ESTIMATION OF CROP YIELD VARIATION AND ITS USE IN FARM FINANCIAL PLANNING

By

MARK E. ANIBAL

AN ABSTRACT

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Agricultural Economics

1989

ABSTRACT

ESTIMATION OF CROP YIELD VARIATION AND ITS USE IN FARM FINANCIAL PLANNING

By

Mark E. Anibal

Actual farm records were used to evaluate variation of crop yields for corn, wheat, and soybeans grown in Michigan. The assumption of yields being normally distributed with independent and constant variance was tested for all farms, farms characterized by soil groupings, and individual farms of a selected soil group.

Sufficient evidence was not found to reject the assumption that yield variation is normally distributed with independent and constant variance when all farms were grouped together. Grouping farms by soil potential showed strong support for yields being independent and normally distributed.

The assumption of constant variance had strong support for wheat, but was weak for corn. Analysis of individual farms for a selected soil group, suggested that farms having data sets with significant amounts of negative kurtosis or negative outliers could be better modeled by alternative distributions.

Estimates for yield variation parameters were incorporated into planning tools used for predictive analysis.

ESTIMATION OF CROP YIELD VARIATION AND ITS USE IN FARM FINANCIAL PLANNING

By

MARK E. ANIBAL

PLAN B PAPER

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Agricultural Economics

1989

ACKNOWLEDGEMENTS

I wish to thank Dr. Gerald Schwab for serving as my major professor and research supervisor. Without his guidance and supportive input this thesis would not have been completed. Also, special thanks goes to Dr. J. Roy Black for the major part of guiding this research and his help in applications to farm financial planning.

I would also like to thank the members of my committee: Dr. Sherrill Nott, for his input with farm financial planning; and Dr. Oran Hesterman, for his help in assessing soil groupings.

Also, thanks goes to Chris Wolf for computer support to this project and also to Linda Peters for doing the word processing to make this final draft a reality.

Lastly, special thanks to my wife, Marilyn, for her patience and continual encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Managing a modern farm business in a complex and risky world is a challenging task. The farm manager is faced with many uncertain events in planning for the future. Ever-changing conditions in the decision environment pose significant concern for the decision-maker. Risk and uncertainty is not a new phenomenon for farmers who have been taking risks for years. Some of these risks and uncertainties include: 1) timing of rainfall for maximum yield potential; 2) timely equipment purchases; 3) adequate prices for products sold; 4) machine downtime at critical periods; and 5) government regulations. Farm managers must make decisions with imperfect knowledge as the future is unpredictable. Specific decisions can have a number of possible results depending on many factors which are beyond the farm managers control. The level of knowledge, however, can vary considerably from complete uncertainty to fairly reliable predictions. It is then beneficial for the farm manager to incorporate the consideration of risk in the decision-making process.

1.1 Risk Management In Agriculture

Gaining an understanding of risk and its possible sources is necessary for understanding how risk should be managed at the producer level. The definition of risk used in this research study combines the common usage in decision-making contexts and the definition outlined by Frank Knight.[1] The definition of risk used in decision- making contexts is the chance of adverse outcomes associated with a particular action. Knight defines risk as measurable uncertainty where uncertainty is knowledge of all or partial outcomes of an action without the ability to quantify the likelihood of occurrence for the outcomes. The definition of risk used in this study is the probability of adverse outcomes associated with an action where probability is a measure of uncertainty.

Risk is encountered by agricultural producers from many sources. Nelson, Casler, and Walker outline major sources of agricultural risk that farmers face in operating a business.[2] It is important to note that these sources of risk differ between enterprises and may change over time. These major sources of risk are outlined as follows:

1) Production Risk - This source of risk is due to the variability in production of agricultural products caused by unpredictable factors that affect their production such as weather, disease, pests, genetic variations, and timing of management and cultural practices. Examples of production measures would be crop yields, animals per litter, animal rate of gain, feed conversion rate, death loss, labor hours required, planting date, etc.

2) Price Risk - This source of risk depicts the variability and unpredictability of market prices farmers receive for products and those

---

[1] Frank H. Knight. Risk, Uncertainty, and Profit, Boston:The Riverside Press, Cambridge, 1921, pp.46-48.

[2] A. Gene Nelson, George L. Casler, and Odell L. Walker. Making Farm Decisions In A Risky World: A Guidebook, (Oregon State University, July 1978), pp.1-3 - 1-4.

paid for production inputs. Random price variations result from random supply and demand interaction which is influenced by buyer - seller expectations, speculation, government programs, and consumer demand.

3) Business risk - This source of risk relates to the financial structure of the business, assets controlled, and debt obligations to creditors. This type of risk has become more important as larger capital investments are required in present day agriculture with a higher proportion of the financing coming from borrowed capital. Variable cash flows increase the risk of meeting debt payments and other financial obligations.

4) Technology Risk - The development of new technology can make current production methods obsolete. Timing the adoption of new technologies is a risk for farmers. An example might be purchasing the machinery for ridge - planting vs conventional planting practices.

5) Casualty Loss Risk - This refers to the loss of assets to fire, wind, hail, flood, and theft which is a traditional source of risk. Inflation has greatly increased the importance of this risk as the value of potential losses can increase yearly.

6) Legal Risk - Governmental laws and regulations are a major source of uncertainty for agricultural producers stemming from changing social attitudes. Typical examples are environmental protection; controls on feed additives, insecticides, and herbicides; and land use planning. There is also risk of law suits from liabilities due to farm accidents.

7) Human Risk - This refers to the unpredictability of the character, health, and behavior of individuals. The disabling of the farm manager can disrupt the continuity of an efficient farm operation

as one example of risk. The possibility of losing employees during critical production periods and the dishonesty or undependability of business associates are other examples of human risks.

Risk management considers the impact of all these sources of risk in managing the farm business. The focus of this research study will be concerned specifically with production risk.

Management of production risk becomes of importance as monetary returns to the business are directly related to this source of risk. Monetary returns are partial reward for taking the chance that actual return might be below the expected return. In other words producers must take the risk of earning less (income less than the expected return) to have a possibility of earning more. Risk then becomes synonymous with a negative variation in income and gives a relative measure to judge between risky options. This principal is illustrated in Figure 1.1. Option A in Figure 1.1 is less risky than Option B as the variance in income is less than Option B. Conversely, Option B has a higher income on average than Option A. The decision-maker must then choose a desired level of expected income considering the variability in expected income. Risk management is then concerned with maximizing the returns to risk at the minimum acceptable level of risk (income variance) exposure.

1.2 Focus Of Research Problem

The main objective for this research is to quantify the production risk for corn, wheat, and soybeans grown in Michigan. This will be accomplished by evaluating the continuous probability distribution that appropriately explains the variation in yield over time for the three different crops.

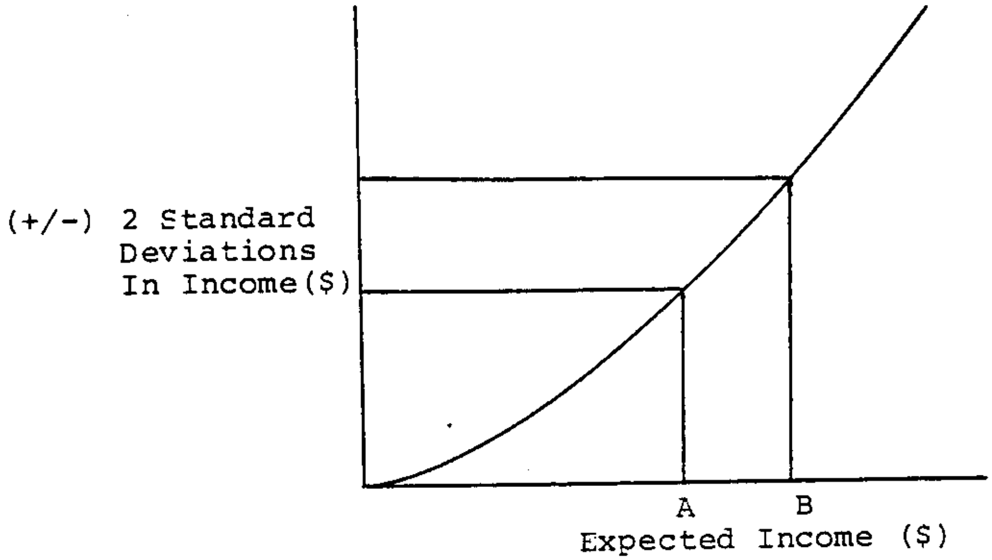


Figure 1.1.
Variation With Increase In Expected Income

Another objective of this research study is to develop the methodology for estimating continuous probability distributions and evaluating their effectiveness in modeling the yield variation.

It is also of interest to estimate the parameters of the distribution and to test their validity for the state and by soil groups within the state. Functional form of the continuous probability distribution will also be evaluated by soil groups and a detailed analysis will be given for a selected soil group.

The last concern of this study is the incorporation of continuous probability distributions into decision-making models that will be useful for risk management.

CHAPTER 2

HYPOTHESIZED DISTRIBUTIONS

Risk was previously defined as the probability of adverse outcomes associated with an action. The discussion up to this time has only considered risk in terms of variation in outcomes. Risk, as defined, considers not only the variation in outcomes but also the probability or likelihood of their occurrence. In evaluating risk, consideration must be given both to the distribution of probabilities of possible outcomes and the variation of outcomes. It is necessary to understand the concept of a probability distribution so that the nature of risk can be understood.

2.1 Probability Distributions

The probability distribution is a theoretical model that assigns probabilities to the possible values of a random variable. The distribution describes the nature of outcomes for a random variable where a random variable is a variable defined as X taking discrete or continuous values.

2.1.1 Discrete

The throwing of a die is one simplistic example. Each die has six sides with a number from one to six on each side. Table 2.1 shows the associated probability of occurrence for all possible outcomes.

Table 2.1
Probability Distribution Of A Fair Die

Discrete Values Of X	1	2	3	4	5	6
Probability Of X	1/6	1/6	1/6	1/6	1/6	1/6

This table represents the probability distribution for the use of a die. This table can be illustrated graphically as in Figure 2.1.

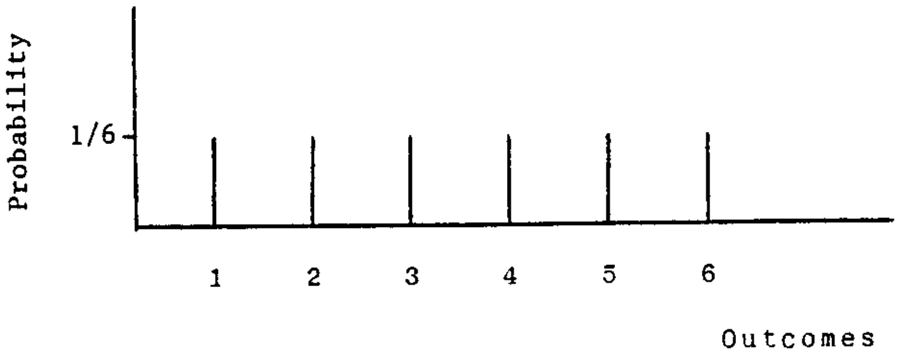


Figure 2.1
Line Diagram For Distribution Given In Table 1.1

The previous example illustrates a discrete random variable. The probability distribution $f(x_i)$ of a discrete random variable is a list of distinct values x_i of X together with the associated probabilities $f(x_i) = P[X = x_i]$ for i = 1 ... k (see Table 2.2).

Table 2.2
Probability Distribution Of A Discrete Random Variable

Distinct Values Of X	x_1	x_2	...	x_k
Probability	$f(x_1)$	$f(x_2)$	...	$f(x_k)$

The capital letter X denotes a random variable and the lower case letter x represents a value of the random variable. The $f(x_i)$ denotes the probability of the value x_i of X where the numbers represented by $f(x_i)$ are all between 0 and 1 and $\Sigma f(x_i)=1$.

As data sets can be described by measures of center (mean = $\bar{x}$) and spread (variance = s^2), probability distributions can be described with similar measures as probabilities may be viewed as long-range relative frequencies. If a large series of values for a random variable X are observed, the relative frequencies approach the probabilities as the number of observations increase. Therefore, the mean of a random variable X or the corresponding probability distribution of X can be denoted as:

$$\text{Mean of } X = \Sigma \text{ (Value times Probability)}$$

The mean of the random variable X is more accurately defined as the expected value or expectation of X because each value of X is weighted by its corresponding probability. The formal definition of the expected value is:

$$E(X) = \Sigma[x_i \cdot f(x_i)]$$

where E(X) represents the mean of the random variable X. The products $x_1 \cdot f(x_1)$ are summed for all distinct values of X to compute the E(X).

The expected value of a probability distribution is a measure of the center for the distribution. The E(X) has a physical interpretation as the center of gravity of a mass function. This represents the point at which the distribution would balance as the distribution has equal weight on both sides of the expected value. This property is illustrated in Figure 2.3.

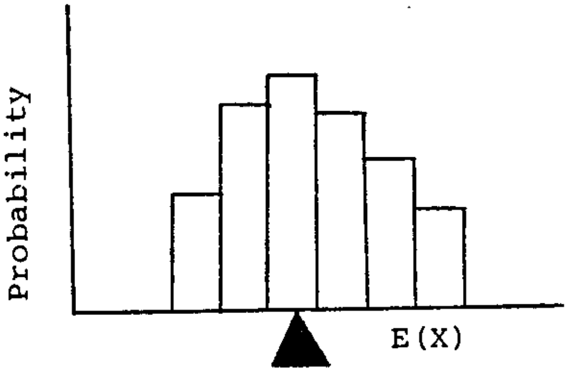


Figure 2.3
E(X) Represents The Center Of Mass

The expected value gives the center of the distribution but a measure of the spread describing the deviations from the mean is also needed. The magnitude of the deviations need only be considered (X - E(X)) in measuring the spread of a distribution. The square of the deviation $(X - E(X))^2$ is used to remove the sign and is called the variance. Variance is formally defined as:

$$Var(X) = E[\ (X - E(X))^2] = E(X^2) - E(X)^2$$

where Var(X) is the variance of a random variable X, $E(X^2) = \Sigma[\ x_i f(x_i)\ ]$ for all distinct values of X, and E(X) is the expected value of X. The variance of X is also denoted by σ^2 or σ_X^2. To express the measure of spread in the same units as X is expressed, the square root of the variance (standard deviation) is used. The standard deviation is defined as:

$$SD(X) = \sqrt{Var(X)}$$

where SD(X) is the standard deviation of the random variable X and Var(X) is the variance of X. Standard deviation is also denoted by σ or σ_X.

2.1.2 Continuous

The other type of probability distribution is of a continuous nature as the random variable can assume all values in an interval. The probability distribution of a continuous random variable is a smoothed form of the relative frequency histogram. Figures 2.4a and 2.4b illustrate the probability distribution as a smoothed form of the relative frequency histogram. The probability distribution represents the manner in which the total probability 1 is distributed over the range of possible values of the random variable X. If a mathematical function is known to describe the behavior of the random variable the probability distribution is known as the probability density function. The density of the distribution between two defined points is the result of the functional relationship. The properties of a probability distribution for a continuous random variable are as follows:

a) The total area under the curve is 1.

b) $P[a< X <b]$ = area under the curve between a and b.

c) $f(x)$ is positive or 0.

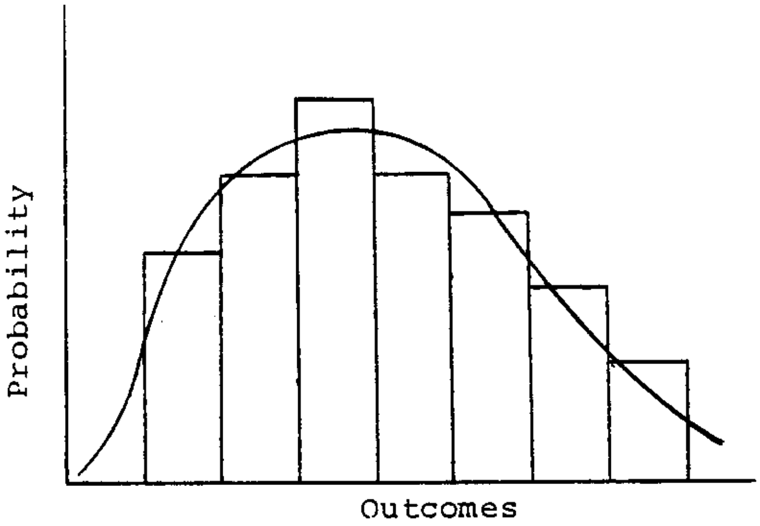


Figure 2.4a
Smoothed Form Of A Probability Distribution

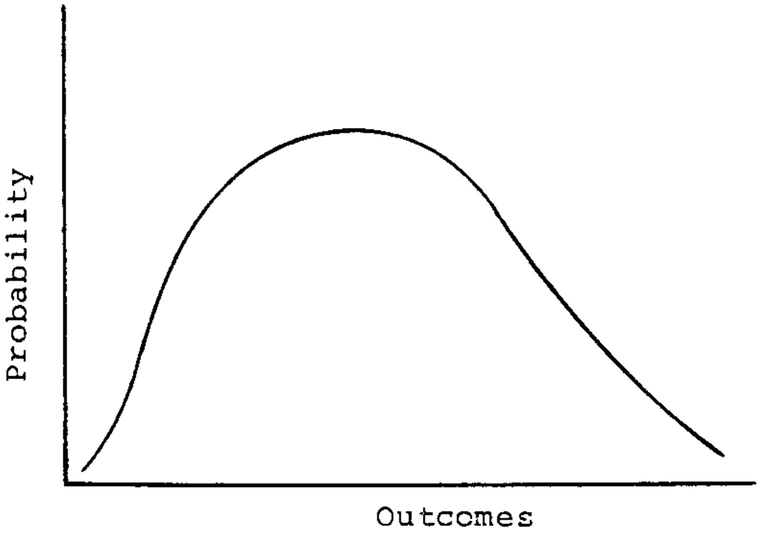


Figure 2.4b
Probability Distribution For Histogram In Figure 2.4a

When considering a random variable that is characterized by a continuous distribution, the probability that X = x will always be zero. The probability of X = x is only meaningful for an interval which lies around x (see Figure 2.5).

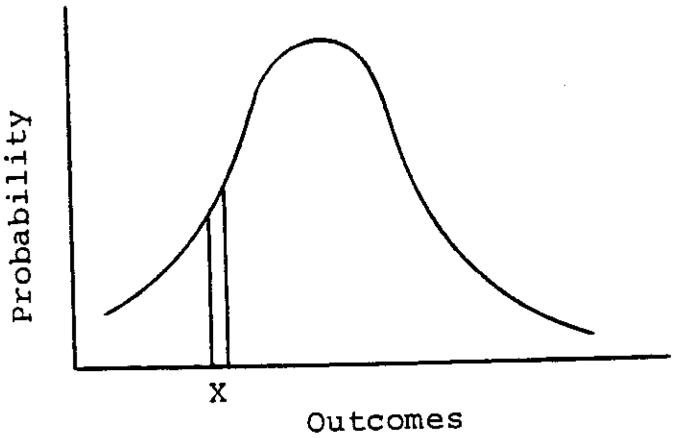
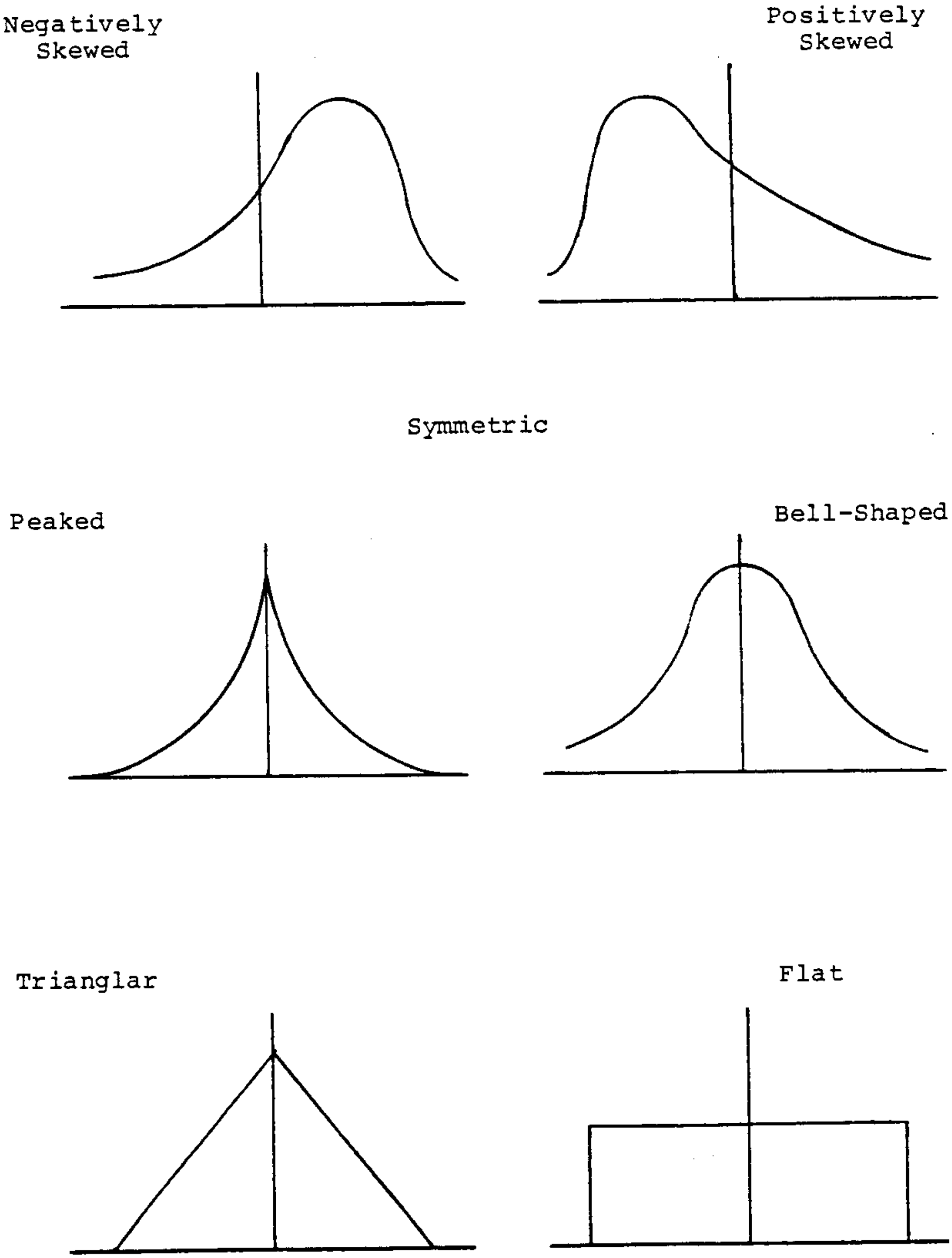


Figure 2.5
Representation Of Probability By Area

Probability distributions of a continuous random variable can exhibit a wide variety of shapes of which a few are illustrated in Figure 2.6.

