. We are also indebted to Jean Mitchell for a careful and cheerful job of typing.

Jock R. Anderson

John L. Dillon

J. Brian Hardaker

AGRICULTURAL DECISION ANALYSIS

CHAPTER ONE
INTRODUCTION TO DECISION ANALYSIS

THE HUMAN ESTATE embodies the privileges and responsibilities of decision making, for it is the capacity for rational choice that principally distinguishes man from the animals. We are all continually confronted by decisions, but fortunately we can resolve most of them with a minimum of deliberation and anxiety. A decision problem exists only when we feel that the possible consequences are important and yet are unsure of what is the best thing to do. When a person is uncertain about the consequences of his decision, we can say he faces a risky choice.

Risky choice is inherently difficult to rationalize, but procedures have been developed to allow the process to be systematized. These procedures are collectively known as decision analysis. It is our aim to show how decision analysis can be used to lead to better decisions in agriculture.

Two potential misconceptions about our treatment should be immediately dismissed. First, a good risky decision does not guarantee a good outcome; rather, it is one consistent with the decision maker's beliefs about the risks surrounding the decision and with his preferences for the possible outcomes. A good decision is a considered choice based on a rational interpretation of the available information. Whether such a decision turns out right or wrong is partly a matter of luck and in any case can never be determined until after the event, and often not even then. Second, the normative procedures we will present serve to *aid* decision makers through the complexities of their decision problems. The role of the decision maker himself is in no way denigrated; as we shall see, his beliefs and preferences are of absolutely paramount importance.

As with any scientific development, decision analysis has led to the definition of some clarifying concepts and some jargon. We will introduce the main concepts and jargon of decision analysis somewhat formally by listing the components of decision problems, and then less formally by way of a simple example.

1.1 COMPONENTS OF DECISION PROBLEMS

Decision problems in general are characterized by several possible components, although particular decision problems may not feature all those listed here. The abstract presentation in this section tends to disguise the subjective and artistic aspects of decision analysis. Later we will attempt to remove this disguise.

A decision problem is like a chain insofar as it makes little sense to single out particular components or links as the most important. All are important if analysis is to be successful. However, we can say that unless the first two components (acts and states) are appropriately specified, any finesse in specifying others may well be wasted.

Acts. The pertinent acts (or actions) available to the decision maker, between which he must choose, are denoted here by $a_1, a_2, \ldots, a_j, \ldots$. These acts must be defined to be mutually exclusive and should also be exhaustive for the job to be well done. Obviously, a decision maker can only be as good as the decisions he considers, so good decision analysis must be based on skilful definition of the acts. One "act" that crops up in many decision problems (and perhaps should appear in many more) is "do nothing" or "defer action." Decision problems featuring an intrinsically continuous decision variable, such as fertilizer rate, may sometimes require specification of an infinite set of acts but typically can be represented approximately but adequately by a small finite set of discrete acts.

States. The possible events or states (states of nature or states of the decision maker's world), denoted here by $\theta_1, \theta_2, \ldots, \theta_i, \ldots$, must also be defined by a mutually exclusive and exhaustive listing. If the decision maker does not know for certain which state will prevail, the decision problem is said to be risky. Some state variables are intrinsically continuous (e.g., rainfall), but often a discrete representation of such variables (such as "good," "average," or "poor" for rainfall) proves adequate. Skill, experience, and judgment are all important in specifying states in optimal detail. States may be of simple or compound description. For example, a state of nature might be defined in terms of rainfall during growing season, rainfall at flowering, and prices after harvest to account adequately for the several elements of uncertainty impinging on a decision. Sometimes the specification of states will hinge on decisions made by one or more opponent decision makers. With the exception of international trade and some other marketing decisions, such situations of conflict are relatively rare in agricultural management. Except for brief reference in Chapter 5, we will not consider decisions involving conflicting parties.

Prior probabilities. We believe that probabilities can always be attached to the occurrence of states in managerial decision problems. Such probabilities are subjective or personal in nature, if only because they are judged relevant to personal decision making. Later (Chapter 2) we will present methods for specifying and eliciting the decision maker's probabilities or degrees of belief. We will follow the convention of denoting the prior probability of the ith state or event by $P(\Theta_i)$. By the prior probability of a state we mean the decision maker's initial judgment about the state's chance of occurrence.

Consequences. Depending on which state occurs, choice of an act leads to some particular consequence, outcome, or payoff. Measuring the consequences of act and state pairs involves a variety of difficulties if we take the (correct) view that consequences should be assessed in such a manner that all aspects of the decision maker's preferences are captured and correctly balanced. When this is done, the consequences are said to be measured in terms of utility. Useful methods for assessing utility are reviewed later (Chapter 4). For the moment we will denote the utility resulting from the ith state when we have taken the jth act by $U(a_j \mid \Theta_i)$.

Choice criterion. Determination of an objective function is necessarily closely tied in with the measurement of consequences. We will be exclusively concerned with the criterion of maximizing expected utility (preference). Discussion and defense of this choice criterion is deferred until Chapter 4.

Experiments. Many decision problems feature the possibility of conducting one or more "experiments" (e.g., running a survey, reading a predictive device, buying a forecast, consulting an expert, etc.). An experiment results in predictions that we will denote $z_1, z_2, \ldots, z_k, \ldots$. The impact of an experiment in decision analysis is through the additional information it gives about the probabilities of the states. The vehicles for transmitting this information are conditional probabilities that are usually termed likelihoods.

Likelihood probabilities. These are probabilities that pertain to a specific experiment and have the conditional form "the probability of observing prediction z_k given that a particular state, the ith, prevails." Accordingly, they are denoted as $P(z_k \mid \Theta_i)$. Sometimes likelihoods are unambiguously determined by the nature of a sampling process (e.g., as shown in Chapter 3 for binomial sampling), but often in managerial applications they too must be based on personal or subjective judgment. When an experiment is in the offing, the initial probabilities held about the occurrence of states are designated as prior probabilities (i.e., they are the relevant probabilities prior to conducting the experiment). As we will see in Chapter 3, these prior probabilities are modified through the likelihoods and Bayes' theorem to become *posterior probabilities* (i.e., probabilities relevant after or posterior to observing a prediction). Note that although the prior probability $P(\Theta_i)$ is designated as being unconditional, this is a false impression since all probabilities are really conditional—priors being conditional on all the knowledge, experience, and judgment that has gone into their specification. Indeed, it is not unusual for the posterior probabilities in one problem to be the priors in the next.

Strategies. A strategy in the present context is simply a recipe for future action conditional on observed experimental outcomes. In this sense, a strategy is a rule or schedule that specifies in advance for each possible future information signal (prediction or experimental outcome) the action that will be taken in response to it. For instance, the *t*th strategy s_t might be defined as taking the *j*th act when the *k*th prediction is observed.

Summarizing the above, we have outlined the following components of a risky decision problem:

a_j = the *j*th *act* or risky prospect
Θ_i = the *i*th *state* of nature
$P(\Theta_i)$ = the *prior probability* of occurrence of Θ_i
$U(a_j \mid \Theta_i)$ = the *utility* that results if a_j is chosen and Θ_i occurs
z_k = the *k*th possible *forecast* from an experiment
$P(z_k \mid \Theta_i)$ = the *likelihood probability* of z_k occurring given that Θ_i prevails
s_t = the *t*th *strategy,* implying choice of some particular act if some particular experimental outcome occurs

Now that we have listed the bare bones of a generalized decision problem, we will bring the components together more concretely through an example of such simplicity that the overall structure is at all times quite obvious. Real-world complexities will not trouble us for the moment. The mode of analysis employed is known as the "normal form" of analysis, but (as we shall see in Chapter 5) it is not the most efficient mode nor the one usually followed in solving decision problems.

1.2 A SIMPLISTIC EXAMPLE

Suppose a farmer has just harvested his grain crop and his immediate decision problem relates to the disposition of his marketable surplus, i.e., sale now or temporary storage (at a cost) in the hope of a higher price later. For expository simplicity, assume that the acts he identifies (perhaps with our help) are a_1 = "sell now" and a_2 = "store and speculate;" and the states are Θ_1 = "market normal" and Θ_2 = "market in short supply." The farmer's subjective probabilities are assumed to be 0.8 and 0.2 for states Θ_1 and Θ_2 respectively.

Consequences are assumed to have been appropriately budgeted by the farmer and scaled to account for his preferences, so they can now be referred to as utilities. Suppose the relevant decision matrix is as shown in Table 1.1.

TABLE 1.1. Relevant Decision Matrix

θ_i	$P(\theta_i)$	a_1	a_2
θ_1	0.8	100	90
θ_2	0.2	100	150
Expected utility		100	102*

*Value associated with the prior optimal act.

The expected utility of each act is now calculated in the ordinary way as the probability weighted average of the utilities associated with each state for that act. Thus the expected utility of a_2 is $(90)(0.8) + (150)(0.2) = 102$, which is greater than the expected utility of 100 for a_1. The expected utility of a_2 is marked with an asterisk to denote that it is the value associated with the *prior optimal act*.

Observe that this small example is truly a decision problem because neither act is clearly superior to the other. An act a_3 with utilities (95, 95) would never be preferred to a_1 with utilities (100, 100). Similarly, an act a_4 with utilities (90, 149) is inferior to a_2. In technical parlance, a_2 is said to dominate a_4 stochastically. Dominated acts such as a_3 and a_4 should be excluded from decision analysis as soon as they are apparent. The concept of dominance is discussed further below in connection with the evaluation of strategies, and in rather more detail in Chapter 9.

Now suppose our farmer had access to a perfect predictor of which of the uncertain states would prevail. In this case he would sell now with the best possible payoff of 100 if the predictor said that θ_1 would eventuate. If the perfect forecast was for θ_2, he would adopt a_2 with a payoff of 150. The expected value of his decision problem under such perfect information is thus $(100)(0.8) + (150)(0.2) = 110$. The amount by which this exceeds the value of the prior optimal act is called (for obvious reasons) the *expected value of perfect information* (EVPI). Thus EVPI $= 110 - 102 = 8$ units of utility, and this measures the degree of uncertainty faced by the decision maker.

Suppose our decision maker can obtain some further information about the out-of-season grain market from a market research agency that is prepared to forecast either that "the market will be normal" (z_1) or that "the market will be in short supply" (z_2) this year. Our man has had some past (not always happy) experience with the research agency and judges that if in fact the market is normal, there is a 60% chance the market will be forecast correctly. The agency's record is slightly better for abnormal forecasts, and his degree of belief in a correct forecast of short supply is 0.7. These subjective conditional probabilities can now be formalized as in Table 1.2. Note that the likelihood probabilities sum to one *across* the table since they are specified for each of the possible forecasts conditional on a particular state prevailing.

TABLE 1.2. Likelihood Probabilities

	$P(z_k \mid \theta_i)$	
θ_i	z_1	z_2
θ_1	0.6	0.4
θ_2	0.3	0.7

Armed now with the prospect of these predictions, the strategies open to our farmer can be spelled out. A strategy merely prescribes what acts are to be taken conditional on each possible prediction. In general, the number of strategies available will be the number of acts raised to the power of the number of predictions or experimental outcomes, so in this case there are four possible strategies. These are defined in the left columns of Table 1.3 where strategy s_4, for example, implies that a_2 will be chosen if the prediction is z_1 and a_1 will be chosen if it is z_2.

TABLE 1.3. Evaluation of Strategies in the Decision Problem Example

Strategy	Act Adopted If:		Conditional Expectation $U(s_t \mid \theta_i)$		Unconditional Expectation
s_t	z_1	z_2	θ_1	θ_2	$U(s_t)$
s_1	a_1	a_1	100	100	100
s_2	a_2	a_2	90	150	102
s_3	a_1	a_2	96	135	103.8*
s_4	a_2	a_1	94	115	98.2

*Optimal strategy.

In the so-called normal form of analysis, strategies are evaluated in two steps. The first of these is an evaluation conditional on each state actually occurring. For example, let us consider the conditional evaluation of strategy s_4. On condition that θ_1 occurs, the utility of this strategy is $U(a_2 \mid \theta_1)$ if z_1 is the forecast and $U(a_1 \mid \theta_1)$ otherwise. But the probabilities of observing z_1 and z_2 (also conditional on θ_1 occurring) are the likelihoods $P(z_1 \mid \theta_1)$ and $P(z_2 \mid \theta_1)$ respectively. Consequently, the expected utility of s_4, conditional on Θ_1, is given by

$$U(s_4 \mid \Theta_1) = U(a_2 \mid \Theta_1) P(z_1 \mid \Theta_1) + U(a_1 \mid \Theta_1) P(z_2 \mid \Theta_1)$$
$$= (90)(0.6) + (100)(0.4) = 94$$

You can check your understanding by calculating $U(s_3 \mid \Theta_2)$. All the conditional expectations are reported in the central columns of Table 1.3.

The final arithmetic step in evaluation is to bring our knowledge of the prior probabilities of the states to bear on the analysis. Weighting of the conditional expectations by their respective priors leads to (unconditional) expected utilities for each strategy. Thus the expected utility of s_t is given by

$$U(s_t) = U(s_t \mid \Theta_1) P(\Theta_1) + U(s_t \mid \Theta_2) P(\Theta_2) \tag{1.1}$$

For example, $U(s_4) = (94)(0.8) + (115)(0.2) = 98.2$. All unconditional expectations are reported in the right side of Table 1.3 where it can be seen that the largest is 103.8 for s_3. This search thus reveals that the third strategy is the one that maximizes expected utility. This is often referred to as *"the Bayes strategy."*

Before leaving the conditional expectations of Table 1.3, we should observe that strategies can be checked for dominance in an analogous manner to that noted for acts. One strategy (s_α) is said to dominate another (s_γ) if for all states $U(s_\alpha \mid \Theta_i) \geqq U(s_\gamma \mid \Theta_i)$ and the strict inequality ($>$) holds for at least one state. Applying this rule to the present example indicates that the only case of dominance is that of s_3 over s_4. Thus s_4 could have been dismissed from further consideration before taking the final expectations.

The value of the Bayes or optimal strategy, 103.8, can be compared fruitfully with the value of the prior optimal act, 102. It can never be less if information is costless, and in this case it exceeds the prior value by 1.8 utility units. If the forecast is free, the implication is that it is certainly worth having; the farmer should therefore avail himself of the (as yet uncertain) forecast before reaching his decision about disposal of his crop. If, however, he must pay for the forecast, its cost must be formally accounted for in the analysis before we can say whether it would be worth purchasing. Techniques for this type of decision making, where predictions are not costless, are elaborated in Chapter 5.

The normal form of analysis used in our simple example can also be illustrated graphically. This can be done easily and in two dimensions because there are only two states. The axes of Figure 1.1 measure conditional utility expectations, and the conditional evaluations of strategies previously computed are plotted as the points s_1, s_2, s_3, and s_4. In graphical

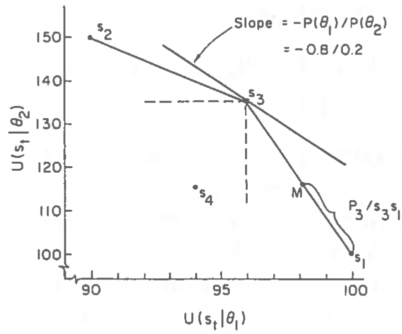


FIG. 1.1. Graphical representation of strategy-conditional expectations in a two-state decision problem.

terms, dominance for the two-state case is indicated by quadrants. A strategy dominates all others that lie to its southwest (i.e., within its southwest quadrant), including those on the quadrant boundaries.

Conversely, a strategy will be dominated by all others in its northeast quadrant. Hence it is apparent in Figure 1.1 that s_3 dominates s_4. Strategies that are not dominated are called *admissible strategies* of which s_1, s_2, and s_3 are examples here. These strategies are also *pure strategies* in distinction to *mixed strategies* formed as probabilistic mixtures of two or more pure strategies. For example, a particular mixed strategy would be to adopt s_3 if a fair coin shows heads and s_1 if it shows tails. Intuition suggests and some simple algebra would confirm that this strategy can be represented through its conditional expectations as the midpoint of a straight line joining s_3 and s_1. Indeed, a straight line joining two pure strategy coordinates defines the characteristics of all their mixed strategies. A mixed strategy M based on choosing s_3 with probability P_3 and choosing s_1 with probability $1 - P_3$ is represented graphically as the point dividing the line $s_1 s_3$ in the proportions $P_3{:}1 - P_3$, as indicated in Figure 1.1.

Applying the test of dominance to mixed strategies leads to the conclusion that if mixed strategies are admissible, all strategies to the southwest of straight lines joining s_2 to s_3 and s_3 to s_1 are dominated. Then the only strategies of potential interest to the decision maker are those actually on these lines, which together are called the *admissible boundary*. Which of the admissible strategies is best depends on the probabilities of the states.

Recall that in locating the utility-maximizing strategy, we first found the unconditional expectation through equation (1.1). If we rearrange (1.1) as

$$U(s_t \mid \Theta_2) = U(s_t)/P(\Theta_2) - [P(\Theta_1)/P(\Theta_2)]\, U(s_t \mid \Theta_1)$$

we have it in terms of the axes of Figure 1.1, in the form of an equation for a straight line of slope $-P(\Theta_1)/P(\Theta_2)$ with intercept depending on $U(s_i)$. The Bayes strategy is thus found graphically by finding the member of the family of straight lines of this slope tangential to the admissible boundary and thus having the largest possible intercept. As would be expected from our previous arithmetic analysis, this line of slope $-0.8/0.2$ passes through s_3. Note that any admissible strategy will be a Bayes strategy for some feasible set of probabilities.

Graphical representation provides clear demonstration of the influence of different prior probability judgments on the selection of an optimal strategy (and acts if it is observed that s_1 is simply a_1 and s_2 is a_2). As $P(\Theta_1)$ increases, the slope of the expected utility line increases even more rapidly until s_3 and s_1 are simultaneously optimal, when in fact, from the geometry of Figure 1.1 and the knowledge that $P(\Theta_1) + P(\Theta_2) = 1$, we have

$$P(\Theta_1) = \frac{U(s_3 \mid \Theta_2) - U(s_1 \mid \Theta_2)}{U(s_3 \mid \Theta_2) - U(s_1 \mid \Theta_2) - [U(s_3 \mid \Theta_1) - U(s_1 \mid \Theta_1)]}$$
$$= (135 - 100)/[(135 - 100) - (96 - 100)] = 35/39$$

When two pure strategies are simultaneously optimal, so are all their mixed strategies. But since all have the same expected utility, we can say that the Bayes strategy can always be a pure strategy. In the present example s_1 will be the Bayes strategy for all $P(\Theta_1)$ greater than 35/39, and (following analogous reasoning) s_2 will be optimal for all $P(\Theta_1)$ less than 5/7.

Apart from expository value the main virtue of this normal form of analysis (which leaves the use of prior probabilities until the last step) is the ready indication it gives of the sensitivity of the optimal strategy to prior probability judgments. By their nature these judgments may be somewhat clouded. Conversely (though not well illustrated in our simple example), the cost of the normal form of analysis is the volume of arithmetic required in assessing all strategies. Methods to be presented in Chapter 5 are more efficient for practical decision analysis.

1.3 SUMMARY

Decision analysis is a logical procedure for making risky choices. It is a mechanism for bringing together all the pertinent aspects of a decision environment. Most important, it fully recognizes the personal element in decision making—personal beliefs about the risks involved and personal preferences for possible consequences.

For any particular decision maker and any particular decision probem, decision analysis involves (1) defining relevant acts and states and heir consequences, (2) eliciting prior degrees of belief or probabilities for ;tates and degrees of preference or utility for consequences, (3) taking ac- :ount of whatever further predictive information may be available as a ɔasis for revising the initial probabilities, and (4) selecting the optimal strategy on the basis of maximizing expected utility.

These steps amount to no more than spelling out the processes followed by managers in making risky choices, processes that are usually attempted in intuitive fashion. However, many risky decisions are too complex and important to be handled satisfactorily by intuition. Decision analysis, by its formal procedure, enables a manager better to ensure that his risky choices are in line with his preferences and beliefs and that full value is extracted from all the information available to him.

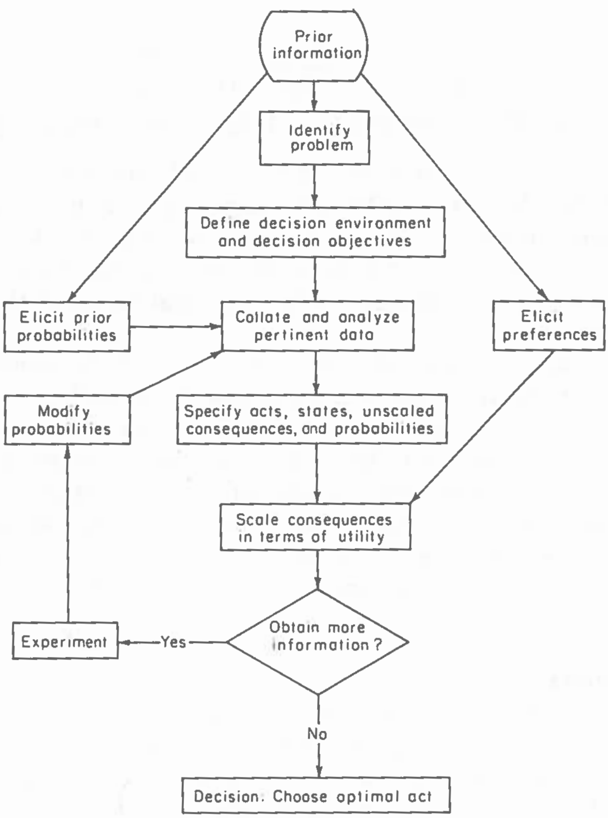


FIG. 1.2. Overview of the typical structure of decision analysis.

Figure 1.2 gives a broad overview of the typical structure of decision analysis. The flow chart shows the usual chronological ordering of procedures but is not inviolate. The stage labeled "obtain more information?" involves evaluation of strategies and will be the subject of detailed attention in Chapter 5.

A cycle of decision analysis terminates when an optimal (best bet) decision is found. The process does not cease here since decision makers must go on to implement the indicated act, to discover the realized consequences, and to bear responsibility for these outcomes through a further cycle of decision making. Whether decision makers *should* take the excursion through a formal analysis in the course of any particular decision is a good question for which, unfortunately, there can be no general and correct answer. However, the more important the decision, the more likely it is that decision analysis could profitably be used.

PROBLEMS

1.1. What is the probability that:
- (a) Tossing the first coin you draw from your pocket will result in a head?
- (b) Having tossed a coin nine times and getting nine heads, the next toss will be a tail?
- (c) It rained in London on April 21, 1931?
- (d) It will rain here tomorrow?
- (e) You will write more than 30 letters next year?
- (f) You will not be working for the government in five years time?
- (g) You will visit Russia in the next 10 years?
- (h) You will write less than 31 letters next year?
- (i) You will write no letters next year?
- (j) You will die next year?

1.2. Comment on the following statements:
- (a) The only real decisions are those involving uncertainty.
- (b) Probabilities have no other purpose than to assist in decision making.
- (c) If any two rational managers face the same decision problem, they should make the same decision.
- (d) Good outcomes imply good decisions.
- (e) Only luck can prevent regrets about good decisions.
- (f) For any decision, the acts, states, and payoffs can always be specified by the decision maker.

1.3. What relationships exists between pure strategies, mixed strategies, Bayes strategies, and admissible strategies?

1.4. Solve the example of Section 1.2:
- (a) Under the assumption that $P(\Theta_1) = 0.6$ and $P(\Theta_2) = 0.4$.
- (b) Under the assumption that $P(z_1 \mid \Theta_1) = 0.7$ and $P(z_1 \mid \Theta_2) = 0.4$.

Depict your solutions graphically and compare with Figure 1.1.

1.5. (a) List some of the predictive devices available to farm managers. How might the likelihoods of these devices be established? Should these likelihoods be the same for different farm managers?

(b) Suppose a new variety of some crop has just become available. Specify the components that make up the decision problem confronting a farmer who wonders whether to use the new variety.

1.6. Woles Food Stores has the opportunity of purchasing a shipment of 100,000 avocados. Estimates are of a profit of 4 cents per avocado sold and of a loss of 22 cents per overripe avocado. Records indicate that in previous purchases the frequency of overripe fruit has been as follows:

Overripe	Frequency
(%)	
13	15
14	36
15	29
16	12
17	6
18	2

(a) Should the shipment be purchased?

(b) Suppose a prediction costing $60 is available that accurately forecasts whether there is more than 15% of overripe fruit but does not discriminate any further. Should this prediction be purchased?

1.7. Specify and solve a not too unrealistic decision problem involving two states and three undominated acts, assuming that a predictive device is available. Calculate all the various quantities we have mentioned in this chapter.

1.8. Ned Kelly of "Rawdeal" via Wallan Wallan is considering how many months of fodder reserve to keep on hand for his cattle. Using relevant cost data and his subjective probability estimates for various lengths of drought, Ned has calculated the expected cost and its 70% range for various levels of fodder reserve. These are as follows:

Size of Reserves	Expected Cost	70% Range
(mos)	($)	(±$)
0	170	100
1	146	98
2	124	92
3	108	84
4	98	73
5	94	63
6	93	54
7	95	46
8	97	42
9	100	38
10	104	37
11	108	37
12	112	37

With expected cost on one axis and 70% range on the other, draw a graph showing the various possible fodder reserve policies. Can Ned's possible choices be classified in any way? What can be said of policies in the range from 6 to 10 months?

1.9. O. N. Theball of Nogobung Holdings Ltd. has decided to purchase a bull from Gundy Pastoral Company. The alternatives are to buy a stud bull at $1500 or a grade bull at $300, both unproven. Taking account of the fact that the bull may be infertile, Theball judges that the chances of each bull throwing various classes of progeny are as follows:

Sale Price	Probability Judgment	
of Progeny	Stud bull	Grade bull
($/hd)		
120	0.07	0.02
110	0.55	0.21
100	0.20	0.48
90	0.15	0.24

If fertile, a bull is culled after producing 200 progeny. If infertile, the bull is culled immediately. In either case, Theball estimates a salvage value of $200. Progeny are sold as yearlings, and their production cost apart from depreciation on the bull is $80 per head.

(a) Given that Theball is a man for whom a dollar is a dollar is a dollar, which bull should he buy?

(b) What if Gundy Pastoral offers a refund of $1000 on the purchase of the stud bull if it should prove infertile?

(c) Now suppose an infallible prepurchase test of fertility is also available at a cost of $20. What is the expected value of using this test? Should the test be used? Which type of bull should be purchased?

SELECTED FURTHER READING

The reference material available in the general field of decision analysis is diverse in mathematical difficulty and subject matter application. The following are chosen as being in the same vein as the present treatment.

Halter, A. N., and G. W. Dean. 1971. *Decisions under Uncertainty with Research Applications*. Cincinnati: South-Western.

Comprehensive, with some advanced treatments and a specifically agricultural orientation.

Lindley, D. V. 1971. *Making Decisions*. London: Wiley.

Very much a beginner's introduction, especially to probability concepts.

Menges, G. 1974. *Economic Decision Making: Basic Concepts and Models*. London: Longman.

Of interest for its coverage of both Bayesian and non-Bayesian approaches to decision analysis.

Raiffa, H. 1968. *Decision Analysis*. Reading, Mass.: Addison-Wesley.

Highly recommended for its coverage and brilliant exposition but unfortunately employs some nonstandard notation.

Schlaifer, R. 1969. *Analysis of Decisions under Uncertainty*. New York: McGraw-Hill.

Lengthy, detailed, practical treatment in a didactic style. Orientation is very strongly to business management. One for the affluent student.

Winkler, R. L. 1972. *An Introduction to Bayesian Inference and Decision*. New York: Holt, Rinehart and Winston.

An outstanding text in the area of statistics for business decisions. Includes a bibliography of some 450 items with excellent chapter-by-chapter directions for further reading.

We will make no attempt to present a comprehensive review of the literature concerning decision analysis. Bibliographical sources with a specifically agricultural flavor include:

Anderson, J. R., and J. B. Hardaker. 1972. An appreciation of decision analysis in management. *Rev. Mktg. Agric. Econ.* 40(4): 170–84.

Lists 70 references.

Dillon, J. L. 1971. An expository review of Bernoullian decision theory. Is utility futility? *Rev. Mktg. Agric. Econ.* 39(1):3–80.

Lists a further 377 references of earlier vintage.

An introduction to the voluminous literature on decision making from a philosophical point of view is given by:

Mitroff, I. I., and F. Betz. 1972. Dialectical decision theory: A meta-theory of decision making. *Mgmt. Sci.* 19(1):11–24.

Weber, J. D. 1973. *Historical Aspects of the Bayesian Controversy* (Tucson: Div. Econ. Bus. Res., Univ. Arizona).

CHAPTER TWO
PROBABILITY

AT DIFFERENT TIMES and in different places man has had different concepts of beauty, justice, right, and truth. Probability, a somewhat analogous abstraction, has a similarly checkered history of evolving conception. We will take brief note of this history in the course of introducing the concept of subjective probability, which in our view is the only type of probability useful in decision analysis. Methods for eliciting subjective probabilities of various types are reviewed, and warnings are given to assist in avoiding several possible sources of bias.

2.1 SUBJECTIVE AND OTHER CONCEPTS OF PROBABILITY

Our brief sketch cannot do justice to the voluminous literature concerning alternative concepts of probability. Historically, the most important early concept viewed probabilities as *relative frequencies*. The modern axiomatic form of this concept of so-called *objective probabilities* views them as (infinite) limits of relative frequencies. A concept based on an infinite number of trials is clearly not operational in "finite" decision making. Moreover, when we consider that the states in decision problems can rarely be regarded as repeatable events (e.g., the seasons of this calendar year come only once), we see that any probability concept based on relative frequency is doomed to irrelevance for purposes of decision analysis. A historical record of rainfalls, for example, may have strong influence in formulating subjective probabilities (i.e., those not restricted to repeatable situations). However, we should recognize that use of historical distributions in decision analysis involves a strong subjective presumption that the historical structure is unchanged and is relevant to the specific planning period under review. Unqualified use of such "objective" relative frequencies as probabilities in decision analysis is a mechanical, a simplistic, and probably an inefficient and inaccurate procedure of specification.

The *logical concept* views probability as the logical relationship between a proposition and a body of evidence. The physical laws of aerodynamics and the properties of fair coins and fair tosses interact to give a logical probability of 0.5 for heads. Unfortunately, all the physical laws, properties, and interactions appropriate to defining the occurrence of the states of nature in real-world decision problems can never be known. Thus this concept must also fall by the wayside in searching for an operational notion for decision analysis.

We are then left with a *personal concept* of probability. The degree of belief or strength of conviction an individual has about a proposition (such as the occurrence of a state) is his *subjective probability* for it—with two important provisos. These are that such personal or psychological probabilities are consistent with the axioms, rules, and calculus of probabilities, and that they are consistent with the degrees of belief really held. We do not wander through life with our minds packed with numbers that we have identified as degrees of belief or subjective probabilities, but we certainly make decisions as if such numbers exist. To make our decision analysis explicit, we must determine these numbers. Because they must be formulated or judged, subjective probabilities are also known as judgmental probabilities.

We all undeniably have inner feelings of uncertainty. The process of translating these into degrees of belief or initial probabilities is beyond present understanding—even by psychologists specializing in this field. Decision analysts must thus be content to take degrees of belief as their starting point. The first problem arises in ensuring that a person's expressed subjective probabilities are reasonably consistent with his true degrees of belief. The methods of elicitation discussed in Section 2.3 have been designed to try in an informal way to keep consistency high. More formal methods employ the idea of scoring rules to keep assessors "honest." Since such rules have little applicability in practical decision analysis, discussion is deferred to Section 2.5.

Those familiar with the way scientists profess to think will not be surprised to learn that the notion of subjective probability has not been accepted without some strong arguments, and considerable controversy persists in some quarters. Some people have been unable to tolerate the sacrifice of "scientific objectivity" inherent in the personal approach. But "objectivity" in science is a myth, in life an impossibility, and in decision making an irrelevance. Its loss need not be regretted.

Subjective probability is the only valid concept for decision making, just as (purposes of communication aside) decision making is the only valid use for probabilities. The probabilities used in decision making ideally should be those of the person who bears responsibility for the decision, although this is certainly not to say that he should not avail himself of all pertinent information (such as communicating with experts, consulting historical records, making use of probability calculcus, etc.) in formulating his beliefs. Accordingly, an individual's degrees of belief will generally change as his experience and knowledge expand. Likewise, there is no reason why two individuals should not hold at any one time differing degrees of belief for the same states. We would anticipate, however, that their degrees of belief would correspond more closely as their store of common experience increases. Subjective probabilities cannot be "right" or "wrong," although a rational person would presumably wish to refine his probability judgments, eliminating as far as possible any biases arising from misconceptions or misinterpretations of the information available to him.

2.2 BIASES IN PROBABILITY JUDGMENT

Research has shown that people make probability judgments in much the same way as they make estimates of other quantities such as distance, i.e., by using certain perceptual clues. For example, distance is judged in part by the clarity with which an object can be seen. Near objects are normally clearer than distant objects. Clarity is therefore a useful perceptual clue in judging distance, but it can sometimes be misleading. In fog or mist, distances are often overestimated, sometimes with disastrous results; in clear conditions (e.g., in mountain air) it is easy to underestimate distance. If we are aware of these traps, we can try to allow for them. The same applies to possible biases in probability judgments. As Tversky and Kahneman (1975) have shown, the most important of such biases relate to the phenomena of *representativeness* and *anchoring*. We will look at both these sources of bias in turn.

Representativeness

Many probability judgments require an assessment of the chances that A is a member of the set B. For example, a farmer may need to evaluate the probability that a spell of dry weather is the start of a prolonged drought. Typically, such judgments are made by assessing the extent to which the object or occurrence under review is representative of the class to which it is to be related. So the farmer in our example might judge how representative the current dry spell is of the first few weeks of droughts he has experienced in the past. While representativeness is obviously a relevant clue in forming probability judgments, there is a danger of placing too much reliance on it to the neglect of other kinds of evidence. In this way, representativeness can lead to three important types of bias respectively involving the neglect of prior probabilities, misconceptions of chance, and disregard of predictive accuracy.

Neglect of prior probabilities. If A is regarded as highly representative of the set B, many decision makers allocate a high probability to A being a member of *B* regardless of prior evidence about the chances involved. Thus the farmer might assign a high probability to the possibility of a drought starting because "the present dry spell is just like the start of the last big drought," disregarding the fact that few spells of dry weather actually develop into long droughts.

Misconception of chance. People tend to expect a sequence of events generated by a random process to represent the characteristics of that process, even when the sequence is short. For instance, our misguided farmer might say that because (on an average) one year in five in his area is a drought year and because the last four years have been wet, there is bound to be a drought this year. In reality, although weather cycles do occur, their effect is small and the probability of rainfall is largely independent of weather observed in recent years and months.

Disregard of predictive accuracy. People often make predictions by selecting the outcome that is most representative of the evidence available to them. The confidence they have in their prediction depends primarily on the representativeness of the evidence, with too little regard commonly taken of its predictive accuracy. The very high prices sometimes paid for stud stock might in part result from overestimation of the reliability of certain characteristics of a sire or dam as predictors of progeny performance.

Anchoring

A second important source of bias (of particular importance in the context of the probability elicitation methods to be described below) is known as the anchoring effect. Most people find the introspective effort required to make probability judgments quite difficult. Consequently, once some particular value occurs to them (or is suggested by someone else), they tend to anchor on this value. Research in cases where judgmental probability distributions can be calibrated against the actual distributions of outcomes has shown that anchoring can lead to assessed distributions that are too "tight," i.e., with too small a variance. The explanation appears to be that people commonly first estimate the mean or modal value and then anchor on this, discounting the probabilities of values far from the anchor point. By contrast, if an assessor is encouraged to consider probabilities of extreme values first, the estimated distribution may be calibrated as too broad. Only further experience and research can show how these kinds of bias can be minimized.

In the processing of probability information, an important kind of bias due to anchoring has been called *conservatism*. It arises when decision makers are reluctant to revise their prior probability judgments in the light of additional evidence. As shown in Chapter 3, decision makers who wish to revise their beliefs in a manner consistent with the probability calculus can eliminate this bias by using Bayes' theorem.

It must be emphasized that subjective probabilities should first and foremost be consistent with the decision maker's beliefs and with the probability calculus. Any rational person, however, will also strive to achieve consistency in his whole network of degrees of belief. A careful comparison of interrelated beliefs will often reveal inconsistencies that, if not corrected, may well result in biased probability judgments.

2.3 ELICITATION OF SUBJECTIVE PROBABILITIES

Probability distributions describe the stochastic or probabilistic behavior of random variables. States of nature may depend upon one or more random variables. Assessment of the (subjective) probabilities of states thus usually requires knowledge of the decision maker's degrees of belief about the underlying random variables. A variety of methods is available. In reviewing those of most relevance to decision analysis, we will assume a working familiarity with probability calculus. Only particularly pertinent operations will be reviewed here. We should recall, however, that all probability calculus devolves from basic axioms: (1) probabilities cannot lie outside the range of zero to one, (2) the probability that two or more mutually exclusive events will occur is the sum of their respective probabilities, and (3) the probability of the exhaustive (universal) set of events is one. Thus if θ_i, i = 1, 2, ... n, denotes the ith mutually exclusive event of an exhaustive set, we must have

$$0 \leqq P(\theta_i) \leqq 1$$

$$P(\theta_i \text{ or } \theta_j) = P(\theta_i) + P(\theta_j)$$

$$\sum_{i=1}^{n} P(\theta_i) = 1$$

One Random Variable—The Discrete Case

The idea of gambling or participating in lotteries has become inextricably intertwined with schemes for determining probabilities. Something of the flavor of this can be sensed in Savage's (1954) interpretation of probability in terms of the price a person is willing to pay for a gamble. Suppose we offer participation in a lottery that pays a small sum of money M if the event "rain next week" holds true and pays nothing if it does not. If you are willing to pay just m units of money to participate in this lottery, your subjective probability for the event is m/M. This approach depends on preference for money being linear over the payoff range from zero to M. It is *not* a recommended method.

We could always simply ask a decision maker to state his degrees of belief directly; e.g., "What is your probability that it will rain next week?" Such a question may elicit a fast response that may or may not be a good measure of the true degree of belief. Unless perhaps it is the middle of the rainy reason or the dry season, a good correspondence between derived probability, degree of belief, and inner feelings of uncertainty is unlikely to arise without at least some careful contemplation and introspection, a requirement that is not effortless.

A conceptual device that has been used to aid consistent contemplation and judgment is the *reference lottery*. Suppose that at first you say the probability of rain next week is 0.3. Now imagine a lottery that pays a desirable prize if, in fact, it does rain next week and pays nothing otherwise. Further suppose there is a perfectly random device (such as a masked bag containing three marked marbles out of a total of ten) that also pays the same prize with probability 0.3 and nothing otherwise. Would you pay the same price to participate in both lotteries? If not, you must modify the probability and cycle through the process (perhaps employing more conceptual marbles) until indifference is reached.

The reference lottery approach generalizes in a straightforward manner to the case where a discrete random variable can take several values. But then the reference lottery becomes more complicated since we need as many prizes as states (or alternatively we need as many two-event lotteries as states, a more cumbersome approach). Naturally, decisions about indifference become correspondingly more demanding of judgment and patience as the number of states increases.

Probabilities as we have discussed them are essentially manifestations of verbal behavior and an appealing and helpful idea is to involve another sense such as sight in what we call the *visual impact method*. A chart or form is prepared on which the discrete values of the random variable are identified in a systematic manner along with respective spaces for counters. A reasonable number of counters (say 50 matches) is then allocated visually over the spaces according to degrees of belief. Probabilities are assessed as the ratios of observed cell frequencies to total counters. For example, if 13 out of 50 counters are allocated to the space for the third possible state, $P(\Theta_3) = 0.26$. This can be checked by means of appropriate reference lotteries.

We have been implicitly discussing the determination of what might loosely be called "unconditional" probabilities. The methods we have outlined are also directly applicable to *assessment of conditional probabilities* such as the likelihood probabilities that are often useful in decision analysis. An example of a possible likelihood would be the (conditional) probability that the International Wool Secretariat might issue a forecast that "demand for wool will be 'strong' next year" given that "wool price is 'low' next year." Alternatively, in some decision analyses it may prove simplest to derive posterior probabilities, i.e., $P(\text{state} \mid \text{forecast})$, directly. Such conditional probabilities are amenable to the same techniques as prior probabilities. As we shall see, assessment of judgmental joint probabilities also often requires knowledge of conditional probabilities. Conditional elicitation is not quite as simple as unconditional because it demands that the assessor project himself emotionally into the given or "if so-and-so" condition.

One Random Variable—The Continuous Case

Continuous random variables, even though they may be represented discretely in a decision analysis, are usually best estimated as continuous probability distributions. Methods for deriving continuous distributions may be classified according to whether assessment is direct, either as a cumulative distribution function (CDF) or as a probability density function (PDF), or indirect by means of an equivalent past or future hypothetical sample (Winkler, 1967, 1971). Direct elicitation of continuous PDFs is neither easy nor satisfactory. Indirect methods using sampling concepts seem more appropriate to statistical than to general managerial applications. We therefore concentrate on methods of determining the CDF.

CDFs may be defined as either $P(x \leqq X^*)$ or $P(x \geqq X^*)$, where X^* is some particular value of the uncertain quantity x. We find it convenient to use only the first definition, i.e., $P(x \leqq X^*)$. Such a CDF is therefore simply a function that gives the probability that x is less than or equal to a particular value X^*. A CDF function can be represented graphically with $P(x \leqq X^*)$ plotted on the vertical axis and X^* on the horizontal axis.

The visual impact method can be adapted to the estimation of the probability distribution of a continuous random variable. The range of the variable is first determined and then divided into a number of mutually exclusive and exhaustive intervals. Counters are allocated to the classes as before, but this time the resulting probability ratios are used to determine points on the CDF. A curve can be smoothed through these points, permitting values of $P(x \leqq X^*)$ to be read off for any selected values of X^*. Joint probabilities such as $P(X^{**} < x \leqq X^*)$ can also be read off the CDF very easily. These can be combined with cumulative probabilities such as $P(x \leqq X^*)$ to obtain conditional probabilities of the form

$$P(x > X^{**} \mid x \leqq X^*) = P(X^{**} < x \leqq X^*)/P(x \leqq X^*)$$

Such an implied conditional probability can then be compared with the corresponding directly obtained probability (Given that x is not greater than X^*, what is the probability that it is greater than X^{**}?) and thereby can provide a check with the earlier probabilities.

The second method involves direct determination of the CDF using what Raiffa (1968) has dubbed the *judgmental fractile method*. A $0.b$ fractile, denoted $f_{0.b}$ (read "point b fractile"), is that value of a random variable x for which $P(x \leqq f_{0.b}) = 0.b$. Fractile values therefore consist of points on the CDF function defined above, the lowest possible value of an uncertain quantity being the 0.0 fractile, and the highest the 1.0 fractile.

The judgmental fractile method is based on finding equally likely probability intervals. In using this method, the assessor has only to express judgments relative to the intuitively appealing "ethically neutral" probability 0.5. However, before initiating this procedure, it is a good idea to reflect on the uncertain quantity x in terms of the lowest possible value $f_{0.0}$, the highest possible value $f_{1.0}$, and the most likely (modal) value in order to avoid the bias of anchoring on the median $f_{0.5}$. The procedure then continues as follows:

1. Find the value $f_{0.5}$ such that it is equally likely that x is less than or greater than $f_{0.5}$.
2. Next, find $f_{0.25}$ such that it is equally likely that x is less than $f_{0.25}$ and between $f_{0.25}$ and $f_{0.5}$. Similarly, find $f_{0.75}$.
3. In a similar way, continue to find values that subdivide previously determined intervals into equally likely parts until there are enough fractiles to indicate the shape of the CDF. Nine fractiles are usually sufficient. This would imply fractiles at probability values of 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, and 0.875, together with the initially specified (but possibly revised) extreme fractiles at probability values of 0.0 and 1.0.

Questioning proceeds indirectly, e.g., in step (1): Consider the value X^*. Is x more likely to be above or below X^*? Several values may be tried before declaring that some particular $X^* = f_{0.5}$. Once a value is decided upon, it can be subjected to a 0.5 reference lottery check to sharpen judgment. When questioning is complete, the fractiles can be arranged graphically and a curve smoothed through the coordinates. Figure 2.1 illustrates the results of such a process. The value of the uncertain quantity at the point of steepest slope is the mode of the distribution, and if the elicitation has been consistent, this should agree closely with the earlier opinion as to the most likely value. This method is also directly applicable to derivation of conditional distributions.

Some applications will require probabilities in the form of a PDF rather than a CDF. The slope of the CDF is proportional to the height of the corresponding PDF, but this is not a very useful connection for numerical translations. It is usually easiest to plot a histogram with bar heights proportional to probabilities $P(X_1 < x \leqq X_2)$ where these are read off the CDF as differences $P(x \leqq X_2) - P(x < X_1)$. The histogram can then be

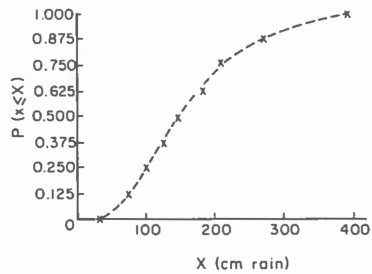


FIG. 2.1. Subjective CDF for next year's June–October rainfall in New Delhi.

smoothed, taking care to keep the area under the PDF equal to one. The process of reading intervals off the CDF so as to sketch a histogram can also provide a convenient opportunity for calculating moments of the distribution if these should be required. For example, the mean $E(x) \cong \Sigma[P(\text{in interval})(\text{midpoint of interval})]$, whereas the exact calculation $E(x) = \int xf(x)\,dx$ may not be feasible. In similar fashion, higher moments about the mean may also be approximated. Moments are discussed further in the next subsection.

Sometimes the analyst may be confronted with the reverse situation; i.e., he is given distributions in graphical PDF form but requires CDFs, perhaps for checking stochastic dominance as discussed in Chapter 9. For those with good eyes and ample patience, one method of going from a graphical PDF to the corresponding CDF is to divide the area under the PDF, plotted on graph paper, into a good number of vertical strips. The probability in each strip is estimated as the ratio of the number of squares in the strip to the total squares in all strips. The CDF is then defined by cumulating these probability elements from left to right and plotting them at values of the uncertain quantity x corresponding to the right (upper) limit of each strip. Counting squares (and parts of squares) to estimate areas in irregular trapezoidal strips is very tedious. The difficulty and tedium can be greatly reduced by constructing a histogram of rectangular strips to approximate a smooth PDF. However, if there is ready access to a good laboratory balance, an even better way is to plot the PDF on good-quality paper of regular thickness, dissect the cut-out PDF into vertical strips, weigh each strip, and determine probabilities as before using cumulative proportional masses in place of areas.

Our presentation has thus far given no attention to mathematical descriptions of elicited distributions and, with a few exceptions, this is the stance we will maintain in this book, as befits our generally nonmathematical approach. However, a great diversity of mathematical forms is available for the mathematically inclined (see, e.g., Johnson and Kotz, 1970). One form we will have occasion to use (in Chapter 8) is the triangular distribution (so called for the shape of its PDF, which has the apex at the mode or most likely value and the sides of the triangle spanning the range of the random variable). The triangular distribution is an appropriate mathematical form if the segments of the CDF on either side of the mode (the value of the random variable where the CDF has the greatest slope) can be well described by quadratic functions of the random variable. If we are prepared to assume that a distribution can be adequately described by a triangular approximation, elicitation is greatly simplified. Only three numbers are required to describe the triangular distribution completely, namely, the modal, the lowest possible, and the highest possible values. As well as having only these simple requirements, because of its simple linear form the triangular distribution is analytically very convenient.

Description of Distributions by Moments

Moments are useful descriptors of probability distributions, especially since they can facilitate some types of decision analysis (Section 4.6), but also because of the succinct description contained in just a few numbers. Moments are also useful for fitting theoretical distributions such as the beta distribution.

A moment is a summation of probabilities times distances or deviations from a specified point, with these distances raised to a power specified as the number of the moment. The kth moment M_k thus involves the kth power. We will consider only one moment about the origin (zero) and several moments about the mean. The first moment about the origin of an uncertain quantity x, $M_1'(X)$, is the expected value $E(x)$ or the arithmetic mean. For the discrete variable, taking values $x_1, x_2, \ldots, x_n$ with probabilities $P_1, P_2, \ldots, P_n$, $E(x) = \sum_i P_i x_i$; or more explicitly, $M_1'(x) = \sum_i P_i (x_i - 0)^1$. The complete expected value operator E works for continuous distributions as $E(\cdot) = \int_{-\infty}^{\infty}(\cdot)f(x)dx$ where $f(x)$ is the PDF, so $E(x) = \int_{-\infty}^{\infty} xf(x)\,dx$. Moments about the origin higher than the first are rarely encountered in decision analysis, so we turn now to the important class of moments about the mean, the so-called central moments.

The first moment about the mean is always zero, as seen, for instance, in the discrete case,

$$M_1(x) = \sum_i [x_i - E(x)]^1 P_i = \sum_i x_i P_i - E(x) = E(x) - E(x) = 0$$

The second moment about the mean $M_2(x)$ is known more familiarly as the variance $V(x)$ and naturally involves the second power or square of deviations from the mean. Thus $M_2(x) = E[x - E(x)]^2$ = either

$\int[x - E(x)]^2 f(x)dx$ or $\sum_i[x_i - E(x)]^2 P_i$, depending on the nature of the distribution. Squaring has the property of canceling negative signs, so variances are intrinsically nonnegative measures of dispersion. This contrasts with the case of the third central moment, where the cubing of deviations preserves the sign of the original deviations, $M_3(x) = \sum_i[x_i - E(x)]^3 P_i$ or $\int[x - E(x)]^3 f(x)dx$. If the deviations tend to be larger and/or to occur with higher probability above (below) the mean, the sign of $M_3(x)$ will tend to be positive (negative). Thus the third moment provides one measure of the skewness of a distribution; e.g., a distribution that has a relatively long tail to the right (left) will be described as being positively (negatively) skewed and will have a positive (negative) third moment. The fourth central moment we will define explicitly is $M_4(x) = \sum_i[x_i - E(x)]^4 P_i$ or $\int[x - E(x)]^4 f(x)dx$. This moment is strictly nonnegative and greatly magnifies the effect of large deviations from the mean.

The units of higher moments make them somewhat awkward to interpret, so for descriptive purposes they are usually transformed to more comprehensible units. For example, the positive square root of the variance, the standard deviation σ, is in the natural units of the uncertain quantity. Various dimensionless measures are also sometimes used, such as the coefficient of variation, $CV = \sigma/E(x)$, to measure dispersion; relative skewness, $\alpha_3 = M_3(x)/V(x)^{1.5}$, to measure departures from symmetry; and relative kurtosis, $\alpha_4 = M_4(x)/V(x)^2$, to measure peakedness and tail thickness. The reference point in interpreting measures such as these is usually taken to be the normal distribution, which is completely characterized by the first two moments, $E(x) = \mu$ and $V(x) = \sigma^2$. Being symmetric, its odd moments, e.g., $M_3(x)$, are all zero (so $\alpha_3 = 0$). Its higher even ($k = 2, 4, \ldots$) moments are given by $M_k(x) = (k - 1)(k - 3) \ldots (5)(3)(1)\sigma^k$; e.g., $M_4(x) = 3\sigma^4$. Thus for the normal case $\alpha_4 = 3$, and distributions with $\alpha_4 > 3$ (thick tails relative to the normal) are described as leptokurtic, while those with $\alpha_4 < 3$ are called platykurtic.

To briefly illustrate the descriptive use of moments, we give an example of a simple discrete distribution and mention some pragmatic procedures for continuous distributions. We will defer until Section 6.3 a discussion of using observed data to estimate moments directly, i.e., without estimating a distribution explicitly. Computation of moments for a discrete distribution simply involves substitution into the formulas given above. For instance, consider the random variable that takes values 5, 10, and 20 with probabilities 0.6, 0.3, and 0.1 respectively. The expected value is 8, so that the deviations from the mean are −3, 2, and 12 respectively. Calculation of the third moment, for example, then involves cubing these deviations, weighting by the probabilities, and summing to give in this case $M_3 = 159$ and $\alpha_3 = 1.65$, indicating positive skewness.

The integrations required for continuous distributions are more complex and are only really feasible when the mathematical form of the distribution is known and relatively simple or well documented. Computation of moments for arbitrary nonnormal continuous distributions (such as emerge from the judgmental fractile method) must rely either on the approximate method based on discrete approximation mentioned in the previous subsection or on sampling of variates from the distribution combined with direct estimation using the methods outlined in Section 6.3. However, it is possible to obtain estimates of the mean and variance directly from points on a subjective CDF. Pearson and Tukey (1965) and Perry and Greig (1975) have shown that the mean is very reliably estimated by

$$E(x) \cong f_{0.5} + (0.185)(f_{0.95} + f_{0.05} - 2f_{0.5}) \qquad (2.1)$$

where the 0.05, 0.5, and 0.95 fractiles are read from a CDF. Variance estimation by such a simple formula is not so reliable, but if the distribution is fairly bell-shaped, a good quick estimate is obtained by

$$V(x) \cong [(f_{0.95} - f_{0.05})/3.25]^2 \qquad (2.2)$$

More Than One Random Variable

The determination of distributions involving several random variables has received less attention and is intrinsically more difficult than assessment of distributions of single random variables. In considering distributions of several variables, the first step is to establish any stochastic independence among the variables. We believe this can usually be done by careful consideration and introspection without direct recourse to numerical enumeration. If this is true, it means that full specification of a joint distribution need only be attempted when dependence is judged relevant. As we will see, evaluating judgmental joint distributions is rather irksome and is to be avoided whenever possible.

A test of intuitively assessed independence can be made by simply applying the definition of the term. That is, suppose unconditional or marginal distributions have been determined for two discrete random variables A and B, then several products $P(A_i)P(B_j)$ can be compared with independent assessments of the joint probabilities $P(A_i, B_j)$. If these estimates correspond fairly closely, independence can reasonably be assumed. From the definition of conditional probability, $P(A_i, B_j) = P(A_i \mid B_j)P(B_j)$; and if A and B are independent, $P(A_i \mid B_j) = P(A_i)$; i.e., conditional probability is identical to marginal probability. Such a test procedure can be applied to groups of random variables to identify independent sets of variables. As a practical matter, this is difficult if the sets are not small.

Sections marked with Skippy are relatively difficult and can be skipped without loss of continuity.

If dependence exists, the methods discussed earlier for direct assessment can be extended to the multivariate case. For distributions that can be handled in discrete form, the visual impact method is the most useful way of directly measuring PDFs. Consider first the simplest case of a pair of variables, each composed of n discrete intervals. The fastest method of assessment is to draw up a two-way table with $n \times n$ compartments and to allocate counters according to direct judgment of the joint PDF. For example, suppose we need a farmer's joint distribution for the yield and price of his next (dryland) wheat crop. After identifying the ranges of possible yield and price, we could draw up a convenient 5×5 table to encompass the farmer's distribution and present him with 100 counters to allocate to the 25 cells according to his beliefs. The end result of his head scratching and juggling of counters might be the allocation shown in Table 2.1. Checking the values of $P(A_i)P(B_j)$ against $P(A_i, B_j)$ for these data indicates that the variables are not independent. For example, for a yield of 2–3 t/ha and a price of 5–6 ¢/kg we have $P(A, B) = 0.12$ and $P(A)P(B) = 0.132$.

TABLE 2.1. Elicited Frequency Data for Joint Distribution of Two Variables

Wheat Yield	Wheat Price (¢/kg) 4–5	5–6	6–7	7–8	8–9	Marginal Totals
(t/ha)						
0–1	1	3	4	2	1	11
1–2	2	7	5	4	1	19
2–3	7	12	15	6	0	40
3–4	6	7	5	2	0	20
4–5	3	4	2	1	0	10
Marginal totals	19	33	31	15	2	100

Data such as those of Table 2.1 can be used to compute moments of the subjective distribution by (1) taking the fraction of counters in each cell as an estimate of the probability associated with each joint (or marginal) interval and (2) associating all such probabilities with the respective midpoints of the intervals; e.g., the midpoint of the first yield interval Y_1 is 0.5 t/ha, and for the first price interval R_1 it is 4.5 ¢/kg. An estimate of the mean of the marginal distribution of yield is thus given by

$$E(Y) \cong \sum_i P_i Y_i$$

$$= (0.11)(0.5) + (0.19)(1.5) + \dots + (0.10)(4.5) = 2.5$$

where P_i denotes the probability associated with the interval whose midpoint is Y_i. For price we likewise have $E(R) \cong \sum_j P_j R_j = 6$. Marginal variances are likewise calculated approximately as

$$V(Y) \cong \sum_i (P_i)[Y_i - E(Y)]^2$$

$$= \sum_i (P_i)(Y_i^2) - [E(Y)]^2 = 1.23$$

$$V(R) \cong \sum_j (P_j)[R_j - E(R)]^2$$

$$= \sum_j (P_j)(R_j^2) - [E(R)]^2 = 1.05$$

For the covariance of yield and price, we have

$$\text{Cov}(Y, R) \cong \sum_i \sum_j (P_{ij})[Y_i - E(Y)][R_j - E(R)]$$

$$= \sum_i \sum_j (P_{ij})(Y_i)(R_j) - E(Y)E(R)$$

$$= (0.01)(0.5)(4.5) + \cdots + (0)(4.5)(8.5) - (2.5)(6)$$

$$= -0.285$$

where P_{ij} denotes the probability associated with the joint interval centered on Y_i and R_j. Given the covariance, we can readily compute the correlation coefficient ρ_{YR} as

$$\rho_{YR} = \text{Cov}(Y, R)/[V(Y)V(R)]^{0.5} \cong -0.25$$

The method illustrated above extends adequately to the general case of k variables by taking $k(k - 1)/2$ pairwise bivariate distributions and thus determining the covariances. However, this generally does not give the full joint distribution that may be required in some problems, though it is satisfactory for analyses where the variance and covariance are the highest moments required of the distribution of consequences.

The complete distribution has to be approached by assessing marginal and conditional distributions separately. For the three variables A, B, and C with n intervals in each, one marginal distribution (say that of C) is assessed; then for each interval of this variable the conditional distribution of another variable (say B) is assessed. Finally, for each combination of intervals of these two variables, the conditional distribution of the remaining variable (in our case, A) is assessed. The joint probabilities are then found as $P(A_i, B_j, C_k) = P(A_i \mid B_j, C_k)P(B_j \mid C_k)P(C_k)$. The amount of judgment explodes rapidly as the number of variables increases. For the full joint distribution of k variables, one (arbitrary) marginal distribution and Σn^i progressively more complex conditional distributions must be assessed where the summation is over $i = 1, 2, \ldots, k - 1$.

Suppose a Punjab irrigation farmer is planning his crop strategy in the face of uncertain availabilities of his two key irrigation inputs—canal water and electric power for his tubewell pumps. Ideally, he would like to irrigate his crop four times. This could be done least expensively with water from the canal rather than from his tubewells. On the basis of his knowledge of the storage situation in the mountain reservoirs and ignoring the low-probability event of further clashes with Pakistan this year (in which case reservoir levels would be dropped for safety), he estimates the (unconditional) probabilities of various total numbers of canal irrigations that he might be able to make as:

Possible canal irrigations	1	2	3	4
Probability	0.05	0.10	0.25	0.60

Hydroelectricity generated from water in the same mountain reservoirs is an important contributor to the grid from which he must draw his electric power for tubewell pumping. Restrictions may be imposed on the supply of electricity especially when water is in short supply. He estimates the marginal probabilities pertinent to this season as no restrictions, 0.25; slight restrictions, 0.50; and severe restrictions, 0.25.

Even in years of ample water supply, restrictions are quite likely because of the politically inspired oversubscription in rural electrification schemes. However, since there is apparently some causal link between canal water supply and electricity supply, we can very properly inquire as to the interdependence of these distributions because it may prove to be an important factor in crop planning. Probably the simplest method is to elicit conditional distributions for canal irrigations given each electricity situation. Suppose these are as given in Table 2.2. Joint probabilities are then found by multiplying these conditionals for number of irrigations by the probability of the respective power situations. This gives the joint probabilities of Table 2.3.

TABLE 2.2. Conditional Distributions for Canal Irrigations

	Number of Possible Irrigations			
Power Situation	1	2	3	4
No restrictions	0.00	0.00	0.20	0.80
Slight restrictions	0.00	0.10	0.30	0.60
Severe restrictions	0.15	0.25	0.20	0.40

TABLE 2.3. Joint and Marginal Distributions of Two Variables

	Number of Possible Irrigations				
Power Situation	1	2	3	4	Marginal
No restrictions	0.00	0.00	0.05	0.20	0.25
Slight	0.00	0.05	0.15	0.30	0.50
Severe	0.0375	0.0625	0.05	0.10	0.25
Marginal	0.0375	0.1125	0.25	0.60	1.00
cf. Initial marginal	0.05	0.10	0.25	0.60	1.00

Once the joint probabilities are calculated, they can be summed across electricity states to give marginal probabilities as indicated in Table 2.3. In turn these marginals can be compared with the initial marginal probabilities on possible irrigations (also shown in Table 2.3). The close correspondence apparent in our example is much closer than will usually be found on a first attempt, especially when counters are used to determine discrete distributions (since these tend to introduce rounding errors). The slight inconsistencies are best smoothed out by systematically checking through the elements of the joint probability table, after which the marginal distributions will "look after themselves." The lack of independence between the irrigation and power events facing our farmer is readily apparent by comparing products of marginals with corresponding joint elements; e.g., the joint probability of four irrigations and no power restrictions is 0.20, which is not equal to the product, $(0.60)(0.25) = 0.15$, of the relevant marginals.

Though still not easy, joint assessments are greatly simplified if it can be assumed that the variables follow a multivariate normal distribution. In this case the distribution is specified by its k marginal means and $k(k + 1)/2$ variances and covariances, and there is no need to worry about higher moments and co-moments. More research is needed to assess the feasibility of alternative methods of estimating joint distributions, to determine how adequate the normal assumption can be for nonnormal distributions, and to explore the conditions under which only partial measurement can be sufficient.

Probability Distributions for Enterprise Revenues

Some applications, such as quadratic risk programming for farm planning (Chapter 7), require estimates of joint distributions of enterprise net revenues. Usually, however, only means, variances, and covariances are required so that assessment need not be too demanding. If several or many historical observations on enterprise revenues are available and judged pertinent, these could be analyzed, perhaps using the variate difference method demonstrated by Carter and Dean (1960), to provide the desired statistics. However, because markets are usually changing and new varieties and new diseases are often emerging, the available historical data may be of dubious relevance—in which case assessment must rely heavily on judgment.

The most analytical approach to specifying revenue distributions is to begin with the underlying distributions of the uncertain components. Usually these components will be yield and price. We have already glimpsed some of the tedium of measurement for one such enterprise in the wheat example given in the previous subsection. We could proceed to derive the moments of the wheat revenue distribution by applying standard formulas, but even this is not completely straightforward because of the nonindependence of yield and price. To illustrate what is involved, we need to define the following notation:

G = wheat gross margin in \$/ha
Y = wheat yield in t/ha
R = wheat price in ¢/kg
C = variable costs of wheat production in \$/ha. It is assumed that it is known with certainty that $C = 40$.

We can now proceed from the previous joint distribution on Y and R, making appropriate allowance for the units used as follows:

$$\begin{aligned} G &= 10YR - C \\ E(G) &= 10E(Y)E(R) + 10\,\mathrm{Cov}(Y,R) - C \\ &= (10)(2.5)(6) + (10)(-0.285) - 40 = \$107.15 \end{aligned}$$

Since the variable costs are known with certainty, they do not influence the variance of G, which is given approximately by

$$\begin{aligned} V(G) &\cong 10^2[E(R)^2V(Y) + E(Y)^2V(R) + 2E(R)E(Y)\,\mathrm{Cov}\,(Y,R)] \\ &\cong 100[(2.5)^2(1.05) + (6)^2(1.23) + 2(2.5)(6)(-0.285)] \\ &\cong 4200 \end{aligned}$$

Hence the standard deviation of wheat gross margin is approximately \$65/ha. An exact formula for the variance of such a product has been given by Bohrnstedt and Goldberger (1969). It involves the addition of the following higher order terms (multiplied by 10^2 in our case) to the approximate expression

$$2E(Y)E\{[Y - E(Y)][R - E(R)]^2\} + 2E(R)E\{[Y - E(Y)]^2[R - E(R)]\} + E\{[Y - E(Y)]^2[R - E(R)]^2\} - \mathrm{Cov}\,(Y,R)$$

A computationally more convenient form of this adjustment term is

$$E(Y^2R^2) - E(Y^2)[E(R)]^2 - [E(Y)]^2E(R^2) - 4E(YR)E(Y)E(R) + 5[E(Y)]^2[E(R)]^2 - \text{Cov}(Y,R)$$

In the present example, these terms add only about 140 to our above estimate of 4200 for $V(G)$ and thus make only a small adjustment of +3% to the approximate value.

Correlations between Enterprise Returns

Even after having gone through the problem of estimating revenue distributions for several enterprises, the analyst is still left with the difficulty of establishing the appropriate correlations between enterprise returns. These are most likely to be positively correlated because of high positive correlations between enterprise prices or between enterprise yields or both. However, even if such joint distributions have been obtained (i.e., joint yield distributions, joint price distributions, and joint yield-price distributions for each enterprise), the foregoing formulas indicate that it is not a trivial chore to derive joint revenue distributions. Indirect methods (discussed below) may prove to be most expedient; alternatively, the assessment can be approached directly and subjectively as follows.

Suppose a farm planning problem calls for consideration of three winter crops: wheat under standard technology, new high-yielding wheat needing high fertilizer rates, and oats. The farmer is a traditional wheat grower and feels most comfortable in his knowledge of the revenue distribution for the standard-wheat crop. The judgmental fractile method indicates that his unconditional distribution for revenue ($/ha) from standard wheat can be depicted by the seven equally spaced, 1/8(1/8)7/8, fractiles 35, 75, 95, 105, 130, 150, 170. Plotting these on normal probability paper (Section 2.4 below) suggests that this distribution can be well represented as normal, with mean 105 and standard deviation 60. Such approximate normality suggests proceeding with estimation of the joint distribution as though it was multivariate normal. This presumption is checked in a useful way by estimating conditional distributions for the other enterprise revenues, again using the convenient judgmental fractile method; e.g., Given that revenue from standard wheat is going to be $100/ha, is it more likely that revenue from oats will be more than $90/ha or less than this amount? Continuing in this manner, suppose the following seven fractiles are pinned down for oats revenue ($/ha), conditional on standard wheat revenue being $100/ha, viz., 30, 50, 60, 80, 85, 105, 175. This conditional distribution is also fairly normal, with graphically approximated parameters of mean 75 and standard deviation 35.

In an identical manner, conditional judgmental fractiles for new-wheat revenue (given standard-wheat revenue is \$100/ha) are established as 145, 150, 165, 180, 195, 205, 225, which (graphically) indicates a normal distribution with parameters (180, 35). We still need a link between the oats and new-wheat revenues. This is approached by identifying a further conditional distribution for new wheat, given that oats revenue is \$60/ha, for example. The fractiles are 75, 130, 150, 175, 180, 210, 250; and the fitted conditional normal parameters are (165, 70).

To summarize, we now have eight estimates of distributional means and standard deviations but, in total, nine parameters of a three-variable joint normal distribution are required, viz., 3 marginal means, 3 unconditional variances, and 3 correlations (or covariances). More information is obviously required. The easiest data to obtain are unconditional central tendency measures of revenue from oats and new wheat. Suppose these (obtained as unconditioned medians from the first step of a judgmental fractile elicitation) are \$80/ha and \$190/ha respectively.

Now we have just enough information to obtain estimates of the parameters of the subjective multivariate distribution. Solution requires knowledge of at least one formula for the parameters of a conditional distribution for one variate of a bivariate normal distribution. If X_1 and X_2 are jointly normally distributed with means $E(X_1)$ and $E(X_2)$; standard deviations σ_1 and σ_2; and correlation ρ_{12}; the conditional distribution of X_1, given $X_2 = X_2^*$, is characterized (Raiffa and Schlaifer, 1961, p. 250) by mean and variance

$$E(X_1 \mid X_2 = X_2^*) = E(X_1) + \rho_{12}(\sigma_1/\sigma_2)[X_2^* - E(X_2)] \qquad (2.3)$$

$$V(X_1 \mid X_2 = X_2^*) = \sigma_1^2(1 - \rho_{12}^2) \qquad (2.4)$$

These formulas provide the method of solution in the present instance but, with the data so far presented, simultaneous solution of nonlinear equations would be involved. Solution is greatly simplified and reliability improved with two further pieces of information, namely, the unconditional standard deviations for revenues from oats and new wheat. Since the presumption of normality has now been well substantiated, this final procedure can be more direct than in the judgmental fractile method. We only need to elicit the range symmetric about the mean such that it is equally likely that revenue will fall either inside or outside the range, e.g., Do you feel that it is more likely that revenue from oats per hectare will be between \$60 and \$100 or outside this range? In this way, suppose the "equally likely inner ranges" for oats and new-wheat revenues are established as 80 ± 30 and 190 ± 50 respectively. Since 50% of the area under the normal PDF is contained in the range of the mean ± 0.6745 times one standard deviation, unconditional standard deviations are found as

30/0.6745 ≅ 44.5 for oats revenue and 50/0.6745 ≅ 74 for new-wheat revenue. At this stage we can usefully tabulate the parameter information so far obtained. This is done in Table 2.4.

TABLE 2.4. Parameters for Normal Distributions of Enterprise Revenues ($/ha)

Enterprise	Condition (if any)	Mean	Standard Deviation
Standard wheat	...	105	60
Oats	...	80	44.5
New wheat	...	190	74
Oats	Standard wheat = 100	75	35
New wheat	Standard wheat = 100	180	35
New wheat	Oats = 60	165	70

The formulas for the conditional normal mean and variance given in equations (2.3) and (2.4), when taken with our parameter data, give two estimating equations for each of the correlation coefficients (which are the only parameters now missing). For example, the correlation between oats and standard-wheat revenues can be found by solving each of the equations:

$$75 = 80 + \rho(44.5/60)(100 - 105)$$
$$35^2 = 44.5^2(1 - \rho^2)$$

The first implies that $\rho = 1.35$ and the second that $\rho = \pm 0.62$. Alternative solutions for the other correlations between enterprise revenues, found in a similar manner, are listed in Table 2.5.

TABLE 2.5. Estimates of Correlation Coefficients between Enterprise Revenues

Enterprise Pairs	Based on Conditional Mean	Based on Conditional Variance	Preferred Estimate
Standard wheat and oats	1.35	±0.62	0.62
Standard and new wheat	1.62	±0.88	0.88
Oats and new wheat	0.75	±0.32	0.75

The first two estimates of Table 2.5 based on conditional means are clearly wrong since correlations by definition cannot be outside the range ±1; but because they are positive, they do aid the subjective step of opting for the positive solutions of the first pairs of estimates based on the conditional variance. Subjective interpretation is even more important in selecting a preferred estimate in the third case, where the same logic would suggest 0.32 as the best estimate. However, because of the subjectively high positive correlation felt by the analyst, the larger value of 0.75 is selected as "best." The game is not quite over since it remains to check that the full matrix of correlations (with units on the diagonal) is positive-definite, i.e., that the correlations are internally consistent and feasible. The necessary and sufficient condition for this to be so is that the following $k - 1$ determinants of the correlation matrix $[\rho_{ij}]$ are positive:

$$\begin{vmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{vmatrix}, \ldots, \begin{vmatrix} 1 & \rho_{12} & \cdots \rho_{1k} \\ \rho_{12} & 1 & \cdots \rho_{2k} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \rho_{1k} & \rho_{2k} & \cdots 1 \end{vmatrix}$$

If any of these is not positive, the correlation matrix must be revised until it checks as feasible while agreeing with the analyst's (probably vague) feelings about the relationships. These conditions are all fulfilled in our example.

Ways around Joint Specification

The above discussion of joint distributions serves to emphasize that such assessment is rather tedious even for only three related uncertain quantities and that procedures are neither simple nor guaranteed to produce satisfactory results. Together these facts can make us understanding (but hopefully not tolerant) of the frequent failure to capture probabilistic dependencies in decision analyses (or even to attempt to do so). On the positive side we can be encouraged to follow the suggestions of Pratt, Raiffa, and Schlaifer (1965) to avoid the difficulties of joint specification whenever possible.

Two broad approaches to such circumnavigation are available; viz., using transformations of variables to achieve independence and relating jointly dependent variables to common explanatory variables to achieve conditional independence.

Relative to the use of transformations, suppose that harvested grain yield and absolute yield loss to birds are jointly dependent uncertain quantities. It may be, for example, that transformation to harvest potential (prebirds) and percentage loss to birds gives independent quantities. However, rather than transformations, the use of common explanatory variables to achieve at least approximate independence seems likely to be generally more useful in agricultural applications. Many of the uncertain quantities of interest to agricultural managers have some common influences. For instance, many dryland crop yields are closely dependent on common rainfall experience. Once this influence is accounted for, remaining sources of variation, such as crop-specific diseases, may lead to little statistical dependence.

For composite distribution situations such as the enterprise revenue case above, it may be possible to relate revenues to a set of variables, including measures of weather and general market indicators, such that the disturbances in the regressions for individual enterprises are effectively independent from each other. An explanatory structure of this nature can be a useful step in an indirect approach to specification of such multivariate distributions as are required in quadratic risk programming. Once the structure has been identified, the underlying variables (such as weather and an index of farm prices) can be simulated, the joint variables of interest can be generated in Monte Carlo fashion, and sufficient "data" can be generated for estimation of means, variances, and correlations by conventional unbiased sampling procedures (Section 6.3); viz., $E(x) = \sum x/n$, $V(x) = \sum[x - E(x)]^2/(n - 1)$, and $\mathrm{Cov}(x,y) = \sum[x - E(x)] \cdot [y - E(y)]/(n - 1)$ for a sample of size n.

2.4 HISTORICAL DATA AS AN AID TO ELICITATION

Our experience obviously has an important influence on our feelings of uncertainty and degrees of belief. To begin the elicitation procedure, it is often useful to array pertinent historical data in a form that can help us sort out our beliefs. When data abound (usually the case with rainfall records, for example), they may be formed into frequency diagrams such as a histogram or processed into statistics such as sample moments and/or fractiles (Schlaifer, 1959, 1969); or a frequency curve fitting exercise may be entertained such as fitting the moment-based Pearson curves (Elderton and Johnson, 1969; Phillips, 1971). The best approach will depend upon the problem, the amount of data, and the statistical services available.

Subjective interpretations of historical data are invariably required. A useful starting point is the fact that most distributions are smooth and unimodal. Those that are not probably reflect the influence of an assignable cause that should be explicitly accounted for in the distributional analysis. Thus apparent irregularities in historical data records can usually be safely smoothed out, either by hand or with a numerical device. Normal probability paper is very useful for manual treatment of data for which the assumption of normality seems a reasonable approximation.

Consider the following example of reducing a body of historical data to probability distribution form. Suppose we are interested in describing the environment of a pea-growing enterprise. A very important factor in the process is the rainfall during the growing period because disease problems increase in very wet seasons. A nearby recording station has data available for growing-season rainfalls for the years 1909 to 1967. These 59 observations are presented in Table 2.6.

TABLE 2.6. Historical Record of Pea Growing-Season Rainfall

	Decade Year									
Decade	0	1	2	3	4	5	6	7	8	9
					(mm)					
1900	...	...	...	...	...	...	...	...	...	290
1910	341	376	550	408	369	343	305	428	510	474
1920	452	364	272	248	447	479	264	254	419	366
1930	256	533	435	341	630	498	276	405	719	419
1940	424	251	349	483	270	212	434	535	255	387
1950	214	584	399	288	252	344	581	321	240	554
1960	519	407	532	599	584	458	361	513	...	...

A straightforward approach to processing these data is to form them into frequencies by specifying several intervals and counting the number of occurrences in each. For example, with 50-mm intervals beginning at 201 mm we obtain the frequencies shown in Table 2.7. Some analysts might proceed to fit various theoretical distributions to frequencies such as those of Table 2.7 and then test for goodness of fit with a chi-square test. However, it is usually best to "have a good look" at the data first. A frequency histogram is a good start. Using the data of Table 2.7, we have the frequency histogram of Figure 2.2, which is only vaguely suggestive of a normal distribution.

TABLE 2.7. Frequencies of Pea Growing-Season Rainfall in 50-mm Intervals

Interval	Frequency	Cumulative Frequency
(mm)		
201–250	4	4
251–300	11	15
301–350	7	22
351–400	7	29
401–450	10	39
451–500	6	45
501–550	7	52
551–600	5	57
601–650	0	57
651–700	1	58
701–750	1	59

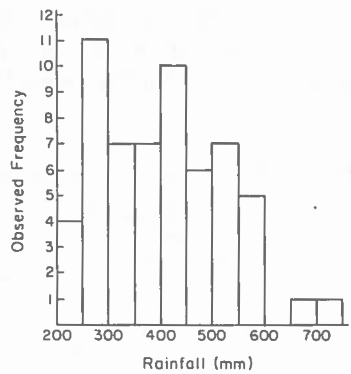


FIG. 2.2. Histogram of pea growing-season rainfall, 1909–67.

The irregularities apparent in even a fairly long record are quite obvious in Figure 2.2 notwithstanding the reasonable presumption that the present random process has a smooth unimodal frequency function. How then might a smoothing of the data best take place? We recommend that it be done in CDF form. This step also provides a ready check on the normality of the distribution if normal probability paper is used, as has

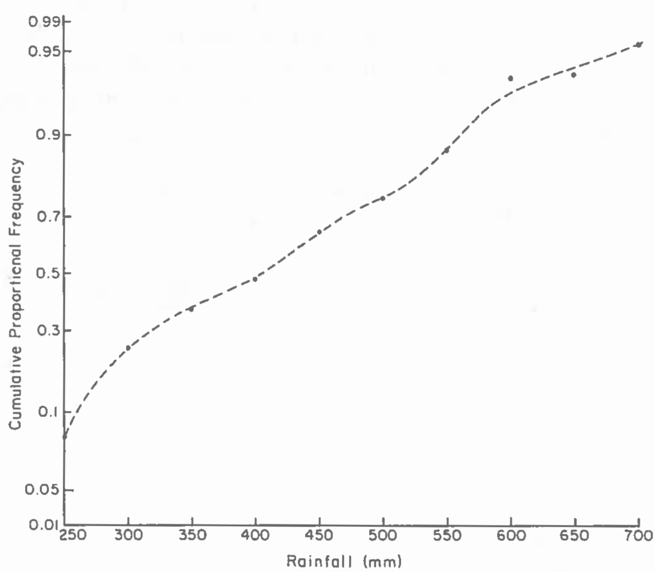


FIG. 2.3. The CDF for pea growing-season rainfall plotted on normal probability paper.

been done in Figure 2.3 where the cumulative frequencies of Table 2.7 are plotted at the upper bound of each interval.

If the cumulative frequencies follow a straight line when plotted on normal probability paper, a normal probability process is operative. This is not strongly evidenced in Figure 2.3. As usually found with rainfall data, the concave curvatures at each end of the graph indicate that the operative distribution is positively skewed, in turn suggesting that an appropriate transformation may just about normalize the distribution. As will be seen later, normality in distributions can greatly simplify decision analysis. Square root and cube root transformations often work for rainfall data. To try this possibility, square roots were extracted for the 59 observations, and these again were classified into frequencies as shown in Table 2.8.

TABLE 2.8. Frequencies of Square Root Transformed Pea Growing-Season Rainfall

Interval	Frequency	Cumulative Frequency
($mm^{0.5}$)		
14.01–15	2	2
15.01–16	6	8
16.01–17	6	14
17.01–18	3	17
18.01–19	6	23
19.01–20	6	29
20.01–21	9	38
21.01–22	6	44
22.01–23	4	48
23.01–24	5	53
24.01–25	4	57
25.01–26	1	58
26.01–27	1	59

The graph of these cumulative frequencies on normal probability paper is shown in Figure 2.4. It is rather more linear than the previous one and looks to us like a reasonably fair representation of the data, so that in the absence of other subjective information we are prepared to accept a normal approximation. To fit a normal distribution, all we need to do is place a straight line through the data of Figure 2.4 to make a fit of best subjective compromise. This is done by the unbroken line. The parameters of the normal distribution can then be read directly off the graph; the mean corresponds to the 0.5 fractile and is indicated by the unbroken line as 19.95 $mm^{0.5}$. To find the standard deviation, we can use the fact that 84.1% of the area under a normal PDF lies below the mean plus one standard

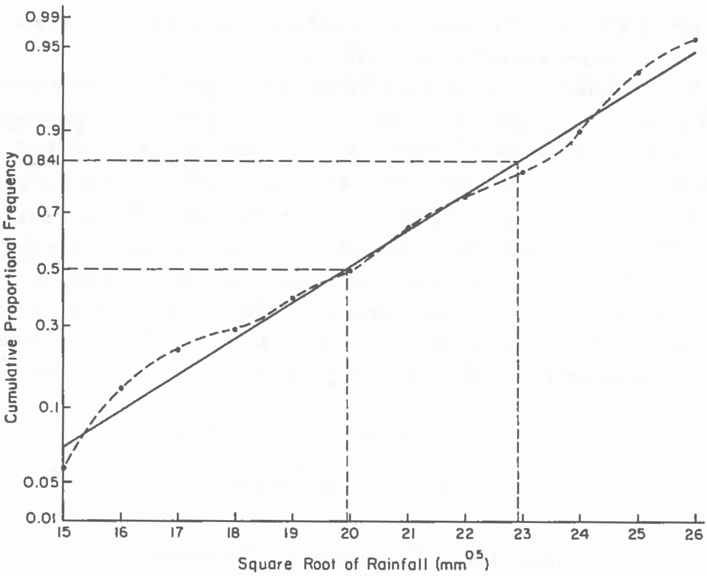


FIG. 2.4. The CDF for square-root transformation of pea growing-season rainfall plotted on normal probability paper.

deviation. From Figure 2.4, the 0.841 fractile is read off the fitted straight line as 22.95 mm, so one standard deviation is 22.95 − 19.95 = 3.00 $mm^{0.5}$. Thus our analysis indicates that, as far as can be judged from the available historical data, pea growing-season rainfall has a probability distribution that is approximately normalized by a square root transformation, the derived distribution having a mean of 19.95 $mm^{0.5}$ and a variance of 9 mm.

Sparse Data Situations

Sometimes only a few pertinent historical observations may be available, i.e., a *sparse data situation*. Though more data would be desirable, all is not lost in such a situation since we can make use of a handy distributional rule that works irrespective of the form of the underlying probability distribution. Suppose only N observations are available on a continuous random variable. When arrayed in ascending order of size, the Kth observation is a reasonable estimate of the $K/(N + 1)$ fractile. As Mood and Graybill (1963, p. 405) show, this rule is reasonable in the sense that the *expected* fraction of all values of the random variable falling below the Kth order statistic is $K/(N + 1)$. See also Barnett (1975).

Fractile estimates from the sparse data rule can be plotted and a CDF smoothed subjectively through these coordinates to incorporate whatever additional information may be available (e.g., the nonnegativity of many variables and the unimodal nature of most empirical distributions). The CDF of a unimodal two-tailed distributions has an S shape. Consider, for example, four observations on a continuous nonnegative random variable: 6, 10, 4, 6. In applying the rule, these would be arrayed in order as 4, 6, 6, 10, which then serve as estimates of the 1/(4 + 1), 2/(4 + 1), 3/(4 + 1), and 4/(4 + 1) fractiles respectively. Clearly, this is minimal information from which to sketch a CDF. However, if it is all we have to go on and our decision analysis requires an estimate of the distribution, we should proceed to smooth a subjective CDF through these points, perhaps resulting in something akin to Figure 2.5.

Such a sparsely based CDF does not pretend to be more than it is—a smoothed treatment of some scanty information. As for any subjective probability, different judgments will be made by different analysts. It is an empirical question whether such differences (which tend to be small) also result in different decisions. People are likely to differ most in assessing the tails of CDFs, and since relatively little probability is located there, such differences have little impact in many decision analyses (some exceptions are discussed in Chapter 9). At any rate, sparse data CDFs will be a better stochastic representation than those obtained by assuming, for example, that the few observations represent an exhaustive equally likely discrete listing of states. We have found in some sampling experiments from known parent distributions that even with as few as only three to five observations, there is about a 50% chance that an estimated CDF will be very similar to the parent. Obviously, chances of success increase as sample size increases, but even with a "good" size of a dozen or so, estimates are sometimes quite wide of the mark; although in the absence of any other historical or subjective information, they are still the best one can do (Anderson, 1974).

Sometimes a data situation may not be quite as sparse as it seems. For example, if a process is highly conditioned by an uncertain quantity about which there is a good deal of probabilistic information, the latter may be

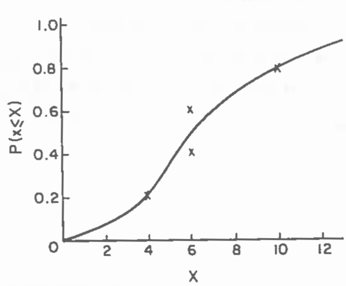


FIG. 2.5. Subjective CDF based on sparse data-rule fractiles.

able to be brought to bear on the former. A situation of this type is where an experiment may have been conducted in only a few years. If the process is very dependent on rainfall, then knowledge of rainfall distributions may greatly assist in extending and interpreting the limited experimental data (see Problem 2.9).

2.5 DIGRESSION ON SCORING RULES

Personal probabilities are only required to be consistent with probability calculus and to correspond with individual's judgments, i.e., to be "honest." *Scoring rules,* which have been developed to provide a means for keeping assessors honest, are payoff schemes that depend on the estimated distribution and on the state that results.

Suppose that for a set of mutually exclusive and exhaustive states, the assessor has beliefs $p = \{p_1, p_2 \ldots, p_m\}$ but these probabilities are expressed as $r = \{r_1, r_2 \ldots, r_m\}$. This judgment is honest if $r = p$. A scoring rule assigns a score or reward $S_i(r)$ if the ith state occurs so that the subjective expected score is $S(r, p) = \sum p_i S_i(r)$. *Strictly proper scoring rules* are those for which $S(r, p)$ is maximized only when $r = p$. One family of strictly proper scoring rules is $S_i(r) = a \log_e r_i + b$, $a > 0$. Note, however, that in this case $r_i = 0$ implies $S_i(r) = -\infty$ and that $S_i(r)$ depends only on r_i. Other strictly proper scoring rules avoid these rather severe features and may also provide more sensitive scoring, e.g., the quadratic scoring rule, $S_i(r) = 1 + 2r_i + \sum_j^m r_j^2$.

The rules discussed presume that assessors will want to maximize the expected score; otherwise it is necessary to introduce the additional complexity of accounting for preference in modifying the scoring rules. To simplify matters, the range of scores should be kept so that utility of score is linear. The use of scoring rules seems to have most relevance to clinical research in psychology and to meteorological forecasting. Here the states can be realized frequently, and the cycle of scoring and determination can be repeated in an adaptive manner to provide educational feedback to the assessor or perhaps a measure of performance to someone appraising his work. The cycle time for probabilistic states in agricultural decisions is typically of such length as to effectively preclude the adaptive use of scoring rules as aids to elicitation. Accordingly, we will not take up the idea any further here.

PROBLEMS

2.1. Using both the visual impact method and the judgmental fractile method, formulate your CDF and PDF for two of the following uncertain quantities:

(a) Your annual salary or net income five years from now.

(b) The average price three years from now of some farm product with which you are familiar.
(c) The rate of inflation in your country next year.

2.2. Comment on the following statements:
(a) Objectivity in science is a myth, in life is an impossibility, and in decision making is an irrelevance.
(b) It was clear how much was yet to be expected from his [Leonard Jimmie Savage's] clarifying spirit for the success in our task: to relieve science and mankind from the strange superstitious prejudice that the obvious subjective probability feelings could or should be related to, or even replaced by, some hypothetical notion that, in some indefinable sense, could be called objective (de Finetti, 1972, p. vi).

2.3. Design an experiment aimed at investigating anchoring as a source of bias in the elicitation of probabilities.

2.4. The joint probability distribution of the random variables X and Y is represented by the following table:

	Y			
X	5	6	7	8
1	0.01	0.18	0.24	0.06
2	0.06	0.09	0.12	0.03
3	0.02	0.03	0.04	0.12

(a) Determine the marginal distributions of X and Y.
(b) Are X and Y independent?
(c) Graph the joint distribution in three dimensions.
(d) What is the conditional distribution of X, given $Y = 7$.
(e) What is the expected value of Y, given $X = 3$?
(f) Repeat this exercise after interchanging the top left probability 0.01 with the bottom right probability 0.12.

2.5. Complete the following probability table, given that X and Y are independent random variables.

	Y			
X	18	20	23	$P(X)$
−10	0.05			
0			0.20	0.50
4				
$P(Y)$	0.20			1.00

2.6. Derive your joint probability distribution for next year's yield per acre and price per unit for some farm product in some district with which you are familiar.

2.7. Suppose that, instead of being confronted with all the data of Table 2.6, only the last nine rainfalls are available to you for the purpose of probability specification. Make your best estimate of the distribution based on these nine observations. Compare and contrast your estimate with that of Figure 2.4 after resketching the fitted line of Figure 2.4 on ordinary graph paper and with the rainfall not (square root) transformed.

2.8. Reconsider the judgmental fractiles discussed in connection with the joint assessment of revenues from standard wheat, oats, and new wheat in Section 2.3. Proceed with your own assessment of (a) the adequacy of the normality assumption and (b) the parameters of a fitted multivariate normal distribution. Report your results critically.

2.9. The traditional recommendation on seeding rate for peas is 100 kg/ha. Tom Bayes, the district research officer, feels that this is too conservative and decides to run an experiment to check his hunch. He conducted his rate-of-seeding experiment in the 1965–66 and 1966–67 seasons, but the results of the last experiment were so disastrous that he was severely discouraged from continuing the work and in fact was transferred to another area. The 1966–67 crop suffered badly, especially in the high-density treatments, because of fungal attack on the peas induced by the very wet season (see the pertinent rainfall data in Table 2.6). Bayes' mean experimental yields in kg/ha were:

		Seeding rate (kg/ha)			
Season	Rainfall	90	110	135	180
	(mm)				
1966	361	4830	6740	7860	9540
1967	513	4830	4840	3930	2800

(a) Does it make good sense simply to average the yields across seasons for each seeding rate with an implicit probability of 0.5 for each "state?" Why?

(b) Would this be a good occasion to use the sparse data smoothing rule to estimate a yield distribution for each seeding rate? Why?

(c) Given the information of Table 2.6 and Figures 2.2–2.4, what probabilities would you attach to the results for each year?

(d) If seed peas cost 40¢/kg and revenue net of harvest cost is 11¢/kg, (1) should the seeding recommendation be modified and (2) was it a good idea to stop the experimental work?

2.10. (a) Using the "quick and dirty" formulas of equations (2.1) and (2.2), estimate the means and variances of the random variables depicted in Figures 2.1, 2.3, and 2.5.

(b) Compute the second, third, and fourth central moments of the marginal distribution of wheat price reported in Table 2.1 and describe the distribution in terms of the dimensionless descriptors CV, α_3, and α_4.

REFERENCES AND SELECTED FURTHER READING

Anderson, J. R. 1974. Sparse data, estimational reliability, and risk-efficient decisions. *Am. J. Agric Econ.* 56(3): 564–72.

Barnett, V. 1975. Probability plotting methods and order statistics. *Appl. Statist.* 24(1):95–108.

Bohrnstedt, G. W., and A. S. Goldberger. 1969. On the exact variance of products of random variables. *J. Am. Statist. Ass.* 64(328): 1439–42.

Carter, H. O., and G. W. Dean. 1960. Income, price and yield variability for principal California crops and cropping systems. *Hilgardia* 30(6): 183–93.

de Finetti, B. 1964. Foresight: Its logical laws, its subjective sources. In *Studies in Subjective Probability,* H. E. Kyburg and H. E. Smokler, eds., pp. 93–158. New York: Wiley.

An important historical essay in the development of the theory of subjective probability by one of its leading exponents and modern pioneers.

———. 1972. *Probability and Induction: The Art of Guessing.* New York: Wiley.

A series of papers following on from the previous reference. Written in a conversational style but advanced in content.

———. 1974. *Theory of Probability.* New York: Wiley.

In inimitable style, this two volume text provides the best and most thorough introduction to subjective probability.

Elderton, W. P., and N. L. Johnson. 1969. *Systems of Frequency Curves.* Cambridge: Cambridge Univ. Press.

Presents a thorough discussion of the fitting of Pearson and other types of frequency curves and surfaces to empirical data.