The measures of expectation and variance discussed for discrete random variables apply also to continuous random variables. However, an additional measure known as the median is necessary for continuous probability distributions to find the center of area under the curve. The median defines a point having 50 percent of the area to the right and left.

The example given in the previous chapter of increasing variance in income for a higher expected income assumed symmetric probabilities around the expected income. In such cases, deviations in either direction from the mean income have equal probability of occurring. Figure 2.7 illustrates a production risk for two farms growing corn. Both farms have the same mean yield response with different standard deviations in yield.



Negatively Skewed

Positively Skewed

Symmetric

Peaked

Bell-Shaped

Trianglar

Flat

Figure 2.6
Variety Of Continuous Probability Distributions

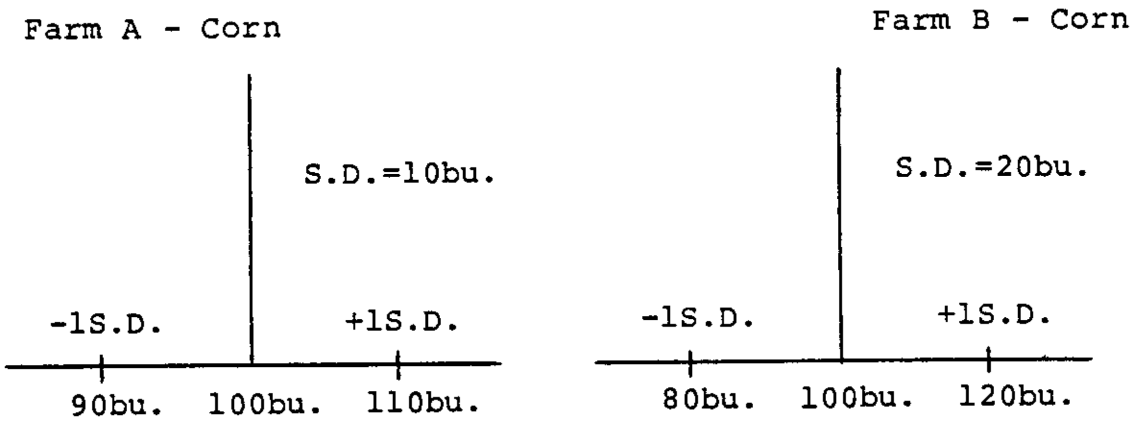


Figure 2.7
Comparison Of Variation For Farms With Same Mean, Different Variance

If we assumed symmetric probabilities, then farm A would exhibit less production risk than farm B. Figure 2.8 introduces non-symmetric hypothetical probability distributions for corn yield outcomes of two farms with the same mean.

Farm A has a probability distribution that is skewed to the right placing more weight or a higher likelihood on yield outcomes below the mean yield value. Farm B, conversely, has a probability distribution that is skewed to the left giving more weight to outcomes above the mean yield response.

The additional information provided by the probability distributions drastically changes the evaluation of risk. Farm A has a greater probability of yield responses below the mean, but the deviations are smaller than the distribution of yields for farm B. Farm B has a higher probability of yield responses above the mean but the deviations in yields from the mean are larger. An intuitive answer on the preferred situation is not obvious from looking at the possibilities.

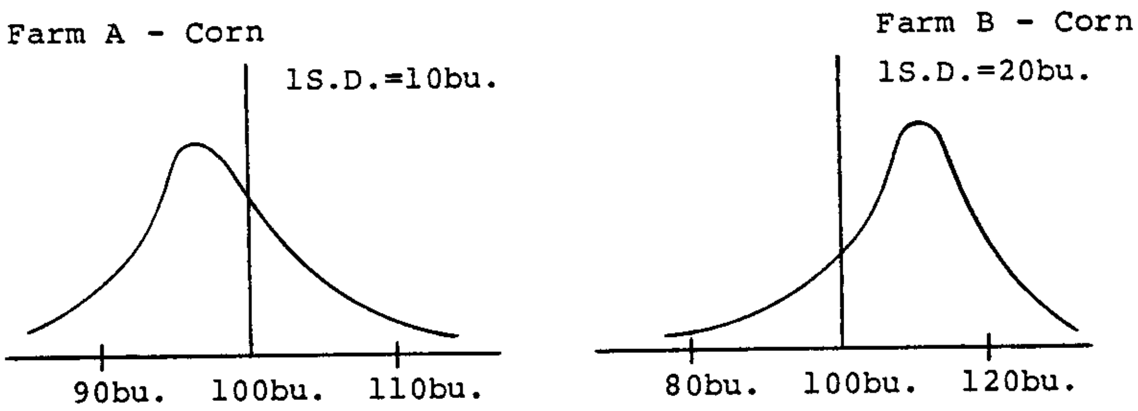


Figure 2.8
Probability Of Variation For Farms With Same Mean, Different Variance

Introduction of different yield means for two hypothetical farms with different distributions complicates the risk analysis even more. Figure 2.9 considers two farms whose mean yields are different and also differ in their standardized deviations in yield response. Farm C has the lower mean value of the two farms, but the standard deviation in mean yield response is smaller. Farm D has both a higher mean yield value and a larger standard deviation. Farm C exhibits less risk because it has a smaller standard deviation in yield response even though farm D has a higher mean value. Considering hypothetical probability distributions for both farms changes the entire impact of the variability. Figure 2.10 shows that farm C (with the lower mean) has a higher probability of values being above the mean. Farm D has a higher likelihood that yield outcomes will be less than the mean.

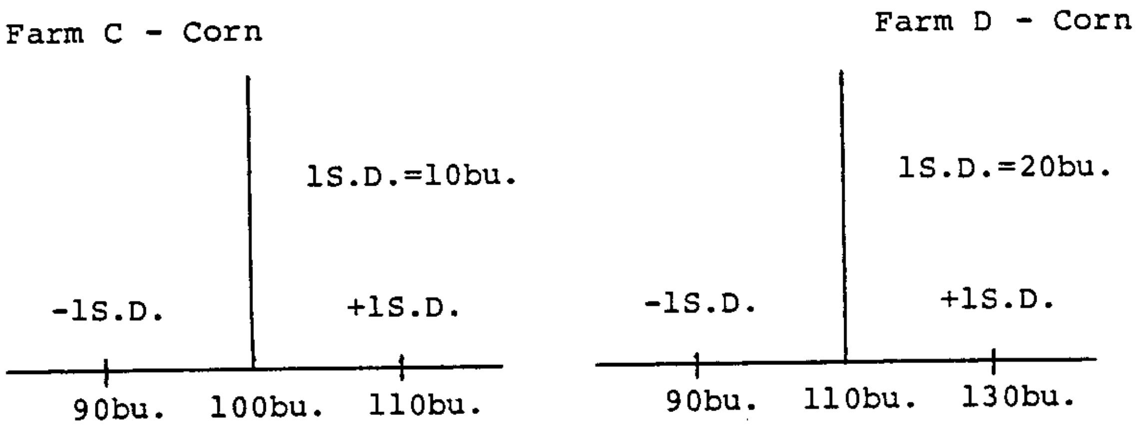


Figure 2.9
Comparison Of Variation For Farms With
Different Means, Different Variance

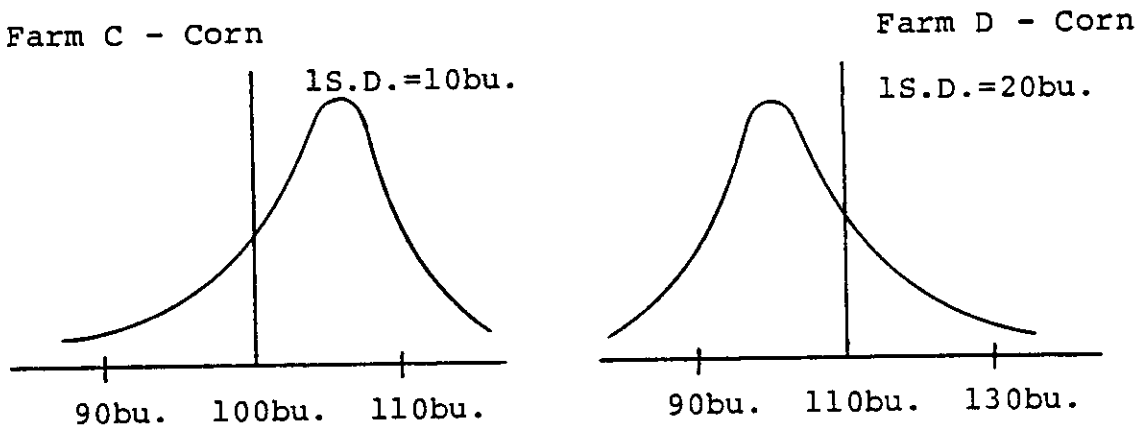


Figure 2.10
Probability Of Variation For Farms With
Different Means, Different Variance

Given the additional information provided by the probability distribution, a broader picture is given to evaluate the nature of the risk. Conclusions are difficult at best with the incorporation of the probability distributions to assist in evaluating risk.

The previous two examples give an appreciation for the contribution that probability distributions make in evaluating risk. Probability distributions describe the behavior of random variables that can assume all values in an interval. These production risks illustrated with corn yield data are examples of random variables.

2.2 Current Approaches to Risk Management.

Extension attempts to teach decision-making under risk and uncertainty has been relatively limited in past years. Publications developed for use by Extension professionals have employed static assumptions and the concept of expected value. Agricultural economists have used expected yields, expected costs, and expected prices as substitutes for certain values in analyzing decision alternatives. Payoff matrixes and decision trees are the core of this approach because they are viewed to be the simple procedures people use in real life decision-making.

The payoff matrix is a simple format that conveniently summarizes the components of the decision problem. Constructing the payoff matrix for particular decision problems consists of three steps:

1) List the alternative actions that are relevant to the problem.
2) List the possible events (states of nature).
3) Budget the payoff for each action/event combination and enter payoffs in the table.

Table 2.3 illustrates the setup of a payoff matrix. The alternative actions for a decision are listed across the top of the matrix. The events corresponding to the decision are listed down the side of the matrix. The elements of the matrix are the appropriate

Table 2.3
Construction Of A Payoff Matrix

	Alternative Actions				
Events	A(1)	A(2)	A(3)	...	A(j)
E(1)	P(1,1)	P(1,2)	P(1,3)	...	P(1,j)
E(2)	P(2,1)	P(2,2)	P(2,3)	...	P(2,J)
.	.	.	.		
.	.	.	.		
.	.	.	.		
E(k)	P(k,1)	P(k,2)	P(k,3)	...	P(k,j)

payoffs. An example is given to demonstrate the use of the payoff matrix. The decision to be made is the level of fertilizer application. The possible actions to take are to fertilize lightly, moderately, and heavily. Rainfall during the growing season will have the greatest effect on the returns to fertilizer application and the possible events are considered to be low, average, and high rainfall. The payoffs are estimated for each event/action combination and are shown in Table 2.4.

The payoff matrix summarizes the decision problem but does not indicate which action should be taken by the decision-maker. Several decision rules have been proposed for use in risky and uncertain situations. If the decision-maker has no knowledge about which event may occur, two approaches may be taken. The first is the pessimistic approach which prefers to avoid uncertainty. The decision-maker first selects the worst outcome for each action, then he/she selects the best

Table 2.4
Payoff Matrix For Fertilizer Application Decision

	Actions - Fertilize		
Events - Rainfall	Lightly	Moderately	Heavily
	-- Net Returns ($) From 400 Acres --		
Low	8,000	5,500	2,000
Average	10,000	12,000	11,000
High	11,000	15,000	18,000

of these worst outcomes and chooses the corresponding action. This is known as the Maximin rule assuring the decision-maker of receiving no less than the return indicated. This criterion applied to the payoff matrix in Table 2.4 would select light fertilization as the action related with the best of the worst outcomes. The second approach is optimistically oriented. The decision-maker in preferring uncertainty chooses the action with the highest payoff. This criterion is called the Maximax rule because it focuses on the possibility of achieving the best possible outcome and ignores the possibility of an event with a poor outcome. Applying the optimistic approach to the payoff matrix in Table 2.4 would select heavy fertilization as an action because it has the best possible outcome.

Many times, however, the decision-maker will have some information about the chances, or probabilities, of the occurrence of various events. There are three different types of probabilities which are based on the manner they are estimated or derived. These three types are empirical, deductive, and subjective. Empirical probabilities are based on the frequencies of empirical observations. Suppose, for example, that a producer has twenty observations available for the amount of spring rainfall. The manager is interested in knowing the probability of receiving more than three inches of rain. Table 2.5 gives the relative frequencies for six possible amounts of rainfall. Using the relative frequencies, the manager could determine that the chance of receiving more than three inches of rainfall would be 12 (6 + 4 + 2) out of 20 years or 12/20 = .6.

Table 2.5
Example Of A Empirical Probability Distribution

Rainfall	Number of Years	Probability
0 - 1	1	1/20
1 - 2	2	2/20
2 - 3	5	5/20
3 - 4	6	6/20
4 - 5	4	4/20
5 - 6	2	2/20
Total	20	1.00

The second type of probability is obtained by deductive reasoning, thus it is called a deductive probability. It is not necessary to use a frequency approach as the outcomes are systematic. Consider a box with 4 balls. One of the balls is white and the other 3 are black. One can deduce that the probability of obtaining the white ball on a random draw from the box would be 1 out of 4. Unfortunately, most of the phenomena considered in farm decision-making are not subject to such logical deduction.

The third type of probability is called subjective probability as it measures the decision-makers strength of conviction about the chance of occurrence for a particular outcome. In estimating these probabilities, we assume that the decision-maker is rational in examining his/her own experience, the data available, and consulting others. The subjective probability school of thought argues that all probabilities are subjective. The perspective claims there is no logical difference between probabilities assigned subjectively and empirical probabilities discussed earlier. In estimating empirical probabilities, certain underlying assumptions exist causing a decision-maker to decide whether the past frequencies reflect the future and whether there are enough observations. Using one's judgement in evaluating the validity of the empirical probabilities makes them subjective in nature.

Once the probabilities have been obtained, either empirically or subjectively, the expected value rule can be used. This decision rule seeks to maximize expected monetary value using the available knowledge of the likelihood of occurrence of the various events. This is exactly the same as the expected value discussed in the previous section, except it is in terms of expected dollars. Table 2.6 shows the payoff matrix for the fertilization decision with probabilities. Using the expected value criterion, the heavy fertilization action would be chosen.

Table 2.6
Fertilizer Example Payoff Matrix With Probabilities Added

Events Rainfall	Probability	Actions - fertilize Lightly	 Moderately	 Heavy
		--Net returns ($) from 400 acres--		
Low	.2	8,000	5,500	2,000
Average	.3	10,000	12,000	11,000
High	.5	11,000	15,000	18,000
Expected Monetary Value		10,100	12,200	12,700

The probabilities represent a discrete random variable for three levels of fertilizer application. The probability distribution for each of the actions are shown graphically in Figures 2.11a, 2.11b, 2.11c. The payoff matrix and graphical representation demonstrate the principle of increasing income variation with the increase of the expected income.

The decision tree uses the same information as the payoff matrix but places it in a structure that clearly portrays the various aspects of the decision problem. Decision trees are particularly helpful for analyzing more complicated problems where a sequence of decisions need to be considered. Figure 2.12 depicts the fertilizer problem in a decision tree format. The square nodes are used to denote decisions where the circular nodes denote the events. When the number of nodes,

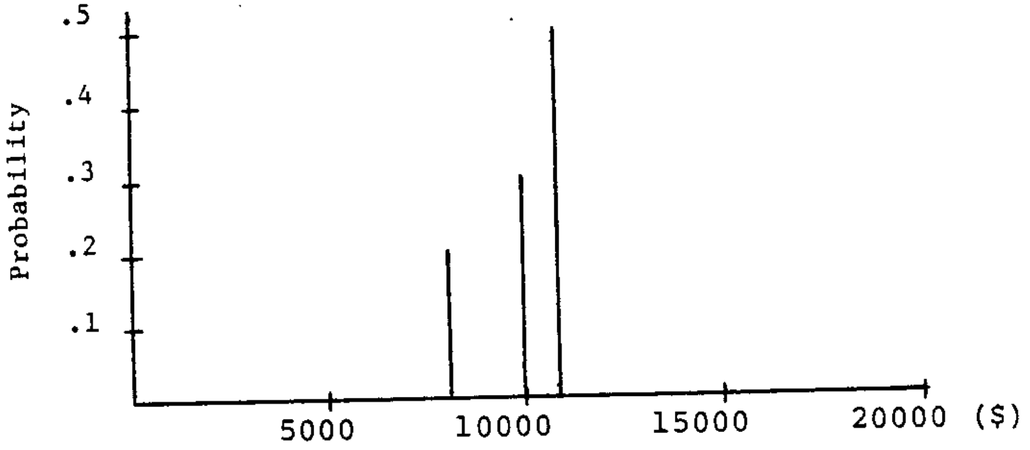


Figure 2.11a
Fertilize Lightly

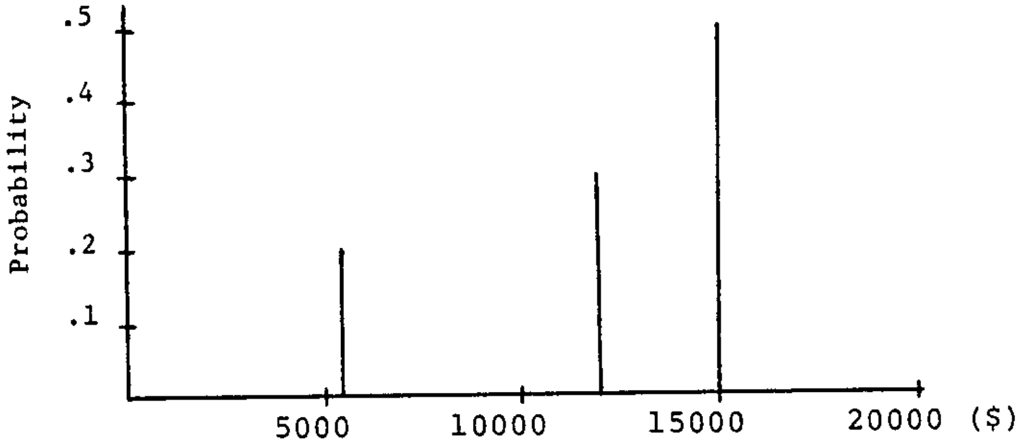


Figure 2.11b
Fertilize Moderately

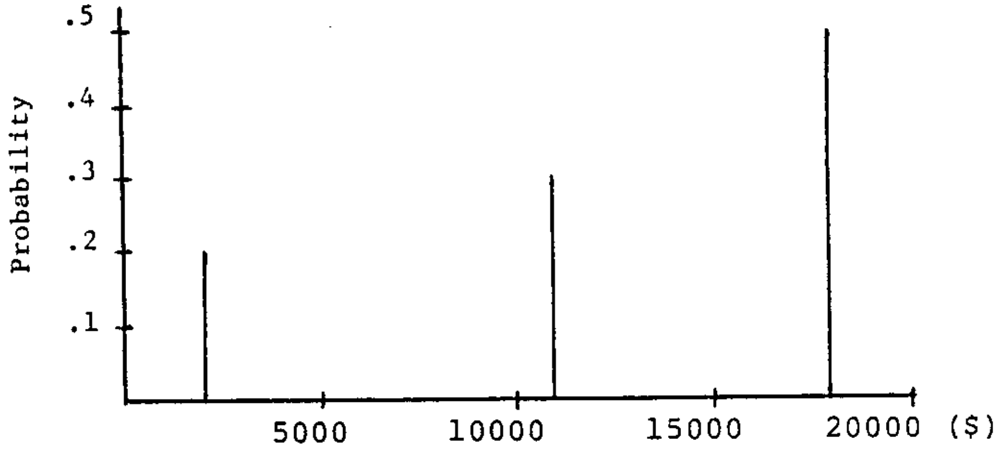


Figure 2.11c
Fertilize Heavily

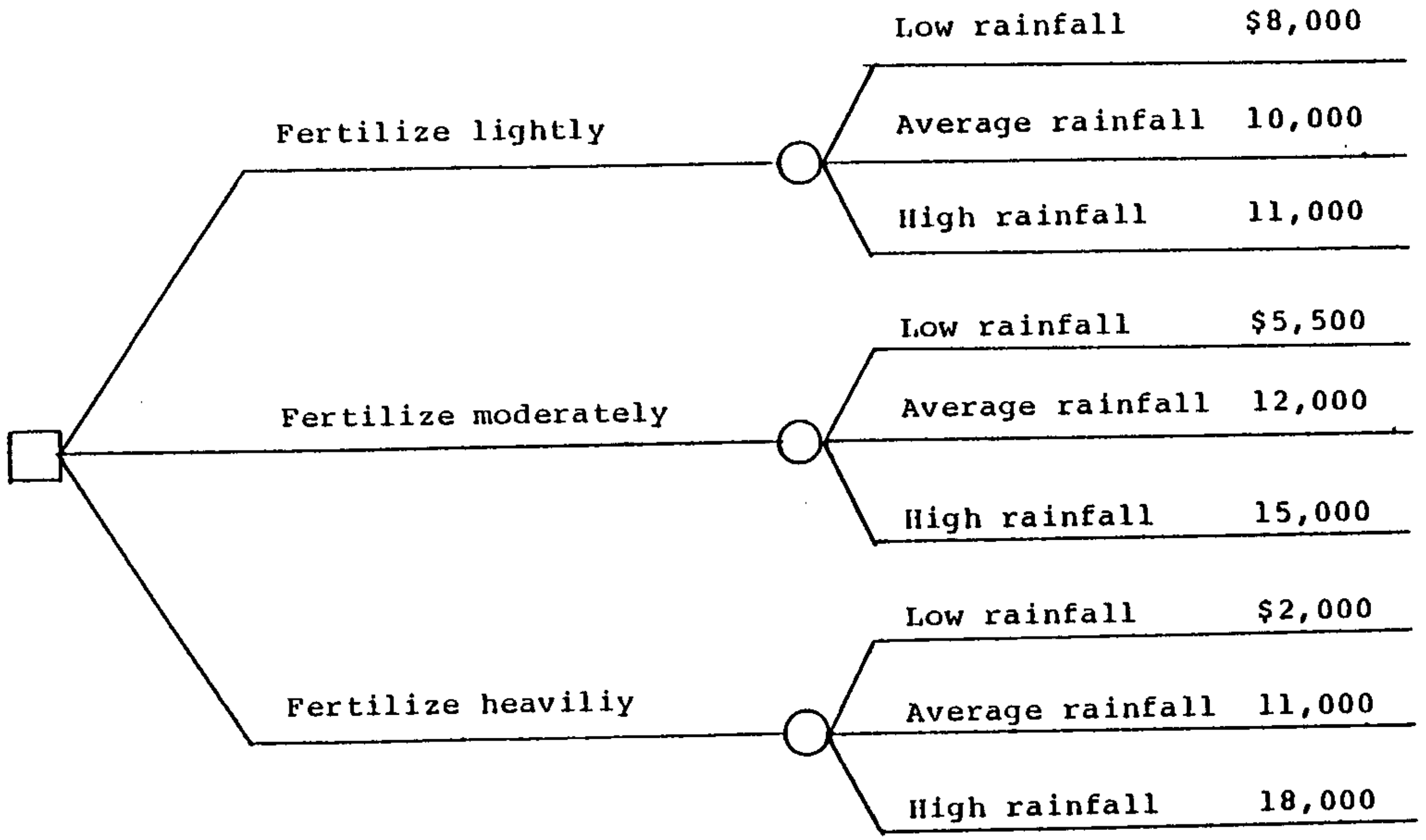


Figure 2.12
Decision Tree For Fertilizer Example

alternative actions, and number of events multiply, the decision tree can become much more complicated. It is best to begin with a coarse tree and to develop branches in further detail.

The use of payoff matrixes and decision trees gives the decision-maker information concerning the variation of income and the likelihood of their outcomes. The nature of the situations most commonly modeled concerning production, price, and business risk, involve continuous random variables rather than discrete random variables. If data were available to measure the variables that affect the events of a particular decision, the decision process would benefit from the use of a continuous random variable to model the variation in outcomes. The current use of discrete random variables in modeling decision-making gives the farm manager a narrow picture of the possible outcomes in the future. The intent of these decision tools, for the sake of simplicity, is to consider only the significant events.

The question that can now be raised is when should other decision methods be used instead of a payoff matrix. The answer to this is ambiguous at best but a few guidelines can be given. If the variables influencing the decision are continuous random variables, one must have data available to estimate the probability distribution and the amount of data gathered should be directly proportional to the value of the proposed decision. The obvious fact of reality is that producers will be very unlikely to inconvenience themselves to gain the additional information provided by a probability distribution. The responsibility of evaluating continuous random variables found in production, price, and business risks should be taken on by agricultural researchers and Extension professionals. Estimation of probability distributions and parameters for these continuous random variables could benefit both agricultural research and the producer. There is a growing gap between the economics literature on risk response and farm management techniques needed by producers as seen by Robert Jolly.[3] He feels there is a need for yield distributions for the application of stochastic efficiency. John Antle sees the goals of research to improve farm management should be to develop the means of measuring and evaluating the properties of price and output distributions faced by farmers so they can be incorporated into their management decisions.[4] The analysis of crop insurance can also be benefited by research considering the variation in crop yields as outlined by Rob King.[5]

2.3 Previous Studies

Review of the literature indicates that very limited research has been conducted to estimate probability distributions for crop yields. Availability of data is considered to be the main factor for the absence of this type of research. One notable exception has been research conducted by Richard Day.[6] Day analyzed cotton, corn, and oat yields taken from the Delta Branch of the Mississippi State Experiment Station

---

[3] Robert W. Jolly. "Risk Management in Agriculture Production," American Journal of Agricultural Economics, 65(December 1983):1107-1113.

[4] John M. Antle. "Incorporating Risk in Production Analysis," American Journal of Agricultural Economics, 65 (December 1983):1099-1106.

[5] Rob King. "Crop Insurance Research Needs," paper presented at seminar sponsored by Southern Regional Research Project S-180, San Antonio, TX, March 28, 1983.

[6] Richard H. Day. "Probability Distributions of Field Crop Yields," Journal Farm Economics, 47 (1965):713-741.

in order to estimate probability distributions for these crops. Time series data was available from 1921-1957 for cotton and corn and 1928-1957 for oats. The hypothesis Day tested concerned the probability distributions having evidence of positive skewness as he predicted less than average yields were more likely than greater than average yields. The findings reported by Day were extreme departures from normality, and dependence of skewness on the rate of nitrogen application.

2.4 Hypothesis Of Distribution Functional Form

This research study considers only the form of the probability distribution and the parameters (mean, variance) that describe the distribution. Although many factors determine the nature of crop yields, no attempt is made to quantify their impact as the appropriate data were not available. This study seeks only to describe the behavior of crop yield probability distributions based on empirical evaluation of past yield observations.

The underlying hypothesis of this study is that crop yield distributions are independent and normally distributed with constant variance, therefore, the crop yield continuous probability distributions are normally distributed. Should this hypothesis not hold, an alternative hypothesis is that crop yields could be described by a negatively skewed distribution as pictured in Figure 2.13. It seems plausible that large negative deviations from the mean would have a greater likelihood than large positive deviations above the mean. The normal and non-normal distributions that will be considered can be described in three classes; non-negative continuous distributions, unbounded continuous distributions, and bounded continuous distributions.

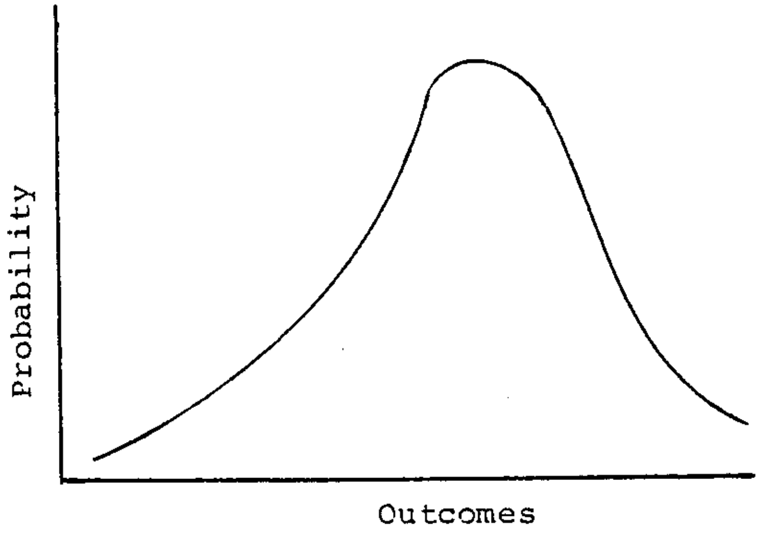


Figure 2.13
Negatively Skewed Distribution

2.4.1 Non-Negative Continuous Distributions

These types of distributions are used to model continuous random variables defined for $0 < X < \infty$. The shape of a typical non-negative distribution is shown in Figure 2.14.

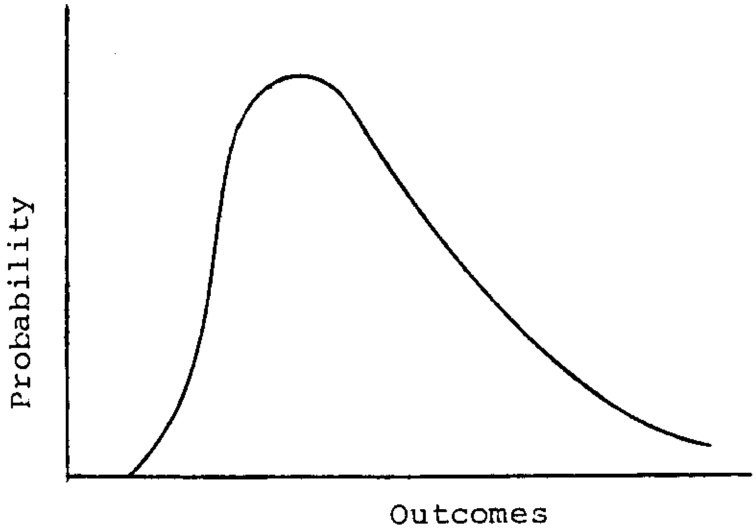


Figure 2.14
Typical Non-Negative Distribution

The form of the non-negative distributions being considered in this study can be depicted by location (γ), scale (β), and shape (α) parameters. The location (γ) parameter defines a location point on the horizontal or x-axis for a range of values described by the distribution (usually the midpoint or lower endpoint of the the range). Scale (β) parameters determine the unit of measurement for the range of values in the distribution. The basic form of the distribution is determined by the shape (α) parameter and is distinctly different from location and scale parameters.

All the non-negative continuous distributions used in this study will model positive skewness as Figure 2.14 indicates. The non-negative continuous distributions considered by this research study are the Weibull, Lognormal, and Inverse gaussian and are shown in Figures 2.15, 2.16, 2.17, respectively. All three distributions have versatile shape properties and are well suited for modeling positively skewed distributions.

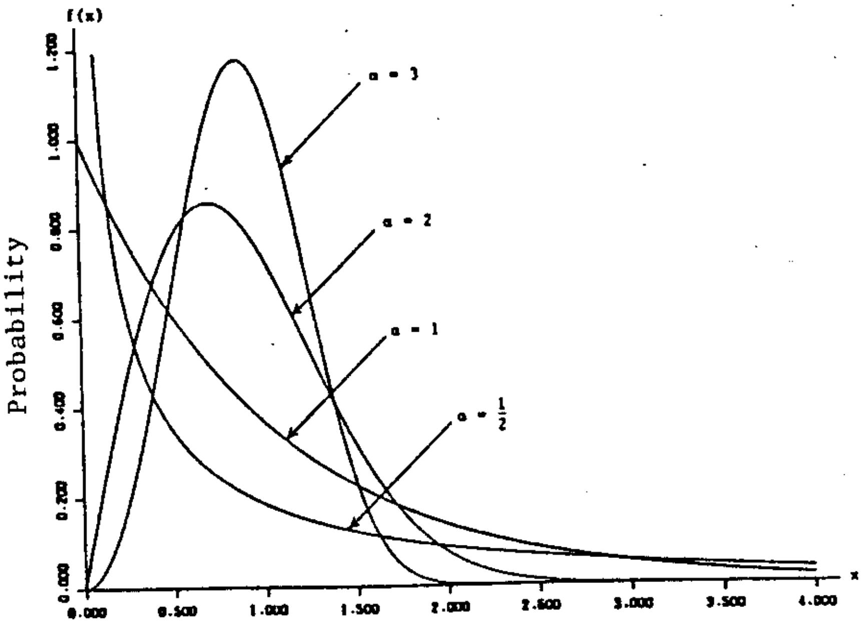


Figure 2.15
Weibull (0, 1, α) Density Function

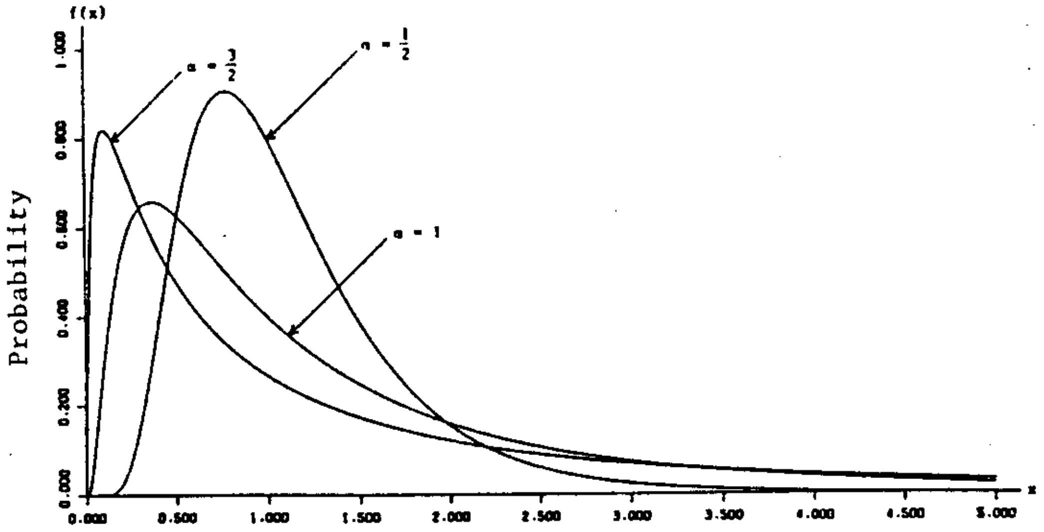


Figure 2.16
Lognormal (o, o, α) Density Function

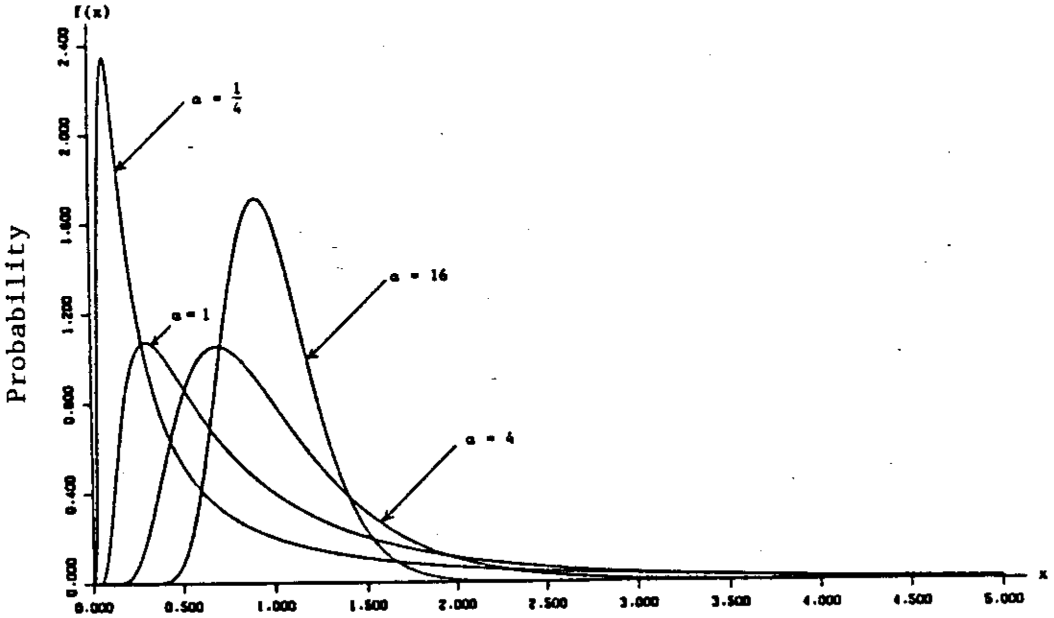


Figure 2.17
Inverse Gaussian (o, 1, α) Density Function

2.4.2 Unbounded Continuous Distributions

These distributions are used to model random variables which are defined for $\infty < X < \infty$. This raises a problem as crop yields can not be negative. In order to gain the benefit of using these distributions, the distributions are used with the understanding that the random variable is defined for $0 < X < \infty$. The probability distribution for a typical unbounded continuous distribution is shown in Figure 2.18. Each unbounded distribution has location and scale parameters. The specific unbounded continuous distributions used in this study are the normal, logistic, extreme value type A, and extreme value type B, and are shown in Figures 2.19, 2.20, 2.21, 2.22, respectively.

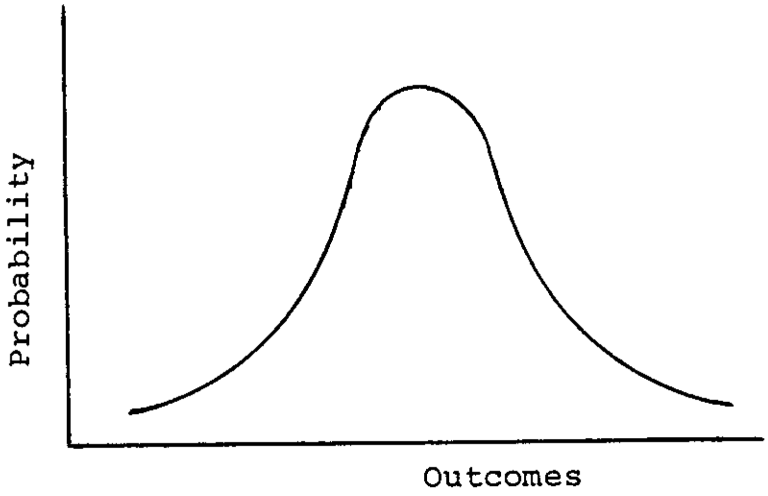


Figure 2.18
Typical Unbounded Distribution

The normal distribution is believed to be the underlying probability distribution as it is symmetric around the expected value. The logistic probability distribution is symmetric around the expected value, but the tails of the distribution are larger than the normal. The logistic function is more peaked than the normal and the corresponding area is shifted to the tails. The extreme value type A distribution models a symmetric (normal, logistic) distribution which has been skewed left due to an outlying negative data value. This distribution is in keeping with the alternative hypothesis of negative skewness. Extreme value type B distributions model symmetric functions which have been skewed right from outlying positive data values.