Francisco, E. M., and J. R. Anderson. 1972. Chance and choice west of the Darling. *Australian J. Agric. Econ.* 16(2): 82–93.

An empirical study looking at farmers' appreciation of probabilities and their elicitation. Uses the visual impact method.

Hampton, J. M., P. G. Moore, and H. Thomas. 1973. Subjective probability and its measurement. *Jl. R. Statist. Soc.* Ser. A, 136(1): 21–42.

Presents an application-oriented critical survey of the literature on subjective probability measurement. Valuable also for its extensive list of references.

Hendrickson, A. D., and R. J. Buehler. 1972. Elicitation of subjective probabilities by sequential choices. *J. Am. Statist. Ass.* 67(340): 880–83.

Develops a sequential elicitation procedure with built-in incentives for honesty.

Hogarth, R. M. 1975. Cognitive processes and the assessment of subjective probability distributions. *J. Am. Statist. Ass.* 70(350): 271–91

A highly recommended review with an extensive bibliography.

Huber, G. P. 1974. Methods for quantifying subjective probabilities and multiattribute utilities. *Decn. Sci.* 5(4):430–58.

Complementary to Hampton et al. (1973).

Johnson, N. L., and S. Kotz. 1970. *Distributions in Statistics,* 3 vols. New York: Houghton Mifflin.

An excellent reference source for discrete and univariate continuous distributions.

Julian, P. R., and A. H. Murphy. 1972. Probability and statistics in meteorology: A review of some recent developments. *Bull. Am. Met. Soc.* 53(4): 957–65.

Outlines the potential role of subjective probability in weather forecasting.

Lindley, D. V. 1971. *Making Decisions.* London: Wiley.

A thorough introductory treatment of most of the probability calculus we have assumed herein.

Mardia, K. V. 1970. *Families of Bivariate Distributions.* London: Griffin.

Outlines the algebra, potential usefulness, and fitting of a great variety of bivariate probability distributions.

Mood, A. M., and F. A. Graybill. 1963. *Introduction to the Theory of Statistics*. New York: McGraw-Hill.

Moore, P. G. 1972. *Risk in Business Decision*. London: Longman.

Chapter 9 presents practical discussion of probability elicitation.

Ord, J. K. 1972. *Families of Frequency Distributions*. New York: Hafner.

Pearson, E. S., and J. W. Tukey. 1965. Approximate means and standard deviations based on distances between percentage points on frequency curves. *Biometrika* 52(3–4): 533–46.

Perry, C., and I. D. Greig. 1975. Estimating the mean and variance of subjective distributions in PERT and decision analysis. *Mgmt. Sci.* 21(12): 1477–80.

Phillips, J. B. 1971. Statistical methods in systems analysis. In *Systems Analysis in Agricultural Management*, J. B. Dent and J. R. Anderson, eds. pp. 34–52. Sydney, Australia: Wiley.

Discusses computer procedures for the fitting of probability distributions.

Phillips, L. D. 1970. The "true probability" problem. *Acta Psychol.* 34 (2/3): 254–64 (*Subjective Probability* issue).

Useful discussion of psychological aspects of probability including the experimental work on conservatism of Ward Edwards and others.

Pouliquen, L. Y. 1970. *Risk Analysis in Project Appraisal*. Baltimore: Johns Hopkins Press.

Chapter 7 gives very practical discussion of the choice of a subjective probability distribution. Those discussed include the step rectangular, uniform, beta, trapezoidal, triangular, and normal distributions.

Pratt, J. W., H. Raiffa, and R. Schlaifer. 1965. *Introduction to Statistical Decision Theory*. New York: McGraw-Hill.

The assessment of joint distributions is discussed in Chapters 21 and 22. Unfortunately this comprehensive monograph was only published in a limited preliminary edition and is difficult to obtain.

Raiffa, H. 1968. *Decision Analysis*. Reading, Mass.: Addison-Wesley.

Chapter 10 contains an excellent overview of the alternative concepts of and controversies concerning probability.

Raiffa, H., and R. Schlaifer. 1961. *Applied Statistical Decision Theory*. Boston: Harvard Business School.

Ramsey, F. P. 1964. Truth and probability. In *Studies in Subjective Probability*, H. E. Kyburg and H. E. Smokler, eds., pp. 61–92. New York: Wiley.

A reprint of the pathbreaking work (initially published posthumously in 1931) of this early pioneer theoretician. Mainly of historical interest.

Savage, L. J. 1954. *Foundations of Statistics*. New York: Wiley.

The classical work in rigorous enumeration of the principles of subjective probability.

———. 1971. Elicitation of personal probabilities and expectations. *J. Am. Statist. Ass.* 66(336): 783–801.

Published posthumously, this comprehensive and difficult treatment is the master's last work on the subject.

Savage, L. J., I. Hacking, and A. Shimony. 1967. A panel discussion of personal probability. *Philosophy Sci.* 34: 305–32.

With two philosophers, Savage discusses aspects of the personalistic view where he feels modification and improvement may be needed.

Schlaifer, R. 1959. *Probability and Statistics for Business Decisions*. New York: McGraw-Hill.

Rather statistical in orientation but particularly notable for its discussion of data analysis in probability assessment. More than any other work, this text led to the acceptance of Bayesian procedures in business decision making.

———. 1969. *Analysis of Decisions under Uncertainty*. New York: McGraw-Hill.

Good coverage of procedures for the assessment of judgments about chances and the systematic processing of evidence and expert opinion. Far more oriented to practical application in a management context than his 1959 text.

Spetzler, C. S., and C-A. S. Staël von Holstein. 1975. Probability encoding in decision analysis. *Mgmt. Sci.* 22(3):340–58.

Outlines elicitation and interview procedures for the determination of subjective probabilities. Recommended.

Staël von Holstein, C-A. S. 1970. *Assessment and Evaluation of Subjective Probability Distributions*. Stockholm: Econ. Res. Inst.

A thorough and modern review of the whole field of subjective probability with complete discussion of elicitation methods and scoring rules.

———, ed. 1974. *The Concept of Probability in Psychological Experiments*. Dordrecht: Reidel.

Contains excellent contributions by de Finetti, Kahneman, Tversky, de Zeeuw, Winkler, and Murphy on the problems of experimental research on subjective probability.

Tversky, A., and D. Kahneman. 1975. Judgment under uncertainty: heuristics and biases. In *Utility, Probability, and Human Decision Making*, D. Wendt and C. Vlek, eds., pp. 141–62. Dordrecht: Reidel.

Winkler, R. L. 1967. The assessment of prior distributions in Bayesian analysis. *J. Am. Statist. Ass.* 62(320): 776–800.

———. 1971. Probabilistic prediction: Some experimental results. *J. Am. Statist. Ass.* 66(336): 675–85.

These last two very readable references describe some of Winkler's interesting work on empirical elicitation of degrees of belief.

———. 1972. *An Introduction to Bayesian Inference and Decision*. New York: Holt, Rinehart and Winston.

In this complete yet most readable and mathematically undemanding of statistics texts, Chapters 3 and 4 discuss the elicitation of subjective priors and likelihoods for discrete and continuous distributions. Very highly recommended.

The following references give a (subjective) sampling of some of the rollicking debate that has occurred on subjective probabilities:

Bross, I. D. J. 1969. Applications of probability: Science vs. pseudoscience. *J. Am. Statist. Ass.* 64(325): 51–57. Reply by I. J. Good, pp. 61–66.

Hartley, H. O. 1963. In Dr. Bayes' consulting room. *Am. Statist.* 17(1): 22–24. Comment by R. Schlaifer 17(2): 36. Reply by H. O. Hartley 17(4): 31.

CHAPTER THREE
REVISION OF PROBABILITIES

ALTHOUGH we have emphasized the subjective nature of probability judgments, we have also noted that a reasonable person would wish to take account of any relevant evidence in forming his probability assessments. In a similar fashion, it is rational to update one's probabilities as new evidence or additional information becomes available. Such probability revision can be accomplished using Bayes' theorem. The essential feature of Bayes' theorem in a decision analysis context is that it provides a logical mechanism for learning from experience.

3.1 REVIEW OF BAYES' THEOREM

Bayes' theorem is a noncontroversial elementary theorem of probability derived originally by the eighteenth-century English clergyman Thomas Bayes. This theorem, which is at the heart of procedures for probability revision in decision analysis, is developed in introductory courses of statistics to which we assume you have already been exposed. We will simply sketch the theorem and illustrate its use.

We have already introduced our notation of $P(\Theta_i)$ for prior probabilities of states, $P(z_k \mid \Theta_i)$ for likelihoods, and $P(\Theta_i \mid z_k)$ for posterior probabilities. In these terms the discrete form of Bayes' theorem may be expressed by

$$P(\Theta_i \mid z_k) = \frac{P(\Theta_i)P(z_k \mid \Theta_i)}{\sum_i P(\Theta_i)P(z_k \mid \Theta_i)} = \frac{P(\Theta_i, z_k)}{P(z_k)} \qquad (3.1)$$

Stated verbally, the first of these formulas says that the posterior probability of the ith state, given that the kth prediction has been made, is equal to the product of (1) the prior probability of the state and (2) the likelihood probability of the prediction given the state, divided by all such products summed over all the states. As the second formula indicates, the numerator of the right side is by definition just the joint probability of Θ_i and z_k, while the denominator is the marginal or unconditional probability of occurrence of the particular prediction z_k. We can think of Bayes formula most generally as posterior probability (density) being proportional to prior probability (density) times likelihood. The proportionality factor is the denominator in the above formulas and serves to normalize the numerator so as to ensure that the revised probabilities lie between 0 and 1 and sum to unity over all the states. This normalizing process is rather simple in the discrete case but may involve some awkward integrations in the continuous case.

The logical validity of Bayes' theorem can be demonstrated most simply by reverting to a simple two-state situation and imposing a specific prediction z_k. Visually, as shown in Figure 3.1, we might distinguish the two state probabilities as contiguous rectangles and the prediction probability as an oval lying across their common boundary. Bayes' theorem hinges on the definition of conditional probability, viz., $P(A \mid B) = P(A \text{ and } B)/P(B)$. Thus, relative to Figure 3.1, we can write definitionally,

$$P(z_k \mid \Theta_1) = P(z_k \text{ and } \Theta_1)/P(\Theta_1) \tag{3.2}$$

so that

$$P(z_k \text{ and } \Theta_1) = P(\Theta_1)P(z_k \mid \Theta_1) \tag{3.3}$$

and analogously,

$$P(z_k \text{ and } \Theta_2) = P(\Theta_2)P(z_k \mid \Theta_2) \tag{3.4}$$

The marginal probability of z_k is found by summing its exclusive and exhaustive joint probabilities with the states,

$$P(z_k) = P(z_k \text{ and } \Theta_1) + P(z_k \text{ and } \Theta_2) \tag{3.5}$$

Again definitionally,

$$P(\Theta_1 \mid z_k) = P(z_k \text{ and } \Theta_1)/P(z_k) \tag{3.6}$$

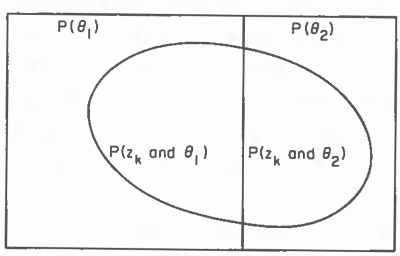


FIG. 3.1. Visual representation of probabilities with two states Θ_1 and Θ_2 and a specific prediction z_k.

where the numerator is specified in equation (3.3) and the denominator in (3.5). Making these substitutions into (3.6), and also (3.3) and (3.4) into (3.5), we have the simplest form of Bayes' theorem,

$$P(\theta_1 \mid z_k) = \frac{P(\theta_1)P(z_k \mid \theta_1)}{P(\theta_1)P(z_k \mid \theta_1) + P(\theta_2)P(z_k \mid \theta_2)}$$

which generalizes straightforwardly to the result given in (3.1).

Bayes formula calculations can be illustrated most simply by harking back to our introductory example in Section 1.2 and computing the posterior probabilities for the states, given that the z_1 prediction has been received. The calculations simply follow the formula but are most conveniently laid out in tabular form as shown in Table 3.1. Note that when summed over the states, both the prior and the posterior probabilities must add up to 1.0, but there is no such requirement on the likelihood and joint probabilities. For the two-state case there is need only to calculate one posterior probability per prediction since the other is complementary; in general, however, all should be calculated so that their sum to unity provides a check on our arithmetic.

TABLE 3.1. Illustration of Bayes Formula Calculations with Two States and a Given Forecast

State θ_i	Prior $P(\theta_i)$	Likelihood $P(z_1 \mid \theta_i)$	Joint $P(z_1 \text{ and } \theta_i)$	Posterior $P(\theta_i \mid z_1)$
θ_1	0.8	0.6	(0.8)(0.6) = 0.48	0.48/0.54 = 0.89
θ_2	0.2	0.3	(0.2)(0.3) = 0.06	0.06/0.54 = 0.11
	1.0		$P(z_1)$ = 0.54	Check = 1.00

We now consider a slightly more extensive application of Bayes' theorem. Suppose Kanzo Makame, a Japanese businessman, is contemplating investing in a new pearl farming venture in the Patterson Islands. The venture is risky since the technology is unproven under local conditions. Kanzo decides to summarize his feelings of uncertainty in terms of three levels of success: θ_1 = good, θ_2 = fair, and θ_3 = poor. He attaches the following prior probabilities to these states: $P(\theta_1) = 0.4$, $P(\theta_2) = 0.2$, and $P(\theta_3) = 0.4$. These probabilities are respectively represented by the segments *AB, BC,* and *CD* of the horizontal axis of the "probability square" in Figure 3.2.

Kanzo has a friend, Joe Nagasaki, who is a renowned expert on pearl farming. Kanzo plans to ask Joe for his opinion of the proposed scheme. He expects to be able to classify Joe's answer as either z_1 = favorable or z_2 = unfavorable. Kanzo expresses his confidence in Joe's judgment by means of

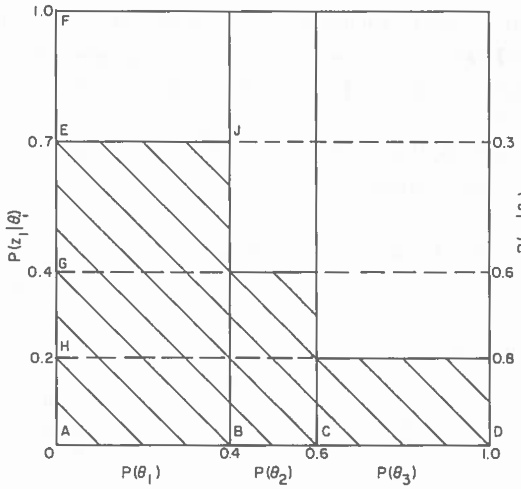


FIG. 3.2. Probability square illustrating the application of Bayes' theorem.

the following likelihood probabilities:

$$P(z_1 \mid \Theta_1) = 0.7 \qquad P(z_2 \mid \Theta_1) = 0.3$$
$$P(z_1 \mid \Theta_2) = 0.4 \qquad P(z_2 \mid \Theta_2) = 0.6$$
$$P(z_1 \mid \Theta_3) = 0.2 \qquad P(z_2 \mid \Theta_3) = 0.8$$

These probabilities are also shown in Figure 3.2, being measured on the vertical axes. Thus relative to Θ_1, represented by the left section of the square, $P(z_1 \mid \Theta_1)$ is represented by the length AE, and $P(z_2 \mid \Theta_1)$ by the length EF. Likewise, we have $P(z_1 \mid \Theta_2) = AG$ and $P(z_1 \mid \Theta_3) = AH$. Kanzo wants to know what posterior probabilities he should apply to the three states in the light of Joe's comments. Moreover, he wants to see how he should revise his opinions *before* he knows Joe's views; so he must apply Bayes' theorem for both possible predictions, selecting the relevant posterior probabilities *after* the prediction is obtained.

Recalling from equation (3.1) that Bayes' theorem can be written as

$$P(\Theta_i \mid z_k) = [P(\Theta_i)P(z_k \mid \Theta_i)]/P(z_k)$$

and noting that in this case there are three possible states and two possible forecast signals, it is apparent that the formula must be applied $3 \times 2 = 6$ times to obtain the six posterior probabilities. For example, for Θ_1 and z_1 we have $P(\Theta_1) = 0.4$, represented by the distance AB in Figure 3.2, and $P(z_1 \mid \Theta_1) = 0.7$, represented by the distance AE. The joint probability $P(\Theta_1)P(z_1 \mid \Theta_1) = (0.4)(0.7) = 0.28$ and is represented by the area $ABJE$.

The denominator in Bayes formula, $P(z_1)$, is equal to the sum of the joint probabilities $P(\theta_i)P(z_1 \mid \theta_i)$ for all three states and is thus equal to the whole shaded area in Figure 3.2. Hence $P(z_1) = (AB)(AE) + (BC)(AG) + (CD)(AH) = (0.4)(0.7) + (0.2)(0.4) + (0.4)(0.2) = 0.44$. Thus $P(\theta_1 \mid z_1) = 0.28/0.44 = 0.636$. Table 3.2 shows a convenient way of setting out all the calculations.

TABLE 3.2. Calculation of Posterior Probabilities

θ_i	$P(\theta_i)$	$P(z_k \mid \theta_i)$		$P(\theta_i)P(z_k \mid \theta_i)$	
		z_1	z_2	z_1	z_2
θ_1	0.4	0.7	0.3	0.28	0.12
θ_2	0.2	0.4	0.6	0.08	0.12
θ_3	0.4	0.2	0.8	0.08	0.32
			$P(z_k) =$	0.44	0.56
			$P(\theta_1 \mid z_k) =$	0.636	0.214
			$P(\theta_2 \mid z_k) =$	0.182	0.214
			$P(\theta_3 \mid z_k) =$	0.182	0.572
			$\sum_i P(\theta_i \mid z_k)$	1.000	1.000

From these results Kanzo can see that if Joe's forecast is favorable, his chance of the venture proving to be an outstanding success is better than 60%, coupled with a less than 20% probability of marginal success. On the other hand, the probabilities are almost reversed if Joe's prognosis is gloomy.

We have assumed thus far that only one piece of predictive information is to be accounted for in the probability revision. But Bayes' theorem can be used in repeated fashion to aggregate any number of pieces of probabilistic information. Thus, after observing t pieces of information $I_1, \ldots, I_t$, the tth revision of Bayes' theorem is

$$P(\theta_i \mid I_1, \ldots, I_t) = \frac{P(\theta_i \mid I_1, \ldots, I_{t-1})P(I_t \mid \theta_i, I_1, \ldots, I_{t-1})}{\sum_i P(\theta_i \mid I_1, \ldots, I_{t-1})P(I_t \mid \theta_i, I_1, \ldots, I_{t-1})} \tag{3.7}$$

Here $P(\theta_i \mid I_1, \ldots, I_{t-1})$ represents the posterior probability of θ_i relative to the first $t - 1$ pieces of information and constitutes the prior probability relative to the tth piece. On the other hand, $P(\theta_i \mid I_1, \ldots, I_t)$ is a posterior probability relative to all t pieces of information. In this form of Bayes' theorem the likelihood $P(I_t \mid \theta_i, I_1, \ldots, I_{t-1})$ is the probability of observing I_t, given that θ_i is the true state and given the $t - 1$ pieces of information already received.

The form of Bayes' theorem in equation (3.7) is instructive since it emphasizes the dynamic or sequential nature of probability revision in the light of accruing information. However, rather than processing information piece by piece as it becomes available, it may be accumulated and then processed by a single application of Bayes' theorem. If this procedure is followed, the relevant form of Bayes' theorem is:

$$P(\Theta_i \mid I_1, \ldots, I_t) = \frac{P(\Theta_i)P(I_1, \ldots, I_t \mid \Theta_i)}{\sum_i P(\Theta_i)P(I_1, \ldots, I_t \mid \Theta_i)} \tag{3.8}$$

Algebraically, equations (3.7) and (3.8) are equivalent. In theory they should lead to the same posterior probability when applied to the same set of information. In practice, however, they may often lead to different estimates because the likelihoods used in (3.7) may be out of line with those used in (3.8). The difficulty is that likelihoods, like all other probabilities, are subjective. In making these judgments, particularly when more than one or two pieces of information are involved, different answers are likely to arise because of the difference in form of the "one-shot" likelihood of (3.8) $P(I_1, \ldots, I_t \mid \Theta_i)$ and the "extended-form" likelihoods of the type $P(I_t \mid \Theta_i, I_1, \ldots, I_{t-1})$ used in (3.7). As a practical matter when dealing with a series of pieces of information, both approaches could be applied. Any major difference in the results would signal an inconsistency on the basis of which the likelihoods used should be reassessed until a satisfactory degree of consistency is achieved.

In the extended form of Bayes' theorem given in equation (3.7), the likelihoods $P(I_t \mid \Theta_i, I_1, \ldots, I_{t-1})$ are conditional upon both the state Θ_i and the previously observed information. A simplifying assumption often made is that each piece of information is independent of any other, i.e., $P(I_t \mid \Theta_i, I_1, \ldots, I_{t-1}) = P(I_t \mid \Theta_i)$. If this assumption holds for all values of i and t, the information sources are said to be *conditionally independent*.

The assumption of conditional independence is often convenient since it may simplify the determination of likelihoods. In particular, this will be the case if it can be assumed that the information-generating process follows one of a number of well-known statistical models, such as a Bernoulli, Poisson, or independent normal process, in which conditionally independent trials are assumed.

Overall, the most important feature of Bayes' theorem is that it provides a logical mechanism for the consistent processing of additional information. Many experiments by psychologists have shown that man's intuition is an inefficient basis for such processing. When acting on the basis of intuition, the great majority of people exhibit conservatism. They do not extract as much information from the available evidence as they should. In this sense Bayes' theorem is an efficient procedure for probability revision. It ensures that full value or weight consistent with the logic of the laws of probability is given to any additional information that becomes available.

3.2 YES/NO TRIALS AND THE BINOMIAL DISTRIBUTION

While there are many particular types of probability distributions, one of these—the binomial distribution—is particularly pertinent to managerial decision making. It is relevant because management is often concerned with processes that have outcomes classifiable into only two categories such as yes or no, dead or alive, defective or not defective, success or failure, etc. These are known as *Bernoulli processes* if (1) there are only two possible types of outcome, (2) each type of outcome has a constant chance of occurrence, and (3) each outcome is independent of previous outcomes. For example, the continuous tossing of a fair coin constitutes a Bernoulli process because on each trial (i.e., on each toss) there are only two possible types of outcome, head or tail, each of which has a constant probability 0.5 of occurring, and the outcome of each toss is uninfluenced by what happened in previous tosses. In general, the two possible outcomes of a Bernoulli process are referred to as "success" and "failure."

Business examples of Bernoulli processes are common. Each person interviewed in a market survey might be classified as a buyer or nonbuyer. Each ewe in a flock may be in lamb or not. Vouchers examined by an auditor may be correct or incorrect. Fruit harvested by a mechanical picker may be damaged or not damaged. Of course, business processes may not meet exactly the definition of a Bernoulli process; e.g., the chance of damage to fruit by a mechanical picker may not be constant but may increase over time due to wear and tear on the machine. Usually, however, the effect of such departures from the theoretical ideal is only minor in terms of problem solution and can be ignored.

The relevance of the binomial probability distribution is that it gives the probability $P(r \mid n, p)$ of having r successes in a series of n outcomes from a Bernoulli process where p is the probability of success on any given trial. The general formula for calculating this probability is

$$P(r \mid n, p) = [n!/r!(n - r)!]p^r(1 - p)^{n-r}$$

where $n!$ denotes factorial n, i.e. the product $n(n - 1)(n - 2)\ldots(2)(1)$. Conversely, since success and failure are the only possible outcomes, r successes in n trials imply $n - r$ failures, hence $P(n - r \mid n, 1 - p) = P(r \mid n, p)$.

As an example of a binomial probability, suppose past experience leads us to judge that 5% of the heifers in our herd will not be in calf. If we sell 15, what is the probability that exactly 13 will be in calf (or, equivalently, that only 2 will not be in calf)? The answer is

$$P(r = 13 \mid n = 15, p = 0.95) = (15!/13!2!)(0.95)^{13}(0.05)^2 = 0.1348$$

Similarly, the probability that all 15 will be in calf is

$$P(r = 15 \mid n = 15, p = 0.95) = (15!/15!0!)(0.95)^{15}(0.05)^0 = (0.95)^{15} = 0.4633$$

Theoretically, the binomial distribution is only relevant for trials from infinite populations. This requirement will generally not be met in business situations. While there is certainly an infinite sequence of possible coin tosses, there will not be an infinite set of vouchers for an auditor to check for correctness or incorrectness, nor will there be an infinite set of fruit to be damaged or not damaged by a mechanical picker, and so on. However, unless the series of trials or sample we are considering constitutes a substantial proportion (say 20% or more) of the population, the binomial distribution serves as an adequate approximation—it does not introduce untoward errors into the analysis.

The context in which binomial probabilities are most relevent in business decisions is when the decision problem involves (or more accurately, is subjectively judged to involve) a Bernoulli process about which we seek further information by taking a sample of outcomes. Information from this sample can then be used via Bayes' theorem to revise our prior probabilities. To this end, tables of the binomial distribution for various values of r, p, and n have been published. Winkler (1972), for example, gives a table of $P(r \mid n, p)$ values for $n = 1, 2, \ldots, 20, 50, 100$; $r = 0, 1, 2, \ldots, n$; and $p = 0.01, 0.02, \ldots, 0.99$.

As an example of the use of binomial probabilities in a decision context, suppose we are considering the purchase of 200 Hereford heifers that have recently been artificially inseminated with semen of a new beef breed. Our crudely stated prior judgment is that the fraction of heifers not in calf may run from 1% to 5% with the probabilities $P(\Theta_i)$ shown in the second column of Table 3.3. In this context the possible event Θ_i corresponds to the probability p of a Bernoulli process where "success"

TABLE 3.3. Probability Revision with Binomial Probabilities

Θ_i	$P(\Theta_i)$	$P(r = 0 \mid \Theta_i)$	$P(\Theta_i)P(r = 0 \mid \Theta_i)$	$P(\Theta_i \mid r = 0)$
1%	0.1	0.9044	0.0904	0.1231
2%	0.2	0.8171	0.1634	0.2225
3%	0.3	0.7374	0.2212	0.3012
4%	0.3	0.6648	0.1994	0.2716
5%	0.1	0.5987	0.0599	0.0816
			$P(r = 0) = 0.7343$	1.0000

relates to a heifer not being in calf. Further suppose a pregnancy test is run on a sample of 10 heifers and indicates that all 10 are pregnant. Using tables of the binomial distribution, we can then ascertain $P(r = 0 \mid n = 10, p)$ for p equal to each value of Θ_i. These are our likelihood probabilities previously designated $P(z_k \mid \Theta_i)$ and are shown in the third column of Table 3.3.

In the binomial context, z_k corresponds to possible values of r, and Θ_i to the values of p. Note that in the present case r equals the number of nonpregnant heifers since this is what p refers to. Applying Bayes' theorem in the form

$$P(\Theta_i \mid r) = P(\Theta_i)P(r \mid \Theta_i)/\textstyle\sum_i P(\Theta_i)P(r \mid \Theta_i)$$

we can derive our revised probabilities for the probability distribution of the number of nonpregnant heifers. The calculations can be carried out as shown in Table 3.3, where the values for $P(r = 0 \mid \Theta_i)$ are from published tables. The revised probabilities $P(\Theta_i \mid r = 0)$ would then be used as required for decision analysis.

3.3 BAYES' THEOREM AND CONTINUOUS DISTRIBUTIONS

In many decision problems it is reasonable and convenient to assume that the uncertain quantity of interest can be represented by a continuous rather than a discrete probability distribution. In practice, measurement procedures are such that observed random variables are always discrete. However, where measurement is precise enough, it is not unrealistic to assume continuity. Moreover, continuous probability models often have the advantage of being mathematically easier to work with than discrete models, frequently eliminating the large amount of arithmetic associated with manipulation of a discrete random variable that can take many possible values. The assessment of probability density functions (PDFs) for such continuous variables has already been discussed (Section 2.3). We now turn to the application of Bayes' theorem to such probability models.

Suppose we have some uncertain quantity X for which a prior PDF, $f(X)$, has been assessed. We now observe some value Y of the sample statistic $\tilde{Y}$ bearing on X. Bayes' theorem for this continuous case can be written as

$$f(X \mid Y) = \frac{f(X)f(Y \mid X)}{\int f(X)f(Y \mid X)\,dX} \tag{3.9}$$

The PDFs $f(X \mid Y)$ and $f(X)$ represent the posterior distribution and the prior distribution of X respectively, while $f(Y \mid X)$ is the likelihood function. Equation (3.9) can therefore be written in words as

$$\text{posterior density} = \frac{(\text{prior density})(\text{likelihood})}{\int (\text{prior density})(\text{likelihood})}$$

The parallel with the discrete form of the formula is apparent, the only difference being that the summation sign in the denominator for the discrete case is replaced by an integration sign.

Although conceptually equation (3.9) provides a convenient way to revise a PDF to account for sample information, in practice difficulties may arise in integrating the denominator. Unless $f(X)$ and $f(Y \mid X)$ are fairly simple mathematical functions, it may be necessary to resort to advanced mathematical methods to solve the equation. To avoid this difficulty, Bayesian statisticians have developed the concept of *conjugate prior distributions*. Once certain assumptions are made about the nature of the statistical sampling process, the likelihood function is uniquely determined. For any likelihood function it may then be possible to define a family of conjugate prior distributions. Each member of this set of distributions can be combined, without too much difficulty, with the given likelihood function to give the posterior distribution. Winkler (1972, Ch. 4) gives examples.

Three properties sought in conjugate families are that the distributions should be mathematically tractable, rich enough to describe the prior information adequately, and readily interpretable by an informed decision maker. The prime requirement of tractability is satisfied (1) if it is reasonably straightforward to compute the posterior distribution with Bayes' theorem; (2) if the resulting posterior distribution is also a member of the same conjugate family so that repeated use of the theorem is possible; and (3) if moments of the prior and the posterior distribution, such as the means and the variances, are computable.

Conjugate distributions have been found for many common sampling distributions. For example, if sampling is from a Bernoulli process, the set of conjugate distributions is the family of beta distributions. A second example is if sampling is from a normal distribution with known variance, in which case the conjugate family is the family of normal distributions. The latter case is discussed in the following section.

3.4 BAYES' THEOREM AND THE NORMAL DISTRIBUTION

Decision problems will often involve an array of possible states that are continuous or virtually continuous. Although such situations can be handled by breaking the state possibilities into discrete intervals, it may sometimes be more convenient to carry out the analysis in continuous form. This is particularly so if there are only two alternative acts, if the state variable is judged to follow a normal distribution, and if it is possible to obtain further information by taking a sample (running an experiment) that is not too small.

We might be interested in the expected per store level of sales of a new food product that could be sold in 10,000 retail outlets. We will define x as the sales (in kg) per store over the period of interest, where x is a continuous random variable with mean μ and standard deviation σ. We suppose that the decision problem calls only for an estimate of μ, the mean level of sales per store, and that the decision maker can specify his prior probability as a normal distribution $\mathcal{N}(\mu_0, \sigma_0)$, where μ_0 denotes his prior mean for the distribution of the population mean μ and σ_0 is the standard deviation of this normal distribution. Since σ is assumed known, σ_0 implies an "equivalent sample size" n_0, associated with the prior information and defined by the relationship

$$\sigma_0^2 = \sigma^2/n_0 \tag{3.10}$$

Although the decision maker may not know what a normal distribution is, he may suggest a symmetrical distribution with some mean and a percentage range about the mean; e.g., "the average level of sales per store will be about 200 kg and there is a 60% chance that it will lie between 180 and 220 kg." Making use of the tables for the standard normal distribution $\mathcal{N}(0, 1)$ found in most statistics texts, such "mean and range" statements can be used to derive σ_0 and hence full specification of the prior distribution. The above statement, for example, implies $\sigma_0 = 23.75$ since P (average sales per store $\leqq 220) = 0.8$; and from tables of the cumulative standard normal distribution we find that for $F(d) = 0.8$, $d = 0.842$, so that $\sigma_0 = (220 - 200)/0.842 = 23.75$. Thus the prior distribution is $\mathcal{N}(\mu_0, \sigma_0) = \mathcal{N}(200, 23.75)$.

Now suppose it is possible to test a sample and hence gain further information about the likely value of the state variable. We might take a sample of (i.e., run an experiment with) n stores to test sales of the new food product (or we might merely ask the managers of n stores for their best estimate of sales of the new food product). This would give us a sample mean m and a sample standard deviation s. As long as n, the size of the sample, is no less than about 30, the distribution of m in repeated samples of the same size will be normal. Suppose we test our new food product in 100 stores and find $m = 210$ and $s = 110$. With this sample information we can now revise our prior normal distribution. Conveniently, if the prior distribution is normal and if the distribution of the sample mean is normal, the posterior distribution will also be normal. In such cases the mean μ_1 and standard deviation σ_1 of the posterior distribution for the population mean are easily ascertained (Schlaifer, 1959) as

$$\mu_1 = (\sigma_0^{-2}\mu_0 + \sigma_m^{-2}m)/(\sigma_0^{-2} + \sigma_m^{-2}) = (\mu_0 n_0 + mn)/(n_0 + n) \tag{3.11}$$

$$\sigma_1 = [\sigma_0^2\sigma_m^2/(\sigma_0^2 + \sigma_m^2)]^{0.5} = \sigma/(n_0 + n)^{0.5} \qquad (3.12)$$

where the standard error of the sample mean σ_m is calculated as

$$\sigma_m = \sigma/n^{0.5} \qquad (3.13)$$

if the population standard deviation is known from previous experience. If σ is not known, an estimate of it is given by s, the sample standard deviation, so that for our present example we have $\sigma_m = s/n^{0.5} = 110/10$. Likewise, if required, we can also use s instead of σ in equation (3.10) to estimate the equivalent prior sample size. Thus we have $n_0 = s^2/\sigma_0^2 = 21.45$.

Equations (3.11) and (3.12) correspond to the application of Bayes' theorem with a normal prior and normal sample mean distribution. Note that n, the sample size, influences σ_m and hence μ_1 and σ_1. As shown later in Section 5.6, this fact gives a basis for the evaluation of alternative sized samples.

Using equations (3.11) and (3.12), we can obtain the revised distribution for the mean level of sales per store of our new food product:

$$\mu_1 = [(200)(21.45) + (210)(100)]/(21.45 + 100) = 208.2$$
$$\sigma_1 = 110/(21.45 + 100)^{0.5} = 10$$

Data for this particular example are given in Table 3.4.

TABLE 3.4. Data for Store Example

Distribution of Mean Sales per Store	Mean	Standard Deviation
Prior	$\mu_0 = 200$	$\sigma_0 = 23.75$
Sample ($n = 100$)	$m = 210$	$\sigma_m = 11$
Posterior	$\mu_1 = 208.2$	$\sigma_1 = 10$

When such revised probabilities have been calculated, they may then be used as required in decision analysis. As shown in Section 5.6, significant shortcuts in analysis are available if the prior and posterior distributions are normal.

PROBLEMS

3.1. Comment on the following statements:

(a) Not only prior probabilities but also likelihoods and even posterior probabilities may be chosen on the basis of introspective judgment.

(b) The basic failing of Bayesian decision theory is that it assumes the decision situation is one of uncertainty, not one of the unknown. States given

a prior probability of zero, either because of prior belief or because of ignorance of their possible existence, are forever condemned to oblivion.
(c) The one set of probabilities may serve as both posteriors and priors.
(d) "Either one starts with initial probabilities and everything proceeds according to the logic of probability, or else one starts with nothing and remains forever unable to say anything" (de Finetti, 1972, p. 194).
(e) Bayes' theorem provides a means of combining hunches with hard facts.
(f) "Gut-feeling" decision makers know only two degrees of probability, zero and one. Decision analysts, in contrast, recognize every degree of probability except zero and one.

3.2. Ron Fisher, an agricultural extension officer, believes his district will face a locust plague in 2 years out of 20. Past research has shown that if a plague occurs, locusts are to be seen locally by early May 5 times out of 10; but they are only to be seen 1 time out of 10 if a plague does not occur. Given that Ron sees locusts in early May, what is the probability that his district faces a plague?

3.3. (a) Assume that weather on any day in your region can be classified as fine, showery, or wet. What likelihoods would you associate with possible forecasts of fine, showery, or wet made this morning by your weather bureau for tomorrow? How do you justify your choices? Would farmers be better served if forecasts were made in probabilistic terms?
(b) Design an experiment to test for conservatism in the intuitive processing of information.

3.4. A certain disease is carried without any obvious sign by 0.1% of cattle. Veterinary procedures have been developed to test for carriers of the disease. The test results are positive about 80% of the time for carriers, and are negative about 90% of the time for noncarriers. If a positive result is obtained from testing one beast drawn at random from a herd of 10,000 head, what is the probability that it is indeed a carrier of the disease? As a first step, make an intuitive estimate.

3.5. Mr. O. K. Shambles of "Willigobung" faces a decision problem involving three states of nature for which he has prior probabilities $P(\theta_1) = 0.15$, $P(\theta_2) = 0.30$, and $P(\theta_3) = 0.55$. To gain further information, he consults two experts α and β who each conduct a quite separate experiment for him. The first yields a forecast z_α and the second a forecast z_β. The conditional probabilities for these are respectively

$$P(z_\alpha \mid \theta_1) = 0.30 \qquad P(z_\alpha \mid \theta_2) = 0.50 \qquad P(z_\alpha \mid \theta_3) = 0.10$$
$$P(z_\beta \mid \theta_1) = 0.10 \qquad P(z_\beta \mid \theta_2) = 0.80 \qquad P(z_\beta \mid \theta_3) = 0.60$$

What probabilities should Shambles use in selecting the action to take?

3.6. Your prior probabilities $P(r)$ for spring rain (mm) this year are $P(0 \leq r < 30) = 0.0$, $P(30 \leq r < 90) = 0.2$, $P(r \geq 90) = 0.8$. Now suppose the local weather station makes a probabilistic prediction that there is a 30% chance of $r < 30$ and a 70% chance of $30 \leq r < 90$. Discuss how you would revise your priors.

3.7. Mr. A. R. (Spud) Murphy, B.Ag.Ec., manager of Tingha Tasty Chip Pty. Ltd. and a crackerjack decision analyst, believes that p, the proportion of retailers who will stock a potential new Tasty Chip product line, is either 0.3, 0.35, 0.4, or 0.45. Spud's prior probabilities for these four values are respectively 0.2, 0.3, 0.3, and 0.2. However, he considers it worthwhile to have the Agricultural Business Research Institute of the university undertake a market survey appraisal. The ABRI takes a sample of 19 retailers of whom 6 say they will stock the new line of chips. Calculate Spud's posterior probabilities. What if none in the sample had expressed any interest? What if all 19 had said yes?

3.8. Charlie Dobos, production manager of Tip Top Paprika, has to estimate the average yield per greenhouse of early season seedlings. This estimate is crucial to his planning of land preparation. The seedlings are produced in 500 plastic greenhouses of standard size and under standard controlled conditions. Charlie's prior distribution for average yield is normal with a mean of 10,000 and a standard deviation of 500. A sample of 30 greenhouses drawn at random shows an average yield of 9000. The standard deviation of yields in the sample is 1200. What should Charlie's posterior distribution be? What probability should he assign to the average yield being less than 10,000?

3.9. Gundy Pastoral Company has a one-year option on a development block of 5000 ha in an area of northwestern Australia with an assured water supply. The only apparent potential use of the land is for grain sorghum production aimed at the Japanese market. The option arrangement was entered on the basis of a preliminary investment appraisal that indicated an acceptable expected profit. In making this appraisal the management group of Gundy Pastoral agreed on the following probability distribution for long-term average annual grain yield per hectare.

Yield (t/ha):	6	7	8	9	10	11	12
Probability:	0.05	0.20	0.30	0.20	0.15	0.08	0.02

These yield judgments were based on information from a somewhat analogous climatic area, but with different soil type, in northern Queensland. The uncertainty about yield arose from doubts as to whether local agronomic problems could be overcome, there being no previous experience of sorghum cropping in the region.

As part of the option arrangement, Gundy Pastoral committed itself to run some sorghum trials in the area using the best-bet technology suggested by a consultant. Five blocks each of 40 ha were planted at widely scattered locations across the option area. These gave yields of 7, 9, 9, 10, and 11 t/ha respectively.

Given the above information and assuming you are Gundy Pastoral's management group, what revised yield probabilities would you use in deciding whether or not to take up the option? How did you obtain these posterior probabilities?

REFERENCES AND SELECTED FURTHER READING

Bayes, T. 1958. Essay towards solving a problem in the doctrine of chances. *Biometrika* 45: 293–315. (London, 1763).

de Finetti, B. 1972. *Probability and Induction: The Art of Guessing*. New York: Wiley.

Chapter 9 discusses the role of Bayes' theorem in learning from experience.

Edwards, W. 1968. Conservatism in human information processing. In *Formal Representation of Human Judgment,* B. Kleinmuntz, ed., pp. 17–52. New York: Wiley.

Discusses the evidence from numerous laboratory experiments indicating that individuals are not efficient intuitive processors of information.

Finkelstein, M. O., and W. B. Fairley. 1970. A Bayesian approach to identification evidence. *Harvard Law Rev.* 83(3): 489–517.

Completely nonagricultural but a most interesting discussion of the possible application of Bayes' theorem in the judicial system.

Hadley, G. 1967. *Introduction to Probability and Statistical Decision Theory.* San Francisco: Holden-Day.

Chapter 8 of this excellent text outlines the use of experiments in practical decision analysis.

Moore, P. G. 1972. *Risk in Business Decision.* London: Longman.

In an applied business context, Chapters 7, 10, and 11 discuss Bayes' theorem, binomial sampling, and normal sampling.

Peterson, C. R., and L. R. Beach. 1967. Man as an intuitive statistician. *Psychol. Bull.* 68(1): 29–46.

Discusses the findings of psychologists on the phenomenon of conservatism in the intuitive processing of information.

Pratt, J. W., H. Raiffa, and R. Schlaifer. 1965. *Introduction to Statistical Decision Theory.* New York: McGraw-Hill.

Chapters 11 and 16 provide an advanced treatment of the application of Bayes' theorem to Bernoulli and normal processes.

Raiffa, H., and R. Schlaifer. 1961. *Applied Statistical Decision Theory.* Boston: Harvard Business School.

The pioneer work on the use of conjugate distributions in probability revision. Provides full algebraic treatment of the application of Bayes' theorem to Bernoulli, Poisson, and normal processes. Despite the "applied" in its title, it has no claim to easy reading for the mathematically uninitiated!

Schlaifer, R. 1959. *Probability and Statistics for Business Decisions.* New York: McGraw-Hill.

Part 3 discusses probability revision in a variety of business contexts.

Winkler, R. L. 1972. *An Introduction to Bayesian Inference and Decision.* New York: Holt, Rinehart and Winston.

Chapters 3 and 4 cover the application of Bayes' theorem to discrete and continuous distributions. Very highly recommended.

Winkler, R. L., and A. H. Murphy. 1973. Information aggregation in probabilistic prediction. *IEEE Trans. Syst. Man. Cybern.* SMC-3(2): 154–60.

Outlines the use of Bayes' theorem in the aggregation of multiple pieces of information, discusses conditional dependence, and reviews a variety of experiments on the aggregation process.

CHAPTER FOUR
UTILITY

How DO YOU FEEL now about earning an extra dollar? An extra ten dollars? Would you feel the same if you had just inherited a million dollars from a rich unknown relative? Such nonconstant preference (for money or any other attributes pertinent to evaluating decision consequences) is the subject of the modern theory of utility. This theory plays a key role in decision analysis. We first discuss the general concept of utility, and then review more formally the central theorem of utility analysis known as Bernoulli's principle or, as it is often called, the expected utility theorem. Finally, we treat several aspects of the empirical scaling and algebraic representation of preference or utility.

4.1 CONCEPTS OF UTILITY AND NONLINEAR PREFERENCE

We will focus attention on preference for money outcomes because money is such a convenient common denominator of many aspects of consequences in farm and business decisions. However, the ideas developed are equally applicable to other measurable attributes.

To illustrate the concept of utility, consider the following starkly simple decision problem with two acts and only two subjectively equally likely states (Table 4.1). Possible outcomes are shown as dollar gains or losses.

TABLE 4.1. Simple Decision Problem

θ_i	$P(\theta_i)$	a_1	a_2
		$	$
θ_1	0.5	200	1000
θ_2	0.5	0	−800
Expected money value		100	100

The expected money value of both acts is identical at $100, so that if money payoffs measure consequences for us adequately, we should feel fairly indifferent between choosing a_1 or a_2. But suppose you are faced with this choice, would you be indifferent between a_1 and a_2? Our guess is that you would prefer a_1, although there is certainly nothing inherently right or wrong in either choice. It is simply a matter of personal evaluation. For most, the potential disaster of losing $800 will outweigh the desirability of possibly gaining $1000. Others may be able to afford the adventure more easily and may aspire to the purchasing power inherent in gaining $1000.

Whatever the choice and for whatever reasons, nonindifference illustrates that for you money payoffs do not "tell the whole story" in evaluating consequences. That the whole story should be properly told is important in risky decision making. Because risky choice implies choice between probability distributions of consequences, mental balancing of a number of possible consequences simultaneously is required. In contrast, in decision making under subjective certainty it is only necessary to rank consequences to determine the most desirable choice. Compared to riskless choice, risky choice is intrinsically more difficult. In appraising risky alternatives, utility analysis provides the practical means whereby preferences are crystallized and consistent choice simplified. It should be noted, however, that although risky decisions are more difficult, all of us make them quite painlessly most of the time (although some of the possible consequences, e.g., of road-traversing decisions, are not too painless!).

We shall use the concept of utility in a conditionally normative manner. Having determined the decision maker's preferences for relatively uncomplicated risky consequences, these will be employed via a utility or preference function to aid his decision making in more complex situations, thereby ensuring consistency between the decision maker's preferences and his choices. In this sense utility analysis is conditional on expressed preferences. As we will soon see, a *utility function* is simply a device for assigning numerical utility values to consequences in such a way that a decision maker should act to maximize subjective expected utility if he is to be consistent with his expressed preferences. This implies the use of *Bernoulli's principle* or, as it is otherwise known, the *expected utility theorem*.

4.2 BERNOULLI'S PRINCIPLE

Daniel Bernoulli postulated his principle well over 200 years ago in recognition of the fact that an extra dollar is worth more to a poor man than to a rich man. But its potential went unrecognized until the work of von Neumann and Morgenstern in the 1940s. They showed that Bernoulli's principle is a logical deduction from a small number of postulates or axioms that many people agree are reasonable, at least to the extent that they would wish their own choices to conform with them. More recently, the axioms have been formulated in a variety of ways and under a variety of names, and several people have provided alternative and increasingly elegant proofs of the theorem.

The following set of three axioms is a sufficient basis for deducing Bernoulli's principle for the case of risky prospects with single-dimensioned consequences. For risky prospects with multidimensioned consequences, a slight but reasonable extension of the axioms is necessary (Fishburn, 1970). By a *risky prospect* we mean an act or a possible choice that has a probability distribution of outcomes. The *j*th risky prospect will be denoted by a_j.

1. *Ordering and transitivity.* A person either prefers one of two risky prospects a_1 and a_2 or is indifferent between them. This presumption that people can order prospects is not trivial, as perhaps is illustrated by even the simply structured prospects we have already discussed. The logical extension of ordering is to transitivity of orderings of more than two prospects, e.g., a_1, a_2, and a_3. This implies that if a person prefers a_1 to a_2 (or is indifferent between them) and prefers a_2 to a_3 (or is indifferent between them), he will prefer a_1 to a_3 (or be indifferent between them). Experimental psychologists have demonstrated that subjects are not always perfectly transitive in their choices, particularly where discrimination between prospects is difficult, such as when alternatives are complex. Human imperfections and limitations such as these provide all the more reasons for formalizing preferences via decision analysis so as to minimize inconsistencies.
2. *Continuity.* If a person prefers a_1 to a_2 to a_3, a subjective probability $P(a_1)$ exists other than zero or one such that he is indifferent between a_2 and a lottery yielding a_1 with probability $P(a_1)$ and a_3 with probability $1 - P(a_1)$. This implies that if faced with a risky prospect involving a good and a bad outcome, a person will take the risk if the chance of getting the bad outcome is low enough. Continuity seems to be a reasonable requirement to demand of an orderly thinking person, but the axiom may give operational difficulties when the prospects consist of disparate alternatives. For example, it has been argued that the axiom breaks down when the unfavorable outcome is very bad, e.g., death. And yet we risk death every time we cross the street or drive a car, and often for meager reward.
3. *Independence.* If a_1 is preferred to a_2, and a_3 is any other risky prospect, a lottery with a_1 and a_3 as its outcomes will be preferred to a lottery with a_2 and a_3 as outcomes when $P(a_1) = P(a_2)$. In other words, preference between a_1 and a_2 is independent of a_3. Again, only the sort of practical difficulties of comprehension that lead to intransitivity seem likely to cause problems with this axiom. It says that preferences persist independently of successive probability resolutions in evaluating compound lotteries.

Bernoulli's principle may be deduced from such axioms and may be stated as follows: a utility function exists for a decision maker whose preferences are consistent with the axioms of ordering and transitivity, continuity, and independence; this function U associates a single real number (utility value) with any risky prospect and has the following properties, where we denote the utility value of a_j by $U(a_j)$.

1. If a_1 is preferred to a_2, then $U(a_1) > U(a_2)$ and vice versa.
2. The utility of a risky prospect is its expected utility value. This is obtained by evaluating the expected value of the utility function in terms of the risky prospect's consequences, i.e.,

$$U(a_j) = E[U(a_j)] \tag{4.1}$$

the expectation being based on the decision maker's subjective distribution of outcomes. In the case of discrete outcomes

$$U(a_j) = \textstyle\sum_i U(a_j \mid \Theta_i) P(\Theta_i) \tag{4.2}$$

and in the case of a continuous distribution of outcomes

$$U(a_j) = \int U(a_j \mid \Theta) f(\Theta) d(\Theta) \tag{4.3}$$

Higher moments of utility, e.g., its variance, are not relevant to decision making. Note furthermore that the axioms logically imply use of the decision maker's subjective probability distribution for utility evaluation of the risky prospect's outcomes. Thus the axioms lead to both personal probability and Bernoullian utility.

3. The scale on which utility is defined is arbitrary, analogous to a temperature scale. In particular, the properties of a utility function that are relevant to choice or decision analysis are not changed under a positive linear transformation; e.g., the function U' will serve as well as the function U where $U' = aU + b$, $a > 0$. There is thus no absolute scale of utility and, tempting as they may be at times, comparisons of utility values between individuals are quite meaningless. Similarly, it makes no sense to speak of one prospect yielding, e.g., twice as much utility as another prospect to a person. We can only say that one prospect exceeds the other in utility.

Bernoulli's principle provides the means for ranking risky prospects in order of preference, the most preferred being the one with the highest (expected) utility. It thus brings together in an explicit way the decision maker's degrees of belief and his degrees of preference—which, of course, are the important subjective inputs in a decision analysis. This process will be illustrated in Chapter 5, but first we must examine methods for eliciting the utility functions of decision makers. Before doing that, we emphasize again the remarkable nature of the expected utility theorem. It says first that if a person accepts the perfectly reasonable axioms of ordering and transitivity, continuity, and independence, this necessarily implies the existence of both a utility function that reflects his preferences for consequences and a subjective probability distribution that reflects his personal judgment about the chances he faces. Second, it says that he should choose between risky prospects to maximize his expected utility. If you accept the axioms, you must also logically accept the criterion of maximizing expected utility. Moreover, the theorem implies a unified theory of utility (preference) and subjective probability (degree of belief). That such a simple and reasonable set of axioms could involve such powerful implications is surely amazing.

4.3 ELICITATION OF PREFERENCES

Elicitation of preferences is as easy or as difficult as choice between simple risky prospects. Some people find such choice easy, while others experience great difficulty in making "public" their choices between risky prospects. Interviewers need to be sympathetic to a slow respondent or to one experiencing difficulty. A few helpful words to make a hypothetical situation more subjectively realistic are often useful. However, an interviewer must take care not to intrude his own preferences into the questioning. We have found that it is nearly always possible to elicit preferences from a decision maker whose attitude is not hostile. We have also found that interviewers learn best by doing and that they improve very rapidly with a little experience.

Unidimensional Utilities

Consequences that can be represented by a single attribute are the simplest to elicit. This is the case of a single dimension of utility so that the utility function has only one argument. A variety of operational procedures has been used for elicitation, the best of which involve choice between two-state risky prospects and systematic manipulation of some consequence component(s) until indifference is reached. Experience has shown that manipulation of probabilities while keeping consequences fixed is not very satisfactory because of the difficulties people have in mentally grappling with probabilities other than those involving only a single decimal digit. In addition, some people exhibit preference for particular probability values (as "favorite numbers" as it were) and this distorts the utility assessment if probabilities are varied rather than consequences.

As with probabilities, it is best to phrase utility questions in terms of two equally likely states; thus both the methods we suggest are based on using the (ethically neutral) probability 0.5.

ELCE METHOD

The simplest recommended method for elicitation is based on considering an Equally Likely risky prospect and finding its Certainty Equivalent (ELCE). Before outlining the ELCE method we therefore need to explain the concept of certainty equivalence. As the name implies, a *certainty equivalent* is the amount exchanged with certainty that makes the decision maker indifferent between this exchange and some particular risky prospect. For example, a particular manager might be indifferent between (1) taking a risky prospect having a 0.4 chance of gaining $10,000 and a 0.6 chance of losing $2000 and (2) a sure prospect giving a gain of $1560. His certainty equivalent for the risky prospect is thus $1560. A certainty equivalent accounts simultaneously for the probabilities in the risky prospect and the preferences for the consequences.

Since we are concentrating on decision problems expressed in financial outcomes, it is useful to compare a certainty equivalent (CE) with the expected money value (EMV) of a prospect. When the CE is less than the EMV, the decision maker is said to display an aversion to risk; if his CE is greater than the EMV, the decision maker is said to exhibit risk preference. The difference between the mean of a risky prospect and its CE (i.e., EMV − CE in the case of a money prospect) is called the *risk premium* for the prospect. The case where CE = EMV (i.e., the risk premium is zero) is the special and relatively rare case of indifference to risk.

The concept of risk premium is illustrated in Figure 4.1. The concave curve represents the utility function of a (risk-averse) manager who, as one of his alternatives, is faced with a risky prospect involving a possible gain of $2000 with probability 0.6, or a possible loss of $3400 with probability

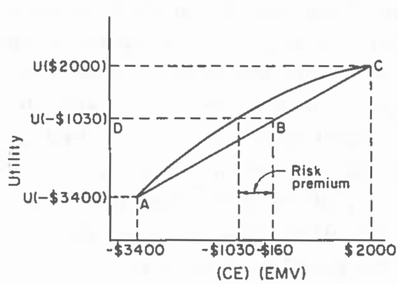


FIG. 4.1. Illustration of the concept of risk premium for a risk-averse decision maker.

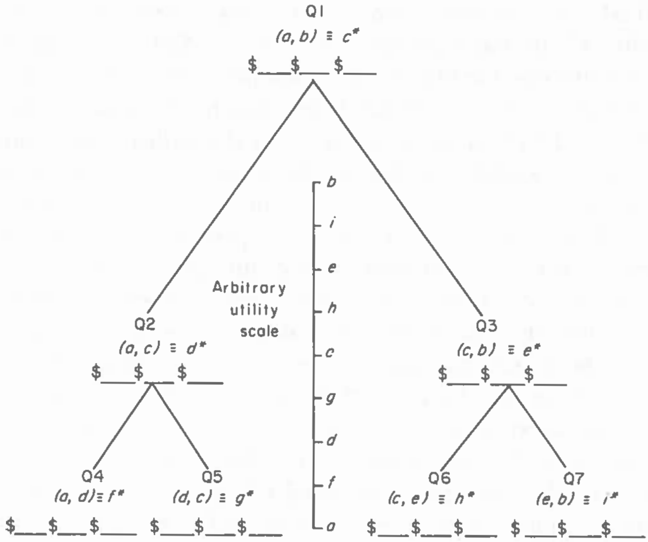
0.4. The EMV of this risky prospect is thus −\$160. However, the expected utility of the risky prospect is $(0.4)U(-\$3400) + (0.6)U(\$2000)$, which corresponds on the utility axis to the point B on the straight line AC such that $AB/BC = 0.6/0.4$. Point B, reading to the money axis, also corresponds to the EMV of the prospect. On the utility axis, point B corresponds to point D which, reading back to the utility curve AC and down to the money axis, tells us the expected utility of the prospect is (approximately) $U(-\$1030)$. While the EMV of the prospect is −\$160, because of his risk aversion the risky prospect is only equivalent to a sure −\$1030 for our manager, i.e., his CE for it is −\$1030. The difference of \$870 between the EMV and his CE is our manager's risk premium for the risky prospect.

In using the ELCE method, the first step in dealing with preferences for a single attribute is to find the CE for a hypothetical 50/50 lottery with the best and the worst possible outcomes of the decision problem as the two risky consequences. The next step is to find the CE for each of the two 50/50 lotteries involving the first-established CE and the best and worst possible outcomes. This process of establishing utility points is continued until sufficient CEs are elicited to plot the utility function. Figure 4.2 depicts the type of recording sheet we have found most useful in using the ELCE method. It also shows the linked nature of the sequence of 50/50 lotteries used. The CEs, denoted with an asterisk, should not be directly requested. Rather, a series of sure payoffs should be presented for each of which the respondent is asked to express his preference (yes or no) relative to the current 50/50 lottery. In this way the questioning can "zero in" upon the CE. We have found it best not merely to present the series of 50/50 lotteries verbally but also to pencil the numbers on a pad in a form such as given in Table 4.2. Here we have used + (−) to indicate that the sure payment is (not) preferred. In practice, instead of recording the whole series of trial CEs (1), (2), . . . , we would use an eraser and just show a single potential CE at a time.

TABLE 4.2. Series of 50/50 Lotteries

50/50 Lottery	Sure Payment (1)	Successive Iterations of Sure Payment (2)	(3)	(4)	(5)	(6)
\$10,000						
	\$4000 (+)	\$1000 (−)	\$3000 (+)	\$2500 (+)	\$2250 (−)	\$2300 (=)
\$0						

The check questions indicated in Figure 4.2 are most important to keep the expression of preferences consistent with inner preferences and



Steps:
(1) Set the range a ($ worst) to b ($ best) for which preference is to be estimated.
(2) In $Q1$, find the CE c^* of the 50/50 lottery (a, b).
(3) In subsequent questions successively bifurcate the utility range into equal preference intervals.
(4) Graph points defining the curve as they are found.
(5) If checks reveal significant inconsistency, do again from the start.

Checks:
After $Q3$ is $(d, e) \equiv c$?
After $Q6$ is $(g, h) \equiv c$?
After $Q7$ is $(f, i) \equiv c$?

FIG. 4.2. Scheme for using the ELCE method of eliciting a decision maker's utility curve.

one with another. Where checks do not correspond closely, it is necessary to return to the first question and repeat all of them until consistency is obtained. Judgment of consistency and adequate closeness is greatly facilitated by concurrently plotting the certainty equivalents after each question, as indicated in the first quadrant of Figure 4.3. For this purpose an arbitrary scale is required and can be set by assigning arbitrary utilities $U(a) < U(b)$ to a and b where $a < b$. For example, $U(a) = 0$ and $U(b) = 8$ (or 100 or 1.0) are convenient for plotting purposes. The utility values of the CEs are determined on this scale by direct application of Bernoulli's principle to each indifference relationship. For the example of Table 4.2, $(0.5)U(a) + (0.5)U(b) = U(c^* = \$2300)$, and substituting the arbitrary

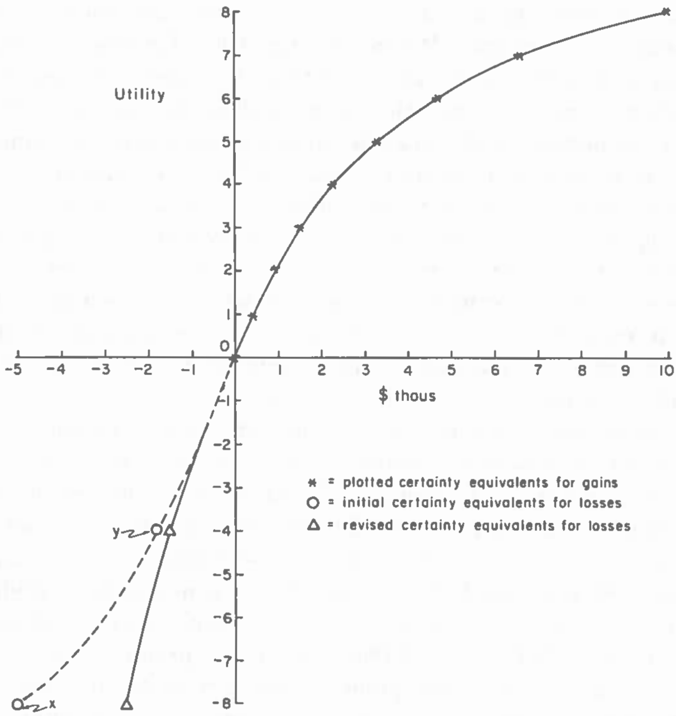


FIG. 4.3. Elicitation plot of the utility curve for gains as per Figure 4.2 (first quadrant) and extensions to utility for losses (third quadrant).

scale values gives $U(c^*) = (0.5)(0) + (0.5)(8) = 4$; similarly, $U(d^* = \$900) = 2$, etc., as plotted in Figure 4.3.

Depending on the degree of accuracy deemed appropriate, the degree of consistency apparent in the emerging sketch, and the degree of impatience perceived in the respondent, it may be quite adequate to cease the interview after successfully passing the check question following question 3 of the ELCE method. It is quite easy to draw through the five points a smooth curve that will be determined after question 3, although it is even easier when nine points are determined, as they are following question 7. At whatever stage the interview is terminated, consistency will be appraised by the smoothness of a curve passing as close as possible to all the determined points. Typically, the curve will be concave (from below), reflecting the fact that most people are averse to risk. Since it seems that most risk-averse people become less so as they become wealthier, this additional typical feature imposes some constraints on the form of the curve, which (as discussed in Section 4.4) can also be exploited in elicitation.

Before we leave the ELCE method, one commonly encountered difficulty should be mentioned. When eliciting utility functions for losses as well as gains, it is frequently observed that the curve obtained after an initial cycle of questioning reveals a convex shape for losses implying an attitude of risk preference that may be superficial and false. Seemingly, our responses to hypothetical situations involving losses are not very reliable. This difficulty is usually avoided by working with assets rather than gains and losses (positive and negative increments to present assets), and for this reason we recommend structuring decision problems and preference interviews where possible in terms of net assets. However, where risk preference for losses is apparent, it should be checked and, if inappropriate, the utility curve corrected to be a more accurate representation of attitudes that are probably risk averse for both gains and losses.

Suppose the questions in Figure 4.2 are extended to encompass money losses. One way of linking the utility segment for losses to the already derived segment for gains is to find payoff x in the 50/50 lottery, $(b, x) \equiv a$; i.e., (\$10,000, x) ≡ \$0. Suppose the farmer's first response is to select x = −\$5000 so that on the arbitrary scale U(−\$5000) = −8. Analogously with the previous question 1, the interval (0, −8) is now split by finding the CE for the 50/50 lottery $(a, x) \equiv y$. Suppose indifference is initially determined at (\$0, − \$5000) ≡ −\$1800, i.e., a risk premium of −\$2500 − (−\$1800) = −\$700. These new points x and y are added to Figure 4.3 in the third quadrant; when the smooth utility curve is extrapolated (by the broken line), it is seen to be convex, implying risk preference for losses.

Is this risk preference plausible and reasonable? The best way of checking such convexity is to repeat some gain questions conditional on a loss already having been incurred; e.g., to check the convexity implied by point y (at −\$1800), ask the farmer to suppose he has just incurred a loss of \$5000. What then would be his CE for the 50/50 lottery (\$5000, \$0)? Let this be \$1500, i.e., a risk premium of \$1000. Correcting these payoffs for the just-lost \$5000, the lottery is seen to be equivalent to (\$0, −\$5000) ≡ (\$1500 − \$5000) = −\$3500, which contrasts markedly with the earlier answer of −\$1800 and exemplifies the inconsistencies of the first round of questions. Plotting this revised CE would indicate a concave segment between −\$5000 and \$0. However, since this is probably on a different scale from the curve in the first quadrant, this is not shown on Figure 4.3. But is the −\$5000 point consistent and correct on the "old" scale? Very likely not if this latest evidence of risk aversion for losses is any guide. What we should do now, after having explained the farmer's earlier inconsistencies to him, is to return to the search for x and to check the apparent risk attitudes by rephrasing the questions in the style "Suppose now that you had just paid out a sum of \$$x$, . . ." and elicit CEs for the equivalent lotteries. In this way we can expect to emerge with a curve that probably looks much more like a concave extrapolation of the segment in the first quadrant, implying risk aversion over the whole range of outcomes considered. The revised curve has been sketched on the assumption that, on renewed reflection, $x = -\$2500$ and $y = -\$1640$.

ELRO METHOD

A shortcoming of the ELCE method, and one that may sometimes be serious for people with a strong aversion to gambling per se, is that preference is expressed through certainty equivalents. An alternative method has been devised that overcomes this difficulty, although at the cost of a more complicated questioning procedure. The method is based on preferences between acts with Equally Likely but Risky Outcomes (ELRO).

Suppose we are interested in deriving a utility function for money outcomes over the range a to z, $a < z$. We will normally need at least five utility values to be able to graph a smooth function. We therefore start the ELRO method by selecting a reference interval, involving two money outcomes x and y somewhere near the middle of the range a to z, with $x < y$ and $y - x$ approximately equal to one-tenth of $z - a$. This should give us an ample number of points for our graph, although at this stage we cannot be sure of this. If we find we end up with too few points, we must set a narrower reference interval and start again. The next step is to set a scale for our utility function by defining $U(y) - U(x) = 1$ (or any other arbitrary utility value). Similarly, we need an arbitrary origin for which it is convenient to define $U(a) = 0$.

We now present the decision maker with the hypothetical 50/50 lottery given in Table 4.3. As before, the value marked with an asterisk (b in

TABLE 4.3. First Hypothetical Lottery

θ_i	$P(\theta_i)$	a_1	a_2
θ_1	0.5	x	y
θ_2	0.5	b^*	a

TABLE 4.4. Second Hypothetical Lottery

θ_i	$P(\theta_i)$	a_1	a_2
θ_1	0.5	x	y
θ_2	0.5	c^*	b

this case) is varied until indifference is found between the risky prospects a_1 and a_2. Then, by Bernoulli's principle, $0.5U(x) + 0.5U(b) = 0.5U(y) + 0.5U(a)$, which can be rearranged as $U(b) - U(a) = U(y) - U(x)$. However, as defined above, $U(y) - U(x) = 1$ and $U(a) = 0$. Hence $U(b) = 1$. The hypothetical lottery in Table 4.4 uses this newly established value of b in place of a to find a further indifference value c.

By similar reasoning to that above, it can be shown that $U(c) = 2$. We now have the utilities of three money consequences, a, b, and c, and from this point on the questioning can proceed in a simpler fashion. The sequence of further lotteries presented to the decision maker is given in Table 4.5. In each case the entry marked with an asterisk is varied until

TABLE 4.5. Sequence of Further Lotteries

		Lottery 3		Lottery 4		Lottery 5		
θ_i	$P(\theta_i)$	a_1	a_2	a_1	a_2	a_1	a_2	Etc.
θ_1	0.5	a	b	b	c	c	d	
θ_2	0.5	d*	c	e*	d	f*	e	

indifference is established. The sequence is extended until a money value greater than or equal to z, the upper limit of the range of interest, is reached. For each lottery in the sequence the utilities of three values have previously been established, so the utility of the remaining amount can be calculated. In fact, it is easy to show that on the scale we are using, $U(d) = 3$, $U(e) = 4$, $U(f) = 5$, etc.

Check questions can be included in the ELRO procedure as they were for the ELCE method. For example, the subject should be indifferent between the alternatives in Table 4.6. To see why, note that we have

TABLE 4.6. Alternatives Offered

θ_i	$P(\theta_i)$	a_1	a_2
θ_1	0.5	a	c
θ_2	0.5	d	b

established $U(a) = 0$, $U(b) = 1$, $U(c) = 2$, and $U(d) = 3$. Then by Bernoulli's principle we have $U(a_1) = (0.5)(0) + (0.5)(3) = 1.5$ and $U(a_2) = (0.5)(2) + (0.5)(1) = 1.5$. Consistency of judgment can also be assessed by plotting each utility point as for the ELCE method.

Multidimensional Utilities

Money is not everything; and the consequences of many decision problems are, unfortunately from an analytic viewpoint, not well represented in terms of only a single attribute such as monetary gain or loss. Large organizations such as corporations and governments must typically consider several dimensions in assessing consequences. Failure to do so would doubtless make for a short life for any such organization. At the same time, because simultaneous consideration must be given to both the probabilities of various consequences and to tradeoffs between different levels of the various attributes, it is much more difficult for decision makers to think about risky alternatives involving more than one attribute.

We will outline three methods of multidimensional utility assessment—the benchmark approach, the "quasi-separable" utility function approach, and the additive utility function approach. We will also briefly indicate the possible relevance of lexicographic utility orderings in situations involving multidimensional consequences.

BENCHMARK APPROACH

The essence of the benchmark approach (Raiffa, 1968; von Winterfeldt and Fischer, 1973) is that for every multiattribute consequence, a consequence is found that is indifferent to it and has constant values in all dimensions except one that is preferentially independent of the others (in the sense that preference for values in that attribute are independent of constant values in the other attributes). For example, suppose we are concerned with consequences characterized by three attributes x, y, and z. We will use subscripts to denote particular values of x, y, and z. For a given decision maker, attribute x is preferentially independent of the others if when the consequence (x_1, y_1, z_1) is preferred to (x_2, y_1, z_1), then (x_1, y_i, z_j) will always be preferred to (x_2, y_i, z_j) for all values of y_i and z_j. In other words, x is preferentially independent of y and z if the decision maker's preferences for values in x are independent of constant values in y and z; if his preferences for x conditional on particular levels of y and z change as y and z vary, then x is not preferentially independent.

If at least one of the attributes is not preferentially independent of the rest, choice between multiattribute consequences is purely a matter of the decision maker's intuition and cannot be formalized. Most commonly, however, this will not be the case, and the benchmark (or some other formalized) approach can be used.

We will first consider the simplest case where consequences can be adequately described by just two attributes. Suppose these are measured by x and y and that x is preferentially independent of y. The ith consequence is defined as the pair (x_i, y_i). For example, in a commercial farming situation x might measure annual net profit and y might measure annual net peak indebtedness; or in a peasant farming context x might measure consumption of nonfarm goods and y might measure consumption of farm production such as grains, etc. Now we select a "benchmark" level of y, say y^+, where this is chosen as a value that is easily conceptualized, such as the "normal" annual peak indebtedness. The approach to dealing with multidimensional consequences consists of relating any consequence back to the benchmark. In the two-dimensional case we must find the value of x_i, denoted by x_i^+, which when paired with y^+ makes the decision maker just indifferent between this hypothetical consequence and a particular consequence (x_i, y_i). Obviously, this will not be too easy a task and will involve asking a series of questions such as, Now, on very careful consideration, would you rather have (x_i, y_i) or (x_i', y^+)? where x_i' is used to denote some trial estimate of x_i^+. It is not all that easy to respond to such questions, as a little introspection readily reveals.

With two-dimensional consequences it is feasible and often worthwhile to systematize the procedure of trading off one attribute for another by developing an empirical indifference map. The precision sought in such an exercise will, as always, depend on the requirements of a particular analysis but may often not be too demanding. Elicitation can be aided by imposition of the usual features ascribed to indifference curves such as smoothness, convexity, and nonintersection. In this way a rough but reasonable indifference map might be sketched from a dozen or so points spanning the preference region of interest in a decision analysis. The result may look something like that depicted in Figure 4.4, where the series of points marked *a, b, c* are the elicited combinations of *x* and *y* that plot out the indifference curves *AA, BB,* and *CC* respectively.

An indifference map such as that of Figure 4.4 permits interpolation of the benchmark equivalent of any consequence. For example, as illustrated in Figure 4.4, the consequence *D* can be seen by interpolation to have a benchmark equivalent of (x_d^+, y^+). Once the x_i^+ equivalence is determined, it remains to complete the final phase of elicitation, namely, the scaling of preference for the x_i^+. This can be executed in the conventional ways noted earlier, but phrasing the questions (of, say, the ELCE method) conditional on *y* being set at the benchmark level y^+. In this way it is possible, albeit in a somewhat roundabout way, to associate a single utility value with any (x_i, y_i) pair, and so to proceed to resolve the analysis in

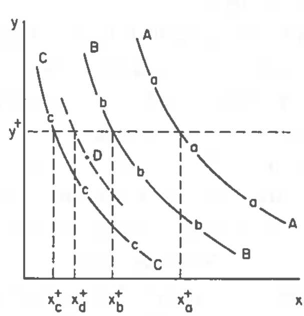


FIG. 4.4. An elicited indifference map for consequences with attributes *x* and *y*.

terms of expected utility. Thus for the jth act a_j with the two-dimensional consequence (x_{ij}, y_{ij}) occurring when Θ_i prevails, we have

$$U(a_j) = \sum_i U(a_j \mid \Theta_i) P(\Theta_i) = \sum_i U(x_{ij}, y_{ij}) P(\Theta_i)$$
$$= \sum_i U(x_{ij}^+, y^+) P(\Theta_i) = \sum_i U(x_{ij}^+ \mid y^+) P(\Theta_i) \qquad (4.4)$$

where $U(x_{ij}^+ \mid y^+)$ denotes the utility value for x_{ij}^+ derived from the utility function $U(x \mid y^+)$ for attribute x conditional on attribute y being held constant at the benchmark value y^+.

As an example, suppose we have to choose between acts a_1 and a_2 with

TABLE 4.7. Two-dimensional Consequences

	a_1		a_2	
Θ_i	x	y	x	y
Θ_1	33	180	100	15
Θ_2	190	70	150	155

TABLE 4.8. Equivalent Consequences

	a_1		a_2	
Θ_i	x^+	y^+	x^+	y^+
Θ_1	51	120	11	120
Θ_2	110	120	190	120

two-dimensional consequences specified as in Table 4.7. Taking $y^+ = 120$ as our benchmark, suppose our set of equivalent consequences is ascertained to be as shown in Table 4.8. With use of the ELCE or ELRO method, the utility function $U(x \mid y = 120)$ for x with y held constant at 120 can then be determined. Applying equation (4.4), we have

$$U(a_1) = U(51 \mid y = 120) P(\Theta_1) + U(110 \mid y = 120) P(\Theta_2)$$
$$U(a_2) = U(11 \mid y = 120) P(\Theta_1) + U(190 \mid y = 120) P(\Theta_2)$$

Finally, comparison of $U(a_1)$ and $U(a_2)$ indicates which act maximizes expected utility.

The same basic approach can be applied to multidimensional consequences when more than two attributes are involved although, understandably, the specification of tradeoffs becomes correspondingly more difficult. Thus suppose we are concerned with consequences characterized by four attributes w, x, y, and z. If only one attribute (say w) is preferentially independent of the others, benchmark values are defined for x, y, and z (say x^+, y^+, and z^+) and for every consequence (w_i, x_i, y_i, z_i) an outcome (w_i^+, x^+, y^+, z^+) is found that is indifferent to it. Because only w is preferentially independent, this determination of indifference has to be made in a single step by simultaneous tradeoff across the other attributes. If two attributes (say w and x) are preferentially independent, indifference can be established for each consequence (w_i, x_i, y_i, z_i) relative to the benchmarks x^+, y^+, and z^+ by a two-step procedure of establishing indifference first with some consequences (w_i^+, x_i, y^+, z^+) and then of this consequence with $(w_i^{++}, x^+, y^+, z^+)$.

Simplest of all is the case where all (or all except one) the attributes are preferentially independent. In this case, if there are n attributes and thus $n - 1$ benchmark values, we can proceed via a sequence of $n - 1$ tradeoffs, each involving only tradeoffs between two attributes. This is much simpler for decision makers than having to make simultaneous tradeoffs involving three or more attributes. Thus if w, x, and y (or w, x, y, and z) are preferentially independent, we can use the three-step procedure of establishing (w_i, x_i, y_i, z_i) indifferent to (w_i^+, x_i, y_i, z^+) indifferent to $(w_i^{++}, x_i, y^+, z^+)$ indifferent to $(w_i^{+++}, x^+, y^+, z^+)$. Again then, a conditional preference function $U(w \mid x^+, y^+, z^+)$ can be determined and used for evaluation by carrying out expected utility analysis along the lines of equation (4.4).

Utility scaling of multiattributed consequences is still more difficult when some or all of the attributes are qualitative in nature. The method described above can be adapted for use in such cases if each dimension of the qualitative consequences can be categorized into a number of ordered classes, e.g., true or false; good, fair, or poor. Failing this, each multiattributed consequence must be scaled directly. Such scaling may be facilitated by first ranking the consequences in order of preference and then assigning utility values of, say, zero and one to the worst and best outcomes respectively. Each other multiattributed outcome can then be scaled by considering a lottery of the form given in Table 4.9. In this lottery a is the most preferred multiattributed consequence and z is the least

TABLE 4.9. Lottery for Scaling Multiattributed Outcomes

Θ_i	$P(\Theta_i)$	a_1	a_2
Θ_1	p	a	c
Θ_2	$1 - p$	z	c

preferred, while c is any other multiattributed outcome to be assigned a utility value. The probability p is varied until indifference between a_1 and a_2 is established so that $U(c) = pU(a) + (1 - p)U(z)$. Hence $U(c) = p$. Note, however, that this method is vulnerable to biases caused by preference for particular probability values, as discussed above in relation to elicitation of unidimensional utilities.

"QUASI-SEPARABLE" UTILITY FUNCTION APPROACH

The procedures for estimating multidimensional utilities outlined above become tedious if there are many possible consequences, each having more than a couple of attributes. They have the advantage, though, of requiring only relatively weak assumptions about independence between particular attributes. However, if the two assumptions of joint preferential independence and utility independence (sometimes called weak and strong conditional utility independence respectively) can be made, the analysis of large problems (involving say hundreds of possible consequences each with many attributes) can be much simplified through use of the "quasi-separable" utility function approach developed by Keeney (1968, 1972, 1973a). The essence of this approach is that it uses the above assumptions to enable decomposition of the multiattribute utility function into component parts.

To explain the meaning of joint preferential independence and utility independence, consider consequences with attribute dimensions $x_1, x_2, \ldots, x_n$. Attributes x_i and x_j are *jointly preferentially independent* of the other attributes if the location and shape of the decision maker's indifference curves for combinations of x_i and x_j are independent of the level of other attributes. The easiest way to check this requirement is by direct questioning of the decision maker: first, to establish two or more consequences differing only in x_i and x_j such that he is indifferent between them and, second, to check that this indifference is not upset when the levels of the other attributes are changed.

Attribute x_i is *utility independent* of the other attributes if the decision maker's preferences for lotteries involving only x_i with other attributes held constant do not depend on the level of the other attributes. A convenient way to check this requirement is to see whether the decision maker's certainty equivalent for 50/50 gambles between particular values of x_i (say x_i^* and x_i^-, representing the most desired and least desired levels of x_i respectively) stays constant as the levels of the other attributes are changed. If so, x_i is utility independent of the other $n - 1$ attributes. While utility independence implies preferential independence, the reverse is not true.

Given that the requirements of joint preferential and utility independence are met for all n attributes, the utility function $U(x_1, x_2, \ldots, x_n)$ can be specified as a function f such that

$$U(x_1, x_2, \ldots, x_n) = f[U_1(x_1), U_2(x_2), \ldots, U_n(x_n)] \tag{4.5}$$

where $U_i(x_i)$, scaled from zero to one, is a utility function over the ith attribute and depending only on that attribute. When scaled from zero to one, the function f is either of the additive form

$$U(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} k_i U_i(x_i) \tag{4.6}$$

or of the multiplicative form

$$U(x_1, x_2, \ldots, x_n) = \left\{ \prod_{i=1}^{n} [Kk_i U_i(x_i) + 1] - 1 \right\} / K \tag{4.7}$$

where k_i is a scaling factor between zero and one for $U_i(x_i)$ and K is another scaling constant. Because of the scaling requirements, the k_i values determine K. If $\sum k_i = 1$, then $K = 0$ and f takes the additive form of equation (4.6). If $\sum k_i \neq 1$, then $K \neq 0$ and f takes the multiplicative form of (4.7). As shown by Keeney (1974), these implications follow from the specific quasi-separable form of (4.5), which is

$$\begin{aligned} U(x_1, x_2, \ldots, x_n) = {} & \sum_{i=1}^{n} k_i U_i(x_i) + K \sum_{i=1}^{n} \sum_{j>i} k_i k_j U_i(x_i) U_j(x_j) \\ & + K^2 \sum_{i=1}^{n} \sum_{j>i} \sum_{m>j} k_i k_j k_m U_i(x_i) U_j(x_j) U_m(x_m) + \cdots \\ & + K^{n-1} \prod_{i=1}^{n} k_i U_i(x_i) \end{aligned} \tag{4.8}$$

For $\sum k_i = 1$ so that $K = 0$, (4.8) reduces immediately to the additive form of (4.6). When $\sum k_i \neq 1$ so that $K \neq 0$, multiplication of each side of (4.8) by K, followed by the addition of one to each side and factoring of the right side, gives the multiplicative form of (4.7).

All of this looks quite complicated, but the essence is that instead of trying to assess the n-dimensional utility function $U(x_1, x_2, \ldots, x_n)$ directly (a virtually impossible task), it is only necessary to assess n one-dimensional functions $U_i(x_i)$ and the n scaling factors k_i. These k_i values represent the utility assigned to a consequence with all its attributes except the ith set at their least preferred amount within the relevant range and the ith set at its most preferred amount within the relevant range.

To determine k_i, we elicit from the decision maker the probability p_i such that he is indifferent between (1) the consequence with x_i at its most preferred amount and all other attributes at their least preferred amount and (2) a lottery with chance p_i of yielding the consequence with all attributes at their most preferred amount and a chance $1 - p_i$ of yielding the consequence with all attributes at their least preferred amount. The value of p_i must equal k_i, as shown by the following simple example. We will again use x_i^* and x_i^- to indicate respectively the most and least preferred levels of x_i within the relevant range. Suppose there are three attributes x_1, x_2, and x_3. To determine, say, k_1, we elicit p_1 such that

$$U(x_1^*, x_2^-, x_3^-) = p_1 U(x_1^*, x_2^*, x_3^*) + (1 - p_1) U(x_1^-, x_2^-, x_3^-)$$

Hence $U(x_1^*, x_2^-, x_3^-) = p_1$ since scaling of the utility function from zero to one implies $U(x_1^*, x_2^*, x_3^*) = 1$ and $U(x_1^-, x_2^-, x_3^-) = 0$. From equation (4.7) we have

$$KU(x_1^*, x_2^-, x_3^-) + 1 = [Kk_1 U_1(x_1^*) + 1][Kk_2 U_2(x_2^-) + 1][Kk_3 U_3(x_3^-) + 1]$$

so that, again by virtue of the scaling of the utility functions from zero to one, $Kp_1 + 1 = Kk_1 + 1$ since $U_1(x_1^*) = 1$ and $U_2(x_2^-) = 0 = U_3(x_3^-)$. Therefore, $p_1 = k_1$, or for the general case, $p_i = k_i$ and $0 < k_i < 1$ since necessarily $0 < p_i < 1$. Use of (4.6) gives the same result.

If $\sum k_i \neq 1$, it is necessary to determine K. To do this, consider equation (4.7) with each attribute set at its most preferred level. We have

$$U(x_1^*, x_2^*, x_3^*) = \left\{ \prod_{i=1}^{n} [Kk_i U_i(x_i^*) + 1] - 1 \right\} / K$$

which reduces to

$$K + 1 = (Kk_1 + 1)(Kk_2 + 1) \ldots (Kk_n + 1) \tag{4.9}$$

since $U(x_1^*, x_2^*, x_3^*) = 1 = U_i(x_i^*)$. Given the already ascertained k_i values, we can solve the above equation for K, making use of the fact (Keeney, 1972) that if $\sum k_i > 1$, we must have $-1 < K < 0$; and if $\sum k_i < 1$, then $0 < K < \infty$.

As an example of the "quasi-separable" utility function approach to multiattribute problems, suppose we are acting as decision analysts for a person considering the purchase of a farm. He has narrowed the choice to three alternatives a_1, a_2, and a_3 and is concerned with the following attributes:

x_1 = average annual net return on equity capital assessed at market value (%)
x_2 = initial equity (%)
x_3 = distance from the capital city (km)
x_4 = driving distance from the beach (km)

Of these attributes only x_1 is uncertain, and it will vary according to whether long-term conditions for agriculture are good (Θ_1), fair (Θ_2), or poor (Θ_3)—taking into account questions of climate, international markets, technological change, etc. Our client judges these x_1 possibilities as shown in Table 4.10 with x_2, x_3, and x_4 being fixed within each alternative for each type of economic circumstance.

TABLE 4.10. Multiattribute Payoff Matrix for Property Purchase Decision Problem

		a_1				a_2				a_3			
θ_i	$P(\theta_i)$	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
θ_1	0.2	12	60	80	160	7	80	220	40	12	96	300	200
θ_2	0.5	7	60	80	160	5	80	220	40	10	96	300	200
θ_3	0.3	5	60	80	160	1	80	220	40	8	96	300	200

Looking at the data of Table 4.10, we can readily agree that to make a choice is not easy if real account is to be taken of all the attribute dimensions. Our first step is to establish whether the attributes are jointly preferentially independent and utility independent for our client. For joint preferential independence we have to check if the tradeoffs between any two attributes are independent of the levels of other attributes. In the present case we have six pairs (x_1, x_2), (x_1, x_3), (x_1, x_4), (x_2, x_3), (x_2, x_4), and (x_3, x_4) to check. As the procedure is the same for any pair, consider (x_2, x_3) as an example.

To begin, we ask questions to find the level of initial equity x_2 such that the combination of equity and some distance to the city (say $x_3 =$ 200 km) is indifferent to, say, the combination ($x_2 = 60$, $x_3 = 100$), in both cases the level of the other attributes being held constant at a desirable level (say $x_1 = 12$, $x_4 = 40$). In other words, given that the average return on equity and distance to the beach are fixed at 12% and 40 km respectively, what initial equity percentage in combination with a distance to the city of 200 km would be indifferent to an equity level of 60% and a distance of 100 km to the city? Suppose after considering various values, our client chooses an initial equity of 70%. This same questioning is then repeated but with x_1 and x_4 at less preferred levels (say $x_1 = 1$ and $x_4 = 200$). If the answer is again 70% initial equity, we may conclude x_2 and x_3 are jointly preferentially independent. We will assume that after applying this procedure to all the attribute pairs, we are able to conclude they are all jointly preferentially independent to our client (though they may not be to us).

For utility independence we have to check each attribute individually, using the same procedure for each. Consider x_4, distance to the beach. We first set the other attributes at convenient levels (say the most desirable ones so that $x_1 = 12$, $x_2 = 96$, $x_3 = 80$) and using the ELCE or ELRO method elicit the conditional utility function for x_4 over its relevant range from 40 to 200 km. Then we set the other attributes at their least desirable levels of $x_1 = 1$, $x_2 = 60$, $x_3 = 300$ and again elicit the conditional utility function. This procedure may be repeated with x_1, x_2, and x_3 set at various other levels. If the conditional function $U(x_4 \mid x_1, x_2, x_3)$ remains unchanged as x_1, x_2, and x_3 vary, then x_4 is utility independent. Again suppose that each attribute is found to be utility independent for our client. However, if we found x_4 to be utility interdependent with some other attribute (say x_3), we could use the benchmark method to put these two interdependent attributes onto one utility scale (say for x_3 conditional on some benchmark level of x_4) and then continue using independent procedures for (x_1, x_2), and $(x_3 \mid x_4^+)$.

In the process of successfully checking for utility independence, we have ascertained a utility function for each attribute. These individual functions, scaled from zero to one and drawn as a graph or fitted algebraically, are the required functions $U_i(x_i)$ of equation (4.8). Suppose they are as depicted in Figure 4.5.

To estimate the k_i scaling values, consider net return on equity, x_1, as an example. We have to elicit the probability p_1 such that our client is indifferent between (1) the consequence with x_1 at its most preferred level and the other attributes at their worst level—i.e., the consequence ($x_1^* = 12$, $x_2^- = 60$, $x_3^- = 300$, $x_4^- = 200$)—and (2) the lottery with a chance p_1 of yielding the consequence with all attributes at their preferred level—i.e. ($x_1^* = 12$, $x_2^* = 96$, $x_3^* = 80$, $x_4^* = 40$)—and a chance $(1 - p_1)$ of yielding the consequence with all attributes at their worst level—i.e. ($x_1^- = 1$, $x_2^- = 60$, $x_3^- = 300$, $x_4^- = 200$). Suppose this value of p_1 is 0.6.

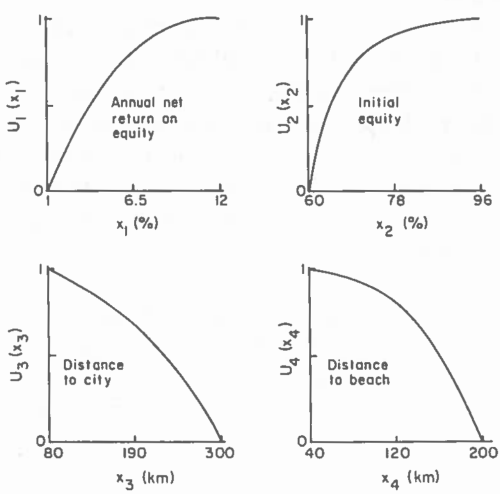


FIG. 4.5. Utility functions for individual attributes in a multiattribute decision problem using the quasi-separable approach of equation (4.8).

Therefore, as shown above, $k_1 = 0.6$. Repeating this process for the other attributes, suppose we elicit the set of k_i values shown in Table 4.11.

TABLE 4.11. Set of k_i Values

Attribute	k_i
x_1 = return on equity	0.6
x_2 = initial equity	0.4
x_3 = distance to city	0.3
x_4 = distance to beach	0.3

Since the sum of these k_i values is not equal to one, we know the utility function is multiplicative rather than additive and we must determine the value of K as per equation (4.9). Thus we know K is the solution to

$$K + 1 = [K(0.6) + 1][K(0.4) + 1][K(0.3) + 1][K(0.3) + 1]$$

or, equivalently, $0.0216K^3 + 0.234K^2 + 0.93K + 0.6 = 0$ such that $-1 < K < 0$ since $\sum k_i > 1$. Solving this equation gives $K = -0.79$. The required multiattribute utility function can therefore be specified as per equation (4.7) as

$$U(x_1, x_2, x_3, x_4) = \{[1 - 0.474U_1(x_1)][1 - 0.316U_2(x_2)] \cdot [1 - 0.237U_3(x_3)][1 - 0.237U_4(x_4)] - 1\}/-0.79$$

where the single-attribute utility functions $U_i(x_i)$ are as depicted graphically in Figure 4.5. Reading $U_i(x_i)$ values from the graphs [in larger problems it would be better to fit $U_i(x_i)$ algebraically and use a computerized approach] and substituting into the above utility function, we obtain the utility values under each state of nature for our client's three alternatives. These $U(a_j \mid \theta_i)$ values are as listed in Table 4.12. The expected utility of each act is then found in the usual way as $U(a_j \mid \theta_i)P(\theta_i)$. As shown in Table 4.12, these utility values indicate a_2 as the optimal choice in preference to a_3 in preference to a_1.

TABLE 4.12. Utility Evaluation of Property Purchase Decision Problem

θ_i	$P(\theta_i)$	a_1	a_2	a_3
θ_1	0.2	0.82	0.91	0.81
θ_2	0.5	0.76	0.86	0.80
θ_3	0.3	0.68	0.67	0.77
Expected utility		0.75	0.81	0.79

ADDITIVE UTILITY FUNCTION APPROACH

The approach most used in evaluating multidimensional consequences has been the additive utility function of equation (4.6). It is by far the simplest approach since it only involves determination of the unidimensional utility function $U_i(x_i)$ for each attribute x_i and the scaling factor k_i associated with each $U_i(x_i)$. For instance, each attribute dimension x_i is scaled as in the quasi-separable approach by eliciting the utility function $0 \leqq U_i(x_i) \leqq 1$ for each x_i. Then the most preferred of the set of preferred values $x_1^*, x_2^*, \ldots, x_i^*, \ldots$ is given a k_i value of one, and the most preferred value for each of the other attribute dimensions is scaled between zero and one against the rating of unity given to the most preferred of the preferred values. These ratings are k_is. Alternatively, procedures can be used that find k_i implicitly while the $U_i(x_i)$ are being found (Fishburn, 1967).