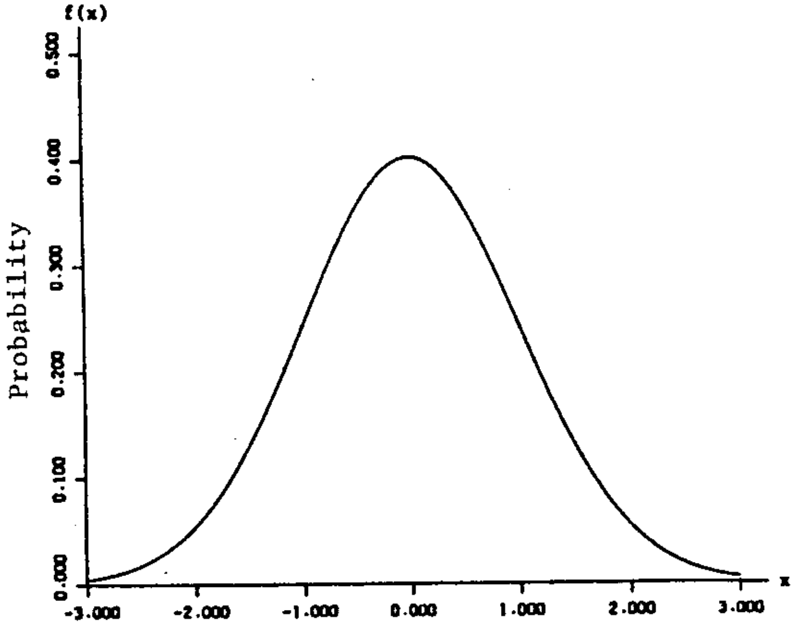


Figure 2.19
Normal (0, 1) Density Function

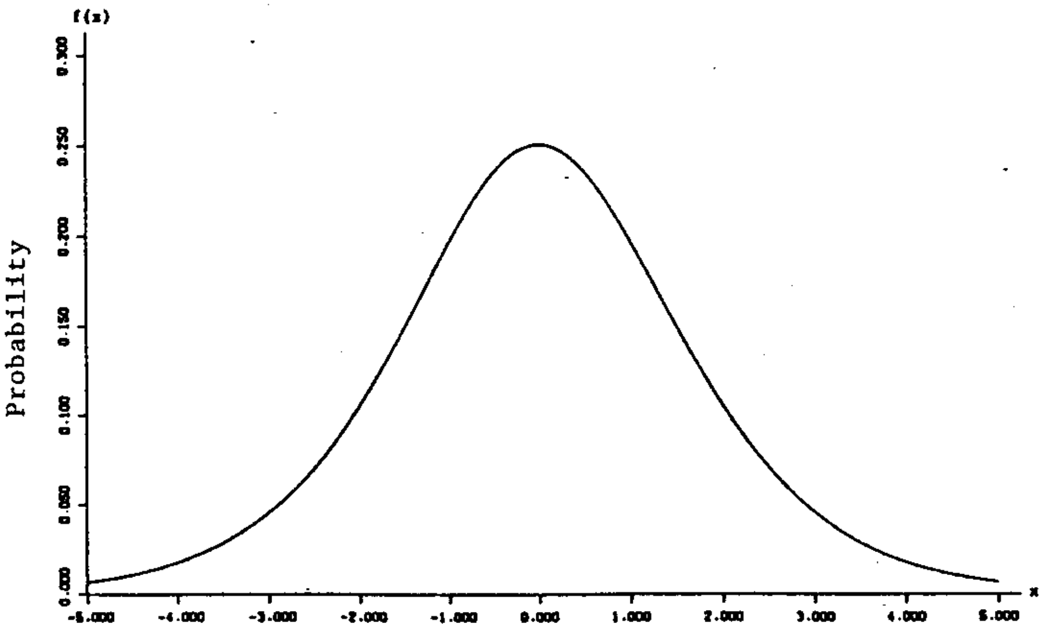


Figure 2.20
Logistic (0, 1) Density Function

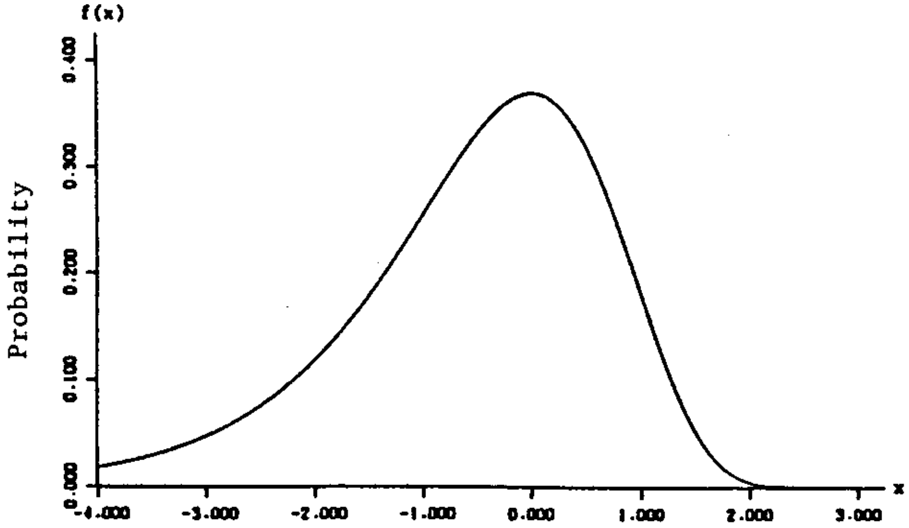


Figure 2.21
Extreme Value Type A (0, 1) Density Function

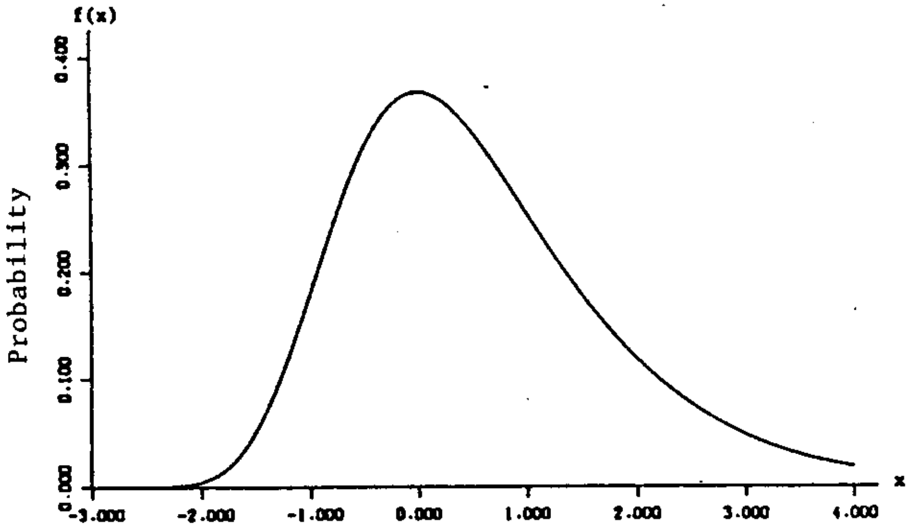


Figure 2.22
Extreme Value Type B (0, 1) Density Function

2.4.3 Bounded Continuous Distributions

This type of distribution can represent any value in a finite interval (a,b). A lower bound and an upper bound must be specified for the distribution. The lower bound must be less than X(1) and the upper bound greater than X(n). The beta distribution is the only bounded continuous distribution used in this study and is shown in Figure 2.23a, 2.23b. The beta distribution has two shape parameters allowing it to be flexible enough to model negative skewness, symmetric distributions, and positive skewness.

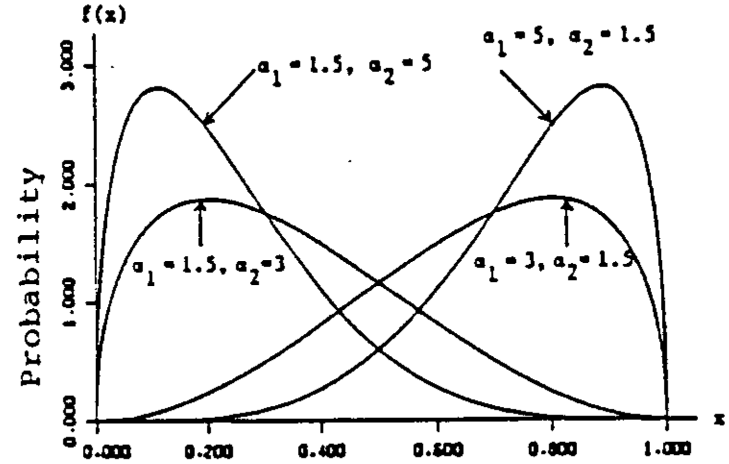


Figure 2.23a
Beta (0, 1, α_1, α_2) Density Function

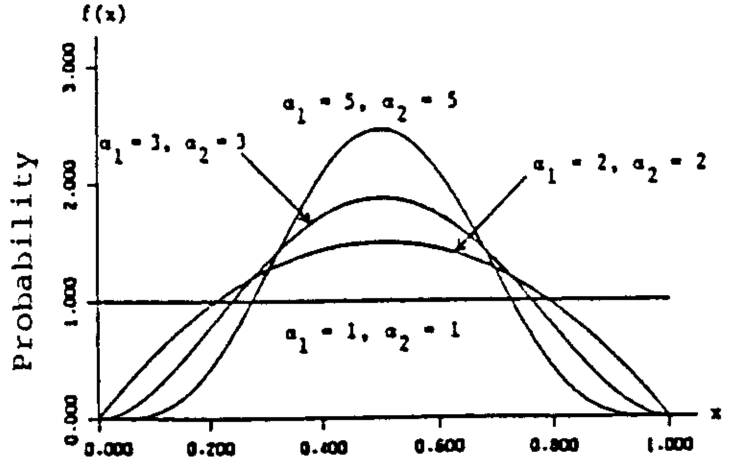


Figure 2.23b
Beta (0, 1, α_1, α_2) Density Function

CHAPTER 3

DATA HISTORY AND PREPARATION FOR ANALYSIS

Probability distributions for continuous random variables can be created by fitting a curve to relative frequency histogram. The relative frequency histogram should be based on a large number of observations in order for the probability distribution to be stable in shape, therefore, approximating the true underlying distribution. The observations must also be consistent in their method of measuring a random variable of interest. It would be desirous that a large number of observations for crop yields from a consistent soil classification be available for estimating a continuous probability distribution for a particular crop.

Also, in establishing evidence that the distribution holds for a specified geographical location, a large number of cases would be needed from the area to support this assertion. Data with these characteristics are nearly impossible to obtain or too costly to collect.

3.1 Data Source And Background

The data for this research study were obtained from individual farm records kept on the Telfarm record-keeping project at Michigan State University. Telfarm was started as a combined effort back in 1925 between the Cooperative Extension Service of the State of Michigan and farmers located in the State of Michigan. The project was created to provide data for research purposes.

Agricultural producers would voluntarily give annual summarized data to Telfarm. In turn, Telfarm would send financial and production analysis reports to the cooperators. The summarized data consisted of yearly totals for variable cost, fixed costs and production figures (crop acreage, total production). Supervised year-end data collection was provided by Extension services to assist farmers in calculating financial data or estimating crop quantities in inventory which was produced on the farm but not sold that year.

In 1963, the structure of the program changed. The mutual effort between the Cooperative Extension Service and state farmers changed to a computerized record keeping-analysis service provided by Telfarm for a fee to subscribers. The only thing that did not change was the benefit of farm data for the Cooperative Extension Services and research purposes. Procedures for collecting the data are basically the same today as they were back in 1963. Monthly financial data sheets are provided to subscribers for collecting income and expense transactions during the month. These data sheets are mailed to the Telfarm processing center at the end of each month by subscribers and an updated financial report is returned to them within 10 days. Annual production data are sent by farmers on a production worksheet provided by Telfarm. The worksheet provides for acres planted, average yield per acre, and total production of feed, cash crops and other commodities.

After all the financial and production data have been collected for a year, annual summaries and reports are processed and sent to subscribers. The data from the annual summaries are used by the

Agricultural Economics Department of Michigan State University to produce Agricultural Economics Reports for use by Michigan farmers in analyzing their respective business. The reports provide farm business analysis summaries for dairy cattle, hogs, beef cattle, and cash crops farms in the State of Michigan.

3.2 Data Retrieval

Copies of the annual summaries for each farm are made for research purposes and stored on microfilm. The data for this research study were compiled directly from the microfilms for each year of the individual farms.

The first step in the data retrieval process was to obtain a listing from Telfarm of all farms that had data in their system. From this listing, each farm with more than 20 years of data was selected as a potential candidate for analysis. The retrieval process involved recording the acreage and yield values for each crop for each year in the time series for individual farms.

The criteria for excluding data from analysis was based on missing years. Each farm was allowed 2 missing years for a 20 year time series independent of when the missing years occurred. Sometimes, due to unknown circumstances, annual reports where not available for a farm. If more than 2 years were missing in the process of recording data, the farm would be dropped from any further consideration.

After all the data were recorded from the microfilms, it was keypunched directly into a permanent storage file on a mainframe computer system.

3.3 Error Checking

Checking the data for errors was approached in three steps. The first step was a double checking of data values during the keypunch process to prevent errors in copying the data from the worksheets to the storage file. The second step involved manually checking a printout of the data storage file for proper spacing of columns and checking acreage and yield data of each individual crop for obvious errors. The third and last step was to identify extreme values in the residual plots which were taken from the individual regression analysis of each farm and each crop (to be covered in the next chapter). Outliers in the residual plots, defined as data points more than (+/-) two standard deviations from the estimated line, were double-checked with values recorded on the work sheets to verify keypunch accuracy.

3.4 Data Validation

Data values that were verified for keypunch accuracy but still considered suspicious of possible errors were checked for consistency with other crops grown on the farm during the same year and approximately the same season. This gave validity for a particular direction of yield response based on the performance of other crops grown by the farm in the same year.

If the crop yield seemed inconsistent with other crops grown during that particular year and season then personal confirmation or disaffirmation of the yield values were obtained from the respective farm managers. Letters were sent to farm operators informing them of our research on yield variation using their farm data and the need to verify the accuracy of specific data values for their farm. In order to help amplify the questionable nature of the data values, a plot of yield vs. time was sent with the letter to help the farmer visualize the departure of the outliers from the rest of the observations. Phone calls were made to each farm manager who received a letter to confirm or obtain the correct yield value(s). Of the 24 questionable data values, 10 or 41 percent were confirmed as correct yield values. Of the remaining 14 incorrect yield values, 3 corrected yield values were obtained, 9 yield values were unable to be corrected due to unvailable records, and 2 yield values were incorrect due to errors in recording process. Also, of the 10 correct data values, 5 values were positive outliers. This observation does not support the hypothesis that any significant departure of yield values from the average would be in the negative direction.

The corrected values obtained for the questionable data were incorporated into the data sets replacing the incorrect values. Missing values were assigned to the data points that were unavailable from historic records which resulted in several farms being dropped because of the data exceeding the limit of 1 missing value per 10 years of data. Farms that were found to have recording errors were omitted from further consideration as more errors might be possible.

3.5 Data Representation Of The Population

Corn is the predominate crop grown in Michigan in terms of acreage used. Figure 3.1 shows the production of corn by counties. Figure 3.2 gives the estimated location of farms providing the data samples. There are a total of 87 farms with corn data meeting the required data

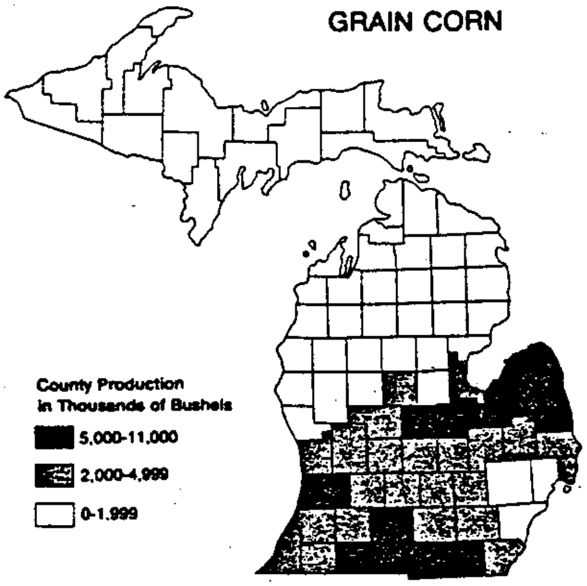


Figure 3.1
Corn Production In Michigan

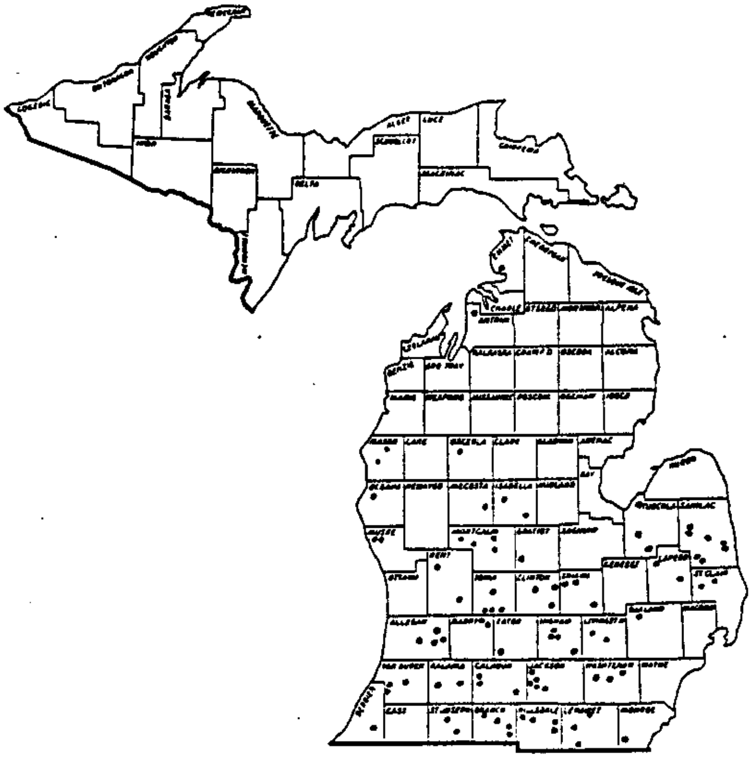


Figure 3.2
Location Of Farms With Corn Data

criterion. Of the 87 farms, 75 have time series of 21 years (1963 - 1983) and the remaining 12 farms have 42 year (1942 - 1983) time series.

Wheat is also an important crop grown in Michigan. Figure 3.3 shows wheat production by county for Michigan. The total number of farms studied for wheat were 35. Twenty-seven farms having 21 years of data (1963 - 1983) and 8 farms with 42 data points (1942 - 1983). Figure 3.4 shows the approximate location of the farms with wheat yield data.

Soybeans are a relatively new crop in Michigan and have seen expanded acreages over the last 10 years. Because of the relatively recent popularity of soybeans it was difficult to find farms to represent the population and still meet the data quality criterion. Figure 3.5 shows the production of soybeans by county for Michigan. In order to obtain more farms for analysis, the desired number of years of data was reduced to > 10 years for soybeans still allowing only 1 missing yield value per 10 years. The total number of farms available for soybeans were 7 farms with data points ranging from 10 to 21. Figure 3.6 shows the estimated location of the farms used to study soybean yield variation.

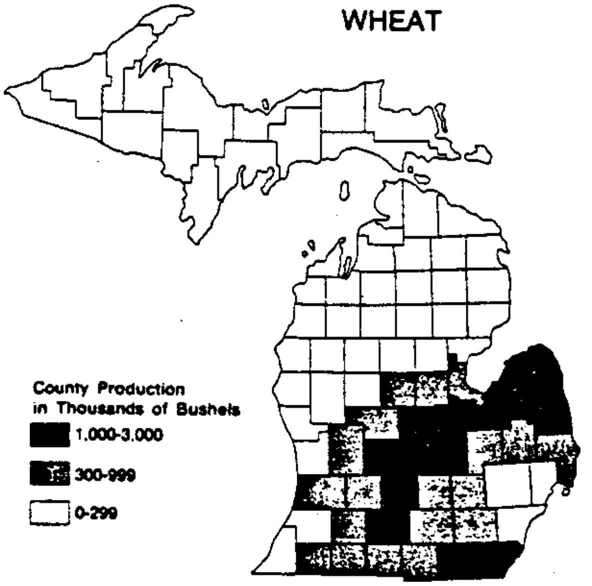


Figure 3.3
Wheat Production In Michigan

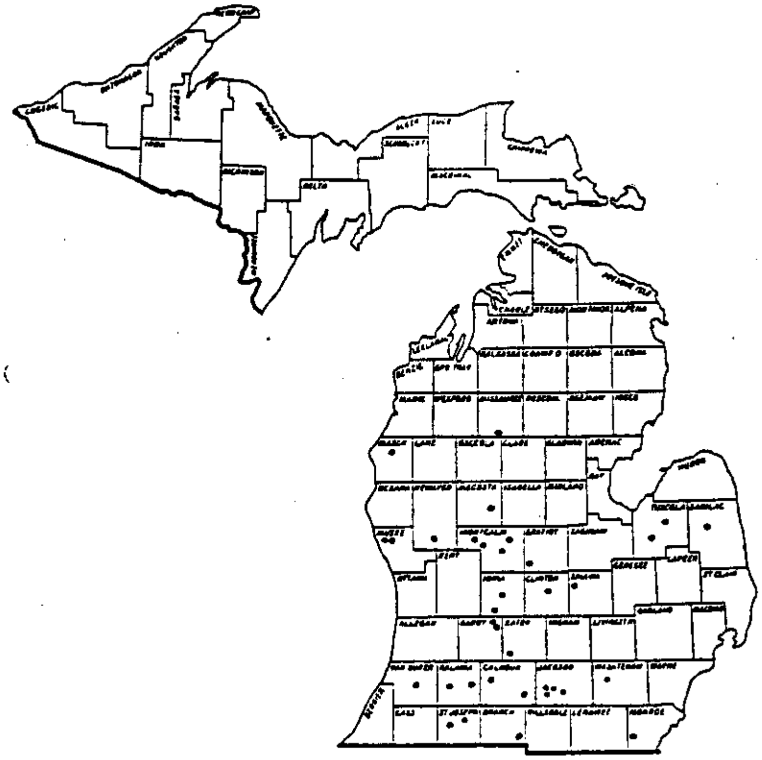


Figure 3.4
Location Of Farms With Wheat Data

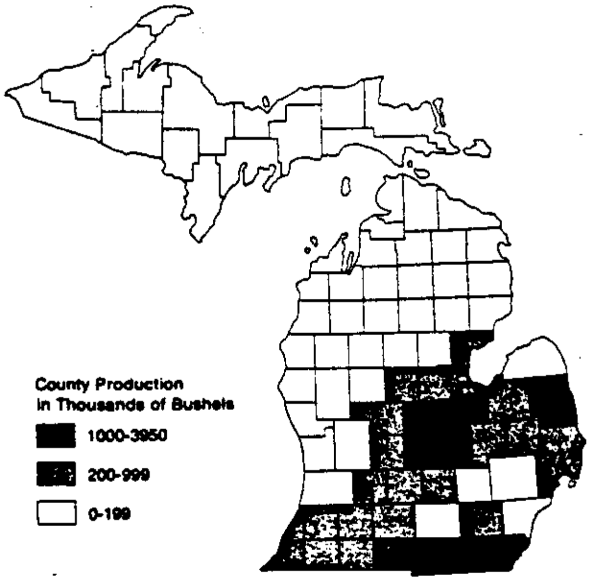


Figure 3.5
Soybean Production In Michigan

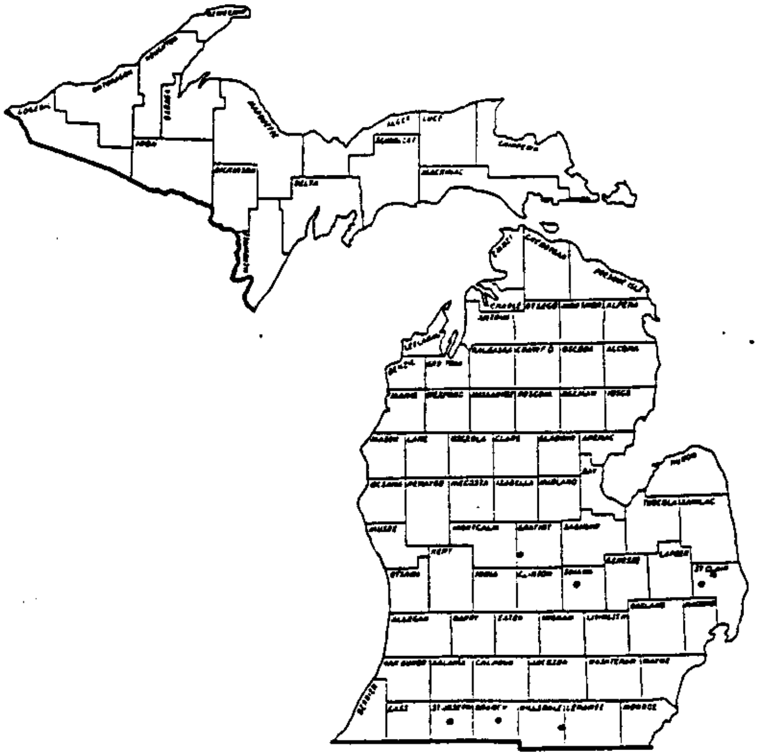


Figure 3.6
Location Of Farms With Soybean Data

CHAPTER 4

EVALUATION OF DISTRIBUTION ASSUMPTIONS

The contents of this chapter cover the research methods and procedures used to test the stated hypothesis that individual farms can be characterized by independent, normally distributed crop yields with constant variance.

4.1 Regression of Yield Vs. Time.

Due to the increase in crop yields over time from the impact of technology and management, data values for a specified farm can not be analyzed from an equivalent basis (see Figure 4.1). Trend over time in yield response must first be quantified, assuming a trend is present, then incorporated in detrending data values (see Figure 4.2). Removing the trend from the data set gives a consistent basis of technology and management for each data value.