Though the necessary requirements for an additive utility function to be true will rarely be met, the assumption of additivity may not be too bad since what is required of the multidimensional utility function is the power to discriminate between alternative acts $a_1, a_2, \ldots, a_j, \ldots$. In by far the majority of multidimensional situations, as shown by Yntema and Torgerson (1967), main (i.e., additive) effects tend to swamp interaction (i.e., multiplicative) effects. Thus while an additive utility function will not accurately specify $U(a_j)$, it will generally serve reasonably well to discriminate between acts in much the same way as would a more correct but far more complicated nonadditive function (Huber, 1974a). For example, if we apply the linear utility function

$$U(x_1, x_2, x_3, x_4) = \sum_{i=1}^{4} k_i' U_i(x_i)$$

to the multiattribute property purchase problem discussed above, using $k_i' = k_i / \sum k_i$ so as to scale the utility function from zero to one and reading $U_i(x_i)$ values from Figure 4.5, the expected utilities of the three alternatives a_1, a_2, a_3 are respectively 0.59, 0.71, and 0.61. The recommended choice would still be a_2 in preference to a_3 in preference to a_1.

LEXICOGRAPHIC UTILITY

Lexicographic utility orderings prevail when we have multiattribute situations in which the decision maker is not prepared to allow tradeoffs between attributes. Rather, he attaches dominant priorities to attributes in some specified order reflecting a hierarchy of wants. In terms of the Bernoulli axioms this implies nonacceptance of the continuity axiom. Instead of being measurable as a single real number, the utility of an act with multidimensional consequences must then be expressed as a priority-ordered vector showing the expected utility in each attribute dimension. Choice proceeds on the basis of a lexicographic comparison of these priority-ordered vectors. For example, it has been suggested that some farmers evaluate alternatives on the basis of a top priority survival goal (requiring actual payoff to exceed some critical level with some specified probability) and a profit-maximizing goal. Only if alternatives are equal in terms of the safety-first survival requirement would choice hinge on a consideration of the expected utility of profit.

Suppose that our client in the property purchase problem above is not prepared to allow tradeoffs between the four attribute dimensions. Instead he places them in a dominant priority order of annual net return x_1, initial equity x_2, distance to the city x_3, and distance to the beach x_4. Reading individual attribute utility values from Figure 4.5, we have the utility payoff matrix of Table 4.13. Our client's choice thus rests between the three lexicographically ordered (expected) utility vectors

$$U(a_1) = (0.83; 0.00; 1.00; 0.50)$$
$$U(a_2) = (0.50; 0.93; 0.54; 1.00)$$
$$U(a_3) = (0.96; 1.00; 0.00; 0.00)$$

Comparison of these vectors on the basis of the required dominant priority of x_1 over x_2 over x_3 over x_4 indicates a_3 is the preferred choice since $0.96 > 0.83 > 0.50$. For choice to hinge on the least priority attribute x_4, each vector must have equal utility values for x_1 (e.g., 0.96), for x_2 (e.g., 1.00), and for x_3 (e.g., 0.54). If this was so, a_2 would be the preferred act.

TABLE 4.13. Lexicographic Utility Vectors for Property Purchase Decision Problem

		a_1				a_2				a_3			
θ_i	$P(\theta_i)$	U_1	U_2	U_3	U_4	U_1	U_2	U_3	U_4	U_1	U_2	U_3	U_4
θ_1	0.2	1.00	0	1	0.5	0.85	0.93	0.54	1	1.00	1	0	0
θ_2	0.5	0.85	0	1	0.5	0.67	0.93	0.54	1	0.98	1	0	0
θ_3	0.3	0.67	0	1	0.5	0.00	0.93	0.54	1	0.91	1	0	0
Expected utility		0.83	0	1	0.5	0.50	0.93	0.54	1	0.96	1	0	0

4.4 CONSTRAINTS ON THE UTILITY FUNCTION

Little mention has been made of any general properties that we might find desirable to impose on an elicited preference curve. If we always prefer more money to less, it follows from the first property of Bernoulli's principle that utility increases monotonically with money or, equivalently, that the marginal utility for money (i.e., the first derivative of the utility function) is strictly positive. This feature must be actively borne in mind when empirical recourse is made to otherwise perhaps innocuous looking algebraic representations such as $U(x) = x + bx^2$, $b < 0$, which reaches a maximum at and has negative marginal utility beyond $x = -1/2b$. At best, such a function could only represent preferences for $x < -1/2b$.

Apart from his response to elicitation questions, a decision maker may wish to ensure that his utility function expresses certain qualitative aspects of his preferences. For example, he might suggest a general aversion to risk with specific regions of preference for risk corresponding perhaps to particularly important aspiration levels. Friedman and Savage (1948) and others have postulated variously shaped utility functions to rationalize observed simultaneous participation in insurance and lottery markets. However, except for providing a general frame of reference for checking an elicited function, it is not clear how such qualitative assessments may be incorporated in utility analysis. If they really exist, such qualitative aspects will be reflected in the elicited utility curve.

As a generalization, most people seem to be averse to risk over the range of payoffs appropriate to their usual managerial decision making. Recall that this means that for any risky prospect, the risk premium (EMV − CE) is positive. Visually, this implies that the utility curve of a risk-averse person displays diminishing marginal utility (i.e., the second derivative is negative).

When the argument of a utility function is specified to be wealth or asset position rather than gains or losses, it is often possible to specify some additional constraints regarding risk aversion. Consider facing a risky prospect involving an equally likely gain or loss of \$100. Suppose you were prepared to pay up to \$20 to avoid facing this prospect, in which case your risk premium = EMV − CE = 0 − (−20) = \$20. Now, if you were rather wealthier than you are currently, would you be prepared to pay the same sum to avoid this fixed-size lottery? If you would pay less (and this would put you in the same category as most people), this indicates *decreasing risk aversion* with increasing wealth.

In eliciting utility curves for wealth or net assets, it is useful to determine qualitatively if aversion to risk is decreasing. If it is, the curve can be readily checked by computing the implied risk premiums for a sequence of symmetric equally likely "lotteries" involving pairs of increasing asset positions and by checking that they diminish.

More formally, Pratt (1964) has shown that degree of risk aversion can be measured by a coefficient $r(W)$, defined as the negative ratio of the second and first derivatives of the utility of wealth function $U(W)$; i.e.,

$$r(W) = -U_2(W)/U_1(W) \tag{4.10}$$

where the subscripts denote derivatives and W denotes wealth or assets. The Pratt coefficient is the simplest measure of curvature that is not changed by an arbitrary positive linear transformation of the utility function. Also, since $r(W)$ is a pure number, it allows interpersonal comparisons of the degree of risk aversion at particular wealth levels. The Pratt coefficient is positive for risk aversion; and for decreasing risk aversion $r(W)$ diminishes with increasing wealth, i.e., $r_1(W) < 0$. Only a few functional forms can represent utility with these characteristics (e.g., $U = \log_e W$; $U = W^c$, $0 < c < 1$). In particular, it should be noted that the quadratic utility of wealth function ($U = W + bW^2$, $b < 0$) even over its range of respectability, implies increasing rather than decreasing risk aversion. Constant aversion to risk implies a particular class of utility functions $U(W) = 1 - e^{-cW}$, where $c > 0$ is Pratt's measure of risk aversion. This negative exponential utility function was the form employed by Freund (1956) in one of the earliest agricultural applications of utility analysis.

Further qualitative restrictions on preferences, such as decreasing aversion to size of risks as discussed by Zeckhauser and Keeler (1970), can also be employed to circumscribe even more narrowly the field of algebraic contenders with consequent impact on algebraic specification of empirical utility functions for wealth. Algebraic forms suitable for specifying the utility of gains and losses (x) are also rather restricted by the need to handle losses satisfactorily. In particular, functions of the form $U = f(\log x)$ are not defined for $x \leqq 0$.

4.5 ALGEBRAIC REPRESENTATION OF THE UTILITY FUNCTION

It is often convenient to have utility functions for gains and losses expressed in algebraic form. Two particular situations where this is so can be distinguished: (1) the case where an algebraic representation of a curve smoothed through a set of elicited utility points is needed to permit the utility value of any outcome in the range to be computed and (2) the case where, in addition to the requirements above, knowledge of the algebraic form of the fitted function enables easy calculation of the function's derivatives for use in certain analytical decision procedures described in Section 4.6.

A number of procedures can be employed to meet the first case. Perhaps the easiest, in terms of ready availability of suitable computer routines, is to fit a polynomial function by least squares regression. Note, however, that conventional tests of goodness of fit (such as R^2 values) are of limited relevance since the purpose of the analysis is to establish the formula of a curve whose shape has been specified, not to approximate a relationship revealed by a set of data embodying a random component. Consequently, goodness of fit may best be judged by plotting the fitted function and assessing visually how well it matches the elicited utility values. In some cases a good fit can be obtained by use of spliced polynomials, as described by Fuller (1969).

In the second case, where a mathematically tractable function is required, more care is needed in choice of an appropriate functional form. Qualitative constraints on the form of the function should be considered as well as the requirements for the decision analysis to be performed subsequently. Again least squares regression may often be the most convenient estimation procedure to use, but in this case the analyst may have to be satisfied with a less close fit to the utility points. Subjective visual appraisal again is the best test of goodness of fit, but in some applications sensitivity analysis may be called for to evaluate the consequences of using different algebraic forms for the utility function.

As a practical matter, if algebraic specification is needed, the essential aspect is to obtain an estimate that is (subjectively) judged to fit the elicited utility points satisfactorily over the relevant range of gains and losses. A variety of different functional forms may be suitable (e.g., polynomial, logarithmic, or exponential), in which case the easiest to manipulate should be used. Often this will imply the use of a polynomial of second or third degree.

Polynomial Specification

The use of a polynomial to represent the utility function for gains and losses can often be justified on the grounds that it is a Taylor series approximation to the unknown true utility function over the relevant range. Thus if $U(x)$ has a finite nth derivative $U_n(x)$ for all x and $U_{n-1}(x)$ is continuous everywhere, Taylor's theorem states that for any x^* and every $x \neq x^*$ a point ξ exists in the interval joining x and x^* such that $U(x)$ may be approximated within a specified bound of error as

$$U(x) = U(x^*) + \sum_{k=1}^{n-1} [U_k(x^*)(x - x^*)^k/k!] + U_n(\xi)(x - x^*)/n! \qquad (4.11)$$

Collecting like powers of x, $U(x)$ may thus be approximated as a polynomial

$$U(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \cdots \qquad (4.12)$$

Since $U(x)$ is only defined up to a positive linear transformation, equation

(4.12) may be written as

$$U(x) = x + bx^2 + cx^3 + \cdots \tag{4.13}$$

Further, by the expected utility theorem we have

$$U(x) = E(x + bx^2 + cx^3 + \cdots) = E(x) + bE(x^2) + cE(x^3) + \cdots \tag{4.14}$$

so that utility can be expressed relative to a risky prospect with a consequence x following some probability distribution $f(x)$. For any such random variable, $E(x^n)$ can be expressed in terms of the first n moments about $E(x)$. For example if $M_k(x)$ denotes the kth moment $E[x - E(x)]^k$ about the mean, the first three terms of (4.14) can be written as

$$U(x) = E(x) + b\{M_2(x) + [E(x)]^2\} + c\{M_3(x) + 3E(x)M_2(x) + [E(x)]^3\} + \cdots \tag{4.15}$$

This expression approximates the utility of any risky prospect $f(x)$ as a function of its mean and its moments about the mean. Often this will be more convenient than direct use of the probability distribution as implied by (4.3). As explained in Section 4.6, an alternative approach based on the Taylor series expansion of $U(x)$ about $E(x)$ is also possible and often useful.

If a polynomial such as equation (4.13) is to be fitted, which will often be a good first step to obtain a continuous functional form for $U(x)$, the question arises as to how many terms should be included. What degree should the polynomial be? An immediate answer is that as many powers of x should be used as are needed to fit the elicited utility points satisfactorily. On a pragmatic basis, such a procedure then specifies how many moments of $f(x)$ are taken into account by the decision maker, who may also be presented with a variety of $f(x)$ distributions to ascertain which moments are relevant to him. If he considers only mean and variance, i.e., $E(x)$ and $M_2(x)$, a quadratic polynomial is relevant; if skewness is also considered so that $M_3(x)$ plays some part in his choices, a cubic polynomial is relevant. As an empirical matter, it seems that the ith moment is more influential than the $(i + 1)$th and that, for most decision makers, moments beyond the third one play no great role in choice. Also, particularly if the utility points follow an S-shaped pattern, it may often be expedient to fit a pair of spliced quadratic or cubic polynomials.

If the utility function is quadratic, we have

$$U(x) = x + bx^2 \tag{4.16}$$

and the restriction $dU/dx > 0$ necessitates

$$x > -1/2b \quad \text{if} \quad b > 0 \qquad x < -1/2b \quad \text{if} \quad b < 0$$

Within these ranges, x is the certainty equivalent of all risky prospects whose utility is equal to $U(x)$. The second derivative of the quadratic shows that $b > 0$ implies increasing marginal utility as x increases; $b < 0$ implies decreasing marginal utility as x increases; and if $b = 0$, $U(x)$ is linear and marginal utility is constant as x increases.

If x is a risky prospect, the quadratic may be written as

$$U(x) = E(x) + b[E(x)]^2 + bM_2(x) \tag{4.17}$$

where $M_2(x)$ is the variance of x. Since $M_2(x)$ is necessarily positive and $\partial U/\partial M_2(x) = b$, increasing marginal utility for x (i.e., $b > 0$) implies that variability in x is attractive. Conversely, diminishing marginal utility for x (i.e., $b < 0$) implies that variability in x is disliked. Thus if $b > 0$, the decision maker is a risk preferrer; if $b < 0$, he is a risk averter; if $b = 0$, he is risk indifferent.

The expected value of the risky prospect x is $E(x)$. In the quadratic case its utility is

$$U[E(x)] = E(x) + b[E(x)]^2 \tag{4.18}$$

Comparison with equation (4.17) shows $U(x) - U[E(x)] = bM_2(x)$. Since $M_2(x)$ is necessarily positive, with quadratic preference the utility of a risky prospect is greater than, equal to, or smaller than the utility of its expected value according to whether the decision maker is a risk preferrer, risk indifferent, or a risk averter. As would be expected on more general grounds, these relations imply that the certainty equivalent of a risky prospect will, with the quadratic, be greater than, equal to, or smaller than its actuarial or mean value $E(x)$ according to whether risk preference, indifference, or aversion respectively prevails.

A similar type of relationship can be derived for the cubic function

$$U(x) = x + bx^2 + cx^3 \tag{4.19}$$

For dU/dx to be positive over the whole range of the cubic function, it must have $3c - b^2 > 0$. If these restrictive requirements are met, the shape of the function necessarily involves an initial stage of decreasing marginal utility followed beyond the inflexion point at $x = -b/3c$ by a final stage of increasing marginal utility. If marginal utility is first increasing and then decreasing, the only cubic functions that can be used will not be increasing everywhere; i.e., their relevant range will be restricted. However, such patterns of changing marginal utility are unlikely to be encountered in careful practice.

Though simple polynomial utility functions (particularly the quadratic) have often proved useful in empirical application, they have been strongly criticized on three theoretical grounds. First, polynomials are not everywhere monotonically increasing. Operationally, however, this is seldom a problem. Empirical utility functions are estimated over a particular range of gains or losses, and no one would recommend their use beyond that range. Nor would anyone with his wits about him use a polynomial approximation outside the range where it is monotonically increasing.

Second, a polynomial of degree n implies that only the first n moments of the probability distribution of outcomes are taken into account. Thus the quadratic allows only the mean and variance of a risky prospect to play any part in choice. This can never lead to error if the decision maker's utility function is "truly" quadratic or if the risky prospect's distribution is normal. Otherwise, use of the quadratic may lead to error, but not necessarily so. While the utility of each risky prospect may be wrongly assessed, the overall set of prospects may still be correctly ranked depending on how influential the higher moments are.

Third, polynomial functions for utility of wealth fail to meet the intuitive requirement of decreasing risk aversion with increasing wealth. While this may be a significant fault in utility of wealth functions, it is not too serious in terms of utility functions for gains and losses about a given wealth position because we can rationalize a risk-averse polynomial utility function for gains and losses as simply a local approximation of a decreasingly risk-averse utility function for wealth. To see this, we write the utility of wealth in the manner of equation (4.11) as a Taylor series expansion about a given level of wealth (say current wealth W_0) as

$$U(W) = U(W_0) + U_1(W_0)(W - W_0) + U_2(W_0)(W - W_0)^2/2! + U_3(W_0)(W - W_0)^3/3! + \cdots \tag{4.20}$$

wealth current income

Recognizing that $W - W_0 = x$ (the extent of gain or loss from current wealth), if we now make the particular positive linear transformation of subtracting $U(W_0)$ and dividing by $U_1(W_0)$, we have $U(W)$ expressed as a function of x, $U(x)$, where

$$U(x) = x + [U_2(W_0)/U_1(W_0)]x^2/2 + [U_3(W_0)/U_1(W_0)]x^3/6 + \cdots \tag{4.21}$$

This utility function is precisely analogous to the standardized polynomial of equation (4.13) but with $b = [U_2(W_0)/U_1(W_0)]/2$, which, from (4.10) we see is equivalent to $b = -r(W_0)/2$; i.e., b equals minus half the Pratt coefficient evaluated at current wealth. As current wealth increases, decreasing risk aversion will be appropriately reflected in the local polynomial utility function expressed in terms of gains and losses. Operationally, this poses no difficulty since, with use of the ELCE or ELRO method $U(x)$ can be easily obtained directly at a particular point in time and wealth for a given decision maker. From a practical point of view, the approach fits in well with the pragmatic procedure of regarding a polynomial as an approximation to the unknown true utility function, recognizing that a new utility function for gains and losses should be assessed whenever the decision maker's asset situation changes significantly.

Relative to equation (4.21), note further that if the derivatives beyond the second are sufficiently small to be ignored, the utility function approximation will be quadratic.

Quadratic Utility and (E, V) Analysis

With $U(x)$ quadratic, discussion is often presented in terms of mean-variance or (as it has come to be designated) (E, V) analysis. Of course, E is $E(x)$ and V is $M_2(x)$. For convenience, we will write equation (4.17) as

$$U = E + bE^2 + bV \tag{4.22}$$

Equation (4.22) implies a utility surface in the three dimensions U, E, and V. For constant values of U, the function can be represented by a series of isoutility contours or indifference curves in (E, V) space. Thus on setting $U(x)$ equal to some constant U^*, rearrangement gives

$$V = U^*/b - E/b - E^2 \tag{4.23}$$

as the (E, V) locus of all mean-variance combinations that yield the same level of utility. Corresponding to the relevant range of the quadratic, the relevant range of the isoutility loci is $E \geqq -1/2b$ for $b > 0$, and $E \leqq -1/2b$ for $b < 0$, with $V \geqq 0$ in both cases. Note also that the intercept of an isoutility curve with the E axis (i.e., where $V = 0$) is the certainty equivalent of all mean-variance combinations on that indifference curve. If the isoutility curves are drawn in mean-standard deviation space (i.e., with axes measuring E and $V^{0.5}$), they will be concentric circles with the center at $E = -1/2b$, $V = 0$.

The decision maker's tradeoff or substitution rate between mean and variance is given by the slope of the isoutility (E, V) curve, which is

$$dE/dV = -b/(1 + 2bE) \tag{4.24}$$

Since $(1 + 2bx) = dU/dx$ and must be positive, its expected value $(1 + 2bE)$ must be positive also. Hence dE/dV will be positive, zero, or negative within the relevant range according to whether b is negative, zero, or positive. As is intuitively obvious, a risk averter will need increases in mean value to compensate for increased variance if his utility is to remain unchanged.

The second derivative of the isoutility curve is

$$d^2E/dV^2 = [2b^2(1 + 2bE)^{-2}](dE/dV) \tag{4.25}$$

The term in square brackets is always positive, and dE/dV is positive or

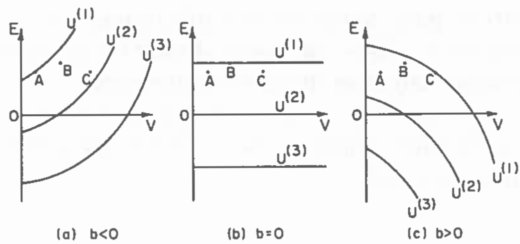


FIG. 4.6. The (E, V) indifference or isoutility curves for $U = E + bE^2 + bV$.

negative over the relevant range according to whether b is negative or positive. Hence for a risk averter the (E, V) indifference curves have increasing slope (i.e., the tradeoff rate increases) as V increases, and for a risk preferrer the (E, V) indifference curves have increasing negative slope as V increases. The greater the degree of risk aversion or preference (i.e., the greater $|b|$ is), the steeper the indifference curves.

The above relationships are portrayed diagrammatically in Figure 4.6: (a) with risk aversion, (b) with risk indifference, and (c) with risk preference. In each case three isoutility or (E, V) indifference curves are shown for utility levels $U^{(1)} > U^{(2)} > U^{(3)}$.

Just as the quadratic utility function may be depicted in (E, V) space, so also may the set of risky prospects from which a choice is to be made. These prospects will have moments beyond the second, but in the context of (E, V) analysis such higher moments are assumed irrelevant to choice. When depicted in (E, V) space, the preferred risky prospect is indicated as the one that lies on the highest isoutility curve. Consider, for example, the alternative prospects A, B, and C shown in Figure 4.6. For the risk averter A is optimal, B is optimal with risk indifference, and C is optimal for the risk preferrer. More detailed discussion of the depiction of risky prospects in (E, V) space is deferred until discussion of portfolio analysis in Section 7.1.

4.6 UTILITY ANALYSIS USING MOMENTS OF DISTRIBUTIONS

The approach of depicting decision problems and preferences in terms of moments of probability distributions is more general than revealed above in our discussion of polynomial utility functions.

If utility depends only on a single attribute, the utility function can be respecified as an expected utility function defined in terms of the moments of the probability distribution of the single attribute. In the case of multidimensional consequences, the utility function can be respecified in terms of a single conditionally referenced attribute as outlined in Section 4.3. This "moment method" is exactly equivalent to the direct method for some utility functions; for others it serves as a fair but useful approximation. The computation of most decision analyses is substantially reduced through using the moment method, particularly when many states must be considered. In consequence, even when it provides only an approximation in utility evaluations, the moment method deserves attention for the increased feasibility it can impart to analytical procedures.

The basis of the moment method is a Taylor series expansion. We take first the utility function for gains and losses $U(x)$ as in equation (4.11), but with $x^* = E(x)$ so that the expansion is taken about the mean. Thus we have

$$U(x) = U[E(x)] + U_1[E(x)][x - E(x)] + U_2[E(x)][x - E(x)]^2/2! + U_3[E(x)][x - E(x)]^3/3! + \cdots \tag{4.26}$$

Using the expected utility theorem and taking the expectation of equation (4.26), the utility of the risky prospect x is

$$U(x) = U[E(x)] + U_1[E(x)]E[x - E(x)] + U_2[E(x)] \cdot E[x - E(x)]^2/2 + U_3[E(x)]E[x - E(x)]^3/6 + \cdots$$

Recalling that $E[x - E(x)] = 0$ and that the kth moment about the mean $M_k(x) = E[x - E(x)]^k$, this equation can be written as

$$U(x) = U[E(x)] + U_2[E(x)]M_2(x)/2 + U_3[E(x)]M_3(x)/6 + \cdots \tag{4.27}$$

Thus the utility of a risky prospect x is equal to the utility function evaluated at the mean of x plus a series of products of moments of x, corresponding derivatives of the utility function and inverse factorials, other than that involving the first derivative. The approximation is better for "tight" distributions (i.e., for those with relatively small variance) and improves with the number of terms included. The number of terms it is necessary to retain in any application depends on the individual problem and especially on the closeness of the utilities of alternative acts.

It is usually found that because $U_k[E(x)]/k!$ becomes smaller at a rather faster rate than $M_k(x)$ becomes larger as k increases, terms beyond those involving the third moment add insignificantly to the precision of the approximation of equation (4.27). More specifically, it seems that terms beyond the first three add insignificantly to the precision of the approximation when Pratt's coefficient of risk aversion does not exceed a magnitude of the order of one tenth of the inverse of the standard deviation of the risky prospect's probability distribution. Accordingly, we will only use products up to the term involving $M_3(x)$ in the following examples in which equation (4.27) is applied to several algebraic forms of utility functions.

Quadratic. The quadratic function is the simplest nonlinear form to manipulate, but it has already been indicated as having some limitations both as an empirical and as a theoretical utility function. By a suitable positive linear transformation, any quadratic is equivalent to $U = x + bx^2$ where, for risk aversion, $b < 0$. Its derivatives are $U_1 = 1 + 2bx$, $U_2 = 2b$, and the third and beyond vanish so that substituting into equation (4.27) provides the exact expected utility function involving only the mean and variance,

$$U = U[E(x)] + 2bM_2(x)/2 = E(x) + b[E(x)]^2 + bV(x) \qquad (4.28)$$

where, to accord with standard notation, $V(x) = M_2(x)$. This is precisely the result obtained earlier in (4.17).

Cubic. All polynomial functions lead to simple results analogous to that noted for the quadratic. A kth order polynomial leads to an expected utility function involving moments only up to the kth since corresponding derivatives beyond this vanish. The cubic case serves to illustrate this fact: $U = x + bx^2 + cx^3$ has $U_1 = 1 + 2bx + 3cx^2$, $U_2 = 2b + 6cx$, and $U_3 = 6c$. Substitution into equation (4.27) gives

$$\begin{aligned} U &= U[E(x)] + [2b + 6cE(x)]V(x)/2 + 6cM_3(x)/6 \\ &= E(x) + b[E(x)]^2 + c[E(x)]^3 + bV(x) + 3cE(x)V(x) + cM_3(x) \\ &= E(x) + b\{[E(x)]^2 + V(x)\} + c\{[E(x)]^3 + 3E(x)V(x) + M_3(x)\} \end{aligned} \qquad (4.29)$$

The foregoing polynomial functions are potentially useful for encoding preferences for gains and losses; but as noted in Section 4.4 they have various limitations as utility of wealth functions. The converse situation applies for the following three nonpolynomial functions, and accordingly we will specify them as utility of wealth functions. The connection with moments of returns on gains and losses is preserved by defining [as we did in discussing equation (4.20)] the random variable wealth W as the sum of (known) wealth before a decision W_0 and the gain or loss outcome x of a risky decision problem (i.e., $W = W_0 + x$). It follows that $E(W) = W_0 + E(x)$, $V(W) = V(x)$, $M_3(W) = M_3(x)$, etc., so that in expanding about $E(W)$, (4.27) holds if everywhere $E(x)$ is replaced by $W_0 + E(x)$.

Logarithmic. The utility function $U = \log_e W$ has derivatives $U_1 = 1/W$, $U_2 = -1/W^2$, and $U_3 = 2/W^3$, so expected utility can be approximated by

$$U = \log_e E(W) - \{1/[E(W)]^2\} V(W)/2 + \{2/[E(W)]^3\} M_3(W)/6$$
$$= \log_e [W_0 + E(x)] - (1/2) V(x)/[W_0 + E(x)]^2$$
$$+ (1/3) M_3(x)/[W_0 + E(x)]^3 \qquad (4.30)$$

Note that the ratio of moments in the second term is the square of the coefficient of variation of wealth and in the third term is a measure of relative skewness.

Power. A utility analog of the Cobb-Douglas response function is $U = W^c$, $0 < c < 1$, with derivatives $U_1 = cW^{c-1}$, $U_2 = c(c-1)W^{c-2}$, and $U_3 = c(c-1)(c-2)W^{c-3}$, so expected utility can be approximated using the wealth variant of equation (4.27) as

$$U = [W_0 + E(x)]^c + c(c-1)[W_0 + E(x)]^{c-2} V(x)/2$$
$$+ c(c-1)(c-2)[W_0 + E(x)]^{c-3} M_3(x)/6 \qquad (4.31)$$

Negative exponential. The constant risk aversion function $U = 1 - e^{-cW}$ has derivatives $U_1 = ce^{-cW}$, $U_2 = -c^2e^{-cW}$, and $U_3 = c^3e^{-cW}$. As long as the risk aversion coefficient c is small relative to the variance and higher moments (say, so that $c^2V(x) < 0.01$), expected utility may be approximated by

$$U = 1 - e^{-cE(W)} - c^2e^{-cE(W)} V(W)/2 + c^3e^{-cE(W)} M_3(W)/6$$
$$= 1 - e^{-c[W_0 + E(x)]}[1 + c^2V(x)/2 - c^3M_3(x)/6] \qquad (4.32)$$

Such Taylor series approximations will be most useful for analysis of problems featuring continuous random variables or many states, as shown in Chapters 6 and 7. The approach is also of general use in the evaluation of limits to the value of information as discussed in Section 5.3. Since the only moment changed by subtracting a cost of information is $E(x)$, the use of a Taylor series expansion can provide a ready means of obtaining good approximations to information value limits.

4.7 SUMMARY REMARKS

The idea of encoding a decision maker's preferences in the form of a utility function is a simple one, but it is very important for ensuring consistency of preference in conditionally normative analyses of choices. We have presented some practical suggestions for eliciting utility functions, and with use of such techniques we see little difficulty in dealing with nonlinear preference of individuals. An "individual" might also refer to a group of decision makers acting in complete accord on the basis of agreed common preferences. Unfortunately, for other more realistic group decision situations the approach we have adopted cannot be simply applied. We explain why in Section 5.8.

Observation at the individual level has wider implications. The bulk of empirical evidence so far gathered suggests that most decision makers are risk averse. Generalizing from such empirical observations leads to at least two important and related statements about nonelicited preference and appraisal of individual economic efficiency. First, we could never establish preferences for all the decision makers in a community, and the mind boggles even at determining the utility functions of all the farmers in one village. But if we can be fairly safe in assuming that most decision makers are risk averse, it obviously makes no sense to assume, say, an unqualified goal of profit maximization in an analysis intended to aid attainment of farmers' goals. The neoclassical theory of the firm with its assumptions of certainty and linear utility is inadequate for normative analysis of risky production where preference for profits is nonlinear. An approach involving decision analysis methods is clearly more appropriate. If for some reason individual preference cannot be determined, an arbitrary assumption of a theoretically sound decreasingly risk-averse utility function for assets (such as "Everyman's function" $U = \log_e W$) is more defensible than arbitrarily assuming linear preference and ignoring risk.

A closely related issue concerns the descriptive use of economic theory founded on the notion of certainty. For instance, it has been a popular pastime of agricultural economists to estimate empirical production functions to judge resource-use efficiency in a profit-maximizing sense. This frame of reference may have some relevance from a national point of view, but in the light of our comments there is no reason why a farmer should want to operate at any position other than his subjective utility-maximizing position. Without special encouragement this may diverge from an average profit-maximizing position or some other perhaps claimed as optimal by purveyors of new technology.

Supplier

PROBLEMS

4.1. First determine your utility function and then that of a friend for money gains and losses in the range from −\$500 to \$500.

4.2. (a) If X denotes net return per technical unit (e.g., per hectare or per sow), under what conditions will $U(kX) = kU(X)$ where k is a constant? What does this imply about the way in which financial payoff matrices must be expressed for decision analysis?

(b) In the neoclassical theory of the competitive farm firm, fixed costs play no role in determining the optimal level of inputs for profit maximization. Show that if the farmer's aim is to maximize utility with respect to net profit, fixed costs must be taken into account in determining optimal input levels if he is not indifferent to risk or is not constantly risk averse.

(c) It is sometimes argued that while we can say $U(A)$ is greater than $U(B)$ if A is preferred to B, it is fallacious to say A is preferred to B because $U(A)$ is greater than $U(B)$. What comment would you make?

4.3. A farmer's utility function for money gains and losses is approximately represented by $U(X) = 2.05X - 0.01X^2$, $(X \leqq 80)$, where X denotes thousands of dollars. The farmer is currently wondering whether to spend more on fertilizer for his 1000 ha of crop than last season's \$4/ha. Pertinent information is shown in the following payoff matrix of possible dollar profits per hectare.

		Possible Actions			
Type of Season	Probability	Spend \$4/ha	Spend \$8/ha	Spend \$12/ha	Spend \$16/ha
			(\$/ha)		
Poor	0.1	−8	−12	−16	−20
Fair	0.2	−2	− 8	−12	−16
Good	0.5	2	4	6	8
Excellent	0.2	12	20	24	24

(a) How much should the farmer spend on fertilizer?
(b) What is his risk premium for each of his possible acts?
(c) Verify that the expected utilities when computed by the moment method are identical with those estimated by the direct method.

4.4. Plot the utility function specified in Problem 4.3 for the range $-20 \leqq X \leqq 30$ and show graphically the risk premium for the possible action of spending \$16/ha on fertilizer.

4.5. If you were offered a choice between bets A and B, which would you choose?

Bet A: You win \$1,000,000 for sure.
Bet B: You win \$5,000,000 with probability 0.10.
You win \$1,000,000 with probability 0.89.
You win \$0 with probability 0.01.

Now choose between bets C and D.

Bet C: You win \$1,000,000 with probability 0.11.
You win \$0 with probability 0.89.
Bet D: You win \$5,000,000 with probability 0.10.
You win \$0 with probability 0.90.

This problem is known as Allais' paradox (Borch, 1968).

If you chose bet A, you should have also chosen C. Prove this. If you chose bet B, you should have also chosen bet D. Prove this. Comment on the role of utility analysis as a means of achieving consistency in choice relative to given preferences in complicated decision problems.

4.6. She'llbejake Ltd., the construction subsidiary of Gundy Pastoral, does subcontracting on government beef-road contracts. The construction company's utility function is approximately represented by $U(X) = 2X - 0.01X^2$, $(X \leqq 100)$, X being thousands of dollars.

(a) Suppose She'llbejake is considering bidding on a contract. Preparation of a bid would cost $8000, and this would be lost if the bid failed. If the bid succeeded, She'llbejake would make $40,000 net gain. If She'llbejake judges the chance of a successful bid as 0.3, what should the company do?

(b) What chance of a successful bid would make the company indifferent between bidding and not bidding for the contract?

4.7. Comment on the following statements:

(a) Because it can take a broad view and weather any resultant financial storm, the government should be indifferent to risk in appraising alternative investment projects.

(b) Since most farmers are risk averse, they will use variable inputs at sub-optimal levels from the point of view of maximizing expected profit. It would therefore be in the national interest to subsidize variable inputs such as fertilizer.

(c) As a farmer becomes wealthier, he will operate his farm more intensively.

(d) The familiar S-shaped curve showing the rate of adoption of new agricultural techniques is explained by the fact that a few farmers are risk preferrers, while the bulk are risk averse.

4.8. (a) What are the implications of the fact that the uniqueness of the utility function is defined only up to a positive linear transformation?

(b) What influences do you think might lead to a change in a person's utility function?

(c) Might it be said that every utility function is lexicographic to some extent?

4.9. The management of Gundy Pastoral Company is considering the establishment of a cattle feedlot on the outskirts of Darwin. While the project looks quite profitable with a gross margin of $30 per head and a fixed cost of $20,000 over the relevant range of throughput, there would be a significant olfactory pollution problem. How would you go about ascertaining the management's utility for various alternative feedlot sizes if consequences were to be evaluated in terms of both annual profit and pollution? The range of sizes being considered is from an annual throughput of 1000 head up to 8000 head. In purely physical terms (without taking account of the effect on Gundy Pastoral's public image) it is thought that the olfactory pollution problem would treble with every doubling of annual throughput.

4.10. Suppose the property purchase decision problem of Section 4.3 had also involved the following alternatives:

	a_4				a_5			
θ_i	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
θ_1	12	90	100	200	3	96	80	150
θ_2	6	90	100	200	2	96	80	150
θ_3	1	90	100	200	1	96	80	150

(a) What should be our client's order of preference across the five alternatives?

(b) What if the relevant k_i values were $k_1 = 0.3$, $k_2 = 0.2$, $k_3 = 0.1$, and $k_4 = 0.3$?

(c) What if our client had a lexicographic ordering with dominant priorities of x_4 over x_2 over x_1 over x_3?

4.11. Comment on the algebraic implications and draw a graph of each of the following utility functions where W represents wealth and the functions are proposed for $0 < W < 1000$.

(a) $U(W) = (W + 1000)^2$

(b) $U(W) = -(1{,}000 - W)^2$

(c) $U(W) = \log_e W$

(d) $U(W) = \log_e(W + 100)$

(e) $U(W) = 1 - e^{-W/100}$

(f) $U(W) = W^\beta, 0 < \beta < 1, \beta = 0.5$

4.12. (a) Plot the (E, V) and $(E, V^{1/2})$ indifference systems implied by the utility function of Problem 4.3.

(b) Consider a risky prospect that yields a gain of \$6150 with probability 0.8 and a gain of \$11,500 with probability 0.2.

(1) Calculate the first three moments of payoff from this prospect.

(2) Scale the utility-of-gain function $U = x - (5)10^{-5}x^2$ so that $U(6150) = 90$ and $U(11{,}500) = 150$ and use the moment method of approximation to compute the expected utility of the prospect.

(3) Scale the utility-of-wealth function $U = \log_{10} W$ with initial wealth $W_0 = 50{,}000$ so that $U(W_0 + 6150) = 90$ and $U(W_0 + 11{,}500) = 150$ and use the moment method of approximation with the first three moments to compute the expected utility of the prospect. (Recall that the first derivative of $U = a + b \log_{10} W$ is $U_1 = 0.4343b/W$.)

REFERENCES AND SELECTED FURTHER READING

Anderson, J. R. 1973. Risk aversion and polynomial preference. *Australian Econ. Pap.* 12(21): 261–62.

Presents a justification for the use of polynomial approximations as utility functions for gains and losses about a given wealth position.

Anderson, J. R., and J. L. Dillon. 1968. Economic considerations in response research. *Am. J. Agric. Econ.* 50(1): 130–42.

Uses (E, V) analysis for the appraisal of the value of conducting further research.

Arrow, K. J. 1971. *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland.

Chapter 1 of this collection of essays by one of today's foremost contributors presents a masterly review of the history of alternative approaches to risky choice. Risk aversion and its implications are discussed in Chapter 3. Later chapters develop the implications of utility analysis for risky decisions in research, medicine, insurance, and public investment.

Baumol, W. J. 1963. An expected gain-confidence limit criterion for portfolio selection. *Mgmt. Sci.* 10(2): 174–82.

One of the first discussions of a lexicographic approach based on safety-first considerations.

Borch, K. H. 1968. *The Economics of Uncertainty*. Princeton: Princeton Univ. Press.

Chapters 2 to 5 provide an easily read but elegant treatment of Bernoulli's principle.

Champernowne, D. G. 1969. *Uncertainty and Estimation in Economics,* Vol. 3. Edinburgh: Oliver and Boyd.

A wide-ranging discussion of uncertainty and Bernoulli's principle with Chapters 19 and 20 of particular relevance.

Cochrane, J. L., and M. Zeleny, eds. 1973. *Multiple Criteria Decision Making*. Columbia: Univ. South Carolina Press.

An interesting potpourri of papers on multidimensional decision making.

Day, R. H., D. J. Aigner, and K. R. Smith. 1971. Safety margins and profit maximization in the theory of the firm. *J. Pol. Econ.* 79(6): 1293–301.

Considers the lexicographic approach of profit maximization subject to a safety-first requirement.

de Neufville, R., and R. L. Keeney. 1972. Use of decision analysis in airport development for Mexico City. In *Analysis of Public Systems,* A. W. Drake et al., eds., pp. 497–519. Cambridge: MIT Press.

Discusses the same problem as Keeney (1973b) but with some consideration of the broader social and political aspects.

Dean, G. W. 1974–75. Firm theory incorporating growth and risk: Integration into farm management research. Suppl. *Int. J. Agr. Aff.,* double number, pp. 110–20.

Argues for the use of lexicographic utility orderings in investment decisions pertinent to farm-firm growth.

Diamond, P. A., and J. E. Stiglitz. 1974. Increases in risk and in risk aversion. *J. Econ. Theory* 8(3): 337–60.

Extends the results of Rothschild and Stiglitz (1970) to include the comparison of risks based on expected utility-preserving spreads.

di Roccaferrera, G. M. F. 1973. Behavioral aspects of decision making under multiple goals. In *Multiple Criteria Decision Making,* J. L. Cochrane and M. Zeleny, eds., pp. 635–56. Columbia: Univ. South Carolina Press.

Of interest for its survey of a variety of behavior-based approaches to multiple goal decisions. Includes motivation, measurement, and goal programming models.

Dillon, J. L., and J. R. Anderson. 1971. Allocative efficiency, traditional agriculture, and risk. *Am. J. Agric. Econ.* 53(1): 26–32.

Reviews empirical evidence regarding profit-maximizing behavior in traditional farming and discusses efficiency in the context of utility maximization.

Dillon, J. L., and A. G. Lloyd. 1962. Inventory analysis of drought reserves for Queensland graziers: Some empirical analytics. *Australian J. Agric. Econ.* 6(1): 50–67.

Farmers' alternative choices on fodder reserves are appraised in the framework of (E, V) analysis.

Drèze, J. H. 1974. Axiomatic theories of choice, cardinal utility and subjective probability: A review. In *Allocation under Uncertainty,* J. H. Drèze, ed., pp. 3–23. London: Macmillan.

An excellent review of utility and associated theory.

Dyckman, T. R., S. Smidt, and A. K. McAdams. 1969. *Management Decision Making under Uncertainty.* London: Collier-Macmillan.

Appendix 11A of this text provides a proof of Bernoulli's principle that is less complex than most other published proofs.

Easton, A. 1973. *Complex Managerial Decisions Involving Multiple Objectives.* New York: Wiley.

A very readable compendium of practical business approaches to multiple objective decisions.

Epstein, R. A. 1967. *Theory of Gambling and Statistical Logic.* New York: Academic Press.

Chapter 3 discusses utility in the context of its fundamental importance in the theory of gambling.

Ferguson, C. E. 1965. The theory of multidimensional utility analysis in relation to multiple-goal business behavior: A synthesis. *Southern Econ. J.* 32(2): 169–75.

Provides a good outline of lexicographic utility theory and argues its potential as a theory of multiple-goal business behavior.

Fischer, G. W. 1975. Experimental applications of multi-attribute utility models. In *Utility, Probability, and Human Decision Making,* D. Wendt and C. Vlek, eds., pp. 7–46. Dordrecht: Reidel.

A thorough review indicating the general adequacy of the additive utility model.

Fishburn, P. C. 1967. Methods of estimating additive utilities. *Mgmt. Sci.* 13(7): 435–53.

Presents 24 methods for estimating additive multidimensional utility functions.

———. 1970. *Utility Theory for Decision Making.* New York: Wiley.

A mathematically formal presentation by one of the most prolific current workers on utility and decision analysis.

———. 1973. Bernoullian utilities for multiple-factor situations. In *Multiple Criteria Decision Making,* J. L. Cochrane and M. Zeleny, eds., pp. 47–61. Columbia: Univ. South Carolina Press.

Complementary to Keeney's papers but more formal in presentation, this establishes the conditions under which attribute interaction may be allowed for through utility functions involving sums and weighted products of component utilities.

———. 1974. Von Neumann-Morgenstern utility functions on two attributes. *Ops. Res.* 22(1): 35–45.

Examines five different forms of the two-dimensional utility function.

Freund, R. J. 1956. The introduction of risk into a programming model. *Econometrica* 24(2): 253–63.

A pioneering application of utility analysis in farm planning. Uses the negative exponential utility function.

Friedman, M., and L. J. Savage. 1948. Utility analysis of choices involving risk. *J. Polit. Econ.* 56(4): 279–304.

A classic essay in positive economics that rationalizes people's purchase of both insurance and lottery tickets in terms of the expected utility hypothesis.

Fuller, W. A. 1969. Grafted polynomials as approximating functions. *Australian J. Agric. Econ.* 13(1): 35–46.

Halter, A. N., and G. W. Dean. 1971. *Decisions under Uncertainty with Research Applications.* Cincinnati: South-Western.

Chapter 8 includes empirical examples of the use of (E, V) analysis and lexicographic utilities in agricultural decision making.

Hillier, F. S. 1969. *The Evaluation of Risky Interrelated Investments.* Amsterdam: North-Holland.

The use of a Taylor series approximation leading to the moment method of analysis and the formulation of hyperbolic and exponential utility functions is outlined in Chapter 5.

Hirshleifer, J. 1970. *Investment, Interest, and Capital.* Englewood Cliffs, N.J.: Prentice-Hall.

Parts A and B of Chapter 8 provide some good discussion of the postulates underlying Bernoulli's principle and of the general concept of a utility function.

Huber, G. P. 1974a. Multiattribute utility models: A review of field and field-like studies. *Mgmt. Sci.* 20(10): 1393–402.

———. 1974b. Methods for quantifying subjective probabilities and multiattribute utilities. *Decn. Sci.* 5(4): 430–58.

Hull, J., P. G. Moore, and H. Thomas. 1973. Utility and its measurement. *Jl. R. Statist. Soc.* Ser. A, 136(2): 226–47.

A survey article covering a variety of approaches to both single and multidimensional utility function measurement. Complements Huber (1974a and b).

Keeney, R. L. 1968. Quasi-separable utility functions. *Nav. Res. Logist. Q.* 15(4): 551–65.

———. 1972. Utility functions for multiattributed consequences. *Mgmt. Sci.* 18(5): 276–87.

———. 1974. Multiplicative utility functions. *Ops. Res.* 22(1): 22–34.

———. 1973a. Decomposition of multiattribute utility functions. Tech. Rept. 80, MIT Operations Res. Center, Cambridge, Mass.

———. 1973b. A decision analysis with multiple objectives: The Mexico City airport. *Bell J. Econ. Mgmt. Sci.* 4(1): 101–17.

The first four of the above articles by Keeney show the development of the approach to multiattribute utility based on the assumptions of preference independence and utility independence. The fifth article, a companion to de Neufville and Keeney (1972), brilliantly illustrates the application of this theory to a significant real-world problem.

Keeney, R. L., and A. Sicherman. 1975. An interactive computer program for assessing and analyzing preferences concerning multiple objectives. Rept. RM-75-12, Intl. Instit. Applied Systems Analysis, Laxenburg.

Keeney, R. L., and H. Raiffa. 1976. *Decisions with Multiple Objectives: Preferences and Tradeoffs.* New York: Wiley.

An exposition of multiattribute utility analysis by its two foremost exponents.

Lindley, D. V. 1971. *Making Decisions.* London: Wiley.

In contrast to the approach of using a utility function for gains and losses about the current wealth position, Chapter 5 presents an easily read discussion of utility couched in terms of the utility function for wealth.

MacCrimmon, K. R. 1973. An overview of multiple objective decision making. In *Multiple Criteria Decision Making,* J. L. Cochrane and M. Zeleny, eds., pp. 18–44. Columbia: Univ. South Carolina Press.

Under the headings of weighting methods, sequential elimination methods, mathematical programming methods, and spatial proximity methods, this article provides an excellent review of 19 different approaches to multiattribute decision making. Both descriptive and prescriptive approaches are covered.

McCall, J. J. 1971. Probabilistic microeconomics. *Bell J. Econ. Mgmt. Sci.* 2(2): 403–33.

A thorough review of the role of utility in risky microeconomics with particular emphasis on measures of risk and the theory of the firm. Includes an extensive bibliography.

Meyer, R. F., and J. W. Pratt. 1968. The consistent assessment and fairing of preference functions. *IEEE Trans. Syst. Sci. Cybern.* SSC-4(3): 270–78.

An advanced mathematical discussion of fitting decreasingly risk-averse utility functions with the aid of nonlinear programming methods.

Officer, R. R., and A. N. Halter. 1968. Utility analysis in a practical setting. *Am. J. Agric. Econ.* 50(2): 257–77.

Reports practical experience with using three elicitation procedures including the ELCE and ELRO methods. Also presents evidence that farmers' decisions were more accurately predicted by expected utility maximization than by expected profit maximization.

Porter, R. C. 1959. Risk, incentive and the technique of the low-income farmer. *Indian Econ. J.* 7(1): 1–27.

This relatively early article uses an (E, V) framework of analysis to argue strongly the relevance of utility considerations for agricultural development policies.

Pratt, J. W. 1964. Risk aversion in the small and in the large. *Econometrica* 32(1–2): 122–36.

The seminal work on risk aversion, but most of the discussion is rather mathematical.

Pyle, D. H., and S. J. Turnovsky. 1970. Safety-first and expected utility maximization in mean-standard deviation portfolio analysis. *Rev. Econ. Statist.* 52(1): 75–81.

Relates lexicographic safety-first principles to (E, V) analysis.

Raiffa, H. 1968. *Decision Analysis.* Reading, Mass.: Addison-Wesley.

Provides a more discursive and nonmathematical treatment of utility and risk aversion than we have given.

———. 1969. Preferences for multi-attributed consequences. RM-5868, The RAND Corp., Santa Monica, Calif.

A perceptive, easily read, and highly recommended treatment of the utility approach to multiattribute choice.

Richard, S. F. 1975. Multivariate risk aversion, utility independence and separable utility functions. *Mgmt. Sci.* 22(1): 12–21.

Generalizes the notion of risk aversion to the multivariate case.

Rothschild, M., and J. E. Stiglitz. 1970. Increasing risk. I. A definition. *J. Econ. Theory* 2(3): 225–43.

Compares riskiness of distributions (with identical means) on the basis of the restrictive idea of mean-preserving spreads.

Roumasset, J. 1976. *Rice and Risk: Decision Making among Low-Income Farmers.* Amsterdam: North-Holland.

Chapter 2 shows the relationships that exist between the various safety-first criteria that have been suggested by Baumol (1963), Day et al. (1963), and Pyle and Turnovsky (1970).

Schlaifer, R. 1969. *Analysis of Decisions under Uncertainty.* New York: McGraw-Hill.

Chapter 3 discusses the determination of preference curves. Use is made of "the standard reference contract method" which is analogous to the ELCE procedure.

Slovic, P., and A. Tversky. 1974. Who accepts the Savage axioms? *Behav. Sci.* 19(6): 368–73.

A salutary experimental study showing that subjects persistently violated the independence axiom even after it was carefully explained to them. Contradicts the view that the axioms of decision theory are similar to the principles of logic, in the sense that no reasonable person who understands them would wish to violate them.

Tsiang, S. C. 1972. The rationale of the mean-standard deviation analysis, skewness preference and the demand for money. *Am. Econ. Rev.* 62(3): 354–71.

Provides detailed discussion of the criticisms that have been made of (E, V) analysis and shows that the two-moment method is a reasonable approximation as long as risk is small relative to the decision maker's wealth.

von Winterfeldt, D., and G. W. Fischer. 1975. Multiattribute utility theory: Models and assessment procedures. In *Utility, Probability, and Human Decision Making,* D. Wendt and C. Vlek, eds., pp. 47–85. Dordrecht: Reidel.

A thorough survey of the theory and estimation of multidimensional utility functions. An excellent complement to Keeney's papers. Lists an extensive bibliography with some bias to psychological research.

Yntema, D. B., and W. S. Torgerson. 1967. Man-computer cooperation in decisions requiring common sense. In *Decision Making,* W. Edwards and A. Tversky, eds., pp. 300–14. Harmondsworth, Middlesex, England: Penguin.

Shows that the assumption of an additive utility function generally provides a reasonable approximation in multidimensional consequence situations.

Zeckhauser, R., and E. Keeler. 1970. Another type of risk aversion. *Econometrica* 38(5): 661–65.

Draws out the implications of decreasing aversion to size of risk for preference fitting.

Zeleny, M., ed. 1976. *Multiple Criteria Decision Making, Kyoto 1975.* Berlin: Springer-Verlag.

Presents a mixed set of papers plus an extensive bibliography.

From a historical point of view, the following are the three seminal works on utility:

Bernoulli, D. Specimen Theoriae Novae de Mensura Sortis (St. Petersburg, 1738). Trans. L. Somer. 1954. Exposition of a new theory on the measurement of risk. *Econometrica* 22(1): 23–36.

Ramsey, F. P. 1964. Truth and Probability. In *Studies in Subjective Probability,* H. E. Kyburg and H. E. Smokler, eds., pp. 61–92. New York: Wiley.

von Neumann, J., and O. Morgenstern. 1947. *Theory of Games and Economic Behavior,* 2nd ed., Ch. 3 and App. Princeton: Princeton Univ. Press.

The Bernoulli family is famous as the most intellectually gifted family in history. For generations in the seventeenth and eighteenth centuries they made basic contributions to science, particularly mathematics, physics, and economics. Some of this amazing story is told in:

Bell, E. T. 1937. *Men of Mathematics.* London: Gollanz.

CHAPTER FIVE
PROCEDURES FOR DECISION ANALYSIS

WE DISCUSSED the concept of probability in Chapter 2 and the revision of probabilities in Chapter 3. The concept of utility was our main interest in Chapter 4. Probability and utility will now be brought together in a discussion of the procedures recommended for the analysis of risky decisions. In contrast to the normal form of analysis illustrated in Section 1.2, we will emphasize the so-called extensive form of analysis. In the extensive approach, prior probabilities enter at the start rather than at the end of the analysis. This cuts down the number of calculations required.

Our focus on problems involving probabilities is deliberate. In our view any attempt to analyze decision problems under uncertainty without recourse to subjective probabilities is at best an unimportant theoretical exercise; at worst it is an adventure into a world of arbitrary and inconsistent criteria for "games against Nature." As implied by our acceptance of the axioms that lead to Bernoulli's principle, we believe that probabilities can always be elicited for uncertain consequences.

In this chapter we will not give explicit attention to decision problems in which there is no possibility for gathering further information through experimentation. Such situations are merely a special and simpler case of the more general case involving experimentation. If there is no opportunity for obtaining further information, the optimal act can be found directly by applying Bernoulli's principle in terms of the prior probabilities. Our neglect of such situations does not reflect the relative empirical importance of such decision problems; decisions very often have to be made without the opportunity of seeking further information.

Although the focus of our treatment of decision analysis is on individual choice, we recognize that in reality many decisions involve more than one person. Unfortunately, there are formidable theoretical and practical difficulties involved in extending prescriptive decision analysis to multiperson situations. In Section 5.8 we try to expose some of these difficulties. We are able to offer few remedies but do try to explain why the difficulties persist.

5.1 THE GENERAL APPROACH

Previous chapters have introduced the major components of decision problems and have elaborated the specification of probabilities and utilities. Bayes' theorem and Bernoulli's principle have been outlined. The central role that they play in decision analysis will now be indicated.

Decision analysis where no experiment is possible consists of applying Bernoulli's principle and finding the act that maximizes expected utility. When there is the possibility of conducting an experiment, likelihoods are specified (subjectively or statistically) and posterior probabilities are determined with the aid of Bayes' theorem. After proper account is taken of the costs of the possible experiment, the Bayes strategy can again be determined through application of Bernoulli's principle, based this time on the appropriate posterior probabilities for each possible experimental outcome. Note that, as previously emphasized, by an experiment we mean any procedure leading to predictive information about the states. It may be a field experiment or the purchase of a forecast or the conduct of a survey, etc.

Having determined the Bayes strategy, its utility can be calculated (taking into account the cost of the experiment). If the utility of the Bayes strategy is greater than that of the prior optimal act, the experiment should be run, thereby leading to one specific posterior distribution and a specific posterior optimal act (i.e., terminal act); otherwise the prior optimal act should be taken. Of course, both the prior optimal act and the posterior optimal act could be the nonaction "do nothing."

The fundamental character of this process of decision analysis, as de Finetti (1972, p. 204) comments, "is that it is at once optimal in the anticipatory formulation of the decision problem, which asks what is to be done in the light of whatever evidence may arise, and also in the retrospective formulation, which asks later what it is best to do in the light of this evidence that has now arisen."

To represent this normative process more succinctly, we will concentrate on the case of unidimensional preference for a single financial attribute measured by money units. This might be net present value, net income, cash surplus, etc. The scheme of decision analysis procedures that we discussed in Chapter 1 (see Figure 1.2) can now be extended to incorporate procedures to evaluate experimentation. This is done in Figure 5.1 where c is used to denote the cost of the experiment under consideration. As noted in Figure 5.1, the evaluation of the Bayes strategy prior to experimentation is usually termed *preposterior analysis,* a logical name in terms of Bayesian jargon but perhaps unfortunate in its similarity to "preposterous"! Following on from preposterior analysis, the decision may be to purchase the experiment. If this is done, posterior analysis based on the particular forecast received may be carried out, using updated priors and

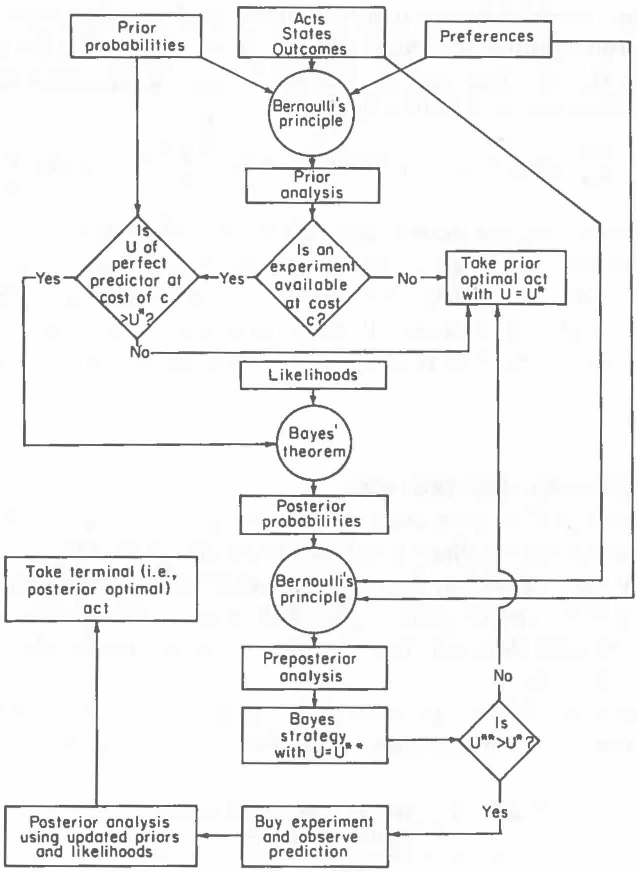


FIG. 5.1. Scheme for evaluating the worth of obtaining more information in decision analysis.

likelihoods if they have changed and revised consequences if they too have changed. Such posterior analysis then leads to choice of the terminal act.

Although shown in Figure 5.1, one important step has so far not been discussed. This is a procedure that should be applied at the very start of any consideration of the worth of an experiment. It is indicated in Figure 5.1 as the question, Is the utility achievable with a perfect predictor that costs *c* greater than the utility of the prior optimal act? Equivalently, this question can be phrased as, Is the money value of the expected value of perfect information (EVPI) greater than the cost of the experiment? In practice this question (which may pose tedious computational problems) can be turned around and asked more simply (in what amounts to prepreposterior analysis) as, If an experiment costing *c* yielded a perfectly reliable forecast of the states, would utility be higher with this perfect information than it is with the prior optimal act? Algebraically, if we denote the money consequences of the ith state and the jth act by x_{ij}, this question can be expressed as determining whether or not

$$\sum_i [\max_j U(x_{ij} - c)]P(\Theta_i) > \max_j \left[\sum_i U(x_{ij})P(\Theta_i)\right]$$

If the answer is yes, the more computationally tedious preposterior evaluation should be undertaken. Otherwise the experiment is immediately shown not to be worthwhile since not even a perfect forecast device at the same price would cover its cost. Perhaps a less expensive experiment might be worth considering; but if none exists, the prior optimal act should be taken.

5.2 A REVISITED ILLUSTRATION

To illustrate the application of decision analysis, we now return to the farmer's marketing problem discussed in Section 1.2. The example is so simple that the impression that decision analysis is computationally trite and economically trivial must be guarded against. The technique really is useful in complex decision making and is being successfully applied in many diverse fields.

Suppose our farmer, perhaps with the help of his farm adviser, has budgeted the dollar consequences of his decision problem as in Table 5.1.

TABLE 5.1. The Decision Problem with Payoffs in Dollars

Θ_i	$P(\Theta_i)$	a_1	a_2
Θ_1	0.8	\$6,800	\$6,150
Θ_2	0.2	\$6,800	\$11,500
Expected money value		\$6,800	\$ 7,220

The expected money values of the acts are readily calculated using the elicited prior probabilities, and at this point a_2 looks to be the better prospect; but this, of course, is not necessarily the whole story since we still have to take account of the farmer's preferences. Earlier, it was suggested that preference be checked by comparing the certainty equivalent of a lottery involving the extreme payoffs with its expected money value. However, it is a good idea to span somewhat more than this range in utility specification. Suppose we determine the farmer's preferences over the range of gains from \$6000 to \$12,000 and find that the following certainty

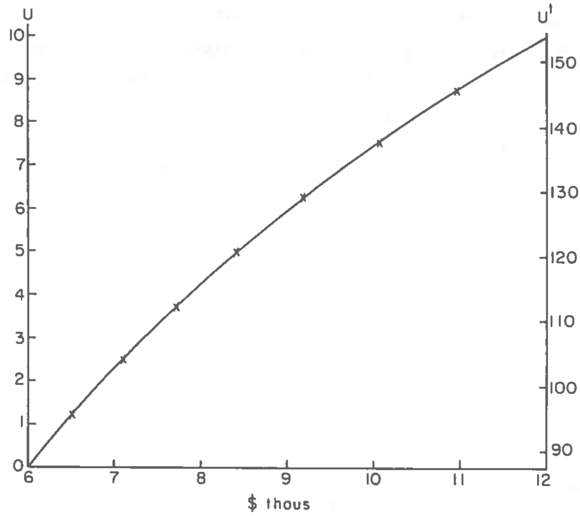


FIG. 5.2. Farmer's utility curve for grain sale income.

equivalents are established for equally likely reference lotteries: ($6000, $12,000) ≡ $8400; ($6000, $8400) ≡ $7100; ($8400, $12,000) ≡ $10,100; etc. The smooth utility curve through these certainty equivalents is sketched in Figure 5.2 by assigning the arbitrary values $U(12{,}000) = 10$ and $U(6000) = 0$. Thus $U(8400) = 5$, etc. Hence utilities corresponding to the payoffs in the decision problem can be read as $U(6150) = 0.4$, $U(6800) = 1.9$, and $U(11{,}500) = 9.4$; and the expected utilities associated with a_1 and a_2 are 1.9 and 2.2 respectively, so that a_2 is the utility-maximizing act.

To emphasize (purely for expository purposes) that this scale is arbitrary, consider the particular positive linear transformation, $U' = aU + b$, where a and b are chosen so that $U'(6150) = 90$ and $U'(6800) = 100$. The a and b values are found by solving simultaneously: $90 = a(0.4) + b$ and $100 = a(1.9) + b$, with the result $a = 6.67$ and $b = 87.33$. Then, for example, on the new scale $U(11{,}500) = (6.67)(9.4) + 87.33 = 150$ which, not coincidentally, gives us the same utility payoffs for this problem as were assumed in Section 1.2. As before, a_2 is indicated as being the prior optimal act. Note that there would normally be no reason to change the utility scale in the manner illustrated. Working with any one scale is adequate, and henceforth in the example we will use the U' scale.

Our primary purpose here is to illustrate the general procedure outlined in Section 5.1, but first we illustrate the equivalence of the normal and extensive forms of analysis for problems involving the opportunity of experimentation. The normal form was illustrated in Section 1.2. It evaluates all feasible strategies by taking account of additional information and prior information sequentially. In contrast, the extensive form of analysis is a one-step procedure based on direct use of the posterior probabilities. The procedure is illustrated in Table 5.2 where, as in Section 1.2, it is assumed that a costless forecast of likely market conditions is available to our farmer. The top part of Table 5.2 lays out the Bayes formula calculation of posterior probabilities for each possible market forecast; the bottom part of the table leads to identification of the Bayes strategy and its characteristics.

TABLE 5.2. Calculation of Bayes Strategy for Example of Section 1.2 Using the Extensive Form of Analysis

θ_i	$P(\theta_i)$	$P(z_k \mid \theta_i)$: z_1	$P(z_k \mid \theta_i)$: z_2	Joint Probabilities: z_1	Joint Probabilities: z_2
θ_1	0.8	0.6	0.4	0.48	0.32
θ_2	0.2	0.3	0.7	0.06	0.14
			$P(z_k)$	0.54	0.46

	Utilities: $U(a_1 \mid \theta_i)$	Utilities: $U(a_2 \mid \theta_i)$		Posterior Probabilities: $P(\theta_i \mid z_1)$	Posterior Probabilities: $P(\theta_i \mid z_2)$
θ_1	100	90		0.889	0.696
θ_2	100	150		0.111	0.304
Prior expected utility	100	102*			
		$E[U(a_j \mid z_k)]$:	a_1	100*	100
			a_2	96.7	108.2*
		Bayes strategy:		a_1	a_2

Expected utility of Bayes strategy = (0.54)(100) + (0.46)(108.2) = 103.8

*Denotes utility of optimal act.

It can be seen that the extensive analysis of Table 5.2 is equivalent to the normal form of analysis of Section 1.2. Both lead to the same strategy. In this simple case both approaches involve about the same amount of arithmetic. For larger problems the extensive form is computationally more efficient because instead of evaluating each feasible strategy individually, it selects the Bayes strategy automatically.

Returning to our analytical illustration, suppose that the experiment (i.e., the market research agency's forecast) is not free but will cost our farmer \$100. Should he acquire it? First, we make a quick check on its feasibility under the extreme assumption that the forecast is always correct. We need to express the consequences in utility of returns that are net of the prospective \$100 cost. The required utility values, read from the U' scale of Figure 5.2, are given in Table 5.3.

TABLE 5.3. Required Utility Values from Figure 5.2

		Net Money Consequences		Net Utility Consequences	
θ_i	$P(\theta_i)$	a_1	a_2	a_1	a_2
θ_1	0.8	\$6,700	\$ 6,050	98.5	88
θ_2	0.2	\$6,700	\$11,400	98.5	149
Expected value		\$6,700	\$ 7,120	98.5	100.2*

*Denotes utility of optimal act.

The utility value of our farmer's problem under conditions of perfect prediction is $(0.8)(98.5) + (0.2)(149) = 108.6$. This exceeds the expected utility of 102 units from the prior optimal act (evaluated before deducting the \$100 cost) and indicates that the experiment is worthy of further consideration in a preposterior analysis.

Since the posterior probabilities are available from our earlier discussion, preposterior analysis taking account of the cost of the forecast can now be done quite simply by using an abridged form of Table 5.2 as shown in Table 5.4. The value of the Bayes strategy (102.2 utility units) is just greater than the 102 for the prior optimal act. Therefore, considering its cost and encoded value, the experiment is worth acquiring. The experimental information should be sought forthwith.

TABLE 5.4. Calculation of Bayes Strategy for Example of Section 1.2 with Allowance for Cost of Information

	$U(x_{ij} - c)$		$P(\theta_i \mid z_k)$	
θ_i	a_1	a_2	z_1	z_2
θ_1	98.5	88	0.889	0.696
θ_2	98.5	149	0.111	0.304
$E[U(x_{ij} - c \mid z_k)]$: a_1			98.5*	98.5
$E[U(x_{ij} - c \mid z_k)]$: a_2			94.8	106.5*

Bayes strategy: a_1 a_2

Expected utility of Bayes strategy $= (0.54)(98.5) + (0.46)(106.5) = 102.2$

*Denotes utility of optimal act.

Suppose our farmer makes his payment and receives forecast z_1. Now he knows something more about the occurrence of the states, his updated beliefs being represented by the posterior probabilities $P(\Theta_i \mid z_1)$, for which Table 5.4 tells us that the posterior expected utilities for a_1 and a_2 are 98.5 and 94.8 respectively. Accordingly, he should now take the terminal decision of following a_1. Assuming he does this, if Θ_2 occurs (a chance of 0.111), he will doubtless have some regrets based on hindsight (and will tend to revise his likelihoods for future predictions). But he certainly cannot be criticized for having chosen a_1. It was the best possible decision he could make in terms of the information available to him when he had to choose.

5.3 LIMITS TO THE VALUE OF INFORMATION

We have seen in the above example how prospective information available at a cost c can be evaluated first through the extreme assumption that it is perfect and second, if it passes the first test, via preposterior analysis. These evaluations were based on comparisons of (expected) utilities indicating the best actions. Recalling that the scale on which utility is measured is arbitrary, we must be cautious about expressing the value of additional information in utility terms. Is a difference (e.g., of 108.6 − 102 = 6.6 utility units) between the value of the problem if the predictor were perfect and the value of the prior optimal act a large or a small difference? We cannot say. Because the utility scale is arbitrary, the difference, denoted by U(EVPI), could be made to look large by an appropriate scale change. The same limitation prevents statements about the size of the difference $U^{**} - U^{*}$ between the utility of the Bayes strategy U^{**} and the utility of the prior optimal act U^{*}. However, we can speak meaningfully about the ratio of these two differences. This ratio gives a measure of the efficiency of the imperfect predictor relative to the perfect predictor. Converting to percentages for conversational convenience, we could refer to percentage efficiency as $(U^{**} - U^{*})\,100/U(\text{EVPI})$ since this ratio is independent of linear transformations of the utility scale. In the above example the relative efficiency of the predictor is $(102.2 - 102)(100)/(108.6 - 102) = 3\%$. It would be negative if the experiment was not worth running. Under no circumstances could the ratio exceed 100%.

While nonlinear preference rather circumscribes statements about information values, it obviously has an influence in answering such questions as, How much could I afford to pay for this particular experiment? or, What would be the maximum I could afford to pay for any conceivable experiment (including a perfect predictor)? The second of these is the simplest to answer so we approach it first. Denote the maximum money value of perfect information by v_p; then v_p is found as the money value that satisfies

$$\sum_i [\max_j U(x_{ij} - v_p)]P(\Theta_i) = \max_j \left[\sum_i U(x_{ij})P(\Theta_i)\right]$$

i.e., it is the cost of forecast that just equates the utility value of a perfect predictor and the utility value of the prior optimal act. The value of v_p depends on the structure of the whole decision problem including outcomes, preferences, and prior probabilities.

It would be possible, for instance, to find the prior probability distribution that gave the maximum possible solution v_p, but such a parametric exercise would not generate anything of interest to the decision maker. His degrees of belief are givens that are already encoded in the prior distribution.

Except for simple utility functions such as the quadratic polynomial, finding the value of v_p is a rather tedious trial-and-error job that is much more suited to numerical analysis on a computer than to hand computation, although intermediate plotting and interpolation speed the process. However, answers that are fairly close to the correct solution can be found by the fast and direct method of approximating v_p as the difference between the certainty equivalent of the expected utility of the decision problem under a costless perfect predictor and the certainty equivalent of the prior optimal act. For the present example the value of v_p is approximately \$575.

The question of the maximum money value of a particular experiment is more difficult to answer. Using the subscript s to denote sample or experiment, the money value of imperfect forecast information, denoted by v_s, may be found as the value that satisfies

$$\sum_k \max_j \left[\sum_i U(x_{ij} - v_s)P(\Theta_i \mid z_k)\right]P(z_k) = \max_j \left[\sum_i U(x_{ij})P(\Theta_i)\right]$$

Essentially, this amounts to cycling through preposterior analysis with varying experimental costs until a solution is found such that the utility of the Bayes strategy equals the utility of the prior optimal act. Use of a computer makes such an iterative calculation procedure feasible, whereas in a problem with many states and experimental outcomes it would not otherwise be contemplated. Analogously to the approximation procedure noted above for finding good trial values of v_p, fairly close answers to v_s can usually be obtained as the difference between the certainty equivalent of the Bayes strategy for a costless predictor and the certainty equivalent of the prior optimal act. For the present example, the maximum our farmer could afford to pay for the forecast information is approximately \$110, conditional on his elicited beliefs, preferences, etc., as specified in Table 5.1.

As noted in Section 4.6, the moment method of utility evaluation can be useful in ascertaining v_p and v_s. Computationally, it will usually be an easier approach than working with the expectation of the utility function and takes advantage of the fact that the only moment of a risky prospect changed by subtracting a cost of information is its mean.

5.4 GENERAL MODEL OF DISCRETE DECISION ANALYSIS

So far we have been rather discursive and have avoided the temptation to simplify our presentation by using too much algebraic representation. But we now bring degrees of belief and degrees of preference together in a general model of decision analysis using algebraic shorthand. We will, however, still stick to discrete states and acts since agricultural problems can usually be most easily handled in discrete form. Discussion of analysis in continuous form is deferred until Section 5.6.

Our basic symbols are:

a_j = the jth act
Θ_i = the ith state
$P(\Theta_i)$ = the prior probability of Θ_i
x_{ij} = the unscaled consequence associated with Θ_i and a_j
$U(\cdot)$ = the utility of $(\cdot)$
z_k = the kth forecast signal
$P(z_k \mid \Theta_i)$ = the likelihood of z_k given Θ_i
$P(\Theta_i \mid z_k)$ = the posterior probability of Θ_i given z_k
c = the cost of the forecast device generating the set $\{z_k\}$ of possible forecasts

A single asterisk will be used to denote optimality with respect to prior probabilities, e.g., a_j^* and a double asterisk with respect to posterior probabilities, e.g., a_j^{**}. A prime will be used to denote a perfect forecast (e.g., z_k') and its associated optimal act (e.g., a_{ji}').

The utility of the jth act without any forecast information is

$$U(a_j) = \sum_i U(x_{ij})P(\Theta_i) \tag{5.1}$$

and the prior optimal act a_j^* will be the act such that

$$U^* = U(a_j^*) = \max_j U(a_j) = \max_j \left[\sum_i U(x_{ij})P(\Theta_i)\right] \tag{5.2}$$

Now suppose a forecast device is available at a cost of c. Given knowledge of the likelihoods $P(z_k \mid \Theta_i)$, Bayes' theorem implies the posterior probability $P(\Theta_i \mid z_k)$, where

$$\begin{aligned} P(\Theta_i \mid z_k) &= P(\Theta_i)P(z_k \mid \Theta_i) \Big/ \sum_i P(\Theta_i)P(z_k \mid \Theta_i) \\ &= P(\Theta_i)P(z_k \mid \Theta_i)/P(z_k) \end{aligned} \tag{5.3}$$

For each possible forecast signal from the particular predictive mechanism at hand, there will be a revised set of probabilities. In consequence, for each signal there will be an optimal act. The array of optimal acts, each conditional on the receipt of a particular forecast possibility, constitutes the optimal strategy.

The utility of a_j given the forecast signal z_k at a cost of c is

$$U(a_j \mid z_k) = \sum_i U(x_{ij} - c)P(\Theta_i \mid z_k) \tag{5.4}$$

Denoting the posterior optimal act given z_k by $a_{j\,k}^{**}$, this will be the act such that

$$U(a_{j\,k}^{**}) = \max_j U(a_j \mid z_k) \tag{5.5}$$

The Bayes strategy is then specified as the set of preposterior optimal acts, one for each of the possible forecast signals. We may denote the Bayes strategy by the vector $\{a_{j\,k}^{**}\}$ where j may take on any value from its range of possible values and k runs 1, 2, 3, . . . up to the number of possible forecasts. Note that the number of forecasts may be less than, equal to, or greater than the number of possible states.

The probability that each posterior optimal act will be taken is the probability that its associated forecast occurs. The utility of the Bayes strategy can therefore be calculated as

$$U^{**} = U(\{a_{j\,k}^{**}\}) = \sum_k U(a_{j\,k}^{**})P(z_k) \tag{5.6}$$

Equations (5.3)–(5.6) constitute the core of preposterior decision analysis. Such analysis and the steps leading up to it are shown in simple flow chart form in Figure 5.3. This diagram may be contrasted with Figure 1.2 (which gives a broader overview) and Figure 5.1 (which emphasizes the question of information purchase).

In preposterior terms, the utility of the forecast device that generates the set of possible predictions $\{z_k\}$ is

$$U(\{z_k\}) = U^{**} - U^{*} = U(\{a_{j\,k}^{**}\}) - U(a_j^{*}) \tag{5.7}$$

If this value is negative, the additional information expected from the forecast device is not worth purchasing. The maximum price that may be economically paid for the forecast is given by the value of c for which $U(\{z_k\}) = 0$. In Section 5.3 this value was denoted v_s.

If there are a number of forecast devices available (e.g., various sized binomial samples from a Bernoulli process), the evaluation procedure of equation (5.7) can be carried out for each device (e.g., sample size) so as to assess relative merits. Obviously, the preferred device will be the one for

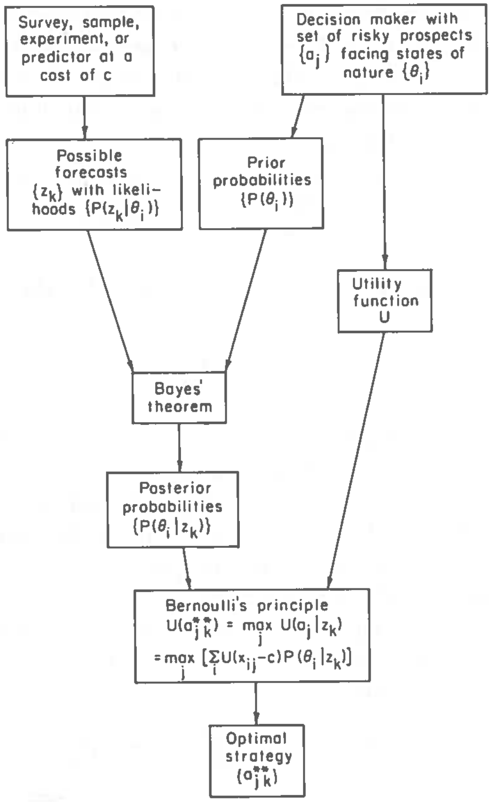


FIG. 5.3. Block diagram summarizing decision analysis procedure involving Bayes' theorem.

which the difference $U^{**} - U^{*}$ is largest, remembering that the cost of each forecast device is already allowed for in this assessment.

Now consider the case of a perfect forecast mechanism. Since a perfect predictor is never wrong, it implies a posterior probability of unity for some state and zero for the rest. Thus, using a prime to denote perfection, there is a one-to-one correspondence between the kth perfect forecast signal z_k' and some state, say Θ_i, so that we can denote the kth perfect forecast by z_i'. Further, by Bayes' theorem $P(z_i') = P(\Theta_i)$.

With a perfect forecast device the optimal act can always be chosen. For the forecast z_i' this will be the act a_{ji}' such that

$$U(a_{ji}') = \max_j U[(x_{ij} - c) \mid z_i'] \tag{5.8}$$

Since the perfect forecast z_i' will occur with probability $P(\Theta_i)$, the utility of a strategy based upon a perfect forecast mechanism is

$$U(\{a'_{ji}\}) = \sum_i U(a'_{ji})P(\theta_i) \tag{5.9}$$

The utility value of the perfect predictor, i.e., U(EVPI), is

$$U(\text{EVPI}) = U(\{z'_i\}) = U(\{a'_{ji}\}) - U(a^*_j) \tag{5.10}$$

The maximum price that should be paid for the perfect predictor is given by the value of c for which $U(\{z'_i\}) = 0$. This value of c equals v_p, the maximum money value of perfect information as discussed in Section 5.3. Finally, the percentage efficiency of the less-than-perfect predictor generating the set of potential forecasts $\{z_k\}$ can be evaluated as $U(\{z_k\})100/U(\{z'_i\})$.

To illustrate the above procedures, consider the decision problem shown in Table 5.5. It relates to a manager who has to choose between contracting to purchase 1000, 1200, or 1600 cattle for fattening on summer pasture. His profit will depend on whether the pasture growing season turns out to be good, fair, or poor. For these events his degrees of belief are 0.4, 0.2, and 0.4 respectively. The budgeted net dollar gain or loss from each alternative under each state is shown in Table 5.5. The three available actions all happen to have the same expected money value of $12,400.

TABLE 5.5. Cattle Purchase Decision Problem

θ_i (type of season)	$P(\theta_i)$	x_{ij} values		
		a_1 (buy 1000)	a_2 (buy 1200)	a_3 (buy 1600)
			($ thous)	
θ_1 (good)	0.4	20	25	34
θ_2 (fair)	0.2	10	12	16
θ_3 (poor)	0.4	6	0	−11

Given that the manager's utility function for gains and losses is adequately represented by

$$U(x) = x - 0.005x^2 \qquad \text{for } x \leqq 80$$

where x denotes thousands of dollars, which alternative should he choose?

Applying equation (5.1) and using the manager's utility function to convert dollar values to utility values, the utility of each alternative can be calculated as follows:

$$\begin{aligned} U(a_1) &= \sum_i U(x_{i1})P(\theta_i) \\ &= 0.4U(\$20{,}000) + 0.2U(\$10{,}000) + 0.4U(\$6000) \\ &= 0.4(18) + 0.2(9.5) + 0.4(5.82) = 11.428 \end{aligned}$$

$$U(a_2) = \sum_i U(x_{i2})P(\Theta_i)$$

$$= 0.4U(\$25{,}000) + 0.2U(\$12{,}000) + 0.4U(\$0)$$

$$= 0.4(21.875) + 0.2(11.28) + 0.4(0) = 11.006$$

$$U(a_3) = 0.4U(\$34{,}000) + 0.2U(\$16{,}000) + 0.4U(-\$11{,}000)$$

$$= 0.4(28.22) + 0.2(14.72) + 0.4(-11.605) = 9.590$$

Comparing these utility values, the prior optimal act is obviously a_1 (i.e., buy 1000), and as per equation (5.2) we have $U^* = U(a_j^*) = U(a_1) = 11.428$.

Now suppose a forecast of the coming season is available at a cost of $c = \$200$. The forecast consists of a signal z_1 or z_2 or z_3. The likelihoods $P(z_k \mid \Theta_i)$ of these various signals relative to the possible states of nature have been estimated by our manager on the basis of past experience and subjective judgment. They are shown beside his prior probabilities in the top left side of Table 5.6. Below them are listed the utility payoffs $U(x_{ij} - c)$. For example, with $i = 2$ (i.e., a fair season) and $j = 3$ (i.e., buy 1600), $x_{ij} = \$16{,}000$ so that $U(x_{ij} - c) = U(\$15{,}800) = 14.552$.

Calculation of the prediction probabilities $P(z_k)$ and of the posterior probabilities $P(\Theta_i \mid z_k)$ as per equation (5.3) is shown on the right side of Table 5.6 above the derivation of the optimal strategy $\{a_{jk}^{**}\} = (a_{31}, a_{12}, a_{13})$ and the expected utility of this strategy. The optimal strategy is to purchase 1600 cattle if the forecast signal is z_1; but if the signal is either z_2 or z_3, then 1000 cattle should be bought. The calculations leading to these results in the bottom right side of Table 5.6 are based on (5.4), (5.5), and (5.6). For example, using (5.4), the utility of a_2 given a forecast signal z_3 at a cost of \$200 is

$$U(a_2 \mid z_3) = \sum_i U(x_{i2} - 200)P(\Theta_i \mid z_3)$$

$$= 0.17U(\$24{,}800) + 0.17U(\$11{,}800) + 0.66U(-\$200)$$

$$= 0.17(21.725) + 0.17(11.104) + 0.66(-0.200) = 5.448$$

As per (5.7), we can calculate the utility value of the forecast device as

$$U(\{z_k\}) = U(\{a_{jk}^{**}\}) - U(a_j^*) = 13.843 - 11.428 = 2.415$$

Since this utility value is positive, the forecast is worth purchasing by the manager at its price of \$200.

The expected utility of the optimal strategy $\{a_{ji}'\}$ based upon a perfect predictor costing \$200 is given by equation (5.9) as

$$U(\{a_{ji}'\}) = \sum_i U(a_{ji}')P(\Theta_i) = \sum_i \{\max_j U[(x_{ij} - c) \mid z_i']\}P(\Theta_i)$$

$$= (28.088)(0.4) + (14.552)(0.2) + (5.632)(0.4) = 16.398$$

TABLE 5.6. Calculation of Optimal Strategy for Cattle Purchase Decision Problem

Data

State of nature θ_i	Prior probability $P(\theta_i)$	Likelihood $P(z_k \mid \theta_i)$ z_1	z_2	z_3
Good ($i = 1$)	0.4	0.7	0.2	0.1
Fair ($i = 2$)	0.2	0.4	0.4	0.2
Poor ($i = 3$)	0.4	0.1	0.5	0.4

Utility payoffs $U(x_{ij} - c)$	a_1	a_2	a_3
Good	17.840	21.725	28.088
Fair	9.320	11.104	14.552
Poor	5.632	−0.200	−11.827

Calculation of Posterior Probabilities and Strategy Values

Forecast $z_k \rightarrow$		Joint probabilities z_1 and θ_i	z_2 and θ_i	z_3 and θ_i
(Likelihood × prior) =		0.28	0.08	0.04
		0.08	0.08	0.04
		0.04	0.20	0.16
$P(z_k)$ = sum	=	0.40	0.36	0.24
Posterior probabilities		$P(\theta_i \mid z_1)$	$P(\theta_i \mid z_2)$	$P(\theta_i \mid z_3)$
Good ($i = 1$)		28/40 = 0.70	8/36 = 0.22	4/24 = 0.17
Fair ($i = 2$)		8/40 = 0.20	8/36 = 0.22	4/24 = 0.17
Poor ($i = 3$)		4/40 = 0.10	20/36 = 0.56	16/24 = 0.66
$U(a_j \mid z_k) = \sum_i U(x_{ij} - c) P(\theta_i \mid z_k) \rightarrow$ $U(a_1 \mid z_k)$	=	14.915	9.129	8.334
$U(a_2 \mid z_k)$	=	17.408	7.110	5.448
$U(a_3 \mid z_k)$	=	21.390	2.756	−0.559
$U(a_{jk}^{**}) = \max_j U(a_j \mid z_k) \rightarrow$ $U(a_{jk}^{**})$	=	21.390	9.129	8.334
Optimal strategy = $\{a_{jk}^{**}\}$	=	a_{31}	a_{12}	a_{13}
Expected utility of optimal strategy = $\sum_k U(a_{jk}^{**}) P(z_k)$	=	(21.39)(0.4) +	(9.13)(0.36) +	(8.33)(0.24)
		= 13.843		

the optimal strategy with a perfect predictor being $(a'_{31}, a'_{32}, a'_{13})$. Using (5.10), the utility value of the perfect predictor is

$$U(\{z'_i\}) = U(\{a'_{ji}\}) - U(a^*_j) = 16.398 - 11.428 = 4.970$$

The efficiency of the predictor generating $\{z_k\}$ relative to a perfect predictor, both being assumed to cost \$200, is thus (2.415/4.970)(100) = 48.6%. Finally, solving (5.10) for the value of c that makes $U(\{z'_i\})$ equal to zero indicates that the maximum price our manager could economically pay for a perfect predictor is \$6070.

5.5 DECISION TREE REPRESENTATION

It is often very useful to represent risky decision problems in the schematic form of a *decision tree* or *decision flow diagram*. This is especially so for the analysis of problems involving a sequence of decisions (including information-gathering possibilities) which follow a natural or chronological order. All the previously noted components of general decision problems can be represented in decision flow diagrams. As might be anticipated, the jargon consists of a blend of botanical and decision theory terms. Acts branch from *decision nodes,* conventionally denoted by squares; and states or events branch from *chance* or *event nodes,* denoted by circles. Experiment-buying acts are treated as other acts.

Decision tree analysis naturally pivots around application of Bayes' theorem and Bernoulli's principle. The analysis may be carried out in two ways. We will call the first of these the *certainty equivalent approach* because it uses certainty equivalents and does not require specification of the utility function. The second is known as the *utility function approach* because it assumes knowledge of the utility function. Both methods are equivalent and should lead to identical decisions. However, for the analysis of problems on an ad hoc basis, the certainty equivalent approach is the simpler of the two methods.

Certainty Equivalent Approach

Procedures with the certainty equivalent approach to decision tree analysis do not require a knowledge of the decision maker's utility function. Rather than separate assessment of preferences and beliefs as implied by calculating expected utilities, the certainty equivalent approach relies on the decision maker simultaneously taking intuitive account of his degrees of belief and degrees of preference. The approach involves the following five steps:

1. Draw the decision tree in chronological sequence from left to right with acts branching from decision nodes denoted by squares and events branching from event or chance nodes denoted by circles.

2. Assign the relevant subjective probabilities to event branches (including possible forecast signals), checking that the probabilities are coherent and using Bayes' theorem where revisions are required.
3. Attach net dollar payoffs to the terminal branches, making sure that account has been taken of the outcomes and costs of all preceding branches.
4. Working back leftward from the terminal branches, replace the chance events at each event node by their certainty equivalents; then choose between antecedent acts on the basis of these certainty equivalents, the act with the highest certainty equivalent being the preferred alternative at each decision node. This process of *backward induction* is sometimes called "averaging out and folding back." As backward induction is carried out, write the certainty equivalent at each node to make the whole process clearly explicit.
5. Mark off or delete inferior acts as they are located so that when the base of the tree is reached, the optimal path through the tree is clearly evident.

We will illustrate the certainty equivalent approach via the cattle purchase decision problem of Section 5.4. For simplicity look first at the problem, assuming no forecast information is available. Figure 5.4 shows the data of Table 5.5 in the form of a decision tree. Any payoff matrix or decision problem can be depicted in such extensive fashion with a sequence of act-event nodes and branches. Node *A* is the current decision fork; nodes *B*, *C*, and *D* are event forks. In a more complicated problem we would have a series of act-event sequences rather than a single act-event sequence as in Figure 5.4. Note that "actions" are at the discretion of the decision maker while "events" are not.

Proceeding by backward induction, the decision maker replaces the set of possible events at each event node by his certainty equivalent, i.e., by the sure amount that he judges is equivalent to the risky prospect specified by the event fork. For example, our decision maker's certainty equivalent for the risky prospect of a 0.4 chance of \$20,000, a 0.2 chance of \$10,000, and a 0.4 chance of \$6000 at event fork *B* of Figure 5.4 might be \$12,200. He would be indifferent between an assured payment of this amount and the opportunity of participating in the risky prospect specified by event node *B*. Nominated by the decision maker on the basis of introspective judgment, such certainty equivalents encompass both his degrees of preference and his degrees of belief about the risky prospects being assessed. They could, of course, be computed more precisely from the probabilities and the utility function if this is known.

Replacement of the risky prospect at each event node by its certainty equivalent gives a simpler but equivalent problem of choice between certainty equivalents at the antecedent decision node. Obviously, the alterna-

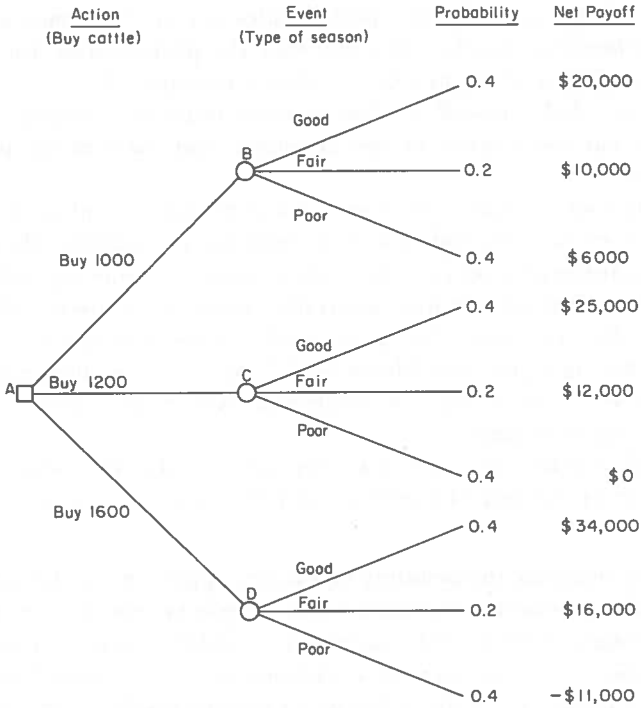


FIG. 5.4. Decision tree for certainty equivalent analysis of the cattle purchase decision problem of Section 5.4.

tive with the largest certainty equivalent would be the preferred choice. For example, if the certainty equivalents of event nodes *B*, *C*, and *D* in Figure 5.4 were respectively $12,200, $11,700, and $10,100, the preferred choice at *A* would be to purchase 1000 cattle.

The decision of whether to purchase forecast information can also be assessed via the certainty equivalent approach. As before, suppose the set of possible forecasts for our cattle purchase problem consists of z_1, z_2, and z_3 as specified in Table 5.6. To encompass the purchase of information, the decision tree of Figure 5.4 has to be extended as shown in Figure 5.5. The forecast is again assumed to cost $200.

In Figure 5.5, event forks such as *J*, *K*, and *M* must first be replaced by their certainty equivalents. Suppose for *J*, *K*, and *M* these are respectively $16,000, $19,000, and $24,000. This implies that if a forecast is purchased and it turns out to be z_1, then "buy 1600" with a certainty equivalent of $24,000 is the preferred action; i.e., it is the z_1 component of the optimal strategy. The decision tree to the right of event fork *F* may thus be

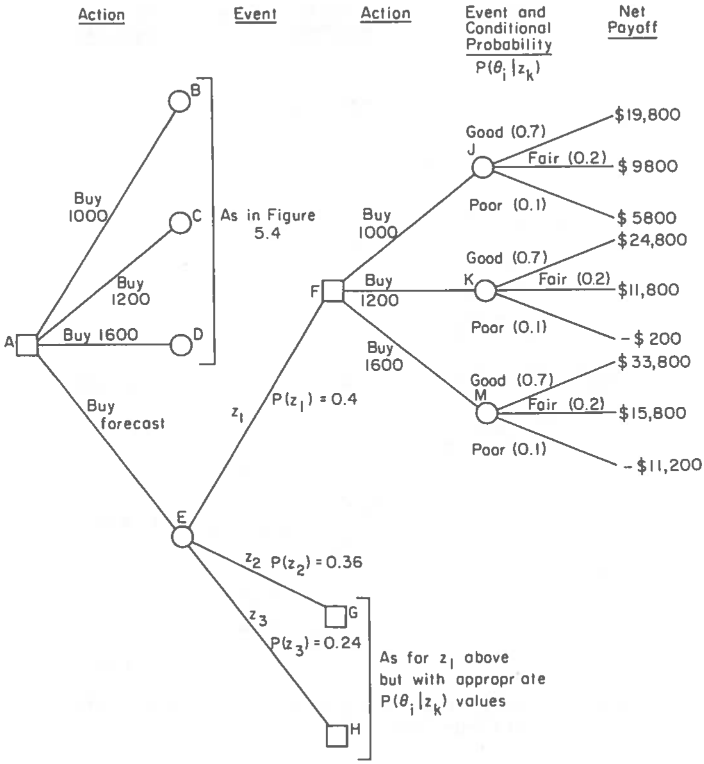


FIG. 5.5. Extension of Figure 5.4 to encompass possible purchase of additional information.

replaced by its certainty equivalent value of \$24,000. Likewise, suppose the analogous certainty equivalents to be placed at *G* and *H* are respectively \$9500 (for "buy 1000" if z_2 is the forecast received) and \$8500 (for "buy 1000" if the forecast is z_3). Given the analysis already made for Figure 5.4, Figure 5.5 may now be replaced by the equivalent but simpler decision problem of Figure 5.6. All that remains is to replace event fork *E* by its certainty equivalent. If this is greater than \$12,200 (the certainty equivalent of the optimal act without extra information), the forecast should be purchased and the optimal strategy would be to "buy 1600" if z_1 occurs and to "buy 1000" if z_2 or z_3 is the signal received.

Though simpler because of its simultaneous consideration of degrees of belief and preference, the decision tree approach is formally equivalent to that of the general model of decision analysis outlined in Section 5.4. A disadvantage of the approach is that it is not conveniently applicable if it is desired to delegate choice. Because of their introspective basis, the deci-

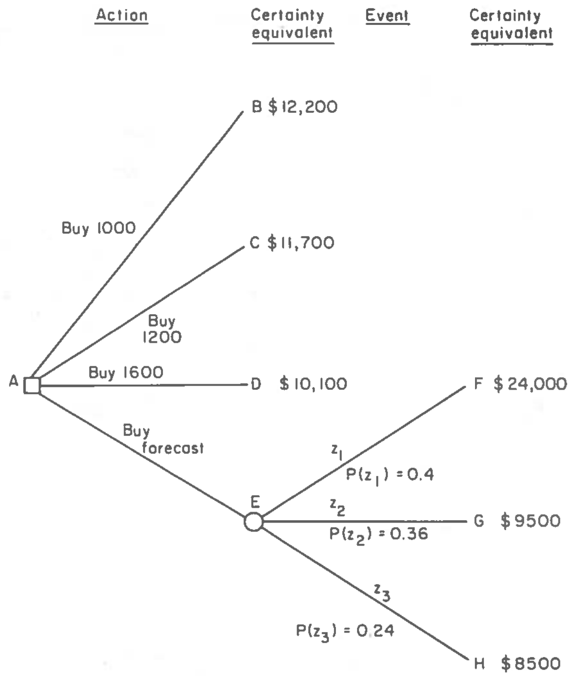


FIG. 5.6. Simpler decision problem equivalent to that of Figure 5.5.

sion maker has to specify the required certainty equivalents himself. With the explicit utility approach of the general model, decision making may be more easily delegated to a subordinate (perhaps a computer) who, making use of his superior's utility function and probabilities, would make exactly those decisions his superior would have made. As shown below, the general expected utility model can be applied in decision tree format.

Utility Function Approach

Given knowledge of the decision maker's utility function, the utility function approach involves the following five steps:

1. Draw the decision tree in chronological sequence from left to right.
2. Assign the relevant subjective probabilities to event branches.
3. Making use of the known utility function, convert the net dollar payoff of each terminal branch to its utility value.
4. Work back leftward from the terminal branches, taking expectations at chance nodes and utility maximizing decisions at decision nodes.
5. Continue the process of backward induction as per (4) until the base of

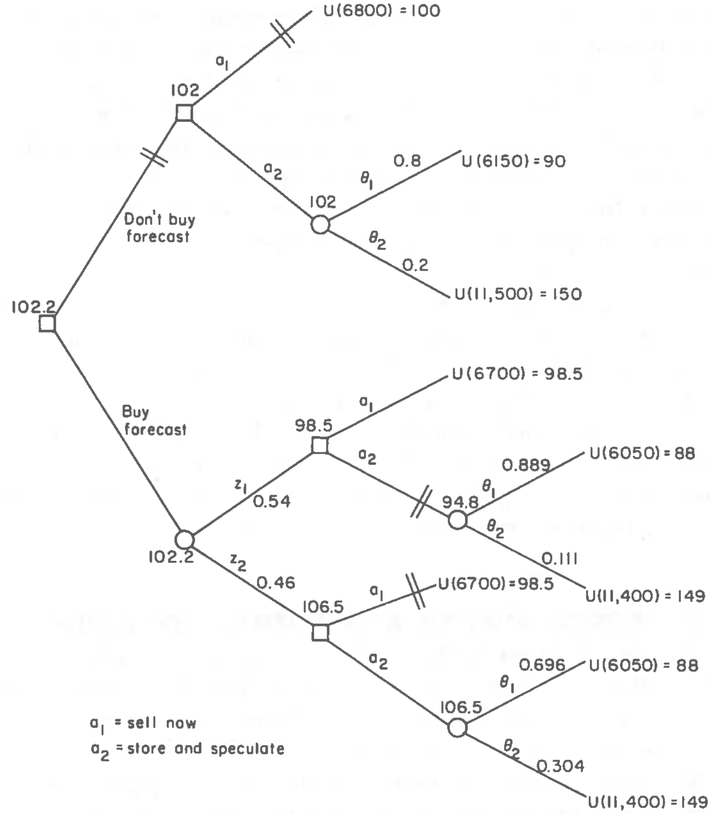


FIG. 5.7. Illustration of the utility function approach to decision tree analysis of risky choice using the example of Sections 1.2 and 4.2.

the tree is reached, the optimal path through the tree being specified as the sequence of preferred acts from left to right.

Figure 5.7 illustrates the application of this utility function decision tree procedure in terms of our simple "sell now" or "store and speculate" example of Sections 1.2 and 5.2. The result is identical to that obtained earlier, namely, that the farmer should purchase the forecast and follow the strategy (a_{11}, a_{22}) of taking a_1 if z_1 occurs and taking a_2 if z_2 occurs.

Given knowledge of the decision maker's utility function, there is no computational advantage in using a decision tree layout. Utility values and posterior probabilities still have to be calculated. At the same time, there are advantages per se in drawing a decision tree; the disciplined thinking such an exercise demands will generally lead the decision maker or his consultant analyst to a better appreciation of the decision problem. An obvious difficulty, however, is to prevent the decision tree from becoming a "bushy mess," as Raiffa [1968, p. 240] nicely puts the problem. As acts and events multiply, the tree explodes rapidly—particularly so if multiperiod possibilities are included. Art as well as science seems necessary to keep the picture comprehensible. A workable procedure appears to be to begin with a rather coarse tree specifying only the major branches, checking which of these might be lopped, developing the unlopped branches in further detail, and repeating the cycle.

As the problem of bushiness implies, drawing an adequate decision tree is not always easy. Indeed, it is usually difficult, and first attempts at specifying real-world decision problems in the form of a tree (or a payoff matrix) often lead to despair. Persistence and practice are required. Difficulties are compounded if the decision problem is not well specified. In consequence, except for the simplest of problems, different people may well draw somewhat different trees for the same problem, just as different artists will depict the same scene in different ways.

5.6 DECISION ANALYSIS WITH NORMAL PROBABILITIES

In Section 3.4 we noted that if the state variable is judged to follow a normal distribution, analysis in continuous rather than discrete form may often be convenient. This is especially so if there are only two actions and the posterior distribution is also normal, as will be the case if the distribution of the sample mean is normal. To illustrate the procedures, we will again use the new food product example of Section 3.4. Recall that we were concerned with the expected level of sales per store of a new food product that could be sold in 10,000 stores. Our prior distribution was $\mathcal{N}(\mu_0, \sigma_0) = \mathcal{N}(200, 23.75)$. Using the information from a sample of $n = 100$ stores, which gave a sample mean of $m = 210$ and a sample standard deviation of $s = 110$, our posterior distribution was $\mathcal{N}(\mu_1, \sigma_1) = \mathcal{N}(208.2, 10)$.

Suppose the problem confronting us is whether to market the new product, given that the setup cost for production, advertising, and distribution is \$410,000 and that our net return per unit (kg) sold is 20 cents. For simplicity we will assume that our utility function is linear over the range relevant to the problem.

Expected profit based on our prior distribution is

$$E(\pi_0) = E[-410{,}000 + (0.20)(10{,}000)S]$$

where S denotes sales per store. Taking the expectation and using the fact that $E(S) = \mu_0 = 200$, we have

$$E(\pi_0) = -410{,}000 + (0.20)(10{,}000)(200) = -\$10{,}000$$

Obviously, based on our prior information we should not go ahead with the new product. This will always be our best decision as long as our prior normal mean μ_0 is less than the break-even value u_b. In the present case u_b is 205 kg/store since to break even we must have 410,000 = (0.20) (10,000) u_b.

Given our prior distribution, how valuable would costless perfect information be? With a normal distribution this is easily calculated via the formula (La Valle, 1970; Raiffa and Schlaifer, 1961)

$$\text{EVPI}_0 = C\sigma_0\mathcal{N}(D_0) \tag{5.11}$$

In this special formula for the normal distribution case, C is the slope of the expected profit function. Thus for our present example, C is the change in profit per unit change in μ, the mean sales per store, and is equal to (0.20)(10,000) = \$2000; $\mathcal{N}(D_0)$ is the value of the "loss function" for the prior standard normal curve, evaluated at the quantity $D_0 = |\mu_0 - u_b|/\sigma_0$. Values of $\mathcal{N}(D)$ are given in Appendix Table A.1.

For our example, using equation (5.11), we have

$$\begin{aligned}\text{EVPI}_0 &= (2000)(23.75)[\mathcal{N}(|200\text{–}205|/23.75)]\\ &= (47{,}500)\mathcal{N}(0.2105)\\ &= (47{,}500)(0.3025) = \$14{,}369\end{aligned}$$

Thus relative to our prior optimal act, we would expect to be \$14,369 better off if we had a perfect predictor available. Alternatively, we could afford to pay up to \$14,369 for a perfect predictor without being worse off.

On the basis of our posterior distribution, the decision should be to market the new product since $\mu_1 = 208.2$ is greater than $u_b = 205$. Our expected profit is

$$E(\pi_1) = -410{,}000 + (0.20)(10{,}000)(208.2) = \$6400$$

The EVPI relative to our posterior distribution is

$$\begin{aligned}\text{EVPI}_1 &= C\sigma_1\mathcal{N}(D_1) = (2000)(10)\mathcal{N}(|208.2 - 205|/10)\\ &= (20{,}000)(0.2592) = \$5184\end{aligned}$$

For preposterior analysis with normal distributions the expected gross value of sample information (EVSI) is calculated in a manner similar to EVPI but with σ^* as the standard deviation where

$$\sigma^* = \sigma_0^2/(\sigma_0^2 + \sigma_m^2)^{0.5} \tag{5.12}$$

Thus

$$\text{EVSI} = C\sigma^*\mathcal{N}(D^*) \tag{5.13}$$

where $D^* = |\mu_0 - u_b|/\sigma^*$. To illustrate, suppose we wish to evaluate the worth of a sample of size $n = 64$, and from past experience we estimate the standard deviation of the population σ to be equal to 100. With this information we are able to estimate the standard deviation of the sample mean σ_m as $\sigma_m = \sigma/n^{0.5} = 100/8 = 12.5$. Substituting for σ_0 and σ_m in equation (5.12), the value of σ^* is $\sigma^* = (564.06)/(720.31)^{0.5} = 21.017$. Hence, using (5.13),

$$\begin{aligned} \text{EVSI} &= (2000)(21.017)\mathcal{N}(|200-205|/21.017) \\ &= (42{,}034)\mathcal{N}(0.2379) = (42{,}034)(0.2912) = \$12{,}240 \end{aligned}$$

There would thus be no value in a sample of size 64 if the cost of sampling 64 stores was greater than \$12,240. At a cost less than this such a sample would be worthwhile.

In similar fashion, preposterior analysis could be carried out for other sample sizes. Given values of EVSI for various values of n, a preposterior estimate of the optimal sample size n^* can be obtained. This will be the value of n that gives the biggest net return when we deduct the cost of the sample from its EVSI.

As the above example illustrates, decision analysis with normal probability distributions can be relatively simple. At least this is so when there are only two alternative acts (such as "market" or "don't market") and when C, the slope of the expected profit function, is a constant. If either of these two provisos is not met, the situation becomes rather more complicated and analysis in discrete form may be much simpler. The analysis is also somewhat less simple if utility is nonlinear, but this is not too much of a difficulty since the mean and variance are the only parameters of the normal distribution. Making use of this fact, utility evaluation can be carried out fairly readily by using the moment method as outlined in Section 4.6.

5.7 A WORD ON SENSITIVITY ANALYSIS

A worthwhile illustration of sensitivity analysis would require a much more complex example than we can develop here. Procedures are anything but standardized; they consist of commonsense but systematic appraisals of the sensitivity of optimal choice to various components of decision problems with a view to simplification and size reduction. This will be most useful when outcomes are determined by complex interactions among many variables, some of which may be quite unimportant in their influence and can accordingly be held fixed at an appropriate setting.

To carry out sensitivity analysis, the feasible ranges of possibly relevant variables can be swept through a multifactorial design type of appraisal to rank sensitivities. It may seem natural to judge sensitivity in terms of utility, but again we must guard against improper interpretations of an arbitrary scale. Quantitative interpretations can only be soundly based by extracting certainty equivalents of a key attribute corresponding to the expected utility evaluations. Then, for instance, conditional certainty equivalents might be plotted against the parameter space of a component variable to appraise (subjectively) its sensitivity. For all this, our guess is that sensitivity analysis will be only rarely undertaken in agricultural applications of decision analysis.

5.8 MULTIPERSON DECISIONS

Throughout this discussion we have assumed that only one person is directly involved in the decision process, or that at most there is one decision maker aided by a decision analyst. The emphasis on individual choice has been deliberate, for there are formidable difficulties in multiperson decision analysis. In this section we make a brief sortie into this area to expose some of the deficiencies of both theory and technique for normative appraisal of multiperson decisions. We offer few remedies, but try to indicate why the deficiencies persist.

A multiperson or collective decision situation arises when the activities of two or more individuals are interrelated through the decision. MacCrimmon (1973), from whom the above definition is derived, attempts to provide a framework for modeling multiperson decisions based chiefly on various structural elements of the decision situation. These elements of his model framework, with some minor changes in terminology, are indicated in Figure 5.8. They comprise the individuals participating in the decision and functioning, either singly or in combinations, as information units, choice units, or action units. Information units process information from the environment and pass on relevant data to the choice unit, which in turn hands on instructions to the action unit to implement the decision reached. The information, choice, and action units collectively form the decision entity. In one-person decisions all these functions are performed by the same individual.

Within this general framework a large number of types of interrelationships among individuals can be recognized. In many organizational or social structures there may be many multiperson decision entities, some of which have a more or less permanent existence while others are more ephemeral. Within any one decision entity more than one individual may be involved in information processing, in the choice process itself, or in implementing the decision or decisions. For example, in the context of investment planning in a large business organization, the accounts department might be classified as the information unit, the board of directors as

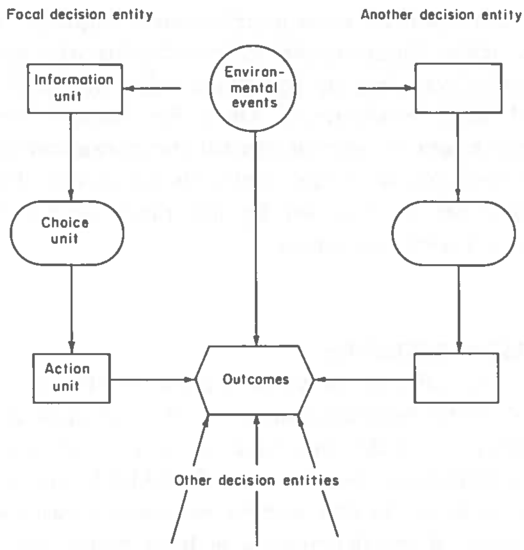


FIG. 5.8. Framework of a structural model of multi-person decisions.

the choice unit, and the professional managerial staff as the action unit. The decisions reached could be influenced by the interrelationships between individuals within each of these task units. Similarly, interrelationships between individuals working in different task units could also have a bearing on the eventual outcome. Some of the interpersonal relationships may contribute to "good" decision making, but others may be counterproductive. Most of us can recognize both types from our own experience.

In addition to the above-mentioned relationships, discrete decision entities may be interrelated in various ways. For example, a number of decision entities may experience the same environmental events. Thus all the farmers in the face of a drought affecting their region would be stimulated to make interrelated decisions. Similarly, as Figure 5.8 illustrates, the outcomes experienced by any particular decision entity may be influenced not only by the actions of that decision entity itself and by the occurrence of uncontrolled environmental events but also by the actions of other decision entities. The profits of a firm producing for a given market will depend on the production decisions of the firm's competitors and on the purchasing decisions of the consumers of its product. A decision entity may take more or less account of the impact that the outcomes of its decision or decisions have on others and may behave in a more or less selfish or more or less altruistic way. Recognition of interrelationships through outcomes may lead to bargaining or cooperation between decision entities.

Other types of interrelationships are possible, but perhaps enough has been said to make it evident that there is a very large number of possible multiperson decision situations. Even with the aid of the framework provided in Figure 5.8, a complete categorization of all such situations would be extremely difficult. Such an attempt would also be at variance with the main focus of this book, which is on individual decision making. In this brief excursion into multiperson decision making we will therefore confine ourselves to only a few cases, which should nevertheless be sufficient to indicate the kinds of analytical problem encountered. We will also try to point to those avenues that seem most likely to lead toward eventual resolution of these analytical difficulties.

Forming a Consensus of Beliefs

Suppose you are faced with an important decision and by now being convinced of the advantages of decision analysis, you begin by decomposing your problem into separate assessments of beliefs and of preferences. In forming the probability judgments that measure your beliefs about the relevant uncertain quantities or events, you will naturally draw on any relevant data; but what if such data are few and far between? Perhaps there are people you can consult who know more about the issues under consideration than you do and whose opinions you respect. You might consult one such person and adopt his probabilities as your own. But if one expert is worth consulting, it seems obvious that it would be still more useful to consult several. The trouble is, they may not all agree. How should you account for their differing views (i.e., differing probability distributions) in forming your probability assessment for your problem? And if you can get your group of experts together as a panel, what is the best way of encouraging them to resolve their differences?

Unfortunately, as is the case in most multiperson decision situations, there is no obviously "correct" way to proceed. All we can do is to examine some suggestions that may be judged appropriate in given circumstances. We will consider the case where each expert is prevailed upon to express his beliefs in quantified form as a proper subjective probability distribution. The problem of how to combine two or more of these into a single distribution we will call the *consensus problem* (Winkler, 1968; Winkler and Cummings, 1972). Note that, in the context of the model in Figure 5.8, we are dealing with a decision entity with a multiperson information unit.

Perhaps the most obvious and straightforward solution to the consensus problem is for the decision maker to combine the experts' distributions in a subjective fashion, but this may not always be plausible. For example, if the decision maker is relatively ignorant of the issues under consideration, he may not be well equipped to evaluate the opinions expressed by several experts. Alternatively, in a hierarchical decision entity the choice unit may demand a consolidated statement of probabilities from the information unit. To deal with such situations, a number of more formal procedures have been developed and some testing of these has been reported. Four such procedures are the weighted average method, the conjugate method, group reassessment, and feedback and reassessment methods.

In the *weighted average method*, as its name implies, one simply takes a weighted sum of the individual estimates $f(x) = \Sigma_k w_k f_k(x)$, where w_k is the weight attached to the kth individual's assessed distribution $f_k(x)$, such that $w_k \geqq 0$ and $\sum_k w_k = 1$. This reduces the consensus problem to one of selecting appropriate weights, and a decision maker might be willing to adopt one of the following rules for this purpose.

1. Weights proportional to the decision maker's subjective assessment of the competence of each expert, assuming that he feels able to make such an assessment.
2. Weights, proportional to a rating made by each expert, of his own competence in relation to the particular issue.
3. Equal weights, implying that the decision maker attaches equal credence to the views of each expert or has no basis for distinguishing between them.
4. Weights based on the degree of success each expert has achieved in the past in assessing distributions for which actual outcomes subsequently became known. (See Section 2.5 for a brief review of scoring rules that might be useful in this context.)

Other possibilities could obviously be added to this list, such as weights based on the experts' ratings of themselves and each other, but these more complex procedures seem unlikely to have general appeal. Procedures (1) and (2) above seem most representative of the Bayesian approach; the first (subjective weights) seems so because it is consistent with the emphasis on the sovereignty of the decision maker, while the latter procedure (self rating) is consistent with the idea of using additional relevant information whenever it is available.

The *conjugate method* of forming a consensus makes use of the concept of conjugate distributions introduced in Section 3.3 as a means of simplifying the application of Bayes' theorem to continuous distributions. In this method it is in effect assumed that the prior information is equivalent to sample information from the data-generating process of interest. In the present context, if the statistical nature of the data-generating process can be identified, it may be plausible to regard each expert's judgment as equivalent to a sample from that process (i.e., equivalent to a prior conjugate of the process). It is then possible to combine the probability distributions from the experts in a similar manner to the way we would combine the information from several successive samples by repeated application of Bayes formula.

Although the conjugate method has some appeal from the point of view of its statistical sophistication, we are still left with a problem of choosing weights to apply to each expert's judgment. These are needed to reflect the amount of information embodied in each "equivalent sample"; but this time there is the added complication (compared with the weighted average method) that the sum of the weights must also be determined. Using the conjugate method, the larger the sum of weights, the more information is contained in the consensus distributions. If the equivalent samples can be regarded as completely independent (i.e., based on independent experiences), the weights can be set equal to one, and the sum of weights for a panel of K experts will be K. This is the greatest possible amount of information from the panel. At the other extreme, if all the experts share the same experience of the process of interest, their assessed distributions should in theory be identical and the sum of weights should be one. At this lower limit the first expert consulted provides all the available information and the remainder add nothing. In reality there will always be some degree of independence of experience, but not complete independence, so weights summing to between one and K must somehow be selected. This complication, together with the fact that the calculations involved in applying the conjugate method are often far from simple, appears to rule out the procedure for all but specialist use.

If the experts disagree, it might seem that the most logical action would be to get them together to sort out their differences, i.e., *group reassessment*. In this way any information available to some but not all of the experts could be shared and any flaws in logic in interpreting the information could be exposed. Despite these important advantages, such group methods have the serious drawback that the consensus finally reached may be more a reflection of the relative strengths of the personalities of the panel members than of their abilities as analysts. Too often, one forceful individual, who may not necessarily be the best informed, can succeed in imposing his point of view on the group.

To try to avoid the risk of this dysfunction in such group procedures, various impersonal *feedback and reassessment methods* have been developed, including the so-called Delphi technique (Dalkey, 1967; Pill, 1971). In the Delphi procedure, subjects receive feedback information about the judgment of the other panel members, but only in an anonymous form. In variants of the procedure the other experts may be identified during the feedback process, but social contact between the panel members is kept to a minimum to reduce the influence of any socially dominant individual.

We emphasize that there is no single correct way of amalgamating the beliefs of a number of people. The behavioral methods such as group reassessment or the Delphi procedure have considerable appeal in that they allow each expert to revise his views in the light of the opinions of others. When such procedures cannot be used, the more mechanical weighted average or conjugate methods seem appropriate, and they may also be needed to resolve any differences in judgments still outstanding after reassessment.

Group Choice Situations

We now turn to the class of decision situations in which two or more individuals must function as a single choice unit. Such a multiperson choice unit might be a husband and wife owning a family farm, a board of directors of a company, or a cabinet in government. Ideally, for decision analysis purposes, we would prefer to derive a group utility function reflecting the preferences of the group in the same way as for an individual. Moreover, it would obviously be desirable for any such group utility function to satisfy certain conditions implying rationality and consistency as well as conditions relating to some concept of "fairness" to all members of the group. Unfortunately, here again we run into difficulties. Arrow (1963), in his proof of what is known as the impossibility theorem of social choice, postulated a number of apparently innocent conditions that such a social (group) welfare function should satisfy and showed that, in general, no such function can exist to meet all the conditions. Although Arrow's logic has never been faulted, his assumptions have been questioned (see, e.g., Little, 1952; Fishburn, 1970). In particular, Arrow's theorem is based on rankings, not cardinal utilities. As Keeney (1975) and Keeney and Kirkwood (1975) show, with cardinal utility a group preference function may sometimes be formed.

The well-known Pareto principle states that a group decision is Pareto optimal if no other decision exists that is preferred by at least one group member and is no less acceptable to all the others. The validity of this principle will be discussed further in relation to social choice; but now we observe that the principle is thrown into doubt, to put it mildly, as soon as we consider the possibility that group members may differ in both their preferences and their probability assessments when facing a group decision under uncertainty. In such a case it would seem most consistent with the principles of decision analysis for the group members to find some way of combining their probability distributions as well as their utility functions before proceeding to an analysis of the decision using group beliefs and preferences. Regrettably, no matter what procedure is used for combining beliefs and preferences (as long as these are kept separate and no one is allowed to dictate to the others), it is always possible to construct an example in which members (analyzing the problem as individuals) would all agree about what to do but in which analysis by using group beliefs and preferences would lead to a different conclusion (Raiffa, 1968, p. 230).

From this rather insecure foothold, we return to tackle the issue of how to form a group utility function. We know that, because any individual utility function is derived using an arbitrary origin and scale, most kinds of comparisons of utility between individuals are invalid. Thus if some outcome is assigned a utility value of 10 using *A*'s function and a utility value of 20 using *B*'s function, we cannot say that the utility to *A* and *B* together is $10 + 20 = 30$. Some brave efforts have been made to overcome the invalidity of interpersonal comparisons of utility. It has been suggested that such points as present wealth or survival level of income might be used to fix an origin, while common scales have been proposed in terms of a "bliss" point (the same for everyone) or on the basis of "just perceptible differences" in utility (Ng, 1975). None of these attempts can be said to be wholly convincing, generally accepted, or operationally feasible. Utility remains (and perhaps it should) as something unique to the individual that can neither be expressed nor measured on an absolute scale.

In operational terms the noncomparability of individual utilities means that we must try to find some way of arriving at a group utility function other than as a purely arithmetic combination of individual functions. If the group is well integrated, with all sharing a strong common interest, the members may be able to agree collectively on a set of preferences that will imply a unique group utility function. In other words, by the social process that takes place within the group, each individual when acting as a group member will take into account the preferences of the others. This will lead to a set of "higher level" individual preferences that will presumably be more in conformity than were the original ones, and in an "ideal" group they would converge to the point where a single group utility function could be defined. Moreover, if the group has some long-term existence (as would be almost essential for it to be well integrated), it may more readily reach a conformity of view because of the opportunities for give-and-take bargaining among members over a range of issues through time (Coleman, 1966).

Conformity of views within the group will not always be reached; but if it is, the decision analyst may be able to proceed using group preferences obtained "under the boardroom door." By this we mean that he obtains a group utility function by questioning the group as such, without concerning himself with the process by which agreement is reached. Indeed, such an appraisal may be used for groups with divergent preferences, provided the analyst can reasonably leave it to the group itself to decide unaided how differences of opinion among its members should be resolved. The chief danger in this case is that the group may be using some informal or unstructured decision process that yields frequent inconsistencies.

Another possibility is that the group agree to use some well-defined decision rule to specify its preferences. Most rely on some voting method, usually some form of "majority rule." Variants of majority rule, familiar in the field of politics, include "simple" majority or "Condorcet," inverse rank score or "Borda," transferrable votes, and repeated or exhaustive ballots (Black, 1958). One such voting method might be used, either directly or indirectly, to obtain a group utility function. If used directly, group voting would be necessary on the simple risky choice questions described in Chapter 4 for deriving a utility function. More plausibly, however, using the indirect approach, the group would elect a spokesman whose (higher level) preferences would be treated as group preferences.

If the procedures described above seem tenuous and unreal, it is worth noting that "democratic" government operates broadly along these lines. A social utility (welfare) function is defined through the ballot box, albeit perhaps very imperfectly. Research into "better" voting procedures is a valuable but difficult field of endeavor.

Groups function as choice units in very many different forms and circumstances; therefore, it is unlikely that any one rule or procedure for establishing group preferences will be universally acceptable. On the other hand, at least some groups might judge it desirable to try to rationalize their decision making along the lines of decision analysis. We could suggest that they choose a method of establishing a group utility function that the members find fair and reasonable. While the circularity of reasoning in this last remark illustrates the difficulties inherent in group choice, we know that groups often do function successfully and make decisions without violent disruption or dissolution. It is the task of the decision analyst to be acquainted as well as possible with methods used by the group and where appropriate to rationalize these so that prescriptive analyses of group decision problems can be successful. On this, see Hogarth (1976).

Risk Sharing

Very often opportunities arise in business for risks to be shared between two or more individuals or firms. A businessman faced with a risky investment opportunity may take one or more partners, or he may be able to raise part of the risk capital on the share market. Farmers generally have limited access to the risk capital market; but they do have opportunities for insuring certain types of risk, and partnership agreements are quite common in agriculture. We will attempt to show when risk sharing is appropriate and how such deals might be struck. We begin with a general statement of risk sharing, followed by a simple example. For simplicity we confine the theoretical treatment to the case where only two decision makers are involved, although the general results can be extended readily to the case where several individuals participate.

Let a be a risky prospect that yields a (one-dimensional) payoff of x_i if state Θ_i results, $i = 1,2,\ldots,m$. Let the probability of each state Θ_i be $P(\Theta_i)$ and assume that these probabilities are agreed upon by both parties involved. (For example, they may be based on abundant "objective" evidence.) Then a *partition* of prospect a describes how the payoff x_i will be divided between two individuals A and B if Θ_i results. We can represent such a partition as shown in Table 5.7, where $x_{iA} + x_{iB} = x_i$ for all i. As can be seen, a partition of a between A and B can be represented by the final two columns as an $m \times 2$ matrix of payoffs to each individual given each state. This matrix, which we will call X, defines a partition of the prospect a.

TABLE 5.7. Partition of a Risky Prospect

State	Probability	Total Payoff	Payoff to A	Payoff to B
Θ_1	$P(\Theta_1)$	x_1	x_{1A}	x_{1B}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Θ_i	$P(\Theta_i)$	x_i	x_{iA}	x_{iB}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Θ_m	$P(\Theta_m)$	x_m	x_{mA}	x_{mB}

We next assume that the preferences of A and B for the payoffs are encoded as utility functions U_A and U_B respectively and that these are defined so that $U_A(0) = U_B(0) = 0$; i.e., both functions assign zero utility to zero payoff (no change in wealth or income). For any partition X of a we can compute the pair of utilities $[U_A(X), U_B(X)]$, where

$$U_A(X) = \sum_i P(\Theta_i)\, U_A(x_{iA})$$

and

$$U_B(X) = \sum_i P(\Theta_i)\, U_B(x_{iB})$$

This joint utility evaluation can be plotted, for example, as point T in Figure 5.9, the location of T depending upon the particular partition of a.

By considering all possible partitions of a, we can construct the set of joint utility evaluations for all X. The set will be convex, or it can be made so, because any indentations in its boundary can be filled by forming linear combinations of any two appropriate points in the set by randomiza-

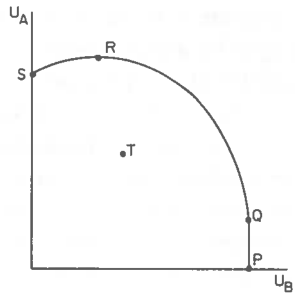


FIG. 5.9. Joint utility evaluations for partitions of a risky prospect.

tion. Thus in Figure 5.9 the random mixture $p(P) + (1 - p)Q$, where p is a probability, locates a point on the straight line joining P and Q. A convex set of joint utility evaluations is shown in Figure 5.9. Note that the particular shape of the set shown, apart from its convexity, is quite arbitrary.

For the set of utility evaluations shown in Figure 5.9 we may deduce that A acting alone would at best operate at point S. Given the opportunity of sharing with B, he would, if motivated only by self-interest, prefer to operate at point R. For B, on the other hand, points P and Q both yield the same utility; and unless B considered A's interests, he would be indifferent between all the points on the line segment PQ. In other words, if the prospect is offered to A, he would be sensible to elect to share it with B to arrive at point R. If the offer is to B, he has no incentive to share with A, although he can make A better off with no loss to himself by sharing at point Q.

A more difficult problem arises if the prospect is offered to A and B jointly in such a way that neither can dictate terms to the other. To determine what partition of a would then be appropriate, we can first eliminate all clearly suboptimal points in the joint utility evaluation set. We define a partition as *Pareto optimal* or *efficient* if there is no other partition preferred by one individual that is not less preferred by the other. Thus in Figure 5.9 point T is not Pareto optimal since it is "dominated" by points in the set above and to the right of it. Similarly, point S is dominated by point R, and point P is dominated by point Q. By an extension of this reasoning we can deduce that the efficient partitions of a all lie on the northeast boundary of the set of joint utility evaluations between points Q and R.

Thus far the going has been straightforward, but now it gets tougher. How are we to determine which point on the boundary QR is "best?" We know we cannot legitimately compare utility gains for one individual with utility losses for another. One possibility is that the two individuals might not act in a totally self-interested way, but each would take account of the other's preferences. In other words, as we proposed in relation to group choice, each would form a "higher level" utility function so that the distance between Q and R would become less and they might eventually reach agreement on a joint (higher level) optimal partition of the prospect.

In the absence of such convergence of preferences, some bargaining process would presumably take place between A and B. If the prospect is offered in a form such that neither can "go it alone," then both have a vested interest in finding some compromise. The determination of the best strategy for each to follow in such bargaining lies in the field of game theory. Each must decide what he ought to do in the light of how his "opponent" might behave. The resolution of such competitive games is beyond the scope of this discussion, and we refer interested readers to an appropriate text such as Luce and Raiffa (1957) or Levin and Des Jardins (1970).

If the two risk sharers fail to reach an agreement by either of these two methods, they might decide to refer their problem to an arbitrator. He would need the wisdom of Solomon to find a fair compromise; but given the need to reach some solution, he might well use some arbitrary rule such as to select the point on the efficient boundary where the certainty equivalents of the two participants are equal.

We now consider a simple example of risk sharing in agricultural management, relating to insurance. Suppose a farmer has some fixed asset (such as a shed) that is worth x dollars and could be completely destroyed by fire during the next year with probability p. He can insure any proportion q of the value of the asset against this risk at a cost of c per dollar of cover. Assuming he is risk averse, what amount of insurance should he buy?

To emphasize the risk sharing aspect of this problem, the data are set out in Table 5.8, following the general layout introduced in Table 5.7. Note that for each state, "burns" or "doesn't burn," the partitioned payoffs sum to the total payoff. If the asset is destroyed, the farmer's loss is x dollars less the sum insured, and he must also pay the insurance premium, i.e., a payoff of $-x + qx - cqx$. The corresponding payoff to the insurer consists of his payment to the farmer of the insured sum in return for the premium he receives, i.e., a payoff of $qx + cqx$. If there is no fire, a premium of cqx dollars is paid by the farmer to the insurer.

We will analyze the problem only from the farmer's point of view. He

TABLE 5.8. Risk Sharing through Insurance

Event	Probability	Total Payoff	Partitioned Payoff: Farmer	Partitioned Payoff: Insurer
Burns	p	$-x$	$-x + qx - cqx$	$-qx + cqx$
Doesn't burn	$1 - p$	0	$-cqx$	cqx

must decide how much cover to buy at a specified cost with no opportunity for bargaining. His attitude to risk will obviously condition his optimal decision. Insurance companies, on the other hand, are generally large enough for us to assume that they would be indifferent to risk in relation to all but very large contracts. In this case the expected money value to the insurer can be assumed to be positive, so that the company can be expected to agree to accept any proportion of the farmer's risk that he requests.

Let $U(\cdot)$ denote the farmer's (risk-averse) utility function. Then the expected utility to the farmer when he insures a proportion q of the value of the asset is

$$U = pU(-x + qx - cqx) + (1 - p)U(-cqx) \qquad (5.14)$$

To find the best level of insurance cover, we need to differentiate equation (5.14) with respect to q, set the result equal to zero, and solve for q subject to the conditions that $0 \leq q \leq 1$ and $d^2U/dq^2 < 0$. For example, adopting the utility function $U = (W_0 + x_f)^{0.8}$ where W_0 is the farmer's initial wealth and x_f is the payoff to him after insuring (both in thousands of dollars), we find that

$$q = [B(W_0 - x) - AW_0]/\{[B(c - 1) - Ac]x\} \qquad (5.15)$$

where $A = [p(1 - c)]^5$

$B = [c(1 - p)]^5$

By way of illustration, if $W_0 = 3$, $x = 1$, $p = 0.001$, and $c = 0.00105$, substituting in equation (5.15) yields $q = 0.35$. The utility function chosen has a Pratt coefficient from (4.10) of $r(W) = 0.2/(W_0 + x_f)$, implying decreasing risk aversion with increasing wealth or payoff. As might be expected, therefore, the optimal amount of insurance falls with increase in W_0 or in c, but more cover is justified as x increases relative to W_0.

Some Other Multiperson Decision Situations

Two other kinds of multiperson decisions warrant mention here. The first is team decision making. The economic theory of teams, developed primarily by Marschak and Radner (1972), relates to the prescriptive analysis of choice in organizations in which there is unanimity of beliefs and preferences among the organization (team) members. Team theory addresses three main questions: (1) what information members should collect about the uncertain environment confronting the team, (2) how members should communicate this information to other team members, and (3) how each member should act on the information he receives. Team theory therefore avoids some of the more intractable problems of multiperson choice and focuses on the information-processing aspects of organizational choice. While a complete sharing of goals and beliefs may be an overoptimistic assumption for most real organizations, it may often be a reasonable first approximation yielding policy prescriptions that can be adapted to suit the real-world circumstances.

A second kind of multiperson decision that deserves special comment is social choice, by which we mean decisions by planners or policy makers that affect large numbers of people. Decisions such as where to locate a new large-scale dam, what degree of tariff protection to afford to farmers, or whether to institute land reform fall into our category of social choice. In practice, political considerations may have an appreciable bearing on such decisions. The decisions reached may be more or less rational in terms of the preferences of power-seeking politicians, which will be "social" to the extent that they reflect the voting inclinations of the electorate. Politics aside, rational analysis of social choice options is obviously desirable. Project appraisal techniques, such as investment and benefit-cost analysis, are widely used and well accepted in several areas of social policy making. There is an extensive literature on these techniques of social decision analysis, and we cannot attempt to treat the topic adequately here. Nevertheless, a few comments seem to be worthwhile.

It is clear that many social choices relate to actions whose consequences are highly uncertain. Our treatment of individual choice highlights the need to take explicit account of such uncertainties. Too often, social choice options are analyzed under assumed certainty, or the recognition of uncertainty is confined to a limited amount of sensitivity analysis. Seldom are appropriate distributions for the relevant uncertain quantities included explicitly in the analysis. Accounting for risk in such an analysis adds an extra dimension of complexity to an already difficult task, but we believe that the extra effort will usually be well justified in terms of more complete and valid representation of the possible consequences of alternative actions. The above discussion of the consensus problem is relevant to the formulation of probability distributions for the analysis of social choice under uncertainty.

The problem of interpersonal utility comparison again arises in relation to social choice. A new airport may make air travel easier and cheaper for travelers, but those living nearby will suffer greater noise and other disturbance. The Pareto principle has sometimes been invoked in relation to social choices, implying that no decision can be unequivocally supported if it means that some people will be better off and at least one other worse off. This is usually just an argument for the status quo. If used as the only basis for action, it would, for example, never have allowed land reform, universal education, or the abolition of slavery (Sen, 1970). In most realistic social choices the utility gains of one group must somehow be matched against the utility losses of the other; but how? Moreover, as our example illustrates, there is often the added complication that such social decisions have multiattributed consequences for individuals. Residents near a new airport may be adversely affected by noise, increased traffic densities, more air pollution, and lowered land values, but they may benefit from new roads or better public transport facilities. The assessment of utility for several such attributes acting together was discussed in Chapter 4. Although operational procedures were outlined, their implementation is far from easy, and the problem of aggregating the resulting individual utility values remains unsolved.

Despite theoretical and operational difficulties, the importance of many social choice issues is such that some analysts have tackled the evaluation of multiattributed social consequences (e.g., Edwards, 1972; Keeney, 1973; Sinden, 1974). It may be that with more research and experience, ways will be found of more adequately extending decision analysis to deal with this important category of human choice.

PROBLEMS

5.1. Mr. O. K. Shambles of Willigobung has to choose between contracting to purchase 1000, 1200, or 1600 cattle for fattening on summer pasture. His profit depends on whether the pasture growing season is good, fair, or poor—for which events his strengths of conviction are 0.3, 0.4, and 0.3 respectively. The budgeted consequences, in terms of dollar profits per animal, are as follows:

Type of Season	Buy 1000	Buy 1200	Buy 1600
Good	18	20	25
Fair	10	8	6
Poor	8	2	−8

If desired, Shambles can purchase a forecast of the type of season for $300. His likelihoods for this forecast are as follows:

Type of Season	Likelihoods		
	Good	Fair	Poor
Good	0.6	0.3	0.1
Fair	0.2	0.5	0.3
Poor	0.1	0.3	0.6

(a) What is the prior optimal act?
(b) What is U(EVPI) and v_p?
(c) What is Shambles' optimal strategy?
(d) Is the forecast worth purchasing?
(e) How efficient is the predictor?
(f) What is the maximum price that should be paid for a perfect predictor?
(g) What is the maximum price that Shambles should pay for the actual predictor?

In answering the above questions, use the extensive form of analysis (as per the layout of Table 5.6) and assume *first* that Shambles' utility function is linear and *second* that it is adequately respected by $U = 2.05X - 0.01X^2$, $X \leq 80$, where X is thousands of dollars gain or loss.

5.2. For the case where Shambles is assumed to have a linear utility function, lay out the solution to Problem 5.1 in the form of a decision tree.

5.3. Jack Doughan, 30 years old, with wife Joana and three children, runs a 650-ha wheat and sheep property in an area with an average annual rainfall of 635 mm. He took over the farm following the rather premature death of his father. At that time, probate difficulties and the need to pay a brother and sister their share of the estate meant that he went very heavily into debt. Subsequently, the introduction of wheat quotas and a decline in wool prices have caused him financial difficulty. The economic position of the farm is shown by the following current annual budget.

FARM BUDGET: 650 ha
(Current prices)

Capital Position	
Assets (fair sale values):	
Livestock	$ 15,000
Plant and improvements	20,000
Land	80,000
Total	$115,000
Liabilities:	
Debt capital	$ 50,000
Equity capital	65,000
Total	$115,000
Profitability	
Gross Margins:	
60 ha wheat @ $43	$ 2,580
30 ha oats @ $25	750
40 ha barley @ $38	1,520
1200 merino ewes @ $5	6,000
300 cross-bred ewes @ $6.20	1,860
30 beef cows @ $80	2,400
Total gross margin	$ 15,110
Miscellaneous income	190
Total	$ 15,300

Overhead Costs	
Farm fixed costs	$ 6,000
Interest on borrowed funds	3,500
Total	$ 9,500
Farm Operating Profit	$ 5,800

Because of the pronouncements of certain well-known agricultural economists, the Doughans feel that perhaps they should "get big or get out." Unfortunately, most of the neighboring properties are well-established medium- to large-sized businesses, and it is unlikely that any of these will come on the market even if they could afford such a purchase. However, the only neighboring small-scale operator has put his place up for sale. It is a property of 200 ha made up entirely of good arable land, all of it capable of growing wheat. The asking price is $42,500 on a walk-in walk-out basis, but Jack believes he could get the place for around $40,000 cash if he acted quickly. He believes he could finance the purchase with a private mortgage from his brother for the full amount, amortized at 7% over 20 years. His anxiety, shared by Joana, is that if he incurs this debt, he might be unable to meet the debt servicing charges.

Jack thinks his main alternative course of action is to sell out. He has no special qualifications for any other kind of employment, although he did reasonably well at school and graduated at a sufficiently high level to virtually guarantee him entrance to a university if he decided to continue his education. At the worst he believes he would have no difficulty in finding an unskilled job in the city. However, he has always had an interest in accounting and feels that if he must leave the land, this is the profession he would prefer. In fact, provided he obtains professional accounting qualification and can raise at least $30,000 in capital, there is at least a chance of his obtaining a seat on the board of directors of a small but successful pharmaceutical company owned by Joana's father.

Jack believes he might be able to enter the accounting profession via a three-year university degree in economics or commerce, followed by one year of professional work with a firm of certified public accountants. He wonders, however, whether he will be able to make the grade after so many years away from formal study. In addition, there is the question of how he will support the family during the three-year period. To obtain some evidence on his chances of success in full-time university work, Jack is considering enrolling as a part-time economics student.

During the time (late August) that he is mulling over these various alternatives, Jack is presented with yet another possibility. He is offered a job as clerk-bookkeeper with a real estate firm in the nearby town. The job offers no prospects of advancement, but there would be a steady wage of $4000 per year and good security of employment. The position will be held open for three weeks, but if Jack does not accept within that time it will then be advertized and presumably filled. At this point, Jack and Joana feel that their decision problem is becoming too complex for them to make a rational decision, and they enlist your help. Jack asks you whether the problem might be formalized so that he and Joana can obtain a clearer picture of the available options and their associated risks. You feel in need of a break and spend a couple of days early in September with the Doughans to collect such information as you need about their beliefs and preferences in order to try to help them. You have in mind a decision tree approach to their problem.

Consultation with the Doughans leads to the following timetable of relevant actions and events:

Last date for decision on job in local town:	September 20 this year
Farm to be sold following acceptance of job in local town by:	January 1 next year
Last date for decision to buy extra land:	November 1 this year
Last date for decision to enroll in part-time courses:	January 1 next year
Results of part-time courses available:	December 1 next year
Last date for decision to enter full-time university course:	
Following part-time trial:	January 1 year after next
No part-time courses:	January 1 next year

Jack insists the farm must be sold before he can enter the university. He also feels that because of his age and the ages of his children, any decision to leave the farm cannot be delayed beyond January 1 next year.

Further, it is deduced that the most appropriate decision criterion is expected long-run income, defined as the operating profit of the farm business or annual gross wage or salary from paid employment, plus interest on nonfarm investments at 8%, plus or minus adjustments for nonmonetary benefits or costs. These latter include the following items:

(a) The value of perquisites, tax concessions, etc., in farming, is assessed at $500 per year.

(b) A penalty cost is attached to wages or salaries following full-time university study, which Joana feels is necessary as recompense for the effort involved and the reduced family living standard during Jack's full-time university study. Joana estimates that they would require an extra $500 per year on Jack's long-term earnings for every year of full-time study. Thus the approximate deductions are: three years of study (successful), $1500; or one or two years (unsuccessful), $750.

(c) Jack is very concerned about the risk of bankruptcy if he buys the extra 200 ha using borrowed funds. (He regards this risk on his present farm area and debt level as negligible.) Taking a very pessimistic view, he suggests that bankruptcy should be evaluated in terms of the loss of all his equity capital, subsequent reduction of his income to the rates for unskilled employment in the city, and an annual penalty cost of $1000 per year on his long-term salary, representing redemption of his debt to his brother.

(d) With the exception of (c) above, Jack and Joana feel that risks associated with the various options are more or less similar, and do not wish to make any special allowances for risk.

Further consultation with the Doughans yields the following information on (A) potential capital payments and receipts, (B) the likely farm budget if Jack should buy the extra 200 ha of land, and (C) their probability judgments for various relevant possibilities.

A. POTENTIAL CAPITAL PAYMENTS AND RECEIPTS

Net receipts from sale of farm	
Fair price:	$ 65,000
Poor price:	40,000
Purchase of extra land, all capital borrowed:	40,000
Removal costs to local town, including house purchase:	12,000
Cost of part-time study, including fees, books, and farm income forgone:	500
Cost of full-time university study, including fees, books, removal, living costs, rent, etc., at $5000 per year (with account taken of interest received on invested capital):	
Three years of successful study:	15,000
Expected cost if unsuccessful:	7,500
Removal costs to city, including house purchase:	30,000

B. FARM BUDGET: 850 ha
(Current prices)

Capital Position	
Assets (fair sale values):	
Livestock	$ 21,000
Plant and improvements	23,000
Land	111,000
Total	$155,000
Liabilities:	
Debt capital	90,000
Equity capital	65,000
Total	$155,000
Profitability	
Gross Margins:	
100 ha wheat @ $43	4,300
30 ha oats @ $25	750
80 ha barley @ $38	3,040
1200 merino ewes @ $5	6,000
750 cross-bred ewes @ $6.20	4,650
50 beef cows @ $80	4,000
Total gross margin	$ 22,740
Miscellaneous income	260
Total	$ 23,000
Overhead Costs:	
Farm fixed costs	$ 8,000
Interest on borrowed funds	6,300
Total	$ 14,300
Farm Operating Profit	$ 8,700

C. THE DOUGHANS' PROBABILITY JUDGMENTS

Best price offered for farm:
P(fair price) = 0.65
P(poor price) = 0.35

Long-run profitability of farming:

		Operating profit 650 ha	Operating profit 850 ha
P(farm profits improved)	= 0.30	$9,000	$13,000
P(farm profits same)	= 0.50	$5,800	$ 8,700
P(farm profits worse)	= 0.20	$4,000	$ 5,200

(They also believe that there is a chance of 0.05, included in the 0.2 probability, that the economic situation will become so bad that Jack will be forced into bankruptcy if he increases his liabilities by buying the extra land.)

Result in part-time courses:
P(good result) = 0.05
P(fair result) = 0.60
P(poor result) = 0.35

Result in full-time university degree:
(a) Prior probabilities:
P(pass) = 0.61
P(fail) = 0.39
(b) Posterior probabilities, given result in part-time courses:
P(pass | good result) = 0.85
P(fail | good result) = 0.15
P(pass | fair result) = 0.80
P(fail | fair result) = 0.20
P(pass | poor result) = 0.25
P(fail | poor result) = 0.75

Long-run wage in unskilled city job:
P($6000 per yr) = 0.30
P($5000 per yr) = 0.40
P($4000 per yr) = 0.30

Long-run salary as qualified accountant:
(a) With $30,000 or more available:
P($14,000 per yr) = 0.40
P($8000 per yr) = 0.40
P($5000 per yr) = 0.20
(b) Without $30,000 available:
P($14,000 per yr) = 0.10
P($8000 per yr) = 0.70
P($5000 per yr) = 0.20

Jack and Joana agree that if you need any further information to assess their problem you should use your own best estimates.

Using a decision tree approach, prepare a report for the Doughans indicating what you believe to be their optimal strategy. Prepare this in a fashion that will be understandable to Jack and Joana and, if relevant, indicate where you have used your own judgment.

5.4. As well as a linear utility function, Midas Nut Corporation has a macadamia plantation of 10,000 trees. In the past Midas has used manual labor for harvesting at a cost of 50 cents per tree. The same could be done this year. However, a contractor is also available this year with a new mechanical picker of Hawaiian design that can harvest the nuts from 500 trees per day. The contractor's charge is $100 per day. Midas has made enquiries in Hawaii and found that ten farmers who used the picker last year had the following discrete outcomes in terms of the percentage of their trees damaged:

Farmer	Percentage of Trees Damaged	Farmer	Percentage of Trees Damaged
1	0.05	6	0.05
2	0.07	7	0.07
3	0.05	8	0.03
4	0.07	9	0.10
5	0.05	10	0.03

The cost of a damaged tree is $5. In further discussion with the contractor, he agrees to a trial demonstration on three trees at no cost to Midas.

(a) What is the prior optimal choice of Midas?
(b) What is the EVPI?
(c) What is the optimal strategy for Midas?
(d) How efficient is the proposed trial?
(e) How much could Midas pay for the demonstration?
(f) Assume that the contractor is not prepared to offer a free demonstration but will demonstrate on any number of trees at a cost of $20 plus $1 per tree. Evaluate the relative merit of purchasing demonstrations involving one, two, and three trees. Could you determine the optimum-sized demonstration?

5.5. Farmers face a two-month period of bloat danger in a particular region where there are normally some 10,000 head of cattle being fattened on improved pasture. The judgment of relevant experts is that in any year the probability of high, medium, or low bloat incidence is 0.2, 0.5, and 0.3 respectively. Possible farmer response to this danger is to use no controls or to use pluronics or capsules. Assuming that the conditional expectation of annual cost per 100 animals of these three approaches is as shown below, discuss the economics of research aimed at developing a predictor that would give a timely forecast of the type of bloat incidence to be expected in a particular season.

Bloat Incidence	Cost ($/100 head)		
	No control	Pluronics	Capsules
High	3000	900	600
Medium	1000	250	200
Low	100	150	150

5.6. Dr. D. P. Podszol, an employee of Gundy Pastoral and the owner of a linear utility function, has to advise on the purchase of 1000 ha of land for development. There are two alternative areas. Area I has an unknown proportion of its total hectarage affected by salt, whereas area II is guaranteed free of toxicity. Area I is \$10,000 cheaper than area II. It costs \$200/ha to reclaim salt-affected land so that it is comparable to unaffected land. If taken for analysis, soil samples would cost $\$(50 + 10n)$, where n is the number of samples. From previous experience and visual inspection of area I, Podszol judges the chance of area I having various proportions of salt-affected land to be:

Proportion Affected	Chance
0.01	0.30
0.03	0.35
0.05	0.20
0.10	0.10
0.25	0.05

Based on the above information, Podszol is inclined to recommend purchase of area II. He is, however, somewhat doubtful and decides that he will analyze eight soil samples from randomly selected sites.

(a) What is the prior optimal choice?
(b) What is the EVPI?
(c) What strategy should Podszol follow relative to potential results from the soil sampling?
(d) Suppose two of the eight soil samples show salt toxicity. Would a further sampling of four sites be worthwhile?

5.7. In situations where samples of varying size may be taken to gain additional information, pertinent quantities referred to in the decision analysis literature (e.g., Winkler, 1972) are EVPI, EVSI, CS (the cost of sampling), and ENGS (the expected net gain from sampling). For a sample of size n, ENGS is defined as $\text{ENGS}(n) = \text{EVSI}(n) - \text{CS}(n)$. Show graphically how you would expect EVPI, EVSI, and ENGS to vary as n increases from zero. Assume that EVPI is costless. Does your diagram indicate an optimal sample size n^*?

5.8. If sampling costs \$100 plus \$10 per unit sampled, evaluate the expected net gain from sampling for the example of Section 5.6 with sample sizes of 25, 64, 100, and 144. What is the approximate optimal sample size? What if the cost of sampling were $\$(100 + 40n)$?

5.9. By solving for the percentage change required in each parameter to change the optimal act, investigate the sensitivity of the prior and posterior solutions of the new product problem of Section 5.6 to:

(a) Net return per unit sold.
(b) Setup cost.
(c) Our prior mean estimate μ_0.
(d) Our prior standard deviation σ_0.

5.10. Fred Vice, a district extension officer, is faced with the problem of whether to recommend the spraying of wheat crops in his region for weed control in the coming season. Spraying costs \$4/ha and the net profit from each extra kilogram of wheat produced per hectare is \$0.02. Information from other analogous districts along with his own local experience has lead Fred to estimate that mean yield after spraying would be 1800 kg/ha with a 50/50 chance of lying between 1400 and 2200 kg/ha. Since making this estimate, Fred has received information on spraying trials conducted on 25 representative areas of wheat in his district. Yield in these trials appeared to be normally distributed with a mean of 1700 kg/ha and a standard deviation of 300 kg/ha. Mean yield on the control areas was 1500 kg/ha, which tallies with average yield in the region.

(a) What is the break-even yield for spraying?
(b) What is Fred's prior distribution for mean yield?
(c) What is the prior optimal recommendation?
(d) What is the prior EVPI?
(e) What is the posterior distribution for mean yield?
(f) What is the posterior optimal recommendation?
(g) What is the posterior EVPI?
(h) In this problem does it make sense to talk of EVSI?

5.11. A farmer has available an investment that will yield an expected annual net cash flow of $2000 with a standard deviation of $2000. He can retain any proportion q of the investment for himself, sharing the remaining proportion of $1 - q$ for a guaranteed annual payment to the partner of $1000(1 - q)$ dollars. After this payment, annual returns are shared in the proportion $q{:}1 - q$. The farmer's utility function for x (the annual net cash flow in dollars that he receives) is $U = x - 0.0002x^2$. Find the optimal proportion q that he should retain. What if his utility function is $U = \log(10{,}000 + x)$?

5.12. (a) Relative to some major public decision of interest to you, comment on the following statement by Howard et al. (1972, p. 1202):

> When the [public] consequences of deploying new technology are uncertain, who will make the choice? While many individuals or groups may share responsibility, decision analysis separates the roles of the executive decision maker, the expert, and the analyst. The analyst's role is to structure a complex problem in a tractable manner so that the uncertain consequences of the alternative actions may be assessed. Various experts provide the technical information. . . . The decision maker acts for society. . . . The analysis provides a mechanism for integration and communication so that the technical judgements of the experts and the value judgements of the decision maker may be seen in relation to each other, examined and debated. Decision analysis makes not only the decision but the decision process a matter of record. For any complex [public] decision that may affect the lives of millions, a decision analysis showing explicitly the uncertainties and decision criteria can and should be carried out.

(b) Outline the procedures you would follow in applying the points made in Section 5.7 to one of the examples or problems presented in this chapter.

REFERENCES AND SELECTED FURTHER READING

All the references listed for Chapter 1 elaborate procedures for decision analysis but the best are Raiffa (1968) and Winkler (1972). More formal mathematical or philosophical presentation is available in:

de Finetti, B. 1972. *Probability and Induction: The Art of Guessing*. New York: Wiley.
La Valle, I. H. 1970. *An Introduction to Probability, Decision and Inference*. New York: Holt, Rinehart and Winston.
Raiffa, H., and R. Schlaifer. 1961. *Applied Statistical Decision Theory*. Boston: Harvard Business School.

In addition to the items listed later in relation to decision trees, a sampling of the application of decision analysis in a variety of fields is given by the following:

Anderson, J. R. 1976. Modeling decision making under risk. Paper for ADC Conference on Risk, Uncertainty and Agricultural Development, Dept. Agric. Econ. Busin. Mgmt, Univ. New England, Armidale.
Balch, M., D. McFadden, and S. Wu, eds. 1974. *Essays on Economic Behavior under Uncertainty*. Amsterdam: North-Holland.
Betaque, N. E., and G. A. Gorry. 1971. Automating judgmental decision making for a serious medical problem. *Mgmt. Sci.* 17(8): B421–34.
Brown, R. V. 1970. Do managers find decision theory useful? *Harvard Busin. Rev.* 48(1): 78–89.
Byerlee, D. R., and J. R. Anderson. 1969. Value of predictors of uncontrolled factors in response functions. *Australian J. Agric. Econ.* 13(2): 118–27.
Carlson, G. A. 1970. A decision theoretic approach to crop disease prediction and control. *Am. J. Agric. Econ.* 52(2): 216–23.
Eidman, V. R., G. W. Dean, and H. O. Carter. 1967. An application of statistical decision theory to commercial turkey production. *J. Fm. Econ.* 49(4): 852–68.
Ellis, H. M., and R. L. Keeney. 1972. A rational approach for government decisions concerning air pollution. In *Analysis of Public Systems*, A. W. Drake et al., eds. Cambridge: MIT Press.
Forst, B. E. 1974. Decision analysis and medical malpractice. *Ops. Res.* 22(1): 1–12.
Grayson, C. J. 1960. *Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators*. Boston: Graduate School Busin. Admin., Harvard Univ.
Moore, P. G., and H. Thomas. 1973. The rev counter decision. *Opl. Res. Q.* 24(3): 337–51.
O'Mara, G. T. 1971. A decision-theoretic view of the microeconomics of technique diffusion. Stanford Univ. Ph.D. diss. Ann Arbor: Univ. Microfilms.

Some references noting the inadequacies of game theory approaches to risky choice are:

Dillon, J. L. 1962. Applications of game theory in agricultural economics: Review and requiem. *Australian J. Agric. Econ.* 6(2): 20–35.
Halter, A. N., and G. W. Dean. 1971. *Decisions under Uncertainty with Research Applications*. Cincinnati: South-Western.
Luce, R. D., and H. Raiffa. 1957. *Games and Decisions*. New York: Wiley.
Officer, R. R., and J. R. Anderson. 1968. Risk, uncertainty and farm management decisions. *Rev. Mktg. Agric. Econ.* 36(1): 3–19.

The use of decision trees in decision analysis is outlined and illustrated in the texts by Raiffa (1968) and Winkler (1972). A more detailed treatment of decision trees is given by:

Brown, R. V., A. S. Kahr, and C. Peterson. 1974. *Decision Analysis for the Manager*. New York: Holt, Rinehart and Winston.
Pratt, J. W., H. Raiffa, and R. Schlaifer. 1965. *Introduction to Statistical Decision Theory*. New York: McGraw-Hill.
Schlaifer, R. 1969. *Analysis of Decisions under Uncertainty*. New York: McGraw-Hill.
Thomas, H. 1972. *Decision Theory and the Manager*. London: Pitman.

The following references provide excellent illustrations of the use of decision trees in a variety of contexts.

Byrnes, W. G., and B. K. Chesterton. 1973. *Decisions, Strategies and New Ventures.* London: Allen and Unwin.
Ginsberg, A. S., and F. L. Offensend. 1968. An application of decision theory to a medical diagnosis-treatment problem. *IEEE Trans. Sys. Sci. Cybern.* SSC-4(3): 355–62.
Howard, R. A., J. E. Matheson, and D. W. North. 1972. The decision to seed hurricanes. *Science* 176(4040): 1191–202.
Rae, A. N. 1971. An empirical application and evaluation of discrete stochastic programming in farm management. *Am. J. Agric. Econ.* 53(4): 625–38.
———. 1972/73. Horticultural decision-making under non-certainty. *Scient. Hort.* 24: 207–23.

It may sometimes be desired to use continuous probability distributions in decision analysis. If the distributions involved can be satisfactorily approximated by either a normal, rectangular, or beta distribution, analysis is much easier than otherwise. The following references give a more extensive treatment of continuous analysis than we have given in Section 5.6.

Jedamus, P., and R. Frame. 1969. *Business Decision Theory.* New York: McGraw-Hill.

A well-written middle-level introductory text. Covers the normal distribution for continuous analysis with two acts.

Morgan, B. W. 1968. *An Introduction to Bayesian Statistical Decision Processes.* Englewood Cliffs, N.J.: Prentice-Hall.

A well-written middle-level introductory text. Covers the normal distribution for continuous analysis with two acts.

Sasaki, K. 1968. *Statistics for Modern Business Decision Making.* Belmont, Calif.: Wadsworth.

Excellent introduction but somewhat formal. Covers the normal and beta distributions, and also the application of a Bayesian approach to regression analysis.

Winkler, R. L. 1972. *Introduction to Bayesian Inference and Decision.* New York: Holt, Rinehart and Winston.

In lengthy but understandable style, Chapter 4 and the latter half of Chapter 6 provide excellent coverage of continuous analysis with the normal and beta distributions.

The Bayesian general model assumes the number of possible states and the number of alternative acts to be fixed. No allowance is made for gaining information pertinent to revision of the size of the payoff matrix. Such revision implies a process of search and learning, typically at some cost. Discussion of the value of such search and learning and the criteria for allocation of search effort are presented by:

Morris, W. T. 1968. *Management Science: A Bayesian Introduction.* Englewood Cliffs, N.J.: Prentice-Hall.
Radner, R. 1964. Mathematical specification of goals for decision problems. In *Human Judgments and Optimality,* M. W. Shelly and G. L. Bryan, eds., pp. 178–216. New York: Wiley.

The extension of decision analysis to multiperson decisions opens up a vast field of literature including, for example, welfare economics and benefit-cost analysis. No attempt has been made to survey this whole field, and the following references generally relate fairly directly to the discussion of Section 5.8. See also Luce and Raiffa (1957, Ch. 14).

Arrow, K. J. 1963. *Social Choice and Individual Values,* 2nd ed. New York: Wiley.

Arrow's shattering contribution to social choice theory, first published in 1951, in which the impossibility theorem is presented.

———. 1967. Values and collective decision-making. In *Philosophy, Politics and Society,* 3rd Ser., P. Laslett and W. G. Lunciman, eds., pp. 215–32. Oxford: Blackwell.

A good short exposition of the central idea of his 1951 book.

Black, D. 1958. *The Theory of Committees and Elections.* Cambridge: Cambridge Univ. Press.

A comprehensive treatment of voting procedures for group choice.

Borch, K. 1974. *The Mathematical Theory of Insurance.* Lexington, Mass.: Heath.

An annotated selection of papers on insurance published by Borch since 1960. The treatment is generally in terms of utility theory, but usually from the viewpoint of the insurance company rather than the client.

Coleman, J. S. 1966. The possibility of a social welfare function. *Am. Econ. Rev.* 56(5): 1105–22.

Dalkey, N. C. 1967. Delphi. P-3704, The RAND Corp., Santa Monica, Calif.

Discusses the problems of bias in committee decision making and describes the Delphi method, noting its advantages. See also RM-5888-PR from RAND by the same author (1969) for a good general description of Delphi.

Davies, J. H. 1973. Group decision and social interaction: A theory of social decision schemes. *Psychol. Rev.* 80(2): 97–125.

Deals with amalgamation of individual preferences into a collective choice. Proposes a general theory of group decision making. Behavioral rather than normative in approach.

de Groot, M. H. 1974. Reaching a consensus. *J. Am. Statist. Ass.* 69(345): 118–21.

Provides a model of how a group might reach a subjective probability consensus by each member weighting the opinions of the others.

Edwards, W. 1972. Social utilities. In *Decision and Risk Analysis: Powerful New Tools for Management,* A. Lesser, ed., pp. 112–29. Hoboken, N.J.: The Engineering Economist.

An application of utility analysis to the selection of experiments for Skylab. Deals with group utility assessments under assumed certainty.

Fishburn, P. C. 1970. The irrationality of transitivity in social choice. *Behavl. Sci.* 15(2): 119–23.

Examples are presented that illustrate the untenability of the transitivity condition in relation to group decision making.

———. 1972. Social choice functions in group decision theory. In *Decisions and Risk Analysis: Powerful New Tools for Management,* A. Lesser, ed., pp. 161–95. Hoboken, N.J.: The Engineering Economist.

A survey article with a good list of 100 references, not quite as formal as most of this author's work.

———. 1973. *Theory of Social Choice.* Princeton: Princeton Univ. Press.

A formal and rigorous treatment including a comprehensive discussion of the various voting procedures.

Gustafson, D. H., R. K. Shukla, A. Delbecq, and G. W. Walster. 1973. A comparative study of the differences in subjective likelihood estimates made by individuals, interacting groups, Delphi groups and normal groups. *Organiz. Behav. Hum. Perform.* 9(2): 280–91.

Compares the effectiveness of different methods of forming a probability consensus, and points out the dysfunction of interacting group processes for judgment tasks.

Hogarth, R. M. 1976. *Methods for Aggregating Opinions.* Fontainbleau, France: INSEAD.

Keeney, R. L. 1973. A decision analysis with multiple objectives: The Mexico City Airport. *Bell J. Econ. Mgmt. Sci.* 4(1): 101–17.

An outstanding illustration of the application of the multiattribute utility approach to a social choice problem.

———. 1975. A group preference axiomatization with cardinal utility. Rept. RM-75-47, Intl. Instit. Applied Systems Analysis, Laxenburg.

Presents axioms under which a group utility function is feasible.

Keeney, R. L., and C. W. Kirkwood. 1975. Group decision making using cardinal social welfare functions. *Mgmt. Sci.* 22(4): 430–37.

Considers the problem of constructing a group utility function whose arguments are the individual utility functions of group members.

Keeney, R. L., and H. Raiffa. 1972. A critique of formal analysis of public decision making. In *Analysis of Public Systems,* A. W. Drake et al., eds. Cambridge: MIT Press.

A good general discussion of the pros and cons of decision analysis in public decision making.

Levin, R. I., and R. B. Des Jardins. 1970. *Theory of Games and Strategies.* Scranton, Pa.: International Textbook Co.

A more up-to-date treatment of game theory in a management context.

Little, I. M. D. 1952. Social choice and individual values. *J. Polit. Econ.* 60(5): 422–32.

An early critique of Arrow's impossibility theorem.

MacCrimmon, K. R. 1973. Theories of collective decision. Invited review paper for 4th Res. Conf. Subjective Probability, Utility and Decision Making, Rome. Univ. British Columbia, Vancouver.

Introduces a framework for a general model of multiperson decision making.

Madansky, A. 1964. Externally Bayesian groups. RM-4141-PR, The RAND Corp. Santa Monica, Calif.

Suppose a group of individuals, each of whom is a Bayesian, is required to make a joint decision, and although the individuals all agree on the utility function for the problem, they disagree on the prior distribution of the states of nature. An axiom of group rationality is introduced—viz., that to an outsider the decisions of the group appear like the decisions of a Bayesian—and its implications are explored.

Marschak, J., and R. Radner. 1972. *Economic Theory of Teams.* New Haven: Yale Univ. Press.

The leading work on team theory.

Moskowitz, H. 1975. Some observations on theories of collective decisions. In *Utility, Probability, and Human Decision Making*, D. Wendt and C. Vlek, eds., pp. 381–95. Dordrecht: Reidel.

A useful review.

Mueller, D. C. 1976. Public choice: A survey. *J. Econ. Lit.* 14(2): 395–433.

A wide-ranging review emphasizing the difficulties of consensus in public choice. Lists an extensive bibliography.

Ng, Y-K. 1975. Bentham or Bergson? Finite sensibility, utility functions, and social welfare functions. *Rev. Econ. Stud.* 42(4): 545–69.

Pill, V. 1971. The Delphi method: Substance, context, a critique and an annotated bibliography. *Socio-Econ. Plann. Sci.* 5(1): 57–71.

The Delphi methodology and its historical development are described. An annotated bibliography of 40 references is provided.

Raiffa, H. 1968. *Decision Analysis*. Reading, Mass.: Addison-Wesley.

Chapter 8 provides a good outline of risk sharing and group decision making.

Sen, A. K. 1970. The impossibility of a Paretian liberal. *J. Polit. Econ.* 70(1): 152–57.

Sinden, J. A. 1974. A utility approach to the valuation of recreational and aesthetic experiences. *Am. J. Agric. Econ.* 56(1): 61–72.

Theil, H. 1963. On the symmetry approach to the committee decision problem. *Mgmt. Sci.* 9(3): 380–93.

Proposes a system of weights for forming a group utility function. The problem of the arbitrary origins of individual utility functions is tackled by working with disutilities where the best outcome for each individual is assigned zero disutility. Weights and hence scales are determined such that the disutility that *A* imposes on *B* if *A*'s view prevails is equal to the disutility that *B* imposes on *A* if *B*'s view prevails.

Wilson, R. 1968. The theory of syndicates. *Econometrica* 36(1): 119–32.

The criterion of Pareto optimality is used to determine the construction and conditions for existence of a group utility function and a consensus of the members' probability assessments. The results obtained are mainly negative, establishing that only under very restrictive circumstances will the group's behavior be consistent with the axioms of Bernoullian decision theory.

Winkler, R. L. 1968. The consensus of subjective probability distributions. *Mgmt. Sci.* 15(2): B61–75.

Several methods of combining probability distributions are described and compared. It is shown that under experimental conditions, the different methods may well produce varying results.

Winkler, R. L., and L. J. Cummings. 1972. On the choice of a consensus distribution in Bayesian analysis. *Organiz. Behav. Hum. Perform.* 7(1): 63–76.

An excellent review of the consensus problem and used as the main basis of our discussion in the text.

CHAPTER SIX
PRODUCTION UNDER RISK

IN THIS CHAPTER, which could just as well be entitled "Risky Response Analysis" or "Resource Allocation under Risk", we blend the (riskless) neoclassical theory of crop and livestock production (Heady and Dillon, 1961) with the concepts of decision theory developed in earlier chapters. Our purpose is to explore the impact of risk on agricultural producers who are utility maximizers and face production functions exhibiting diminishing returns. Needless to say, we will establish that risk generally has a significant impact on the way resources should be allocated—a result that raises serious questions about the relevance of any theory of production that specifically ignores risk.

Agricultural firms typically are competitive, face known input prices but uncertain product prices, and face uncertainty in some of the factors that influence the quantity and quality of the output they produce. Such firms usually produce several crop and/or livestock products through the control of many inputs or decision variables. For simplicity we will first discuss the case of a firm using a single variable input to produce a single product whose level of output and price are uncertain. Extension to more complex cases is at least conceptually simple. Discussion of this is followed by an outline of empirical procedures pertinent to the evaluation of risky input-output decisions. Some policy implications arising from the risky nature of agricultural production are then covered.

Throughout, we will assume a production function exhibiting diminishing returns and will follow the approach of setting up risky decisions on resource use in the context of expected utility maximization. This implies the straightforward application of Bernoulli's principle and is a deliberate attempt at analytical consistency. Other approaches to resource allocation under risk, not recognizing the basic axioms of preference under risk, have been suggested. For instance, Day et al. (1971) and Roumasset (1976) have discussed risky decisions on input and output levels for the firm in the context of safety margins and bounds on the probability of loss.

Such approaches generally imply a lexicographic ordering of preferences. We will return to these safety-first notions in our discussion of chance-constrained programming in Chapter 7. Here, however, we will stick to the expected utility maximization model. As our analysis shows, consideration of the classical production function in a unidimensional utility context leads us to some easily appreciated economic implications of production risk. In general these results are as intuition would suggest: if a producer is risk averse, risk acts as a friction to production and induces a lower level of resource use than would otherwise prevail; if a producer prefers risk, the reverse occurs.

6.1 ONE-FACTOR ONE-PRODUCT CASE

Our notation and initial assumptions are as follows: the level of the firm's single input decision variable is denoted by v and its assumed known price per unit by p_v; the uncertain level of output is denoted by y and its uncertain price per unit by p_y; fixed costs are denoted by F. Thus the random variable profit, denoted by x to accord with the notation of earlier chapters, can be defined as

$$x = p_y y - p_v v - F \tag{6.1}$$

where y is some function of v.

We will again assume that the decision maker's preferences for risky profits are encoded in the utility function $U = U(x)$. The optimal level of the decision variable is found according to Bernoulli's principle as the value of v that maximizes expected utility where expectations are taken over the distributions of y and p_y. Note that the distribution of y will generally be conditioned by v since y is some function of v and that, under the reasonable assumption of perfect competition among agricultural producers, p_y will be independent of v.

For all but trivial or unrealistically restrictive specifications of the response and probability functions, utility maximization cannot proceed directly. Two alternative approaches may be employed. Numerical explorations of the decision variable's range for stochastically simulated values of y and p_y could be used to generate achievable values of expected utility. In this way we could approximate the value of the decision variable v that maximizes $U(x)$ given the uncertainty in y and p_y. Alternatively and of direct interest to us here, the problem can be approximated by means of the moment method of Section 4.6—especially when Taylor series terms involving the third and higher derivatives can be ignored. Analysis is further simplified by the reasonably realistic assumption that output y of the individual firm and product price p_y are stochastically independent (i.e., the probability distribution of one does not depend in any way on that of the other). Under these several simplifying assumptions, the mean of profit is given by

$$\begin{aligned} E(x) &= E(p_y y - p_v v - F) = E(p_y)E(y) - p_v v - F \\ &= E(p_y)g(v) - p_v v - F \end{aligned} \tag{6.2}$$

where $g(v) = E(y)$ is an empirical function relating the mean of y to v. Likewise, the variance of profit is given by

$$\begin{aligned} V(x) &= V(p_y y - p_v v - F) \\ &= [E(p_y)]^2 V(y) + [E(y)]^2 V(p_y) + V(p_y)V(y) \\ &= [E(p_y)]^2 h(v) + [g(v)]^2 V(p_y) + V(p_y)h(v) \end{aligned} \tag{6.3}$$

where $h(v) = V(y)$ is an empirical function relating the variance of y to v.

By Taylor series approximation from equation (4.27),

$$U = U[E(x)] + U_2[E(x)]V(x)/2 \tag{6.4}$$

Maximization of this expected utility with respect to v, even under simple functional relationships for $g(v)$ and $h(v)$, is usually messy and best done numerically with the aid of a computer. However, the logic of the operation can be seen as follows. The first-order condition for a maximum of $U = U[E(x), V(x)]$ is that the derivative

$$dU/dv = [\partial U/\partial E(x)][dE(x)/dv] + [\partial U/\partial V(x)][dV(x)/dv] \tag{6.5}$$

is equal to zero which implies

$$0 = dE(x)/dv + \{[\partial U/\partial V(x)]/[\partial U/\partial E(x)]\}dV(x)/dv \tag{6.6}$$

In this expression, the ratio of square-bracket terms measures the rate of utility substitution between $E(x)$ and $V(x)$. That this is so can be seen by considering an isoutility curve in (E, V) space. Along such a curve, utility in terms of $E(x)$ and $V(x)$ is constant so that its total differential dU is zero. The differential dU is related to the differentials $dE(x)$ and $dV(x)$ through the partial derivatives of utility with respect to $E(x)$ and $V(x)$ as

$$dU = [\partial U/\partial E(x)]dE(x) + [\partial U/\partial V(x)]dV(x) \tag{6.7}$$

which is simply rearranged for $dU = 0$ as

$$[dE(x)/dV(x)]_U = -[\partial U/\partial V(x)]/[\partial U/\partial E(x)] \tag{6.8}$$

Magnusson (1969) has called this expression, which is the negative of the curly bracket term in equation (6.6), the *risk evaluation differential quotient* (REDQ). We will discuss this quantity in more detail after completing our analysis of (6.6).

Expressions for $dE(x)/dv$ and $dV(x)/dv$ can be obtained from the middle forms of equations (6.2) and (6.3). Substituting these expressions into (6.6), we have

$$0 = E(p_y)dE(y)/dv - p_v - \text{REDQ}\left(\{[E(p_y)]^2 + V(p_y)\}\cdot dV(y)/dv + 2V(p_y)E(y)dE(y)/dv\right) \qquad (6.9)$$

This expression can be explained most simply by making the further assumption that p_y is not random but has a fixed value so that $E(p_y) = p_y$ and $V(p_y) = 0$. Equation (6.9) then collapses to

$$p_v = p_y dE(y)/dv - \text{REDQ}[p_y^2 dV(y)/dv] \qquad (6.10)$$

which says that the optimal level of v occurs when the marginal factor cost (i.e., input price) is equal to the value of the marginal expected product minus a *marginal risk deduction* that depends on the utility function and the marginal variance of revenue. The additional complexity of (6.10) relative to the condition for profit maximization under certainty (that marginal value product equals marginal factor cost) is readily apparent and epitomizes the analytical costs of attempting to bring uncertainty explicitly into normative analysis of factor use. The same effect can be seen by rearranging (6.10) to give

$$p_y = p_v dv/dE(y) + \text{REDQ}p_y^2 dV(y)/dE(y) \qquad (6.11)$$

which says optimality is achieved when marginal revenue is equal to marginal cost (with respect to changing expected output) plus a marginal addition due to risk.

Consider now some of the additional complexities introduced by accounting for risk. For simplicity, we continue to concentrate on equation (6.10) with its assumption that only y is stochastic. One way of rearranging (6.10) is to expand the variance term. By definition, $V(y)$ may be written as $V(y) = E(y^2) - [E(y)]^2$ so that, with permissible differentiation within the expectation operator under some continuity and convergence properties that typically hold good,

$$dV(y)/dv = 2[E(y dy/dv) - E(y)E(dy/dv)]$$

which, by the definition of covariance, can be written as

$$dV(y)/dv = 2\text{Cov}(y, dy/dv)$$

i.e., $dV(y)/dv$ is equal to twice the covariance between output and marginal product, and may be substituted into (6.10) to yield

$$p_v = p_y dE(y)/dv - \text{REDQ}[2p_y^2 \ \text{Cov}(y, dy/dv)] \qquad (6.12)$$

The covariance between output and marginal product may be positive, zero, or negative; and determination of its sign and magnitude is an empirical matter. Figure 6.1 illustrates crudely both the positive and the negative covariance cases at input levels v^* and v^{**} respectively by presenting three sample response curves from a population of risky response curves. At v^* the slope of the response curves increases with increasing level of output, suggesting a positive covariance and correlation; the converse situation is depicted at v^{**}. It is a personal issue whether one finds it simpler to conceptualize the covariance just mentioned or to conceptualize the variation of $V(y)$ with v, i.e., the *marginal risk* $dV(y)/dv$. Since we find the latter simpler to interpret, our remarks are biased accordingly.

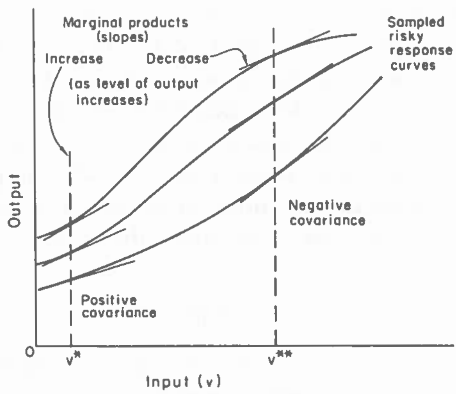


FIG. 6.1. Sketch suggestive of positive and negative cases of covariance between output and marginal product.

Returning to equation (6.10) and ignoring for the moment the term involving REDQ and $dV(y)/dv$,

$$p_v = p_y dE(y)/dv \tag{6.13}$$

i.e., optimality is achieved when the factor price equals the value of the marginal expected product. This relationship is depicted by the intersection of the two unbroken lines in Figure 6.2 and clearly has a strong analogy with the riskless criterion of achieving optimality in resource use by equating marginal value product with factor price. We can ignore the term dropped in (6.13) only if either REDQ = 0 or $dV(y)/dv = 0$; i.e., if either the decision maker is neutral toward risk or if the level of risk (i.e., variance of output) is not influenced by the decision variable (assuming still that output price is not risky). Thus if a process has risky output but with a constant level of risk, decision making on resource use will effec-

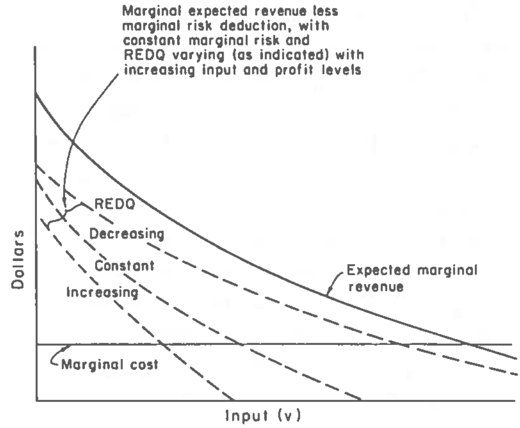


FIG. 6.2. Schematic representation of marginal quantities in the single-input risky production situation corresponding to equation (6.10).

tively be the same as under certainty; the only difference is the use of $E(y)$ instead of y.

In general, however, REDQ and the marginal risk cannot be ignored. To this extent the picture that emerges will differ from that described by equation (6.13). What happens depends on the signs and the rates of change of REDQ and marginal risk as v changes. As we have already discussed in introducing $\mathrm{Cov}(y, dy/dv)$, the nature of marginal risk is not very amenable to naked contemplation and must be assessed empirically. Beyond noting that it typically will be positive, we cannot be sure if it increases or decreases with v.

We can make rather better progress with the REDQ term, especially if we are prepared to make some arbitrary assumptions about the preferences $U = U(x)$. For instance, with risk aversion we have $\partial U/\partial V(x) < 0$ which, combined with the usual requirement of $\partial U/\partial E(x) > 0$, implies REDQ > 0 for the "normal" case of aversion to risk. In the particular case of risk-averse quadratic preference

$$U = E(x) + b[E(x)^2 + V(x)] \qquad b < 0$$

so that taking partial derivatives and substituting into equation (6.8) gives

$$\mathrm{REDQ}_{\mathrm{quad}} = -b/[1 + 2bE(x)] \tag{6.14}$$

which is positive within the relevant range of $E(x) < -1/2b$. As noted in the context of absolute risk aversion in Section 4.4, REDQ for the quadratic increases with $E(x)$ and also with v if $dE(x)/dv > 0$, which is rather counterintuitive. With constant marginal risk and quadratic preference, the marginal risk deduction $\mathrm{REDQ}[p_y^2 dV(y)/dv]$ is indeed a deduction and increases with v. This is depicted by the lower broken curve in Figure 6.2 where the broken lines represent the right side of (6.10), i.e., marginal expected revenue corrected for risk under various preference assumptions. The intersection of these broken curves with the marginal cost line gives optimal levels of input corresponding to the solution of (6.10) for each preference assumption. The lowest of the broken curves in Figure 6.2 depicts the quadratic case and progressively becomes more distant from the unbroken marginal expected revenue line as input increases.

Such counterintuitive results are avoided by using alternative preference functions such as the logarithmic function that we have earlier shown to be decreasingly risk averse. From Section 4.6 the (E, V) approximation of the log function is given by

$$U = \log_e[W_0 + E(x)] - 0.5V(x)/[W_0 + E(x)]^2$$

and the corresponding risk evaluation by

$$\mathrm{REDQ}_{\log} = 0.5[W_0 + E(x)]/\{[W_0 + E(x)]^2 + V(x)\} \qquad (6.15)$$

which is positive and diminishes with respect to $E(x)$ and also v if $dE(x)/dv > 0$, providing that the coefficient of variation, $[V(x)]^{0.5}/[W_0 + E(x)]$, is less than one. The highest broken line in Figure 6.2 depicts the risk-corrected marginal expected revenue curve under logarithmic preference and constant marginal risk.

The second-order condition for a solution to be a maximal solution is that $d^2U/dv^2 < 0$. We have deliberately avoided discussion of this question; however, it is given extensive attention by Magnusson (1969). The conditions are not only complex to derive but are also rather obscure in the absence of several very specific assumptions and, we believe, are not too important in practice for analysts who are aware of the need to check that indeed maximal and not minimal solutions are found. Moreover, since any serious empirical work will probably resort to numerical exploration of conditional expected utility surfaces, the second-order conditions will automatically be taken into account. Likewise, the question of boundary conditions (i.e., constraining optimal inputs to the nonnegative and "sensible" range) will probably not cause any difficulties in empirical work, so we have ignored these too.

We have thus far not specified on what bases v, y, x, etc., are measured. Implicit in writing $U = U(x)$ is the notion that variations and achievements of x are the only elements of the producer's decision problem that impinge on his preferences and satisfactions. In most cases of dealing with the farm firm, x should be some aggregate measure of net financial gain or net change in equity. It follows that if x denotes total net change, account must be taken of the total quantities of input and output and not merely of quantities expressed on a per machine, per man, per hectare, etc., basis.

Many empirical relationships relevant to risky production analysis will be expressed on some alternative basis; e.g., crop response and variable inputs may be expressed in mass per unit area. In such cases it is necessary to aggregate the process to find the impact on the firm by introducing the size of the enterprise. For instance, if y and v are in kg/ha, F is in \$/ha, and the enterprise consists of A ha of the process presently being discussed, the profit function should be written as $x = A(p_y y - p_v v - F)$, where p_y and p_v are prices in \$/kg. It follows that the right sides of equations (6.2) and (6.3) for $E(x)$ and $V(x)$ should be multiplied by A and A^2 respectively, so that the first-order condition of (6.10) becomes

$$Ap_v = Ap_y dE(y)/dv - \text{REDQ}[(Ap_y)^2 dV(y)/dv] \tag{6.16}$$

Our difficulties are not yet completely resolved since we have so far shirked detailed discussion of the fixed cost F, which is assumed completely independent of output. It is apparent from equations (6.8) and (6.16) that fixed costs may exert an influence in the determination of optimal v only through their effect of reducing $E(x)$ and in turn, with decreasing risk aversion, of increasing REDQ. These fixed costs are the current manifestation of sunk costs. In a theoretical sense F is therefore the value of an annuity that could have been acquired had the sunk costs not been sunk. Abstracting from the uncertainties inherent in finding and realizing such an annuity, our difficulty devolves to identifying what costs are "sunk" with respect to the decisions being considered. Two possible approaches are: (1) Given that this process is to be operated, what is the best level of v? (2) Given a conditional optimal level of v, should this or some other process be run? The costs regarded as fixed and sunk may well differ between these cases.

Consider the following simple problem in risky response analysis. A farmer, whose relevant risk preferences for profits, x are approximately encoded in $U = x - (3)10^{-5}x^2$, is intending to grow 100 ha of maize. The empirical relations specifying the risk he believes he faces are depicted in Figure 6.3 and given by

$$E(y) = 5000 + 32.5N - 0.125N^2 \tag{6.17}$$

$$V(y) = 700{,}000 + 10{,}000N \tag{6.18}$$

where E and V denote the mean and variance operators respectively and y = kg/ha of maize priced at \$0.04/kg and N = kg/ha of nitrogen priced

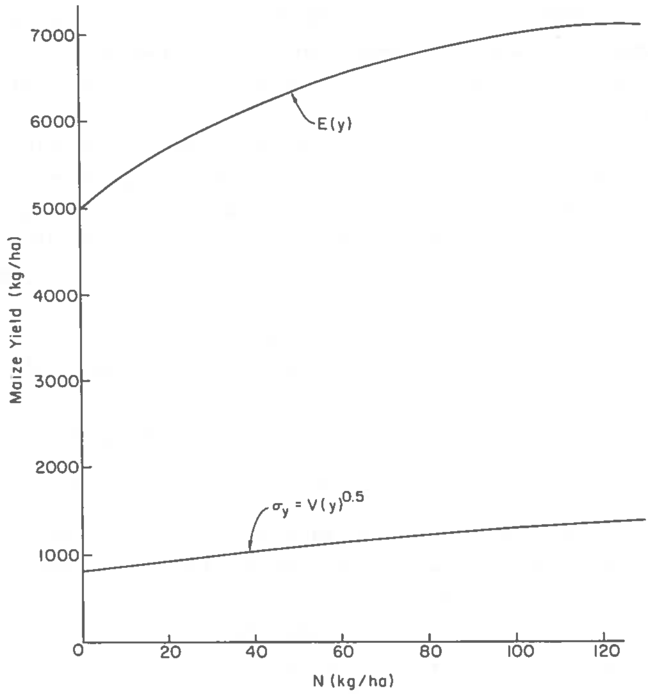


FIG. 6.3. Graphical depiction of empirical maize-response data of equations (6.17) and (6.18).

at \$0.30/kg. These empirical equations correspond to $g(v)$ and $h(v)$ of equations (6.2) and (6.3) respectively and are regression estimates based on eight years of observations from a maize-nitrogen experiment. They exemplify what is called a "gross relationship" in Section 6.3. Assuming that fixed costs amount to \$100/ha, how much nitrogen should our farmer use per hectare?