The two-variable linear regression model was used to evaluate and quantify trends in yield responses for each individual farm and crop.[7] The model specification and assumptions are as follows:

$$Y_{i,j,t} = \alpha + \beta \mathrm{Time} + e_{i,j,t}$$

---

[7] Robert S. Pindyck and Daniel J. Robinfeld, Econometric Models and Economic Forecasts, 2 ed. (New York:McGraw-Hill, Inc., 1981) p.47.

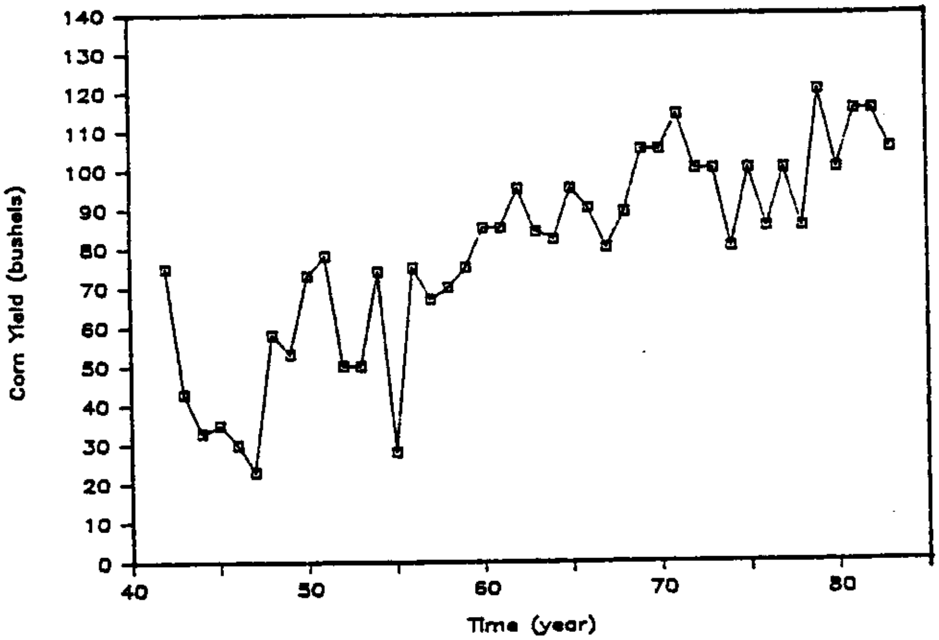


Figure 4.1
Corn Yield vs Time

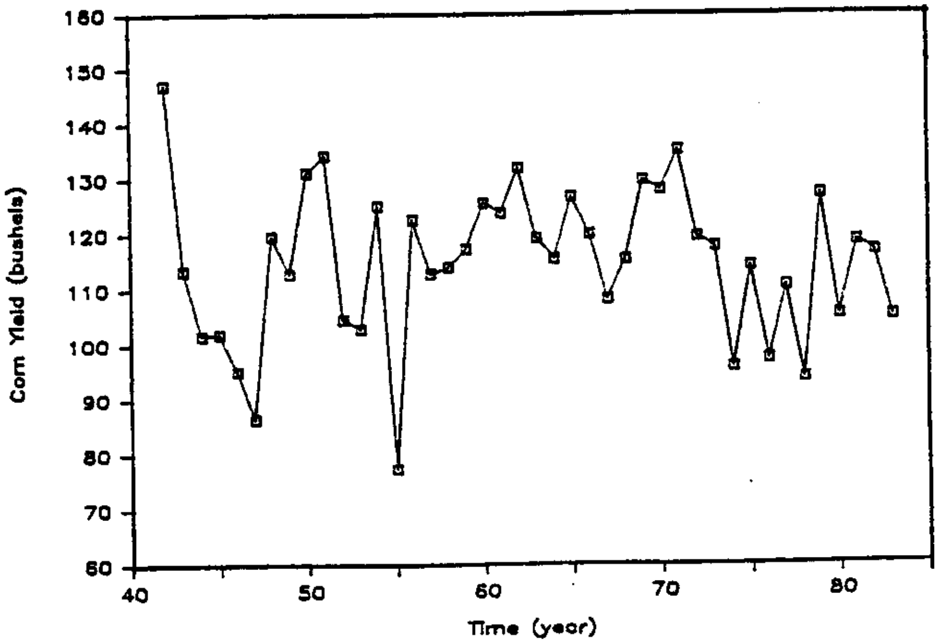


Figure 4.2
Detrended Corn Yield vs Time

where:

$Y_{i,j,t}$ is the yield/planted acre on i^{th} farm, for j^{th} crop, in t^{th} year.

Time is the representative year (i.e., 78,79).

$e_{i,j,t}$ is "random" term that characterizes the variability in yield not associated with trend.

Assumptions:

I. The relationship between Y and X is linear.

II. The X_i 's are nonstochastic variables.

III. a. The error term has zero expected value and constant variance for all observations; that is, $E(e_i) = 0$ and $E(e_i^2) = \sigma^2$.

b. The random variables e_i are uncorrelated in a statistical sense; i.e., errors corresponding to different observations have zero correlation. Therefore $E(e_i\ e_j) = 0$, for $i=j$.

c. The error term is normally distributed.

The assumptions of the two-variable linear regression model encompass the hypothesized yield distribution characteristics (independent, normally distributed, constant variance) set forth in the previous chapter. Therefore, testing the validity of the model assumptions for each farm and each crop is equivalent to testing the hypothesis that yield responses are independent and normally distributed with constant variance. The regression equations estimated for each farm for corn, wheat, and soybeans are given in Appendix A-1.

4.2 Test For Yield Independence

After the regression was run, the error terms of the regression were tested for first order serial correlation which is equivalent to $E(e_i \; e_j) = 0$ for $i = j$. This is consistent with testing for yield dependence between adjacent years (i.e., Y_t and Y_{t-1}). The Durbin-Watson test was used to test for the presence of first order serial correlation in the error terms, the null hypothesis being no presence of serial correlation.[8] The Durbin-Watson statistic is defined as:

$$DW = \frac{\sum_{t=2}^{N} (e_t - e_{t-1})^2}{\sum_{t=1}^{N} e_t^2}$$

Two limits are usually given, labeled d_ℓ and d_μ. In evaluating the DW statistic, values greater than 0 but less than d_ℓ would be evidence for positive serial correlation of the error terms and the null hypothesis would be rejected (see Figure 4.3).

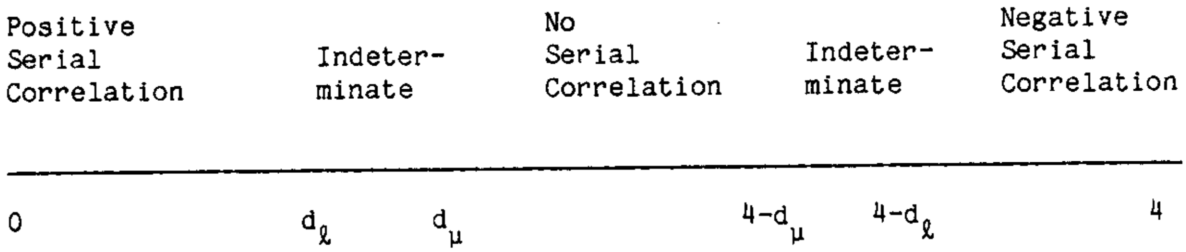


Figure 4.3
Scale For Evaluating The DW Statistic

For values between d_ℓ and d_μ the results are indeterminate, meaning that there can not be conclusive evidence for positive serial correlation or no serial correlation. Statistic values between d_μ

---

[8] J. Durbin and G.S. Watson. "Testing For Serial Correlation In Least Squares Regression," Biometrica 37 (1950):409-428.

and $4-d_\mu$ accepts the null hypothesis of no serial correlation. Again, another indeterminate range is present between $4-d_\mu$ and $4-d_\ell$. DW statistics above $4-d_\ell$ would reject the null hypothesis in favor of evidence for negative serial correlation. It would be assumed that if any evidence of serial correlation is present the most logical result would be positive serial correlation. The reasoning comes from the possibility of weather cycles being present. Heady and Pope[9] found no evidence of autocorrelation in estimating yield response equations for crops in the Corn Belt states. Their conclusion was that weather patterns should be described as being generated from a stochastic process containing no exogenous variables. The literature review for the study was inconclusive in supporting crop yields being random or non-random.

The results of the DW statistics for corn, wheat, and soybeans can be found in Appendix A-1. Seven cases of positive serial correlation and one case of negative serial correlation were found at the 5 percent level of significance for corn (87 cases). This is very little evidence of yield dependance for corn. One case of both positive and negative serial correlation were found for wheat (35 cases). There was no evidence of serial correlation observed for soybeans (7 cases).

4.3 Estimation Of Mean And Variance

Discussion in Section 4.1 indicated that the crop yield values must be detrended for a consistent analysis to be achieved. The fitted line

---

[9] C. Carden Pope, III and Earl O. Heady. The Effects of Technological Progress and Weather on Corn Belt Crop Yields (Ames: Center for Agricultural and Rural Development, (1982)), pp.9-11, 66-71.

estimated from least squares regression was used to detrend the data values. It is important to note that the estimated line need not be significant in order to use the residuals for analysis. This result is taken from work by Park[10] and Glejser[11]. Figure 4.4 shows data values for an actual farm and the line fitted by linear regression. The detrending process simply pivots the fitted line on the right endpoint (most recent year) until it is horizontal as shown in Figure 4.5. This is accomplished by using the Y-estimate of the fitted line for the last year as a reference point. The residuals (data value - Y-estimate) of all the other data values are individually added to the Y-estimate for the last year to obtain a detrended yield value. All yield values are then on an equivalent basis and pure variation can be analyzed from a common mean value. It is this y-estimate of the line for the last year in the time series that is used as an estimate for the mean. This estimate for the mean for the yield distribution was observed to be greater than the sample mean in most all cases as slopes were predominantly positive. The mean ($\hat{Y}$) for the distribution is reported in Appendix A-1 for corn, wheat, and soybeans, respectively.

For purposes of estimating variance of the crop yield distribution, the residuals from the linear regression were used as being equivalent to using the detrended yield values. Consider again Figures 4.4 and 4.5. The variance from the fitted line (residuals) are the same. The detrending process simply moves the line for common reference. The

---

[10] R.E. Park. "Estimation with Heteroscedastic Error Terms," Econometric 34 (October 1966):888.

[11] H. Glejser. "A New Test for Heteroscedasticity," Journal of the American Statistical Association 64(1969):316-323.

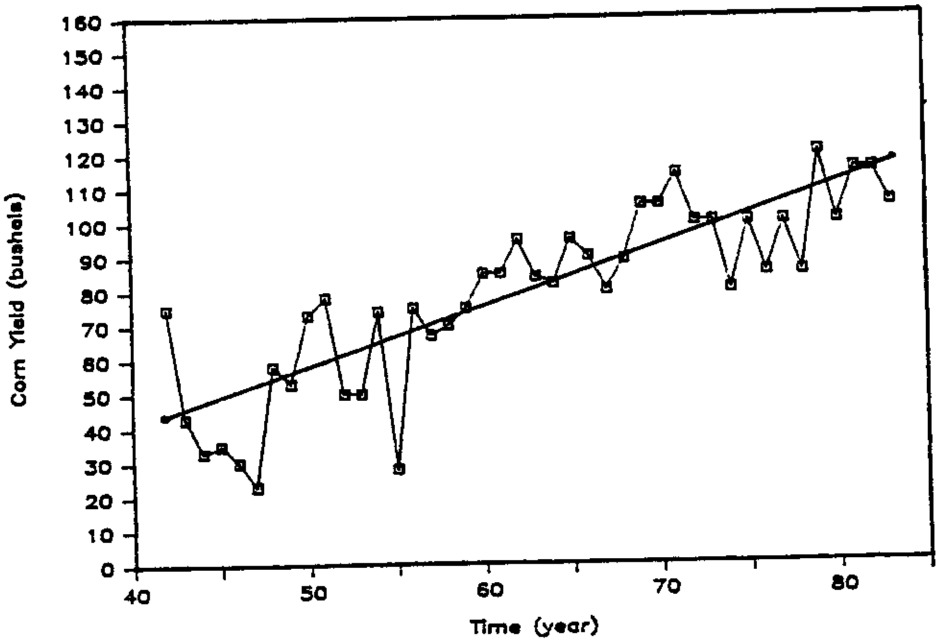


Figure 4.4
Regression Line Fitted To Corn Yield Data

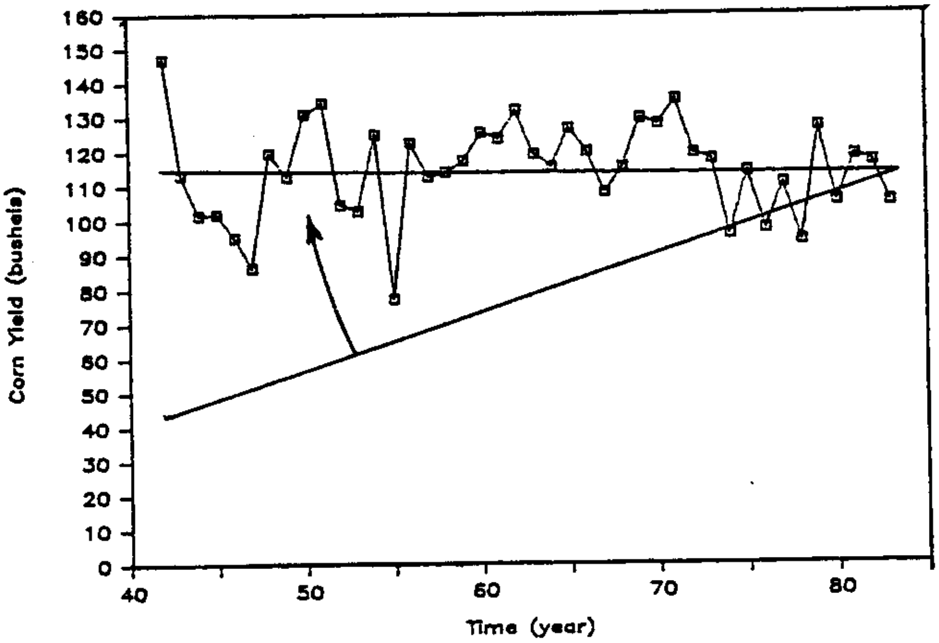


Figure 4.5
Detrending Yield Data With Y-Estimate Of Last Year

variance was computed using the following formula:

$$s^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N - 1}$$

where:

s is the estimate of variance.

X_i is a residual value.

$\overline{X}$ is the mean of the residuals ($\overline{x} = 0$).

N is the number of residuals.

The denominator is adjusted by subtracting 1 to provide an unbiased estimate of the population variance. The standard deviations (square root of variance) are reported in Appendix A-1 for corn, wheat, and soybeans, respectively.

4.4 Test For Constant Variance

The nature of the variance in crop yields over time must be investigated for evidence to support the constant variance assumption of the two-variable linear regression model. Testing the statistical model previously defined for the absense of heteroscedasticity is equivalent to testing a farm for constant variance $E(e_i^2) = \sigma^2$ for a specific crop. Heteroscedasticity is the presence of increasing variance (see Figure 4.6) or decreasing variance (see Figure 4.7).

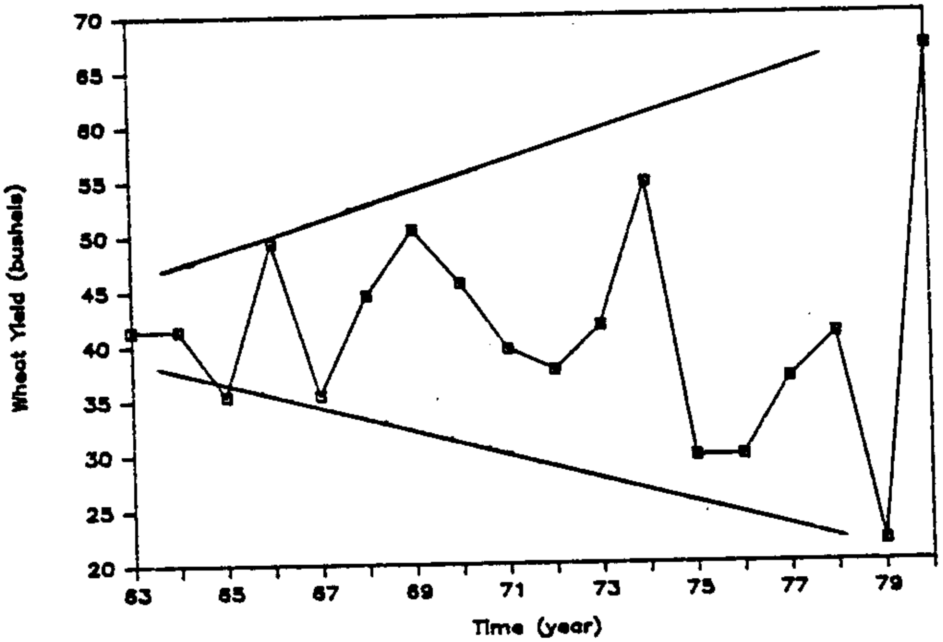


Figure 4.6
Increasing Variance

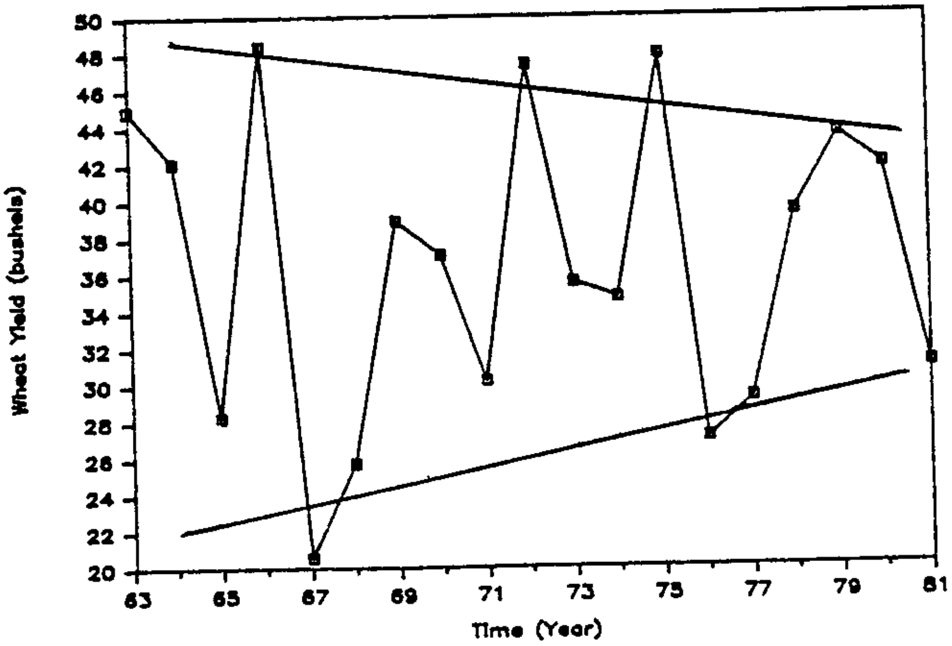


Figure 4.7
Decreasing Variance

The procedure outlined by Stewart and Wallis for evaluating the presence of heteroscedasticity involves regressing the absolute value of the residuals from the two-variable linear regression model against time.[12]

$$|e_t| = \alpha_1 + \alpha_2 \text{ Time}$$

where:

e_t is a residual value in the t^{th} year.

α_1, α_2 are unknown parameters

Time is the representative year.

The null hypothesis of homoscedasticity can be tested by evaluating if $\alpha_2 = 0$ with a t-test. The estimated values α_2 and its significance level are presented in Appendix A-1 for corn, wheat, and soybeans, respectively.

The structure of the hypotheses state:

H_0 :constant variance ($\alpha_2 = 0$)

H_1 :decreasing variance ($\alpha_2 < 0$)

H_1 :increasing variance ($\alpha_2 > 0$)

Using a significance level of 15 percent, 33 of the 87 farms were found to have evidence of heteroscedasticity for corn. The 10 percent significance level showed 25 farms having evidence of heteroscedasticity and 16 farms at the 5 percent level. Scattergrams for the absolute value of residuals showed little evidence to confirm the rejection of the null hypothesis of constant variance for those farms having a

---

[12] Mark B. Stewart and Kenneth F. Wallis. Introductory Econometrics, 2 ed. (New York:Halsted Press, 1981) p.248.

significant value for the different levels of significance. In general, tests for heteroscedasticity are low in power and must be utilized carefully. In view of evaluating the scattergrams for the farms rejecting the hypothesis of constant variance, there appears to be little evidence for heteroscedasticity for corn. Of the 35 farms for wheat, 16 farms were found to be significant at the 15 percent level, 8 farms at the 10 percent level, and 2 at a 5 percent level of significance. Soybeans have 2 farms that are significant at the 5 percent level. The conclusions for wheat and soybeans are the same as for corn, the scattergrams of the absolute value of residuals give little supportive evidence of heteroscedasticity for those farms that tested as being significant.

4.5 Test For Skewness

The skewness of a distribution, also known as the third moment, is a measure of its symmetry. Skewness is defined by:

$$\sqrt{\beta_1} = \sum_{i=1}^{N} [X_i - E(X))^3]/(s^2)^{3/2}$$

where:

$\sqrt{\beta_1}$ is the coefficient of skewness.

X_i is the i yield value.

$E(X)$ is the mean of the distribution.

s^2 is the estimated variance.

If $\sqrt{\beta_1}$ is negative the distribution is skewed to the left and, if $\sqrt{\beta_1}$ is positive the distribution is skewed to the right. The coefficient of skewness for each farm are given in the Appendix for corn, wheat, and soybeans, respectively. Ten cases of skewness significant at the 5 percent level were found for corn (87 cases). Seven cases of skewness were observed for wheat (35 cases) and one case for soybeans (7 cases) at the 5 percent level of significance. Evidence for non-symmetrical distributions for corn and soybeans appear to be insufficient although wheat is questionable as 1 out of every 5 farms have significant skewness. Another simple way to evaluate skewness is to compare the median and the mean of the distribution. If the distribution is skewed to the left, or negatively skewed, the median will be larger than the mean. Conversely, if the skewness is to the right, or positively skewed, the median will be smaller than the mean.