Solution is approached most simply by substituting into equation (6.16) the given parameters, the derivatives of (6.17) and (6.18), and the result of (6.14). This gives

$$(100)(0.3) = (100)(0.04)(32.5 - 0.25N) - \{(3)(10^{-5})/[1 - (2)(3)(10^{-5})E(x)]\}[(100)(0.04)]^2(10{,}000) \quad (6.19)$$

Even for this simple case of constant marginal risk, algebraic solution of (6.19) to find the optimal level of N is not trivial because $E(x)$ appears in the denominator of the curly bracket term and, since $E(x)$ is a quadratic function of N, (6.19) is cubic in N. However, graphical solution of the style

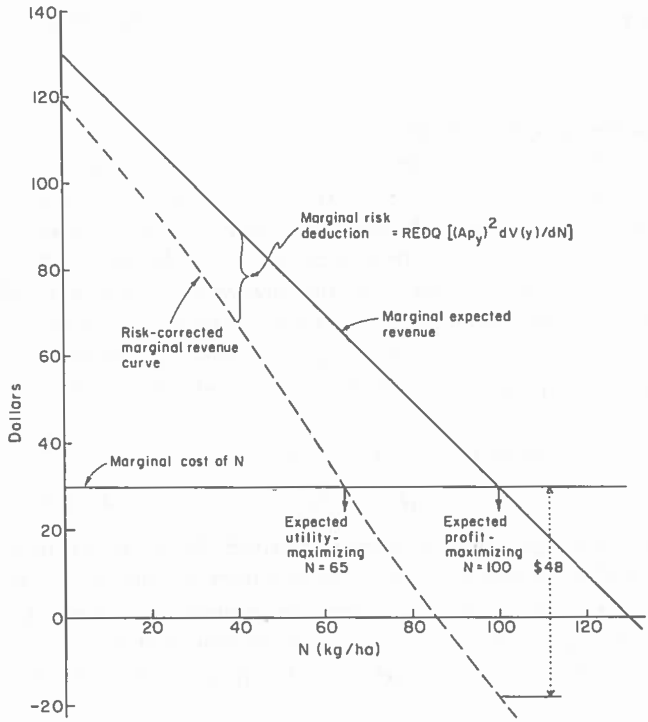


FIG. 6.4. Graphical solution of the maize-fertilizer problem of equation (6.19).

sketched in Figure 6.2 is simple and is illustrated in Figure 6.4. This indicates an optimal N of about 65 kg/ha, which is about two-thirds of the 100 kg/ha rate that maximizes expected profit. Obviously, in this particular case risk constitutes a substantial friction relative to resource use in the risk-free situation. With the assumed level of risk aversion the marginal risk deduction ranges from about \$12/ha at $N = 0$ to about \$48/ha at $N = 100$. Rather less dramatic risk deductions are implied by some alternative assumptions about preference. For instance, assuming that $U = \log_e (100{,}000 + x)$ produces an optimal level of N of about 98 kg/ha that is only trivially different from the risk-indifferent solution.

6.2 SOME EXTENSIONS OF THE SIMPLEST CASE

Our introductory model of production under risk can be extended in several ways, either singly or in combination. For simplicity we choose to survey the possible extensions by a quick listing of one-at-a-time extensions.

Beyond the Mean and Variance

In using the moment method to reach pragmatic solutions of our more general problem, we cut short the Taylor series specification of the utility function so that it included only the mean and variance of profit. This can be rationalized by assuming that derivatives of the preference function beyond the second are very small or that higher order moments of the distribution of x are relatively small. In reality, however, skewness or the effect of the third moment about the mean will often be important to a decision maker. In our present context the effect of skewness can be appraised as follows.

From the Taylor series of equation (4.27)

$$U = U[E(x)] + U_2[E(x)]V(x)/2 + U_3[E(x)]M_3(x)/6 \qquad (6.20)$$

if we ignore terms involving derivatives beyond the third. To proceed with our risky response analysis, we need to supplement the empirical relationships $E(y) = g(v)$ and $V(y) = h(v)$ with the empirical relationship $M_3(y) = t(v)$. To maintain simplicity, we will again assume that p_y is certain. Then, analogously to (6.5), the first-order condition for maximum utility is defined by $dU/dv = 0$ where

$$\begin{aligned} dU/dv = {} & [\partial U/\partial E(x)][dE(x)/dv] + [\partial U/\partial V(x)][dV(x)/dv] \\ & + [\partial U/\partial M_3(x)][dM_3(x)/dv] \end{aligned} \qquad (6.21)$$

or

$$\begin{aligned} 0 = {} & dE(x)/dv - \text{REDQ}\,dV(x)/dv \\ & + \{[\partial U/\partial M_3(x)]/[\partial U/\partial E(x)]\}dM_3(x)/dv \end{aligned} \qquad (6.22)$$

Analogously to the earlier definition of REDQ, we can define the *marginal skewness quotient* (MSQ) as

$$\begin{aligned} \text{MSQ} &= [dE(x)/dM_3(x)]_{U,V(x)} \\ &= -[\partial U/\partial M_3(x)]/[\partial U/\partial E(x)] \end{aligned} \qquad (6.23)$$

and rewrite (6.22) with substitution from the profit equation (6.1) as

$$\begin{aligned} p_v = {} & p_y dE(y)/dv - \text{REDQ}[p_y^2 dV(y)/dv] \\ & - \text{MSQ}[p_y^3 dM_3(y)/dv] \end{aligned} \qquad (6.24)$$

which is identical to (6.10) except for the subtraction from the right side of a marginal skewness allowance. Though it is essentially an empirical question, it seems intuitively reasonable that yield will generally tend to become less negatively skewed with increasing v. This implies $dM_3(y)/dv > 0$. Likewise, utility may be expected to increase with $M_3(x)$ so that, from (6.23), MSQ would be negative. On these grounds, therefore, we may expect the effect of accounting for skewness to be opposite in direction from that earlier speculated as typical in the case of accounting for variance effects. The greater the degree of positive skewness or the less the degree of negative skewness, i.e., the greater is $M_3(x)$ or the less the tail of the probability distribution of x extends to the left, the more attractive will be the production process and the greater the optimal level of input, other things being equal.

Several Decision Variables

Our earlier assumption of dealing with only a single variable factor of production served to keep the algebra relatively uncomplicated and to make the effects of risk relatively transparent. Exactly analogous results obtain for more general production relationships involving k decision variables $v_1, \ldots, v_i, \ldots, v_k$. Rather than develop these in the same detail as employed above, we choose merely to present the result. Thus corresponding to equation (6.10), for the ith decision variable we have

$$p_i = p_y \partial E(y)/\partial v_i - \text{REDQ}[p_y^2 \partial V(y)/\partial v_i] \qquad i = 1, \ldots, k \qquad (6.25)$$

which is identical to (6.10) except for the addition of subscripts and the inclusion of partial derivatives.

Solution for the optimal vector of v_i involves the simultaneous solution of these k invariably nonlinear equations and may not prove to be straightforward, perhaps involving several boundary solutions. Similarly, the second-order conditions for a maximum solution are rather more complex than for the single-factor case. In practice, however, the easiest method of seeking a solution will usually be to follow a systematic numerical exploration of the expected utility surface (in $k + 1$ dimensional space); such a procedure will automatically take care of boundary and second-order conditions.

Nonindependence of Output and Its Price

While the output price p_y was being treated as risky in Section 6.1, we made the simplifying assumption that the distributions of p_y and y were stochastically independent. This was defended as being realistic since under the competitive conditions that characterize farming, and ignoring such subtleties as the influence of poor seasons on price premiums for quality, there will be little dependence between the yield and price that an individual farmer experiences. This would not be true for situations where all farming areas of a country tend to experience the same climatic conditions at the same time, so that droughts, hurricanes, etc., tend to influence the whole nation rather than isolated areas. Accordingly, it is worth looking briefly at the effect of nonindependence in our simple case. Since we are currently defining utility only as a function of mean and variance, the dependence between p_y and y is captured through the simple correlation between them denoted by the coefficient ρ. If we were to be concerned with skewness as well, it would be necessary to consider the co-third moments also.

On dropping the assumption of independence but assuming that p_y and y are normally distributed, equations (6.2) and (6.3) become respectively

$$E(x) = E(p_y)E(y) + \mathrm{Cov}(p_y, y) - p_v v - F \tag{6.26}$$

$$\begin{aligned} V(x) = {} & [E(p_y)]^2 V(y) + [E(y)]^2 V(p_y) + V(p_y)V(y) \\ & + \rho^2\{\rho V(p_y)\, V(y) + 2E(p_y)E(y)[V(p_y)V(y)]^{0.5}\} \end{aligned} \tag{6.27}$$

It seems reasonable to assume that the distribution and hence the moments of p_y are not influenced in any way by the level of the decision variable v. Even with this assumption, substitution of the derivatives of (6.26) and (6.27) into (6.6) yields the cumbersome expression:

$$\begin{aligned} p_v = {} & E(p_y)dE(y)/dv + \mathrm{Cov}(p_y, dy/dv) \\ & - \mathrm{REDQ}\{[E(p_y)]^2 dV(y)/dv + 2E(y)V(p_y)dE(y)/dv \\ & + V(p_y)dV(y)/dv + \rho^3 V(p_y)dV(y)/dv \\ & + 2\rho^2 E(p_y)[V(p_y)V(y)]^{0.5} dE(y)/dv \\ & + \rho^2 E(p_y)E(y)[V(p_y)]^{0.5}[dV(y)/dv]/[V(y)]^{0.5}\} \end{aligned} \tag{6.28}$$

The influence of the correlation coefficient on the optimal conditions is quite pervasive. At the same time it is virtually impossible to comment on the general implications of ρ for decision making about v without some very specific assumptions about the signs and magnitudes of the several terms in equation (6.28). Beyond noting the complexity of the expression, we can distill one interesting result from (6.28) by assuming risk neutrality; i.e., REDQ = 0. Even risk-indifferent decision makers should not simply set $p_v = E(p_y)dE(y)/dv$ to find their optimal v. Under nonindependence they need to adjust expected price times marginal expected product by the covariance term $\mathrm{Cov}(p_y, dy/dv)$. Once again it is an empirical question as to whether this is a positive or negative correction.

Risky Factor Prices

A further way of generalizing our simple model is to allow for risk in factor price. We have not done this in the belief that agricultural decision making typically proceeds with fairly certain knowledge of the prices to be paid for factors. Turnovsky (1969) has shown that under rather general conditions, demand for a factor responds negatively to increases (1) in its expected price, (2) in the variance of its price, and (3) in the decision maker's aversion to risk. Such extension of the analysis to risky factor prices is also reviewed briefly by Magnusson (1969).

Response Efficiency over Time

All agricultural response processes necessarily involve the effluxion of time. Indeed, it is the passing of time that permits the intrusion of uncertainty into agricultural production. The models of production we have used in this chapter have abstracted from the influence of time and treated response as if it was instantaneous. Dillon (1976) has demonstrated that in a riskless world the influence of time on response efficiency can be both pervasive and complex, and that like risk aversion it constitutes a friction to resource use. Complexity is further increased if the dimension of risk is added to that of time because, among other complexities, a multiperiod preference function is required. Such considerations are discussed in Chapter 8, though not in the context of response theory models.

Several Products

We have avoided discussion of multiproduct production for two good reasons. First, from a somewhat different slant Chapter 7 deals with this topic in considerable detail. Second, to pursue the ideas of Section 6.1 cast in a multifactor multiproduct setting with flexibly defined production relationships is to flirt with algebraic folly that achieves little in the way of nicely transparent results. It is sufficient to note that compared to the riskless multiproduct case, risk aversion implies that optimal resource use will favor less risky products relative to more risky ones and that overall resource use will be decreased. With this generalization we can terminate our discussion of extensions to the simplest case.

6.3 EMPIRICAL ANALYTICS

Our discussion of risky response analysis has not yet tackled the estimation of several types of quantitative relationships we have been presuming to exist. In fact, there have been only a few reported attempts to quantify most of the risky relationships discussed. Most of the response analysis literature (e.g., Heady and Dillon, 1961) is (often implicitly) concerned only with specification of expected response functions that give $E(y)$ as some function of the decision variables $v_1, v_2, \ldots, v_k$. A few people (Colyer, 1969; Doll, 1972; Fuller, 1965; McArthur and Dillon, 1971) have attempted to estimate $V(y) = h(v_1, v_2, \ldots, v_k)$, and some have also attempted specification of $M_3(y) = t(v_1, v_2, \ldots, v_k)$ (Anderson, 1973; Day, 1965). However, little success has been reported with the most general risky specification $D(y) = f(v)$ where $D(y)$ denotes every aspect (e.g., fractiles, parameters, etc.) of the probability distribution of y, although the efforts of Anderson (1974b) and de Janvry (1972) marginally fall into this category.

Analytical Approach

We can identify two broad approaches to specifying the intrusion of risk into productive processes; one is analytical, the other somewhat gross. In the analytical model the production process is conceptualized in terms of a three-way categorization of input variables. The three types of variables are classified as follows:

1. Input variables that are under the producer's control, i.e., decision variables. These are denoted by v_i, $i = 1, 2, \ldots, k$.
2. Input variables that are outside the producer's control and are stochastic and whose values are unknown at the time of decisions about the v_i. These are denoted by s_j, $j = 1, 2, \ldots, r$.
3. Input variables that are outside the producer's control but whose values are known at decision time. These are denoted by q_w, $w = 1, 2, \ldots, m$. They may involve variables that are fixed (such as soil type) and stochastic variables that are realized and known at decision time (such as fallow rainfall).

Using the above notation for input variables, the production function can be specified as

$$y = f(v_1, \ldots, v_i, \ldots, v_k; s_1, \ldots, s_j, \ldots, s_r; q_1, \ldots, q_w, \ldots, q_m) \quad (6.29)$$

where the q_w variables, being given, serve merely to condition the response of y to the v_i and s_j variables. The uncertainty associated with y in this specification arises solely from the influence of the s_j variables.

In terms of equation (6.29) we can think of the analytical approach to risk specification of the production process as a two-step procedure of first describing the effect of the v_i and s_j variables in the process (i.e., measuring the production or response function) and, second, assessing the joint probability distribution associated with the s_j variables. Agricultural processes are typically influenced by so many stochastic variables that it is virtually impossible for such an empirical procedure to be carried out successfully. At best, therefore, the approach might be called pseudoanalytical. A particular danger is the likelihood of omitting some of the variables through empirical difficulties (not to mention the risk of functional misspecification), thereby understating the extent of risk faced by the decision maker.

In his empirical work along the above lines, de Janvry (1972) combined several rainfall variables into a single weather index, which probably had the effect of underestimating the risk involved. Likewise, Byerlee and Anderson (1969) confined their attention to a single uncertain factor of production: the amount of growing-season rainfall. The empirical relationship they determined for wheat response to nitrogen and growing-season rainfall was

$$(Y_N - Y_0) = 1.37N - 0.0836N^2 + 0.0421Ns_1 - 0.00286Ns_1q_1$$

where N is applied nitrogen in kg/ha; s_1 is growing-season rainfall in mm; q_1 is soil nitrogen measured at planting time in ppm; and $Y_N - Y_0$ in kg/ha is yield with applied N, less yield with zero applied N. The stochastic variable s_1 ranged from about 75 to 375 mm (mean about 200 mm) and thereby leads to corresponding variation in response to N. However, through the neglect of other uncertain factors (such as frost incidence) and the imperfection of a blunt measure like growing-season rainfall in describing the intraseason plant-moisture regime, this relationship doubtless underestimates the risk inherent in this crop-fertilizer response process.

An advantage of employing a pseudoanalytical approach of the type outlined above is that valuation of additional information on one or more of the uncertain factors may be possible. The Bayesian framework for determining the EVPI and EVSI is applicable as long as the role of the uncertain factor(s) in the process can be quantified and appropriate priors and likelihoods attached. Byerlee and Anderson (1969) give an example of such evaluation based on linear preference. They show that under risk indifference a necessary condition for information on the s_j variables to have a positive economic value is that they interact with the v_i variables in other than a purely additive manner. However, under a nonneutral attitude to risk, information (i.e., probability modification) bearing on the extent of risk faced will in general (even in the additive case) have a positive value that can be compared with its cost.

Gross Approach

The inescapable difficulties in the pseudoanalytical approach make the alternative "gross" approach relatively attractive when the emphasis is on resource allocation rather than on valuation of information. In the gross procedure, we simply compound all the variation, without identifying the effect of individual sources like the s_j, and quantify the composite probability distributions so that they may be functionally related to the decision variables. The empirical illustration of Section 6.1 concerning risky response of maize to nitrogen fertilizer as per equations (6.17) and (6.18) is an example of the gross approach. Generally it will be quite sufficient to restrict attention to $E(y)$, $V(y)$, and perhaps $M_3(y)$ in order to undertake most of the decision analyses discussed in this chapter. Such moments may come directly from subjective assessments, using the methods outlined in (2.1) and (2.2) of Section 2.3; or they may be based on pertinent historical data accepted as reasonable.

When there are sufficient data, say 10 or more observations, the suggested moment-estimation formulas of R. A. Fisher are applicable; viz., for the set of observations $y_1, y_2, \ldots, y_n$ define

$$Z_j = \sum_{i=1}^{n} y_i^j \qquad j = 1, \ldots, 4$$
$$S_2 = Z_2 - Z_1^2/n$$
$$S_3 = Z_3 - 3Z_1Z_2/n + 2Z_1^3/n^2$$
$$S_4 = Z_4 - 4Z_1Z_3/n + 6Z_1^2Z_2/n^2 - 3Z_1^4/n^3$$

so that

$$E(y) = Z_1/n$$
$$V(y) = S_2/(n-1)$$
$$M_3(y) = nS_3/[(n-1)(n-2)]$$

and, if required,

$$M_4(y) = n[(n+1)S_4 - 3(n-1)S_2^2/n]/[(n-1)(n-2)(n-3)] + 3S_2^2/(n-1)^2$$

These calculations are illustrated in a gross approach to the analysis of some broiler feed-trial response data in Table 6.1.

When data are sparse (say $n \leqq 10$), we are on much less firm ground to make good estimates of moments beyond the mean. The best suggestion we can make is to revert to the sparse-data smoothing rule elaborated in Section 2.4. Application of this rule will yield smooth CDFs from which moments may readily be calculated (as in Anderson, 1973, 1974a). An illustration is provided by the sparse fertilizer-response data plotted in Figure 6.5 and the smooth CDFs sketched through these data.

Moment computation can in general be done most conveniently by first breaking these curves into a large number of discrete elements and then calculating the moments using the probability elements and the associated midpoints of the probability intervals. However, if only the mean and variance are required, it is preferable to employ the much simpler procedure of reading off the respective 0.05, 0.5, and 0.95 fractiles and

TABLE 6.1. Example of Moment Calculations for Gross Approach to Risk Specification of a Broiler-Feed Response Function Based on Feed-Trial Data

A: Observed Weight Gain (kg) per 30 Broilers on 18% Protein Ration

Group of 30 broilers	Kg feed per 30 broilers: 30	60	90	120
1	13.26	24.33	32.85	45.78
2	13.47	23.01	32.13	44.61
3	13.35	24.39	34.26	43.08
4	13.22	24.21	35.46	40.92
5	13.71	24.06	33.15	39.69
6	13.53	24.06	31.02	43.17
7	13.53	24.78	34.65	42.54
8	13.56	23.67	32.58	44.64
9	13.53	23.28	32.64	44.97
10	13.68	23.64	33.21	41.64
11	13.44	23.49	33.96	39.51
12	13.65	23.73	32.70	45.90

B. Calculation of Moments of Weight Gain (G)*

	30	60	90	120
Z_1	...	286.65	398.61	516.45
Z_2	...	6,580.19	13,256.71	22,281.56
Z_3	...	163,769.53	441,412.70	963,628.83
Z_4	...	3,916,904.09	14,715,562.92	41,772,929.13
S_2	...	...	15.878	54.843
S_3	...	...	3.221	−34.446
S_4	...	...	56.071	454.753
$E(G)$	13.50	...	33.22	43.04
$V(G)$	0.0202	...	1.4435	4.986
$M_3(G)$	−0.0007	...	0.3514	−3.758
$M_4(G)$	−0.0009	...	6.2511	45.979

*Statistics not shown are deleted to allow practice computations in Problem 6.3.

using equations (2.1) and (2.2). Applying this method to the 0, 18, and 36 kg P/ha curves of Figure 6.5 gives expected yields of 1380, 1930, and 2070 kg/ha respectively and corresponding standard deviations of 630, 800, and 780 kg/ha.

Once we have such moments conditioned on particular combinations of the decision variables, it is necessary to relate these to the decision variables. Usually this can be done quite successfully via least-squares regression. In other applications, especially those involving stochastic dominance orderings as discussed in Chapter 9, it becomes necessary to relate the whole probability distribution to the conditioning decision variables; and unless rather specific families of distribution are fitted, this

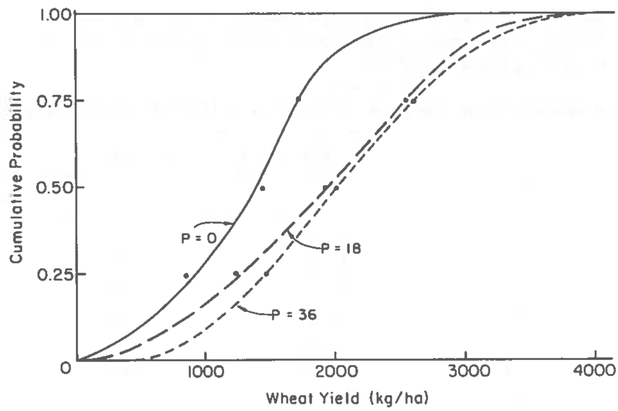


FIG. 6.5. The CDFs for wheat yield at three levels of phosphorus (kg/ha) with nitrogen fixed at 22 kg/ha.

is a demanding task that we will not take up at this point (see Anderson 1974b, Section 3.2).

We can attempt to summarize the considerations involved in a choice between analytical and gross approaches most conveniently by means of a schematic chart. Figure 6.6 depicts the broad options open to an analyst attempting to quantify production risks on the basis of experimental response data. Although seemingly complex, it still represents a simplification of the real situation because partial failures to quantify the effects of subsets of the respective variables are not accommodated in this generalized scheme.

Before parting from the gross approach we should note another of its useful applications. Where pertinent time-series data on both risky outputs and prices are available, the gross method may be applied to the conditional total revenue data (i.e., output times price). This will avoid the necessity of accounting for price and output effects separately, which may introduce additional complexity as indicated by the discussion of non-independence in Section 6.2. Such use of the gross method parallels that given earlier in connection with enterprise net revenues in Section 2.3.

6.4 POLICY IMPLICATIONS

Compared to its members, society or the state can take a long-term view and has a very diverse portfolio of activities. On these grounds it may be argued that as a corporate entity the state should be indifferent to risk relative to the allocation of resources within small sectors of the economy

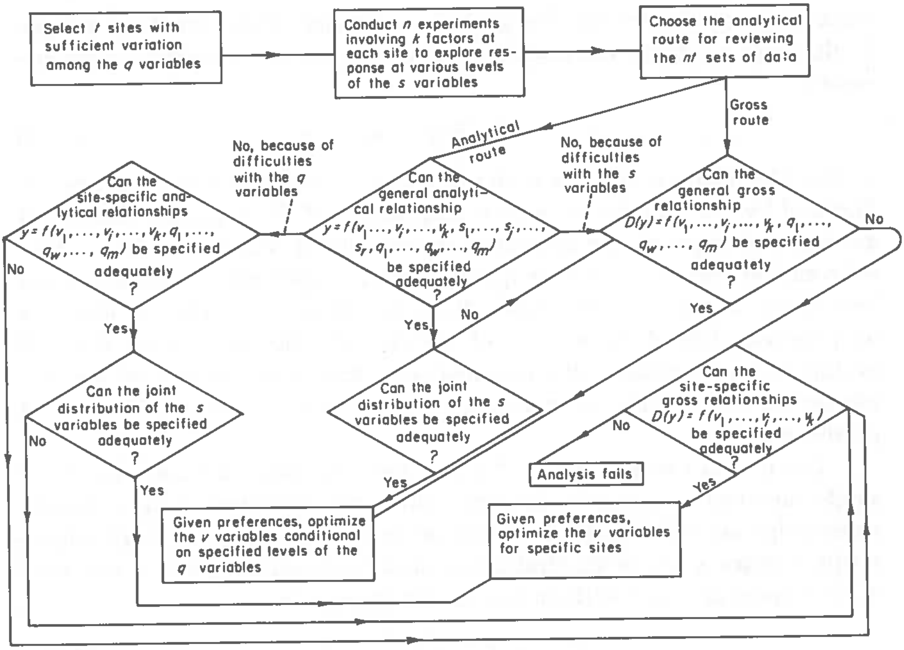


FIG. 6.6. Scheme for empirical efforts in risky response analysis based on experimental data.

(Arrow and Lind, 1970). Assuming this view, the state would wish individuals such as farmers to operate so as to maximize expected profits which (barring support measures like subsidies and tariffs) will correspond to the social optimum. Our discussion throughout this book has emphasized the individuality of efficient utility-maximizing operation. Taking the state's view, it is pertinent to ask what measures might be considered to encourage individuals to make their decisions more in line with the criterion of expected profit maximization. Given the wide-ranging measures already in force in most states to interfere with resource allocation, any such appraisal must be rather tongue in cheek since in reality the situation is inexorably second best.

Product Bounties

The state could attempt to induce resource use at the expected profit-maximizing level by providing either input subsidies or output bounties or some combination of these to producers. First consider a bounty of $\$B/$ unit of output. For simplicity we ignore the fact that both bounties and subsidies will inevitably have unfortunate distributional implications since large producers receive greater assistance than small producers. Under such a scheme the profit equation (6.1) of our simplest model becomes

$$x = (p_y + B)y - p_v v - F \tag{6.30}$$

so that the decision maker's preference $U(x)$ and action v can now be influenced by government through manipulation of B. A purely benevolent government might simply determine the amount of B such that the utility-maximizing level of v now corresponds with the expected profit-maximizing level in the absence of the bounty. Even this simplest of schemes founders on practical difficulties because of the individuality of at least the risk evaluation and probably the marginal risk also, and because of the administrative desirability of maintaining the same bounty rate across all producers.

Discussion will be clearer if a few more symbols are introduced. A single superscript asterisk denotes utility maximization and a double superscript asterisk denotes expected profit maximization. A single superscript b denotes utility maximization under a bounty scheme. The producer's optimal return without the bounty is given by

$$x^* = p_y Y^* - p_v v^* - F \tag{6.31}$$

and his optimal return under a bounty scheme is given by

$$x^b = (p_y + B)y^b - p_v v^b - F \tag{6.32}$$

if product and input prices are uninfluenced by the bounty, as would be the case in the (unlikely) event that demand for y and the supply of v were perfectly elastic. The cost of the scheme to the state for this producer is the random variable By^b which averages $BE(y^b)$. The producer's monetary gain from the bounty is $x^b - x^*$, which will make some positive contribution to the producer's utility if it is positive, as it will be most of the time. For this one producer, the state measures its gross gain by the expectation of $x^b - x^*$ and its net gain G, by deducting the average bounty cost and the average administration cost per producer M as

$$G = E(x^b) - E(x^*) - BE(y^b) - M \tag{6.33}$$

This gain can then be compared with the gains from other forms of state intervention in the economy so as to aid decision making on the allocation of the state's limited budget. It may well be that this type of participation of the state in the risk bearing of farmers is not very attractive when judged in these terms.

An example of such a calculation is readily obtained from McArthur and Dillon's (1971) analysis of utility-maximizing stocking rates of sheep on pasture. Among other results, they give information of the type presented in Figure 6.7. This depicts the optimal decision variable, v sheep/ha, as a function of price per unit of output (wool) p_y. Under their assumptions about the size (405 ha) and technology of a typical sheep farm, the utility-maximizing v^* is 15.25 sheep/ha when p_y = \$0.99/kg and the corresponding v that maximizes expected profit is v^{**} = 15.56 sheep/ha.

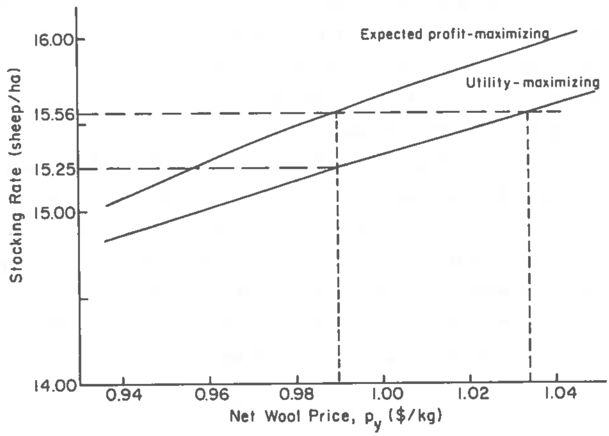


FIG. 6.7. Utility-maximizing and expected profit-maximizing stocking rates in relation to product price.

Figure 6.7 gives us the additional information that, other things being equal, the utility-maximizing rate would rise to this same level of 15.56 if the output price rose to \$1.034/kg. Thus a bounty of \$0.044/kg would encourage resource use at the expected profit-maximizing level in the sense that $v^b = v^{**}$.

A bounty of B = \$0.044/kg can be evaluated under our assumptions by substitution of the pertinent expected values into equations (6.31), (6.32), and (6.33). This indicates an average annual cost to the state of $BE(y^b)$ = \$1040 for the typical producer, which can be compared to the state's gross gain for this producer of $E(x^b) - E(x^*)$ = \$1044. Thus with zero administration costs ($M = 0$) the social net gain G for this farm is only about \$4. Even ignoring the (possibly high) costs of administration, such a scheme would not appear to be very attractive. In practice this small apparent gain would also have to be further discounted because the demand for y and the supply of v will in general not be perfectly elastic.

Input Subsidies

When analysis of risky production decisions is approached via the graphical procedure introduced in Figures 6.2 and 6.4, based on equation (6.10), it is much simpler to appraise possible government participation in individual risk bearing through the use of input subsidies rather than output bounties. For instance, suppose the government is considering whether to encourage the "average" decision maker of Figure 6.4 to increase his use of nitrogen from the utility-maximizing level of $N = 65$ to the expected profit-maximizing level of $N = 100$. To ascertain the required fertilizer subsidy, it is only necessary to determine the input cost level that intersects with the broken line representing the risk-corrected marginal revenue curve at $N = 100$. This occurs at a level of −\$18, which is equivalent to a farm price of −\$0.18/kg of nitrogen. Thus if the government offered a subsidy (\$$S$/kg) of \$0.30 + \$0.18 = \$0.48/kg of nitrogen (which is equivalent to subsidizing fully the marginal risk deduction), the utility-maximizing rate with the subsidy—v^s in notation analogous to equation (6.32)—would be approximately $N = 100$. We must say approximately because in this example the REDQ depends on expected profit $E(x^s)$, which will be adjusted by the subsidy receipts of the individual.

The producer's return is defined analogously to equation (6.32) by

$$x^s = p_y y^s - (p_v - S)v^s - F \tag{6.34}$$

and the state makes subsidy payments of Sv^s per producer for an expected societal gain of

$$G = E(x^s) - E(x^*) - Sv^s - M \tag{6.35}$$

where costs of administering the subsidy payments are denoted by M. Such administrative costs are likely to be lower than bounty administration costs because payments can be made through relatively few input manufacturers. We can see something of the likely magnitudes of such gains by considering the maize-fertilizer example of Section 6.1. The subsidy payment on the assumed 100 ha of crop would amount to \$4800 and the corresponding "gain" to the state as defined by (6.35) would amount to \$650—again under the simplifying assumptions of zero administrative cost and perfectly elastic supply of nitrogen fertilizer and demand for maize. While this is a rather higher apparent return on the state outlay, it may still be much lower than for other opportunities open to the state.

Other Policy Measures

The complexity of risk analysis doubtless explains why such provisions as income averaging for taxation purposes and price and income stabilization schemes provided by some governments to assist in the risk bearing of farmers are somewhat blunt in their effects. More precise tailoring of measures to individual requirements would probably be an administrative nightmare. However, it is important to emphasize that risk can indeed have considerable impact at the individual producer level. The practical difficulties of quantifying this impact across many producers should not obscure the fact that most people are not indifferent to risk. Policies formulated on the naive basis that risk can simply be ignored may well prove inappropriate when risk does play an important role. Much more research is required to identify those situations where the assessment of risk is desirable, the tracking of its implications is worthwhile, and pertinent policy measures are feasible.

PROBLEMS

6.1. (a) Derive expressions for the REDQ for the following preference functions:

$$U = x - (2)10^{-5}x^2$$
$$U = \log_e(100{,}000 + x)$$
$$U = (100{,}000 + x)^{0.5}$$
$$U = 1 - \exp[-(100{,}000 + x)/5000]$$

(b) What alternative results are obtained in the numerical example of Section 6.1 when each of these preference functions is used?

6.2. Discuss the implications for marginal risk and marginal skewness effects of the following empirical relationships extracted from Anderson (1973):

$$E(y) = 17 + 0.15v - 0.0014v^2$$
$$V(y) = 36 + 2.6v - 0.025v^2$$
$$\alpha_3(y) = -0.6 + 0.004v$$

where $M_3(y) = \alpha_3(y)[V(y)]^{1.5}$

6.3. Consider the broiler-feed response data of Table 6.1.
(a) Compute the missing statistics.
(b) Sketch $E(G)$, $V(G)$, $M_3(G)$, and $M_4(G)$ as functions of feed consumption. Comment on the regularity or otherwise of these relationships and on the general implications for risk-averse optimization of feeding.

6.4. Take a production process with which you are familiar and list the factors that make it risky. Review the feasibility of (a) quantifying their influence in the process and (b) specifying the distributions of the risky factors, including any statistical dependencies.

6.5. Consider a particular risky farming enterprise of your choice. Argue in an analytical way a case for or against government intervention in the farmer's role of risk bearing.

6.6. Consider the following response data relating to production of nonirrigated wheat and its response to nitrogen. All variables are in kg/ha.

$$E(y) = 1800 + 15N - 0.08N^2$$
$$V(y) = 280{,}000 + 20{,}000N - 100N^2$$

Assume that $p_y = \$0.07/\text{kg}$, $p_N = \$0.30/\text{kg}$, that fixed costs amount to \$50/ha, and that 100 ha of wheat are involved in the decision.

(a) Does this process seem more or less risky than that described in the example of Section 6.1?

(b) If the farmer's preferences are described approximately by $U = x - (3)10^{-5}x^2$, what rate of N should he use?

(c) In what way does a higher level of fixed costs influence the utility-maximizing rate of N?

6.7. Consider again the empirical example of Section 6.1 and add the assumption that the maize yield distributions are normal and thus completely determined by the mean and variance equations. It follows also that profit from the process is normally distributed. Compute and compare the probabilities of achieving negative profits (i.e., a loss) at the utility-maximizing and expected profit-maximizing rates of fertilizer.

6.8. Extend the analysis of Section 6.1 under the assumption that output price is certain but input price is risky. Note also the practical conditions under which such a situation might be encountered.

6.9. Explore the implications for risky decisions involving two fertilizers, given the following relationships for nonirrigated wheat production:

$$E(y) = 1200 + 9.2N + 42P - 0.076N^2 - 0.69P^2 + 0.15NP$$
$$[V(y)]^{0.5} = 400 + 8.5N + 18P - 0.070N^2 - 0.49P^2 + 0.18NP$$

All variables are measured in kg/ha and the factor additional to that considered in Problem 6.6 is elemental phosphorus which is priced at $p_P = \$0.50/\text{kg}$.

6.10. McArthur and Dillon (1971) present the following equations for mean and variance of net income of a wool producer:

$$E(x) = A\{S[(M - dS)R - C] - F\}$$
$$V(x) = (dAS^2R)^2\sigma_c^2$$

where
A = farm area
C = variable cost per sheep
F = fixed cost per unit area
M = maximum wool cut per sheep
S = stocking rate of sheep per unit area
d = reduction in wool cut per sheep per unit increase in S
R = wool price per unit weight of cut
σ_c^2 = variance of a climatic index

(a) How might such equations be employed to assist producers in selecting a best stocking rate?

(b) Deduce the effect on optimal stocking rate of one-at-a-time increases in C, F, M, and R.

6.11. Add to the empirical example of Section 6.1 the assumptions that maize price is uncertain, $E(p_y) = 0.04$, $V(p_y) = 0.0001$, and that price and yield are normally and independently distributed.

(a) What is the optimal rate of N under the quadratic preference assumption?
(b) Using Monte Carlo sampling, determine empirically the CDF for returns at the optimal level of N.

6.12. Repeat Problem 6.11 (a) with the additional assumption that there is a negative correlation between p_y and y of -0.3.

REFERENCES AND SELECTED FURTHER READING

Anderson, J. R. 1973. Sparse data, climatic variability, and yield uncertainty in response analysis. *Am. J. Agric. Econ.* 55(1): 77–82.

Elaborates empirical estimation of moment-response functions when only a few experimental observations are available.

———. 1974a. Sparse data, estimational reliability, and risk-efficient decisions *Am. J. Agric. Econ.* 56(3): 564–72.

Explores the reliability of "gross" descriptions of risk based on limited time-series data and finds that even as few as 3 or 4 observations can be very useful.

———. 1974b. Risk efficiency in the interpretation of agricultural production research. *Rev. Mktg. Agric. Econ.* 42(3): 131–84.

As part of a broader study of the use of stochastic dominance orderings in decision analysis (see Chapter 9), this study examines risky response analysis without specific utilities and REDQs.

———. 1976. Methods and programs for analysis of risky gross margins. Agric. Econ. Bus. Mgmt. Misc. Pub. 5. Armidale, Australia: Univ. New England.

Arrow, K. J., and R. C. Lind. 1970. Uncertainty and the evaluation of public investment decisions. *Am. Econ. Rev.* 60(3): 364–78.

Reviews various risk situations in which the state should act as if it were risk indifferent. This is particularly the case when the benefits of state investment are widely dispersed among many members of society.

Barnett, V. 1975. Probability plotting methods and order statistics. *Appl. Statist.* 24(1): 95–108.

A comprehensive review and critique of methods such as the sparse data rule for estimating cumulative probability distributions.

Batra, R. N., and A. Ullah. 1974. Competitive firm and the theory of input demand under price uncertainty. *J. Polit. Econ.* 82(3): 537–48.

In the tradition of Sandmo (1971), shows that a marginal increase in uncertainty about product price stimulates a decline in the output and the demand for inputs by the firm.

Blandford, D., and M. Currie. 1975. Price uncertainty—The case for government intervention. *J. Agric. Econ.* 26(1): 37–51.

Modeling price uncertainty in the fashion of Leland (1972) and Sandmo (1971), this article considers the effect of a guaranteed producer-price scheme on the welfare of producers, consumers, and taxpayers.

Bohm, P. 1974. On the effects of policy costs. *Swedish J. Econ.* 76(1): 104–16.

Presents a framework for analyzing the role of different kinds of policy costs for optimal policy and contends that administrative costs are often substantial

enough to influence the extent and nature of best policy solutions. Such considerations would be pertinent to analysis of policies for risk bearing in agricultural production.

Byerlee, D. R., and J. R. Anderson. 1969. Value of predictors of uncontrolled factors in response functions. *Australian J. Agric. Econ.* 13(2): 118–27.

Chisholm, A. H. 1971. A comparison of income averaging procedures for income tax purposes. *Australian J. Agric. Econ.* 15(1): 36–50.

Compares several income averaging procedures on the basis of several performance criteria, one of which is the stability of after-tax income. Some schemes, including those often used, may actually accentuate the effects of low-income years.

Colyer, D. 1969. Fertilization strategy under uncertainty. *Canadian J. Agric. Econ.* 17(3): 144–49.

Gives an example of the "gross" approach to quantifying mean and variance of crop response from a time series of experimental data.

Day, R. H. 1965. Probability distributions of field crop yields. *J. Fm. Econ.* 47(3): 713–41.

A pioneering study of appraising risk in agriculture through the fitting of (Pearson-type, nonnormal) distributions to time-series experimental data.

Day, R. H., D. J. Aigner, and K. R. Smith. 1971. Safety margins and profit maximization in the theory of the firm. *J. Polit. Econ.* 79(6): 1293–301.

de Janvry, A. 1972. Optimal levels of fertilization under risk: The potential for corn and wheat fertilization under alternative price policies in Argentina. *Am. J. Agric. Econ.* 54(1): 1–10.

Demsetz, H. 1969. Information and efficiency: Another viewpoint. *J. Law Econ.* 12(1): 1–22.

Presents a contrary view to Arrow and Lind (1970).

Dillon, J. L. 1976. *The Analysis of Response in Crop and Livestock Production* 2nd ed. Oxford: Pergamon.

This second edition, in contrast to the first, presents a utility-based analysis of risky response, including analysis of time and risk effects together, and constrained multiple response.

Doll, J. P. 1972. A comparison of annual versus average optima for fertilizer experiments. *Am. J. Agric. Econ.* 54(2): 226–33.

Fuller, W. A. 1965. Stochastic fertilizer production functions for continuous corn. *J. Fm. Econ.* 47(1): 105–19.

Heady, E. O., and J. L. Dillon. 1961. *Agricultural Production Functions.* Ames: Iowa State Univ. Press.

Hiebert, L. D. 1974. Risk, learning, and the adoption of fertilizer responsive seed varieties. *Am. J. Agric. Econ.* 56(4): 764–68.

In a model of production under risk, emphasizes the effects of learning (changed beliefs or probability distributions) on risk-optimal decisions about input levels.

Horowitz, I. 1970. *Decision Making and the Theory of the Firm.* New York: Holt, Rinehart and Winston.

Chapters 12 and 13 contain a wide-ranging discussion of the theory of production under risk concentrated on cost-revenue functions and conditions of competition.

James, E. 1975. A note on uncertainty and the evaluation of public investment decisions. *Am. Econ. Rev.* 65(1): 200–205.

Presents more detailed analysis and clarification of Arrow and Lind (1970).

Just, R. E., and R. D. Pope. 1976. On the relationship of input decisions and risk. Dept. Agric. Econ. Working Paper, Univ. Calif., Berkeley.

Summarizes the possibilities for analytical incorporation of stochastic variables in production functions.

Leland, H. E. 1972. Theory of the firm facing uncertain demand. *Am. Econ. Rev.* 62(3): 278–91.

Lin, W., H. F. Carman, C. V. Moore, and G. W. Dean. 1974. Producer response to income taxes: An empirical test within a risk framework. *Natl. Tax J.* 27(2): 183–95.

Examines the impact of progressive income tax on the production and risk-taking behavior of some large commercial farms.

McArthur, I. D., and J. L. Dillon. 1971. Risk, utility and stocking rate. *Australian J. Agric. Econ.* 15(1): 20–35.

Demonstrates parametrically the influence of variations in the components of decision analysis on the optimal use of factors of production.

McCall, J. J. 1971. Probabilistic microeconomics. *Bell J. Econ. Mgmt. Sci.* 2(2): 403–33.

A comprehensive survey of modern developments in the marriage of microeconomics (including production economics), probability calculus, and expected utility analysis.

Magnusson, G. 1969. *Production under Risk: A Theoretical Study.* Uppsala: Almqvist and Wiksells.

The seminal and pioneer work on extending neoclassical production theory to deal with risk. A comprehensive treatment of theory but weak on empirical aspects.

Montaño, C. B., and R. Barker. 1974. Economic returns from fertilizer application in tropical rice in relation to solar energy level. *Philippine Econ. J.* 13(1): 27–40.

An example of the analytical approach to risk specification.

Nowshirvani, V. F. 1971. Land allocation under uncertainty in subsistence agriculture. *Oxford Econ. Pap.* 23(3): 445–55.

Employs the moment method to analyze crop decisions in subsistence agriculture and discovers that accounting for risk leads to ambiguous results on the likely price responsiveness of farmers.

Oury, B. 1966. *A Production Model for Wheat and Feedgrains in France (1946–1961).* Amsterdam: North-Holland.

Discusses the conceptual and empirical difficulties of accommodating climatic risk in econometric studies of crop production.

Ray, P. K. 1967. *Agricultural Insurance: Principles and Organization.* Oxford: Pergamon.

The standard text on agricultural insurance.

Roumasset, J. 1974. Estimating the risk of alternative techniques: Nitrogenous fertilization of rice in the Philippines. *Rev. Mktg. Agric. Econ.* 42(4): 257–94.

Blends an analytical approach to specifying the impact of risk in production with the use of safety-first criteria and de Janvry's (1972) concept of risk of fertilizer use as the probability that the internal rate of return to fertilizer is less than or equal to zero. Finds that using nitrogen on rice is not as risky as some have speculated.

———. 1976. *Rice and Risk: Decision Making among Low-Income Farmers*. Amsterdam: North-Holland.

A more general setting of his 1974 article that emphasizes safety-first rules in lexicographic orderings.

Sadan, E. 1970. The investment behavior of a firm operating under risk. *Am. J. Agric. Econ.* 52(4): 494–504.

Looks at risky resource allocation via the moment method but restricted to an underlying Cobb-Douglas production function.

Sandmo, A. 1971. On the theory of the competitive firm under price uncertainty. *Am. Econ. Rev.* 61(1): 65–73.

An expected utility analysis of behavior of competitive firms facing risky demand for their products—general in the sense that it does not resort to the moment method.

Tisdell, C. 1973. Certainty equivalence and bias in the management of production. *Rev. Mktg. Agric. Econ.* 41(4): 166–86.

Explores a model of risky production centered on the maximization of expected profits. Examines the conditions governing Theil's concept of certainty equivalence wherein optimization of a surrogate profit function that includes stochastic factors entered at their expected values leads to the same solution for the controlled variables as that which optimizes the expected value of the profit function.

Turnovsky, S. J. 1969. The behaviour of a competitive firm with uncertainty in factor markets. *New Zealand Econ. Pap.* 3(1): 52–58.

Wolgin, J. M. 1975. Resource allocation and risk: A case study of smallholder agriculture in Kenya. *Am. J. Agric. Econ.* 57(4): 622–30.

Takes account of risk attitudes in testing the efficiency of farmers' resource use.

CHAPTER SEVEN
WHOLE-FARM PLANNING UNDER RISK

Decisions are never made in a vacuum. They always relate to some context and lead on to further decisions. This raises two aspects of risky choice that we have so far largely ignored. First, the breadth of the context in which a decision is placed and, second, the time-sequential and time-preference aspects of choice. We will leave the question of time-preference considerations to Chapter 8, where we outline decision analysis procedures for risky project appraisal. Here our interest is in short-term situations where a choice has to be made as to the allocation of resources to some combination of risky prospects from among the set of all available risky prospects. Our emphasis will be on whole-farm planning in its annual context, but as noted below, the principles of analysis are appropriate to any risky choice situation involving a single accounting period where the problem is to choose a portfolio (i.e., a mixture) of risky prospects from among some available set of possibilities.

The topic of portfolio choice, especially under conditions of risk, is an extremely broad one. In one guise or another, but particularly under the headings of linear and nonlinear mathematical programming, it has been a focus of interest in recent years. This has led to the development of a variety of mathematical algorithms for portfolio choice. To attempt to treat these approaches in any great detail is infeasible here. We will do no more than outline some of these procedures as they relate to agriculture. But first we will outline the general principles of risky portfolio choice.

7.1 PORTFOLIO ANALYSIS

Except for our discussion of risky response analysis in Chapter 6, in previous chapters we have mainly been concerned with decisions involving choices of an "all of this or all of that" nature. For example, in presenting the general model of decision analysis in Section 5.4, our illustrative problem involved the three specific risky prospects of "buy 1000," "buy 1200," or "buy 1600" head of cattle with no possibility of making a mixed choice of, for example, buying a total of 1400 head by taking up half the "buy 1200" prospect and half the "buy 1600" prospect. Situations where such mixtures of risky prospects are feasible and pertinent constitute the field of portfolio analysis, the aim being to find the portfolio (i.e., the allocation of resources across an array of choice possibilities) that maximizes the decision maker's utility. The most obvious application is to the choice of a stock market portfolio as discussed, for example, by Jean (1970), Markowitz (1959), and Sharpe (1970). The principles involved, however, are just as relevant to the selection of an enterprise mix for a farm, a policy mix for a government, or a portfolio of research projects for a research organization—though in the latter two cases there would be obvious data and criterion problems.

To illustrate the principles of risky portfolio choice, we will use a simple but rather general model involving a unidimensional utility function. To take the simplest case, we will make the restrictive assumption that we are dealing with a decision maker who is content to evaluate consequences in terms of profit and whose utility function for profit is quadratic so that mean and variance are the only moments relevant to his risky choice. In this case the problem of portfolio choice from a set of n risky prospects, each of which may be taken up to any varying degree within the constraint of total available funds, may be specified as follows:

Let e_i = expected net return per unit of investment in prospect i, $i = 1, 2, \ldots, n$

σ_i = standard deviation of the per unit net returns from prospect i

σ_{ii} = σ_i^2, i.e., the variance of the per unit net returns from prospect i

σ_{ij} = covariance of the per unit net returns from prospects i and j

ρ_{ij} = correlation coefficient of the per unit net returns from prospects i and j

Z = total units of investment funds available

q_i = units of investment allocated to prospect i

If borrowing and lending are excluded, we must have $q_i \geqq 0$ and $\sum q_i \leqq Z$; and any specified mixture of the n risky prospects will have an expected net return of

$$E = \sum_{i=1}^{n} q_i e_i \tag{7.1}$$

and a variance of net return of

$$V = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} q_i q_j = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j q_i q_j \tag{7.2}$$

where e_i and σ_{ij} may be estimated as discussed in Section 2.3 with appropriate account taken of whatever historical data may be available. Substitution of the above expressions for E and V into the decision maker's quadratic utility function gives utility as a quadratic function in $q_1, q_2, \ldots, q_n$. The problem is thus to find the set of values $\{q_i\}$ that maximizes

$$U = E + bE^2 + bV = \sum q_i e_i + b(\sum q_i e_i)^2 + b\sum\sum \sigma_{ij} q_i q_j \tag{7.3}$$

subject to $q_i \geqq 0$ and $\sum q_i \leqq Z$. As outlined in Section 7.3, such problems may be solved by quadratic programming or related procedures. Diagrammatically the situation is as depicted in Figure 7.1, where the opportunity set of feasible portfolios lies on and below the curve AB. Only portfolios on the curve AB are efficient in the sense that they constitute combinations having maximum E for given V or minimum V for given E. For any portfolio below this curve, a portfolio on the curve can be found that yields greater utility. The (E, V) frontier AB is thus known as the *efficiency locus* or the *efficient set* in (E, V) portfolio analysis; and as shown by Merton (1972) for this case of a single budget constraint, its equation is that of a parabola such that V is a strictly convex function of E.

The optimal portfolio is the member of the efficient set that yields the highest utility. For the risk-averse decision maker with isoutility curves $U^{(1)}$, $U^{(2)}$, and $U^{(3)}$ depicted in Figure 7.1, optimal choice is obviously represented by the portfolio corresponding to the point C on the efficiency locus AB. This gives him the highest achievable level of utility.

While quadratic programming aimed at utility maximization can lead directly to specification of the optimal portfolio (i.e., to the point C in Figure 7.1), it may be just as reasonable to determine the set of efficient (E, V) portfolios, list a reasonable number of them in an appropriate way, and let the decision maker make his choice on an inspection basis.

One extension of the (E, V) analysis of Figure 7.1 should be noted.

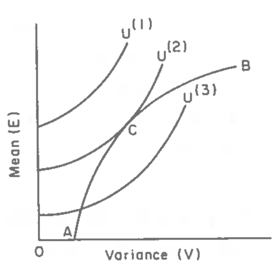


FIG. 7.1. An illustration of (E, V) portfolio analysis $(U^{(1)} > U^{(2)} > U^{(3)})$.

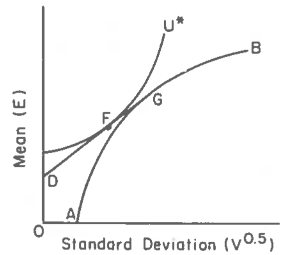


FIG. 7.2. Portfolio analysis with the availability of a riskless prospect.

If any of the n prospects under consideration are riskless (i.e., have zero variance) or, as shown below, if any are perfectly negatively correlated, some portfolios will exist that have $V = 0$. This leads to an enlargement of the opportunity set. Figure 7.2 illustrates the situation in terms of mean E and standard deviation $V^{0.5}$ (so that the frontier AB now becomes a hyperbola). Point D corresponds to the feasible riskless portfolio with highest mean. Since the portfolio D may be mixed with any of the risky portfolios on or below the curve AB, the opportunity set specified in terms of E and $V^{0.5}$ is extended to those portfolios on or below the line DGB where DG is the tangent from D to the curve AB. The efficient set then consists of the line DGB made up of the linear segment DG and the curve GB. As in Figure 7.1, the optimal portfolio is depicted by the point of tangency at F between DGB and an isoutility curve, thereby ensuring the maximum achievable level of utility U^*. Note that if drawn in (E, V) space, because of the change of scale from $V^{0.5}$ to V, the line DG would not be straight but curved toward (i.e., concave to) the V axis.

If borrowing is possible, feasible portfolios lie along the extension of DG through G of Figure 7.2. The efficient set is then represented by the extension of DG to the point corresponding to the maximum feasible level of borrowing; DG extended through G is known as the capital market line. While q_D/Z (i.e., the proportion of total funds invested in the riskless portfolio) ranges from unity at D to zero at G, beyond G it becomes increasingly negative. In stock market terms portfolios on DG are known as lending portfolios and those on the extension of DG through G as leveraged portfolios. Note that leveraged portfolios offer higher expected returns the larger the borrowing is (i.e., the more negative q_D is), but they also have increased variance.

If the returns from a risky portfolio are judged to follow a normal distribution, (E, V) portfolio analysis is relevant even if the decision maker's utility function is not quadratic. The reason is that mean and variance completely specify the normal distribution, which always has odd moments about the mean equal to zero, i.e., $M_k(x) = 0$ for $k = 1, 3, 5, \ldots$, and even moments about the mean given by $M_k(x) = (k-1)(k-3)(k-5)\ldots(3)\cdot(1)V^{k/2}$ for $k = 2, 4, 6, \ldots$. In particular, $M_4(x) = 3M_2(x) = 3V^2$. Accordingly, whatever the form of the utility function, it may be specified in terms of mean and variance via the moment method of Section 4.6. Thus instead of equation (7.3), we would have from (4.27):

$$U = U(E) + U_2(E)V/2 + U_4(E)V^2/8 + U_6(E)V^3/48 + \ldots \qquad (7.4)$$

Moreover, in general we would expect the terms involving $U(E)$ and $U_2(E)$ to be by far the most important.

Relevant to the assumption of normality is the fact that, from the central limit theorem of mathematical statistics, the distribution of the sum of n random variables approaches the normal distribution as n increases. This tendency to normality is stronger the more independent and the more similar the variances of the random variables are. Thus while the assumption of normality for total returns is unlikely to be strictly true, it may be a reasonable approximation, particularly if n is not too small and the risky prospects are diverse.

In an (E, V) context, as Tobin (1965) has shown, only if the decision maker is risk (variance) averse may the optimal portfolio involve a mixture of risky prospects. For a risk preferrer (i.e., one for whom b of equation (7.3) is positive) the optimal portfolio will always consist of only a single risky prospect, since if the prospects are not perfectly correlated, diversification always reduces variability of total returns. Only by putting all his eggs in one basket could a risk-preferring decision maker achieve maximum utility in an (E, V) context.

The relation between correlation and diversification can be shown as follows for the case of two risky prospects i and j. Consider Figure 7.3 where the points i and j respectively denote the portfolios with all the units of investment Z devoted to i or to j. For any mixture of i and j with q_i units

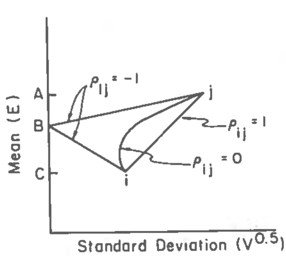


FIG. 7.3. Effect of diversification on the distribution of portfolio returns with varying degrees of correlation.

of Z devoted to i and q_j $(=Z - q_i)$ devoted to j, from equations (7.1) and (7.2) we have

$$E = q_i e_i + (Z - q_i)e_j \tag{7.5}$$

$$V = \sigma_i^2 q_i^2 + 2\rho_{ij}\sigma_i\sigma_j q_i(Z - q_i) + \sigma_j^2(Z - q_i)^2 \tag{7.6}$$

where σ_i denotes the standard deviation of returns from prospect i and ρ_{ij} denotes the correlation coefficient between returns from i and j. Note that the linear form of (7.5) implies that the proportionate holdings q_i/Z in i and $q_j/Z = (Z - q_i)/Z$ in j can be measured by the relative distance along the E axis taken inversely between the E values for i and j. Thus point B in Figure 7.3 corresponds to the portfolio with $q_i/Z = AB/AC$ and $q_j/Z = BC/AC$.

For the case of perfect positive correlation ($\rho_{ij} = 1$), equation (7.6) reduces to

$$V^{0.5} = \sigma_i q_i + \sigma_j(Z - q_i) \tag{7.7}$$

Since e_i, e_j, σ_i, σ_j, and Z are given, we can rearrange (7.7) to give q_i in terms of $V^{0.5}$, substitute into (7.5), and obtain E as a function of $V^{0.5}$. This shows E to be a linear function of $V^{0.5}$ with slope $(e_i - e_j)/(\sigma_i - \sigma_j)$ and therefore passing through the points i and j of Figure 7.3. Hence if i and j are perfectly positively correlated, all portfolio mixtures of them lie on the straight line joining i and j in Figure 7.3.

If i and j are uncorrelated ($\rho_{ij} = 0$), from equation (7.6)

$$V = \sigma_i^2 q_i^2 + \sigma_j^2(Z - q_i)^2 \tag{7.8}$$

which is always less than V from (7.7) since $2\sigma_i\sigma_j q_i(Z - q_i)$ is necessarily positive. Thus if i and j are uncorrelated, their portfolio mixtures have lower variance than if they are perfectly positively correlated, as shown by the curve joining i and j in Figure 7.3. Indeed, the variance-reducing effect of diversification between uncorrelated prospects is so strong that for a risk averter a mixture of them will always dominate the pure prospect with the lower mean. This is shown in Figure 7.3 by the negative slope near i of the curve for $\rho_{ij} = 0$.

If i and j are perfectly negatively correlated ($\rho_{ij} = -1$), from equation (7.6)

$$V^{0.5} = \sigma_i q_i - \sigma_j(Z - q_i) \tag{7.9}$$

If we choose $q_i^* = \sigma_j Z/(\sigma_i + \sigma_j)$, we have $V^{0.5} = 0$. Thus with perfect negative correlation, it is always possible to find a mixed portfolio with zero standard deviation. In this case, as shown in Figure 7.3, the curve for mixtures of i and j consists of two straight lines connecting each of i and j with the zero variance combination depicted by point B. It is unusual to have risky prospects that are perfectly correlated. In general we will have $-1 < \rho_{ij} < 1$, and it is the effect of this that leads to the general shape of the efficient set as shown in Figure 7.1.

Because of its consideration of only the mean and variance of the distribution of returns, there has been strong criticism of (E, V) portfolio analysis. However, as Samuelson (1967) has commented, "in practice, where crude approximations may be better than none, the 2-moment models may be found to have pragmatic usefulness." The analysis can be extended to higher moments if required, using the moment method of equation (4.27), though at some computational cost. In particular, if the analysis is extended to include the third moment, account can be taken of skewness in the distribution of total returns. Just as diversification reduces variance, it also tends to reduce skewness. In general, people seem to prefer positive skewness (i.e., a long tail to the right) and to dislike negative skewness; we would therefore expect a lesser desire to diversify where total returns are positively skewed and a greater desire to diversify where skewness is negative, other things being equal.

We will not give an empirical example of portfolio analysis here; the farm planning example presented in Section 7.3 in the context of quadratic risk programming serves this purpose adequately.

7.2 PLANNING THE WHOLE FARM

Not only must alternative farm enterprises compete for the farmer's limited stock of land, labor, machinery, and capital but they will often also be interdependent if included in the plan. In consequence, many farm planning decisions can only be evaluated properly in terms of the whole-farm situation. The whole-farm planning problem that must be faced, therefore, is to resolve simultaneously (1) which enterprises to adopt on the farm, (2) what method of production to employ in each enterprise, and (3) what amount of resources to allocate to each enterprise.

In the search for operational methods of tackling the whole-farm planning problem, mathematical programming techniques have provided a fruitful line of attack. Of these, linear programming has been the most popular. In linear programming terms, the farm planning problem is to find the optimal values of the variables $x_1, \ldots, x_j, \ldots, x_n$ where x_j represents the level of the jth farming activity. The use of x to denote a decision variable rather than a net return as in earlier chapters is to accord with conventional programming terminology. The activities are chosen to be representative of all possible enterprises that can be conducted on the farm and of all possible ways of undertaking these enterprises. The choice of activity levels is restricted by a set of m linear constraints of the form

$$\sum_{j=1}^{n} a_{hj}x_j \{\leqq \; = \; \geqq\} b_h \qquad h = 1, 2, \ldots, m \tag{7.10}$$

where one and only one of the signs $\leqq$, $=$, or $\geqq$ holds for each constraint, b_h denotes either an accounting identity or (more usually) the available stock of the hth resource, and a_{hj} is the technical input-output coefficient specifying the amount of the hth resource required for a unit of product from the jth activity. These constraints thus reflect the competition between activities for limited farm resources and the interrelationships between activities. The additional restriction that each activity level x_j be nonnegative is also usually imposed since negative areas of crops or negative numbers of livestock are impossible.

Optimality is judged in terms of the maximization of an objective function subject to the constraints of expression (7.10). The objective function is usually farm profit, which can be written as

$$z = \sum_{j=1}^{n} c_j x_j - F \tag{7.11}$$

where z is farm profit, c_j is the per unit net revenue of the jth activity, and F denotes fixed costs. Since by definition the fixed costs do not vary with the levels of the activities, F can be omitted from (7.11) without affecting the choice of an optimal linear programming solution.

In linear programming it is assumed that the a_{hj}, b_h, and c_j are all known constants—an assumption that is fully justified when all the planning coefficients are known for certain. If the uncertainty about most planning coefficients is recognized and if the strong assumption of a linear utility function is made, then the objective function (7.11) can be replaced by the expected profit function

$$E(z) = \sum_{j=1}^{n} E(c_j)x_j - E(F) \tag{7.12}$$

since the expected value of the sum of a number of random variables is the sum of their expectations. Unfortunately, the same treatment in terms of expectations cannot be adopted for the constraints. To write the hth constraint as

$$\sum_{j=1}^{n} E(a_{hj})x_j \{\leqq \; = \; \geqq\} E(b_h)$$

merely requires that the condition specified be satisfied "on the average," with violations occurring perhaps 50% of the time or more. In practice it is seldom reasonable to take such an approach since the cost of violating any particular constraint is not taken into account. Moreover, the use of any measures other than the means of the distributions of random a_{hj} and b_h coefficients tends to be equally arbitrary and unsatisfactory.

The second serious deficiency of the basic linear programming model is that, as the name of the technique implies, both the constraints and the objective function are defined as linear in the variables x_j. Thus "overhead" resource requirements cannot be accounted for in linear programming. Similarly, the simple linear form of the objective function excludes the possibility of accounting directly for a decision maker's nonneutral attitude to risk. However, as discussed in Sections 7.4 and 7.6, the deficiencies of the basic linear programming model can be overcome to some degree by various extensions of the technique.

7.3 QUADRATIC RISK PROGRAMMING

The first attempts to take explicit account of risk in mathematical programming formulations of the whole-farm planning problem were by quadratic risk programming. In these formulations risk is considered only in relation to the activity net revenues c_j, the constraints still being regarded as deterministic. It is usual to assume that the activity net revenues follow a multivariate normal distribution. The relevant statistics are the means, variances, and covariances of the activity net revenues. These are commonly estimated from trend-corrected historical data. Often, however, more subjective methods of assessment (as outlined in Section 2.3) may be more appropriate. The assumption of a multivariate normal distribution for activity net revenues (or appeal to the central limit theorem) implies that the total net revenue z will also be distributed normally (or tend to be) so that utility can be assessed in terms of only the mean and variance of z. Choice of the utility-maximizing set of x_j values can thus be regarded as a particular type of portfolio analysis where the optimal portfolio is some vector $(x_1, \ldots, x_j, \ldots, x_n)$ which maximizes utility subject to the resource constraints of expression (7.10) and subject to $x_j \geqq 0$.

Within the broad framework sketched above, a number of quadratic programming formulations are possible. Some approach utility maximization directly by using the decision maker's utility function expressed in quadratic form as the objective function to be maximized. A somewhat different but common approach is to make use of a parametric programming procedure, such as that developed by Wolfe (1959), as the first stage in a two-stage procedure of first determining the (E, V)-efficient set of portfolios and then ascertaining the utility-maximizing member of this set. We will outline this method as an illustration of quadratic programming. The essence of the first stage is to solve, for β varied parametrically from $-F$ to its maximum relevant value, the quadratic programming problem whose objective function is to minimize

$$s = V(z) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j \tag{7.13}$$

subject to a parametric expected profit constraint

$$\sum_{j=1}^{n} E(c_j)x_j - F = \beta \qquad \beta = -F \rightarrow E_{max} \tag{7.14}$$

and also subject to the constraints of expression (7.10) with $x_j \geqq 0$. In (7.13) σ_{ij} is defined as the covariance of the per unit net revenues of activities i and j, so that the objective is simply to minimize the variance $V(z)$ of profit of the current plan. Similarly, with $E(c_j)$ in (7.14) defined as the expected net revenue per unit of activity j and F denoting fixed costs, the parameter β measures the expected profit $E(z)$ of the current farm plan for values of $\beta = E(z)$ in the range $-F$ to E_{max}, the maximum possible expected profit regardless of variance. When $\beta = E_{max}$, the solution will be identical with that of the corresponding linear programming problem. In terms of portfolio analysis the effect of minimizing (7.13) subject to (7.14) is to ascertain the minimum feasible level of V attainable for any given level of E. These (E, V) values constitute the efficient set.

In solving this parametric quadratic programming problem, an initial optimal solution vector x_0^* is found for $\beta = -F$. Then as β is increased, the levels of the activities in the solution change linearly with β until one of the constraints is met or one of the variables is driven to zero. At this point a "change of basis" occurs, whereupon β can be further increased with the activity levels now varying in a different (linear) manner with β. Further changes of basis occur as further constraints are met or other variables are driven to zero as β is increased. In this way a sequence of critical values of β, denoted $\beta_1, \beta_2, \ldots$, will be obtained, along with a sequence of optimal solutions corresponding to each critical value of β, i.e., a sequence of change-of-basis solutions. At the kth such change-of-basis solution, the expected value E_k and the variance V_k of the total net revenue can be computed respectively as

$$E_k = \sum_{j=1}^{n} c_j x_{kj}^* - F \qquad V_k = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_{ki}^* x_{kj}^*$$

where x_{kj}^* denotes the value of x_j in the solution at the kth change of basis.

For any value of β intermediate between the critical values, the corresponding activity levels to minimize equation (7.13) can be determined by linear interpolation; i.e., for $\beta_k < \beta < \beta_{k+1}$ we have for $j = 1, \ldots, n$:

$$\begin{aligned} x_j^* &= x_{kj}^* + [(\beta - \beta_k)/(\beta_{k+1} - \beta_k)](x_{k+1,j}^* - x_{kj}^*) \\ &= (1 - \lambda)x_{kj}^* + \lambda x_{k+1,j}^* \end{aligned} \tag{7.15}$$

where $\lambda = (\beta - \beta_k)/(\beta_{k+1} - \beta_k)$.

Similarly, it can be shown that

$$\begin{aligned} V(z) &= (1 - \lambda)^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_{ki}^* x_{kj}^* \\ &\quad + \lambda^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_{k+1,i}^* x_{k+1,j}^* + 2\lambda(1 - \lambda) \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_{ki}^* x_{k+1,j}^* \\ &= (1 - \lambda)^2 V_k + \lambda^2 V_{k+1} + 2\lambda(1 - \lambda) H_k \end{aligned} \tag{7.16}$$

where

$$H_k = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_{ki}^* x_{k+1,j}^*$$

while $E(z)$ is simply β.

Using the above parametric procedure, the set of solutions (i.e., activity portfolios) that yields minimum variance for given levels of expected profit subject to the specified constraints can be generated. These solutions represent an efficient set in (E, V) terms in the sense that the (E, V) utility-maximizing solution for any risk-averse decision maker will be in the set. Identification of the optimal portfolio of activity levels for a particular decision maker can be left to individual choice from the efficient set on an inspection basis, or it can be uniquely determined for a specified utility function by evaluating utility expressed in moment terms—as per equation (4.27)—for efficient portfolios to find the one that maximizes utility. This evaluation may be done algebraically or by computerized iteration.

To illustrate this method of quadratic risk programming, we again take a simplified example. Suppose that a farmer has the choice of only three crops—standard wheat, oats, and new wheat—and that the farmer's probability judgments about the revenues of these activities are as discussed in Section 2.3. Variable costs of each crop are assumed to be known with subjective certainty so that the expected net revenues are calculated as shown in Table 7.1 for a unit (one hectare) of each crop.

TABLE 7.1. Derivation of Expected Activity Net Revenues for the Whole-Farm Planning Example

Activity	Expected Gross Revenue	Variable Costs	Expected Net Revenue
		($/ha)	
Standard wheat (x_1)	105.0	33.0	72.0
Oats (x_2)	80.0	26.6	53.4
New wheat (x_3)	190.0	101.2	88.8

Using the data of Tables 2.4 and 2.5, the required variances and covariances of the activity net revenues are

$$[\sigma_{ij}] = \begin{bmatrix} 3600 & 1655 & 3907 \\ 1655 & 1980 & 2470 \\ 3907 & 2470 & 5476 \end{bmatrix}$$

These variances and covariances along with the expected net revenues of Table 7.1 completely specify the multivariate distribution of net revenues. This distribution is multivariate normal since, as shown in Section 2.3, the gross revenues can reasonably be taken as multivariate normal. Note, moreover, that since profit z is a linear combination of the enterprise net revenues (c_js) as per equation (7.11), z will be normally distributed.

We suppose that the farmer has 12 ha of cropland available and that for hygiene reasons not more than 8 ha may be sown to wheat in any year. He is assumed to have \$400 available for working capital, while the per hectare capital requirements of the three crops are: standard wheat \$30, oats \$20, and new wheat \$40. Labor is likely to be limiting only during the harvest period when 80 man-days of family labor are available. Standard wheat and oats both require 5 man-days/ha and new wheat needs 8 man-days/ha. Finally, fixed costs, also assumed to be known with subjective certainty, are set at \$200 per year.

The corresponding linear programming problem is to maximize (expected) profit given by

$$E(z) = 72x_1 + 53.4x_2 + 88.8x_3 - 200$$

subject to

$$\begin{aligned} x_1 + \quad x_2 + \quad x_3 &\leqq 12 \\ x_1 \qquad\quad + \quad x_3 &\leqq 8 \\ 30x_1 + 20x_2 + 40x_3 &\leqq 400 \\ 5x_1 + \ 5x_2 + \ 8x_3 &\leqq 80 \end{aligned}$$

and $x_1, x_2, x_3 \geqq 0$, where x_1 is the area of standard wheat, x_2 is the area of oats, and x_3 is the area of new wheat. The optimal linear programming solution is $x_1 = 1.33$, $x_2 = 4.00$, and $x_3 = 6.67$, yielding an expected net profit of \$701.7. However, when variance of profit is taken into account using the parametric risk-programming procedure, five change-of-basis solutions are generated as listed in Table 7.2 where β_2 corresponds to meeting the land constraint, β_3 to a change of basis involving one of the special variables used in the quadratic programming algorithm, β_4 to meeting the wheat area constraint, and β_5 to meeting the labor constraint. At β_5 the working capital constraint is also nearly met, there being only \$13.30

TABLE 7.2. Change-of-Basis Solutions in Quadratic Risk Programming Analysis of the Farm Planning Problem

k	$E_k = \beta_k$	V_k	x^*_{k1}	x^*_{k2}	x^*_{k3}
1	−200	0	0	0	0
2	535.8	304,500	5.11	6.89	0
3	537.5	305,900	5.20	6.80	0
4	659.5	437,400	3.84	4.00	4.16
5	701.7	500,300	1.33	4.00	6.67

unused. However, it is not possible to reorganize the farm plan to utilize this spare capital effectively, so no β_6 is found.

Solutions for intermediate values of β can be computed by using equations (7.15) and (7.16), the locus of points so generated being depicted by the (E, V) frontier in Figure 7.4. This frontier is the efficient set of activity portfolios. If choice of the optimal portfolio of activity levels is to be made by the decision maker on an inspection basis, he should study the (E, V) frontier of Figure 7.4 and its associated list of activity portfolios and choose the portfolio he thinks best. Obviously, such an approach is likely to be difficult for the decision maker even if he understands the concept of variance. By making analytical use of his utility function, it would be easier to ensure choice of the portfolio that best reflects his preferences.

To see how the optimal point on the (E, V) frontier for a particular

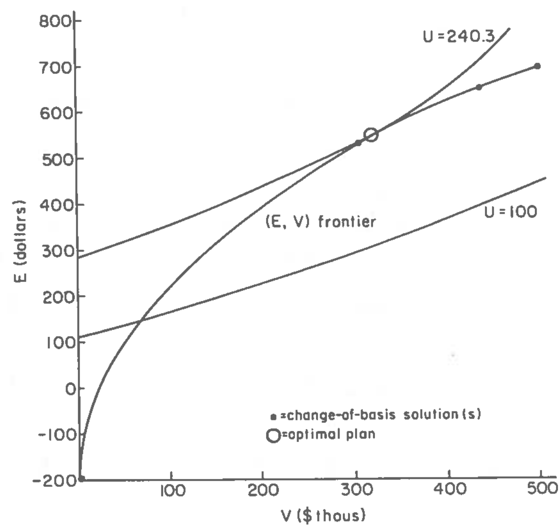


FIG. 7.4. Graphical representation of results of quadratic risk programming.

decision maker can be identified analytically, it is useful to derive the corresponding (E, V) indifference curves from his utility function. To illustrate, we will suppose that in this case the farmer's preferences for the profit over the relevant range can be adequately represented by the quadratic function $U = z - 0.0005z^2$. Then taking expectations and fixing U at, say, U', we obtain

$$V(z) = -2000\{U' - E(z) + 0.0005[E(z)]^2\}$$

Setting $U' = 100$ and solving this expression for a range of values of $E(z)$, we obtain the indifference curve shown in Figure 7.4. It would be possible to use a trial-and-error process to find the value of U' that yields an indifference curve tangential to the (E, V) frontier, the point of tangency corresponding to the optimal farm plan. However, it is more convenient to determine this point algebraically.

For any point on the (E, V) frontier between the kth and the $(k + 1)$th change-of-basis solutions, $E(z) = \beta = (1 - \lambda)E_k + \lambda E_{k+1}$ and $V(z)$ is given by equation (7.16). Hence, substituting these values into the utility function expressed in moment form,

$$U = (1 - \lambda)E_k + \lambda E_{k+1} + b(1 - \lambda)^2E_k^2 + 2b\lambda(1 - \lambda)E_kE_{k+1} + b\lambda^2E_{k+1}^2 + b(1 - \lambda)^2V_k + b\lambda^2V_{k+1} + 2b\lambda(1 - \lambda)H_k \qquad (7.17)$$

Differentiating this expression with respect to λ and setting the result equal to zero gives

$$\lambda = [E_k - E_{k+1} + 2b(E_k^2 - E_kE_{k+1} + V_k - H_k)]/ 2b[(E_k - E_{k+1})^2 + V_k + V_{k+1} - 2H_k] \qquad (7.18)$$

Solving for each pair of adjacent change-of-basis solutions in our case yields four values of λ, only one of which, for $k = 3$, lies in the range $0 \leqq \lambda \leqq 1$; i.e., $\lambda_3 = 0.108$. Since by definition λ must be between zero and one, this value of λ must be the relevant one, allowing the optimal solution (which is quite different from the linear programming solution) to be computed from the third and fourth change-of-basis solutions of Table 7.2 as

$x_1 = 5.05$ ha of standard wheat	$E(z) = 550.6$
$x_2 = 6.50$ ha of oats	$V(z) = 317{,}437$
$x_3 = 0.45$ ha of new wheat	$U = 240.3$

By way of confirmation the (E, V) indifference curve corresponding to $U = 240.3$ is plotted in Figure 7.4 and the point of tangency can be seen.

7.4 LINEAR RISK PROGRAMMING

Practical applications of quadratic risk programming in agriculture have not been numerous. The reasons for this include deficiencies of data and a failure to utilize elicited joint distributions as well as difficulties with quadratic programming algorithms. A number of attempts have been made to develop linear programming models that take account of the stochastic nature of activity net revenues in whole-farm planning. These variations of linear programming aimed at accommodating stochastic net revenues include the incorporation of game theory decision criteria into a programming formulation (McInerney, 1969; Hazell, 1970), the use of constraints on maximum admissible loss (Boussard and Petit, 1967; Boussard, 1971), and the use of mean absolute deviation in place of variance as a measure of risk (Hazell, 1971). Chen and Baker (1974) have shown how a multistage linear programming procedure can be used with a "marginal risk constraint criterion" to approximate the (E, V) frontier of quadratic programming. Their method is more firmly based on decision theory axioms than the game theory and maximum admissible loss approaches, but it is likely to be suitable for large problems only when the number of risky activities is relatively small. Both Hazell and How (1971) and Kennedy and Francisco (1974) have discussed similarities between these kinds of models in a farm planning context. In particular, these authors show that all such models, including quadratic risk programming, can be cast as the minimization of a measure of risk for a range of possible levels of expected profit, subject to the ordinary farm constraints and restrictions. The models vary only in respect to the measure of risk used.

Game Theory Approaches

The game theory models are in effect based on the assumption that all the decision maker's knowledge about the risks involved is embodied in a sample of activity net revenues observed over the previous few years. The sample data, perhaps corrected for trends, are incorporated as additional specially constructed constraints in the linear programming model. For example, McInerney (1967) has shown how the Wald maximin criterion can be incorporated into a whole-farm planning matrix. The problem is formulated to maximize z^-, the farm profit in the event that the most adverse of s possible states of nature occurs. The value of z^- is defined by the additional constraints

$$\sum_{j=1}^{n} c_{rj}x_j - z^- \geqq F \qquad r = 1, \ldots, s \tag{7.19}$$

where c_{rj} is the net revenue per unit of activity j for state r. The approach can be extended to incorporate other game theory criteria such as Laplace and Hurwicz (McInerney, 1969; Kawaguchi and Maruyama, 1972) or to include a constraint, perhaps varied parametrically, on expected farm profit (Hazell, 1970; Low, 1974).

These game theory models can be criticized on the grounds that the decision criteria employed are incompatible with the axioms of rational choice underlying decision analysis (Officer and Anderson, 1968). Game theory criteria imply that nature is malevolent, when this is obviously not so. Moreover, they are based on the assumption that no probabilities can be attached to the defined states to reflect differences in their assessed chances of occurrence. Such approaches are at variance with the philosophy of decision making advocated in this book.

Maximum Admissible Loss Approach

Boussard's (1971) "focus-loss constraint" or "maximum admissible loss" approach involves the addition to the linear programming model of constraints designed to limit the risk of ruin. With some minor variation Boussard's method is illustrated here as an example of the incorporation of risk constraints in linear programming models for farm planning.

First, a maximum admissible loss L is defined for the particular farmer such that L is the difference between his expected profit and the minimum level of profit he needs to sustain inescapable consumption expenses when consumption expenditure is reduced as much as possible. Thus the admissible loss is defined by

$$L = E(z) - z_c = \sum_{j=1}^{n} E(c_j)x_j - E(F) - z_c \tag{7.20}$$

where z_c, known also as the focus loss, is the minimum level of profit necessary to meet inescapable consumption. Note that if the farmer is (within the context of one-year planning) able to borrow to assist with essential consumption expenditure, z_c may be less than the amount required for consumption and may even be negative.

For each activity j there is defined a "normal" net revenue per unit, which may be reasonably taken as $E(c_j)$, and a possible deficiency in net revenue per unit r_j, where r_j is the difference between the normal activity net revenue and the net revenue that will be obtained if "things go wrong." A reasonable interpretation of these definitions is that r_j is selected such that the actual net revenue of the jth activity will be greater than or equal to $E(c_j) - r_j$ with some specified probability that is not too far from unity.

It is also postulated that the activity safety constraints will be satisfied if the possible deficiency of any activity's net revenue does not exceed a specified fraction $1/k$ of the admissible loss; i.e.,

$$r_j x_j \leqq L/k \qquad j = 1, 2, \ldots, n \tag{7.21}$$

The choice of a value for k in these focus-loss constraints is somewhat arbitrary, but Boussard and Petit (1967) have shown that when activity net revenues are normally distributed, $k^2 \geqq n^*$ is a reasonable condition, where n^* is the number of activities in the optimal plan. Unfortunately, n^* cannot be known in advance, but $k = 3$ has been suggested as a convention that appears to give results similar to farm plans adopted by farmers (Boussard, 1971).

Although risk constraints derived in this fashion appear somewhat arbitrary, Boussard and Petit (1967) and Kennedy and Francisco (1974) provide a more formal justification for the case where activity net revenues are independently normally distributed. They show that if the r_j are chosen so that $r_j = \sigma_j t_p$ (where σ_j is the standard deviation of the net revenue of the jth activity and t_p is the negative of the value of the standard normal variate at the cumulative probability level p) and if activity net revenues are normally distributed, from expression (7.21) we have $\sigma_j t_p x_j \leqq L/k$. But $k^2 \geqq n$ ($\geqq n^*$), so $\sigma_j x_j \leqq L/t_p n^{0.5}$. Squaring both sides and summing over the activities leads to

$$\sum_{j=1}^{n} \sigma_j^2 x_j^2 \leqq L^2/t_p^2 \tag{7.22}$$

But assuming zero covariance,

$$V(z) = \sum_{j=1}^{n} \sigma_j^2 x_j^2$$

while by definition $L = E(z) - z_c$. Substituting in (7.22) gives

$$[E(z) - z_c]/[V(z)]^{0.5} \geqq t_p \tag{7.23}$$

This is precisely the requirement that $P(z \leqq z_c) = p$ for a normal distribution. In other words, if activity net revenues are independently normally distributed, the focus-loss constraints of (7.21) have the effect of restricting the chance that profit will fall below the chosen critical level z_c to some maximum probability p. At the worst there is a probability of p that z will be less than z_c. Whether activity net revenues are independently normally distributed in any particular case is an empirical question. In general, independence is not to be expected.

Applying Boussard's approach to our simplified farm planning example used to illustrate quadratic programming and assuming expected fixed cost to be \$200, we again define expected profit as the objective function to be maximized; i.e., maximize

$$E(z) = 72x_1 + 53.4x_2 + 88.8x_3 - 200$$

The technical constraints are also unchanged, but the safety constraints are added. First, the admissible loss L is defined by the constraints

$$L \geqq 0$$

$$72x_1 + 53.4x_2 + 88.8x_3 - 200 - L = z_c$$

We will avoid the difficulty of selecting a value for the critical level of profit z_c, and hence also for $L = E(z) - z_c$, by solving the problem parametrically for all values of z_c from minus infinity to the maximum for which a solution exists.

Next we must add a safety constraint akin to expression (7.21) for each activity. We will arbitrarily take $E(c_j) - r_j$ as being one standard deviation below the mean net revenue for each activity so that $t_p = 1$ and, assuming the c_j are independently normal, $P(z \leqq z_c) = 0.16$. We will also adopt the value of $k = 3$ $(> n^{0.5} = 1.73)$ previously mentioned. Hence we have $r_j = \sigma_j$ and the activity safety constraints corresponding to (7.21) are

$$\begin{aligned} 60x_1 \qquad\qquad\qquad &- 0.333L \leqq 0 \\ 44.5x_2 \qquad\quad &- 0.333L \leqq 0 \\ 74x_3 &- 0.333L \leqq 0 \end{aligned}$$

The change-of-basis solutions to this augmented problem are summarized in Table 7.3. Solutions for intermediate values of z_c can be obtained by linear interpolation. For the specified level of z_c (or L), each of these solutions gives the activity portfolio that maximizes expected profit subject to the technical constraints and to the constraint that there is a chance of no more than $p = 0.16$ that actual profit z is less than the focus-loss level z_c. The behavioral assumptions in this particular parametric application are that $k = 3$ and $p = 0.16$ (or $t_p = 1$).

It can be seen from Table 7.3 that if the critical level of profit is $-\$780$ or less, the risk constraints have no effect, the solution obtained being that which maximizes expected profit. But as z_c becomes less negative, the risk constraints become effective, reducing the area of the most risky crop, new wheat (x_3). At first, the land released is devoted to standard wheat (x_1); but eventually, as the risk constraints are tightened with further increases in z_c, the standard-wheat area is cut back again and the area of the least risky crop, oats (x_2), is expanded. There is no feasible solution to the problem if z_c is increased beyond $-\$64$. In other words, with $k = 3$ there is no ac-

TABLE 7.3. Parametric Solutions to Maximum Admissible Loss Programming of the Farm Planning Problem with $P(z \leq z_c) = 0.16$

z_c	x_1	x_2	x_3	L	$E(z)$
$-\infty$	1.33	4.00	6.67	$+\infty$	701.7
−780	1.33	4.00	6.67	1482	701.7
−146	4.42	4.00	3.58	796	649.8
−64	3.80	5.12	3.08	684	620.5
QP	5.05	6.50	0.45		550.6

tivity portfolio for which $P(z \leqq z_c) = 0.16$ if z_c is larger than $-\$64$. As would be expected, the value of the admissible loss L declines as z_c is increased.

For comparison the quadratic risk programming (QP) solution of Section 7.3 is also listed in Table 7.3. The differences are rather pronounced. They reflect the fact that in this particular case, the behavioral assumptions of the Boussard approach imply a lower level of risk aversion than implied by the utility function assumed in our quadratic programming formulation. Also, while covariance of activity net revenues is allowed for in the quadratic risk programming analysis, it is ignored in the focus-loss approach.

Mean Absolute Deviation Approach

The mean absolute deviation or MOTAD approach is more in tune with a decision analysis view of the whole-farm planning problem than are either the game theory or focus-loss formulations. MOTAD stands for minimization of total absolute deviations. As will be discussed, the method closely parallels the quadratic programming approach, but without the need for a nonlinear programming algorithm. It also readily permits the incorporation of assessed probabilities of occurrence of alternative states of nature.

Given an appropriate sample of activity net revenues from previous years, an unbiased estimate of the mean absolute deviation of expected farm profit is given by

$$M = s^{-1} \sum_{r=1}^{s} \left| \sum_{j=1}^{n} (c_{rj} - \bar{c}_j) x_j \right| \tag{7.24}$$

where s is the sample size, c_{rj} is the net revenue observation for the jth activity in the rth year, and $\bar{c}_j$ is the sample mean net revenue per unit of the jth activity. Hazell (1971) shows that with some manipulation this measure of risk can be incorporated into an augmented linear programming model of a farm planning problem such that mean absolute deviation M can be minimized for a given level of expected profit $E(z)$ varied parametrically over the relevant range. In this way, what might be called the $[E(z), M]$-efficient set of farm plans is generated.

An equivalent but computationally more tidy approach, also exposited by Hazell, is to work with the mean absolute value of negative deviations about the mean, estimated as

$$D = M/2 = s^{-1} \sum_{r=1}^{s} \left| \min \left[\sum_{j=1}^{n} (c_{rj} - \bar{c}_j) x_j, 0 \right] \right| \tag{7.25}$$

The negative deviations can be measured for each year of sample data by new variables y_r defined as

$$y_r = -\sum_{j=1}^{n} (c_{rj} - \bar{c}_j)x_j$$

when the summation yields a negative total and zero otherwise. Following Hazell, the programming problem can then be formulated as the minimization of the sum of the variables y_r, subject to the usual technical constraints and to a parametric constraint on expected total net revenue. Alternatively, the expected farm profit can be maximized with a parametric constraint on the sum of the negative deviations; i.e., maximize

$$E(z) = \sum_{j=1}^{n} \bar{c}_j x_j - F \tag{7.26}$$

subject to

$$\sum_{j=1}^{n} a_{hj}x_j \{\geqq \; = \; \leqq\} b_h \qquad h = 1, \ldots, m \tag{7.27}$$

$$\sum_{j=1}^{n} (c_{rj} - \bar{c}_j)x_j + y_r \geqq 0 \qquad r = 1, \ldots, s \tag{7.28}$$

$$\sum_{r=1}^{s} y_r \leqq \lambda = sM/2 \qquad \lambda = 0 \rightarrow \lambda_{\max} \tag{7.29}$$

with $x_j \geqq 0$ for $j = 1, \ldots, n$ and $y_r \geqq 0$ for $r = 1, \ldots, s$.

In this formulation the usual technical constraints are represented by expression (7.27). In (7.28) there is one variable y_r to measure the negative deviation of the total net revenue for each state r, $r = 1, \ldots, s$. The total deviation for each state is computed in the summation term of (7.28). If this sum is positive, the corresponding y_r variable will be zero. This is assured by the nonnegativity restrictions on the y_r and because the total value of the objective function is limited through the parametric constraint on the sum of the y_r variables in (7.29). Thus only if the sum of the net revenue deviations for any state is negative in (7.28) will the corresponding variable y_r be forced to an equivalent positive value, so λ in (7.29) will measure the sum of the total negative deviations over the s states.

It is a simple matter to extend this model to the case where the sample of net revenue observations is regarded as a set of states of nature, each of which is assigned a probability of occurrence p_r, $r = 1, \ldots, s$, with $\sum_{r=1}^{s} p_r = 1$. The sample mean activity net revenues $\bar{c}_j$ are replaced by expected values $E(c_j)$, computed in the usual way. Then constraint (7.29) in the formulation above becomes

$$\sum_{r=1}^{s} p_r y_r \leqq \Theta = M/2 \qquad \Theta = 0 \rightarrow \Theta_{\max}$$

As shown by Hazell (1971), justification for the use of M as a measure of risk comes from the fact that an unbiased estimate of the population variance is given by $M^2[\pi s/2(s-1)]$ when the population is normal or approximately normal. Thomson and Hazell (1972) have also shown in a Monte Carlo sampling study that the efficient set of plans generated using this measure corresponds closely with the (E, V)-efficient set. They demonstrated that a mean absolute deviation model can be a satisfactory surrogate for quadratic risk programming and indeed, if distributions are skewed, may be superior. Given the relevant data, it is a simple matter to compute the (E, V) locus of the (E, M)-efficient set of farm plans.

To illustrate the above model, we turn again to our simple three-crop example. We now assume that the farmer has enterprise net revenue data available extending over the last five years, as shown in Table 7.4, and that he is prepared to plan his farm on the assumption that uncertainty about next year's results is adequately reflected by these sample data alone; i.e., he assigns equal probabilities to each state. It would be possible to derive hypothetical sets of net revenue data representing a number of possible states (years) in a purely subjective manner to reflect the farmer's beliefs about possible future outcomes. Similarly, the methods introduced in Section 2.4 for probability assessments with sparse data could be employed to make the best possible use of the few available historical data that might be available. The use of such procedures would not affect the way the programming model is applied.

TABLE 7.4. Assumed Sample of Activity Net Revenues for Solving the Farm Planning Problem by the Mean Absolute Deviation Approach

Prior Year	c_1	c_2	c_3
1	99.8	68.3	112.7
2	133.3	130.4	238.4
3	142.7	33.3	93.9
4	154.3	74.4	83.2
5	11.4	25.4	109.7
Mean	108.30	66.36	127.58

Applying the system of expressions (7.26)–(7.29) to our data implies maximizing

$$E(z) = 108.3x_1 + 66.36x_2 + 127.58x_3 - 200$$

subject to

$$\begin{array}{rrrrrrrrl}
x_1 + & x_2 + & x_3 & & & & & & \leqq 12 \\
x_1 + & & x_3 & & & & & & \leqq 8 \\
30x_1 + & 20x_2 + & 40x_3 & & & & & & \leqq 400 \\
5x_1 + & 5x_2 + & 8x_3 & & & & & & \leqq 80 \\
-8.5x_1 + & 1.94x_2 - & 14.88x_3 & +y_1 & & & & & \geqq 0 \\
25.0x_1 + & 64.04x_2 + & 110.82x_3 & & + y_2 & & & & \geqq 0 \\
34.4x_1 - & 33.06x_2 - & 33.68x_3 & & & + y_3 & & & \geqq 0 \\
46.0x_1 + & 8.04x_2 - & 44.38x_3 & & & & + y_4 & & \geqq 0 \\
-96.9x_1 - & 40.96x_2 - & 17.88x_3 & & & & & + y_5 & \geqq 0 \\
& & & y_1 + & y_2 + & y_3 + & y_4 + & y_5 & \leqq \lambda
\end{array}$$

with all variables x_j and y_r nonnegative and for $\lambda = 5D = 2.5M$ varied parametrically from zero to its maximum relevant value. Critical values of λ at which changes of basis occur are denoted by λ_k, $k = 0, 1, \ldots$, solutions for intermediate values of λ being obtainable by linear interpolation.

The change-of-basis solutions for the model are shown in Table 7.5. In comparing these results with those obtained for the quadratic risk programming formulation given in Table 7.2, it should be recalled that because the present model is constructed using only a small sample of data, the coefficients for expected activity net revenues per unit differ between the two formulations. Nevertheless, some broad similarities between the two sets of results can be seen. There are four change-of-basis solutions in each case, with identical activity levels for the first and last solutions. Only the second solution shows appreciable differences in activity levels. Although no generality can be attached to the similarities in the results obtained for the two formulations using this very simple illustrative

TABLE 7.5. Change-of-Basis Solutions in Mean Absolute Deviation Programming Analysis of the Farm Planning Problem

k	x_{k1}	x_{k2}	x_{k3}	E_k	M_k
0	0	0	0	−200	0
1	3.93	0	4.07	744.9	219.8
2	3.57	4.00	4.43	1017.2	334.5
3	1.33	4.00	6.67	1060.4	411.3

example, it conforms with the conclusion of Thomson and Hazell (1972) that the MOTAD approach can be used as a reasonable substitute for quadratic risk programming. Further, the MOTAD approach has the important advantage of requiring only the standard linear programming algorithm.

Rather than using a sample of activity net revenues from prior years, Hazell's (1971) approach can be formulated in terms of assessed distributions for the activity net revenues, as shown by Kennedy and Francisco (1974). Let the farm profit z (conditional upon the farm plan) be a random variable with expected value $E(z)$ and mean absolute deviation $M = E[\,|z - E(z)|\,]$. Regardless of the probability distribution of z, Tchebycheff's inequality can be used to write

$$P[\,|z - E(z)| \geqq K] \leqq M/K \tag{7.30}$$

where P is the probability operator and K is a positive constant. Substituting $E(z) - z_d$ for K where z_d is the disaster level of profit, we obtain

$$P[\,|z - E(z)| \geqq E(z) - z_d] \leqq M/[E(z) - z_d]$$

or, recalling that $|z - E(z)|$ includes both $[z - E(z)]$ if $z \geqq E(z)$ and $[E(z) - z]$ if $z \leqq E(z)$,

$$P(z \leqq z_d) + P[z \geqq 2E(z) - z_d] \leqq M/[E(z) - z_d] \tag{7.31}$$

If nothing is known about the second term on the left side of this inequality other than $0 \leqq P[z \geqq 2E(z) - z_d] \leqq 1$, expression (7.31) may be written as

$$P(z \leqq z_d) \leqq M/[E(z) - z_d] \tag{7.32}$$

Alternatively, if z is symmetrically distributed so that $P(z \leqq z_d) = P[z \geqq 2E(z) - z_d]$, we can write

$$P(z \leqq z_d) \leqq M/2[E(z) - z_d] \tag{7.33}$$

which, in probability terms, is twice as strong as (7.32). The right sides of (7.32) and (7.33) represent upper bounds on the probability of the profit falling below the disaster level z_d. A linear programming constraint can be specified to restrict these upper bounds to some critical probability p or less; e.g., if (7.33) applies, we can specify the constraint $M/2[E(z) - z_d] \leqq p$, which implies the solutions are restricted to portfolios that guarantee no more than a chance of p that profit will fall below the disaster level z_d. Further, assuming zero covariance, the mean absolute deviation of farm profit may be estimated by

$$M = \sum_{j=1}^{n} m_j x_j \tag{7.34}$$

where $m_j = E[\,|\,c_j - E(c_j)\,|\,]$ is the per unit mean absolute deviation of activity j. Similarly, expected profit is given by

$$E(z) = \sum_{j=1}^{n} E(c_j)x_j - E(F) \qquad (7.35)$$

where $E(c_j)$ is the expected net revenue of activity j and $E(F)$ is expected fixed cost. If the further assumption of a symmetric distribution is made, from (7.33) we can specify the constraint

$$\sum_{j=1}^{n} m_j x_j / 2\left[\sum_{j=1}^{n} E(c_j)x_j - E(F) - z_d\right] \leqq p$$

which is more conveniently written as

$$\sum_{j=1}^{n} [2pE(c_j) - m_j]x_j \geqq 2p[E(F) + z_d] \qquad (7.36)$$

The linear programming problem can then be solved to maximize $E(z)$ of equation (7.35) with (7.36) as an additional constraint. Alternatively, and perhaps more informatively, the problem may be specified to minimize M while $E(z)$ is varied parametrically, thereby enabling upper bounds to be placed on the probabilities $P(z \leqq z_d)$ for the range of solutions generated. In this way a range of efficient $[E(z), M]$ combinations can be obtained; for each of these, using the relation $p \geqq M/[E(z) - z_d]$ or $p \geqq M/2[E(z) - z_d]$, the minimum value of p for given z_d or the maximum value of z_d for given p can be found.

A disadvantage of this generalized approach based on Tchebycheff's inequality is that covariance between activity net revenues is ignored. In contrast, in the prior formulation based on a sample of outcomes from previous years, interrelationships between activity net revenues are allowed for by the nature of the sample used. In principle, it seems that it should be possible to exploit this advantage in the case where the full joint distribution of activity net revenues has been assessed. This could be done by selecting representative (or at least pseudorandom) vectors of outcomes from the assessed joint distribution. However, further work is needed to determine how such representative vectors should be constructed. In this regard, the data of Table 7.4 were obtained by pseudorandom normal sampling from the statistics of means, variances, and covariances previously used in the quadratic risk programming example of Section 7.3.

7.5 MONTE CARLO PROGRAMMING

A related approach to the problem of farm planning under risk (not strictly a mathematical programming method) is Monte Carlo programming. In this approach the planning problem is formulated in a broadly similar fashion to that needed for other programming methods, but portfolios of activity levels are selected at random using a computer. Portfolios so generated are first tested for feasibility and are then evaluated in terms of some specific objective function. A large number of such portfolios can be inspected, and a selection of the best can be printed out. The procedure is thus one of search. Two important advantages of Monte Carlo programming are, first, that it is very easy to take account of integer constraints on activities and, second, that almost any form of objective function can be applied. In particular, a utility function defined in terms of the mean and variance of total net revenue is readily computable and, in principle, higher order moments of the distribution could be accommodated. However, in practice, as with other approaches, the problems of assessing the necessary statistics for third and higher order moments may well prove insurmountable.

Once again we turn to our simplified farm planning example. We now suppose that to fit in with fixed field boundaries, the areas of crops in the farm plan must take integer values. The objective function to be maximized is again assumed to be the utility function $U = z - 0.0005z^2$ which, in terms of the first two moments of net profit, implies:

$$U = \sum_{j=1}^{3} E(c_j)x_j - E(F) - 0.0005\left[\sum_{j=1}^{3} E(c_j)x_j - E(F)\right]^2 - 0.0005\sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij}x_j \qquad (7.37)$$

The other technical constraints remain as before so that we must maximize equation (7.37) subject to

$$\begin{aligned} x_1 + x_2 + x_3 &\leqq 12 \\ x_1 + x_3 &\leqq 8 \\ 30x_1 + 20x_2 + 40x_3 &\leqq 400 \\ 5x_1 + 5x_2 + 8x_3 &\leqq 80 \end{aligned}$$

and with x_j allowed to have only nonnegative integer values.

This problem now includes both a quadratic objective function and integer constraints on the variables. It would require a relatively advanced algorithm if it were to be solved by mathematical programming methods. However, the problem can be tackled productively using Monte Carlo programming. In this simple application the procedure was applied by hand. Fifty portfolios (x_1, x_2, x_3) were inspected using the following routine. First, the order of selection of the three activities was determined randomly for each portfolio to be inspected. Second, the amount of land

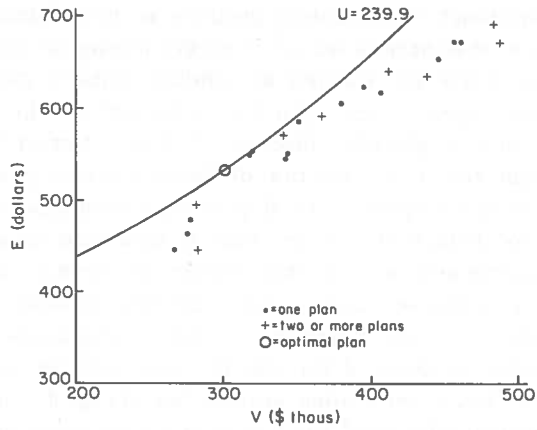


FIG. 7.5. Graphical representation of activity portfolios generated by Monte Carlo programming.

(some integer from 0 to 12 inclusive) to be devoted to the first-selected crop was determined randomly for each portfolio, the level being revised downward if necessary to bring it within the other constraints. Third, the integer area of the second activity to be selected in each portfolio was randomly chosen within the limits of the remaining resources. Fourth, remaining resources were allocated as far as possible to the third activity. Finally, the portfolio so obtained was evaluated in terms of the defined objective function, using the values of $E(c_j)$, $E(F)$, and σ_{ij} specified previously.

The farm plans derived by the above procedure are plotted in (E, V) space in Figure 7.5. The ten best plans in terms of the utility criterion are summarized in Table 7.6. The (noninteger) quadratic risk programming result is also listed for comparison.

TABLE 7.6. Ten Best Portfolios Selected by Monte Carlo Programming of the Integer Farm Planning Problem

Rank	x_1	x_2	x_3	$E(z)$	$V(z)$	U
1	5	7	0	533.8	302,870	239.9
2	6	6	0	552.4	320,040	239.8
3	4	7	1	550.6	318,612	239.7
4	4	6	2	586.0	352,016	237.8
5	7	5	0	571.0	341,750	237.1
6	3	9	0	496.6	282,150	232.2
7	7	4	1	606.4	380,694	232.2
8	8	4	0	589.6	368,000	231.8
9	6	4	2	623.2	395,912	231.1
10	5	4	3	640.0	413,654	228.4
QP	5.05	6.50	0.45	550.6	317,437	240.3

It can be seen that although there is no guarantee the true optimum will be identified by Monte Carlo programming, nevertheless for this simple example the best three plans produced have (E, V) values not too different from the quadratic programming solution. A farmer might reasonably adopt any one of these first three plans.

As the above example illustrates, Monte Carlo programming provides a very flexible means of generating good farm plans. Its flexibility extends from integer specification to such probabilistic features as stochastic returns, constraints, and technical coefficients. An arbitrarily large number of near-optimal plans may be identified. In this regard, the procedures (outlined in Chapter 9) for eliminating plans on the basis of stochastic dominance may be very useful in confining the near-optimal plans to a smaller number of potentially desirable plans.

7.6 STOCHASTIC PROGRAMMING

Thus far we have confined ourselves to the case where only activity net revenues are stochastic. Unfortunately, in practice, risk is seldom if ever confined to those coefficients. Both the technical coefficients and some resource stocks, respectively represented by the a_{hj} and b_h coefficients of equation (7.10), may be stochastic. For example, on a grazing farm both the nutritional requirements per head of livestock and the feed resources available may be affected by weather. On a cropping farm the rate of cultivation, sowing, or harvesting and the time available for these jobs can vary widely from year to year. Programming methods to deal with risk in the resource constraints are usually known under the generic name of stochastic programming.

Following Hadley (1964), it is convenient to classify stochastic programming problems into the two broad groups of sequential problems and nonsequential problems. *Sequential decision problems* involve making two or more related decisions at different points in time. They have the property that the later decision(s) may be influenced both by the earlier decisions and by stochastic parameters whose values become known to the decision maker after the first decision(s) but before the later decision(s). In *nonsequential decision problems,* on the other hand, all decisions are made at one point in time, or if spread through time, there is not the interleaving of decisions and uncertain events. Because of the biological nature of the agricultural production process, most farm planning problems are to a greater or lesser degree sequential in nature; and unfortunately, as we will explain, these sequential problems are generally not very amenable to solution by mathematical programming methods (or by any other optimizing methods). This is not to say that decision analysis is irrelevant to agricultural decisions; we hope its relevance has been amply demonstrated in previous chapters. Inevitably, however, time effects cause complexity.

Nonsequential Stochastic Programming

Two types of nonsequential stochastic programming problems can be identified: problems with random variables only in the right-hand side vector of b_h values and problems where some or all of the technical input-output coefficients a_{hj} are also random variables. We will consider each case in turn.

To illustrate the case where only the b_h elements involve stochasticity, we will again use our simple three-crop farm planning problem. So far our harvest labor constraint has been $5x_1 + 5x_2 + 8x_3 \leqq 80$. We now suppose that the amount of labor that will be available for harvesting is not a known constant but is uncertain. Consequently, the number of man-days in the right side of the above constraint becomes a random variable that we will represent by $v \geqq 0$ with subjective density function $f(v)$. To take account of this, the constraint can be rewritten as the equality

$$5x_1 + 5x_2 + 8x_3 - S = v \tag{7.38}$$

where $-S > 0$ measures any unused harvest labor and $S \geqq 0$ measures any shortage of labor, according to whether $5x_1 + 5x_2 + 8x_3$ is smaller or greater than v. We assume that surplus labor can be employed in other less essential ways and therefore has no cost but that a shortage of labor can only be made good at a penalty cost of \$5 per man-day. Assuming a linear utility function, the objective function must thus include a term representing the expected cost of a labor shortage, the magnitude of this cost being determined by the levels of the three main decision variables x_1, x_2, and x_3. This expected cost is equal to $5\int_0^w Sf(v)\,dv$ where $w = 5x_1 + 5x_2 + 8x_3$ and $S = (w - v) \geqq 0$. In these terms, the programming problem is to maximize expected profit specified as

$$E(z) = 72x_1 + 53.4x_2 + 88.8x_3 - 200 - 5\int_0^w Sf(v)\,dv \tag{7.39}$$

subject to

$$\begin{aligned} x_1 + x_2 + x_3 &\leqq 12 \\ x_1 + x_3 &\leqq 8 \\ 30x_1 + 20x_2 + 40x_3 &\leqq 400 \\ 5x_1 + 5x_2 + 8x_3 - w &= 0 \end{aligned}$$

and $x_1, x_2, x_3 \geqq 0$. Here we have a programming problem with linear constraints but with a nonlinear objective function that is separable in the sense that it is expressed as the sum of separate functions of the single variables x_1, x_2, x_3, and w. Such a problem can be solved using a separable programming algorithm in which the nonlinear functions are represented by piecewise linear approximations. Separable programming algorithms

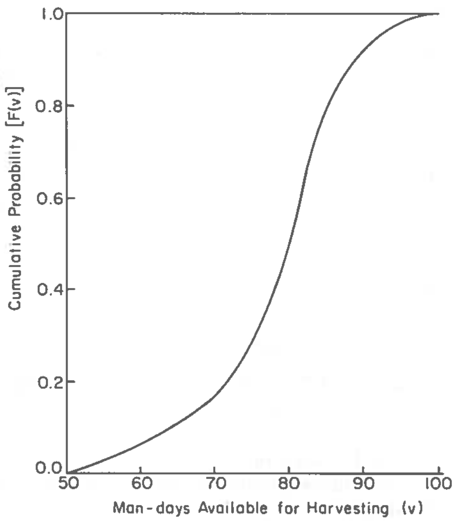


FIG. 7.6. Subjectively assessed CDF for available man-days of labor for the nonsequential stochastic programming example.

can generally only guarantee a local rather than a global optimum, but in the present case the nonlinear objective function is concave. In consequence, as Hadley (1964, Ch. 4) shows, separable programming will permit an approximate global optimum to be found.

To illustrate, we suppose that the CDF for v (the man-days available for harvesting) has been subjectively assessed as shown in Figure 7.6. Using the method of discrete approximation described in Section 2.3 or the Pearson-Tukey procedure of equation (2.1), the expected value of v is found to be approximately 79 days as compared with the 80 days previously assumed to be available. We now define the labor-shortage function

$$g(w) = \int_0^w Sf(v)dv = \int_0^w (w - v)f(v)dv \tag{7.40}$$

and compute $g(w)$ for a range of values of w using discrete approximation. The result is shown in Figure 7.7. This function is then approximated by a series of ten linear segments and the corresponding separable programming problem solved. This gives the solution:

$$\begin{aligned} x_1 &= 0 \text{ ha of standard wheat } (1.33) \\ x_2 &= 4 \text{ ha of oats } (4.00) \\ x_3 &= 8 \text{ ha of new wheat } (6.67) \\ w &= 84 \text{ days of labor demanded } (80.0) \\ E(z) &= 691.1\ (701.7) \end{aligned}$$

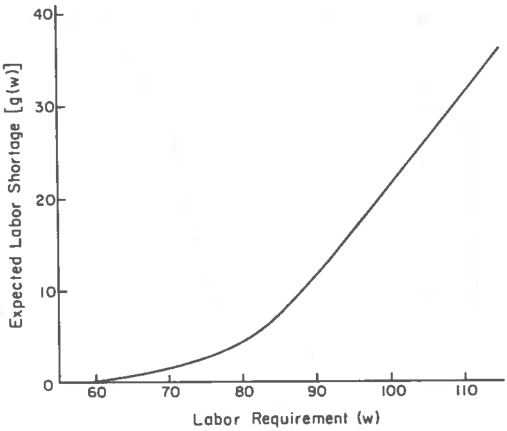


FIG. 7.7. Plotted values of $g(w) = \int_0^w (w - v)f(v)\,dv$ for the nonsequential stochastic programming example.

The figures in parentheses on the right show the solution of the nonstochastic linear programming problem when labor supply is held constant at 80 man-days. It can be seen that introduction of a stochastic labor supply has allowed a more intensive plan to be followed, with the entire wheat area planted with new wheat. Labor demand rises to 84 man-days even though the mean labor supply is now only about 79 man-days. Thus the probability of some deficiency in labor supply is quite high (about 0.73, in fact). In the event of no labor deficiency, the profit would be (4)(53.4) + (8)(88.8) − 200 = \$724; but when a deficiency does occur it must be made good at a cost of \$5 per man-day. The expected labor shortage cost is computed as \$32.9, leaving an expected profit of 724 − 32.9 = \$691.1 as shown.

The above illustration of nonsequential stochastic programming was drastically simplified by assuming a linear utility function so that expected profit could be taken as the objective function. With a nonlinear utility function, the problem becomes more complicated. Suppose, for example, that the decision maker has a utility function of the power form $U = (W_0 + z)^{0.8}$, where W_0 is his current wealth and z is the gain or loss (i.e., positive or negative profit) from the current farm planning decision problem. Under the risky conditions of our example, the farmer will wish to maximize

$$\begin{aligned} U &= E[(W_0 + z)^{0.8}] \\ &= E[\{W_0 + \sum_{j=1}^{3} c_j x_j - F - 5\int_0^w (w - v)f(v)dv\}^{0.8}] \end{aligned} \tag{7.41}$$

subject to the constraints

$$\begin{aligned} x_1 + x_2 + x_3 &\leqq 12 \\ x_1 + x_3 &\leqq 8 \\ 30x_1 + 20x_2 + 40x_3 &\leqq 400 \\ 5x_1 + 5x_2 + 8x_3 - w &= 0 \end{aligned}$$

and $x_1, x_2, x_3 \geqq 0$, where, as before, c_j denotes the uncertain net revenue per unit of activity j and F denotes fixed costs.

Because of its power form, the expectation of equation (7.41) is best approached by approximation. The moment method of analysis outlined in Section 4.6 is appropriate. Taking the approximation in terms of the first two moments,

$$U = [W_0 + E(z)]^{0.8} - 0.08V(z)[W_0 + E(z)]^{-1.2} \tag{7.42}$$

where W_0 is some given value, $E(z)$ is as specified in (7.39). and

$$V(z) = \sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij}x_ix_j + 25\int_0^w (v - \bar{v})^2 f(v)\,dv \tag{7.43}$$

where $\bar{v}$ denotes the mean of v, i.e., 79. No allowance is made for covariance between activity net revenues and the supply of man-days since these can reasonably be regarded as independent. Thus substituting in (7.42) for $E(z)$ from (7.39) and for $V(z)$ from (7.43), the objective function to be maximized subject to the constraints listed above is

$$\begin{aligned} U = &\left[W_0 + 72x_1 + 53.4x_2 + 88.8x_3 - 200 - 5\int_0^w (w - v) f(v)\,dv\right]^{0.8} \\ &- 0.08\left[\sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij}x_ix_j + 25\int_0^w (v - 79)^2 f(v)\,dv\right] \cdot \\ &\left[W_0 + 72x_1 + 53.4x_2 + 88.8x_3 - 200 - 5\int_0^w (w - v) f(v)\,dv\right]^{-1.2} \end{aligned} \tag{7.44}$$

This function is most unpromising from a mathematical programming viewpoint. It is both nonlinear and nonseparable. Consequently, the planning problem must be approached in an indirect way. First it may be noted that the second derivative of the utility function is negative for $z > 0$, implying risk aversion. Thus, under the strong assumption that only the first two moments of z need to be considered, a plausible procedure, therefore, would be to try to use the quadratic risk programming approach of equations (7.13) and (7.14). This implies attempting to minimize variance as given in (7.43), subject to the specified technical constraints and with

$E(W_0 + z)$ varied parametrically over the relevant range. The utility-maximizing point in the efficient set could then be identified by the methods described in Section 7.3. Taking such an approach corresponding to (7.13), we have the quadratic risk programming formulation of the objective function as

$$\begin{aligned} s &= \beta[W_0 + E(x)] - V(z) \\ &= \beta\left[W_0 + 72x_1 + 53.4x_2 + 88.8x_3 - 200 - 5\int_0^w (w - v)f(v)\,dv\right] \\ &\quad - \left[\sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij}x_ix_j + 25\int_0^w (v - 79)^2 f(v)\,dv\right] \end{aligned} \tag{7.45}$$

Both the expected profit and variance of profit terms in this equation are nonlinear. However, the expected profit term is separable and hence causes no difficulty. But in its present form the variance term is both nonlinear and nonseparable and hence not directly amenable to quadratic risk programming methods. To overcome this difficulty, at least three approaches can be suggested:

1. The variance expression might be approximated by a quadratic function of w.
2. The variance expression could be made separable by ignoring covariance so that the expression for $V(z)$ is reduced to

$$V(z) = \sum_{j=1}^{3} \sigma_j^2 x_j^2 + 25\int_0^w (v - 79)^2 f(v)\,dv$$

   This would be justifiable only if it could be shown that the covariance terms had a negligible effect on total variance. A few calculations show that this is not so in the present case.
3. The function can be made separable by introducing six new variables

$$\begin{aligned} y_1 &= (x_1 + x_2)/2 & y_2 &= (x_1 - x_2)/2 \\ y_3 &= (x_1 + x_3)/2 & y_4 &= (x_1 - x_3)/2 \\ y_5 &= (x_2 + x_3)/2 & y_6 &= (x_2 - x_3)/2 \end{aligned}$$

   so that $y_1^2 - y_2^2 = x_1x_2$, $y_3^2 - y_4^2 = x_1x_3$, and $y_5^2 - y_6^2 = x_2x_3$.

The above expressions for the new variables can be written as six additional constraints, while the variance term can be rewritten in the separable form

$$V(z) = \sum_{j=1}^{3} \sigma_j^2 x_j^2 + \sigma_{12}(y_1^2 - y_2^2) + \sigma_{13}(y_3^2 - y_4^2)$$

$$+ \sigma_{23}(y_5^2 - y_6^2) + 25 \int_0^w (v - 79)^2 f(v)\,dv \qquad (7.46)$$

Substituting equation (7.46) for $V(z)$ in (7.45) yields a separable objective function to be maximized for values of β over the relevant range subject to the resource constraints on (7.41) and the transformed-variable constraints listed above.

To solve our nonsequential stochastic programming problem with a nonlinear utility function, we face the prospect of solving a problem with either a quadratic or a separable nonlinear objective function, subject to the usual linear constraints and to a separable nonlinear constraint that is to be varied parametrically. Choice between the above methods would be conditioned by the computer facilities available. Whatever algorithm is used is likely to be able to guarantee only a local optimum; some computational explorations (e.g., using different starting bases) would be needed before we could be confident that a satisfactory solution had been obtained.

The size of the programming matrix will be increased, perhaps substantially, in the attempt to obtain an adequate representation of the problem. The revised objective function incorporating $V(z)$ from equation (7.46), for example, includes ten variables. If each of these is represented by only five linear segments, the full separable programming matrix would require fifty variables—more than sixteen times as many as our initial nonstochastic problem. There would be an equivalent increase in the number of constraints. A proportionate increase in the size of a larger and more realistic farm planning matrix could clearly result in a massive and perhaps uneconomic if not insurmountable computational task.

In view of the difficulties noted above in finding a satisfactory mathematical programming approach to our stochastic planning problem when we introduce a nonlinear utility function, it would be reasonable to turn again to the Monte Carlo programming method of Section 7.5. Recall that we wish to maximize the nonlinear objective function defined by the moment-method approximation of equation (7.44), subject to the system of linear constraints given for (7.41). Monte Carlo programming can be applied to any form of objective function and so can readily be used in this case. The algorithm required would be similar to that described in Section 7.5 except that no integer constraints need be placed on the variables. The integrals in the right side of (7.44) could be represented by piecewise linear approximations. However, the increased complexity of the objective function in this case compared with the previous example makes hand calculation impracticable, and a special computer program would have to be developed.

It is obvious that the problem of risky resource constraints is complicated enough. But even more severe difficulties emerge if the technical input-output coefficients must also be treated as random variables. In our example this might arise in relation to the labor constraint if the rates of harvesting the three crops, and hence their per unit labor requirements, were also to be treated as stochastic. The first difficulty in such cases is the already familiar problem of assessing the relevant joint probability distribution of the a_{hj}. In itself this could be a heroic task. But even setting aside these difficulties, there is the additional complication that with random a_{hj}s, total input requirements across enterprises must also be assessed stochastically. Thus for each constraint containing random input-output variables a_{hj}, we need to be able to determine the density function of the random variables Y_h, $h = 1, 2, \ldots, m$, denoting total requirement of the hth input; i.e.,

$$Y_h = a_{h1}x_1 + a_{h2}x_2 + \cdots + a_{hn}x_n \tag{7.47}$$

The Y_h's are important because it is reasonable to suppose, as before, that the expected loss associated with each constraint will be some function of the deficit in resource supply, measured by $Y_h - b_h$ with $Y_h > b_h$. Whether it is possible to determine the density function of Y_h depends on the form of the density functions of the a_{hj}; but even if it is possible, as yet there is no nonlinear programming method capable of obtaining global optima for problems of this kind.

The nearest mathematical programming approach to such problems is *chance-constrained programming*. In this approach the objective function (such as expected profit) is optimized subject to a set of constraints of the form $P(Y_h \leqq b_h) \geqq p_h$ or, equivalently,

$$P\left(\sum_{j=1}^{n} a_{hj}x_j \leqq b_h\right) \geqq p_h \qquad h = 1, 2, \ldots, m \tag{7.48}$$

and $x_j \geqq 0$, where P is the probability operator and p_h is some critical probability level pertinent to total requirement of the hth resource. This formulation implies that each resource constraint, considered in turn, must at least be satisfied at its specified level of probability. Alternatively, the chance constraint may be formulated to require that any solution must at least satisfy all the constraints taken together at some required level of joint probability. In contrast to inequality (7.48), this implies a single constraint of the form

$$P(Y_1 \leqq b_1, Y_2 \leqq b_2, \ldots, Y_m \leqq b_m) \geqq p \tag{7.49}$$

The choice between the two formulations depends on what is required in relation to the particular decision situation being investigated. Note that under both approaches it is possible for any desired subset of resource coefficients to be regarded as nonstochastic.

The simplest chance-constrained programming models are those in which all but one of the constraints can be treated as deterministic. In the context of our simple farm planning example, this might arise if only the labor requirements per unit of each of the three crops were to be treated as stochastic. The chance-constrained programming formulation of this problem would involve the maximization of expected farm profit (defined and subject to the same technical constraints as before) except that the labor constraint would be rewritten in the form of the probability condition

$$P(d_1x_1 + d_2x_2 + d_3x_3 \leqq 80) \geqq p \tag{7.50}$$

where d_1, d_2, and d_3 are the stochastic labor demands per unit for standard wheat, oats, and new wheat respectively and p is some selected critical probability level.

Since we can anticipate that increasing the critical probability p will reduce farm profit, an alternative more tractable formulation of the problem is possible. We can maximize $P(d_1x_1 + d_2x_2 + d_3x_3 \leqq 80)$, subject to the remaining technical constraints and to a parametric constraint on profit of the form $\sum_{j=1}^{3} c_jx_j - F = z$. The problem would need to be solved for a range of relevant values of z until we obtained a solution that yielded the required value of p. The result would be optimal for the original stochastic programming model. Since all the constraints are now linear, all that remains is to find some programming method to handle the probability expression in the new objective function.

If we define

$$Y = d_1x_1 + d_2x_2 + d_3x_3 \tag{7.51}$$

and if Y is (approximately) normally distributed, it is clear that the probability that labor demand will not exceed supply will depend on both the mean and the variance of Y, represented by $E(Y)$ and $V(Y)$ respectively. Moreover, optimal farm plans must be $[E(Y), V(Y)]$-efficient since it will always be best to select a plan that gives the lowest possible variability of Y for a given value of $E(Y)$. By analogy with the quadratic risk programming model,

$$E(Y) = E(d_1)x_1 + E(d_2)x_2 + E(d_3)x_3 \tag{7.52}$$

$$V(Y) = \sum_{i=1}^{3} \sum_{j=1}^{3} \sigma_{ij}x_ix_j \tag{7.53}$$

where σ_{ij} is the covariance of the per unit labor requirements of activities i and j. Thus a computationally feasible approach, as demonstrated in a slightly different context by Chen (1973), would be to use quadratic programming analogously to equations (7.13) and (7.14) to minimize

$$s(Y) = E(Y) + \beta V(Y) \qquad \beta = 0 \rightarrow \infty \tag{7.54}$$

for a range of values for z in the profit constraint. As Chen shows, with a little manipulation of the results this method allows the optimal farm plan to be identified for any specified value of the critical probability p.

In general, chance-constrained programming becomes somewhat impractical if several stochastic constraints are to be accommodated and if the probability distributions involved are not easily tractable (Kirby, 1970). Moreover, from a decision theory viewpoint chance-constrained programming methods suffer from the same general weakness as the risk-constrained procedures for handling stochastic activity net revenues outlined in Section 7.4. Justification for such models lies largely in their tractability rather than in the realism with which they reflect the decision maker's true preferences. In this sense, chance-constrained programming is a crude and generally unsatisfactory method for whole-farm planning under risk. In particular, it suffers from the arbitrary choice of probability levels.

Sequential Stochastic Programming

Because of the nature of agricultural production, many farm decision problems (both micro and macro) are of a sequential stochastic nature in which later decisions are influenced both by earlier decisions and by stochastic parameters whose values become known after earlier decisions but before the later decisions. Such problems are not easy to handle. The difficulty stems from the fact that at least one of the sequentially related decisions, say the kth, cannot be fully specified until one or more of the random parameters in the system has been observed. For the decision variable x_k to be optimal, it must be based on an analysis of the problem including the observed values of the random variables. This means that, ex ante, the optimal value of x_k will not simply be a number but will be a strategy specified as a function of one or more of the random variables observed before decision k is taken. None of the mathematical programming methods currently available is capable of solving such problems. There is, however, one approximation method that has been developed and applied to a number of agricultural problems. This is the method of discrete stochastic programming (Cocks, 1968; Rae, 1971a, 1971b).

Discrete stochastic programming can be viewed as a (usually linear) programming formulation of a decision tree. In such a formulation the essential feature is an explicit specification of the available acts and possible events in their proper time sequence. In the programming model, act forks are usually represented in terms of continuous decision variables, but event forks can be represented only in terms of a relatively small number of discrete outcomes. The objective function in the linear programming formulation is usually the maximization of expected profit, quadratic or separable programming being needed if we have a nonlinear utility function.

To give some of the flavor of the discrete stochastic programming method, we will consider a simplified example of a rancher who must decide how many steers to run and what purchased feed reserves (if any) to establish for the coming year before he knows for certain what amount of feed will be available from his pastures. If he finds that his farm is overstocked, he can either sell stock or buy extra feed. If the farm turns out to be understocked, he can sell surplus feed from his reserve at a lower price. We assume that the rancher's uncertainty about the feed situation in the coming year is adequately reflected by considering only the following three pasture feed states with the probabilities shown: good, 0.3; medium, 0.4; poor, 0.3. Steers are assumed to be purchased at the beginning of the year for \$100 per head, while feed purchased at that time costs \$35 per unit. It is supposed that the rancher can allocate no more than \$12,500 to these uses, the opportunity cost of his capital being 10%. Fixed costs are set at \$3500, and it is assumed initially that the rancher's utility function is linear. Other assumptions are summarized in Table 7.7.

The decision problem is represented by the decision tree of Figure 7.8, where acts that are clearly nonoptimal have been excluded and continuous decision variables are represented schematically by decision fans. The decision tree reflects the situation that in the event of a good feed year it is likely the rancher will be understocked and have surplus feed to sell; with a medium feed year he may either need to sell or to buy feed, or he may sell steers; and if the feed year is poor, he is likely to be overstocked and will need to buy feed or sell steers.

The corresponding discrete stochastic linear programming matrix is shown in Table 7.8. The matrix is constructed to permit maximization of expected net revenue before deducting the fixed costs. The first two activi-

TABLE 7.7. Further Assumptions for Discrete Stochastic Linear Programming Example

Item	Good Year	Medium Year	Poor Year
Feed available from pasture (units)	100	75	50
Feed required per steer (units)	0.8	1.0	1.2
Price of fat steer (\$)	160	170	180
Selling price of surplus feed (\$)	27.5	30	...
Purchase price of feed (\$)	...	40	60
Selling price of surplus steers (\$)	...	95	75

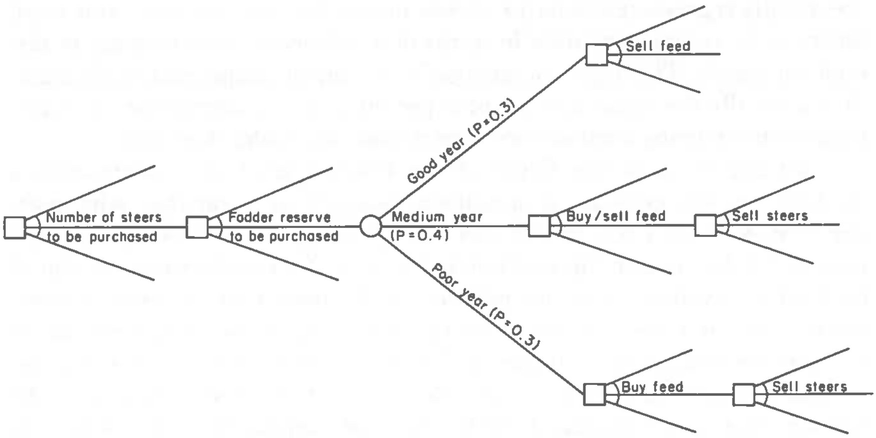


FIG. 7.8. Decision tree representation of the sequential stochastic programming problem.

ties in the matrix, "buy steers" and "buy feed," respectively represent the first two decision fans of Figure 7.8. The entries −110 and −38.5 in the objective row show the costs (including interest on capital) of buying a steer or of establishing a unit of feed reserve respectively. The first constraint represents the competition between these activities for the limited capital stock of $12,500. The capital requirements are shown to be $100 for a steer and $35 for a unit of feed respectively, so that this constraint can be written out in full as $12{,}500 \geqq 100x_1 + 35x_2$.

In the event of a good year, the second constraint links the number of steers fattened (activity 3) to the number available (activity 1), while the third constraint links the feed requirements for fattening steers (0.8 units per head) to the feed available directly from pasture (100 units) and from the feed reserve established (activity 2). Surplus feed reserves can be sold (activity 4). Note that a negative a_{hj} coefficient usually implies the production or purchase of some resource or commodity, while positive coefficients measure consumption per unit level of the activity. The entry for activity 3 in the objective function is the return from fattening a steer in a good year, multiplied by the probability of a good year; i.e., (160) (0.3) = $48. Similarly, for activity 4 surplus feed is sold in a good year for $27.5 per unit, so the entry in the objective function is (27.5) (0.3) = $8.25.

Corresponding interpretations can be placed on the entries in the submatrices for medium and poor years. Note that activities 6 and 10 represent the purchase of feed to make good a deficit, while activities 8 and 11 provide for the sale of steers that cannot be fattened because of lack of feed.

On solving the problem, it turns out that expected profit is maximized by allocating all the available capital to purchase of steers, establishing

TABLE 7.8. Matrix for Discrete Stochastic Linear Programming Example

	Resource Stock or Balance Identity	Form of Constraint	Current Decisions		Good Year		Medium Year				Poor Year		
			Buy Steers	Buy Feed	Feed Steers	Sell Feed	Feed Steers	Buy Feed	Sell Feed	Sell Steers	Feed Steers	Buy Feed	Sell Steers
Activity no.:			1	2	3	4	5	6	7	8	9	10	11
E (Net revenue)	...	...	−110	−38.5	48.0	8.25	68.0	−16.0	12.0	38.0	54.0	−18.0	22.5
Capital	12,500	≧	100	35	...	...	...	...	...	...	...	...	...
Good year:													
Steers	0	≧	−1	...	1	...	...	...	...	...	...	...	...
Feed	100	≧	...	−1	0.8	1	...	...	...	...	...	...	...
Medium year:													
Steers	0	≧	−1	...	...	...	1	...	...	1	...	...	...
Feed	75	≧	...	−1	...	...	1	−1	1	...	...	...	...
Poor year:													
Steers	0	≧	−1	...	...	...	...	...	...	...	1	...	1
Feed	50	≧	...	−1	...	...	...	...	...	...	1.2	−1	...

no feed reserve. The 125 steers purchased is exactly the number that can be fattened in a good year on pasture. In medium and poor years it pays the rancher to buy feed to make up the deficit. The expected net revenue of following this policy is \$4900, implying an expected profit of 4900 − 3500 = \$1400, distributed as: good year = \$2750, medium year = \$2000, and poor year = −\$750.

We now drop the assumption of a linear utility function and assume instead that the rancher is the owner of the risk-averse utility function $U = W^{0.8}$, where W is his terminal wealth at the end of the year. Terminal wealth is calculated as initial wealth W_0, assumed for the purpose of our example to be \$1000, plus profit z from the cattle fattening enterprise. The corresponding programming problem now incorporates the nonlinear utility function

$$U = 0.3(W_1^{0.8}) + 0.4(W_2^{0.8}) + 0.3(W_3^{0.8}) \tag{7.55}$$

where W_1, W_2, and W_3 denote the rancher's terminal wealth after good, medium, and poor years respectively. Thus

$$\begin{aligned}
W_1 &= 1000 - 110x_1 - 38.5x_2 + 160x_3 + 27.5x_4 - 3500 \\
W_2 &= 1000 - 110x_1 - 38.5x_2 + 170x_5 - 40x_6 + 30x_7 \\
&\quad + 95x_8 - 3500 \\
W_3 &= 1000 - 110x_1 - 38.5x_2 + 180x_9 - 60x_{10} \\
&\quad + 75x_{11} - 3500
\end{aligned} \tag{7.56}$$

To solve the problem by mathematical programming, the three new variables W_1, W_2, and W_3 are added to the matrix, defined by constraints derived from equation (7.56). The objective function of (7.55) is nonlinear but is both separable and concave, so an approximate global optimum can be found. The utility-maximizing solution, adjusted to the nearest integer number of steers, requires the purchase of 114 steers and 31.4 units of feed reserve. In a good year surplus feed is sold, while in both medium and poor years deficits are made good by purchasing extra feed. The expected profit from following this strategy is \$1344, distributed as follows: good year = \$2097, medium year = \$1827, and poor year = −\$53.

Two observations can be made about this particular example. First, it took some care in choice of coefficients to produce a different solution when a nonlinear utility function was introduced. This can be attributed mainly to the relatively crude model used. Quite large shifts in the objective function appear to be necessary before an alternative point on the set of feasible solutions becomes optimal. Thus using the same utility function but assuming a more plausible level of initial wealth of say \$10,000 or more results in the same solution as for a linear utility function. Second, even when a change of solution is obtained, the difference between the policy that maximizes expected profit and the policy that maximizes utility does not appear to be very great.

Utilities can be compared only in an ordinal way, but a comparison in terms of certainty equivalents does have some meaning since they are invariant for any positive linear transformation of the utility function. In these terms and for the utility function used, the terminal wealth certainty equivalent of the utility-maximizing policy is \$2299 compared with \$2269 for the policy that maximizes expected profit—an increase of only 1.3%.

Although we cannot say how much importance the rancher attaches to the difference, we might still question how important it is to account for nonlinear preference in discrete stochastic programming formulations, bearing in mind the associated extra computational burden. It seems likely that for decision makers with a not very marked risk aversion or preference, it is most important that the model used should take explicit account of the uncertainties inherent in a given decision problem and that a relatively crude, even linear, representation of the utility function may be adequate. Another possibility is that the decision maker's risk attitudes may be adequately reflected by maximizing expected profit subject to linear risk constraints of the kinds discussed previously.

Although very simple, the above example is sufficient to suggest the problems encountered in more comprehensive and realistic applications of discrete stochastic programming. It will be observed from Table 7.8 that a separate submatrix is required for each set of decision variables following each event. If the number of events considered was to be increased, the size of the matrix would increase considerably. Moreover, in this example only one event fork was needed (see Figure 7.8), whereas most real farm decision problems involve a dynamic sequence of acts and events. The implications of such a sequence in terms of size of matrix can be disastrous; models can quickly reach unmanageable proportions as they run into what is known as the "curse of dimensionality."

The decision analyst faced with these sorts of difficulties has no alternative but to simplify the planning problem. The exact way in which such simplifications can best be made will obviously depend on the particular circumstances, but some general approaches can be suggested.

Decision alternatives that are subjectively judged (perhaps on the basis of preliminary analyses) to be suboptimal can be eliminated. Similarly, random variables for which the associated risks are judged to be small can be treated as deterministic. Alternatively, random variables that are essentially continuous in nature might be approximated by discrete distributions including only two or three possible values. For example, prices might be treated as either "high" or "low"; weather uncertainty might be reflected by "good," "average," and "poor" categories.

In sequential decision problems the length of the planning period might be cut short. The relevant guiding principle is that the planning period should be just long enough to permit proper first-stage decisions to be made. Subsequent decisions can be revised later on a rolling or adaptive planning basis in the light of the then current situation; they are included in the present planning model only because they bear upon the immediate decisions. Extending this principle, we can see that the degree of precision required in modeling future decisions and events is only that which will permit a proper evaluation of the first-stage decisions. Often this means that quite crude representations of decisions and events remote from current decisions can be justified. Sometimes it might be possible to decompose a sequential decision problem into two or more stages, each of which can be analyzed separately from the other. In this way the size of the decision model may be materially reduced, although the overall amount of analytical work required to obtain a solution may not be reduced, and may even increase. Decomposition is particularly important for very large problems involving models that in their entirety are not within the capacity of available computers.

By simplification a previously intractable decision problem may be reduced to a form that can be tackled by one of the programming methods discussed earlier in this chapter. Alternatively, it may be more appropriate to abandon the so-called optimizing facility embodied in programming models and to adopt instead some "trial and error" approach. For any reasonably complex decision problem this will probably imply the construction of a computer simulation model of the system under study. Simulation breaks out of the straightjacket on the form of model inherent in all mathematical programming formulations and permits the analyst to represent the real system in whatever way he judges to be most appropriate. As illustrated in Chapter 8, experiments can be performed on such simulation models to investigate the effect of different decision rules. Costs of model development and analysis can be heavy, and it is a matter of assessing for each particular problem or class of problems how much is worth spending on decision analysis.

7.7 SUMMARY REMARKS

We have noted that mathematical programming models provide perhaps the best representation of the whole-farm planning problem. For nonsequential stochastic problems linear programming was shown to be wholly adequate for the case where only the c_j coefficients are uncertain and the utility function is linear. When the second of these requirements is relaxed, quadratic risk programming probably provides the most accurate means of analysis if adequate computing facilities are available. More generally, however, the computational complexities of quadratic programming are best avoided by using the MOTAD method of linear risk programming. For the nonsequential case with stochastic b_h coefficients we saw that (depending on the form of the expected loss functions) separable programming could be used to approximate either a local or a global optimal solution for a linear utility function. We saw too the difficulties encountered in attempting to extend such an analysis to incorporate a nonlinear utility function combined with stochastic c_j coefficients. The problems of dealing with stochastic a_{hj} coefficients were noted, brief mention being made of the possible use of chance-constrained programming in this context. Finally, we directed attention to the scope and limitations of discrete stochastic programming methods for the resolution of sequential farm planning problems under risk. These methods, though often crude, can provide a means of taking some account of the risks associated with most farm decision making. Despite the limitations imposed by the "curse of dimensionality," our view is that a crude representation of risk is better than ignoring it altogether.

In our discussion of stochastic programming methods we have drawn attention to the limited scope and power of these optimizing techniques. It is very probable that developments of both more powerful algorithms and improved computing facilities will extend the range of problems that can be solved by mathematical programming methods. Nevertheless, for many stochastic farm planning problems it may be sensible to turn away from mathematical programming and to look instead to methods of finding approximate optima such as Monte Carlo programming and simulation. Both the latter techniques have had a shorter history of application, so that important methodological advances can be expected.

PROBLEMS

7.1. (a) Consider Figure 7.2. Derive expressions for the mean and variance of portfolios consisting of q_G units of G and $Z - q_G$ units of D. Show that these expressions imply DG is linear in $(E, V^{1/2})$ space.

(b) Historical data from Taiwan (Cheung, 1969, p. 71) indicates the following approximate means and standard deviations for wheat and rice yields in kg/ha, 1901–1950:

Crop	Mean	Standard Deviation
Wheat	800	700
Rice	1500	320

Argue the relevance of this data to the fact that over the period concerned share rather than cash leasing arrangements were much more common for wheat production than for rice.

7.2. Suppose there are three risky prospects A, B, and C available. Given the data listed below, graph the set of efficient portfolios, assuming that borrowing is not feasible.

$$e_a = 0.10 \quad \sigma_a = 0.06 \quad \rho_{ab} = 0.4$$
$$e_b = 0.07 \quad \sigma_b = 0.04 \quad \rho_{bc} = -1.0$$
$$e_c = 0.03 \quad \sigma_c = 0.01 \quad \rho_{ac} = -0.4$$

7.3. Using the information of Table 7.2 and Figure 7.4, graphically determine the optimal (E, V) portfolio for the three-crop problem of Section 7.3 for each of the following utility functions:

$$U = (1000 + z)^{0.8}$$
$$U = (1000 + z)^{1.2}$$
$$U = \log_e (1000 + z)$$
$$U = 1 - \exp\,[-0.01\,(1000 + z)]$$

7.4. Show how the quadratic risk programming problem solved in Section 7.3 could be tackled by separable programming. Use the method discussed in the text for deriving $x_i x_j$ as the difference between the squares of two new variables. What significance can be attached to the fact that all the covariance terms are positive for this example?

7.5. Discuss the pros and cons of using risk-constrained linear programming. Is any particular form of safety constraint more warranted than other forms? What relationships are there between safety constraints and the continuity axiom, multidimensional utility, and lexicographic orderings?

7.6. Describe how Monte Carlo programming might be applied to a simplified farm planning problem in which some of the technical input-output coefficients are random variables. The flow chart for the required computer program should be drawn. If possible, illustrate your proposals with a simple numerical example.

7.7. Grub Free Grub, a large-scale vegetable producer, can produce n different crops using m resources (land, labor, etc.). Growing 1 ha of crop j requires a_{hj} units of resource h, while b_h units of that resource are available. The yield per hectare of crop j is y_j cases, and variable costs per hectare are c_j. Grub Free sells most of its crops to Woles Supermarkets at a price of p_j per case. However, Grub Free cannot predict precisely what the Woles demand will be. Instead q_j, the quantity demanded of product j, must be treated as an independent random variable with subjectively assessed density $f_j(q_j)$. If the demand for product j turns out to be more than supply, there is a goodwill loss of d_j for each case demanded which cannot be supplied. On the other hand, if supply exceeds demand, the surplus must be sold in the city market at a reduced price of g_j per case. Assume that Grub Free wishes to maximize its expected profit. Show how the problem can be formulated for solution by separable programming.

7.8. Formulate the cattle purchase decision problem represented in decision tree format in Figure 5.5 as a discrete stochastic linear programming problem. Assume a linear utility function. If possible, solve the problem, preferably by computer, and discuss the results you obtain. What mathematical programming methods might be used to solve the same problem but with a nonlinear utility function? Distinguish between utility functions showing risk aversion and functions showing risk preference.

7.9. Show how your matrix in Problem 7.8 would be modified to find the expected value of perfect information. Again, if possible, solve by computer for the linear utility case. Suggest how the EVPI could be computed using mathematical programming methods when the decision maker is risk averse. Similarly, how might mathematical programming methods be used for such a decision maker to find the maximum price that could be paid for a less than perfect predictor?

7.10. Using the empirical covariance matrix of Section 7.3, find the (E, V) locus of the solutions to the risk-constrained linear programming problem of Table 7.3. Plot this locus along with the (E, V) frontier obtained by quadratic risk programming (Figure 7.4) and compare the two. What point on the risk-constrained locus is optimal for the utility function $U = x - 0.0005x^2$?

7.11. If you have access to suitable computer facilities, solve the example problem of risk-constrained linear programming using the Boussard approach for $k = 5$, and then solve again holding k at 5 but with r_j at two standard deviations. Discuss the significance of the results you obtain.

7.12. Extend the discrete stochastic programming example of Section 7.6 to incorporate a predictor that can be observed after the steers have been purchased but before the fodder reserve is established. The predictor, which costs \$100, can yield either a favorable forecast z_1 or an unfavorable one z_2, the likelihoods being $P(z_1 \mid \theta_1) = 0.8$, $P(z_1 \mid \theta_2) = 0.6$, $P(z_1 \mid \theta_3) = 0.3$, where θ_1, θ_2, and θ_3 respectively represent a good, medium, or poor year. Construct the matrix incorporating such a predictor and if possible solve the problem by computer for the linear utility case.

REFERENCES AND SELECTED FURTHER READING

Recent years have seen the development of a vast literature on portfolio analysis, particularly in relation to the stock market. With some overlap, the six references listed below discuss the general model of portfolio analysis, generally in an (E, V) format.

Hirshleifer, J. 1970. *Investment, Interest, and Capital*, Ch. 10. Englewood Cliffs, N.J.: Prentice-Hall.

Jean, W. H. 1970. *The Analytical Theory of Finance*. New York: Holt, Rinehart and Winston.

Markowitz, H. 1959. *Portfolio Selection—Efficient Diversification of Investments*. New York: Wiley.

Merton, R. C. 1972. An analytic derivation of the efficient portfolio frontier. *J. Financ. Quant. Anal.* 7(4): 1851–73.

Mossin, J. 1973. *Theory of Financial Markets*. Englewood Cliffs, N.J.: Prentice-Hall.

Sharpe, W. F. 1970. *Portfolio Theory and Capital Markets*. New York: McGraw-Hill.

The pros and cons of the (E, V) portfolio model and some alternatives are presented by:

Borch, K. 1969. A note on uncertainty and indifference curves. *Rev. Econ. Stud.* 36(1): 1–4.

Feldstein, M. S. 1969. Mean-variance analysis in the theory of liquidity preference and portfolio selection. *Rev. Econ. Stud.* 36(1): 5–11.

Hakansson, N. H. 1972. Mean-variance analysis in a finite world. *J. Financ. Quant. Anal.* 7(4): 1873–80.
Hanoch, G., and H. Levy. 1970. Efficient portfolio selection with quadratic and cubic utility. *J. Busin.* 43(2): 181–89.
Jean, W. H. 1973. More on multidimensional portfolio analysis. *J. Financ. Quant. Anal.* 8(3): 475–90.
Tobin, J. 1969. Comment on Borch and Feldstein. *Rev. Econ. Stud.* 36(1): 13–14.
Tsiang, S. C. 1972. The rationale of the mean-standard deviation analysis, skewness preference, and the demand for money. *Am. Econ. Rev.* 62(3): 354–71. Comments by K. Borch, G. O. Bierwag, and H. Levy; reply by S. C. Tsiang, *Am. Econ. Rev.* 64(3): 428–50.

The risk diversification aspect of portfolios is discussed in:

Heady, E. O. 1952. *Economics of Agricultural Production and Resource Use,* Ch. 17. New York: Prentice-Hall.
Johnson, S. R. 1967. A re-examination of the farm diversification problem. *J. Fm. Econ.* 49(3): 610–21.
Samuelson, P. A. 1967. General proof that diversification pays. *J. Financ. Quant. Anal.* 2(1): 1–12.
Tobin, J. 1965. The theory of portfolio selection. In *Theory of Interest Rates,* F. H. Hahn and F. P. Brechling, eds., pp. 3–51. New York: St. Martin's Press.

Some agricultural production and marketing applications of portfolio analysis are presented by:

Brainard, W. C., and R. N. Cooper. 1968. Uncertainty and diversification in international trade. *Fd. Res. Inst. Stud.* 8(3): 257–85.
Burt, O. R., and R. D. Johnson. 1967. Strategies for wheat production in the Great Plains. *J. Fm. Econ.* 49(4): 881–99.
Cheung, S. N. S. 1969. *Theory of Share Tenancy.* Chicago: Univ. Chicago Press.
Heifner, R. G. 1966. Determining efficient seasonal grain inventories: An application of quadratic programming. *J. Fm. Econ.* 48(3): 648–80.
Kunreuther, H. 1972. Risk-taking and farmers' crop growing decisions. Rept. 7219, Univ. Chicago Center Math. Stud. Bus. Econ.
Kunreuther, H., and G. Wright. 1974. Safety-first, gambling, and the subsistence farmer. Fels Discussion Paper 59. Univ. Pa. Fels Center of Government.
Porter, R. 1959. Risk, incentive, and the technique of the low-income farmer. *Indian Econ. J.* 7(1): 1–27.
Sadan, E. 1970. The investment behavior of a farm firm operating under risk. *Am. J. Agric. Econ.* 52(4): 494–504.
Sutinen, J. G. 1975. The rational choice of share leasing and implications for efficiency. *Am. J. Agric. Econ.* 57(4): 613–21.
Ward, R. W., and L. B. Fletcher. 1971. From hedging to pure speculation: A micro model of optimal futures and cash market positions. *Am. J. Agric. Econ.* 53(1): 71–78.

General references on mathematical programming are legion. A couple of the best ones are:

Gass, S. I. 1975. *Linear Programming: Methods and Applications,* 4th ed. New York: McGraw-Hill.
Hadley, G. 1964. *Non-Linear and Dynamic Programming.* Reading, Mass.: Addison-Wesley.

Coverage of most of the programming approaches we have outlined in this chapter is to be found in the following two texts, the second of which is encyclopedic, contains an extensive list of references, and is very highly recommended:

Hillier, F. S., and G. J. Lieberman. 1974. *Introduction to Operations Research,* 2nd ed. San Francisco: Holden-Day.

Wagner, H. M. 1975. *Principles of Operations Research,* 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.

The algorithm used in our example on quadratic risk programming is given in:

Wolfe, P. 1959. The simplex method for quadratic programming. *Econometrica* 27(2): 382–98.

Some early examples of the application of quadratic risk programming to farm planning are:

Camm, B. M. 1962. Risk in vegetable production on a fen farm. *Fm. Economist* 10(2): 89–98.

Freund, R. J. 1956. The introduction of risk into a programming model. *Econometrica* 24(2): 253–63.

Heady, E. O., and W. Candler. 1958. *Linear Programming Methods,* Ch. 17. Ames: Iowa State Univ. Press.

McFarquhar, A. M. M. 1961. Rational decision making and risk in farm planning—An application of quadratic programming in British arable farming. *J. Agric. Econ.* 14(4): 552–63.

Using a quadratic risk programming approach to derive the (E, V) frontier facing the farmer, some empirical evidence that simple utility maximization performs better than a lexicographic utility or a profit maximization approach as a predictor of farmer behavior is presented by:

Lin, W., G. W. Dean, and C. V. Moore. 1974. An empirical test of utility vs. profit maximization in agricultural production. *Am. J. Agric. Econ.* 56(3): 497–508.

General discussions of linear risk programming, including the use of safety constraints in linear programming models for farm planning, are provided in:

Hanf, E. 1971. Development of a decision rule for farm planning under uncertain conditions. *Papers and Reports, 14th Int. Conf. Agric. Economists,* pp. 325–38. Oxford: Inst. Agr. Aff.

Hazell, P. B. R., and R. B. How. 1971. Obtaining acceptable farm plans under uncertainty. *Papers and Reports, 14th Int. Conf. Agric. Economists,* pp. 338–47. Oxford: Inst. Agr. Aff.

Kennedy, J. O. S., and E. M. Francisco. 1974. On the formulation of risk constraints for linear programming. *J. Agric. Econ.* 25(2): 129–44.

Maruyama, Y., and T. Kawaguchi. 1971. An approach to farm planning under "ambiguity." *Papers and Reports, 14th Int. Conf. Agric. Economists,* pp. 347–55. Oxford: Inst. Agr. Aff.

Examples or discussions of the incorporation of game theory decision criteria into linear programming models for farm planning include:

Hazell, P. B. R. 1970. Game theory—An extension of its application to farm planning under uncertainty. *J. Agric. Econ.* 21(2): 239–52.

Kawaguchi, T., and Y. Maruyama. 1972. Generalized constrained games in farm planning. *Am. J. Agric. Econ.* 54(4): 591–602.

Low, A. R. C. 1974. Decision taking under uncertainty: A linear programming model of peasant farmer behaviour. *J. Agric. Econ.* 25(3): 311–22.

McInerney, J. P. 1967. Maximin programming—An approach to farm planning under uncertainty. *J. Agric. Econ.* 18(2): 279–90.

———. 1969. Linear programming and game theory models—Some extensions. *J. Agric. Econ.* 20(2): 269–78.

Maruyama, Y. 1972. A truncated maximin approach to farm planning under uncertainty with discrete probability distributions. *Am. J. Agric. Econ.* 54(2): 192–200.

Game theory decision criteria are reviewed in relation to the axioms of rational choice in:

Officer, R. R., and J. R. Anderson. 1968. Risk, uncertainty and farm management decisions. *Rev. Mktg. Agric. Econ.* 36(1): 3–19.

The maximum admissible loss approach is described in:

Boussard, J-M. 1971. A model of the behaviour of farmers and its application to agricultural policies. *European Econ. Rev.* 2(4): 436–61.
Boussard, J-M., and M. Petit. 1967. Representation of farmers' behavior under uncertainty with a focus-loss constraint. *J. Fm. Econ.* 49(4): 869–80.

Applications, discussions, and evaluations of the mean absolute deviation approach are to be found in:

Hazell, P. B. R. 1971. A linear alternative to quadratic and semivariance programming for farm planning under uncertainty. *Am. J. Agric. Econ.* 53(1): 53–62. Comment by J. T. Chen and reply by P. B. R. Hazell. *Am. J. Agric. Econ.* 53(4): 662–65.
Hazell, P. B. R., and P. L. Scandizzo. 1974. Competitive demand structures under risk in agricultural linear programming models. *Am. J. Agric. Econ.* 56(2): 235–44.
Schluter, M. G. G. 1974. The interaction of credit and uncertainty in determining resource allocation and incomes of small farms, Surat District, India. Occasional Paper 68, Employment and Income Distribution Project, Dept. Agric. Econ. Cornell Univ., Ithaca.
Thomson, K. J., and P. B. R. Hazell. 1972. Reliability of using the mean absolute deviation to derive efficient *E, V* farm plans. *Am. J. Agric. Econ.* 54(3): 503–6.

In contrast to the above approaches, a risk-constrained programming approach taking account of covariance relationships and with the desirable constraint that no activity can be activated beyond the level at which its marginal utility is zero is presented by:

Chen, J. T., and C. B. Baker. 1974. Marginal risk constraint linear program for activity analysis. *Am. J. Agric. Econ.* 56(3): 622–27.

Discussion of the application of Monte Carlo programming is provided by:

Carlsson, M., B. Hovmark, and I. Lindgren. 1969. A Monte Carlo method study of farm planning problems. *Rev. Mktg. Agric. Econ.* 37(2): 80–103.
Donaldson, G. F., and J. P. G. Webster. 1968. An operating procedure for simulation farm planning—Monte Carlo method. Wye College (Univ. London), Dept. Agric. Econ.
Thompson, S. C. 1967. An approach to Monte Carlo programming. Univ. Reading (England), Dept. Agric. Fm. Mgmt. Sect., Study 3.

The merits of Monte Carlo programming vis-a-vis mixed-integer programming are discussed in:

Candler, W., W. Cartwright, and J. B. Penn. 1973. The substitution of analytic for simulation algorithms: A comment. *Am. J. Agric. Econ.* 54(2): 235–39; response by S. C. Thompson, pp. 240–41.

Separable programming is used to take account of risk in:

Thomas, W., L. Blakeslee, L. Rogers, and N. Whittlesey. 1972. Separable programming for considering risk in farm planning. *Am. J. Agric. Econ.* 54(2): 260–66.

Two seminal papers on chance-constrained programming are:

Charnes, A., and W. W. Cooper. 1959. Chance constrained programming. *Mgmt. Sci.* 6(1): 73–79.
______. 1963. Deterministic equivalents for optimizing and satisficing under chance constraints. *Ops. Res.* 11(1): 18–39.

Review and formal exposition of the method is provided by:

Blau, R. A. 1974. Stochastic programming and decision analysis: An apparent dilemma. *Mgmt. Sci.* 21(3): 271–76.
Jagannathan, R. 1974. Chance-constrained programming with joint constraints. *Ops. Res.* 22(2): 358–72.
Kirby, M. J. L. 1970. The current state of chance-constrained programming. In *Proc. Princeton Symp. Math. Program,* H. W. Kuhn, ed., pp. 93–110. Princeton: Princeton Univ. Press.
Prekopa, A. 1970. On probabilistic constrained programming. In *Proc. Princeton Symp. Math. Program.,* H. W. Kuhn, ed., pp. 113–38. Princeton: Princeton Univ. Press.

An agricultural application of chance-constrained programming is given by:

Chen, J. T. 1973. Quadratic programming for least-cost feed formulations under probabilistic protein constraints. *Am. J. Agric. Econ.* 55(2): 175–83.

Agricultural applications of discrete stochastic programming are:

Cocks, K. D. 1968. Discrete stochastic programming. *Mgmt. Sci.* 15(1): 72–79.
Rae, A. N. 1971a. Stochastic programming, utility, and sequential decision problems in farm management. *Am. J. Agric. Econ.* 53(3): 448–60.
______. 1971b. An empirical application and evaluation of discrete stochastic programming in farm management. *Amer. J. Agric. Econ.* 53(4): 625–38.

Simplifications of sequential stochastic decision problems to achieve computational feasibility are reported in:

Trebeck, D. B., and J. B. Hardaker. 1972. The integrated use of simulation and stochastic programming for whole farm planning under risk. *Australian J. Agric. Econ.* 16(2): 115–26.
Yaron, D., and U. Horowitz. 1972. A sequential programming model of growth and capital accumulation of a farm under uncertainty. *Am. J. Agric. Econ.* 54(3): 441–51.

The issue of required length of time horizon for long-run farm planning under uncertainty is taken up in:

Boussard, J-M. 1971. Time horizon, objective function, and uncertainty in a multiperiod model of firm growth. *Am. J. Agric. Econ.* 53(3): 467–77.

The application of risk programming at the regional or macro level using a mean-standard deviation formulation of the objective function so as to maintain linearity is outlined by:

Hazell, P. B. R., R. D. Norton, and M. Parthasarathy. 1976. The economic significance of risk in agricultural planning. In *Programming Studies for Agricul-*

tural Sector Policy, L. M. Bassoco, R. D. Norton, and J. S. Solis, eds. Oxford: Oxford Univ. Press.

Hazell, P. B. R., and P. L. Scandizzo. 1976. Optimal price intervention policies when production is risky. Development Research Center, World Bank, Washington.

Some of the early uses of simulation methods for whole-farm planning under risk are reported in:

Halter, A. N., and G. W. Dean. 1965. Use of simulation in evaluating management strategies under uncertainty. *J. Fm. Econ.* 47(3): 557–73.

Zusman, P., and A. Amiad. 1965. Simulation: A tool for farm planning under conditions of weather uncertainty. *J. Fm. Econ.* 47(3): 574–94.

The general concepts and procedures of agricultural systems analysis and simulation are covered by:

Anderson, J. R. 1974. Simulation: Methodology and application in agricultural economics. *Rev. Mktg. Agric. Econ.* 42(1):3–55.

Dalton, G., ed. 1975. *Study of Agricultural Systems.* London: Applied Science Publishers.

Dent, J. B., and J. R. Anderson, eds. 1971. *Systems Analysis in Agricultural Management.* Sydney, Australia: Wiley.

CHAPTER EIGHT
INVESTMENT APPRAISAL

In Chapters 6 and 7 we considered the questions of production and whole-farm planning under risk. We avoided the question of time influences by assuming a short (at most one-year) decision horizon. The problem of time influence due to consequences occurring at different times and in different sequences will be discussed in this chapter in the context of investment appraisal.

By an investment we mean the outlay of resources over one or more time periods in the expectation that returns on this outlay will be received over several future time periods. As we will show, while there is adequate theory to handle the problem of time influences, this theory does not as yet appear to be easily translatable into techniques of generally practical usefulness for the normative appraisal of risky investments. At best, for all but relatively simple problems, the theory can only be approximated in practical application.

Throughout our discussion of investment appraisal, to keep the analysis as simple as possible while still maintaining practicality, we will assume discrete time increments of years and will use money as a measure of resource outflows and inflows associated with an investment. These cash flows (which may be zero in some periods) are assumed to occur at time zero and at the end of each year thereafter up to the cessation of the investment. The major symbols to be used are:

t = time subscript, $t = 0,1,2,\ldots,T$
i = project subscript, $i = 1,2,\ldots,I$
r_t = compound rate of interest applicable in period t, $r_t \geqq 0$
C_t = net cash flow at time t
C_{it} = net cash flow per unit of the ith project at time t
α_t = discount factor applicable in period t; $\alpha_t = 1/(1 + r_t)$, $0 \leqq \alpha_t \leqq 1$
PV = net present value, i.e., discounted value at $t = 0$ of a stream of net cash flows

$(PV)_i$ = net present value of the ith project
$(PV)_t$ = present value of the total cash flow occurring at time t
x_i = proportion of the ith project undertaken;
B_t = budget constraint at time t
T = life span of the longest lived project under consideration
$U(\cdot)$ = decision maker's utility function for $(\cdot)$

8.1 SINGLE PROJECT WITH NO RISK

Suppose a decision is to be made as to whether to undertake a particular investment project of fixed size for which (though it is hard to imagine!) there is no risk. The project has a life of T years and an associated sequence of sure *net cash flows* $C_0, C_1, C_2, \ldots, C_T$. These cash flows are the net result of outflows and inflows associated with the project. Just what items should be counted in practice in assessing these outflows and inflows will not be discussed here; such questions are well covered in many of the practical manuals on agricultural project appraisal, e.g., FAO (1971), Gittinger (1972), Little and Mirrlees (1974), Whitlam et al. (1970). Note, however, that C_0 will typically involve initial outlay on the project and C_T will include any salvage or other terminal value for the investment.

A multitude of different criteria has been suggested to appraise such single-project investment possibilities. Many of them are used in practice under an assumption of no risk. Best known of these methods are the criteria based on net present value, internal rate of return, payback period, benefit-cost ratio, and average rate of return. Detailed discussion of these various procedures is to be found in many of the standard texts on project appraisal or financial management. Only two of the methods have any claim to normative respectability—those based on net present value and the internal rate of return. Of these two, for the reasons argued by Harberger (1972), Jean (1970), and others, the net present value approach is best and is to be preferred.

For a single project the essence of the *net present value approach* is that the project should be accepted if its net present value, given by

$$PV = C_0\alpha_0 + C_1\alpha_1 + C_2\alpha_1\alpha_2 + \ldots + C_T\prod_{t=0}^{T}\alpha_t \tag{8.1}$$

is greater than zero. Note that the discount factor $\alpha_t = 1/(1 + r_t)$ and that since r_0 is zero, $\alpha_0 = 1$. Usually it is assumed that the interest rate is constant over all subsequent periods ($r_t = r$ for all $t \geqq 1$) so that the discount factor for each such period is also a constant ($\alpha_t = \alpha$) and equation (8.1) collapses to the Tth degree polynomial:

$$PV = C_0 + \sum_{t=1}^{T} C_t\alpha^t = \sum_{t=0}^{T} C_t\alpha^t \tag{8.2}$$

The assumptions underlying the net present value approach are that there is no risk, that the decision maker has perfect liquidity, and that the capital market is perfect. These assumptions imply that investment and transaction opportunities are available such that any cash flow stream can be converted to any other of the same present value. A traditional point of dispute about the present value approach is the choice or interpretation of the interest rate to be used. Suggestions made include:

1. The borrowing rate, i.e., the compound rate of interest or cost of capital actually prevailing if funds were to be (or are) borrowed.
2. The opportunity cost, i.e., the rate of interest that could be earned in the most attractive alternative investment (of equivalent risk if risk is assumed).
3. The subjective rate of time preference, i.e., the rate that the decision maker considers appropriate for discounting future net flows to their current value.

In determining PV, we will assume the interest rate to be the borrowing rate or cost of capital. The opportunity cost of capital is automatically taken into account in the multiple-project procedures based on the cost of capital that we will consider in Sections 8.2 and 8.3. Subjective time preference is relevant in utility appraisal of time sequences, but not for present value procedures because these depend on the assumption of an idealized investment market.

The attractiveness of the net present value approach is that it reduces a project's time sequence of net cash flows to a single cash value, which under the assumptions made provides a basis for assessment of the project. However, from a decision theory point of view, even in the absence of risk the present value approach contains a major flaw in that it fails to allow adequately for preferences the decision maker may have about the actual pattern of net cash flows over time. We will take up this question in the following section.

8.2 MULTIPLE PROJECTS WITH NO RISK

By a multiple-project situation we mean one in which the decision maker faces the problem of choosing from or between a number of projects. Such situations may be classifiable in various ways. The projects may be divisible or indivisible; i.e., it may or may not be possible to vary a project's size. They may be mutually exclusive, independent, or interdependent.

The economic setting may be one of ample funds or, more typically, a situation of limited funds and capital rationing so that capital budgeting is necessary. We will not consider all these possibilities but only those of mutually exclusive projects (implying relatively simple appraisal procedures) and capital budgeting (implying more complicated appraisal procedures based on either enumeration or mathematical or dynamic programming).

Mutually Exclusive Projects

The simplest case of multiple-project appraisal is that involving choice between two mutually exclusive indivisible projects. By definition only one of the two projects can be undertaken. In present value terms, appraisal is clear-cut. The project with the highest positive net present value should be undertaken. If neither has a positive net present value, neither should be undertaken. The same principle extends to any number of mutually exclusive indivisible projects. Conveniently, any differences between the lengths of life of nonrepeatable projects are automatically handled in the present value approach. The implicit mechanism is that terminal funds from shorter lived projects are taken as being reinvested at the market rate of interest r.

The problem with the present value approach is that it fails to allow adequately for the time pattern of cash flows. Consider the two net cash flow streams listed in Table 8.1. Assuming that a discount rate of 10% is the relevant cost of capital, both these project flows have the same net present value (approximately $105,470). In present value terms a decision

TABLE 8.1. Net Cash Flow Streams for Two Mutually Exclusive Projects of Equal Present Value When $r = 0.10$

	Net Cash Flow ($)	
Year (t)	Project A	Project B
0	−20,000	−20,000
1	60,000	0
2	40,000	0
3	30,000	0
4	20,000	0
5	2,700	0
6	0	20,000
7	0	50,000
8	0	60,000
9	0	70,000
10	0	80,000
PV	105,470	105,470

maker should be indifferent between them. Many if not a majority of decision makers would not find them equivalent. One anxious for early returns would obviously prefer cash flow stream *A*. But one anxious about prospects for obtaining further capital in 10 years time would obviously prefer stream *B* with its record of rising profits and apparent indication of a rosy outlook. To hold such varying assessments about cash flow streams of equal net present value would be irrational if the present value assumptions of no risk, perfect liquidity and a perfect capital market were true. If these assumptions hold, any cash flow stream could by converted by appropriate transactions to any other of the same present value. In reality this cannot be done to anywhere near the degree of perfection implied by the present value approach. For instance, consider a farmer who invests in a large dam and an associated irrigation system on part of his land. There is no way he could trade this investment per se for some other of equivalent present value.

Utility appraisal in multidimensional terms provides the only theoretically sound basis (apart from intuition in obvious cases) for comparing alternative cash flow streams. It is obvious, as soon as one recognizes the fact, that any time sequence of cash flows corresponds to a multidimensional outcome where each time period corresponds to a different dimension. Thus for purposes of comparison between them, the net cash flow streams of Table 8.1 could just as logically be labeled as tabulated in Table 8.2.

Some of the more practical procedures available for assessing the utility of multiattributed alternatives have been outlined in Section 4.3. If their assumptions are met, any of these procedures are appropriate to the utility assessment of alternative investments. For example, we can apply the

TABLE 8.2. Multidimensional Payoffs for Two Alternative Acts

Dimension	Act A	Act B
	($)	
1	−20,000	−20,000
2	60,000	0
3	40,000	0
4	30,000	0
5	20,000	0
6	2,700	0
7	0	20,000
8	0	50,000
9	0	60,000
10	0	70,000
11	0	80,000

benchmark approach to the alternatives of Table 8.1. In doing so, we make the reasonable assumption of preferential independence between periods. This implies, for example, that if the cash flow sequence (−2, 3, 10) is preferred to (−4, 3, 10), then any three-period sequence with −2 in the first period, e.g., (−2, −10, 30), will be preferred to any other three-period sequence with −4 in the first period but the same flows in the second and third periods, e.g., (−4, −10, 30). Accordingly, following the benchmark procedure of Section 4.3 and assuming some hypothetical decision maker, for alternative A of Table 8.1 we might establish that indifference exists between the following cash flow streams expressed in hundreds of dollars:

(−200,	600,	400,	300,	200,	27,	0,	0,	0,	0,	0)
(−180,	600,	400,	300,	200,	0,	0,	0,	0,	0,	0)
(− 30,	600,	400,	300,	0,	0,	0,	0,	0,	0,	0)
(220,	600,	400,	0,	0,	0,	0,	0,	0,	0,	0)
(580,	600,	0,	0,	0,	0,	0,	0,	0,	0,	0)
(1150,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0)

In analogous fashion, for alternative B the actual cash flow stream and its elicited benchmark equivalent might be:

(−200,	0,	0,	0,	0,	0,	200,	500,	600,	700,	800)
(1075,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0)

Since neither alternative A nor B involves any risk, no further calculations to assess utility are required. Because $U(A) = U(\$115{,}000 \mid 0, 0, \ldots, 0)$ is necessarily greater than $U(B) = U(\$107{,}500 \mid 0, 0, \ldots, 0)$, our decision maker should choose project A despite the fact that both projects have the same net present value. Because of the obvious difference in the time patterns of the cash flows of A and B, most decision makers could immediately choose between them on the basis of intuition. It is only when the cash flows of mutually exclusive projects are not too dissimilar or involve risk that multidimensional utility appraisal may be necessary. In these cases, as with single-period decision problems that cannot be easily handled by intuition, decision analysis provides a mechanism for ensuring that a choice is indicated that is consistent with the decision maker's preferences (and beliefs if risk is present).

As Jean (1970) shows, only if the decision maker has a linear utility function (i.e., $U = \sum a_t C_t$) and his discount factor a_t/a_{t-1} for a one-period time shift in terms of utility is equal to the present value discount factor α_t will the present value and utility approaches be equivalent.

Capital Budgeting

Capital budgeting is necessary if the decision maker does not have access to ample funds; if some but not all of the set of available projects are

TABLE 8.3. Characteristics of Seven Indivisible Independent Projects in a Capital Budgeting Problem

Project (i)	C_{i0}	C_{i1}	C_{i2}	$(PV)_i$ $r = 0.10$
	($)	($)	($)	($)
A_1	−1000	1120	1250	1051.19
A_2	−1000	1200	1190	1074.34
A_3	−1000	1230	1150	1068.55
A_4	−1000	1100	1310	1082.58
A_5	−2000	2300	2300	1991.65
A_6	−2000	2200	2500	2066.00
A_7	−3000	3500	3200	2826.33

mutually exclusive; or if possible projects are divisible, postponable, interdependent, or otherwise interrelated. If the set of all possible investment combinations is not too large and the interdependencies not too complicated, capital budgeting may be carried out by enumerating all the possibilities and comparing them to find the best combination. When enumeration is not feasible, mathematical programming procedures are necessary.

As a simple example suppose that investment funds to a maximum of $3000 are currently available with $r = 0.10$ and that there are seven nonpostponable indivisible independent projects $A_1, A_2, \ldots, A_7$ available with the characteristics shown in Table 8.3. Using enumeration, we need to assess all possible project combinations involving an outlay of $3000 or less. These feasible combinations are listed with their present values in Table 8.4. Inspection shows that the best available combination (judged by present value) is the project set $\{A_2, A_3, A_4\}$.

In mathematical programming terms, using maximum net present

TABLE 8.4. Enumeration of Feasible Project Combinations from Table 8.3 When Investment Funds Are Limited to $3000

Project Combination	PV	Project Combination	PV	Project Combination	PV
	($)		($)		($)
1, 2, 3	3194	2, 3	2143	1	1051
1, 2, 4	3208	2, 4	2157	2	1074
1, 3, 4	3202	2, 5	3066	3	1068
2, 3, 4	3225	2, 6	3140	4	1083
1, 2	2125	3, 4	2151	5	1992
1, 3	2120	3, 5	3060	6	2066
1, 4	2134	3, 6	3134	7	2826
1, 5	3043	4, 5	3074		
1, 6	3117	4, 6	3149		

value as the criterion, our example could be most simply formulated as the integer programming problem, maximize

$$PV = \sum_{i=1}^{I} (PV)_i x_i \tag{8.3}$$

subject to the budget constraints

$$\sum_{i=1}^{I} C_{it} x_i \geqq -B_t \qquad t = 0, 1, \ldots, T \tag{8.4}$$

and the integer constraints

$$0 \leqq x_i \leqq 1, \quad x_i \text{ an integer} \qquad i = 1, 2, \ldots, I \tag{8.5}$$

where $(PV)_i$ is the net present value of the ith project, $x_i = 0$ implies the ith project is not undertaken, $x_i = 1$ implies the ith project is undertaken, C_{it} is the net cash flow of the ith project in period t, B_t is the budget constraint in the tth period, and T is the life span of the longest lived project in the set of I possible projects.

The integer condition of expression (8.5) implies that each project can only be rejected or accepted on an all-or-nothing basis. Constraint (8.4) states that the set of accepted projects must be such that the net outflow for all selected projects in year t cannot exceed the prespecified budget constraint B_t on expenditure for that year. Through the constraints of (8.4) and (8.5) the requirements of capital rationing and indivisibility are satisfied simultaneously in determining the set of x_i values that specifies the set of projects maximizing (8.3). Note that for our particular example of Table 8.3, $B_0 = 3000$ and $B_1 = B_2 = 0$. Since the projects are nonpostponable and all cash flows for $t = 1$ or 2 are positive, the only relevant budget constraint, assuming that borrowing is not permitted, is for $t = 0$. For a more general formulation (especially when budget constraints for subsequent periods must be considered) it is necessary to account for any borrowing and saving opportunities by including appropriate additional variables in the model.

The above programming formulation can be extended in various ways. To allow for mutually exclusive projects, it is only necessary to add constraints of the form

$$\sum_{i \subset M} x_i \leqq 1 \tag{8.6}$$

where M is a particular set of mutually exclusive projects. Thus if M consists of the three mutually exclusive projects A_1, A_3, and A_6, constraint (8.6) would be $x_1 + x_3 + x_6 \leqq 1$. Since $x_i = 0$ or 1, this expression would only allow at most one of the three projects to be included in the optimal set.

Likewise, if the acceptance of one project, say x_r, is conditional on the acceptance of another, say x_s, this requirement can be specified by additional constraints of the form

$$x_r \leqq x_s \tag{8.7}$$

so that only if the sth project is undertaken ($x_s = 1$) can the rth be adopted ($x_r = 1$). If projects are divisible so that x_i is not restricted to being an integer but only to $x_i \geqq 0$, where x_i is the fraction or multiple of the ith project undertaken, solution can proceed by way of ordinary linear programming without the computational complications of integer programming. If only some projects are divisible, mixed-integer programming is relevant.

As noted in the reading list, the seminal work on deterministic mathematical (and dynamic) programming of investment was that of Weingartner (1963). More recently, he and others—see Weingartner (1966) and Bernhard (1969)—have extended the procedures to allow for a great variety of additional considerations such as postponability, interperiod cash balance restrictions, payback restrictions, scarce materials, terminal wealth requirements, borrowing limits, increasing interest rate schedules for borrowed funds, nonlinear constraints, etc. We will not elaborate these extensions to the basic model. Nor will we consider the actual computational procedures involved except to note that computer routines for deterministic appraisal are available. As noted in the following section, some progress has also been made in the development of mathematical programming procedures to handle risky investment decisions.

From a decision theory point of view, the difficulty with the deterministic mathematical programming procedures outlined above is that in maximizing total present value subject to specified constraints, no account is taken of the utility assessment of alternative cash flow sequences. This is exactly the same general problem as we discussed relative to the present value assessment of mutually exclusive projects. A difference, however, in the present context is that the relevant net cash flows for multidimensional utility appraisal are not those of individual projects. They are the total flows associated with feasible project combinations. This implies two things. First, that the decision maker's multidimensional utility function for cash flows over T time dimensions must be known or elicited. Assuming that the requirements of joint preferential independence and utility independence (see Section 4.3) are adequately met, a utility function might be specified in the multiattribute form of equation (4.5). This will either be additive as in (4.6) or, more likely, multiplicative as in (4.7). The second implication is that the objective function of (8.3) must be reformulated in terms of utility maximization. Thus the objective should be to maximize the multidimensional utility function $U(C_0, C_1, \ldots, C_T)$, where C_t denotes the total net cash flow in period t. Define

$$C_t = \sum_{i=1}^{I} C_{it} x_i \tag{8.8}$$

Then if the elicited utility function is separable and additive as per (4.6), we have

$$U(C_0, C_1, \ldots, C_T) = \sum_{t=0}^{T} k_t U_t(C_t) \tag{8.9}$$

so that the objective function becomes

$$U = \sum_{t=0}^{T} k_t U_t \left(\sum_{i=1}^{I} C_{it} x_i \right) \tag{8.10}$$

In general, $U_t(C_t)$ will not be linear, so that the objective function (8.10) is not amenable to linear programming. The same situation prevails if the multidimensional utility function is multiplicative as per (4.7). In this case we have

$$U = \left\{ \prod_{t=0}^{T} \left[K k_t U_t \left(\sum_{i=1}^{I} C_{it} x_i \right) + 1 \right] - 1 \right\} \Big/ K \tag{8.11}$$

The appropriate procedure for both the additive case of (8.10) and the multiplicative case of (8.11), given their nonlinearity, is nonlinear integer programming, where the basic constraints are as given in (8.4) and (8.5). By its nature this implies much more complicated computational procedures.

As outlined in Section 4.3, the multiattribute utility approach of (4.8) assumes joint preferential independence and utility independence. Joint preferential independence between time dimensions may not be an unreasonable assumption for most decision makers. However, utility independence between time dimensions cannot generally be assumed. It would imply, for example, that a decision maker's certainty equivalent for the 50/50 gamble of \$10 or \$10,000 in year 3 was the same regardless of whether he received \$500 or \$30,000 in each of the prior two years. Obviously, this would not be the case for most decision makers, so that the multiattribute utility approach of (4.8) could only be used as an approximation. For a less restrictive multiattribute form, see Bell (1975a and b).

If a multiattribute function is not applicable, the benchmark approach could be used to rank alternative cash flow streams as long as at least one time dimension is preferentially independent of the others. A difficulty, however, is that the benchmark approach necessitates an elicitation process for each alternative and is not amenable to such mathematical programming procedures as are possible with a multiattribute function. This makes the benchmark procedure impracticable except for problems where complete enumeration of alternatives is a practical possibility. If no time dimension is preferentially independent, so that not even the benchmark approach can be used, choice cannot be formalized and (as noted in Section 4.3) is purely a matter of the decision maker's intuition.

If it is appropriate to work with a multiattribute utility function, the basic difference from the present value approach is shown by the comparison between the objective function of equation (8.3) and that of (8.10) or (8.11). The present value approach of (8.3) involves discounted cash flow summation by projects and implies utility is linear in money. In contrast, the multidimensional utility approach of (8.10) or (8.11) implies calculating the utility of cash flows summed within a time period and then combining these utilities either additively with weights k_t or multiplicatively with weights K and k_t. Only in the unlikely combination of circumstances that a decision maker's utility function is additive as per (8.9), that $U_t(C_t)$ is linear, and that $k_t/k_{t-1} = \alpha_t$ will the two procedures be equivalent.

Stated bluntly, the problem with the present value approach is that, unlike the utility approach, it has no axiomatic justification. It is based, not on a set of reasonable axioms of choice, but on assumptions of a perfect investment market and of decision makers whose personal preferences imply that their current utility for money due at time t is linear. This criticism of the present value approach is bad enough in the absence of risk. When risk is present (which is invariably true) the situation is far worse. Under conditions of risk, especially when it comes to long-term investments, decision makers are most anxious to act in accord with their preferences and probability judgments. As we will see in the following section, present value appraisals based on risk considerations can give the decision maker unwarranted confidence that his desired preferences are being met. This danger arises because while a decision maker can fairly easily comprehend a set of sure cash flows and can often make reasonable intuitive judgments about choice between them, it is far more difficult for him to conceptualize and make intuitive decisions about risky cash flows.

8.3 RISKY INVESTMENTS

By a risky investment we mean one for which the cash flow stream is uncertain. For example, we might be concerned with a three-period project that has a sure outlay of c at t_0 and either a good (g) or a bad (b) net cash flow at t_1, t_2, and t_3. As can quickly be seen by drawing an event tree of the sequence possibilities, there are $2^3 = 8$ possible cash flow sequences or time traces. These might be conveniently listed as:

(c, g, g, g) (c, b, g, g)
(c, g, g, b) (c, b, g, b)
(c, g, b, g) (c, b, b, g)
(c, g, b, b) (c, b, b, b)

If there were 10 possible outcomes for each of the three periods after t_0, there would be $10^3 = 1000$ possible sequences. With T periods and n possible outcomes per period, there would be n^T possible sequences. In contrast, in the absence of risk there would only be a single sequence or time trace.

The problem of meaningful comprehension of risky time sequences is obvious—and more so when we allow for the fact that appraisal must also involve a probability of occurrence for each possible sequence. It is not surprising, therefore, that much confusion has attended the discussion of risky investment appraisal and that practitioners have developed a variety of ad hoc techniques to allow for risk in project appraisal. These naive approaches include such procedures as using conservative cash flow values in a nonrisky present value framework, using risk-adjusted interest rates to discount more heavily than otherwise, and using the expected cash flow in each period as if it were a sure payment. Though they are behaviorally appealing and easily applied, we will not discuss such naive procedures since they lack theoretical foundation.

Our discussion will focus on three approaches respectively based on mathematical programming under risk, the utility appraisal of risky present value, and multidimensional utility appraisal of risky cash flow sequences. From a decision theory viewpoint, only the multidimensional utility approach is logically respectable. However, just as in the case of nonrisky investment appraisal, it is certainly not as convenient in terms of practical application as the other two approaches, which are somewhat less demanding in their data requirements. All three approaches require information on the decision maker's subjective probability distributions for the C_{it} variables—of which there may be up to IT in total. Moreover, as Hillier (1963) and Wagle (1967) show, even to elicit the probability distribution for a single C_{it} variable may be a demanding task since each such cash flow will usually be based on a number of sources, each with its own stochastic basis. In addition, the utility-based procedures require information on preferences, which may involve complicated or extensive elicitation in the case of the multidimensional utility approach (Bell, 1975b).

Mathematical Programming under Risk

With random cash flows so that C_{it} becomes a random variable following some probability distribution, the mathematical programming approach must be one that accommodates risk. Relative to the objective function (8.3) of our basic linear programming formulation in Section 8.2, an approach to accounting for risk would be to derive the probability distribution of maximum achievable PV for each feasible set of projects. To do this it would be necessary to simulate C_{it} values for each possible investment combination and then solve a different programming problem for each value of C_{it}. An estimate of the probability distribution of maximum PV for each possible investment decision could thus be obtained by solving many such problems. Such a set of distributions would in no way solve the capital budgeting decision problem since it is based on the unrealizable assumption of perfect foreknowledge of the C_{it} values.

Far more useful results are likely to be given by a chance-constrained programming approach to risky investment appraisal. The essence of this approach is that the constraints of expression (8.4) in our basic linear programming model are reformulated in probabilistic form and the objective function (8.3) is reformulated in terms of expected present value. The model is thus specified as maximize

$$E(PV) = E\left[\sum_{i=1}^{I} (PV)_i x_i\right] = \sum_{i=1}^{I} E[(PV)_i]x_i$$

$$= \sum_{i=1}^{I} \left[\sum_{t=0}^{T} E(C_{it})\alpha^t\right] x_i \tag{8.12}$$

subject to the probabilistic budget constraints

$$P\left(\sum_{i=1}^{I} C_{it}x_i \geqq -B_t\right) \geqq p_t \qquad t = 0, 1, \ldots, T \tag{8.13}$$

and the integer constraints

$$0 \leqq x_i \leqq 1, x_i \text{ an integer} \qquad i = 1, 2, \ldots, I \tag{8.14}$$

where P is the probability operator and each p_t is some specified critical probability level. The constraints of (8.13) imply that the budget constraint for each period must at least be satisfied at its specified level of probability. The procedure is analogous to that discussed relative to the chance-constraint formulation of (7.48). To facilitate computation, it has generally been assumed that each of the C_{it} values follows a normal distribution; see e.g., Bernhard (1969) and Näslund (1971). This assumption makes it easier to transform the chance constraints into their deterministic equivalents for linear programming solution. These and other computational aspects, as well as a variety of accounting and resource constraint extensions to the basic model, are discussed by Byrne et al. (1971), Hillier (1969, Ch. 7), and Näslund (1971) in the context of risky investment appraisal.

Under a chance-constrained approach the solution given to the risky investment appraisal problem is a set of x_i values specifying which investments should be accepted to maximize expected present value while at the same time guaranteeing that those constraints specified in chance form will not be violated with more than the specified levels of probability.

Apart from difficulties of computation and data specification [e.g., the budget constraints B_t of expression (8.13) may be unknown or probabilistic] there are two significant criticisms that may be made of the chance-constrained approach. First, no explicit consideration is made either of the penalties involved if the constraints do happen to be violated [e.g., as will happen with probabilities $(1 - p_t)$ relative to (8.13)] or of the possible rewards involved if the chance constraints are made less tight. In other words, no account is taken of the fact that from a normative point of view the choice of an appropriate set of p_t values is itself a decision problem. The second criticism is our standard one: because of its use of present value as the criterion of optimality, the chance-constrained programming approach to investment appraisal fails to allow for the decision maker's preferences about alternative cash flow sequences.

Utility of Present Value

If the net cash flow C_{it} from the ith project in period t is uncertain, it follows that the net present value of any chosen set of projects from among the possibilities available must also be uncertain. Recognizing this fact, Hillier (1969, 1971), Jean (1970), and others have argued that capital budgeting of risky projects should be based on maximizing the expected utility of net present value; i.e., maximize

$$U = E[U(PV)] \tag{8.15}$$

Total present value depends on the set of projects adopted. If we denote the present value of the total cash flow occurring at time t by $(PV)_t$, we have

$$(PV)_t = \sum_{i=1}^{I} C_{it} x_i \alpha^t \tag{8.16}$$

where x_i is the proportion of the ith project undertaken ($x_i = 1$ or 0 in the all-or-nothing case) and the discount factor per period is α. It follows that for a given selection of projects (i.e., for a given set of values $x_1, x_2, \ldots, x_I$),

$$PV = \sum_{t=0}^{T} (PV)_t = \sum_{t=0}^{T} \sum_{i=1}^{I} C_{it} x_i \alpha^t \tag{8.17}$$

Hence in terms of equation (8.15) we have

$$U = E\left[U\left(\sum_{t=0}^{T} \sum_{i=1}^{I} C_{it} x_i \alpha^t\right)\right] \tag{8.18}$$

so that expected utility must be evaluated in terms of a probability distribution that is conditional on the vector $(x_1, x_2, \ldots, x_I)$. Denoting this conditional distribution by $f(PV \mid x_1, x_2, \ldots, x_I)$, the investment decision problem is to choose the set of x_i values to maximize

$$U = \int U(PV) f(PV \mid x_1, x_2, \ldots, x_I) dPV \qquad (8.19)$$

subject to whatever constraints are relevant. If PV is discrete rather than continuous, the discrete probability analog of (8.19) is relevant.