4.6 Test For Kurtosis

The kurtosis of a distribution measures the "tail weight" resulting from the peakedness or flatness of the distribution and is also known as the fourth moment. Kurtosis is defined as:

$$\beta_2 = \sum_{i=1}^{N} [(X_i - E(X))^4]/(s^2)^2 - 3$$

where:

β_2 is a measure of kurtosis.
X_i is the i^{th} yield value.
$E(X)$ is the mean of the distribution.
s^2 is the estimated variance.

The normal distribution has a kurtosis of 3; therefore, β_2 gives the kurtosis relative to the normal distribution. Negative values of kurtosis describe a distribution that is flatter than the normal distribution and positive values denote distributions which are more peaked (narrow) than the normal. The values of kurtosis for corn, wheat, and soybeans are given in Appendix A-1. Nine farms exhibited significant values of kurtosis at the 5 percent level for corn, of which six are positive values and three are negative. Eight significant values were found for wheat and all values were positive. Soybeans have four significant values, two being positive and two being negative.

There is little evidence of the distribution being more peaked or flatter than the normal distribution for corn. It is interesting to note that the significant values of kurtosis are characterized mainly by peakedness. Wheat and soybeans exhibit fairly strong evidence of kurtosis, wheat being characterized by a peaked or narrow distribution.

4.7 Cumulative Probability Plots

At this point, an idea of the nature of the distribution can be obtained from evaluating the mean and median, skewness, and kurtosis. What is really needed is a picture of the sample distribution of data values superimposed against the normal distribution to evaluate how well the normal distribution fits the data.

The cumulative probability distribution for a random variable X is defined as: $F(x_i) = P[X < x_i]$

for any real number x_i.

The assumed approximation of a cumulative distribution for the sample yield values $F_n\ (x_i)$ would be the proportion of the x's that are less than or equal to x_i. Since we would expect that observations larger than x_n are possible, a correction factor is needed. Law and Kelton suggest that the sample distribution function $F_n(X)$ be the proportion of the x's that are less than or equal to x minus the correction factor 0.5/n.[13] Figure 4.8 shows a cumulative probability plot for the Normal probability distribution F(x). Figure 4.9 shows the cumulative probability distribution for the sample distribution function $F_n(x)$ superimposed over the normal cumulative probability. The plot was constructed by ordering the detrended yields and assigning:

$$F_n(x_i) = \frac{i - 0.5}{N}$$

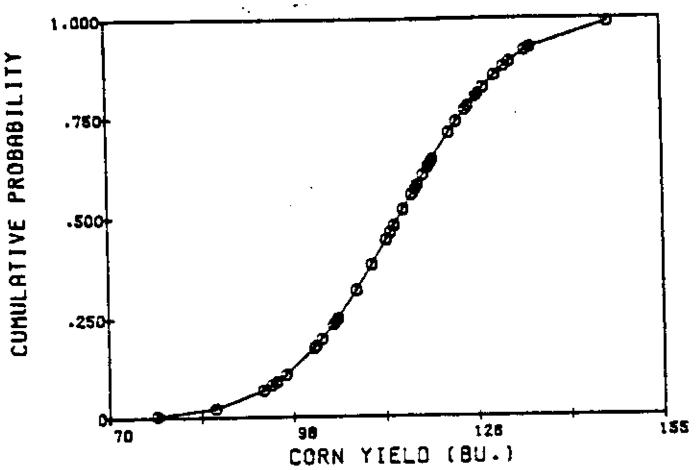


Figure 4.8
Normal Cumulative Probability Distribution

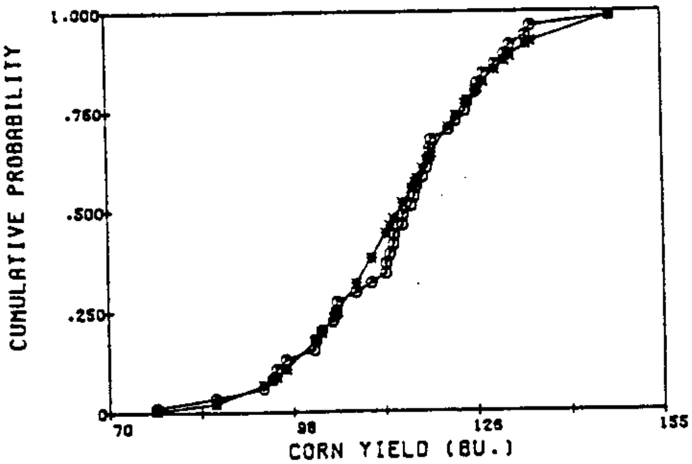


Figure 4.9
Sample Distribution Superimposed Over The Normal Distribution

---

[13] Averill M. Law and W. David Kelton. Simulation Modeling and Analysis, (New York:McGraw-Hill, 1982), p.181.

for all data points. The points $(x_1,F_n(x_1))$, $(x_2,F_n(x_2))$, ..., $(x_N,F_n(x_N))$ were plotted to obtain the cumulative probability plot shown in Figure 4.9.

4.8 Test For Normality

The residuals from the two-variable linear regression for each farm can be tested statistically to evaluate whether they are distributed around the estimated line according to a normal distribution. One factor that must be taken into consideration in choosing a statistical test to evaluate normality is the small sample size for the respective farms. The primary concern is the power of the test given a small sample size. The test statistic which was used for evaluating normality in this study was the Shapiro-Wilks test for normality.[14] In a comparative study of various tests for normality, the Shapiro-Wilks W-test was found to be a superior indicator of non-normality for the various symmetric, asymmetric, short- and long-tailed alternatives and over all sample sizes (n = 10, 15, 20, 35, 50).[15] The W-test evaluates the variation of the sample and is exceptional in sensitivity for continuous distributions over the distance tests KS(Kolmogorov-Smirnov), CM(Cramer-Von Mises), WCM(Weighted CM), D(Modified KS), and CS(Chi-squared). While it is true that a judgement based on both skewness, $\sqrt{\beta_1}$, and kurtosis, β_2, will be sensitive, the W-statistic is

---

[14] S.S. Shapiro and M.B. Wilks. "An Analysis of Variance Test for Normality (Complete Samples)," Biometrica, 52 (December 1965):591-611.

[15] S.S. Shapiro, M.B. Wilk, and Mrs. H.J. Chen. "Comparative Study of Various Tests for Normality," American Statistical Association Journal, 63 (December 1968):1343-1372.

typically as good as the best of either of these tests. However, in some cases the W-test has considerably higher power than skewness and kurtosis measures. The Shapiro-Wilks W-test should then be used in conjunction with skewness and kurtosis in evaluating the form of the distribution.

The data for a W-test should consist of a random sample x_1, x_2, x_3, ..., x_N of size N associated with some unknown distribution function F(x). The structure of the hypotheses state:

H_o: F(x) is a normal distribution function.

H_1: F(x) is a non-normal distribution function.

The W-statistic can be computed as follows:

a) Order the observations from smallest to largest to obtain an ordered sample $X^{(1)} \leq X^{(2)} \leq X^{(3)} \ldots \leq X^{(N)}$

b) Compute the denominator s^2.

$$s^2 = \sum_{i=1}^{N} (X_i - \overline{X})^2$$

where $\overline{X}$ is the sample mean.

c) Compute numerator b^2.

1) If N is even, N = 2k.

$$b^2 = [\sum_{i=1}^{N} a_i (X^{(N-i+1)} - X^{(i)})]$$

where values of a_i are given in Table 5 (Shapiro 1965).

2) If N is odd, N = 2k + 1, the computation is the same for N being even. The sample median $(X^{(k+1)})$ does not enter the computation of b^2.

d) Compute $W = b^2/s^2$.

The decision rule for the W-statistic rejects H_o at the level of significance α if W is less than the quantile as given by Table 6 (Shapiro 1965).

To illustrate the use of the W-statistic, the data used to construct the probability plot in Figure 4.9 will be tested for normality. Using the residuals from the two-variable linear regression analysis of the corn yield data (shown in Table 4.1) is equivalent to using the detrended yields for the W-test.

The first step in computing the W-test is to order the observations from smallest to largest. This is shown in column A of Table 4.2. The next step is to calculate the sum of the variance for the sample or the denominator (s^2) of the W-test. Since the residuals are deviations from the mean yield, each residual is squared and the total of the squared deviations is the sum of the variance. The squared residuals (deviations) are shown in column B and the sum of the variance (s^2) is shown at the bottom of Table 4.2. The third step is to compute the numerator (b^2). The table values of a_i are shown in column C and the $X^{(N-i+1)} - X^i$ values are given in column D. The product of the values in columns C and D are given in column E. These products ($a_i(X^{(N-i+1)} - X^{(i)}$) are summed and the total is squared to obtain the numerator (b^2). The W-statistic then is simply b^2 divided by s^2. The result for the W-statistic is shown in Table 4.2 and is significant at the level $\alpha = .95$.

The results of the W-statistic for each farm are given in Appendix A-1 for corn, wheat, and soybeans, respectively. Six farms were found to have significant W-statistics for corn at the 5 percent level. The assertion made by Shapiro (1965) of the W-statistic being as good as skewness and kurtosis measures and having possibly higher power in some cases was not observed in analyzing the three crops. There are several cases where the skewness or kurtosis measure is significantly different

Table 4.1
Corn Yield Data

YR	YIELD	Y-ESTIMATE	RESIDUAL
42	75	42.56368	32.43632
43	43	44.32481	-1.32481
44	33	46.08592	-13.08592
45	35	47.84784	-12.84784
46	30	49.60816	-19.60816
47	23	51.36928	-28.36928
48	58	53.13041	4.86959
49	53	54.89153	-1.89153
50	73	56.65265	16.34735
51	78	58.41377	19.58623
52	50	60.17489	-10.17489
53	50	61.93601	-11.93601
54	74	63.69713	10.30287
55	28	65.45826	-37.45826
56	75	67.21938	7.78062
57	67	68.98051	-1.98051
58	70	70.74162	-0.74162
59	75	72.50274	2.49726
60	85	74.26386	10.73614
61	85	76.02498	8.97502
62	95	77.78611	17.21389
63	84	79.54723	4.45277
64	82	81.30835	0.69165
65	95	83.06947	11.93053
66	90	84.83059	5.16941
67	80	86.59171	-6.59171
68	89	88.35283	0.64717
69	105	90.11396	14.88604
70	105	91.87508	13.12492
71	114	93.63621	20.36379
72	100	95.39732	4.60268
73	100	97.15844	2.84156
74	80	98.91956	-18.91956
75	100	100.6807	-0.6807
76	85	102.4418	-17.4418
77	100	104.2029	-4.2029
78	85	105.9641	-20.9641
79	120	107.7252	12.2748
80	100	109.4863	-9.4863
81	115	111.2474	3.7526
82	115	113.0085	1.9915
83	105	114.7697	-9.7697

Table 4.2

Shapiro-Wilks Test For Normality

A RESIDUALS	B RESID^2	C TABLE	D X(N-i)-X(i)	E C * D
-37.4583	1403.1212	0.3917	69.8946	27.3777
-28.3693	804.8160	0.2701	48.7331	13.1628
-20.9641	439.4935	0.2345	40.5503	9.5090
-19.6082	384.4799	0.2085	36.8221	7.6774
-18.9196	357.9498	0.1874	35.2669	6.6090
-17.4418	304.2164	0.1694	32.3279	5.4763
-13.0859	171.2413	0.1535	26.2108	4.0234
-12.8478	165.0670	0.1392	25.1219	3.4970
-11.9360	142.4683	0.1259	23.8665	3.0048
-10.1749	103.5284	0.1136	20.9110	2.3755
-9.7697	95.4470	0.1020	20.0725	2.0474
-9.4863	89.9899	0.0909	18.4613	1.6781
-6.5917	43.4506	0.0804	14.3723	1.1555
-4.2029	17.6644	0.0701	9.3723	0.6570
-1.9805	3.9224	0.0602	6.8501	0.4124
-1.8915	3.5779	0.0506	6.4942	0.3286
-1.3248	1.7551	0.0411	5.7776	0.2375
-0.7416	0.5500	0.0318	4.4942	0.1429
-0.6807	0.4634	0.0227	3.5222	0.0800
0.6472	0.4188	0.0136	1.8501	0.0252
0.6917	0.4784	0.0045	1.2998	0.0058
1.9915	3.9661			
2.4973	6.2363			
2.8416	8.0745			
3.7526	14.0820			
4.4528	19.8272			
4.6027	21.1847			
4.8696	23.7129			
5.1694	26.7228			
7.7806	60.5380			
8.9750	80.5510			
10.3029	106.1491			
10.7361	115.2647			
11.9305	142.3375			
12.2748	150.6707			
13.1249	172.2635			
14.8860	221.5942			
16.3474	267.2359			
17.2139	296.3180			
19.5862	383.6204			
20.3638	414.6839			
32.4363	1052.1149			
SUM RESID	SUM RES^2	S	SUM E^2	W-TEST
0	8121.23	8121.23	8007.26	0.98597

from normality, but the W-statistic is not significant at comparable levels of α. Seven farms have siginificant W-statistics for wheat indicating rejection of the normality hypothesis. Soybeans have no significant departures from normality.

4.9 Group Test For Normality

One useful feature of the Shapiro-Wilks W-test is that several independent goodness-of-fit tests may be combined into one overall test of normality.[16] This is necessary when several small samples from possibly different populations are insufficient by themselves to reject the hypothesis of normality, but their combined evidence is enough to disprove normality.

The structure of the hypotheses state:

H_0: The population is normally distributed.

H_1: The population is non-normally distributed.

To combine the results from a series of W-tests, the following steps must be used:

1) Each W-statistic must be converted to values of G as described by table A19.[17] For $7 \leq N \leq 50$, enter table A19 with N to find the coefficients b_n, c_n, and d_n. Then compute:

$$G = b_n + c_n \ln[(W - d_n)/(1 - W)]$$

2) All N-values of G are added together.

---

[16] W.J. Conover. Practical Nonparametric Statistics, 2ed. (New York:John Wiley & Sons, Inc., 1980), pp.363-367.

[17] Pearson and Hartley (1972) as adapted from Conover (1980).

3) The total of all N values of G is divided by $\sqrt{N}$ to get the Z statistic, which is approximately standard normal under the null hypothesis.

4) If Z is less than the α quantile from the standard normal table, the null hypothesis is rejected at the level α.

The grouped test of normality for corn, wheat, and soybeans is given in Appendix A-2. The group test for corn has an estimated level of significance $\hat{\alpha} = .057$. Wheat and soybeans have levels of significance $\hat{\alpha} = .013$ and $\hat{\alpha} = .531$, respectively. The null hypothesis of Normally distributed yields are supported at the 5 percent level of significance for corn and soybeans and the 1 percent level of significance for wheat.

4.10 Summary Of Evaluation

If one could choose the amount of data for this type of research, it would be preferable to have 100-200 farms with 50 or more years of data for each crop. The facts of reality though do not make this kind of data possible for this research study. The conclusions of this research study should then be taken in light of the data available as was discussed in Chapter 3. In summarizing the findings from Chapter 4, it is concluded that corn, wheat, and soybean yield variation could be described by a Normal distribution, independently distributed with constant variance. Very little evidence was present for serial correlation for the three crops to disprove independence of the yield distribution. Even with significance levels as high as 15 percent, little supportive evidence was found for heteroscedasticity considering the power of the tests presently available. The results for skewness, kurtosis, W-statistics, and the grouped W-statistics do not provide substantial evidence to disprove the normality assumption.

CHAPTER 5

EVALUATION OF NORMALITY BY SOIL GROUPS

It was pointed out in Section 1.2 that the focus of this research paper is to evaluate the functional form of yield distributions and estimate the distribution parameters. The findings from the analysis described in Chapter 4 indicate that corn, wheat, and soybean yields for Michigan could be described as independent, normally distributed with constant variance. The question is asked then, how stable are these conclusions?

The ideal analysis would be to fit an appropriate distribution functional form to each farm to evaluate the support of yields being normally distributed as found in Chapter 4. Due to limited resources of time, the distribution assumptions are evaluated for farms having the same soil association in this chapter. The fitted distributions for each farm are considered for a chosen soil association in Chapter 6.

5.1 Classification Of Soil Groups

Different areas of the state of Michigan can be generally described by particular soil associations. For instance, the Saginaw bay and the thumb area of Michigan are generally composed wet clayey soils developed in lacustrine sediments. The area of Eaton, Ingham, Livingston, Oakland, and Lapeer counties are dominated by loamy soils developed in glacial till. The lower part of the state, Cass, St Joseph, Branch, Van

Buren, Kalamazoo, Calhoun, and Barry counties, is largely composed of loamy soils underlain by sand and gravel.

The only knowledge of location for the farms was the mailing address, legal descriptions were not available. Using this information, a Michigan map, and the soil association map of Michigan,[18] each farm was identified as belonging to a regional soil group as defined in Table 5.1. Table 5.1 gives the key for the dominate soil groups used in this study. Tables 5.2 and 5.3 give the soil groups available for corn and wheat, respectively. Since the data for soybeans is limited, soil groups would not be meaningful for analysis. For this reason soil groups are not given for soybeans.

5.2 Evaluation Of Yield Independence By Soil Group

In the previous chapter, yield independence was analyzed for each crop over all farms. The purpose of evaluating yield independence by soil groups is to find evidence of yields being dependent on yield response for the previous year for a particular group of soils. Therefore, the Durbin-Watson statistics for all the farms in a specific soil group can be evaluated for evidence of yield dependence.

The associated statistics for soil groups M1, M3, M4, M5, and M7 for corn can be found in the Appendix (B-1-B-5). Soil groups M1 and M5 are the only soil groups that have significant evidence with 25 percent of the farms having significant Durbin-Watson statistics. The number of farms in the soil groups is small and must be considered in evaluating

---

[18] Michigan State University Cooperative Extension Service. "Soil Asssociation Map of Michigan," (Extension Bulletin E-1550, December 1981).

Table 5.1
Soil Association Map Of Michigan

M1 - Areas Dominated by Clayey Soils
50 -Perrington-Ithaca Association
53 -Morley-Glynwood-Blout Association
58 -St. Clair-Nappanee Association

M2 - Areas Dominated by Wet Clayey Soils
48 -Lenawee-Toledo-Del Rey Association
52 -Ithaca-Pewamo-Belleville Association
60 -Hoytville-Nappanee Association
62 -Blout-Pewamo Association

M3 - Areas Dominated by Loamy Soils
34 -Hillsdale-Riddles Association
41 -Marlette-Capac Association
45 -Bover-Riddles-Marlette Association
56 -Riddles-Teasdale Association
57 -Miami-Conover-Brookston Association
72 -Lapeer-Hillsdale Association

M4 - Areas Dominated by Wet Loamy Soils
42 -Capac-Parkhill Association
61 - Kibbie-Colwood Association
64 -Metamora-Blout-Pewamo-Selfridge Association
68 -Wixom-Londo-Guelph Association
69 -Tappan-Londo Association
70 -Tappan-Londo-Poseyville Association
71 - Tappan-Belleville-Essexville Association
73 -Shebeon-Kilmanagh Association

M5 - Areas Dominated by Sandy Soils
40 -Oakville-Plainfield-Spinks Association
63 -Oakville-Tedrow-Granby Association
65 -Grattan Association
66 -Grattan-Covert-Pipestone Association
67 -Spinks-Perrinton-Ithaca Association

M6 - Areas Dominated by Wet Sandy and Wet Loamy Soils Underlain by Sand and Gravel
38 -Tedrow-Granby Association
39 -Brady-Wasepi-Gilford Association
49 -Tedrow-Tedrow, loamy substratus-Selfridge Association
51 -Pipestone-Kingsville-Saugatuck-Wixom Association
59 -Belleville-Selfridge-Metea Association

M7 - Areas Dominated by Loamy Soils Underlain by Sand and Gravel
35 -Spinks-Oshtemó-Boyer Association
36 -Schoolcraft-Kalamazoo-Elston Association
37 -Kalamazoo-Oshtemo Association
44 -Boyer-Oshtemo-Houghton Association
46 -Boyer-Wasepi Association
54 -Boyer-Fox-Sebewa Association
55 -Oshtemo-Brady-Gilford Association

Table 5.2
Soil Groups For Corn

M1 - Areas Dominated by Clayey Soils
M3 - Areas Dominated By Loamy Soils
M4 - Areas Dominated by Wet Loamy Soils
M5 - Areas Dominated by Sandy Soils
M7 - Areas Dominated by Loamy Soils Underlain by Sand and Gravel

Table 5.3
Soil Groups For Wheat

M1 - Areas Dominated by Clayey Soils
M3 - Areas Dominated by Loamy Soils
M4 - Areas Dominated by Wet Loamy Soils
M7 - Areas Dominated by Loamy Soils Underlain by Sand and Gravel

the results. Each of the farms that were significant at the 5 percent level were checked to see if the functional form of the fitted model could be causing a significant DW statistic. Figure 5.1 shows a linear model being fitted to data which could more accurately be fitted by a non-linear model. The example depicted by Figure 5.1 shows how the DW statistic could be significant but the problem is not necessarily serial correlation, but rather the model choice.

The residual plot and scattergram for each of the farms having a significant DW statistic were evaluated to determine whether this problem was present for corn. There were no cases found of significant DW statistics that might have been corrected by changing the model specification.

The associated statistics for soil groups M1, M3, M4 and M7 for wheat are given in the Appendix (B-6-B-9). Soil group M7 is the only soil class that has evidence of serial correlation and of the 2 farms that have significant DW statistics, model specification could possibly change the results of 1 farm.