Under particular conditions it may be possible to specify the various conditional probability distributions for PV algebraically. If these are tractable, analysis will be facilitated. Usually, however, the form of the PV distributions will not be known. Three alternatives are possible: (1) the distributions may be estimated by simulation procedures as discussed in Section 8.4 below; (2) the distributions might be assumed to be normal on the basis of the central limit theorem since C_{it} and $(PV)_t$, and hence PV, are sums of numbers of random variables; or (3) analysis might proceed in terms of the first two moments of the PV distributions on the basis of a quadratic utility function or of a Taylor series approximation to some other utility function. These alternatives have been discussed by Adelson (1965), Hillier (1963, 1969, 1971), Jean (1970), Reutlinger (1970), and Wagle (1967). A suggestion of Hertz (1964) that has proved popular in practice is that graphs of the probability distributions of present value for the various investment possibilities may be derived by simulation and used on an inspection basis to guide choice. Known as *risk analysis*, this procedure leaves the utility appraisal of present value to the decision maker's intuition. Alternatively, as discussed in Chapter 9, the probability distributions of present value may be analyzed in terms of stochastic dominance criteria to determine the set of stochastically efficient investment possibilities. Choice could thereby be narrowed to those alternatives that are efficient for any risk-averse decision maker whose preferences depend only on present value and not on the time sequence of cash flows yielding the present value.

To illustrate what is involved in the utility appraisal of present value, consider a simple example. Suppose the decision maker's opportunity set consists of three indivisible independent projects ($i = 1, 2, 3$) each extending over three years ($t = 0, 1, 2, 3$) and having uncertain cash flows, as shown in Table 8.5 where H_{it} and L_{it} respectively denote a high or low C_{it} value. These cash flows are those thought to be appropriate by the decision maker on the basis of budgeting, expert advice, and judgment about possible outcomes. They may reflect a discrete simplification of a continuous set of cash flow possibilities. $P(C_{it})$ denotes the marginal probability of C_{it}. In each case $P(C_{i0}) = 1$ since a sure initial outlay is assumed for each project.

TABLE 8.5. Decision Maker's Judgment of Cash Flow Characteristics of Three Indivisible Independent Risky Projects

	$i = 1$		$i = 2$		$i = 3$	
C_{it}	C_{1t}	$P(C_{1t})$	C_{2t}	$P(C_{2t})$	C_{3t}	$P(C_{3t})$
	($\$10^3$)		($\$10^3$)		($\$10^3$)	
C_{i0}	−10	1.0	−8	1.0	−18	1.0
H_{i1}	20	0.7	12	0.6	40	0.5
L_{i1}	0	0.3	2	0.4	20	0.5
H_{i2}	18	0.68	14	0.54	50	0.45
L_{i2}	2	0.32	0	0.46	−10	0.55
H_{i3}	16	0.508	9	0.67	65	0.375
L_{i3}	3	0.492	5	0.33	−20	0.625

The marginal probabilities $P(C_{it})$ are only needed if it is desired to know the probability distribution (or the mean and variance) of each period's total cash flow C_t. This is the case for the certainty equivalent approach discussed below. For direct appraisal of $U(PV)$, the probability distribution of PV is required. This distribution is most easily obtained from the conditional probabilities $P(C_{it} \mid C_{it-1})$. Conveniently, it is these conditional probabilities that are most readily elicited from the decision maker and can also be used to derive the marginal probabilities $P(C_{it})$ if required. For our present example, suppose the elicited conditional probabilities for each project are as listed in Table 8.6. Inspection shows that our decision maker does not regard the possible cash flows as independent over time. For projects 1 and 2, a high (low) result in period $t - 1$ favors a high (low) result in period t. For project 3 the reverse is true. Because of these effects the cash flows from projects 1 and 2 are positively associated; they are negatively associated with those from project 3. Such

TABLE 8.6. Elicited Conditional Probabilities for C_{it} given C_{it-1} for the Projects of Table 8.5

Elicited Probability	$i = 1$	$i = 2$	$i = 3$
$P(H_{i1})$	0.7	0.6	0.5
$P(L_{i1})$	0.3	0.4	0.5
$P(H_{i2} \mid H_{i1})$	0.8	0.7	0.2
$P(L_{i2} \mid H_{i1})$	0.2	0.3	0.8
$P(H_{i2} \mid L_{i1})$	0.4	0.3	0.7
$P(L_{i2} \mid L_{i1})$	0.6	0.7	0.3
$P(H_{i3} \mid H_{i2})$	0.7	0.9	0.1
$P(L_{i3} \mid H_{i2})$	0.3	0.1	0.9
$P(H_{i3} \mid L_{i2})$	0.1	0.4	0.6
$P(L_{i3} \mid L_{i2})$	0.9	0.6	0.4

statistical dependencies between cash flows will often occur due to the influence of common basic factors such as climate or general economic conditions.

From the conditional probabilities of Table 8.6 the marginal probabilities of Table 8.5 can be calculated, e.g., as

$$P(L_{32}) = P(L_{32} \mid H_{31})P(H_{31}) + P(L_{32} \mid L_{31})P(L_{31}) = 0.55$$

and thence

$$P(H_{33}) = P(H_{33} \mid H_{32})P(H_{32}) + P(H_{33} \mid L_{32})P(L_{32}) = 0.375$$

To determine the probability distribution of present value, denoted $P(PV \mid x_1, x_2, x_3)$, we need to derive both the present value of each possible cash flow stream, given as in equation (8.17) by

$$PV = \sum_{t=0}^{T} \sum_{i=1}^{I} C_{it} x_i \alpha^t$$

and the probability of each possible cash flow stream. As shown in Table 8.7, with ample funds there would be seven alternative project combinations and 728 possible cash flow sequences. (If we had assumed three possible cash flows per project per period, the number of sequences would be 21,951—hence the simplicity of our example.) To keep the number of possible sequences manageable for purposes of exposition, we will further assume that total initial outlay is restricted to $18,000. Feasible project combinations are thus as shown in Table 8.8 with the details of their associated cash flow sequences and present values, assuming $r = 0.10$. Note that the period cash flow values for projects 1 and 2 in combination are given by $C_t = C_{1t} + C_{2t}$. This reflects the decision maker's judgment that

TABLE 8.7. Possible Project Combinations and Their Number of Cash Flow Sequences for the Projects of Table 8.5 with Ample Funds

Project Combination (x_1, x_2, x_3)	Number of Sequences
(1, 0, 0)	8
(0, 1, 0)	8
(0, 0, 1)	8
(1, 1, 0)	64
(1, 0, 1)	64
(0, 1, 1)	64
(1, 1, 1)	512
	728

TABLE 8.8. Details of Possible Cash Flow Sequences, Their Present Value and Probability of Occurrence for Feasible Combinations of the Projects of Table 8.5 When Total Outlay Is Limited to $18,000

Project Combination (x_1, x_2, x_3)	Possible Cash Flow Sequences $t = 0$	$t = 1$	$t = 2$	$t = 3$	PV $(r = 0.10)$	Probability $P(PV \mid x_1, x_2, x_3)$
	($\$10^3$)				($\$10^3$)	
(1, 0, 0)	−10	20	18	16	35.078	0.392
	−10	20	18	3	25.311	0.168
	−10	20	2	16	21.856	0.014
	−10	20	2	3	12.089	0.126
	−10	0	18	16	16.896	0.084
	−10	0	18	3	7.129	0.036
	−10	0	2	16	3.674	0.018
	−10	0	2	3	−6.093	0.162
(0, 1, 0)	−8	12	14	9	21.240	0.378
	−8	12	14	5	18.235	0.042
	−8	12	0	9	9.671	0.072
	−8	12	0	5	6.666	0.108
	−8	2	14	9	12.149	0.108
	−8	2	14	5	9.144	0.012
	−8	2	0	9	0.580	0.112
	−8	2	0	5	−2.425	0.168
(0, 0, 1)	−18	40	50	65	108.521	0.010
	−18	40	50	−20	44.658	0.090
	−18	40	−10	65	58.934	0.240
	−18	40	−10	−20	−4.926	0.160
	−18	20	50	65	90.340	0.035
	−18	20	50	−20	26.476	0.315
	−18	20	−10	65	40.752	0.090
	−18	20	−10	−20	−23.108	0.060
(1, 1, 0)	−18	32	32	25	56.318	0.1482
	−18	32	32	21	53.313	0.0165
	−18	32	18	25	44.749	0.0282
	−18	32	18	21	41.744	0.0423
	−18	22	32	25	47.227	0.0423
	−18	22	32	21	44.222	0.0047
	−18	22	18	25	35.658	0.0439
	−18	22	18	21	32.653	0.0659
	−18	32	32	12	46.551	0.0635
	−18	32	32	8	43.546	0.0071
	−18	32	18	12	34.982	0.0121
	−18	32	18	8	31.977	0.0181
	−18	22	32	12	37.460	0.0181
	−18	22	32	8	34.455	0.0020
	−18	22	18	12	25.891	0.0188
	−18	22	18	8	22.886	0.0282

TABLE 8.8. (*continued*)

Project Combination (x_1, x_2, x_3)	Possible Cash Flow Sequences $t = 0$	$t = 1$	$t = 2$	$t = 3$	PV $(r = 0.10)$	Probability $P(PV \mid x_1, x_2, x_3)$
	($\$10^3$)				($\$10^3$)	
(1, 1, 0)	−18	32	16	25	43.096	0.0053
(*continued*)	−18	32	16	21	40.091	0.0006
	−18	32	2	25	31.527	0.0010
	−18	32	2	21	28.522	0.0015
	−18	22	16	25	34.005	0.0015
	−18	22	16	21	31.000	0.0002
	−18	22	2	25	22.436	0.0016
	−18	22	2	21	19.431	0.0024
	−18	32	16	12	33.329	0.0476
	−18	32	16	8	30.324	0.0053
	−18	32	2	12	21.760	0.0091
	−18	32	2	8	18.755	0.0136
	−18	22	16	12	24.238	0.0136
	−18	22	16	8	21.233	0.0015
	−18	22	2	12	12.669	0.0141
	−18	22	2	8	9.664	0.0212
	−18	12	32	25	38.136	0.0318
	−18	12	32	21	35.131	0.0035
	−18	12	18	25	26.567	0.0061
	−18	12	18	21	23.562	0.0091
	−18	2	32	25	29.045	0.0091
	−18	2	32	21	26.040	0.0010
	−18	2	18	25	17.467	0.0094
	−18	2	18	21	14.471	0.0141
	−18	12	32	12	28.369	0.0136
	−18	12	32	8	25.364	0.0015
	−18	12	18	12	16.800	0.0026
	−18	12	18	8	13.795	0.0039
	−18	2	32	12	19.278	0.0039
	−18	2	32	8	16.273	0.0004
	−18	2	18	12	7.709	0.0041
	−18	2	18	8	4.704	0.0061
	−18	12	16	25	24.914	0.0068
	−18	12	16	21	21.909	0.0008
	−18	12	2	25	13.345	0.0013
	−18	12	2	21	10.340	0.0019
	−18	2	16	25	15.823	0.0019
	−18	2	16	21	12.818	0.0002
	−18	2	2	25	4.254	0.0020
	−18	2	2	21	1.249	0.0030

TABLE 8.8. (*continued*)

Project Combination (x_1, x_2, x_3)	Possible Cash Flow Sequences $t=0$	$t=1$	$t=2$	$t=3$	PV $(r=0.10)$	Probability $P(PV \mid x_1, x_2, x_3)$
	($\$10^3$)				($\$10^3$)	
(1, 1, 0)	−18	12	16	12	15.147	0.0612
(*continued*)	−18	12	16	8	12.142	0.0068
	−18	12	2	12	3.578	0.0117
	−18	12	2	8	0.573	0.0175
	−18	2	16	12	6.056	0.0175
	−18	2	16	8	3.051	0.0019
	−18	2	2	12	−5.513	0.0181
	−18	2	2	8	−8.518	0.0272

projects 1 and 2 are independent. They are neither complementary ($C_t > C_{1t} + C_{2t}$ for some t) nor antagonistic ($C_t < C_{1t} + C_{2t}$ for some t).

The probability of occurrence of each present value is equal to the probability of its associated sequence occurring; i.e.,

$$P(PV \mid x_1, x_2, x_3) = P(C_0, C_1, C_2, C_3 \mid x_1, x_2, x_3)$$

These probabilities are derived from the conditional probabilities of Table 8.6. For example, as shown in Table 8.8 for the third sequence of the combination (1, 0, 0),

$$P(C_{10}, H_{11}, L_{12}, H_{13}) = P(H_{13} \mid L_{12})P(L_{12} \mid H_{11})P(H_{11})$$
$$= (0.1)(0.2)(0.7) = 0.014$$

and for the first sequence of the combination (1, 1, 0),

$$P(C_{10} + C_{20}, H_{11} + H_{21}, H_{12} + H_{22}, H_{13} + H_{23})$$
$$= P(C_{10}, H_{11}, H_{12}, H_{13})P(C_{20}, H_{21}, H_{22}, H_{23}) = 0.1482$$

Given knowledge of the decision maker's utility function, the information of Table 8.8 is all that is required for determining optimal choice in terms of $U(PV)$. The appraisal may be carried out in three ways: by direct evaluation of $U(PV)$, by the moment method of equation (4.27), or by a certainty equivalent approach. We will illustrate each procedure assuming the decision maker's utility function to be

$$U = 2.05X - 0.01X^2 \qquad X \leqq 80 \tag{8.20}$$

where X is in thousands of dollars.

DIRECT APPRAISAL OF $U(PV)$

Direct appraisal implies the evaluation of $U(PV)$ via equation (8.19). Thus we have

$$\begin{aligned} U &= \sum U(PV \mid x_1, x_2, x_3) P(PV \mid x_1, x_2, x_3) \\ &= \sum [2.05PV - 0.01\,(PV)^2] P(PV \mid x_1, x_2, x_3) \end{aligned} \tag{8.21}$$

where the summation is over the set of cash flow sequences associated with the project combination (x_1, x_2, x_3), and $U(PV)$ is specified by (8.20). Carrying out these calculations, we obtain

$$U(PV \mid 1,0,0) = 35.73 \qquad U(PV \mid 0,0,1) = 47.87$$
$$U(PV \mid 0,1,0) = 21.00 \qquad U(PV \mid 1,1,0) = 52.09$$

Hence the combination of projects 1 and 2 should be the preferred choice.

MOMENT METHOD APPRAISAL OF $U(PV)$

The moment method of equation (4.27) applied to (8.20) implies

$$U = 2.05E(PV) - 0.01[E(PV)]^2 - 0.01V(PV) \tag{8.22}$$

so that it is necessary to know the mean $E(PV)$ and variance $V(PV)$ of each project combination's probability distribution of present value. Calculated in the usual way from the PV and $P(PV)$ data of Table 8.8, these are as shown in Table 8.9. Since the moment method is exact for quadratic utility functions, evaluation of (8.22) using the means and variances of Table 8.9 gives the same utility values for each of the four investment possibilities as obtained via the direct evaluation method.

TABLE 8.9. Mean and Variance of Present Value for the Project Combinations of Table 8.8

Project Combination (x_1, x_2, x_3)	$E(PV)$	$V(PV)$
	($\$10^3$)	
(1, 0, 0)	20.586	223.315
(0, 1, 0)	11.290	86.529
(0, 0, 1)	32.243	783.009
(1, 1, 0)	31.876	309.930

Relative to the variance of present value for the project combination (1, 1, 0), we have

$$V(PV \mid 1,1,0) = V[(PV \mid 1,0,0) + (PV \mid 0,1,0)]$$

which may be calculated as $E[PV - E(PV)]^2$ or as

$$V(PV \mid 1,0,0) + V(PV \mid 0,1,0) + 2\,\mathrm{Cov}[(PV \mid 1,0,0), (PV \mid 0,1,0)]$$

It is interesting that while their cash flows within periods are positively associated, the correlation between the present values of projects 1 and 2 is virtually zero ($\rho = 0.0003$).

As noted above, the variance of present value used in equation (8.22) is derived directly from the array of total present values of Table 8.8 for each project combination. This variance could also be calculated as a function of the variances of each period's present value of cash flows $(PV)_t$. To do this, it would be necessary to take account of the covariance between the cash flows of different periods. The relationship is

$$V(PV) = \sum_{t=0}^{T} V[(PV)_t] + 2\sum_{s=0}^{T-1}\sum_{t=s+1}^{T} \mathrm{Cov}[(PV)_s, (PV)_t] \qquad (8.23)$$

where

$$V[(PV)_t] = V\left(\sum_{i=1}^{I} C_{it}x_i\alpha^t\right) = a^{2t}V\left(\sum_{i=1}^{I} C_{it}x_i\right) \qquad (8.24)$$

As a general rule, only if the cash flows of different periods are independently distributed will the covariance terms be zero and the variance of total present value be equal to the sum of the variances of each period's present value. By working with the distribution of PV rather than the distributions of $(PV)_t$, we were able to avoid the necessity of calculating the interperiod covariances of cash flows.

CERTAINTY EQUIVALENT APPRAISAL OF $U(PV)$

The essence of the certainty equivalent approach to appraisal is that an unconditional certainty equivalent is established for each period's array of possible cash flows for each alternative combination of projects. The present value of each sequence of certainty equivalents is then used to compare the alternative project combinations. Using this method, the combination with the highest positive present value for its certainty equivalents is the preferred investment. Normally, the certainty equivalent for each period's cash flows for each alternative can be nominated by the decision maker on an intuitive basis. For this reason the certainty equivalent approach is simpler to apply than the direct or moment method approaches to $U(PV)$.

As Jean (1970) has shown, the certainty equivalent approach to risky investment appraisal implies that the decision maker has a utility function $U_t(C_t)$, which he currently holds for each period's cash flow, and that his utility function for the present value of a sequence of certainty equivalents

$v_0, v_1, v_2, \ldots, v_T$ is a separable and additive function of the period utility functions; i.e.,

$$U(PV) = \sum_{t=0}^{T} k_t U_t(v_t) \qquad (8.25)$$

where $PV = \sum_t v_t \alpha^t$. The certainty equivalent approach is thus a multidimensional one akin to equation (8.9). Its simple linear nature implies no utility interaction between time periods so that the utility evaluation of cash flows in each period is independent of the cash flows of other periods. No allowance is made for preference about the time pattern of cash flows.

In applying the certainty equivalent approach to our simple example, we will assume the functions $U_t(C_t)$ are identical and given by

$$U_t(C_t) = 2.05C_t - 0.01C_t^2 \qquad C_t \leqq 80 \qquad (8.26)$$

where C_t is in thousands of dollars. Denoting the mean and variance of C_t by E_t and V_t respectively, by definition of v_t as a certainty equivalent we must have

$$U_t(v_t) = 2.05E_t - 0.01E_t^2 - 0.01V_t \qquad (8.27)$$

Solving equations (8.26) and (8.27) for v_t by putting $C_t = v_t$ in (8.26),

$$v_t = \{-2.05 \pm [4.2025 - 0.04(2.05E_t - 0.01E_t^2 - 0.01V_t)]^{0.5}\}/(-0.02) \qquad (8.28)$$

The values of E_t and V_t, calculated from the data of Table 8.5, are listed in Table 8.10 for each feasible project combination.

TABLE 8.10. Mean and Variance of C_t for the Project Combinations of Table 8.8

Project Combination (x_1, x_2, x_3)	E_1	V_1	E_2	V_2	E_3	V_3
	(\$$10^3$)	(\$$10^6$)	(\$$10^3$)	(\$$10^6$)	(\$$10^3$)	(\$$10^6$)
(1, 0, 0)	14	84	12.88	55.706	9.604	42.239
(0, 1, 0)	8	24	7.56	48.686	7.680	3.538
(0, 0, 1)	30	100	17.00	891.000	11.875	1693.359
(1, 1, 0)	22	108	20.44	104.392	17.284	45.781

Based on equation (8.28), the certainty equivalent sequences for our four feasible project combinations are as shown in Table 8.11 along with their associated present values. The implications for choice are the same as given by the direct and moment methods, although as discussed above, the certainty equivalent approach is based on somewhat different assumptions and need not necessarily indicate the same optimal choice.

TABLE 8.11. Certainty Equivalents for Each Period's Risky Cash Flows and Their Total Present Value for the Project Combinations of Table 8.8

Project Combination (x_1, x_2, x_3)	v_0	v_1	v_2	v_3	PV $(\sum v_t \alpha^t)$
			($\$10^3$)		
(1, 0, 0)	−10	13.527	12.570	9.377	19.730
(0, 1, 0)	−8	7.873	7.304	7.661	10.949
(0, 0, 1)	−18	29.314	11.939	2.970	20.747
(1, 1, 0)	−18	21.332	19.806	17.016	30.545

RISKY INTERRELATED PROJECTS

So far we have assumed the projects under consideration to be economically independent in the sense that the cash flow from a combination of projects is equal to the sum of the cash flows from the individual projects; i.e.,

$$C_t = \sum_{i=1}^{I} C_{it} \tag{8.29}$$

In fact, projects may often be either complementary, so that

$$C_t > \sum_{i=1}^{I} C_{it} \tag{8.30}$$

or antagonistic, so that

$$C_t < \sum_{i=1}^{I} C_{it} \tag{8.31}$$

Indeed, they may be complementary for some t and antagonistic for other t.

Based on the utility of present value, Hillier (1969, 1971) has developed a relatively pragmatic procedure for the appraisal of risky interrelated projects. The essence of his approach is the maximization of $U(PV \mid x_1, x_2, \ldots, x_I)$ over all feasible project combinations; U is a hyperbola-based risk-averse utility function and $(PV \mid x_1, x_2, \ldots, x_I)$ is estimated to allow for pairwise complementarities or antagonisms between projects. Assuming that the probability distributions of present value for different project combinations are normal and hence specified by the mean and variance of $(PV \mid x_1, x_2, \ldots, x_I)$, a Taylor series expansion of the expected utility function involving only the mean and variance of present value is used to determine the optimal investment. Either an approximate solution procedure based on an iterative sequence of linear programs may be used, or a more complicated but exact procedure based on a special algorithm for integer nonlinear programming may be applied. Relevant details are given by Hillier (1969, 1971).

The attractiveness of Hillier's procedures lies in his presentation of a computerized algorithm along with the practical procedures he has developed for specifying management's utility function and estimates of the required means and variances allowing for project interrelationships. In each case these procedures involve only the determination of a small number of parameters, which are readily determined from management's answers to questions of real-world relevance. In this sense, Hillier's approach (though too extensive to elaborate in detail here) probably constitutes the most practical and comprehensive available for the utility of present value appraisal of large-scale risky investment possibilities.

Multidimensional Utility for Risky Cash Flows

As Meyer (1970) has shown, acceptance of the axioms of Bernoullian decision theory implies a multidimensional utility approach to investment appraisal. Only by recognizing cash flows in different periods as multiple attributes with substitution possibilities between them can preferences about the time pattern of cash flows be accommodated. This is perhaps simple enough in deterministic situations involving only a few investment possibilities, as we showed in relation to the two nonrisky alternatives of Table 8.1; choice in such simple deterministic cases may be easily made on the basis of intuition or by the benchmark approach. But even in the deterministic case when there are many alternative possibilities, neither the intuitive nor the benchmark approach is very practicable.

With risky cash flows, except in the very simplest of cases, evaluation in multidimensional terms is likely to be difficult. Whereas in single-period planning problems choice lies between risky prospects, each reducible to a single probability distribution, in risky investment analysis we are dealing with alternative stochastic processes, each generating a sequence of probability distributions over a series of time dimensions. By the nature of the problem, choice between stochastic processes is much more difficult than choice between single probability distributions.

Except that the cash flows involved are probabilistic, the multidimensional utility approach to risky project appraisal may proceed as outlined in Section 8.2 in relation to nonrisky projects. If joint preferential independence and utility independence (as defined in Section 4.3) prevail between time dimensions, the general approach of (4.8) is applicable. While joint preferential independence is quite likely to prevail, utility independence is not. To overcome this problem, Bell (1975a and b) has developed an approach based on a utility function of the general form $U = f[U(C_t, C_{t+1})]$ for $t = 0, 1, 2, \ldots, T$. If cash flows (in at least one time period) are preferentially independent (as will generally be true) the benchmark approach can be used. If no time dimension is preferentially independent, choice between cash flow sequences cannot be formalized and, as discussed by von Winterfeldt and Fischer (1973), is purely a matter of the decision maker's intuition. With any sizable risky investment problem, however, the number of stochastic processes involved will be so large as to make sound intuitive appraisal unlikely. It is no wonder, therefore, that in practice decision makers have tended to use such naive but readily understood measures as risk-discounted or expected present value, expected internal rate of return, expected payback period, etc., as guides to investment choice.

To illustrate the multidimensional approach, we will again use the investment problem specified by the data of Table 8.8. In multidimensional terms, the present value data of Table 8.8 are irrelevant; what is essential are the sets of cash flow sequences $\{C_0, C_1, C_2, C_3\}$ associated with the project combinations (x_1, x_2, x_3) along with their probabilities of occurrence $P(C_0, C_1, C_2, C_3 \mid x_1, x_2, x_3)$.

Inspection of Table 8.8 shows that even for such a relatively small problem as this, which involves only four combinations of three indivisible independent projects each having a time span of three periods and two possible cash flows per period, intuitive appraisal is virtually impossible. It is just not humanly possible to comprehend the data simultaneously in adequate fashion. Intuition, however, might be used to appraise the data of Table 8.10, which summarizes Table 8.8 in terms of the mean and variance of each period's cash flows for each project combination. Certainly such means and variances provide a massive reduction in the amount of data to be comprehended. However, unless the cash flows in each period are normally distributed, description in terms of their means and variances is not equivalent to the full time-state specification of Table 8.8. Nor would the number of project combinations or time periods need to be increased by much to make intuitive appraisal, by use of the time traces of means and variances, impractical. For data in the form of either Table 8.8 or Table 8.10, this difficulty of comprehension occurs because intuitive appraisal in multidimensional terms implies the comparison of time sequences, not element by element but as a block. In Table 8.10, each line has to be compared as an entity with each other line. With Table 8.8, each project combination's block of possible cash flows with their associated probabilities has to be compared with each other block of possibilities. Obviously, intuitive appraisal will generally not be a practical procedure.

Assuming preferential independence between time periods, the benchmark approach of Section 4.3 may be used to evaluate the problem of Table 8.8. The elicitation of a benchmark equivalent, say,

$$(C_0^{+++}, C_1^{+} = 0, C_2^{+} = 0, C_3^{+} = 0) \equiv (C_0, C_1, C_2, C_3) \qquad (8.32)$$

is required for each of the 88 possible cash flows, along with elicitation of the conditional preference function $U(C_0 \mid C_1^+, C_2^+, C_3^+)$. Using this utility function, each of the project combinations may then be evaluated in the usual way in terms of expected utility. Alternatively, the benchmark approach might be applied to the data of Table 8.10 to establish a benchmark certainty equivalent

$$(v_0^{+++}, v_1^+ = 0, v_2^+ = 0, v_3^+ = 0) \equiv (v_0, v_1, v_2, v_3) \tag{8.33}$$

for each mean-variance sequence. Expected utility evaluation would again proceed in terms of the function $U(C_0 \mid C_1^+, C_2^+, C_3^+)$. Such a procedure is somewhat akin to the certainty equivalent approach to $U(PV)$ discussed previously but differs in that the certainty equivalents of equation (8.33) are based on preferences about the entire sequence of means and variances for each combination of projects, not on period-by-period appraisal.

The above discussion assumes elicitation of benchmark equivalents and the conditional preference function on a case-by-case basis. This could be very time consuming. Taking the line that analysis using a rough but easily made approximation is better than either having none at all or the expense of a detailed appraisal, a generalized "everyman's" procedure might be used. For example, the multiplicative ordinal function

$$Q = \Pi_t C_t \tag{8.34}$$

might be used with Q constant to convert each cash flow sequence $(C_0, C_1, \ldots, C_T)$ to a benchmark equivalent $(C_0^{++}, C_2^+, \ldots, C_T^+)$ with C_t^+ equal to some "survival" requirement. Then "everyman's utility function"

$$U(C_0^{++} \mid C_1^+, C_2^+, \ldots, C_T^+) = \log_e (C_0^{++} \mid C_1^+, C_2^+, \ldots, C_T^+) \tag{8.35}$$

could be used for expected utility appraisal. There is nothing particularly special about the suggested forms of equations (8.34) and (8.35); they could be modified in many ways: e.g., current wealth W_0 might be included in (8.35) by using $(W_0 + C_0^{++})$ as the argument. To illustrate this approach, suppose a three-year investment has two possible cash flow streams, (2, 2, 4) and (4, 1, 3), with respective probabilities of 0.3 and 0.7. Applying (8.34) with a "survival" requirement that $C_t^+ = 1$ gives the respective benchmark streams (16, 1, 1) and (12, 1, 1). Using (8.35) and taking the expectation then indicates that the investment has a utility of $(0.3)(2.77) + (0.7)(2.49) = 2.574$.

If joint preferential independence and utility independence can be assumed, the multiattribute approach of equation (4.5) can be used as outlined in Section 4.3. For risky investment appraisal the multiplicative form of (4.7), i.e.,

$$U(C_0, C_1, \ldots, C_T) = \left\{\prod_{t=0}^{T} [Kk_t U_t(C_t) + 1] - 1\right\} \Big/ K \qquad (8.36)$$

is most relevant. In contrast to the simple additive form of (4.6) or (8.9), the multiplicative form at least allows for interaction between the period utility functions $U_t(C_t)$ currently held for C_t. To apply the approach to our example problem of Table 8.8, we will assume specifications as shown in Table 8.12, where C_t is in thousands of dollars. As per (4.9) these k_t values

TABLE 8.12. Assumed Values of k_t and $U(C_t)$

t	k_t	$U_t(C_t)$
0	0.7	$0.090580C_0 - 0.000362C_0^2 + 1.747830$
1	0.6	$0.035181C_1 - 0.000211C_1^2 - 0.069518$
2	0.4	$0.024510C_2 - 0.000196C_2^2 + 0.264706$
3	0.3	$0.021390C_3 - 0.000214C_3^2 + 0.513369$

imply $K = -0.93$. Substituting into (8.36), we have

$$\begin{aligned} U(C_0, C_1, C_2, C_3) = \{&[1 - 0.651(0.090580C_0 - 0.000362C_0^2 + 1.747830)]\cdot \\ &[1 - 0.558(0.035181C_1 - 0.000211C_1^2 - 0.069518)]\cdot \\ &[1 - 0.372(0.024510C_2 - 0.000196C_2^2 + 0.264706)]\cdot \\ &[1 - 0.279(0.021390C_3 - 0.000214C_3^2 + 0.513369)] - 1\}/(-0.93) \end{aligned}$$

Applying this multidimensional utility function to the cash flow sequences of Table 8.8 and taking expectations in terms of $P(C_0, C_1, C_2, C_3 \mid x_1, x_2, x_3)$ as listed in Table 8.8, we obtain the following utility values for the feasible project combinations:

$$U(1,0,0) = 0.810 \qquad U(0,0,1) = 0.636$$
$$U(0,1,0) = 0.845 \qquad U(1,1,0) = 0.646$$

The preferred investment is thus project 2.

As discussed by Hirshleifer (1970) and Jean (1970), risky cash flow sequences may also be depicted as outcomes under an exhaustive and mutually exclusive listing of possible future states of nature, where each state embodies the results of events (cash flows) up to and including the period it describes. The data of Table 8.8, for example, could be arranged in such fashion. Appraisal then proceeds in utility form in the usual way. The arrangement of investment flows in terms of state descriptions and their subsequent utility appraisal has come to be known as the *time-state preference approach*. We will not elaborate it further since (except for the way the cash flows are described) it is equivalent to the approach we have followed.

8.4 SIMULATION AS AN AID TO APPRAISAL

The complexity of models such as the various mathematical programming specifications (and of most real-world applications of decision analysis) has been emphasized at many points in earlier chapters and also in Section 8.3. Various techniques have been suggested and used for coping with such complexities. One of the most important is Monte Carlo simulation, a topic already mentioned briefly on several occasions and presently reviewed in a little more detail—albeit still sketchily because complete exposition is beyond the scope of this book.

We use the term simulation to denote the numerical exploration of a symbolic model, used to mimic the behavior of a modeled system over time. In a decision analysis context, we will almost invariably be interested in Monte Carlo or stochastic simulation, wherein variates are sampled from probability distributions appropriately specified within the model (Mihram, 1972). The probabilistic specification (and the model generally) can be as complex or as simple as dictated by the problem and the detail at hand. Typically, however, a formal modeling approach will be warranted only for relatively complex and economically significant decision problems, and this implied degree of complexity will usually mean that the model needs to be implemented on a computer (Anderson, 1974).

Simulation can aid in investment appraisal in several ways. Perhaps most importantly because of the practical plausibility it adds to investment analysis, it can be used for computation of, say, $U(PV)$ of investment combinations, either directly or indirectly by means of sampling estimations of moments of PV. More informally, it would be useful in exploring possible investment strategies and their uncertain consequences. In less formal modes a simulation model can be linked interactively to the decision maker, who can then use it to explore possible consequences of alternative decisions. Such interactive simulation (often termed gaming) may also prove instructive in discovering preferences among multiattributed consequences over multiple time periods and may well involve, for instance, explicit consideration of the impact of "premature" death, taxation issues, intergenerational transfers, etc. Conceptually, there is no limit to the scope of numerical explorations in stochastic simulations tuned to careful appraisal of investment possibilities.

Simulation will not be very useful either when the investments are extremely simple or when all the distributions are normal, as has often been presumed in investment analysis. In the latter case it is much simpler to compute the means and variances analytically as indicated in Section 8.3 and thereby describe the investment environment succinctly and completely. However, in the many situations when risk is other than normal and (as will often be the case) utility is other than quadratic, appraisal only by means and variances (and covariances) of present values (and their components) may not be adequate to indicate optimal decisions.

Continuous (and discrete) distributions other than the normal are readily accommodated in stochastic simulation and can range from convenient theoretical distributions (e.g., beta, Poisson, gamma, etc.) to quite arbitrary empirical distributions, such as might be determined by the judgmental fractile method. For example, the triangular distribution is one widely used in simulation for its ease of use, its clear nonnormality, and its economy in elicitation and parametric description. We will use it to review the most convenient and most widely used method of Monte Carlo sampling, namely the inverse CDF transformation method (Meier et al., 1969).

The triangular distribution (Sprow, 1967; Pouliquen, 1970) is depicted graphically in both PDF and CDF forms in Figure 8.1 and is completely defined by just the mode m, lowest possible value a, and highest

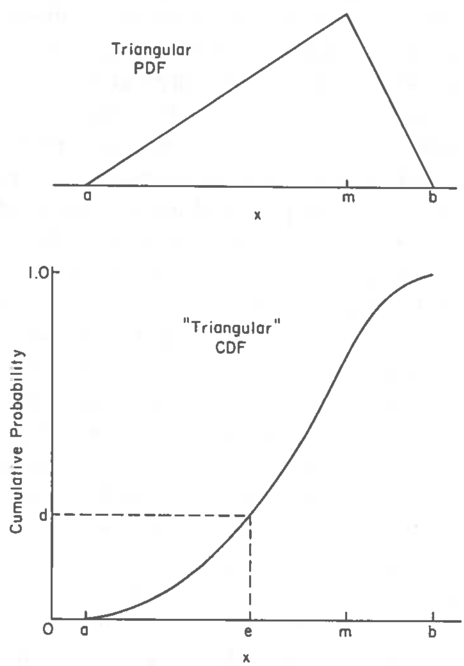


FIG. 8.1. Probability functions for a random variable x which follows a triangular distribution.

possible value b. These parameters can be elicited directly when it is judged that this distribution will provide an adequate description of beliefs. More generally, however, this adequacy of fit should be checked by comparing the "triangular" CDF with an elicited judgmental CDF.

Variates of any distribution can in principle be sampled in the inverse CDF method by projecting a uniform variate u on the cumulative probability scale through the CDF to the scale of the specified random variable. Uniform variates (i.e., rectangularly distributed over the interval 0 to 1) of the pseudorandom type are readily available on all modern computers (or in books as tables of random digits). The projection process is illustrated graphically in the lower part of Figure 8.1 where if d is a particular value of u, the corresponding triangular variate is e. For many distributions this transformation can be done algebraically and therefore very simply. For others such as the normal, the inverse function does not exist and alternative or approximate methods must be used (Mihram, 1972). The CDF for the triangular distribution is of simple form (Sprow, 1967),

$$\begin{aligned} F(x) &= (x-a)^2/[(b-a)(m-a)] \qquad a \leqq x \leqq m \\ &= 1-(b-x)^2/[(b-a)(b-m)] \qquad m \leqq x \leqq b \end{aligned} \tag{8.37}$$

and $F(x)$ can readily be equated with a uniform variate u and equation (8.37) solved for a corresponding triangular variate x:

$$\begin{aligned} x &= a + [u(b-a)(m-a)]^{0.5} \qquad 0 \leqq u \leqq (m-a)/(b-a) \\ &= b - [(1-u)(b-a)(b-m)]^{0.5} \qquad (m-a)/(b-a) \leqq u \leqq 1 \end{aligned}$$

A great advantage of having a simulation model of an investment process is the relative ease afforded to exploring the consequences of stochastic dependence resulting from the joint dependence of some variables on other common variables. Suppose, for instance, that several agricultural investment possibilities are jointly related to seasonal rainfall experience and international commodity price realizations. By relating farm yields to rainfall variables and farm prices to some indices of international economic activity, it is then possible to model directly many of the stochastic dependencies applicable. Such indirect handling of correlations through hierarchical structures of dependency may be simpler and more expedient than direct elicitation of joint distributions. Information on such structures is just as subjective as that on cash flows, probabilities, covariances, etc.; and all can be elicited with such techniques as discussed in Chapter 2.

The manner in which output from a simulation model is handled depends mainly on the extent to which the utility function is known. When it is fully specified, it can be incorporated within the model, and repeated encounters with (replicated samples from) the model can be used to estimate utilities of alternative investment actions directly (Dillon, 1971). At the other extreme when utility is not specified, the model can be used to estimate a variety of performance statistics such as means, variances, and higher moments of present value and (most desirably for the methods reviewed in Chapter 9) estimates of complete CDFs for measures such as present value, internal rate of return, etc.

The inevitable complexity of simulation models has been observed and poses problems of exposition here. We could construct a very simplistic

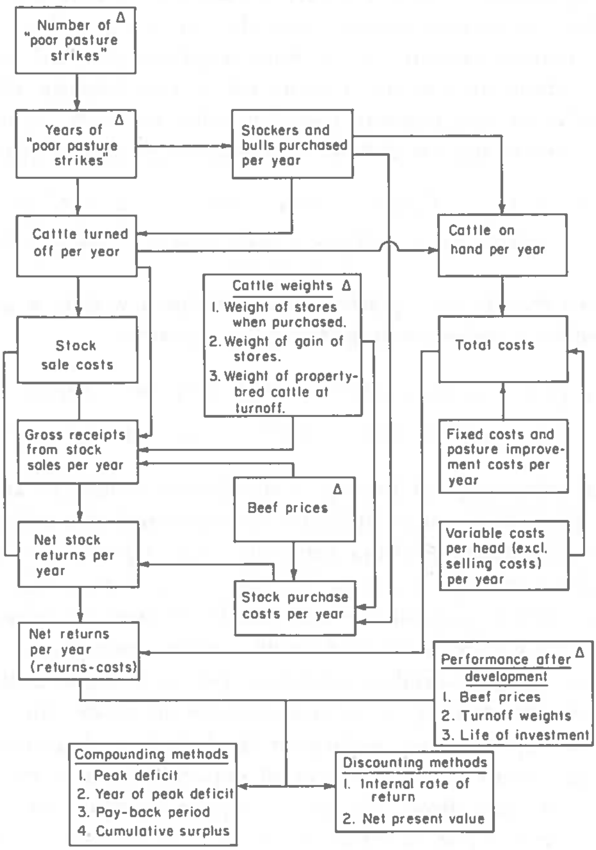


FIG. 8.2. Flow diagram for simulation of an improved pasture investment. (Δ indicates the Monte Carlo selection of the variable from a triangular distribution.)

model to illustrate the mechanics of Monte Carlo sampling but, as indicated, this process is so simple as to achieve little in the way of exemplifying the overall simulation approach. Unfortunately, it is beyond our present purpose to become embroiled in a lengthy description of a more realistic simulation. Our compromise is to look briefly at a practical investment problem and to review the type of information available from a stochastic simulation. The problem concerns a pasture improvement program in the spear grass cattle grazing region of Australia. It has been described in a nonstochastic (best estimate) manner by Haug and Hirst (1967) and analyzed by means of stochastic simulation by Cassidy et al. (1970). These simulators made extensive use of the triangular distribution to represent uncertain quantities in their model, which is summarized by the flow diagram of Figure 8.2.

Relatively complex dependencies are introduced through the sampling of success in pasture establishment, which leads to realized carrying capacities and turnoffs of cattle in subsequent periods. Presumably, within each year the sampled weights of cattle sold were perfectly correlated with each other by using the same uniform variate in computing each triangularly distributed variate. The analysts chose not to sample beef selling prices but rather repeated their simulation for three fixed price levels. A more complete analysis would have also modeled these uncertain prices (again perhaps with triangular distributions, but also allowing for dependence between successive periods), and our illustrative data are speculations as to the types of results that could be so obtained. Two types of illustrative data are presented in Figure 8.3, viz., CDFs for net present values and also for internal rate of return values. Cassidy et al. (1970) also present CDFs for peak deficit, cumulative surplus, year of peak deficit, and payback period.

Curves for CDFs such as those of Figure 8.3 are formed most simply by recording a large number of encounters with the model, say $n \geqq 100$ or 200, ascribing a cumulative probability of $1/n$ to each of the ranked observations and then plotting the data. With n large, this procedure produces results indistinguishable from those obtained with the more correct sparse data rule of Section 2.4. The lower part of Figure 8.3 indicates the sensitivity of the PV results to changes in the rate of discounting. Note that there is a link between internal rate of return (IRR) and PV CDFs (established by recalling that the IRR is defined as a discount rate that sets PV to zero), viz., $P(\text{IRR} < r) = P(PV < 0 \mid r)$.

Analogous results for other investments and feasible combinations of investments can be developed and compared and contrasted with each other (perhaps using the stochastic efficiency concepts of Chapter 9) to aid

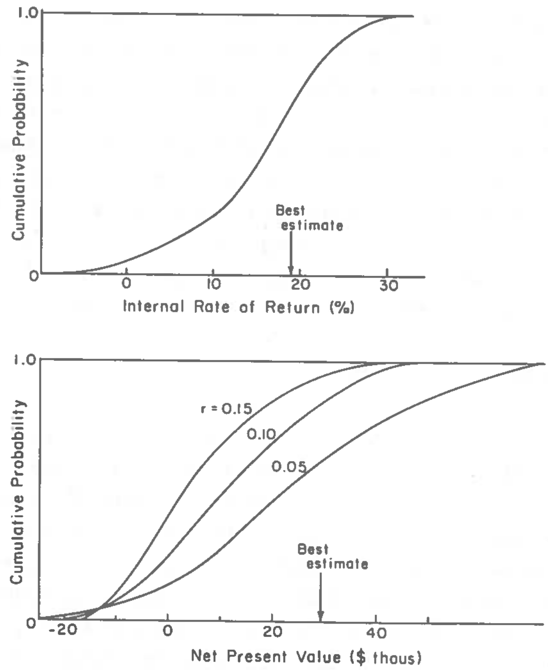


FIG. 8.3. Illustrative CDF results for an investment in pasture improvement.

in investment decision making. When the utility function is known and embedded in the model, the results would include direct estimates of expected utility associated with each alternative examined. Hopefully, this would rank the possibilities unambiguously—in contrast to the often ambiguous indications emerging from consideration of several alternative measures like expected IRR, expected *PV*, expected payback period, etc.

PROBLEMS

8.1. Using the benchmark approach, determine optimal choice between the alternatives of Table 8.1 under the assumption that (a) you are the decision maker, (b) a colleague is the decision maker.

8.2. Given that $r = 0.08$ and the initial outlay is restricted to $24,000, determine the optimal investment on the basis of $U(PV)$ for a decision maker for whom $U = 2X - 0.01X^2$ where X is in thousands of dollars. The available projects are independent, with cash flow and conditional probability characteristics as listed below. Assume that the projects are indivisible.

C_{it}	C_{1t}	C_{2t}	C_{3t}	C_{4t}
		($\$10^3$)		
H_{i0}	−6	−5	−10	−20
L_{i0}	−10	−12	−25	−22
H_{i1}	20	20	50	40
M_{i1}	0	18	30	30
L_{i1}	0	12	20	20
H_{i2}	22	18	75	45
M_{i2}	16	0	60	35
L_{i2}	14	0	40	25

Probability	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$P(H_{i0})$	0.5	0.7	0.1	0.2
$P(L_{i0})$	0.5	0.3	0.9	0.8
$P(H_{i1} \mid H_{i0})$	1.0	0.3	0.2	0.6
$P(M_{i1} \mid H_{i0})$	0.0	0.4	0.5	0.1
$P(L_{i1} \mid H_{i0})$	0.0	0.3	0.3	0.3
$P(H_{i1} \mid L_{i0})$	1.0	0.1	0.1	0.6
$P(M_{i1} \mid L_{i0})$	0.0	0.7	0.8	0.1
$P(L_{i1} \mid L_{i0})$	0.0	0.2	0.1	0.3
$P(H_{i2} \mid H_{i1})$	0.2	1.0	0.6	0.3
$P(M_{i2} \mid H_{i1})$	0.7	0.0	0.2	0.2
$P(L_{i2} \mid H_{i1})$	0.1	0.0	0.2	0.5
$P(H_{i2} \mid M_{i1})$	0.1	1.0	0.3	0.6
$P(M_{i2} \mid M_{i1})$	0.3	0.0	0.3	0.3
$P(L_{i2} \mid M_{i1})$	0.6	0.0	0.4	0.1
$P(H_{i2} \mid L_{i1})$	0.3	1.0	0.5	0.3
$P(M_{i2} \mid L_{i1})$	0.6	0.0	0.3	0.4
$P(L_{i2} \mid L_{i1})$	0.1	0.0	0.2	0.3

8.3. (a) Depict the investment possibilities of Problem 8.2 in the form of a decision tree.
(b) Assuming $U_t(C_t) = 2C_t - 0.01C_t^2$, $C_t = \$10^3$, evaluate the problem of Problem 8.2 using the certainty equivalent approach to $U(PV)$.

8.4. Assuming $U_t(C_t) = 2C_t - 0.01C_t - 0.6C_{t-1} - 0.3C_{t+1}$, $C_t = \$10^3$, evaluate Problem 8.2 using the certainty equivalent $U(PV)$ approach. What if $U_t(C_t) = 2C_t - 0.01C_t^2 - 0.2U(C_{t-1})$? Comment on these forms of utility function.

8.5. (a) Discuss the relationship between portfolio analysis (Chapter 7) and investment analysis.
(b) "Intuition has a legitimate role in risky investment appraisal." Discuss this statement.

8.6. (a) Using the project data of Problem 8.2, illustrate what is meant by joint preferential independence and utility independence.
(b) Solve Problem 8.2, given the following information. Note that the utility functions $U_t(C_t)$ need to be scaled from 0 for the least preferred C_t value to 1 for the most preferred C_t value.

t	k_t	$U_t(C_t)$
0	0.3	$C_0 - 0.010C_0^2$
1	0.1	$C_1 - 0.008C_1^2$
2	0.5	$C_2 - 0.010C_2^2$

8.7. Arrange the data of Problem 8.2 in the form of a time-state preference payoff matrix with states of nature $\Theta_1, \Theta_2, \ldots,$ and consequences (C_0, C_1, C_2).

8.8. Comment on the following statements:
(a) "Decision analysis has little of import to offer the practising investment analyst."
(b) "In the analysis of investment choice it may be far better to simplify the analysis by considering alternatives on a one-at-a-time basis, ignoring their interrelationships, rather than to do violence to risk preferences by assuming a linear utility function. For example, a sophisticated capital budgeting model that is based on maximizing net present value under assumed certainty may be far worse than single project appraisal based on nonlinear risk preference."

8.9. *The central limit theorem* is often approximately applicable in estimating distributions of present value. It states, roughly, that if a random variable arises as the sum of a (large) number of independently and identically distributed random variables, its distribution will tend to normal. Take m samples of size n of uniform $(0 < x < 1)$ random numbers, sum them, estimate empirical distribution functions, and test these for normality. (If a computer is used, use $m = 100$, $n = 3, 5, 7, \ldots, 15$. Otherwise, use $n = 3, 6, 9$ and, with $m = 9$, employ the sparse data smoothing procedure to estimate CDFs.) Draw conclusions about how large n needs to be for the theorem to apply tolerably well.

8.10. Develop and sketch a flow diagram for a relatively simple agricultural investment possibility. Indicate which variables are stochastic and any stochastic dependencies that should be specified in a subsequent model.

8.11. Relative to the triangular distribution of x in Figure 8.1, let $a = \$0$, $m = \$6000$, and $b = \$8000$. Assume that this x is the annual cash flow for each of the 10 years of life of a project requiring an initial outlay of \$50,000 and with a terminal salvage value of \$10,000. Assuming further that cash flows in successive years are independent, compute the mean and variance of net present value $(r = 0.1)$ and comment on the nature of the probability distribution of the present value of this project.

8.12. "Simulation has the advantages of being versatile and easy to understand, so it has been a natural first-generation approach to risk analysis. Unfortunately, it has a number of serious disadvantages ... [including being] inherently imprecise, ... cumbersome, ... expensive ... [and] yields no additional insight into cause-and-effect relationships ... [and] provides no guidance on how ... to select the investments ...". (Hillier 1969, pp. 84–85). Comment.

8.13. Large agricultural projects generally have multiple objectives such as net cash flow, regional effects, linkage effects, balance of payments effects, etc., all of which have risky time traces. How might multidimensional utility analysis be used in ex ante evaluation of such projects? [*Hint:* See Bell (1975b) and Edwards (1976).]

REFERENCES AND SELECTED FURTHER READING

Adelson, R. M. 1965. Criteria for capital investment: An approach through decision theory. *Opl. Res. Q.* 16(1): 19–50.

A seminal paper criticizing the arbitrary nature of discounted cash flow techniques of risky project appraisal. Argues for the utility of present value approach.

Anderson, J. R. 1974. Simulation: Methodology and application in agricultural economics. *Rev. Mktg. Agric. Econ.* 42(1): 3–55.

A review of simulation studies in agricultural management.

Ang, J. S. 1974. State-preference application in capital budgeting. *Engng. Econ.* 19(3): 195–208.

Discusses the practicality of the state-preference approach to investment appraisal and argues the relative merit of a risk-constrained programming approach.

Arrow, K. J., and R. C. Lind. 1970. Uncertainty and evaluation of public investment decisions. *Am. Econ. Rev.* 60(3): 364–78. Comments by E. J. Mishan, R. N. McKean, J. H. Moore and D. Wellington; reply by K. J. Arrow and R. C. Lind. *Am. Econ. Rev.* 62(1): 161–72.

Discusses the question of whether the state should allow for risk in the appraisal of public investment projects. The conclusion (not undisputed) is that the state's utility function should be linear for project appraisal.

Bell, D. E. 1975a. A utility function for time streams having inter-period dependencies. Rept. RR-75-22, Intl. Instit. Applied Systems Analysis, Laxenburg.

Presents formal analysis of time streams as multiattributed consequences.

———. 1975b. A decision analysis of objectives for a forest pest problem. Rept. RR-75-43, Intl. Instit. Applied Systems Analysis, Laxenburg.

Applies the conditional utility independence procedures developed in Bell (1975a) for multiattribute time-stream appraisal.

Bernhard, R. H. 1969. Mathematical programming models for capital budgeting—A survey, generalization and critique. *J. Financ. Quant. Anal.* 4(2): 111–58.

A good article well described by its title. Has no utility orientation.

Borch, K. 1968. Decision rules depending on the probability of ruin. *Oxford Econ. Papers* 20(1): 1–10.

Strongly utility oriented and critical of approaches that fail to recognize the sequential pattern of cash flows.

———. 1970. Specification of objectives in decision problems. *Theory and Decision* 1: 5–21.

Highlights the need to allow for the decision maker's preferences for the time pattern of cash flows. Shows this to be equivalent to establishing preference orderings over sets of stochastic processes.

Bromwich, M. 1970. Capital budgeting–A survey. *J. Busin. Finance* 2 (3): 3–25.

A broad survey with an extensive annotated bibliography.

Byrne, R. F., A. Charnes, W. W. Cooper, and K. Kortanek. 1971. A chance-constrained approach to capital budgeting with portfolio-type payback and liquidity constraints and horizon posture controls. In *Studies in Budgeting,* R. F. Byrne et al., eds., pp. 71–92. Amsterdam: North-Holland.

An example of the chance-constrained present value approach to risky investment appraisal.

Candler, W., and M. Boehlje. 1971. Use of linear programming in capital budgeting with multiple goals. *Am. J. Agric. Econ.* 53 (2): 325–30.

Shows how linear programming may be used interactively between analyst and decision maker in deterministic capital budgeting with multiple goals.

Cassidy, P. A., J. L. Rodgers, and W. O. McCarthy. 1970. A simulation approach to risk assessment in investment analysis. *Rev. Mktg. Agric. Econ.* 38 (1): 3–24.

Chisholm, A. H., and J. L. Dillon. 1971. *Discounting and Other Interest Rate Procedures in Farm Management,* 3rd ed. Armidale, Australia: Univ. New England.

A guide to discounted cash flow procedures. Contains extensive interest rate and related tables.

Dean, G. W. 1974–75. Firm theory incorporating growth and risk: Integration into farm management research. Suppl. *Int. J. Agr. Aff.,* double number, pp. 110–20.

Suggests farm-firm capital budgeting should be based on a lexicographic utility function including present value.

Dent, J. B., and J. R. Anderson, eds. 1971. *Systems Analysis in Agricultural Management.* Sydney, Australia: Wiley.

Contains several agricultural applications of simulation but with little emphasis on investment appraisal.

Dillon, J. L. 1971. Interpreting systems simulation output for managerial decision-making. In *Systems Analysis in Agricultural Management,* J. B. Dent and J. R. Anderson, eds., pp. 85–120. Sydney, Australia: Wiley.

Illustrates the use of utility appraisal within a simulation model.

Edwards, W. 1976. How to use multi-attribute utility measurement for social decision making. Soc. Sci. Res. Instit., Univ. Southern California, Los Angeles.

Eilon, S., and T. R. Fowkes. 1973. Sampling procedures for risk simulation. *Opl. Res. Q.* 24 (2): 241–52.

Investigates alternative sampling schemes in stochastic simulation of investment analyses.

Elton, E. J., and M. J. Gruber. 1974. On the optimality of some multi-period portfolio selection criteria. *J. Busin.* 47 (2): 231–43.

Examines various investment criteria in an expected utility of present or terminal value context.

Fairley, W., and H. D. Jacoby. 1975. Investment analysis using the probability distribution of the internal rate of return. *Mgmt. Sci.* 21 (12): 1428–37.

The probability distribution of the internal rate of return is derived under the assumption that net cash flows can be described by a multivariate normal distribution. An application of the analysis is presented.

FAO. 1971. *General Guidelines to the Analysis of Agricultural Production Projects.* Agric. Plann. Stud. 14. Rome: FAO.

A practical handbook oriented to discounted cash flow procedures.

Gittinger, J. P. 1972. *Economic Analysis of Agricultural Projects.* Baltimore: Johns Hopkins Press.

A handbook for budding project appraisers, largely oriented to public projects, excellent except for its focus on the internal rate of return as the criterion for investment appraisal.

Hanssman, F. 1968. Probability of survival as an investment criterion. *Mgmt. Sci.* 15(1): 33–48.

Harberger, A. C. 1972. *Project Evaluation: Collected Papers.* Chicago: Markham.

Oriented to the economic appraisal of public projects in a macro context. Argues for the present value approach.

Haug, N. F., and G. G. Hirst. 1967. Pasture improvement budgeting in the spear grass zone of Queensland. *Q. Rev. Agric. Econ.* 20(3): 113–26.

Hax, A. C., and K. M. Wiig. 1976. The use of decision analysis in capital investment problems. *Sloan Mgmt. Rev.* 17(2): 19–48.

A recommended exposition.

Hertz, D. B. 1964. Risk analysis in capital investment. *Harvard Busin. Rev.* 42(1): 95–106.

An early expository article that discusses stochastic simulation for investment appraisal, with emphasis on internal rate of return distributions.

Hespos, R. F., and P. A. Strassmann. 1965. Stochastic decision trees for the analysis of investment decisions. *Mgmt. Sci.* 11(10): B244–59.

An extension of Hillier (1963) to decision tree format.

Hillier, F. S. 1963. The derivation of probabilistic information for the evaluation of risky investments. *Mgmt. Sci.* 9(3): 443–57; Suppl. note, 11 (3): 485–87.

First presentation of the concept of a probability distribution of present value.

———. 1969. *The Evaluation of Risky Interrelated Investments.* Amsterdam: North-Holland.

A rather formal but comprehensive study largely based on the utility of present value approach. Chapter 7 presents a chance-constrained programming model.

———. 1971. A basic approach to the evaluation of risky interrelated investments. In *Studies in Budgeting,* R. F. Byrne et al., eds., pp. 3–47. Amsterdam: North-Holland.

A less formal exposition of the material in the previous reference, together with the results of extensive computational experimentation. Highly recommended.

Hirshleifer, J. 1970. *Investment, Interest, and Capital.* Englewood Cliffs, N.J.: Prentice Hall.

Part D of Chapter 8 presents the basic theory of preference involving both time and uncertainty.

James, E. 1975. A note on uncertainty and the evaluation of public investment decisions. *Am. Econ. Rev.* 65(1): 200–205.

Elaborates the implicit assumptions made by Arrow and Lind (1970).

Jean, W. H. 1970. *The Analytical Theory of Finance.* New York: Holt, Rinehart and Winston.

Outstanding in its exposition, this excellent text favors an additive utility function ($U = \sum k_t U_t$) approach to the evaluation of cash flow sequences. Chapters 1, 2, and 5 discuss alternative criteria for cash flow assessment, capital budgeting models, risk over time, and the state preference approach.

Kaplan, S., and N. R. Barish. 1967. Decision-making allowing for uncertainty of future investment opportunities. *Mgmt. Sci.* 13(10): B569–77.

Considers the evaluation of the probability distribution of present value when the interest rate in future periods is uncertain. This difficulty is generally ignored in the literature.

Keeney, R. L., and H. Raiffa. 1976. *Decisions with Multiple Objectives: Preferences and Tradeoffs.* New York: Wiley.

Chapter 9 discusses preferences over time. Highly recommended.

Lee, S. M. 1972. *Goal Programming for Decision Analysis.* Philadelphia: Auerbach.

An introduction to goal programming.

Little, I. M. D., and J. A. Mirrlees. 1974. *Project Appraisal and Planning for Developing Countries.* London: Heinemann.

Deals with project appraisal mainly in a social context. Present social value is the criterion computed by using "accounting prices" that are intended to reflect the real value of the social costs and benefits. Chapter 15 deals briefly with uncertainty in project appraisal.

Machol, R. E., and E. M. Lerner. 1969. Risk, ruin and investment analysis. *J. Financ. Quant. Anal.* 4(4): 473–92.

Defines risk in terms of a chance constraint relative to the probability of ruin and explores its implications.

Mao, J. C. T. 1969. *Quantitative Analysis of Financial Decisions.* London: Collier-Macmillan.

A well exposited text closely related to the major literature in the field.

Massé, P. 1962. *Optimal Investment Decisions.* Englewood Cliffs, N.J.: Prentice-Hall.

An early work highly recommended for its interesting discussion of risky investment, including the problems of flexibility and extreme disasters.

Meier, R. C., W. T. Newell, and H. L. Pazer. 1969. *Simulation in Business and Economics.* Englewood Cliffs, N.J.: Prentice-Hall.

A good general reference on simulation methods.

Meyer, R. F. 1970. On the relationship among the utility of assets, the utility of consumption, and investment strategy in an uncertain, but time-invariant, world. In *Operational Research '69: Proc. 5th Int. Conf. Opl. Res.*, J. R. Lawrence, ed., pp. 627–48. London: Tavistock.

The classic paper on the utility of risky time sequences. Gives the theoretical justification for a multidimensional approach to multiperiod flows.

———. 1975. Preferences over time. In *Proceedings of a Workshop on Decision Making with Multiple Conflicting Objectives,* Vol. 4, Intl. Instit. Applied Systems Analysis, Laxenburg.

An excellent survey. Highly recommended.

Mihram, G. A. 1972. *Simulation: Statistical Foundations and Methodology.* New York: Academic Press.

A good exposition of simulation methods with particular emphasis on Monte Carlo sampling procedures.

———. 1976. Simulation methodology. *Theory and Decn.* 7(1/2): 67–94.

Morris, W. T. 1968. *Management Science: A Bayesian Introduction.* Englewood Cliffs, N.J.: Prentice-Hall.

Of interest for its consideration of Bayes' theorem in relation to investment analysis and the allocation of resources to search and evaluation.

Mossin, J. 1968. Optimal multi-period portfolio policies. *J. Busin.* 41(2): 215–29.

One of the best and earliest considerations of investment in the portfolio context.

Murphy, M. C. 1968. A stochastic approach to investment appraisal. *Fm. Economist* 11(7): 304–18.

Discusses analytical estimation of means and variances of present value and internal rate of return and emphasizes accounting for serial correlation effects.

Näslund, B. 1971. A model of capital budgeting under risk. In *Studies in Budgeting,* R. F. Byrne et al., eds., pp. 48–70. Amsterdam: North-Holland.

A chance-constrained extension of Weingartner's (1963) model.

Naylor, T. H. 1971. *Computer Simulation Experiments with Models of Economic Systems.* New York: Wiley.

A useful compendium of simulation methods and applications but with little attention to investments by individuals.

Pouliquen, L. Y. 1970. *Risk Analysis in Project Appraisal.* Baltimore: Johns Hopkins Press.

Emphasizes estimation of distributions of internal rate of return for World Bank assisted projects.

Rae, A. N. 1970. Capital budgeting, intertemporal programming models with particular reference to agriculture. *Australian J. Agric. Econ.* 14(1): 39–49.

Uses an additive multiattribute objective function.

Raiffa, H. 1969. Preferences for multi-attributed alternatives. RM-5868-DOT/RC, The RAND Corp., Santa Monica, Calif.

Part 9 brilliantly discusses the problem of time influence in investment choice.

Reutlinger, S. 1970. *Techniques for Project Appraisal under Uncertainty.* Baltimore: Johns Hopkins Press.

Illustrates the use of stochastic simulation as an aid to appraisal of World Bank assisted projects. Recommended for its practical orientation and case illustrations.

Sadan, E. 1970. The investment behavior of a farm firm operating under risk. *Am. J. Agric. Econ.* 52(4): 494–504.

Explores some implications of risk aversion in investment decisions by simulating means and variances of present value.

Sandmo, A. 1970. Discount rates for public investment under uncertainty. In *Allocation under Uncertainty,* J. H. Dreze, ed., pp. 192–210. London: Macmillan.

Complementary to Arrow and Lind (1970).

Sprow, F. B. 1967. Evaluation of research expenditures using triangular distribution functions and Monte Carlo methods. *J. Ind. Engng. Chem.* 59(7): 35–38.

Van Horne, J. 1971. *Financial Management and Policy,* 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.

Provides a good treatment of mathematical programming approaches to capital budgeting under risk.

von Winterfeldt, D., and G. W. Fischer. 1973. Multiattribute utility theory: Models and assessment procedures. In *Utility, Probability and Human Decision Making,* D. Wendt and C. Vlek, eds., pp. 47–85. Dordrecht: Reidel.

Includes discussion of time sequences as multidimensional alternatives.

Wagle, B. 1967. A statistical analysis of risk in capital investment projects. *Opl. Res. Q.* 18(1): 13–33.

Complementary to Hillier (1963).

Wagner, H. M. 1969. *Principles of Operations Research.* Englewood Cliffs, N.J.: Prentice-Hall.

Chapter 11 gives interesting discussion of present value and time sequences.

Weingartner, H. M. 1963. *Mathematical Programming and the Analysis of Capital Budgeting Problems.* Englewood Cliffs, N.J.: Prentice-Hall.

The seminal work in its area.

———. 1966. Capital budgeting of interrelated projects: Survey and synthesis. *Mgmt. Sci.* 12(5): 483–516.

Comprehensive survey of techniques available for investment appraisal as of 1966.

Whitlam, G. B., S. R. Harrison, and T. D. Wilson. 1970. Methods of evaluation of farm development projects. Tech. Bull. 7, Econ. Serv. Branch, Queensland Dept. Primary Ind., Brisbane.

Guidebook to the use of discounted cash flow techniques. Ignores risk.

Wilson, R., 1969. Investment analysis under uncertainty. *Mgmt. Sci.* 15(12): B650–64.

Exposits a state preference (i.e., event tree with cash flows) approach to individual project appraisal. Allows for intertemporal utility effects within a programming format.

Williams, F. E. 1973. On the evaluation of intertemporal outcomes. In *Multiple Criteria Decision Making,* J. L. Cochrane and M. Zeleny, eds., pp. 429–38. Columbia: Univ. South Carolina Press.

Shows the need for utility appraisal that allows for preferences over stochastic cash flow sequences.

CHAPTER NINE
DECISION ANALYSIS WITH PREFERENCES UNKNOWN

Our approach to decision analysis has been based on the assumption that the preferences of an individual decision maker could be obtained, quantified, and employed directly in the analysis. Sometimes we have sidestepped the elicitation process and have simply assumed a particular algebraic representation. Likewise, we have sometimes arbitrarily assumed parameters to quantify the extent of risk aversion. Our rationale is that so far our purpose has been to exemplify methods that are applicable when preferences can be elicited and described.

There are many occasions (perhaps a majority) in agricultural management when analysts, for reasons of cost or expediency, cannot obtain appropriately elicited preferences. Arbitrary assumption of a preference function in such circumstance is clearly unsatisfactory since subsequent decision analysis can then only yield a "correct" answer by chance! All is not lost, however, without such arbitrary recourse. As we show in this chapter, provided something about preference may be presumed, it is possible to proceed some distance toward a conventional identification of the best decision.

In this second-best of worlds, we must sacrifice the pursuit of an optimal decision. Instead, we have to search for an efficient set of decisions in the sense that decisions in the set are undominated and hence admissible. We can only hope that this set will be small or closely confined. With all else equal, the more that can be assumed about preference, the smaller the efficient set will tend to be. It should come as no surprise that if nothing can be assumed about preference, nothing can be done to identify decisions that are efficient.

9.1 CONCEPTS OF STOCHASTIC EFFICIENCY

Our discussion of stochastic efficiency will proceed as we progressively introduce more and more restrictive preference assumptions while attempting to retain defensible generality in the validity of the assumptions.

First-Degree Stochastic Efficiency

The first concept of stochastic efficiency to be elaborated and the first to be formalized—by Quirk and Saposnik (1962) and Fishburn (1964)—rests on a most reasonable behavioral assumption that we have already presumed in our discussion of Bernoulli's principle in Section 4.2. This is the basic idea that if x is the unscaled measure of consequence such as profit, decision makers always prefer more to less of x. More formally, this is nothing but the assumption of a monotonically increasing utility function wherein the first derivative is strictly positive, i.e., $U_1(x) > 0$.

The initial efficiency concept (viz., that of first-degree stochastic efficiency) needs to be stated in terms of cumulative probability functions. Consider the case of a pair of continuous CDFs F_1 and G_1 defined within the range $[a, b]$ and respectively associated with two acts or risky prospects F and G. Recall, for instance, that F_1 is related to its PDF $f(x)$ by

$$F_1(R) = \int_a^R f(x)dx \tag{9.1}$$

F is said to dominate G in the sense of first-degree stochastic dominance (FSD) *if $F_1(R) \leqq G_1(R)$ for all possible R in the range $[a, b]$ with at least one strong inequality (i.e., the $<$ holds for at least one value of R).* This efficiency criterion, like all those to be discussed subsequently in this section, is transitive; if F dominates G and G dominates H, F must dominate H.

In graphical terms (as illustrated in Figure 9.1) this rule means that a first-degree stochastically dominant CDF curve must lie nowhere to the left

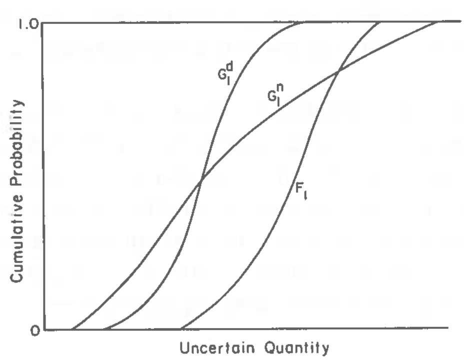


FIG. 9.1. Illustration of FSD (F_1 dominates G_1^d but not G_1^n).

of a dominated curve. Distributions that are dominated in this sense are said to be stochastically inefficient and, conversely, those (in general intersecting) distributions that are not so dominated are said to be stochastically efficient of first degree (FSE). Inefficient distributions are those that would never be preferred by Bernoullian utility maximizers when confronted with the set of efficient distributions. Among the efficient set, identification of the single most preferred distribution depends on knowing more about preference than is so far assumed. The central result here also holds in converse; viz., if F is preferred to G by all utility maximizers with $U_1(x) > 0$, then F dominates G in the sense of FSD.

Discrete distributions are also accommodated in such concepts of efficiency and dominance. Suppose that x takes only a finite number of values x_i, $i = 1, \ldots, n$, all in the interval $[a, b]$. A mass function $f(x_i)$ can be attached with the x_i arranged in ascending order, and thus a CDF is defined as

$$F_1(R) = P(x_i \leqq R) = \sum_{\substack{\text{all} \\ x_i \leqq R}} f(x_i) \tag{9.2}$$

and FSD is defined as before except that now the inequality needs only be examined at the discrete x_i values.

We have already encountered a special case of this discrete distribution rule as early as Section 1.2 when concepts of dominance were initially introduced. Reconsider the simplest conceivable discrete payoff table (Table 9.1) where $A > B$, $C > D$:

TABLE 9.1. Discrete Payoff Table

$P(\Theta_i)$	a_1	a_2
$P(\Theta_1)$	A	B
$P(\Theta_2)$	C	D

We would have said earlier simply that a_1 dominates a_2 and that, accordingly, a_2 can be eliminated from the decision analysis. Such dominance can now be seen as a special case of FSD if the cumulative mass functions for the two acts are sketched as in Figure 9.2. The a_1 cumulative mass function lies nowhere to the left of that for a_2 and so dominates it in the sense of FSD. This is a rather special case of dominance because since the same set of states applies to all the acts, the dominance relation is uninfluenced by the allocation of the probabilities between the states.

First-degree stochastic efficiency is of importance historically and also

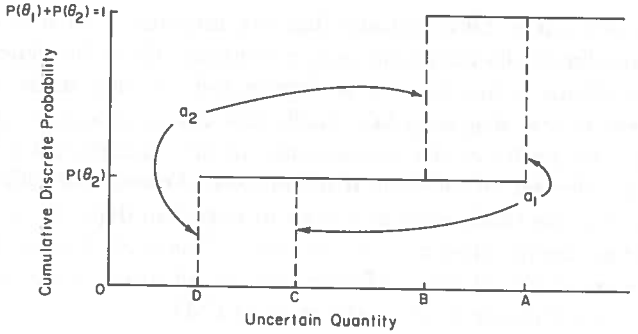


FIG. 9.2. Sketch of a special case of FSD.

didactically as a basis for introducing the general notion and subsequent extensions. As an empirical matter FSE is perhaps not so important because, generally speaking, relatively few acts (distributions) can be eliminated in this way. A related consideration is that it tends to be the rule rather than the exception that CDFs from different families and indeed CDFs from the same family intersect at least once, thereby predisposing against the chance of identifying any FSD. It is therefore of significant operational advantage to seek more restrictive concepts of efficiency so that rather larger numbers of feasible actions can be discarded to leave a smaller efficient set. This is done by the second concept of stochastic efficiency to be introduced.

Second-Degree Stochastic Efficiency

Second-degree stochastic efficiency (SSE) provides a basis for eliminating distributions from the FSE set that are inefficient or dominated in the sense of second-degree stochastic dominance (SSD). The rule for identifying cases of SSD was discovered independently by Fishburn (1964), Hanoch and Levy (1969), Hadar and Russell (1969), and Hammond (1968) and depends only on the additional behavioral assumption that the decision maker is averse to risk. In terms of the utility function over the range $[a, b]$ of possible payoffs, the presumption is that the function is not only monotonically increasing but also strictly concave. As has been elaborated in Section 4.5, this is equivalent to assuming that $U_1(x) > 0$ and $U_2(x) < 0$.

The ordering rule must again be stated in terms of cumulative probability functions and can be understood most intuitively in terms of CDFs. A distribution function F_1 dominates another G_1 if it lies more to the right in terms of differences in area between the CDF curves cumulated from the lower values of the uncertain quantity. This is depicted in the upper dia-

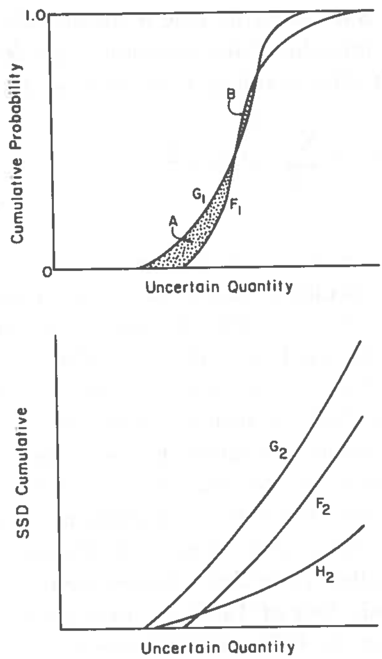


FIG. 9.3. Illustration of SSD where CDFs cross twice (area A > area B).

gram of Figure 9.3 where the area marked A exceeds the area marked B. Such a relationship is assessed most succinctly by defining a further type of cumulative function that measures the area under a CDF over the range of the uncertain quantity. For instance, define the SSD cumulative for a distribution F_1 as

$$F_2(R) = \int_a^R F_1(x)\,dx \tag{9.3}$$

Then the distribution F *is said to dominate* G *in the sense of* SSD *if* $F_2(R) \leq G_2(R)$ *for all possible* R *with at least one strong inequality.* Such a case of dominance is depicted in the lower diagram of Figure 9.3 where F_2 dominates G_2 but not H_2 by SSD. Pairwise comparison of distributions in this manner means that they can be sorted into two sets. Dominated distributions are revealed as inefficient in that they would never be preferred by risk-averse utility-maximizing decision makers. The remaining undominated distributions constitute the SSE set, e.g., the acts corresponding to F_2 and H_2 of Figure 9.3. Analogously to the FSE case, identification of choice within this set depends on knowing more about preference than merely that an unquantified aversion to risk exists.

A version of the SSD ordering rule is applicable to discrete distributions. In the notation introduced for equation (9.2) define $\Delta x_i = x_i - x_{i-1}$; and if x_n is the highest value taken by x, the analog of F_2 is defined by

$$F_2(x_r) = \sum_{i=2}^{r} F_1(x_{i-1})\Delta x_i \qquad r = 2, \ldots, n$$

$$F_2(x_1) = 0 \tag{9.4}$$

Then for SSD we need to have $F_2(x_r) \leq G_2(x_r)$ *for all* $r \leq n$ *with at least one strict inequality*. These calculations are simply exemplified by the illustrative decision problem of Section 5.2. The calculations are summarized in Table 9.2 and indicate that both acts are SSE. Actually, with greater familiarity with the efficiency rules, this should occasion no surprise. One of the mathematical features of all the ordering rules we have discussed and will discuss is that a necessary condition for one distribution to dominate another is that its mean not be less. A second necessary condition for dominance is that the smallest value of a dominant distribution cannot be less than the smallest value of a dominated distribution. Thus knowing that a_2 has a lower smallest value and a higher mean than a_1, it is unnecessary to make the calculations of Table 9.2 to come to the conclusion that both acts are efficient in the FSE and SSE senses.

TABLE 9.2. Efficiency Analysis of the Example of Section 5.2

x_i	(ranked payoffs)	6150	6800	11500
$f(x_i)$	(a_2 probabilities)	0.8	0	0.2
$g(x_i)$	(a_1 probabilities)	0	1.0	0
$F_1(x_i)$	(cumulative probabilities)	0.8	0.8	1.0
$G_1(x_i)$	(cumulative probabilities)	0	1.0	1.0
Δx_i	(payoff first differences)	...	650	4700
$F_2(x_i)$	(SSD cumulatives)	0	520	4280
$G_2(x_i)$	(SSD cumulatives)	0	0	4700

We have attempted to give as much generality as possible to the discussion by saying nothing about the nature of the distributions except whether they are continuous or discrete. The cost of this generality is the implied chore of checking through all the steps of the efficiency criteria discussed. Whenever necessary and possible, such checking is best done graphically for continuous distributions and in tabular layouts like Table 9.2 for discrete distributions. Unfortunately, this work can become exceedingly tedious when large numbers of distributions must be reviewed in the necessary pairwise fashion. Then it is desirable to resort to electronic computational procedures such as discussed in Section 9.2. Another way of simplifying the review procedures depends on severely restricting the types of probability distributions that are analyzed.

Perhaps the theoretical distribution most commonly used in decision analysis is the normal distribution. Its efficiency analysis is relatively simple. Recall that all normal distributions apply to the range $[-\infty, +\infty]$. Unless a pair of distributions has identical variance, the normal CDFs will intersect somewhere within the range. This means that, in general, FSD is not a possibility for normal distributions. A normal distribution can only be dominated by another in this sense when it has a lesser mean and precisely the same variance. We could, however, speak of "approximate FSD" by ignoring very low probability CDF intersections, and this has been pursued by Anderson (1974b, pp. 167–68). The SSD case is rather more interesting than FSD for normal distributions, and the ordering rule here devolves to simple comparisons among the parameters of the specified distributions. Specifically, *a normal F dominates a normal G in the sense of* SSD *if* $E_F(x) \geqq E_G(x)$ *and* $V_F(x) \leqq V_G(x)$ *with at least one strong inequality, where E is the expectation operator and V is the variance operator.*

The above rule is nothing other than the (E, V) criterion we met in the discussion of portfolio selection in Section 7.1. Thus with normally distributed variables, the (E, V)-efficient set (or frontier) is equivalent to the SSE set. Naturally, (E, V) analysis is not confined to portfoliolike problems. For instance, we could return to the maize-nitrogen example of Section 6.1 and conduct an (E, V) analysis of the data in equations (6.17) and (6.18) and the paragraph that follows these equations. With the additional assumption that maize yields are normally distributed, the (E, V)-efficient range of rates of nitrogen is from 0 to 100 kg/ha. This wide range results from the fact that variance of yield increases linearly over the entire range and expected profit increases monotonically from 0 to 100 kg/ha where negative marginal expected profit sets in.

Mean-variance or (E, V) analysis has the great virtue of simplicity. As we have seen in Section 4.5, it is applicable in the case of quadratic utility functions. However, when we are prepared to assume only that decision makers are averse to risk, the (E, V) rule can only be applied properly to random variables that belong to the same family of distributions; and the family is characterized by two parameters, each of which is an independent function of the mean and variance. The normal is clearly such a family. However, the log-normal family of distributions has two parameters that are not independent of its moments, so in this case the (E, V) criterion is an incorrect way to identify the SSE set (Philippatos and Gressis, 1975). In fact, the appropriate criterion is a closely related one, as proved by Levy (1973a). *A log-normal F dominates a log-normal G in the sense of* SSD *if* $E_F(x) \geq E_G(x)$ *and* $V_F(\log x) \leq V_G(\log x)$ *with at least one strong*

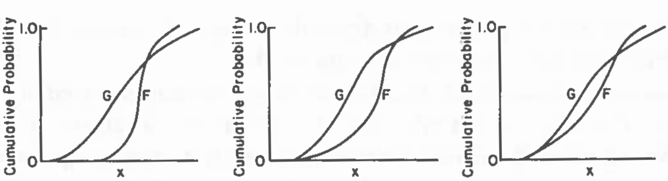


FIG. 9.4. Three ways in which G can be more prone to low outcomes than F when F and G are simply related.

inequality. The log-normal distribution is sometimes encountered in portfolio-type problems and also in oil and water drilling decision problems.

Simplification of the identification of efficient sets of prospects along the lines just discussed has been developed from some slightly different angles by Hammond (1974). He deals particularly with families of distributions whose members have CDFs that do not intersect more than once, i.e., are "simply related" (e.g., the normal). He also develops the concept of proneness to low outcomes and defines an SSD-dominated distribution as one that has a lesser mean and is more prone to low outcomes. A normal distribution is more prone to low outcomes if it has a greater variance. Hammond's results are especially useful in graphical analysis with known means since proneness to low outcomes can then be readily identified. Figure 9.4 depicts three ways in which G can be more prone to low outcomes than F. Note that in the third graph F and G are initially coincident.

Third-Degree Stochastic Efficiency

It is generally accepted that aversion to risk is the "norm" for behavior of the great majority of agricultural (and other) decision makers. Consequently, and because of the practical impossibility of eliciting every decision maker's utility function, we believe that the derivation of SSE sets is of considerable practical importance. We are on much thinner ice when we attempt to go further in narrowing down the efficient set because more restrictive general assumptions about preferences are required. One further generalization seems worth risking to give us a third and final concept of stochastic efficiency.

The concept of third-degree stochastic dominance (TSD) rests on an additional assumption about the underlying utility function, viz., that the third derivative is positive, $U_3(x) > 0$. This restriction is implied by the strongly intuitive requirement that as people become wealthier, they become decreasingly averse to risk. It is a necessary but not a sufficient condition for decreasing risk aversion (in the sense elaborated in Section 4.4). We also speculated in Section 6.2 that $U_3(x) > 0$, thereby predisposing

(Hanson and Menezes, 1971) to a preference for positive skewness in distributions of returns. The TSD ordering rule, due to Whitmore (1970) and Hammond (1974), is a logical extension of that for SSD and requires the definition of a further type of cumulative function, namely the area under the SSD cumulative function:

$$F_3(R) = \int_a^R F_2(x)dx \tag{9.5}$$

The distribution F dominates G in the sense of TSD *if* $F_3(R) \leqq G_3(R)$ *for all possible R with at least one strong inequality and if* $F_2(b) \leqq G_2(b)$, *where b is the upper range, or equivalently,* $E_F(x) \geqq E_G(x)$.

Removal of TSD-inefficient prospects from consideration leaves the third-degree stochastically efficient (TSE) set, which cannot be larger than the SSE set. However, our limited experience suggests that the SSE and TSE sets may generally be very similar. Perhaps this experience may be overly colored by analysis of distributions that are fairly symmetric, whereas we would expect the TSD rule to come into its own when distributions of diverse skewness are compared. Our empirical experience may offer one reason for underplaying the importance and usefulness of TSD. A second reason might be a lack of faith in the validity of the underlying behavioral assumptions implied by specifying $U_3(x) > 0$. We will encounter a third reason in Section 9.2 when the added computational task is confronted. The added cost may not cover the marginal benefit of identifying only a slightly smaller efficient set.

A discrete version of the TSD ordering rule has been inferred by Porter et al. (1973) and requires a new discrete cumulative function,

$$F_3(x_r) = (1/2)\sum_{i=2}^{r}[F_2(x_i) + F_2(x_{i-1})]\Delta x_i \qquad r = 2, \ldots, n$$

$$F_3(x_1) = 0 \tag{9.6}$$

Then F dominates G in the sense of TSD *if* $F_3(x_r) \leqq G_3(x_r)$ *for all* $r \leqq n$ *with at least one strong inequality, and* $F_2(x_{n-1}) \leqq G_2(x_{n-1})$. This rule can be conveniently illustrated by means of a simple numerical example based broadly on a maize insect-control problem discussed by Anderson (1974b). Suppose the decision is between two insecticidal programs F and G, and the payoffs and probabilities are assessed as reported in Table 9.3. The efficiency analysis is summarized in Table 9.4, which reveals that the actions are clearly distinguished by resort to the third-degree rule but are in fact separated at the SSD stage. This means that any utility-maximizing decision maker with (decreasing) aversion to risk, if faced with the problem as depicted, would opt for treatment F.

TABLE 9.3. An Insecticidal Decision Problem

Level of Infestation	Probability	Low-Cost Treatment G	Expensive Treatment F
		($/ha)	
Low	0.2	27	23
Medium	0.5	24	22
Severe	0.3	15	21
	EMV	21.9	21.9

TABLE 9.4. Efficiency Analysis of the Insecticide Problem

x_i	15	21	22	23	24	27
$f(x_i)$	0	0.3	0.5	0.2	0	0
$g(x_i)$	0.3	0	0	0	0.5	0.2
$F_1(x_i)$	0	0.3	0.8	1.0	1.0	1.0
$G_1(x_i)$	0.3	0.3	0.3	0.3	0.8	1.0
Δx_i	...	6	1	1	1	3
$F_2(x_i)$	0	0	0.3	1.1	2.1	5.1
$G_2(x_i)$	0	1.8	2.1	2.4	2.7	5.1
$F_3(x_i)$	0	0	0.15	0.85	2.45	13.25
$G_3(x_i)$	0	5.4	7.35	9.60	12.15	23.85

Except for the discussion below of efficiency in relation to specific families of utility functions, the third-degree case is as far as we propose to go in search of stochastic efficiency criteria. Vickson (1975) has developed a stronger ordering rule (DSD) based on the behavioral assumption of decreasing absolute risk aversion, but this is presently restricted to discrete random variables and is computationally most demanding. It seems that further attempts to narrow the efficient set of prospects would need to call on general behavioral assumptions that are not readily defensible. For instance, Anderson (1974b) has suggested a rule for fourth-degree stochastic dominance and Hammond (1974) has explored high order dominances based on CDFs that intersect a specified number of times.

Convex Stochastic Efficiency

Although there seems little merit in pursuing fourth and higher degree stochastic efficiency, this is not to suggest that work toward further refinement and exploitation of the notions of dominance and efficiency is stagnant. Indeed, the topic is of such importance that it is bound to attract the continued attention of decision theorists. For example, Fishburn (1974a) has generalized the dominance results and rules in a most interesting way. He is especially concerned with a particular class of linear (convex) combinations of probability cumulatives (note, not random variables) wherein the weights $\{\lambda_i\}$ are nonnegative and sum to unity, viz., $\lambda_i \geqq 0$ and $\sum_{i=1}^{n} \lambda_i = 1$. His convex generalization of the FSD and SSD ordering rules may then be crudely stated as follows. *Under the* FSD (SSD) *behavioral assumption(s), if some convex combination of several* CDFs $\sum_{i=1}^{n} \lambda_i[F_1(x)]_i$ *dominates in the sense of* FSD (SSD) *the same convex combination of several other* CDFs $\sum_{i=1}^{n} \lambda_i[G_1(x)]_i$, *then* $U[F_1(x)]_i > U[G_1(x)]_i$ *for at least one of the* i, $i = 1, \ldots, n$. The last phrase of this statement asserts that for all the relevant utility functions (e.g., all risk-averse functions in the case of SSD), the utility of some (at least one) of the F distributions must be greater than the utility of the corresponding G distributions; thus these one or more G distributions must be strictly not preferred, i.e., must be dominated.

The trick is to identify just which distributions are dominated. In general, this is not possible without knowing more about preference than we are able to assume. There is, however, a special case that offers scope for operational identification of such dominated distributions at the cost of reducing the power of the rule. The device used is to take one member of a set of n prospects and to suppose it to consist of $n - 1$ identical prospects labeled G. Convex combinations of the remaining $(n - 1)$ F prospects can be examined for dominance with respect to the combinations of the $(n - 1)$ G prospects (which are exactly the one G prospect itself). The procedure is illustrated in Figure 9.5 for the case of FSD among two rectangular distributions and one other distribution. Without the convex generalization, we would say that all three distributions $(F_1)_1$, $(F_1)_2$, and G_1 lie

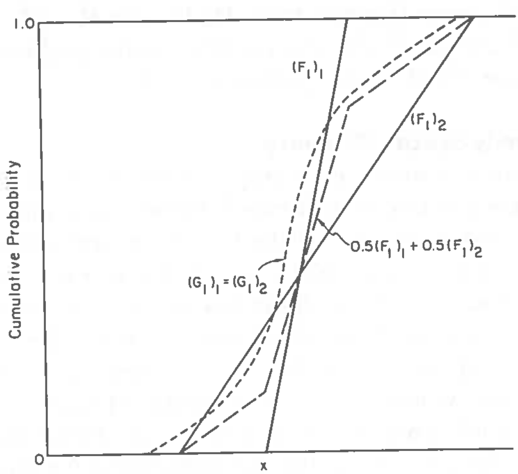


FIG. 9.5. An illustration of convex FSD.

in the FSE set. However, many convex combinations of $(F_1)_1$ and $(F_1)_2$, including $\lambda_1 = \lambda_2 = 0.5$, clearly dominate $(G_1)_1 = (G_1)_2$, so G_1 should be eliminated from the FSE set.

The above simplistic example gives something of the flavor of the additional power that the convex generalization gives to reducing the size of efficient sets. Our speculation is that in typical practical efficiency analyses, the generalization will have a rather small impact on diminishing the size of efficient sets—especially until such time as easily applied computational routines are available. This seems to be some time away because of the difficulties of searching for dominant convex combinations. Perhaps the best possibilities lie in applying the concept and theorems of Fishburn (1974b) to discrete distributions. Work with convex combinations of theoretical continuous distributions is complicated by the fact that the combination distribution does not belong to the same family as a common family of parent distributions such as the normal. For example, in Figure 9.5 $(F_1)_1$ and $(F_1)_2$ are rectangular distributions, but the depicted convex combination is clearly not rectangular.

Other extensions of concepts of stochastic efficiency are also sure to be developed. Fishburn (1974a, b) has employed his convex extensions to explore efficiency in decisions based on voting by several people. Another important extension concerns the development of stochastic efficiency concepts for multiattributed choice situations (Levy and Paroush, 1974a; Kihlstrom and Mirman, 1974), of which an important special case concerns multiperiod choice problems such as portfolio selections (Levy, 1973b; Levy and Paroush, 1974b). The portfolio problem continues to receive attention by theorists searching for more general results concerning stochastic efficiency (Hadar, 1971; Hadar and Russell, 1974a, b). The concepts of efficiency have also been applied to the problem of measuring equality of income distributions (Atkinson, 1970).

Utility Family-Specific Efficiency

To complete our survey of concepts of efficiency, a final and more restrictive approach is briefly mentioned. Hitherto our only preference assumptions were about the signs of the first three derivatives of the utility function. It will come as no surprise that efficient sets may be greatly compacted if rather more specific assumptions are introduced about the utility function, e.g., that the function belongs to a particular family of functions such as the exponential family. The simplest family to study is the quadratic, and here we follow the presentation of Hanoch and Levy (1970) who also examine efficiency for the cubic family of utility functions.

Recall from Section 4.4 that the quadratic utility function

$$U(x) = x - bx^2 \qquad b > 0$$

is only increasing for $x < 1/2b = K$, so we might conveniently linearly transform the function to

$$U(x) = 2Kx - x^2 \qquad K > 0, x < K$$

which in turn leads to the expected utility function

$$U = 2KE(x) - [E(x)^2 + V(x)]$$

Then given two prospects x_1 and x_2 with means μ_1, μ_2 and variances σ_1^2, σ_2^2 respectively, x_1 would be preferred if and only if

$$\begin{aligned} U(x_1) - U(x_2) &= 2K\Delta\mu - (\Delta\mu^2 + \Delta\sigma^2) \\ &= 2\Delta\mu(K - \bar{\mu}) - \Delta\sigma^2 > 0 \end{aligned} \qquad (9.7)$$

where $\Delta\mu = \mu_1 - \mu_2$, $\bar{\mu} = (\mu_1 + \mu_2)/2$, $\Delta\mu^2 = \mu_1^2 - \mu_2^2$, and $\Delta\sigma^2 = \sigma_1^2 - \sigma_2^2$. Since we are presently assuming that we know only that preference is quadratic, we do not know the value of K. However, we know that it would not be sensible to speak of quadratic utility unless K exceeded the upper range x_m of all the considered distributions. Thus if x_m is known, it can be substituted for K to give one empirical efficiency criterion:

$$2\Delta\mu(x_m - \bar{\mu}) - \Delta\sigma^2 > 0 \qquad (9.8)$$

Use of this criterion—as would use of that of expression (9.7) if K were known—would lead to a probably rather smaller efficient set than the simple (E, V)-efficient set since x_m is less than the infinitely large value implied by distributions such as the normal. In less desirable, but perhaps most frequently encountered circumstances, x_m may not be known; and we may be able to assume only that $K \geq \mu_m$, where μ_m is the largest mean of the considered distributions. The criterion then becomes

$$2\Delta\mu(\mu_m - \bar{\mu}) - \Delta\sigma^2 > 0 \qquad (9.9)$$

Practical applications of these criteria may sometimes be through pairwise review of distributions. Then the x_m and μ_m values used in each comparison can be respective maxima from the pair under consideration. With x_m unknown and $\mu_1 > \mu_2$, substitution of μ_1 for μ_m in expression (9.9) yields the simple pairwise criterion:

$$(\Delta\mu)^2 - \Delta\sigma^2 > 0 \qquad (9.10)$$

For example, consider the case where $\mu_1 = 20$, $\sigma_1^2 = 100$, $\mu_2 = 10$, and $\sigma_2^2 = 25$, which reveals no dominance in the (E, V) sense. With quadratic utility and application of criterion (9.10), $(\Delta\mu)^2 - \Delta\sigma^2 = 100 - 75 = 25 > 0$, so that in this sense the second prospect is dominated. In general, use of (9.9) will yield smaller efficient sets than will pairwise application of (9.10).

Hanoch and Levy (1970) also present further refinements of the quadratic case based on additional considerations such as symmetry of distributions, constraints on skewness, etc., but we will not elaborate these since they seem of limited generality. The other major attempt to explore efficiency by appeal to a specific utility family is Hammond's (1974) work on the negative exponential family of constant risk aversion utility functions. Briefly, several of the ordering rules he suggests depend on a limit being placed on the extent of aversion to risk, appeal to the constant risk aversion function as an analytic utility function, and finally, orderings based on his earlier discussed notions of simple relatedness and proneness to low outcomes. This is exemplified by his Corollary 3.1, which may be useful for narrowing down efficient sets under the circumstances specified. *Suppose that F and G are simply related and that F is preferred to G under a constant risk aversion of level c; then if F is more prone to low outcomes than G and risk aversion (as measured by Pratt's coefficient of absolute risk aversion) is less than or equal to c, F is always preferred to (dominates) G.*

9.2 ASSESSMENT OF STOCHASTIC EFFICIENCY

Among the concepts of efficiency surveyed above, several are very easy to implement and present no real problems in practical applications of any scope. These are the rules expressed in terms of a few simple parameters of the reviewed distributions, such as the (E, V) rules for normal and log-normal cases and the quadratic utility criteria. However, various computational difficulties are inherent in most of the other rules. These difficulties and some suggested solutions are briefly discussed here. We will segregate our discussion of continuous and discrete distributions since they present somewhat distinctive problems.

Our discussion of the ordering criteria has so far emphasized that, in general, assessment of stochastic efficiency must proceed by pairwise comparisons among prospects not so far eliminated from the efficient set. If all the prospects considered are "pure prospects," this process says little about the efficiency of possible mixtures of pure prospects. The only way of establishing the efficiency or otherwise of such mixtures is to specify the distributions pertaining to the mixtures and to test these as for other prospects. Considerable practical difficulty is usually encountered in distributional specification for mixed prospects, and the Monte Carlo programming technique described in Section 9.3 is one approach to overcoming this difficulty. The only simple case is in dealing with mixtures or linear combinations of random variables (not of probability cumulatives as is involved in assessing convex stochastic efficiency) that are normally distributed. In this case, equations (7.1) and (7.2) serve to define the parameters (E, V) of the distribution for any specified mixture that may, depending on the degree of correlation present between the component variables, dominate some of the pure prospects (e.g., some crop monocultures may be dominated in the sense of SSD by some diversifications).

For a small numerical example, consider the following three uncorrelated normal distributions described by their means and variances: $F \sim N(20, 100)$, $G \sim N(10, 25)$, and $H \sim N(15, 50)$. All three distributions are seemingly in the SSE set because there is no (E, V) dominance. However, a (0.6, 0.4) mixture of F and G implies the distribution $N(16, 40)$, which dominates H in the (E, V) sense so H should be eliminated from the SSE set if such a mixture is a possibility that should be considered.

Discrete Distributions

With the exceptions mentioned above, discrete distributions are the simplest for assessing stochastic efficiency. Examples of the mode of analysis applicable have already been given in Figure 9.2 and Tables 9.2, 9.3, and 9.4. The method consists of assembling all the combined discrete values of the random variable x for pairs of probability mass functions $f(x_i)$ and $g(x_i)$ and arranging them in ascending order such that if $i < j$ then $x_i < x_j$. If two or more x have the same numerical value, each is considered to be distinct and the rank allocated to ties is lowest for those associated with the potentially dominated distribution. This is the distribution that is more prone to low outcomes; e.g., it has a nonzero probability for the lowest value of x when this is not tied.

Assessment of efficiency for discrete cases of great simplicity (e.g., few states and few distributions) is best done by the illustrated tabular method (Table 9.4) rather than graphically. However, for less simple cases, recourse to a computer provides the most convenient method of assessment. The program implemented can follow the tabular steps in a straightforward manner. There is considerable scope for imaginative programming to make the intrinsically pairwise comparisons very "efficient" from the viewpoint of computational cost. Excellent programs are available (Porter et al., 1973), although these deal with very special cases of discrete distributions. These might be termed discrete sample distributions and consist of a set of n sample observations (usually on an intrinsically continuously distributed random variable), and a discrete probability of $1/n$ is arbitrarily awarded to each observed value. Efficiency analysis is especially easy when "discrete distributions" of equal sample size are compared. These methods of using samples to "discretize" continuous distributions have in fact been widely used to compare general stochastic efficiency criteria with long-used criteria like (E, V) (Levy and Sarnat, 1971; Porter, 1973; Porter and Gaumnitz, 1972).

Letting sample data "speak for themselves" in this way runs counter to the standpoint we adopted in our discussion of probability assessment in Section 2.3. Departure from such an essentially subjective standpoint (i.e., emphasizing elicitation from the individual decision maker) can be defended on the grounds that here we are concerned with less well-structured situations where the decision maker is presumably difficult to pin down (at least insofar as preferences are concerned), so it may also be difficult to elicit his personal probabilities. For consistency we logically should pursue the inferential steps elaborated in Section 2.4 to transform sample data into estimates of distributions, whether these are discrete or continuous. However, let us now turn to the more challenging problems of looking at efficiency in continuous distributions that are treated as continuous.

Continuous Distributions

Assessment of efficiency with only a few continuous distributions is probably best done as far as possible by graphical methods. This is especially convenient when distributions are available in graphical form (as subjective distributions often will be) and is very easy for FSD checking and for SSD checking where distributions are simply related. Graphical methods are not so convenient when the SSD and TSD cumulative functions must be defined. The required integrations can be carried out in such an approach by either counting squares on ruled paper to estimate areas under curves or by cutting out the cumulative areas and weighing them to estimate areas. Sophisticated equipment like planometric instruments or graphical-digital equipment could be used most advantageously but may not be generally available to decision analysts. Usually, however, a computer will be needed for expedient analyses of stochastic efficiency.

When the mathematical form of compared distribution functions is known, several analytical possibilities are available. If they are from the same family, there is a chance that the distributions are simply related so that determination of relative proneness to low outcomes may fairly simply indicate the composition of the SSE set. If the FSE set should be required (and many analysts may feel that it is insufficiently interesting to be worthy of identification), the known CDFs can be compared along their range. This is undemanding if they are known to be simply related because the comparison interval can be rather coarse. Otherwise, however, judgment will be required as to how fine the intervals for comparison need to be.

Even with high-speed computers the infinite number of comparisons implied in the ordering rules for continuous distributions cannot (and indeed need not in practice) be approached. Subsequent efficiency tests require new cumulative functions that in principle can be traced out by successive integrations of the known distribution functions. This might be done analytically for relatively simple functions but more typically will require numerical methods of integration. Whatever the method, it is clear that an element of approximation and judgment is required in performing a necessarily finite number of comparisons of the derived functions. Recourse to numerical integrations will tend to make efficiency analysis of large sets of prospects an expensive analytical procedure (at least in terms of use of computer time), and this is one reason we do not take up such methods here. The other important reason for not emphasizing such methods is that all too frequently we will not know the mathematical form of the distribution functions. Instead, we will typically have a set of smooth curves representing CDFs of indeterminate mathematical form. It would be possible to engage in a mathematical function-fitting endeavor, but this is only a "best fit" (i.e., imperfect) procedure. If it leads to close-fitting functions that are very convenient for efficiency analysis (e.g., normal, log-normal, rectangular, etc.), it may be justified for its saving in analytical costs. However, if inconvenient imperfectly fitting functions are to result, the analytical route seems suboptimal.

The approach we recommend is to take the fact of limited accuracy approximation to heart and settle for a mathematical approximation that is most convenient for subsequent efficiency analysis. We suggest ap-

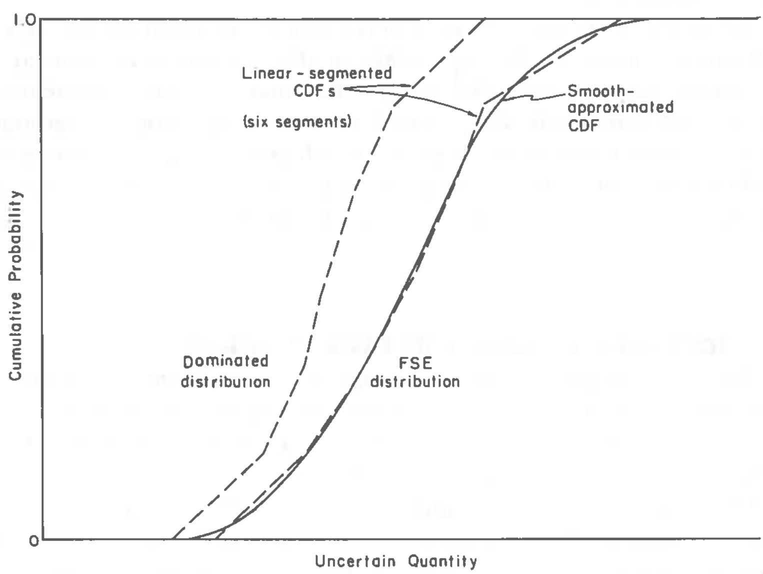


FIG. 9.6. Illustration of linear-segmented CDFs.

proximating each CDF by a predetermined number of linear segments, each spanning an equal interval of cumulative probability (see Figure 9.6). This is equivalent to representing the underlying PDF by a rectangular histogram with rectangles of equal area. The great analytical advantage is the simplicity of integrating a piecewise linear function to define the SSD cumulative function, which then consists of quadratic segments that are also easily integrated to define the subsequent TSD cumulative function. Another advantage of this approach is the relative simplicity of the pairwise comparisons among functions of each type. The mathematical background to this approach is described fully by Anderson (1974b), and a general purpose program for such review is given in Section 9.5. The process works fairly efficiently in applications involving up to about 50 risky prospects, but it may prove expensive on core size and processor time for very large numbers of reviewed distributions. The transitivity of all the efficiency criteria permits an analysis to be conducted by reviewing only a few risky prospects at a time with the currently efficient prospects, permitting progressive revision and enlargement of the efficient set. It must be said that the precision of approximations in this method is controlled largely by the number of linear segments chosen to describe each CDF and a subjective balance must be reached between added precision and added computational burden. Judgment about desired precision may well be influenced largely by the shape of CDFs, especially in the lower tails of the distributions.

Doubtless there will be many developments in methods for assessing stochastic efficiency as the importance of the concept is recognized. Our brief survey has been intended to indicate some general approaches that seem to work fairly well. We can look forward to operational methods for exploiting convex stochastic dominance. Meantime, our recommendation for efficiency analysts is to search for and exploit any method that offers convenience in operation and consistency in approach.

9.3 STOCHASTIC EFFICIENCY IN FARM PLANNING

Several examples of efficiency analysis of problems in agricultural management were introduced in the methodological survey of Section 9.1. Our intention here is to discuss some applications that demand computational assistance of the types reviewed in Section 9.2.

One important field of application of the concepts of stochastic efficiency is analysis of the impact of risk in agriculture when it is simply impossible to elicit utility functions for the farmers involved. This is typically the case in agricultural research and development. For example, our first topic relates to the search for new wheat varieties that will be "efficient" for large numbers of unidentified (mainly peasant) farmers.

Selection of Risk-efficient Crop Varieties

In this example, we illustrate the use of discrete sample data to approximate essentially continuous distributions.

Several methods have been used for identifying crop varieties that have wide environmental adaptability. The basic data for such work are usually obtained from nursery trials conducted in diverse environments, sometimes across many countries, as in the collaborative nursery administered by the International Maize and Wheat Improvement Center (CIMMYT, 1972). The analytical methods used have ranged from comparisons of mean yields to comparisons of statistics based on regressions of varietal yields on environmental indices. In the absence of specifically and carefully elaborated criteria, there can be no one perfect method of appraisal. In the present example we examine the question of adaptability from the point of view of stochastic efficiency—believed to be relevant if the ultimate purpose of identifying widely adapted varieties is to make them available for adoption by farmers who generally are averse to risk. As Finlay and Wilkinson (1963) have observed: "Plant breeders are inclined to ignore the results obtained in low-yielding environments (e.g., drought years), on the basis that the yields are too low and are therefore not very useful for sorting out the differences between selections. This is a serious error, because high-yielding selections under favourable conditions may show relatively greater failure under adverse conditions."

The notions of stochastic efficiency provide a useful framework for posing the essentially empirical question of how different selections perform in diverse risky environments. The analysis is straightforward if it makes good sense to speak of a world probability distribution of wheat yields and if the selection of sites, cooperators, fields, and growing and disease conditions is somehow representative of the relevant domain of production. Unfortunately, for lack of more appropriate measures we are forced to use yield as a surrogate for the argument of the implicit utility function. This assumption, which involves ignoring differences in production costs, is unavoidable in processing international nursery data since each trial is in general grown under differing regimes of irrigation (where practiced), tillage, fertilizers, and weed and pest control that are most difficult to cost. Attention is now concentrated on the data from a particular nursery (CIMMYT, 1972) in which 49 varieties were compared in trials at 60 locations in 37 countries during 1969–70. For each variety, each trial observation is regarded as a distinct component of the discrete sample probability function of that variety. The pairwise comparison of 49 discrete actions involves up to (49)(48)/2 = 1176 FSD comparisons at each of up to (60)(2) = 120 values of the uncertain yield. Such a computational burden can be faced with equanimity only with the aid of a computer.

Complete identification of the efficient varieties is of little interest here—details are given in Anderson (1974b)—so we identify the varieties only by their rank in terms of mean sample yield. The results are summarized in Table 9.5. The inclusion of the high-ranked varieties in the efficient sets is to be expected, but perhaps the more interesting result is the inclusion of some relatively low-ranked varieties in the "risk-efficient" SSE and TSE sets. The efficiency analysis obviously provides information that should be useful to plant breeders. Perhaps such information would be most useful at relatively late stages in breeding programs when materials of somewhat similar adaptation must be screened.

TABLE 9.5. Efficiency Analysis of World Wheat Yield Data

Set of Varieties	Number in Set	Rank Identifications*
Total considered	49	1–49
FSE	27	1–19, 21–24, 26, 27, 34, 35
SSE	6	1, 2, 5, 10, 27, 34
TSE	5	1, 2, 5, 10, 27

*Varieties are identified by their rank of mean sample yield.

Selection of Risk-efficient Fertilizer Rates

This fertilizer example builds on material described more fully by Anderson (1973, 1974a). The starting point is to use the 36 probability distributions of unirrigated wheat yield estimated for each of the design points of a 6 × 6 complete factorial. The treatments are for N at approximately 0, 22, 45, 67, 90, 112 kg/ha and for P at approximately 0, 9, 18, 27, 36, 45 kg/ha.

The estimation of the distributions, which reflect between-year variability, was based on sparse data (see Section 2.4). Examples of these distributions have been given in Figure 6.5, and more are shown in Figure 9.7 in linear-segmented CDF form in which all are subsequently described. Each is described by 20 linear segments spanning equal probability intervals. Yield distributions were transformed to net revenue distributions by the linear expression $R_{ijk} = p_y Y_{ijk} - p_n N_i - p_p P_j$, where R denotes revenue, Y denotes yield, N denotes nitrogen applied, and P denotes elemental phosphorus applied, all per unit area; p_y, p_n, and p_p are the respective unit prices of Y, N, and P; and the subscripts ijk denote respectively the ith level of N, the jth level of P and the kth fractile. Note that

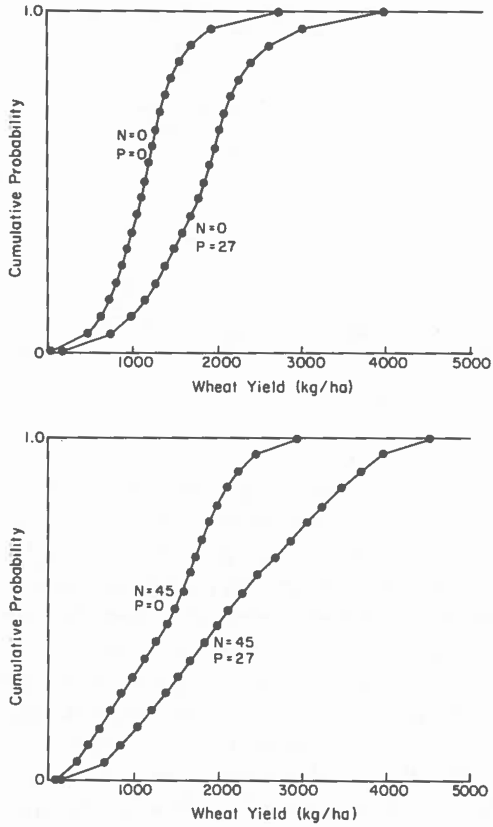


FIG. 9.7. Some examples of linear-segmented CDFs for wheat yields under different fertilizer treatments.

fixed costs and area grown are not included in this expression since they have no influence on the determination of stochastic efficiency.

The procedure for reviewing stochastic efficiency of continuous distributions outlined in Section 9.2 was applied to these 36 discrete actions through the computer program of Section 9.5. Results are most easily displayed in Table 9.6 with rows and columns defined by rates of N and P and the table entries by the degree of stochastic efficiency: zero denotes dominated in the sense of FSD; FSE combinations of fertilizers are indicated by an integer $\geqq 1$; FSE combinations become candidates for SSE review and those that survive are indicated by an integer $\geqq 2$ and in turn become candidates for TSD review; those not then dominated are TSE and are indicated by the integer 3. In the present case the SSE and TSE sets are identical so that no "2" entries appear in Table 9.6.

TABLE 9.6. Stochastic Efficiencies of Specified Fertilizer Combinations

N	P: 0	9	18	27	36	45
	(kg/ha)					
0	0*	1	3†	3†	3†	3†
22	0	0	0	3†	3†	1†
45	0	0	1	3†‡	1	1
67	0	0	0	1	1	1
90	0	0	0	1	1	1
112	0	0	0	0	1	1

*A zero entry denotes inefficiency according to FSD. The nonzero entries denote the highest degree of efficiency attained, so that numbers 1, 2, and 3 indicate FSE, SSE, and TSE respectively.

†Members of the (E, V)-efficient set.

‡The discrete combination with greatest mean return.

These results indicate a fairly consistent pattern wherein (in this case related to crop response on a red-brown earth) a necessary condition for stochastic efficiency of any order is a reasonable dose of phosphorus. Nitrogen is indicated as being a rather risky proposal since most of the risk-efficient combinations involve zero levels of N and the highest risk-efficient rate of N is 45 kg/ha in this nonirrigated situation. Further, given the consistent pattern of risk-efficient rates, it seems reasonable to interpolate within the set. Making such an interpolation suggests that, after resorting to specific assumptions about risk-averse preference functions as detailed by Anderson (1973), all the risk-optimal rates fall in the interpolated efficient set. This should not be surprising, although the specific risk-optimal rates were computed by an approximate procedure using only the first two moments of the yield distribution. In this particular example, as in others from the field of portfolio analysis (Porter, 1973; Porter and Gaumnitz, 1972), it does turn out that the risk-efficient set corresponds very closely with the (E, V)-efficient set, as indicated by the footnote to Table 9.6.

In decision problems involving continuous decision variables, it would be convenient to conduct efficiency analyses in a manner broadly analogous to those of continuous response analysis, so as to locate precise bounds on efficient levels of the decision variable(s). Unfortunately, this is impossible because, as the foregoing sections have illustrated, the analysis of stochastic efficiency is intrinsically a discrete affair involving pairwise comparisons of the cumulative probability or other derived functions. This means that application of the efficiency principles to a problem that is inherently continuous must involve making the problem discrete in such a way that the essence of the original problem is not lost. The alternative to attempting to interpolate among discrete results is to follow a more familiar analytical route and interpolate among the data. For example, if normal distributions have been fitted to response data at several levels of a continuous variable, an analysis might reasonably postulate smooth functional relationships between the continuous variable and the distributional parameters. Subsequent analysis for stochastic efficiency could proceed on the basis of (discrete) predicted or interpolated distributions at levels of the continuous variable other than those at which the observations were available. The interpolation procedure could presumably be carried to any desired intensity to effectively allow a continuous analysis. Such is broadly the approach suggested here.

There are several possibilities for interpolating among data. For instance, in the present example we could attempt to interpolate the coordinates of linear-segmented approximating distributions for rates of fertilizer on a finer grid than the 36 combinations analyzed above. Such an approach has been illustrated for a single fertilizer decision variable by Anderson (1974b). However, for present purposes we prefer to discuss a less direct but we hope more general approach to the same problem.

Interpolation of distributions is here accomplished in two stages: relating sufficient parameters of the 36 estimated distributions to the decision variables N and P and fitting (interpolated) distributions by use of the parameters predicted for any specified combinations of fertilizer nutrients. The selected equations are reported below for the mean response $E(y)$, the variance of response $V(y)$, the lower bound of yield $A(y)$, and the upper bound $B(y)$; N, P, and y are in kg/ha and numbers in parentheses are respective standard errors of the regression coefficients.

$$\begin{aligned}
E(y) &= 1170 + 9.16N + 42.4P - 0.0765N^2 - 0.695P^2 + 0.146NP \\
\bar{R}^2 &= 0.99 \;(19) \quad (0.51) \quad (1.38) \quad (0.0040) \quad (0.025) \quad (0.009) \\
V(y) &= 164{,}200 + 10{,}700N + 26{,}500P - 88.6N^2 - 716P^2 + 1320NP \\
\bar{R}^2 &= 0.94 \;(67{,}070) \quad (1840) \quad (4590) \quad (14.5) \quad (90.8) \quad (31) \\
A(y) &= 106 - 4.33N - 0.76P + 0.040N^2 + 0.357P^2 \\
\bar{R}^2 &= 0.80 \;(62) \quad (1.94) \quad (4.8) \quad (0.016) \quad (0.104) \\
B(y) &= 2840 + 10.0N + 63.5P - 0.137N^2 - 1.50P^2 + 0.831NP \\
\bar{R}^2 &= 0.94 \;(162) \quad (4.45) \quad (11.1) \quad (0.035) \quad (0.220) \quad (0.075)
\end{aligned}$$

For a given combination of N and P within the experimental range, these equations predict with a fairly high degree of accuracy the mean and variance of response and the upper bound of response. The lower bound equation is unfortunately not so precise, reflecting the less consistent pattern of the zero fractile with respect to N and P. These features of yield distributions are readily transformed into the corresponding mean m, variance v, and bounds (a, b) of the relevant net revenue distribution. To fit a beta distribution by the moment method, which presently seems most convenient, it is first necessary to compute the mean and variance, denoted by m^* and v^* respectively, of the corresponding standard beta distribution (range 0 to 1). Hence, $m^* = (m - a)/(b - a)$ and $v^* = v/(b - a)^2$ from which the shape parameters c and d (Mihram, 1972) can be found directly as $(c + d) = [m^*(1.0 - m^*)/v^*] - 1.0$ and $c = (c + d)m^*$.

If the beta distribution could be readily integrated to yield the CDF and successive cumulative functions, the analysis of stochastic efficiency could proceed directly. However, such integrations are very tedious and as argued earlier, the linear-segmented CDF approximation to the beta is suggested as an adequate and expedient approach to completing the analysis. Tabulations of the percentiles of the beta distribution are available for integer values of c and $c + d$ and for fractiles 0.05, 0.10, . . . , 0.95 (Pratt et al., 1965), reproduced in part in Appendix Table A.2. These intervals conveniently fit the use of 20 linear CDF segments of equal probability span. For noninteger values of c and d the fractiles, denoted here by f_s, must be interpolated (linearly seems adequate) from the fractiles of the four standard beta distributions with integer-shaped parameters that embrace the noninteger values of c and d. Fractiles for the respective fitted distribution are then found through the transformation $(b - a)f_s + a$.

The above procedure is demonstrated by reviewing stochastic efficiency first among 40 factorial combinations of N and P involving $N = 0$, 20, . . . , 80 kg/ha and $P = 0, 5, \ldots, 35$ kg/ha. The results are summarized in Table 9.7, using the same notation as in Table 9.6, and reveal that given all the assumptions made, the combinations of fertilizers that are risk efficient are again in the region near where $N = 0$ and $P = 30$ kg/ha. If more precise information is required, it can be obtained by interpolation on a finer grid in the determined region of interest. Happily, the efficiency results of Tables 9.6 and 9.7 are remarkably compatible, given the very different analytical approaches employed.

TABLE 9.7. Stochastically Efficient Fertilizer Rates on a Relatively Coarse Grid of Interpolation

	P				
N	0(5)15	20	25	30	35
			(kg/ha)		
0	0*	1	3	3	3
20	0	0	3	3	3
40	0	0	1	1	1
60	0	0	0	1	1
80	0	0	0	1	1

*A zero entry denotes inefficiency according to FSD. The nonzero entries denote the highest degree of efficiency attained, so that numbers 1, 2, and 3 indicate FSE, SSE, and TSE respectively.

Selection of Risk-efficient Farm Plans

In Chapter 7 we reviewed several alternative models for selecting efficient whole-farm plans. The efficiency criterion that received most attention was the (E, V) rule, and various methods for tracing out the (E, V)-efficient frontier were reviewed for the case where either returns are normally distributed or utility is quadratic. We will now briefly explore how our more general concepts of stochastic efficiency might be used in a farm planning context. Here we take up the suggestion made in Section 7.5 that the dominance rules can be used as the sorting criterion in Monte Carlo programming. We would only want to use this more general criterion when returns are other than normally or log-normally distributed. For brevity, we will refer to the approach to be described as REMP (risk-efficient Monte Carlo programming) (Anderson, 1976).

Monte Carlo programming (MCP) is chosen as the device for generating feasible plans because there is no guarantee that well-defined algorithms such as mathematical programming routines will present feasible plans that are in fact in the stochastically efficient set that is sought. Also, MCP provides a degree of flexibility with which to accommodate the intrusion of a nonnormal risk specification in the model. This risk specification is the important link missing between MCP—e.g., as developed and programmed by Donaldson and Webster (1968)—and stochastic efficiency analysis—e.g., as programmed in Section 9.5. Our general requirements for the probabilistic specification are that (1) nonnormal (skewed, finite range) marginal distributions be permitted for describing returns from individual activities and (2) account be taken of statistical dependence that inevitably exists between farm enterprise returns. We have seen in Section 7.3 how activity means, variances, and covariances are combined to give the mean and variance of total gross margin and total net return. These same methods are used here to capture the average and variability of whole-farm performance. One neat feature is the easy way interdependence is simply accounted for by the correlation coefficients embodied in the covariances.

The other general requirement of specifying arbitrary nonnormality in marginal distributions of returns is not so easily handled, especially since we must bear in mind the necessity of keeping track of such nonnormality in linear combinations of the enterprise distributions. The latter consideration makes for difficulty because there is a very limited number of permissible linear operations on parameters of distributions. Our pragmatic suggestion for meeting this requirement is to specify the upper and lower bounds of the revenue distribution for each enterprise. This satisfies both the finite range and asymmetry requirements neatly, and simple linear combinations of these parameters also define the respective range parameters of the distribution of total gross margin.

Our suggestions so far have said nothing specific about the mathematical form of the marginal enterprise distributions and have yielded just four parameters (mean, variance, and upper and lower range) of the distribution of total gross margin. The general efficiency rules require effectively complete specification of the distribution functions for reviewed plans, so we need to infer such distributions from sets of the four computed parameters. Again, as in the above example of efficiency analysis in fertilizer decision making, we choose the very flexibly shaped beta distribution for this inferential step. The beta family contains distributions of very diverse characteristics (Pratt et al., 1965, Ch. 9) and seems nearly ideal for present purposes. A beta distribution is uniquely determined by the mentioned four parameters and is easily fitted to them. As described in the fertilizer example, it also lends itself readily to linear-segmented approximations of the CDF (by use of the fractiles reported in Appendix Table A.2) and thence to convenient computerized efficiency analysis.

We will illustrate the suggested REMP method by reference to a study by Hazell (1971) that involved a vegetable farm planning problem solved by quadratic risk programming and mean–absolute deviation linear programming. We do not use our own farm planning example of Chapter 7 because it is strongly predicated on an assumption of multivariate normality. While Hazell's enterprise data successfully pass goodness-of-fit tests for normality (e.g., Shapiro and Wilk, 1965) at conventional levels of statistical significance, at least two of his four distributions certainly do not look very normal when processed as sparse data and plotted on normal probability paper. The sparse data procedure of Section 2.4 was applied to all four sets of Hazell's returns data specifically to "estimate" the values of the range extremes for each distribution. One such curve is depicted by the unbroken line in Figure 9.8. The broken line illustrates the beta fitting procedure by depicting a fitted beta distribution with the same mean, variance, and range as the smooth curve. Means, variances, and covariances are supplied by Hazell (1971).

The estimated range data were added to the data of Hazell's quadratic programming formulation of his problem, and the new problem solved as an MCP problem with stochastic efficiency sorting. Generated plans not in the emerging TSE set were discarded and the computer instructed to proceed in this manner until 20 efficient plans were available. This required only a total of 48 feasible plans (i.e., 28 were discarded). The composition of these plans is of little interest to us here. However, it is instructive to compare the plans with Hazell's (E, V)-efficient frontier computed by quadratic programming. The stochastically efficient plans are represented by dots in Figure 9.9.

In terms of (E, V) efficiency, the stochastically efficient plans clearly

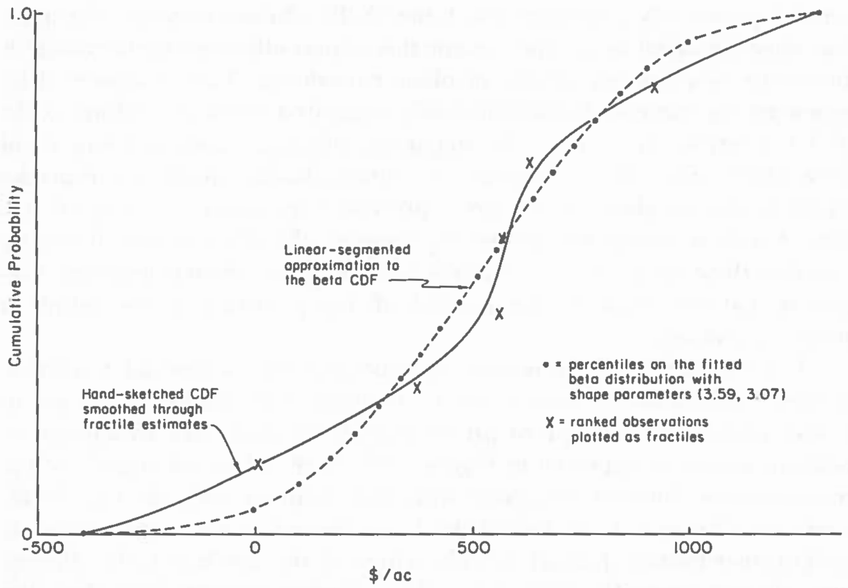


FIG. 9.8. Estimation and description of Hazell's distribution for gross margin from peppers.

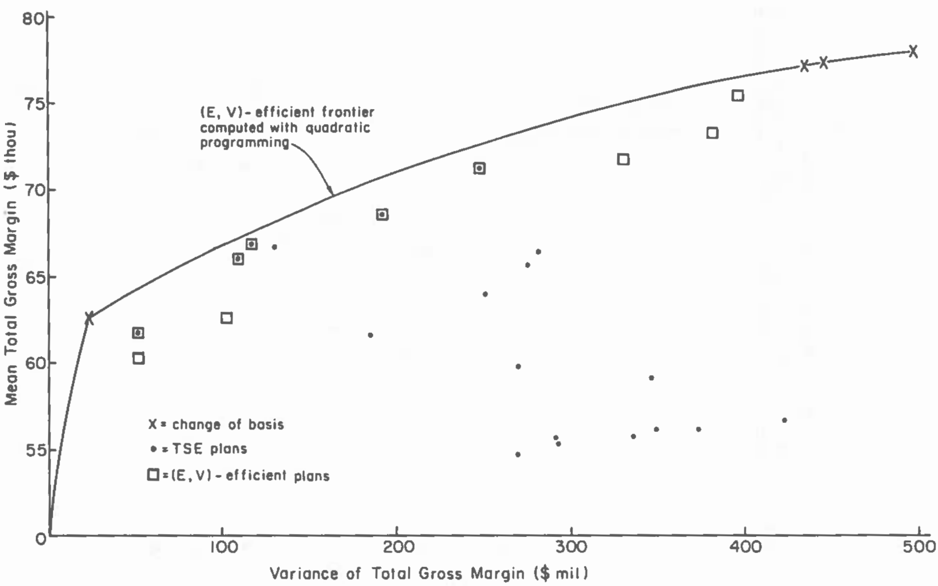


FIG. 9.9. Mean-variance characteristics of two sets of "efficient" farm plans.

perform poorly when compared with the (E, V)-efficient frontier. The question must be asked as to what extent this apparently poor performance is due to the very limited sample of plans considered. This is answered by reviewing the same 48 pseudorandomly generated plans according to the (E, V) criterion rather than the stochastic efficiency criteria. Only 10 of these plans were (E, V) efficient, i.e., undominated in (E, V) terms by others of the 48 plans. These are represented by squares in Figure 9.9. Note that there is considerable overlap between the efficient sets. It can be seen that these 10 plans all lie relatively close to the efficient frontier, suggesting that the result for the general efficiency sorting is not simply a sampling problem.

It is completely inappropriate to condemn the nonnormal stochastic efficiency sorting on the basis of (E, V) analysis. The results can be put in a much more favorable light by presenting similar data with an alternative measure of risk, as reported in Figure 9.10. Here risk is measured by the maximum possible loss associated with each plan in both the (E, V)-efficient and TSE sets. Each loss is the lower bound of the respective gross margin distribution. Judged by this criterion, the stochastically efficient plans appear generally much less risky (or more conservative) than the (E, V)-efficient plans.

The above example serves more to illustrate a possible use of stochas-

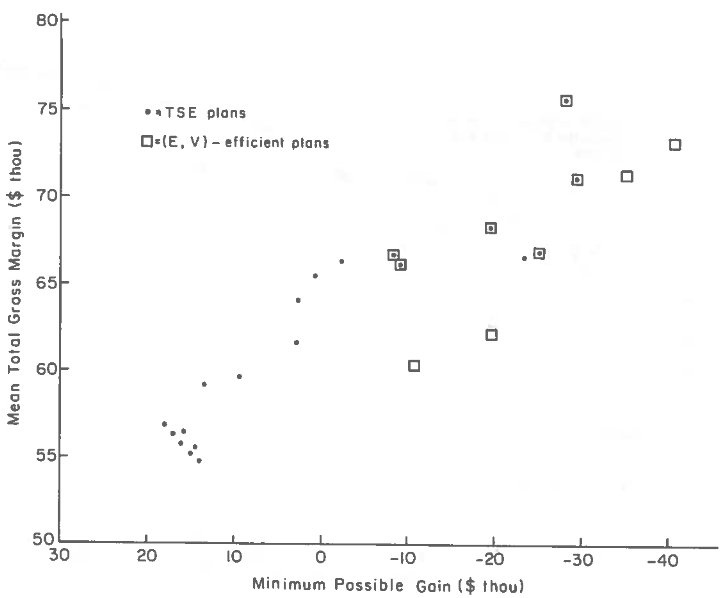


FIG. 9.10. Mean-risk characteristics of two sets of "efficient" farm plans.

tic efficiency criteria than to reveal the potential for risk planning methods of the REMP type. In fact, such methods can be useful for planning in a policy context when governmental intervention in risk bearing is under consideration. For example, Anderson (1975) has shown how stochastic efficiency concepts and the REMP method can be deployed to examine the impact of stabilization policy measures such as income tax arrangements for fluctuating incomes. As with all efficiency analyses, however, there is always a chance that the identified efficient sets will be of such large size and diversity that any sort of precise policy or decision interpretation becomes impossible.

9.4 IMPLICATIONS OF STOCHASTIC EFFICIENCY ANALYSIS

Decision analysis must all too often be attempted with incomplete information. In this chapter we have reviewed what is possible when preferences are unknown except for the most limited general information. As we have shown, there is much that can be said, and this fact has implications for decision theory itself, for research and extension, and for national agricultural policy.

Decision Analysis Methodology

We have emphasized that searching for efficiency without knowing utility functions is very demanding in terms of probability specification. Stochastic efficiency analysis can only be as good as the underlying probability specification, and when this is based on data analysis, there is always an implicit inferential leap. Two specification problems seem to be especially important. When simplification is sought through fitting convenient theoretical distributions, two inferential assumptions are required: that the chosen family is applicable and that the fitted parameters are accurate.

If complete and arbitrary CDFs form the basic data, the ordering rules place great emphasis on the position of the lower tails of the distributions. The very nature of such tails precludes any great confidence being attached to their specification. Questions of estimational risk (Kalymon, 1971; Frankfurter et al., 1971) and its impact have not been examined in this context. Rather, the implicit philosophy has been that the best analysis possible must simply use the best specification of probabilities that is possible. A systematic sensitivity analysis may engender added confidence if efficient sets are fairly robust to data variations. Then again, efficient sets are likely to be sensitive to, say, variations in positioning the lower tails of reviewed distributions. In this case we are either forced back to the philosophical position that our best is the best we can do, or we are forced to declare that efficiency analysis is infeasible and futile.

While our review has been overtly subjectivist, we have not adopted a Bayesian stance in discussing efficiency. Stochastic efficiency analysis should properly be based on the best-judged estimates of distributions. In many cases the best estimate will exist as a posterior distribution emerging from perhaps a sequence of probability revisions. Examples of Bayesian updating of distributions in a related agricultural context have been provided by O'Mara (1971).

Finally, stochastic efficiency analysis has not reached the terminus of its theoretical and empirical development. We can confidently look forward to further useful developments that will contribute to the simplification of decision analysis in the face of inadequate information.

Research and Extension

Increasingly, lip service has been paid to the notion that risk is an important aspect of agricultural technology. While this recognition is valuable in itself, a machinery that deals analytically with risk in the absence of knowledge of farmers' individual attitudes to risk has not hitherto been exploited. Following the conclusion that efficiency analysis is a more or less satisfactory process, what are the implications for agricultural research and extension?

There seem to be some clear guides for research workers. Rather than focusing on estimation of treatment means, whole probability distributions should be explored and estimated to complement conventional "average-oriented" research if risk-averse users are to be well served. This implies that "risk-oriented" research will be generally more demanding and more expensive than "average-oriented" research. Seemingly, this is the price one should pay for work that is potentially relevant in this context.

More particularly, the appraisal of stochastic efficiency implies pinning down the lower tails of probability distributions. This estimational task suggests that agricultural innovations need to be evaluated and reported under the bad as well as typical or average environmental conditions that potential adopters face (e.g., with respect to moisture stress, disease exposure, nutrient suboptimization, etc.).

To the extent that identifiable groups of potentially adopting farmers face different "worst" conditions (if not also different "average" or "good" conditions), efficient technological packages may differ among groups. Consequently, risk-oriented research should deliberately span an appropriate range of environments and environmental conditions, which will usually imply replication over space and time. In the short run especially, there seems to be much unexploited scope for formal documentation of research agronomists' considerable experience of and largely unpublished nowledge of the tails (especially the lower) of relevant probability dis-:ibutions.

People working in agricultural extension should also appreciate the mplications of stochastic efficiency analysis. Extension of technological ıdvice in risky agricultures will certainly be more effective if due recognition is given to the impact of risk and the importance of technologically nduced risk. Such extension will be simplified by dealing with farmers grouped according to the worst environmental conditions faced, and its success will be enhanced by promoting practices that are tailored to be stochastically efficient (to at least the second degree) for the identified groups. Moreover, in judging extension efforts, recognition that a recommendation efficient in terms of average profit may not be risk efficient should temper appraisal of programs.

If an extension effort is mounted on such a scale that it is possible to elicit individual farmers' attitudes to risk and perceived probability distributions, the analysis of stochastic efficiency would become redundant. Until now, however, such a situation has apparently not been attained anywhere—and nowhere does it seem to be an imminent prospect.

Perceptive practitioners of the arts of agricultural research and extension inevitably develop a keen intuition for the importance of risk in most agricultural production. However, their formal training has usually done little or nothing to equip them with an analytical apparatus for dealing directly with this aspect of their work. Clearly, educational programs could do more to sensitize future practitioners to the impact of risk in farming and, consequently, in research and extension. Particular attention needs to be drawn to the fact that traditional experimental methodology is inevitably addressed to estimating only average effects, responses, etc., and accordingly is only directly applicable to risk-indifferent farmer-users. Likewise, probabilistically based educational programs can provide farmers with valuable information on situations they may face under various seasonal conditions. With such information, farmers could appropriately modify their perceptions of risk inherent in various technologies and their associated farm plans.

Applied agricultural research can be judged as potentially worthwhile when it leads to new farming practices that are stochastically efficient relative to existing practices. If the new practices are also stochastically dominant (minimally of degree three and most desirably of degree one), the chance of the research being positively beneficial is correspondingly greater. Of course, the cost of the research should enter the economic evaluation. Conversely, if after a research program has been completed and extended to the farming community and farmers' prior (competing) practices fall in the risk-efficient set, returns from the research must be highly uncertain and may well be negative.

Research planning in the context of stochastic efficiency appears to be intrinsically difficult in at least two senses. First, only retrospective analysis of efficiency is in any way straightforward. However, if research is to be directed toward the development of stochastically efficient new technologies, research planners necessarily must aim to identify technologies that are not only more profitable on the average but are also less prone to low outcomes under unfavorable conditions. Second, to the extent that new technologies embody considerable changes in the rate of utilization of constrained farm resources, it is necessary to assess them in a context of whole-farm planning, as illustrated in the REMP method. Technological assessment on a practice-by-practice basis is thus adequate only for rather minor changes such as modifications in use of varieties and fertilizer.

Agricultural Policy

Formulators of agricultural policy will generally suffer fewer "surprises" in program results if their economic models of farmer behavior include an adequate recognition of farming risks and farmers' attitudes toward them. The main pertinent policy instruments for influencing risk have been crop insurance schemes, weather-oriented income tax arrangements, and minimal price supports for agricultural products. Our discussion of stochastic efficiency places a new slant on such schemes.

By focusing on values in the lower tails of distributions of yields and prices rather than on values near the mean, considerations of stochastic efficiency suggest that relatively low premiums or guarantees may still encourage significant adjustments to farmers' actions. More specifically, a crop "insurance" scheme that effectively truncates yield distributions below a crossover point in the lower tails of two simply related varietal distributions causes the variety that yields higher on the average to be first-degree stochastically dominant. Ensuing adoption by farmers of the now FSE variety could be in the national interest. Typically, a recommended technological practice will dominate traditional practices if it can be "insured" to the extent that under really poor eventualities farmers are not disadvantaged by adoption. Effective truncation of the lower tails of market price distributions through minimal—but low—price and income supports may effectively "take the risk out of" a program under government sponsorship.

9.5 COMPUTER PROGRAM FOR STOCHASTIC EFFICIENCY ANALYSIS

The subprograms listed below are written in the FORTRAN IV language and perform the review of stochastic efficiency discussed in Secion 9.2 and elaborated by Anderson (1974b). These programs are used by lefining a matrix of the coordinates of a set of CDFs and a corresponding vector of numerical identifications and then calling the subroutine SDØM (which in turn calls the other listed subprograms SEL, CSUB, and JASØ2).

The dimensions specified in these programs presently permit up to 20 linear segments (equally spaced in probability) to describe each of up to 40 CDFs. If 20 segments are used, 21 numbers are required to describe the fractiles $f_{0.0}$, $f_{0.05}, \ldots, f_{1.0}$ and these should be in nondecreasing order

The arguments of SDØM are in turn:

NA = the number of distributions to be reviewed (presently $\leqq$ 40 but may be increased by increasing the dimensions specified)

NC = the number of points on each CDF (presently $\leqq$ 21 but may also be increased)

F = a matrix in which distributions are described one per row for each of the first *NA* rows and the points on each CDF are stored in the first *NC* columns of each row, with the zero fractile in the first column

IN = a vector of numerical tags (e.g., 1, 2, ..., *NA*) that respectively identify the distributions stored in *F*

IE = a scalar indicator of use of *ID* when *IE* < 0

ID = a vector for indicating the prespecified stochastic inefficiency of corresponding distributions in *F* by setting the respective elements of *ID* to +1 and all other elements to zero

JT = a scalar dummy argument that takes on the value of the number of elements in the smallest stochastically efficient set

X = a vector of means of the *NA* distributions identified in *IN*. When these are not supplied to the subroutine, the first element of *X* should be set to zero whereupon the means are computed in the subprogram.

Listing of SDØM and Related Subroutines

```
      SUBROUTINE SDOM(NA,NC,F,IN,IE,ID,JT,X)
      DIMENSION F(40,21),S(40,21),T(40,21),ID(40),IN(40),X(40)
C     MODIFY PREVIOUS STATEMENT IF MORE THAN 40 DISTRIBUTIONS
      DIMENSION Z(2,42)
C      MODIFY PREVIOUS 2 STATEMENTS IF MORE THAN 20 SEGMENTS
C             NS EQUAL-PROBABILITY LINEAR SEGMENTS ON CDF'S
      NS=NC-1
      DP=NS
C     DP=ELEMENT OF CUMULATIVE PROBABILITY
      DP=1./DP
      NA1=NA-1
C      COMPUTE MEANS IF NOT SUPPLIED AND STORE IN X
      IF(X(1).NE. 0.)GO TO 102
```

```
      DO 101 I=1,NA
      X(I)=0.
      DO 100 J=1,NS
  100 X(I)=X(I)+F(I,J)
  101 X(I)=DP*(X(I)+(F(I,NC)-F(I,1))*.5)
  102 CONTINUE
      IF(IE.LT.0)GO TO 104
      DO 103 I=1,NA
  103 ID(I)=0
  104 CONTINUE
C     CHECK ALL CDF'S ARE PROPER
      DO 105 I=1,NA
      IF(ID(I).EQ.1)GO TO 105
      DO 105 J=1,NS
      IF(F(I,J+1).GE.F(I,J))GO TO 105
      CDFE=DP*(J-1)
      WRITE(2,550)I,IN(I),CDFF
  550 FORMAT(1X,'PROSPECT',I3,' CALLED ',I4,' DEFECTIVE ABOUT THE ',F5.3
     *,' FRACTILE',/ )
      ID(I)=1
  105 CONTINUE
C     MAKE ALL POSSIBLE COMPARISONS OF FSD(CDF)FUNCTIONS
      DO 160 IC=1,NA1
      IM=IC+1
      DO 160 JC=IM,NA
      IK=IC
      JK=JC
C      ID INDEXES (=1 IF) PROSPECT I IS DOMINATED
      IF(ID(IK).EQ.1.OR.ID(JK).EQ.1)GO TO 160
      IT=0
      DO 110 L=1,NC
      IF(F(IK,L)-F(JK,L))114,110,115
  110 CONTINUE
C     IDENTIFY POTENTIALLY DOMINANT DISTRIBUTION AS IK
  114 CALL CSUB(IK,JK,IT)
  115 L=L+1
      IF(X(IK).LT.X(JK)) GO TO 155
      DO 140 K=L,NC
      IF( F(IK,K). LT.F(JK,K))GO TO 155
  140 CONTINUE
      ID(JK)=1
  155 IF(IT.EQ.0) GO TO  160
      CALL CSUB(IK,JK,IT)
  160 CONTINUE
      CALL SEL(1,NA,ID,NA,IN,JF)
      JT=JF
      IF(JF.LT.2)GO TO 390
C      THROW OUT DOMINATED ACTS
      DO 200 I=1,JF
      K=ID(I)
      X(I)=X(K)
      DO 200 J=1,NC
  200 F(I,J)= F(K,J)
C      START SSD REVIEW
      DO 205 I=1,JF
  205 ID(I)=0
C     COMPUTE SSD FUNCTION AT SEGMENT ENDPOINTS
      DO 210 I=1,JF
      S(I,1)=0.
      DO 210 J=2, NC
      A=J
  210 S(I,J)= S(I,J-1)+ DP*(F(I,J)-F(I,J-1))*(A-1.5)
      JF1=JF-1
C      START PAIRWISE COMPARISONS
      DO 260 IC=1,JF1
      IM=IC+1
      DO 260 JC=IM,JF
      IK=IC
      JK=JC
```

```
C       SKIP DOMINATED ACTS ALREADY IDENTIFIED
        IF(ID(IK).EQ.1.OR.ID(JK).EQ.1)GO TO 260
        IT=0
        DO 215 L=1,NC
        IF(F(IK,L)-F(JK,L))214,213,215
  213   CONTINUE
  214   CALL CSUB(IK,JK,IT)
  215   IZ=2
        IF(X(IK).LT.X(JK)) GO TO 255
        DO 240 K=2,NC
        DO 220 IX=IZ,NC
        IF(F(JK,IX).GT.F(IK,K))GO TO 225
  220   CONTINUE
        IZ=NC
        GO TO 230
  225   IZ= IX-1
        A=IZ
        RI= F(IK,K)-F(JK,IZ)
        SCF=S(JK,IZ)
        IF(F(JK,IX).EQ.F(JK,IZ))GO TO 235
C        AND AVOID POSSIBLE ZERO DIVISION
        SCF=S(JK,IZ)+.5*RI**2*DP/(F(JK,IX)-F(JK,IZ))+
       2RI*DP*(A-1.)
        GO TO 235
  230   RI=F(IK,K)-F(JK,IZ)
C       LINEAR EXTRAPOLATION OF JK SCOLLOPS FOR UPPER IK S FUNCTION
        SCF=S(JK,IZ)+RI
  235   IF(S(IK,K).GT. SCF)GO TO 255
  240   CONTINUE
        IF(IZ.GT.NC-1)GO TO 250
        IZ=IZ+1
C        LINEAR EXTRAPOLATION OF IK FUNCTION FOR UPPER JK S FUNCTION
        DO 245 IX=IZ,NC
        SCF=S(IK,NC) + F(JK,IX)-F(IK,NC)
        IF(SCF.GT.S(JK,IX))GO TO 255
  245   CONTINUE
  250   CONTINUE
C       CHECK POSSIBLE INTERSECTION AT OTHER THAN IK JOIN POINTS
C       MERGE
        NC1=NC+1
        NC2=NC*2
        DO 251 I=1,NC
        J=NC+I
        Z(1,J)=F(JK,I)
        Z(2,J)=0.
        Z(1,I)=F(IK,I)
  251   Z(2,I)=1.
        CALL JASO2(NC2,Z)
        IH=0
        JH=0
        DO 252 M=1,NC2
        IF(Z(2,M).EQ.0.)JH=JH+1
        IF(Z(2,M).EQ.1.)IH=IH+1
        IF(IH.EQ.1)GO TO 253
  252   CONTINUE
  253   NC21=NC2-1
        AIT=0
        AJT=0
        IF(F(IK,IH+1).NE.F(IK,IH))AIT=.5*DP/(F(IK,IH+1)-F(IK,IH))
        IF(F(JK,JH+1).NE.F(JK,JH))AJT=.5*DP/(F(JK,JH+1)-F(JK,JH))
        AAA=AJT-AIT
        IF(AAA.EQ.0.)GO TO 2534
C       SOLN IS NON-QUADRATIC AND INTERSECTN WOULD HAVE BEEN PICKED UP PREV
        BBB=2.*(AIT*F(IK,IH)-AJT*F(JK,JH)) + DP*(JH-IH)
        CCC=DP*(IH-1)*F(IK,IH)-DP*(JH-1)*F(JK,JH)+S(JK,JH)-S(IK,IH)
        CCC=CCC+AJT*F(JK,JH)**2-AIT*F(IK,IH)**2
        RAD=BBB*BBB-4.*AAA*CCC
        IF(RAD.LT.0.)GO TO 2534
```

```
C       NO REAL SOLUTION SO MOVE ON TO NEXT SEGMENT
      RAD=SQRT(RAD)
      SOL1=(-BBB+RAD)/2./AAA
      SOL2=(-BBB-RAD)/2./AAA
      IF(Z(1,M+1).EQ.Z(1,M))GO TO 2534
C     ONLY LOOK FOR INTERSECTIONS WITHIN RANGE
      IF(SOL1.LT.Z(1,M+1).AND.SOL1.GT.Z(1,M).OR.SOL2.LT.Z(1,M+1).
     1AND.SOL2.GT.Z(1,M)   )GO TO  2555
C      C  BECAUSE CURVES INTERSECT, SO NO SSD
 2534 CONTINUE
      IF(M.EQ.NC21)GO TO 254
      M=M+1
      IF(Z(2,M).EQ.0.)JH=JH+1
      IF(Z(2,M).EQ.1.)IH=IH+1
      IF(IH.EQ.NC.OR.JH.EQ.NC)GO TO 254
      GO TO 253
  254 ID(JK)=1
      GO TO 255
 2555 CONTINUE
  255 IF(IT.EQ.0)GO TO260
      CALL CSUB(IK,JK,IT)
  260 CONTINUE
      CALL SEL(2,JF,ID,NA,IN,JS)
      JT=JS
      IF(JS.LT.2)GO TO 390
      DO 300 I=1,JS
      K=ID(I)
      X(I)=X(K)
      DO 300 J=1,NC
      F(I,J)=F(K,J)
  300 S(I,J)=S(K,J)
      DO 305 I=1,JS
  305 ID(I)=0
C       COMPUTE TSD FUNCTION AT SEGMENT ENDPOINTS
      DO 310 I=1,JS
      T(I,1)=0.
      DO 310 J=2,NC
      A=J
      DI=F(I,J)-F(I,J-1)
  310 T(I,J)=T(I,J-1)+.5*DP*DI*DI*(A-2.+1./3.)+DI*S(I,J-1)
      JS1=JS-1
      DO 360 IC=1,JS1
      IM=IC+1
      DO 360 JC=IM,JS
      IK=IC
      JK=JC
      IF(ID(IK).EQ.1.OR.ID(JK).EQ.1)GO TO 360
      IT=0
      DO 313 L=1,NC
      IF(F(IK,L)-F(JK,L))314,313,315
  313 CONTINUE
  314 CALL CSUB(IK,JK,IT)
  315 IZ=2
      IF(X(IK).LT.X(JK)) GO TO 355
C          SUBSIDIARY CHECK ON UPPER SSD FUNCTION FOR TSD
  316 IF(F(IK,NC).LT.F(JK,NC))GO TO 317
C         EXTRAPOLATE JK
      RI=F(IK,NC)-F(JK,NC)
      SCF=S(JK,NC)+RI
      IF(S(IK,NC).GT.SCF)GO TO 355
      GO TO 318
C         EXTRAPOLATE IK
  317 SCF=S(IK,NC)+F(JK,NC)-F(IK,NC)
      IF(SCF.GT.S(JK,NC))GO TO 355
  318 CONTINUE
      DO 340 K=2,NC
      DO 320 IX=IZ,NC
      IF(F(JK,IX).GT.F(IK,K))GO TO 325
```

```
  320 CONTINUE
      IZ=NC
      GO TO 330
  325 IZ=IX-1
      A=IZ
      RI=F(IK,K)-F(JK,IZ)
      TCF=T(JK,IZ)
     2+.5*RI**2*DP*(A-1.)+RI*S(JK,IZ)
      IF(F(JK,IX).EQ.F(JK,IZ))GO TO 335
C       AND AVOID POSSIBLE ZERO DIVISION
      TCF=  TCF   +1./6.* RI**3*DP/(F(JK,IX)-F(JK,IZ))
      GO TO 335
  330 RI= F(IK,K)-F(JK,IZ)
      TCF= T(JK,IZ)+ .5*RI**2 + S(JK,IZ)*RI
  335 IF(T(IK,K).GT. TCF)GO TO 355
  340 CONTINUE
      IF(IZ.GT. NC-1)GO TO 350
      IZ=IZ+1
      DO 345 IX=IZ,NC
      RI= F(JK,IX)-F(IK,NC)
      TCF= T(IK,NC) + .5*RI**2+ S(IK,NC)*RI
      IF(TCF.GT. T(JK,IX))GO TO 355
  345 CONTINUE
  350 CONTINUE
      ID(JK)=1
  355 IF(IT.EQ.0)GO TO 360
      CALL CSUB(IK,JK,IT)
  360 CONTINUE
      CALL SEL(3,JS,ID,NA,IN,JT)
  390 RETURN
      END
      SUBROUTINE SEL(IB,NB,ID,NA,IN,JF)
C     REPORTS NUMERICAL TAGS OF STOCHASTICALLY EFFICIENT SETS
      DIMENSION ID(40),IN(40)
      JF=0
      DO 170 I=1,NB
      IF(ID(I).EQ.1) GO TO 170
      JF=JF+1
      ID(JF)=I
      IN(JF)=IN(I)
  170 CONTINUE
      IF(JF.EQ.0)RETURN
      WRITE(2,420)JF,IB
      K=1
      IF(JF.LE.5)GO TO 180
C   ASSUME ASCENDING NUMERICAL LABELS
      IND=IN(2)-IN(1)
      DO 160 I=3,JF
      IT=IN(I)-IN(I-1)
      IF(IT.LT.IND)IND=IT
  160 CONTINUE
      INS=IN(1)
      DO 175 I=2,JF
      K=I-1
      IF((IN(I)-IN(K)).NE.IND)GO TO 177
  175 CONTINUE
      I=0
      K=JF
  177 CONTINUE
      IF(K.EQ.1)GO TO 180
      WRITE(2,430)INS,IND,IN(K)
      IF(I.EQ.0)RETURN
  180 WRITE(2,410) (IN(J),J=K,JF)
  410 FORMAT(3(1X,I4,23I5,/))
  420 FORMAT(20X,I3,' EFFICIENT PROSPECTS OF DEGREE',I3)
  430 FORMAT(8X,23HINITIAL EFFICIENT RANGE,2X,I5,1H(,I2,1H),I4   )
      RETURN
      END
```

```
      SUBROUTINE CSUB(IK,JK,IT)
C     CHANGE SUBSCRIPTS FOR REVERSE COMPARISON
      IT=IK
      IK=JK
      JK=IT
      IT=1
      RETURN
      END
      SUBROUTINE JASO2(N,X)
C    SORTS ON ROW 1 AND CARRIES ROW 2
      DIMENSION X(2,N)
      N1=N-1
      DO 20 I=1,N1
      NI=N-I
      DO 20 J=1,NI
      IF(X(1,J)-X(1,J+1))20,20,10
   10 T=X(1,J)
      S=X(2,J)
      X(1,J)=X(1,J+1)
      X(2,J)=X(2,J+1)
      X(1,J+1)=T
      X(2,J+1)=S
   20 CONTINUE
      RETURN
      END
```

PROBLEMS

9.1. Consider the following pair of prospects:

$P(x_i)$	x_i
0.2	2
0.3	3
0.5	5

$P(x_i)$	x_i
0.6	3
0.4	6

Are they both efficient in all senses of stochastic efficiency?

9.2. Three alternative rice production technologies have returns that are approximately normally distributed with parameters:

Technology	Mean	Standard Deviation
A	1200	400
B	1000	300
T	500	100

Which technologies are SSE?

9.3. Two similar technological possibilities are open to maize growers in a region. The distributions of yield (t/ha) are described by two sets of subjective fractiles:

Fractile	0	0.25	0.5	0.75	1.0
Technology *A*	0.75	1.25	1.5	2.0	3.5
Technology *B*	1.00	2.00	3.1	4.5	6.0

Suppose the net value of grain is \$100/t and the variable costs associated with these technologies are \$50/ha for *A* and \$100/ha for *B*. Sketch the various cumulative functions required to review stochastic efficiency and declare if any technology is uniquely efficient. (Hint: use linear-segmented CDFs.)

9.4. Examine the (*E*, *V*)-efficient frontiers generated by quadratic programming and presented in Figures 7.4 and 9.9. If we now assume (an unspecified) risk-averse *quadratic* utility function, can the efficient frontiers be further narrowed?

9.5. "Stochastic efficiency analysis may rest on acceptable general assumptions, but results are typically too vague to be useful." Comment.

9.6. Prove to your own satisfaction that the SSD ordering rule is valid. [Consulting Hadar and Russell (1969), for example, would assist considerably.]

9.7. Consider the following three distributions:

Family	Mean	Variance
Normal	10	10
Log-normal	10	10
Poisson	10	10

What can be said of stochastic efficiency for a risk averter faced with these alternatives?

9.8. Net returns x from using input v, $0 \leqq v \leqq 5.5$, are normally distributed with parameters

$$E(x) = 100 + 10v - v^2$$

$$V(x) = 1000 + 100v - 20v^2$$

What is the SSE range of v?

9.9. Rework Problem 2.9d (1) incorporating only the additional assumption of aversion to risk.

9.10. In Problem 4.3 what fertilizer actions are TSE?

9.11. Suppose in Problem 5.4 that Midas accepts the free demonstration on three trees and finds one is damaged. Does Midas have an SSE act open to it?

9.12. Reconsider Problem 5.10f. Is there a unique stochastically efficient recommendation? Do stochastic efficiency concepts appear to have much applicability in an extension context?

REFERENCES AND SELECTED FURTHER READING

Anderson, J. R. 1973. Sparse data, climatic variability, and yield uncertainty in response analysis. *Am. J. Agric. Econ.* 55(1): 77–82.

______. 1974a. Sparse data, estimational reliability and risk-efficient decisions. *Am. J. Agric. Econ.* 56(3): 564–72.

______. 1974b. Risk efficiency in the interpretation of agricultural production research. *Rev. Mktg. Agric. Econ.* 42(3): 131–84.

______. 1975. Programming for efficient planning against non-normal risk. *Australian J. Agric. Econ.* 19(2): 94–107.

———. 1976. Risk efficient Monte Carlo programming—A manual. Agric. Econ. Bus. Mgmt. Misc. Pub. 6. Armidale, Australia: Univ. New England.
Atkinson, A. B. 1970. On the measurement of inequality. *J. Econ. Theory* 2(3): 244–63.
CIMMYT. 1972. Results of the 6th international spring wheat yield nursery. Res. Bull. 23, CIMMYT, Mexico City.
Donaldson, G. F., and J. P. G. Webster. 1968. An operating procedure for simulation farm planning—Monte Carlo method. Wye College (Univ. London), Dept. Agric. Econ.
Finlay, K. W., and G. M. Wilkinson. 1963. The analysis of adaptation in a plant breeding programme. *Australian J. Agric. Res.* 14(6): 742–54.
Fishburn, P. C. 1964. *Decision and Value Theory.* New York: Wiley.
———. 1974a. Convex stochastic dominance with continuous distribution functions. *J. Econ. Theory* 7(2): 143–58.
———. 1974b. Convex stochastic dominance with finite consequence sets. *Theory and Decision* 5(2): 119–37.
Frankfurter, G. M., H. E. Phillips, and J. P. Seagle. 1971. Portfolio selection: The effects of uncertain means, variances and covariances. *J. Financ. Quant. Anal.* 6(5): 1251–62.
Gould, J. P. 1974. Risk, stochastic preference and the value of information. *J. Econ. Theory* 8(1): 64–84.
Hadar, J. 1971. *Mathematical Theory of Economic Behavior.* Reading, Mass.: Addison-Wesley.
Hadar, J., and W. R. Russell. 1969. Rules for ordering uncertain prospects. *Am. Econ. Rev.* 59(1): 25–34.
———. 1974. Stochastic dominance and diversification. *J. Econ. Theory* 7(3): 231–40.
———. 1974b. Stochastic dominance in choice under uncertainty. In *Essays on Economic Behavior under Uncertainty,* M. Balch, D. McFadden, and S. Wu, eds., pp. 133–50. Amsterdam: North-Holland.
Hammond, J. S. 1968. Towards simplifying the analysis of decisions under uncertainty where preference is nonlinear. Harvard Univ. D.B.A. thesis. Ann Arbor: Univ. Microfilms.
———. 1974. Simplifying the choice between uncertain prospects where preference is nonlinear. *Mgmt. Sci.* 20(7): 1047–72.
Hanoch, G., and H. Levy. 1969. The efficiency analysis of choices involving risk. *Rev. Econ. Stud.* 36(3): 335–46.
———. 1970. Efficient portfolio selection with quadratic and cubic utility. *J. Busin.* 43(2): 181–89.
Hanson, D. J., and C. F. Menezes. 1971. On a neglected aspect of the theory of risk aversion. *Western Econ. J.* 9(2): 211–17.
Hardaker, J. B., and A. G. Tanago. 1973. Assessment of the output of a stochastic decision model. *Australian J. Agric. Econ.* 17(3): 170–78.
Hazell, P. B. R. 1971. A linear alternative to quadratic and semivariance programming for farm planning under uncertainty. *Am. J. Agric. Econ.* 53(1): 53–62.
Joy, O. M., and R. B. Porter. 1974. Stochastic dominance and mutual fund performance. *J. Financ. Quant. Anal.* 9(1): 25–31.
Kalymon, B. A. 1971. Estimation risk in the portfolio selection model. *J. Financ. Quant. Anal.* 6(1): 559–82.
Kihlstrom, R. E., and L. J. Mirman. 1974. Risk aversion with many commodities. *J. Econ. Theory* 8(3): 361–88.
Levy, H. 1973a. Stochastic dominance among log-normal prospects. *Int. Econ. Rev.* 14(3): 601–14.
———. 1973b. Stochastic dominance, efficiency criteria and efficient portfolios: The multi-period case. *Am. Econ. Rev.* 63(5): 986–94.
Levy, H., and J. Paroush. 1974a. Toward multivariate efficiency criteria. *J. Econ. Theory* 7(2): 129–42.

———. 1974b. Multi-period stochastic dominance. *Mgmt. Sci.* 21(4): 428–41.
Levy, H., and M. Sarnat. 1971. A note on portfolio selection and investors' wealth. *J. Financ. Quant. Anal.* 6(1): 639–42.
Masson, R. T. 1974. Utility functions with jump discontinuities: Some evidence and implications from peasant agriculture. *Econ. Inquiry* 12(4): 559–66.
Mihram, G. A. 1972. *Simulation: Statistical Foundations and Methodology.* New York: Academic Press.
O'Mara, G. T. 1971. A decision-theoretic view of the microeconomics of technique diffusion in a developing country. Stanford Univ. Ph.D. diss. Ann Arbor: Univ. Microfilms.
Philippatos, G. C., and N. Gressis. 1975. Conditions for equivalence among *E-V*, SSD and *E-H* portfolio selection criteria: The case for uniform, normal and log-normal distributions. *Mgmt. Sci.* 21(6): 617–25.
Porter, R. B. 1973. An empirical comparison of stochastic dominance and mean-variance portfolio choice. *J. Financ. Quant. Anal.* 8(4): 587–605.
———. 1974. Semivariance and stochastic dominance: A comparison. *Am. Econ. Rev.* 64(1): 200–4.
Porter, R. B., and J. E. Gaumnitz. 1972. Stochastic dominance vs. mean-variance portfolio analysis: An empirical evaluation. *Am. Econ. Rev.* 62(3): 438–46.
Porter, R. B., J. R. Wart, and D. L. Ferguson. 1973. Efficient algorithms for conducting stochastic dominance tests of large numbers of portfolios. *J. Financ. Quant. Anal.* 8(1): 71–81.
Pratt, J. 1964. Risk aversion in the small and in the large. *Econometrica* 32(1-2): 122–36.
Pratt, J. W., H. Raiffa, and R. Schlaifer. 1965. *Introduction to Statistical Decision Theory.* New York: McGraw-Hill.
Quirk, J. P., and R. Saposnik. 1962. Admissibility and measurable utility functions. *Rev. Econ. Stud.* 29(2): 140–46.
Rothschild, M., and J. E. Stiglitz. 1970. Increasing risk: I. A definition. *J. Econ. Theory* 2(3): 225–43.
Russell, W. R., and P. E. Smith. 1970. Taxation, risk-taking and stochastic dominance. *Southern Econ. J.* 36(4): 425–33.
Schlaifer, R. 1959. *Probability and Statistics for Business Decisions.* New York: McGraw-Hill.
Shapiro, S. S., and M. B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52(3-4): 591–611.
Vickson, R. G. 1975. Stochastic dominance tests for decreasing absolute risk aversion. I. Discrete random variables. *Mgmt. Sci.* 21(12): 1438–46.
Whitmore, G. A. 1970. Third-degree stochastic dominance. *Am. Econ. Rev.* 60(3): 457–59.

APPENDIX

TABLE A.1. Values of $N(D)$ for the Standard Normal Distribution

D	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.3989	.3940	.3890	.3841	.3793	.3744	.3697	.3649	.3602	.3556
.1	.3509	.3464	.3418	.3373	.3328	.3284	.3240	.3197	.3154	.3111
.2	.3069	.3027	.2986	.2944	.2904	.2863	.2824	.2784	.2745	.2706
.3	.2668	.2630	.2592	.2555	.2518	.2481	.2445	.2409	.2374	.2339
.4	.2304	.2270	.2236	.2203	.2169	.2137	.2104	.2072	.2040	.2009
.5	.1978	.1947	.1917	.1887	.1857	.1828	.1799	.1771	.1742	.1714
.6	.1687	.1659	.1633	.1606	.1580	.1554	.1528	.1503	.1478	.1453
.7	.1429	.1405	.1381	.1358	.1334	.1312	.1289	.1267	.1245	.1223
.8	.1202	.1181	.1160	.1140	.1120	.1100	.1080	.1061	.1042	.1023
.9	.1004	.09860	.09680	.09503	.09328	.09156	.08986	.08819	.08654	.08491
1.0	.08332	.08174	.08019	.07866	.07716	.07568	.07422	.07279	.07138	.06999
1.1	.06862	.06727	.06595	.06465	.06336	.06210	.06086	.05964	.05844	.05726
1.2	.05610	.05496	.05384	.05274	.05165	.05059	.04954	.04851	.04750	.04650
1.3	.04553	.04457	.04363	.04270	.04179	.04090	.04002	.03916	.03831	.03748
1.4	.03667	.03587	.03508	.03431	.03356	.03281	.03208	.03137	.03067	.02998
1.5	.02931	.02865	.02800	.02736	.02674	.02612	.02552	.02494	.02436	.02380
1.6	.02324	.02270	.02217	.02165	.02114	.02064	.02015	.01967	.01920	.01874
1.7	.01829	.01785	.01742	.01699	.01658	.01617	.01578	.01539	.01501	.01464
1.8	.01428	.01392	.01357	.01323	.01290	.01257	.01226	.01195	.01164	.01134
1.9	.01105	.01077	.01049	.01022	.0^{2}9957	.0^{2}9698	.0^{2}9445	.0^{2}9198	.0^{2}8957	.0^{2}8721
2.0	.0^{2}8491	.0^{2}8266	.0^{2}8046	.0^{2}7832	.0^{2}7623	.0^{2}7418	.0^{2}7219	.0^{2}7024	.0^{2}6835	.0^{2}6649
2.1	.0^{2}6468	.0^{2}6292	.0^{2}6120	.0^{2}5952	.0^{2}5788	.0^{2}5628	.0^{2}5472	.0^{2}5320	.0^{2}5172	.0^{2}5028
2.2	.0^{2}4887	.0^{2}4750	.0^{2}4616	.0^{2}4486	.0^{2}4358	.0^{2}4235	.0^{2}4114	.0^{2}3996	.0^{2}3882	.0^{2}3770
2.3	.0^{2}3662	.0^{2}3556	.0^{2}3453	.0^{2}3352	.0^{2}3255	.0^{2}3159	.0^{2}3067	.0^{2}2977	.0^{2}2889	.0^{2}2804
2.4	.0^{2}2720	.0^{2}2640	.0^{2}2561	.0^{2}2484	.0^{2}2410	.0^{2}2337	.0^{2}2267	.0^{2}2199	.0^{2}2132	.0^{2}2067
2.5	.0^{2}2004	.0^{2}1943	.0^{2}1883	.0^{2}1826	.0^{2}1769	.0^{2}1715	.0^{2}1662	.0^{2}1610	.0^{2}1560	.0^{2}1511
2.6	.0^{2}1464	.0^{2}1418	.0^{2}1373	.0^{2}1330	.0^{2}1288	.0^{2}1247	.0^{2}1207	.0^{2}1169	.0^{2}1132	.0^{2}1095
2.7	.0^{2}1060	.0^{2}1026	.0^{3}9928	.0^{3}9607	.0^{3}9295	.0^{3}8992	.0^{3}8699	.0^{3}8414	.0^{3}8138	.0^{3}7870
2.8	.0^{3}7611	.0^{3}7359	.0^{3}7115	.0^{3}6879	.0^{3}6650	.0^{3}6428	.0^{3}6213	.0^{3}6004	.0^{3}5802	.0^{3}5606
2.9	.0^{3}5417	.0^{3}5233	.0^{3}5055	.0^{3}4883	.0^{3}4716	.0^{3}4555	.0^{3}4398	.0^{3}4247	.0^{3}4101	.0^{3}3959
3.0	.0^{3}3822	.0^{3}3689	.0^{3}3560	.0^{3}3436	.0^{3}3316	.0^{3}3199	.0^{3}3087	.0^{3}2978	.0^{3}2873	.0^{3}2771
3.5	.0^{4}5848	.0^{4}5620	.0^{4}5400	.0^{4}5188	.0^{4}4984	.0^{4}4788	.0^{4}4599	.0^{4}4417	.0^{4}4242	.0^{4}4073
4.0	.0^{5}7145	.0^{5}6835	.0^{5}6538	.0^{5}6253	.0^{5}5980	.0^{5}5718	.0^{5}5468	.0^{5}5227	.0^{5}4997	.0^{5}4777

Example: if $D = 3.57$, $N(D) = 0.0^{4}4417 = 0.00004417$.

Source: From *Probability and Statistics for Business Decisions* by R. Schlaifer, © 1959, McGraw-Hill Book Co., Inc. Used with permission of McGraw-Hill Book Co.

TABLE A.2. Fractiles of the Beta Distribution

c+d	c	p: .01	.05	.10	.15	.20	.25	.30	.35	.40	.45
2	1	0100	0500	1000	1500	2000	2500	3000	3500	4000	4500
3	1	0050	0253	0513	0780	1056	1340	1633	1938	2254	2584
4	1	0033	0170	0345	0527	0717	0914	1121	1338	1566	1807
4	2	0589	1354	1958	2444	2871	3264	3633	3986	4329	4666
5	1	0025	0127	0260	0398	0543	0694	0853	1021	1199	1388
5	2	0420	0976	1426	1794	2123	2430	2724	3010	3292	3573
6	1	0020	0102	0209	0320	0436	0559	0689	0825	0971	1127
6	2	0327	0764	1122	1419	1686	1938	2180	2418	2656	2895
6	3	1056	1893	2466	2899	3266	3594	3898	4186	4463	4733
7	1	0017	0085	0174	0267	0365	0468	0577	0693	0816	0948
7	2	0268	0628	0926	1174	1399	1612	1818	2022	2226	2433
7	3	0847	1532	2009	2374	2686	2969	3233	3486	3731	3973
8	1	0014	0073	0149	0229	0314	0403	0497	0597	0704	0819
8	2	0227	0534	0788	1001	1195	1380	1559	1737	1916	2098
8	3	0708	1288	1696	2011	2283	2531	2763	2987	3206	3423
8	4	1423	2253	2786	3176	3501	3788	4052	4301	4539	4771
9	1	0013	0064	0131	0201	0275	0353	0436	0524	0619	0720
9	2	0197	0464	0686	0873	1044	1206	1365	1523	1682	1844
9	3	0608	1111	1469	1746	1986	2206	2413	2614	2811	3007
9	4	1210	1929	2397	2742	3032	3291	3530	3756	3975	4189
10	1	0011	0057	0116	0179	0245	0315	0389	0467	0552	0643
10	2	0174	0410	0608	0774	0926	1072	1214	1355	1498	1645
10	3	0533	0977	1295	1542	1757	1955	2142	2324	2502	2681
10	4	1053	1688	2104	2414	2675	2910	3127	3335	3535	3733
10	5	1710	2514	3010	3367	3661	3920	4156	4378	4590	4796
11	1	0010	0051	0105	0161	0221	0284	0350	0422	0498	0580
11	2	0155	0368	0545	0695	0833	0964	1093	1221	1351	1485
11	3	0475	0873	1158	1381	1576	1756	1926	2092	2255	2419
11	4	0932	1500	1876	2156	2394	2609	2808	2999	3184	3367
11	5	1504	2224	2673	2998	3268	3507	3726	3932	4131	4325
12	1	0009	0047	0095	0147	0201	0258	0319	0384	0454	0529
12	2	0141	0333	0495	0631	0756	0876	0994	1111	1230	1353
12	3	0428	0788	1048	1251	1429	1593	1750	1902	2052	2204
12	4	0837	1351	1692	1949	2167	2364	2548	2724	2896	3067
12	5	1344	1996	2405	2704	2953	3173	3377	3570	3755	3938
12	6	1940	2712	3177	3508	3779	4016	4232	4434	4627	4815
13	1	0008	0043	0088	0135	0184	0237	0293	0353	0417	0486
13	2	0128	0305	0452	0577	0693	0803	0911	1020	1130	1242
13	3	0390	0719	0957	1143	1307	1459	1603	1744	1883	2024
13	4	0759	1229	1542	1778	1979	2162	2332	2496	2656	2815
13	5	1215	1810	2187	2463	2693	2898	3088	3268	3443	3614
13	6	1746	2453	2882	3189	3441	3663	3866	4057	4240	4418
14	1	0008	0039	0081	0124	0170	0219	0271	0326	0385	0449
14	2	0118	0281	0417	0532	0639	0741	0841	0942	1044	1149
14	3	0358	0660	0880	1053	1204	1345	1479	1610	1740	1871
14	4	0695	1127	1416	1635	1822	1991	2150	2303	2453	2602
14	5	1108	1657	2005	2261	2476	2668	2845	3014	3178	3340
14	6	1588	2240	2637	2923	3160	3368	3559	3740	3913	4082
14	7	2129	2870	3309	3618	3870	4090	4290	4477	4656	4829

TABLE A.2 (*continued*)

$c+d$	c	p: .50	.55	.60	.65	.70	.75	.80	.85	.90	.95	.99
2	1	5000	5500	6000	6500	7000	7500	8000	8500	9000	9500	9900
3	1	2929	3292	3675	4084	4523	5000	5528	6127	6838	7764	9000
4	1	2063	2337	2632	2953	3306	3700	4152	4687	5358	6316	7846
4	2	5000	5334	5671	6014	6367	6736	7129	7556	8042	8646	9411
5	1	1591	1810	2047	2308	2599	2929	3313	3777	4377	5271	6838
5	2	3857	4147	4445	4756	5084	5437	5825	6265	6795	7514	8591
6	1	1294	1476	1674	1894	2140	2421	2752	3157	3690	4507	6019
6	2	3138	3389	3650	3925	4220	4542	4902	5321	5839	6574	7779
6	3	5000	5267	5537	5814	6102	6406	6734	7101	7534	8107	8944
7	1	1091	1246	1416	1605	1818	2063	2353	2711	3187	3930	5358
7	2	2644	2864	3094	3339	3604	3895	4224	4613	5103	5818	7057
7	3	4214	4458	4708	4967	5239	5532	5854	6222	6668	7287	8269
8	1	0943	1078	1227	1393	1580	1797	2054	2374	2803	3482	4821
8	2	2285	2480	2685	2905	3143	3407	3709	4067	4526	5207	6434
8	3	3641	3863	4092	4331	4586	4861	5168	5523	5962	6587	7637
8	4	5000	5229	5461	5699	5948	6212	6499	6824	7214	7747	8577
9	1	0830	0950	1082	1230	1397	1591	1822	2111	2501	3123	4377
9	2	2011	2186	2371	2570	2786	3027	3304	3635	4062	4707	5899
9	3	3205	3408	3618	3839	4075	4332	4621	4959	5382	5997	7068
9	4	4402	4616	4835	5061	5299	5555	5837	6159	6554	7108	8018
10	1	0741	0849	0968	1101	1252	1428	1637	1901	2257	2831	4005
10	2	1796	1955	2123	2304	2501	2723	2978	3285	3684	4291	5440
10	3	2862	3048	3242	3446	3665	3905	4177	4496	4901	5496	6563
10	4	3931	4131	4336	4550	4776	5020	5291	5605	5994	6551	7500
10	5	5000	5204	5410	5622	5844	6080	6339	6633	6990	7486	8290
11	1	0670	0767	0876	0997	1134	1294	1487	1728	2057	2589	3690
11	2	1623	1767	1921	2087	2269	2474	2710	2996	3368	3942	5044
11	3	2586	2757	2936	3126	3330	3554	3809	4111	4496	5069	6117
11	4	3551	3738	3930	4131	4345	4577	4837	5139	5517	6066	7029
11	5	4517	4710	4907	5111	5325	5555	5809	6100	6458	6965	7817
12	1	0611	0700	0799	0910	1037	1184	1361	1584	1889	2384	3421
12	2	1480	1613	1755	1908	2077	2266	2486	2753	3102	3644	4698
12	3	2358	2517	2683	2860	3050	3261	3501	3786	4152	4701	5723
12	4	3238	3413	3593	3782	3984	4205	4452	4742	5108	5644	6604
12	5	4119	4302	4489	4684	4889	5111	5357	5642	5995	6502	7378
12	6	5000	5185	5373	5566	5768	5984	6221	6492	6823	7288	8060
13	1	0561	0644	0735	0838	0955	1091	1255	1462	1746	2209	3187
13	2	1360	1483	1615	1757	1914	2091	2296	2546	2875	3387	4395
13	3	2167	2315	2470	2635	2814	3012	3238	3508	3855	4381	5373
13	4	2976	3140	3309	3488	3679	3888	4124	4401	4753	5273	6222
13	5	3785	3958	4136	4322	4518	4731	4968	5245	5590	6091	6976
13	6	4595	4772	4953	5140	5336	5547	5779	6047	6377	6848	7651
14	1	0519	0596	0681	0776	0885	1011	1164	1358	1623	2058	2983
14	2	1258	1373	1496	1629	1775	1941	2133	2368	2678	3163	4128
14	3	2004	2143	2288	2443	2611	2798	3011	3267	3598	4101	5062
14	4	2753	2907	3067	3235	3416	3615	3839	4105	4443	4946	5878
14	5	3502	3666	3835	4011	4199	4403	4631	4899	5234	5726	6609
14	6	4251	4420	4593	4773	4963	5167	5394	5657	5982	6452	7271
14	7	5000	5171	5344	5523	5710	5910	6130	6382	6691	7130	7871

TABLE A.2 (*continued*)

$c+d$	c	p: .01	.05	.10	.15	.20	.25	.30	.35	.40	.45
15	1	0007	0037	0075	0115	0158	0203	0252	0303	0358	0418
15	2	0110	0260	0387	0494	0593	0688	0781	0875	0970	1068
15	3	0331	0611	0815	0975	1117	1248	1373	1495	1617	1739
15	4	0640	1040	1309	1513	1688	1846	1995	2138	2279	2419
15	5	1019	1527	1851	2090	2291	2471	2638	2797	2952	3104
15	6	1457	2061	2432	2699	2921	3117	3298	3468	3633	3794
15	7	1947	2636	3046	3336	3574	3782	3972	4151	4321	4487
16	1	0007	0034	0070	0108	0148	0190	0235	0283	0335	0391
16	2	0102	0242	0360	0461	0553	0642	0729	0817	0906	0998
16	3	0307	0568	0759	0909	1041	1163	1281	1395	1510	1625
16	4	0594	0967	1218	1408	1572	1720	1860	1995	2128	2260
16	5	0944	1417	1720	1944	2132	2301	2459	2609	2755	2900
16	6	1346	1909	2256	2507	2716	2902	3072	3234	3390	3544
16	7	1795	2437	2822	3096	3321	3518	3699	3869	4032	4191
16	8	2287	3000	3415	3707	3944	4150	4337	4512	4679	4840
17	1	0006	0032	0066	0101	0138	0178	0220	0266	0314	0367
17	2	0095	0227	0337	0431	0518	0602	0684	0766	0850	0937
17	3	0287	0531	0710	0850	0975	1090	1200	1308	1416	1525
17	4	0554	0903	1138	1317	1471	1611	1743	1870	1995	2121
17	5	0878	1321	1606	1816	1994	2154	2303	2445	2583	2721
17	6	1251	1778	2104	2341	2539	2714	2876	3030	3178	3324
17	7	1665	2267	2629	2888	3101	3289	3461	3623	3779	3931
17	8	2117	2786	3178	3455	3680	3877	4056	4224	4384	4540
18	1	0006	0030	0062	0095	0130	0168	0208	0250	0296	0346
18	2	0090	0213	0317	0406	0488	0566	0644	0722	0801	0883
18	3	0269	0499	0667	0799	0916	1025	1129	1232	1333	1436
18	4	0519	0846	1068	1237	1382	1514	1639	1760	1879	1997
18	5	0822	1238	1506	1705	1873	2024	2165	2300	2432	2562
18	6	1168	1664	1972	2196	2383	2549	2703	2849	2991	3131
18	7	1552	2119	2461	2707	2909	3088	3252	3406	3555	3702
18	8	1971	2601	2973	3235	3450	3638	3809	3970	4124	4275
18	9	2422	3108	3504	3780	4004	4199	4376	4541	4698	4850
19	1	0006	0028	0058	0090	0123	0159	0196	0236	0280	0327
19	2	0085	0201	0299	0383	0460	0535	0608	0682	0757	0834
19	3	0254	0470	0629	0754	0865	0968	1066	1163	1260	1358
19	4	0488	0797	1006	1166	1304	1429	1547	1662	1775	1888
19	5	0772	1164	1418	1606	1765	1909	2043	2171	2297	2422
19	6	1096	1563	1855	2067	2245	2404	2550	2690	2825	2958
19	7	1454	1990	2314	2547	2740	2910	3067	3215	3357	3497
19	8	1844	2440	2792	3042	3247	3427	3591	3746	3894	4039
19	9	2263	2912	3288	3552	3767	3954	4124	4283	4434	4582
20	1	0005	0027	0055	0085	0117	0150	0186	0224	0265	0310
20	2	0080	0190	0283	0363	0436	0507	0576	0646	0717	0791
20	3	0240	0445	0595	0714	0819	0916	1010	1102	1194	1287
20	4	0461	0753	0951	1103	1233	1353	1465	1574	1682	1789
20	5	0728	1099	1339	1518	1670	1806	1934	2057	2176	2295
20	6	1032	1475	1751	1953	2123	2274	2414	2547	2676	2804
20	7	1368	1875	2183	2405	2589	2752	2902	3043	3180	3315
20	8	1733	2297	2633	2871	3067	3239	3397	3545	3688	3827
20	9	2124	2739	3098	3351	3556	3736	3900	4053	4199	4342
20	10	2540	3201	3579	3843	4056	4241	4408	4565	4713	4858

TABLE A.2 *(continued)*

c+d	c	p: .50	.55	.60	.65	.70	.75	.80	.85	.90	.95	.99
15	1	0483	0554	0634	0722	0824	0943	1086	1267	1517	1926	2803
15	2	1170	1278	1393	1517	1655	1810	1992	2213	2507	2967	3891
15	3	1865	1995	2131	2277	2436	2612	2814	3057	3372	3854	4783
15	4	2561	2706	2857	3017	3189	3377	3592	3845	4170	4657	5567
15	5	3258	3413	3574	3742	3921	4117	4336	4594	4920	5400	6274
15	6	3954	4116	4282	4455	4637	4835	5055	5311	5631	6096	6920
15	7	4651	4816	4984	5157	5339	5535	5751	6001	6309	6750	7512
16	1	0452	0518	0593	0676	0771	0883	1017	1188	1423	1810	2644
16	2	1094	1195	1303	1420	1550	1697	1868	2077	2356	2794	3679
16	3	1743	1866	1995	2133	2283	2450	2641	2872	3173	3634	4532
16	4	2394	2531	2675	2826	2989	3169	3373	3616	3928	4398	5285
16	5	3045	3193	3346	3506	3678	3865	4076	4325	4640	5108	5969
16	6	3697	3852	4010	4176	4352	4543	4756	5005	5317	5774	6597
16	7	4348	4507	4669	4836	5013	5204	5415	5660	5965	6404	7177
16	8	5000	5160	5321	5488	5663	5850	6056	6293	6585	7000	7713
17	1	0424	0487	0557	0635	0725	0830	0957	1118	1340	1707	2501
17	2	1027	1122	1224	1335	1457	1596	1758	1957	2222	2640	3488
17	3	1637	1752	1874	2005	2147	2306	2488	2709	2996	3438	4305
17	4	2247	2378	2514	2658	2813	2985	3180	3413	3712	4166	5029
17	5	2859	3000	3145	3299	3463	3642	3845	4085	4389	4844	5690
17	6	3471	3619	3771	3930	4099	4283	4489	4731	5035	5483	6299
17	7	4082	4235	4391	4553	4724	4909	5116	5355	5654	6090	6866
17	8	4694	4849	5006	5168	5339	5522	5726	5960	6250	6666	7393
18	1	0400	0459	0525	0599	0684	0783	0903	1056	1267	1616	2373
18	2	0968	1058	1154	1259	1375	1507	1661	1850	2102	2501	3316
18	3	1542	1652	1768	1892	2027	2178	2352	2562	2837	3262	4099
18	4	2118	2242	2371	2509	2657	2821	3008	3231	3519	3956	4796
18	5	2694	2828	2968	3114	3271	3444	3639	3869	4164	4605	5434
18	6	3270	3412	3558	3711	3874	4051	4251	4485	4781	5219	6025
18	7	3847	3994	4144	4300	4466	4646	4846	5080	5374	5803	6577
18	8	4423	4573	4725	4883	5049	5229	5428	5658	5945	6360	7094
18	9	5000	5150	5302	5459	5624	5801	5996	6220	6496	6892	7578
19	1	0378	0434	0496	0567	0647	0741	0855	1000	1201	1533	2257
19	2	0915	1001	1092	1192	1302	1427	1574	1754	1995	2377	3160
19	3	1458	1563	1673	1791	1920	2064	2230	2431	2694	3103	3912
19	4	2002	2121	2244	2375	2517	2674	2853	3067	3344	3767	4583
19	5	2547	2676	2809	2949	3100	3265	3453	3676	3960	4389	5199
19	6	3092	3228	3368	3515	3672	3843	4036	4263	4550	4978	5772
19	7	3637	3778	3923	4074	4235	4409	4604	4832	5118	5540	6309
19	8	4182	4327	4474	4628	4789	4964	5159	5385	5667	6078	6814
19	9	4727	4873	5022	5175	5336	5510	5702	5923	6198	6594	7290
20	1	0358	0412	0471	0538	0614	0704	0812	0950	1141	1459	2152
20	2	0868	0949	1036	1131	1236	1355	1495	1668	1898	2264	3018
20	3	1383	1482	1587	1700	1823	1961	2120	2312	2565	2958	3741
20	4	1899	2012	2130	2255	2391	2541	2713	2919	3186	3594	4387
20	5	2415	2538	2666	2801	2945	3105	3285	3500	3775	4191	4983
20	6	2932	3063	3197	3339	3490	3655	3841	4061	4340	4758	5538
20	7	3449	3585	3725	3871	4026	4195	4384	4606	4886	5300	6060
20	8	3966	4106	4249	4397	4554	4725	4915	5136	5413	5819	6553
20	9	4483	4625	4769	4919	5076	5246	5435	5653	5925	6319	7020
20	10	5000	5142	5287	5435	5592	5759	5944	6157	6421	6799	7460

Notes: (1) It is to be understood that a decimal point precedes each four-digit entry. (2) Let $f_{0 \cdot p}(c, c+d)$ denote the pth fractile of the standard beta distribution with parameters c and d. To find fractiles when $c \geqq d$, use the relation $f_{0 \cdot p}(c, c+d) = 1 - f_{0 \cdot (1-p)}(d, c+d)$

AUTHOR INDEX

SUBJECT INDEX