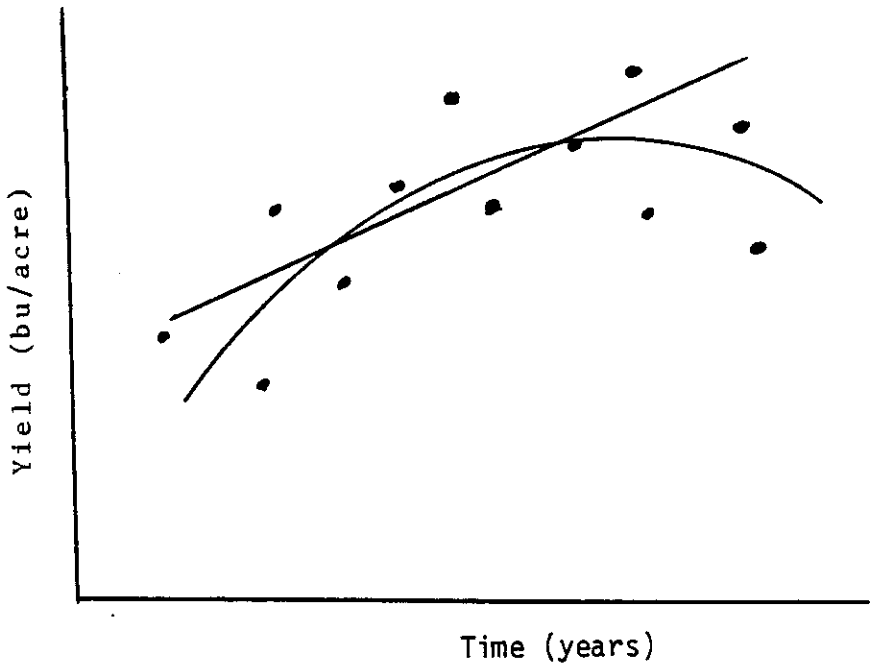


Figure 5.1
Serial Correlation Caused By Model Choice

5.3 Evaluation Of Constant Variance By Soil Group

The assertion of constant variance can also be analyzed by soil groups. The α_2 values (see Section 4.4) for the farms pertaining to a specific soil group can be evaluated for heteroscedasticity. As mentioned in Section 4.4, tests for heteroscedasticity are generally characterized as being low in power. Having this perspective, each farm with a significant α_2 value was evaluated to determine if the scattergram of the absolute value of residuals vs time presented evidence of increasing or decreasing variance. Figures 5.2 and 5.3 show two farms with significant α_2 values but only one farm (Figure 5.2) has evidence of hereroscedasticity.

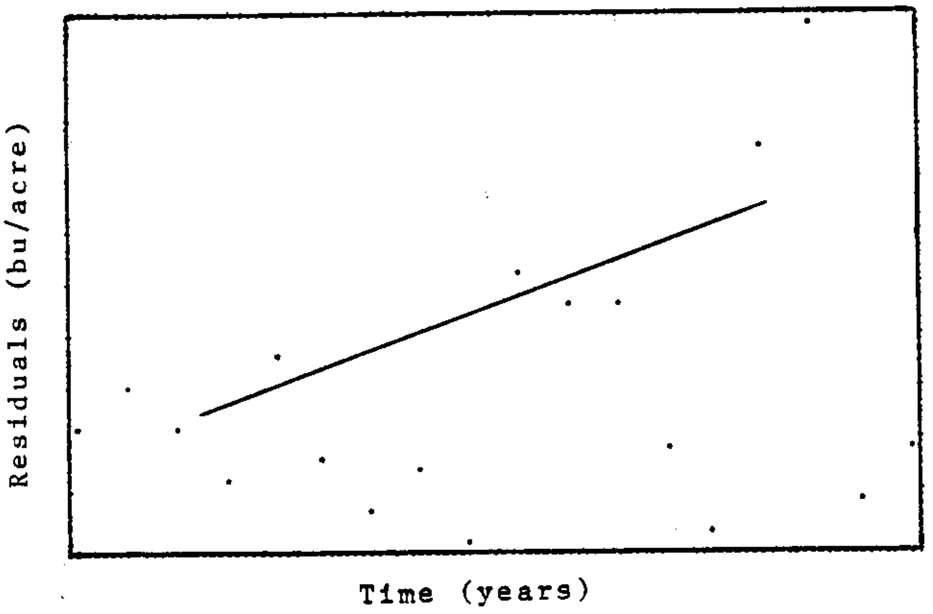


Figure 5.2
Evidence Of Increasing Variance

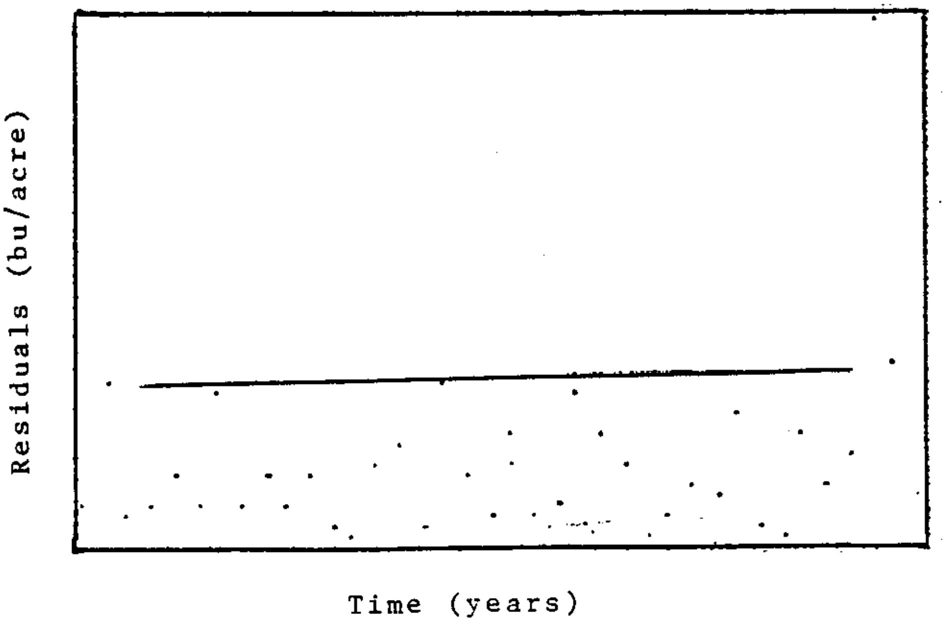


Figure 5.3
Evidence Supporting Constant Variance

The α_2 values for farms growing corn in soil groups M1, M3, M4, M5, and M7 are presented in the Appendix (B-1-B-5). Three of these five soil groups have evidence of heteroscedasticity. Soil group M1 has 4 farms with significant α_2 values, of which 3 farms are confirmed to have heteroscedasticity (2 farms - increasing variance, 1 farm - decreasing variance). Soil group M3 has 18 farms with significant α_2 values, 11 farms being confirmed as having heteroscedasticity (8 - increasing variance, 3 - decreasing variance). Soil group M7 has 6 significant α_2 values of the 18 farms but only 2 are confirmed to have increasing variance.

The α_2 values for farms growing wheat in soil groups M1, M3, M4, and M7 are given in the Appendix (B-6-B-9). Three of the four soil groups show evidence of heteroscedasticity. Soil group M3 has 7 significant α_2 values of which 3 values were confirmed to have increasing variance and 1 value with decreasing variance. Soil groups M4 and M7 both have 1 case of increasing variance.

In summary, there appears to be a problem with heteroscedasticity on loamy soils for corn being characterized mainly by increasing variance. The evidence of heteroscedasticity that exists for both corn and wheat is predominately depicted by increasing variance.

5.4 Test Of Normality Assumption By Soil Group

The Shapiro-Wilks grouped test for normality can be utilized to assess the normality of yield distributions by soil groups. The Appendix (C-1-C-6) gives the results of the group test for normality of corn and wheat, respectively. Evaluating the results of the tests for corn shows no significant evidence to reject the normality assumption.

Wheat, on the other hand, has 1 soil group out of 2 that is significant at the 5 percent level.

An interesting point can be made concerning the group test. Both soil groups have two farms that are significant at the 5 percent level but only one soil group has a significant Z value for the group test. The observable difference is that the soil group testing to reject normality had 1 farm with a very low W-statistic causing the test for normality to fail. In view of the sensitivity of the test statistic to one farm, there does not appear to be conclusive evidence to reject the hypothesis of normality for wheat yields in either soil group.

5.5 Test Of Equivalent Growth In Mean By Soil Group

The conclusions of yields being independent and normally distributed with constant variance from Chapter 4 have been evaluated by soil groups for corn and wheat, but consideration must be given to the stability of the parameters of the distribution by soil group. It is of interest to know whether the estimated means vary significantly in growth over time within a soil group. The two-variable linear regression model allowed each farm to be fitted with its own slope and intercept. In testing whether the growth in means are consistent within a soil group, all farms in the soil group are forced to have the same slope but each farm is allowed to have its own intercept.[19] The model for the second equation is as follows:

$$Y_{i,j,t} = \hat{a}_i D_i + b\text{Time} + e_{i,j,t}$$

---

[19] John J. Johnston, Econometric Methods, 3rd ed. (New York: McGraw-Hill, Inc., 1984) pp.212-214.

where:

$$D_i = \begin{matrix} 1 \text{ if farm } i \\ 0 \text{ otherwise} \end{matrix}$$

Dummy variables are used for each farm, thus, a is the estimated intercept for the i farm but all b's are constrained to be the same. The structure of the hypotheses state:

$$H_0 : b_1 = b_2 = b_3 \ldots = b_N$$

$$H_1 : b_1 = b_2 = b_3 \ldots = b_N$$

The test of the null hypothesis given by:

$$F = \frac{(SSR - SSUR)/N}{SSUR/\Sigma T_i - 2N} \quad F(N, \Sigma T_i - 2N)$$

where:

SSR is the sum of squares of residuals for the restricted model.

SSUR is the sum of squares of residuals for the unrestricted model.

N is the # of farms.

T is the # of observations for the i^{th} farm.

The Appendix (D-1-D-2) gives the F-tests of soil groups for corn and wheat, respectively. Soil group M5 for corn is the only soil group that accepts the null hypothesis of equivalent means. Judging from the results of the other tests for both corn and wheat and the fact that soil group M5 for corn has relatively few farms, it would be safe to conclude that the null hypothesis of equivalent growth in means is rejected at the 5 percent level of significance.

5.6 Test Of Equivalent Variance By Soil Group

It is also of interest to know whether the variances of each farm could be considered to be of the same size. Previously, variance has been analyzed by farm to determine if variation is constant over time. The intent with this test is to evaluate whether variance is constant across farms for a particular soil group. The hypotheses state:

H_o: $\text{variance}_{i,j} = \text{variance}_{2,j} = \ldots = \text{variance}_{I,j}$

H_1: $\text{variance}_{1,j} \neq \text{variance}_{2,j} \neq \text{variance}_{I,j}$

for I farms and J crops.

The test is as follows:

Step 1: Calculate $s^2_{i,j}$ for each farm given the j^{th} crop

where s^2 is the estimated variance.

Step 2: Calculate $s^2_j = (DF_{i,j} / DF_j)s^2_{i,j}$ (weighted variance)

where: $DF_{i,j} + (N_{i,j} - 1)$

$DF_j = DF_{i,j}$

Step 3: Calculate $Q^1 = DF_j \log_e s^2_j - DF_{i,j} \log_e s^2_{i,j}$

where: Q! is distributed as a Chi square with (i-1) degrees of freedom.

Step 4: A more accurate approximation is given by $A = 0^1/c$

where: $c = 1 + [1/(3II-1))][((1/DF_{i,j}) - (1/DF_j))]$

The tests results for each soil group are given in the Appendix (E-1-E-9) for corn and wheat, respectively. Although three of the six tests confirm the null hypothesis of constant variance across farms, the evidence seems to strongly support rejection of the null hypothesis in favor of measures of variance by individual farms for corn. Two of the three tests supporting the null hypothesis have less than 5 farms making the tests presumably less sensitive to variation. The last of the three tests confirming constant variance is very close to being significant at the 5 percent level of significance. The evidence supporting constant variance across farms is very weak as opposed to the alternative hypothesis. Wheat, on the other hand, appears to be supportive of constant variance across farms.

5.7 Summary Of Evaluation

Classification of farms by soil groups for corn and wheat attempted to evaluate whether there is any evidence to reject the hypothesis of normally distributed yields when farms are considered by their soil groupings. The assumption of yield independence holds for all soil groups for both corn and wheat. The constant variance assumption was found to hold for all soil groups and both crops except on the loamy soils for corn. The assumption of normality was supported for all soil groups for both crops except wheat grown on loamy soils. Consideration was also given in evaluating whether the means of the distributions for each farm in a soil group were growing at an equivalent rate. The null hypothesis of equivalent growth in the mean of the distribution was rejected at the 5 percent level. Variances of farms within a soil group were also analyzed to evaluate whether the variances are constant between farms. The evidence supporting constant variance is very weak for corn where wheat has supportive evidence of constant variance across farms.

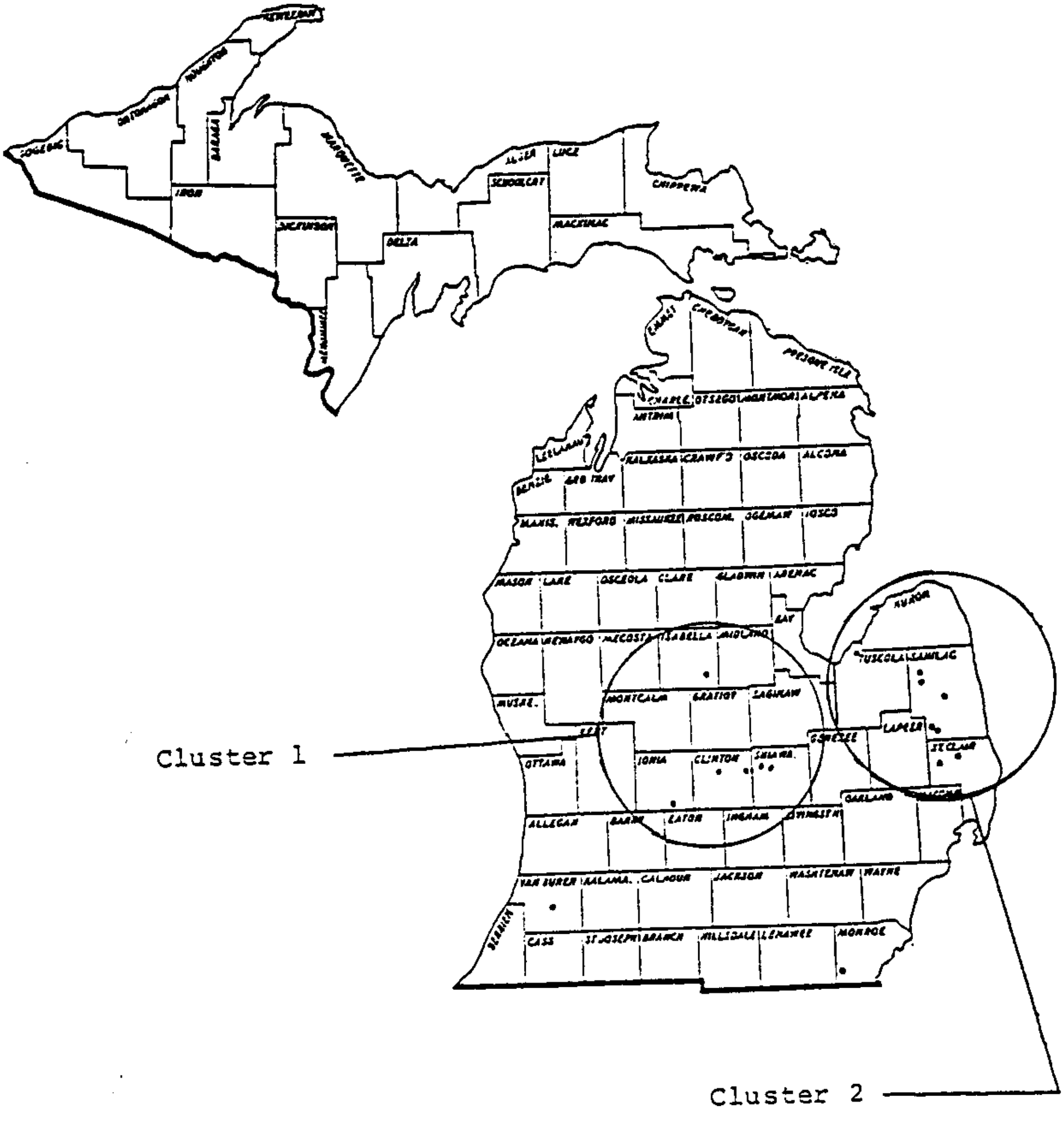
CHAPTER 6

EVALUATION OF FITTED DISTRIBUTIONS FOR A SOIL GROUP

In this chapter, probability density functions are fitted to yield data in an effort to evaluate whether distributions other than the Normal probability density function could more adequately describe the underlying yield distribution. The normality assumption has been supported by statistical investigation for each crop and by aggregate soil groupings for each crop. The purpose of this chapter is to bring the focus of investigation to the farm level for a specified soil group and crop in evaluating alternative distributions.

6.1 Selection Of Soil Group And Crop

Soil group M4 for corn was chosen for analysis as it represents one of the most productive soils in Michigan. This soil group is typified by predominantly wet loamy soils and is concentrated around the Saginaw Bay and thumb areas of Michigan. All farms analyzed have time series yield data from 1963-1983. The approximate location of each farm is given in Figure 6.1. Two clusters of farms emerge when observing the approximate locations. One group of farms is concentrated in Ionia, Clinton, Shiawassee, and Isabella counties. The second group is located in Tuscola, Sanilac, and St Clair counties. The two farms in Van Buren and Monroe counties are excluded from the analysis. It is of interest to see whether these two groupings of farms have obviously different



Cluster 1

Cluster 2

Figure 6.1
Farm Clusters For Distribution Analysis

probability density functions for the same crop and presumably the same dominate soil type.

6.2 Fitting Probability Distributions With Unifit

Unifit is an interactive computer package for fitting probability distributions to observed data. The computer package employs a three-activity approach for determining an appropriate distribution. The activities are: (1) hypothesize one or more families of distributions which might be appropriate, (2) estimate the parameters for each hypothesized distribution, and (3) determine which of the fitted probability models is the best representation of the data using goodness-of-fit tests and graphical displays.

Activity 1 has already been considered in Chapter 2. The distributions discussed in Chapter 2 were used as possible candidates in fitting a probability density function for the corn yield data of the 14 farms. The Descriptive Sample Summary function provided by Unifit was utilized to further decide on specific distributions to fit to the data. The options used in the Descriptive Summary routine provided summary statistics, histograms, and quantile summary and box plots. Summary statistics give the minimum-maximum observation, mean, median, variance, and the coefficients of variation, skewness, and kurtosis. Histograms give the relative frequencies of intervals covering the range of available data. The histogram is essentially a graphical representation of the plot for the density function. In the case of the data for the 14 farms that were being fitted with probability density functions, the number of observations is small and proves difficult to use the resulting histogram. However, this descriptive tool proved helpful in many cases. The quantile summary and box plot is a synopsis of the sample percentiles and is useful in determining whether the probability distribution is symmetric or asymmetric. An example is given in Table 6.1. The box plot confirms the presense of negative skewness.

Table 6.1
Quantile Summary And Box Plot Of Sample

```
SYMBOL AND                SAMPLE  VALUE(S)
 QUANTILE    DEPTH     (LOWER)        (UPPER)     MID-VALUES
----------  -------  --------------------------  ------------
M MEDIAN     10.5             102.957              102.957
Q QUARTILE    5.5     82.2357       116.510        99.3726
O OCTILE       3      67.4518       121.346        94.3990
E EXTREME      1      34.6729       127.336        81.0046


                                                        M
E                          O          Q-----------------M----------Q   O   E
E--------------------------O----------Q                 M          Q---O---E
E                          O          Q-----------------M----------Q   O   E
                                                        M
```

Activity 2 estimates the parameters for each specified probability distribution that is being fit to the data. Several models can be specified for a data sample allowing many models to be compared at the same time.

Activity 3 determines which fitted probability density function has the best fit using goodness-of-fit tests and heuristics. Goodness-of-fit tests formally examine whether a fitted distribution with distribution function F is a good representation of the observed data. The formal null hypothesis states:

H_0 : The X 's are independent and identically distributed random variables with distribution function $\hat{F}$.

The formal goodness-of-fit tests available with Unifit are the Chi-Square, Kolmogorov-Smirnov, and the Anderson-Darling tests. Unifit provides a listing of the results for goodness-of-fit tests of each model as shown in Table 6.2.

Table 6.2
Model Test Comparisons With Sample

```
THE CHI-SQUARE GOODNESS-OF-FIT TEST HAVING 3 INTERVALS.
EACH WITH EQUAL MODEL PROBABILITY 3.33333E-1.

                                                  KOLMOGOROV      ANDERSON
MODEL       DISTRIBUTION          CHI-SQUARE       -SMIRNOV        -DARLING
-----   --------------------   ---------------  --------------  ------------
  1     INVERSE GAUSSIAN            2.50000          .26082          1.52543
  2     LOGNORMAL                   1.30000          .24099          1.34047
  3     WEIBULL                     1.30000          .16591           .66441
  4     EXTREME VALUE TYPE A        3.10000          .13587           .35535
  5     EXTREME VALUE TYPE B        1.30000          .21626          1.13475
  6     LOGISTIC                    1.30000          .13637           .58404
  7     NORMAL                      1.30000          .16165           .65476
  8     BETA                        1.30000          .18952           .74466
```

This provides the opportunity to compare model test results in choosing the best models being fitted to the data. The approach taken in evaluating the best probability density function being fitted to the yield data was to choose the best two models from the model goodness-of-fit test comparisons. These two models were then evaluated by heuristics to determine the best fit.

Heuristics utilizes graphical techniques in comparing the sample and a hypothesized probability distribution to determine how well a fitted distribution represents the observed data. The heuristic methods used in this study was the cumulative frequency comparison, quantile-quantile plot, and the probability-probability plot. The cumulative frequency comparison is a comparison between a sample distribution function which is computed from the observed data and the distribution function of a fitted distribution. The sample distribution function is an approximation to the distribution function of the underlying distribution of the observed data as was discussed in Section 4.7. Figure 6.2 shows a typical cumulative frequency comparison by Unifit.

The quantile-quantile (Q-Q) plot and probability-probability (P-P) plot are graphical displays designed to amplify certain differences between the sample distribution function $F_n(X)$ and the fitted distribution function F(X). The definition of a Q-Q plot is illustrated in Figure 6.3. Corresponding to each ordinate value q are the two quantiles X_q^s and X_q^m . If the family of distributions corresponding to F(X) is the same as the family of distributions corresponding to the true underlying distribution, the F(X) and F_n (X) will be close together and the Q-Q plot will be approximately linear with an intercept of zero and a slope of 1 (see Figure 6.4).

The P-P plot is a graph of the model probability versus the sample probability. This definition is also illustrated in Figure 6.3. If F(X) and $F_n(X)$ are close together, then the P-P plot will also be approximately linear with an intercept of zero and a slope of 1 (see Figure 6.5).

The Q-Q plot will amplify differences which exist between the tails of the model distribution function F(X) and the tails of the sample distribution function F_n (X), whereas the P-P plot will amplify differences between the middle of both distributions. This is illustrated in Figures 6.6 and 6.7.

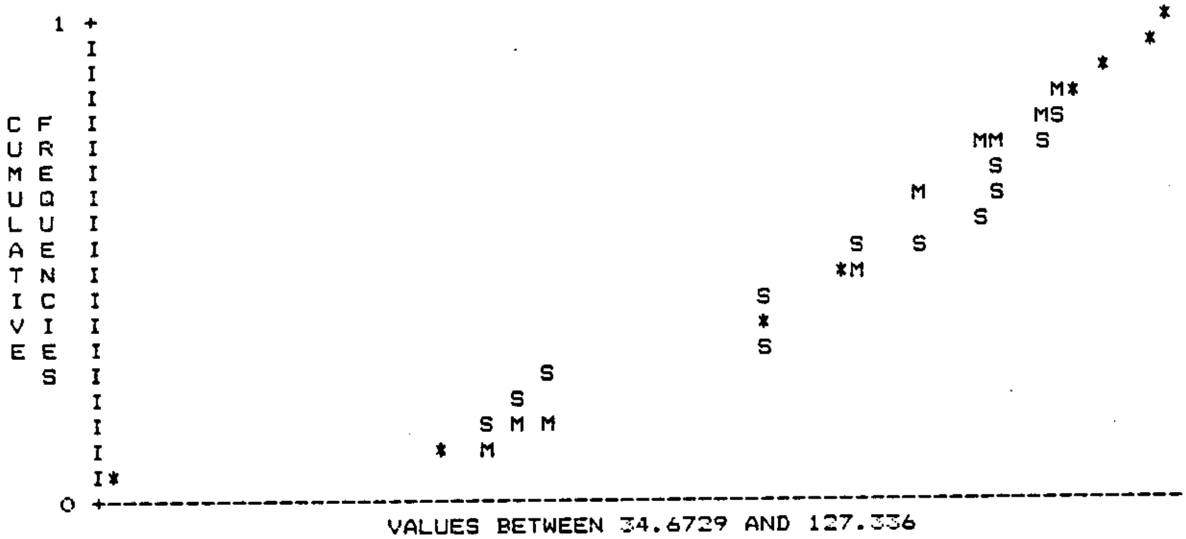


Figure 6.2
Cumulative Frequency Comparison Of Model 4 Extreme Value Type A

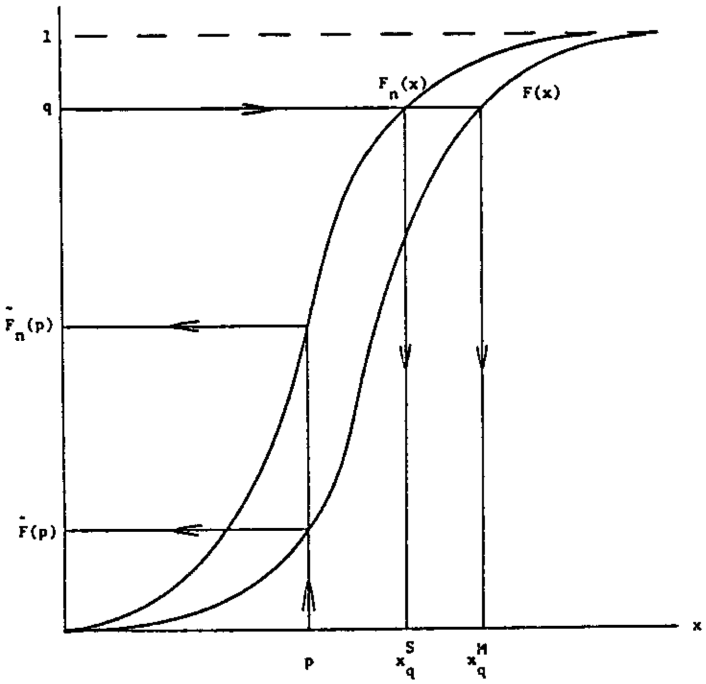


Figure 6.3
Definition Of The Q-Q And P-P Plots

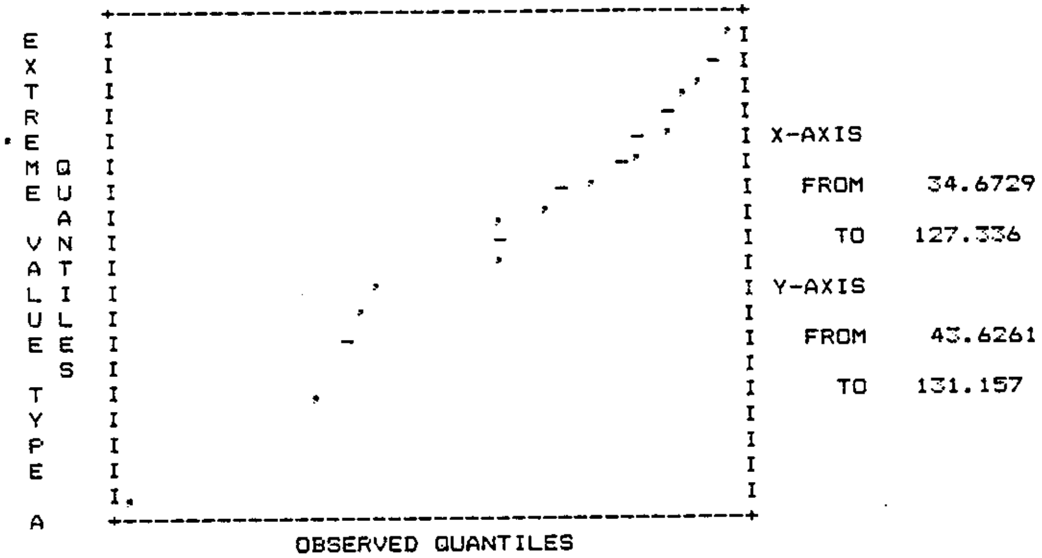


Figure 6.4
Q-Q Plot Of Model 4 And Sample

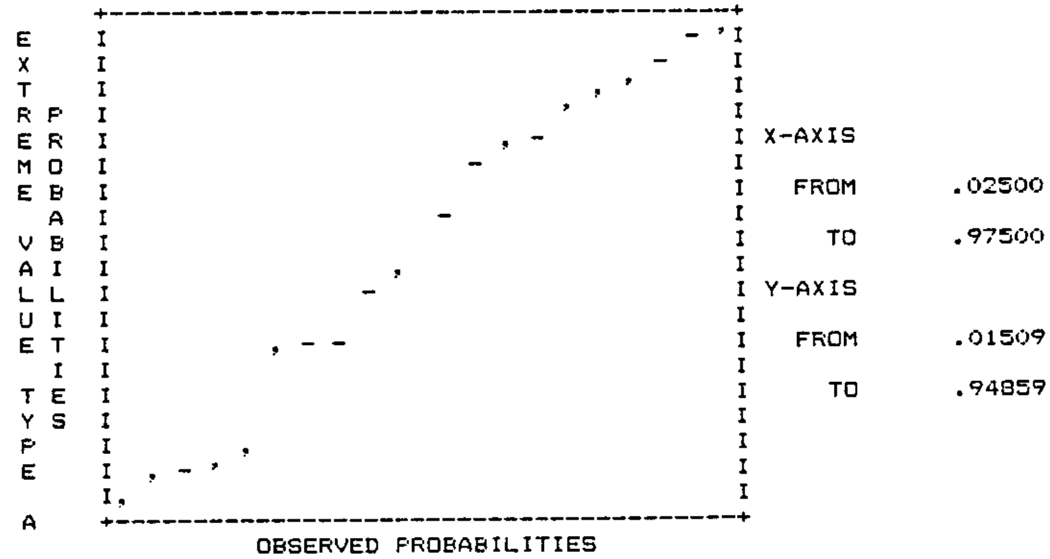


Figure 6.5
P-P Plot Of Model And Sample

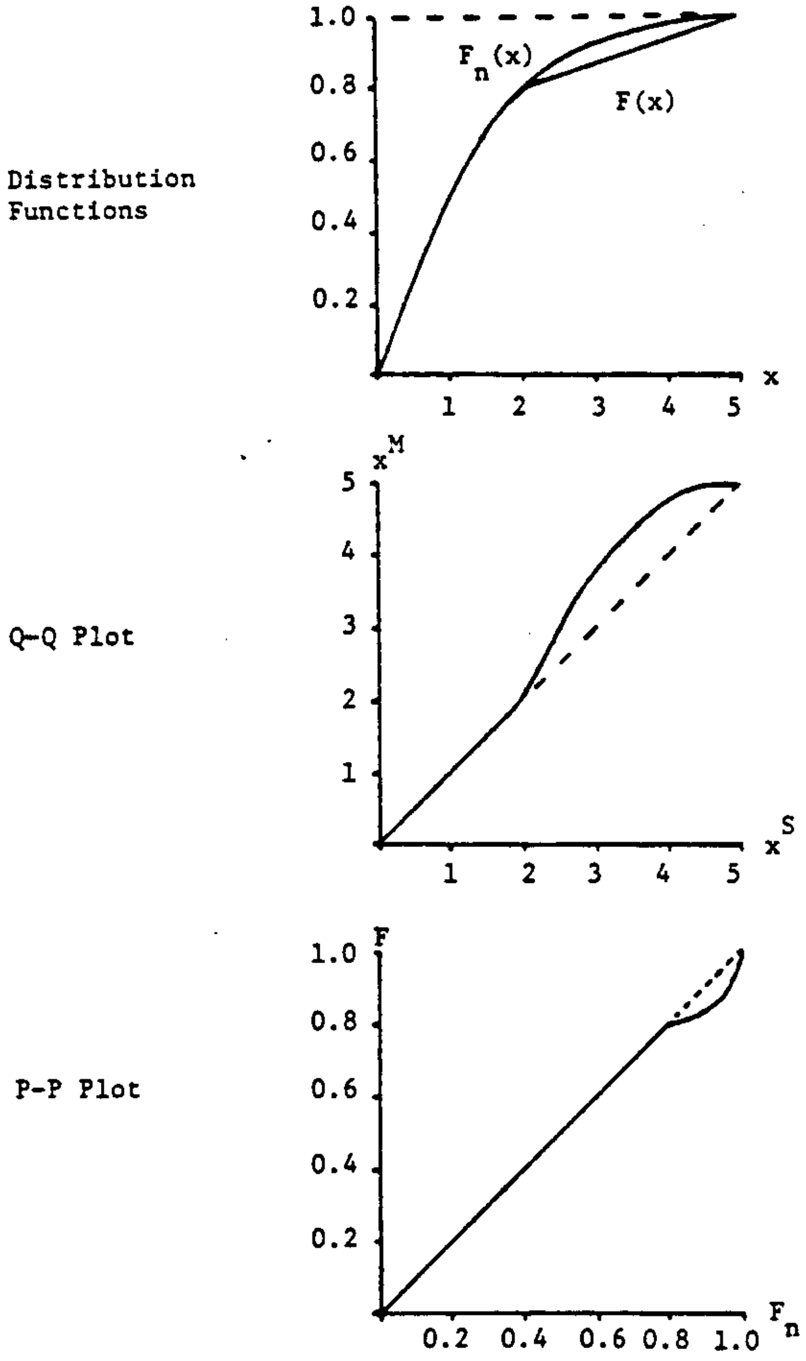


Figure 6.6
The Difference Between The Right Tails Of F(x) And $F_n(x)$
Amplified By The Q-Q Plot.

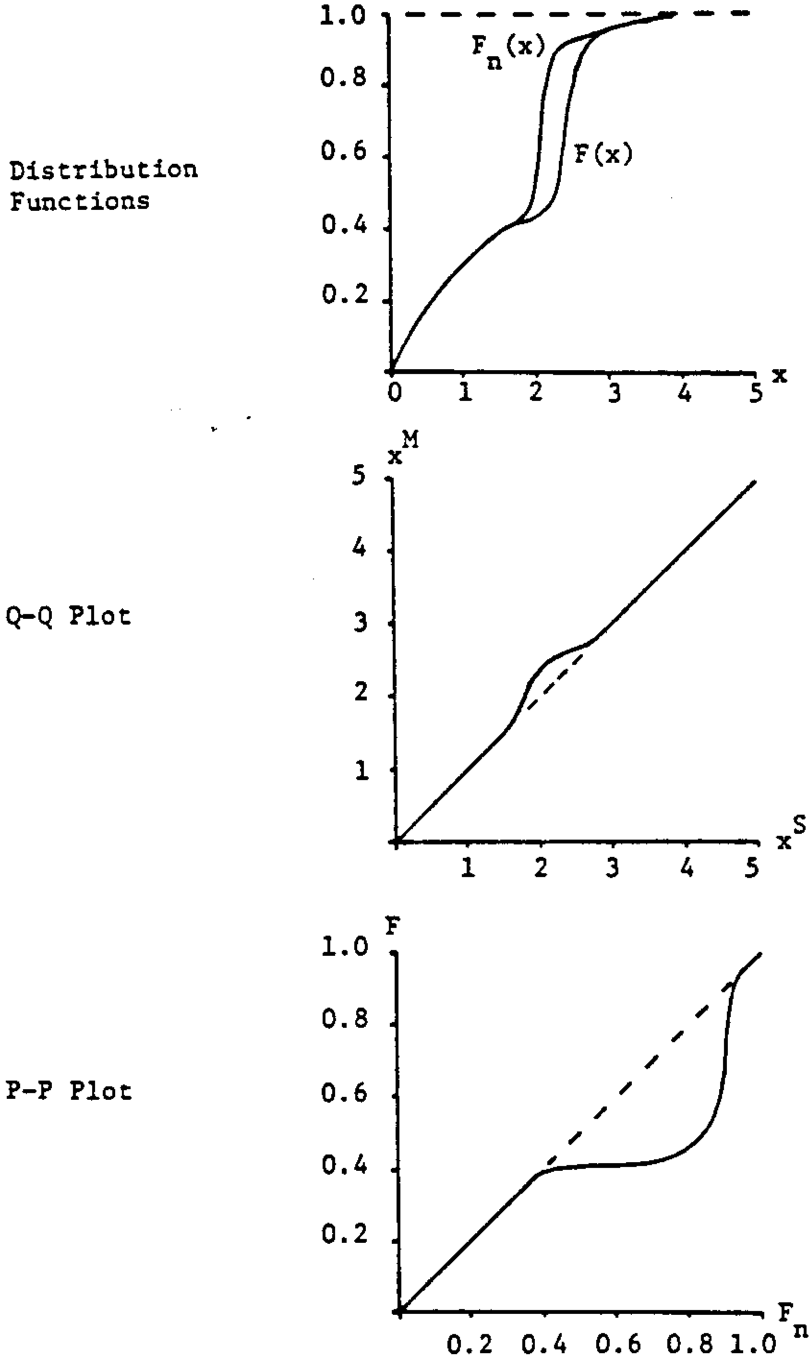


Figure 6.7
The Difference Between The "Middles" Of F(x) And $F_n(x)$
Amplified By The P-P Plot.

6.3 Evaluation Of The Results

Fourteen of the 16 farms were fitted with distributions hypothesized in Chapter 2. From the original eight distributions that were hypothesized, three probability density functions were dominate in describing the yield data. Table 6.3 gives each of the the farms and the probability density function that best fits the data. The three probability distributions that emerge are the Normal, Beta, and the Extreme Value Type A. The Normal seems to best fit farms with relatively small amounts of skewness and kurtosis as would be seemingly appropriate. The Beta distribution is best fit to data which has relatively small amounts of skewness but significant amounts of negative kurtosis even though the Shapiro-Wilks test seems to indicate normality for the sample. The Extreme Value Type A distribution is seemingly best fit to yield data with relatively significant amounts of negative skewness and kurtosis. This situation is usually caused by a negative outlying data value. All the farms fitted with a Extreme Value Type A probability distribution are characterized by a relatively large negative outlying residual as compared with the positive residuals.

Evaluating distributions fit to farms by geographical area did not show supportive evidence of one distribution being dominate for either cluster of farms.

6.4 Summary Of Evaluation

Soil group M4 (wet loamy soils) for corn was chosen to assess more closely the functional form of the probability density function at the farm level for supportive evidence of the normality conclusion. It was found that farms with fairly symmetric data values around the mean but

Table 6.3
Fitted Distributions For Soil Group M^{4}

Cluster 1

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST	FITTED DISTRIBUTION
190027	63-83	0	21	-43.1	34.7	.17	17.575	-.419	.609	.97341	EXTREME VALUE TYPE A
190042	63-83	81	20	-62.7	29.9	.25	24.660	-1.001	.548	.90798	EXTREME VALUE TYPE A
340560	63-83	82,83	19	-43.3	30.2	.24	22.140	-.579	-.912	.92537	EXTREME VALUE TYPE A
370146	63-83	83	20	-38.6	46.9	.23	24.533	.273	-.474	.95653	NORMAL
760086	63-83	76,83	19	-31.7	32.9	.22	19.075	-.148	-.913	.96415	BETA
760540	63-83	83	20	-30.8	23.6	.15	16.721	-.185	-1.083	.94452	BETA

Cluster 2

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST	FITTED DISTRIBUTION
740077	63-83	64,77	19	-40.1	33.7	.16	19.963	-.493	.137	.95872	NORMAL
740087	63-83	79	20	-19.4	24.7	.14	12.720	.115	-.668	.96974	BETA,NORMAL
740099	63-83	0	21	-31.6	21.6	.12	16.137	-.416	-.841	.94640	BETA
740167	63-83	81,83	19	-45.9	35.3	.18	21.172	-.413	-.539	.94081	BETA,NORMAL
740169	63-83	80	20	-50.2	25.8	.18	19.437	-.852	.810	.94227	EXTREME VALUE TYPE A
770005	63-83	0	21	-24.7	24.1	.12	13.843	.015	-.767	.96299	BETA
770626	63-83	81,82,83	18	-28.3	18.9	.15	13.486	-.518	-.363	.93648	EXTREME VALUE TYPE A
790115	63-83	79	20	-36.1	23.7	.10	14.945	-.536	.212	.96542	EXTREME VALUE TYPE A

having significant amounts of negative kurtosis could more accurately be described by a Beta probability density function. Data sets with negative outliers were found to be best fit with a Extreme Value Type A probability distribution.

CHAPTER 7

APPLICATION FOR FARM MANAGEMENT

One approach that has been taken by agricultural economists and extension professionals in assisting producers to make decisions in a risky environment has included the use of expected values. Agricultural economists need to provide farm managers with improved decision support tools that approximate real world conditions. Ikerd and Anderson[20] feel that farm managers need estimates of the probability favoring a profit or loss and the ability to compare one risky alternative with another. For purposes of decision-making, the grower needs an estimate for the probability of the business covering variable costs, meeting debt service commitments, earning an acceptable living, and meeting long-term financial growth objectives. Planning information which does not provide answers to these questions is not sufficient for effective decision-making.

Farm management decision-making could benefit from research involving estimation of the functional form of the probability distribution and parameters for random variables that impact production risk, price risks, financial risks, etc. As equally important, research

---

[20] John Ikerd and Kim Anderson. "Teaching Risk Related Management Strategies to Farmers and Ranchers," Presentation to Workshop For Extension Specialists on Marketing, Risk, and Financial Management (Minneapolis, MN, n.p., 1984).

is also needed in applying the estimated functional form of a probability distribution and its parameters into the farm planning framework for decision-making purposes.

The primary objective of this research project has been to estimate the continuous probability distributions for corn, wheat, and soybean yields in an attempt to quantify the production risk for these crops grown in Michigan. A descriptive approach was taken in analyzing the behavior of yield variability for data taken from actual farm records. Research results of this study indicate that the Normal probability distribution adequately describes the production risk for corn, wheat, and soybeans. Using the conclusions of this descriptive analysis, it is proposed that this information can be incorporated into a predictive analysis framework as a practical application for decision-making purposes. The applications that will be considered in this chapter will be: (1) short-term planning; (2) long-range planning; (3) forward pricing; and (4) crop insurance.

7.1 Short-Term Planning

Short-term planning for this research project is taken in the context of a 1-2 year planning horizon. The formal approach in planning for short-term decisions is commonly labeled as forward planning. Forward planning is used to predict the expected outcomes of financial performance for the farm business under different assumptions regarding future production, price, and financial situations. This type of planning tool allows the producer to ask "what would happen if ..." types of questions.

The incorporation of probability distributions into forward planning tools gives recognition to the range and variance of results that might happen in the future. An analysis of the yield variation for past years can be used to provide more accurate direction for what might happen with future yield variance. In other words the "what would happen if" questions for production risk are being answered from what has happened in the past. It is assumed that the nature of variability from the past will continue in the future if no fundamental structural change occurs that would alter the production environment.

Based on the results given in the previous chapters, the Normal distribution for corn, wheat, and soybean yields can be used to model the nature of variability for any particular year. If data for a particular farm to be modeled is available, the parameters of the distribution can be estimated and the validity of the normality assumption can be evaluated. The observations from Chapter 6 could be applied for those farms that appear to be non-normal. If data are not available, the averages for the mean and standard deviation presented in Tables 7.1, 7.2, and 7.3 could be used as a best estimation. It should be noted that the mean of the distribution should be adjusted for a particular year by the slope of the fitted line[21] considered in Chapter 4.

The forward planning tools being considered are the Enterprise budget and the Whole-farm budget.

---

[21] $Y = A + B\ (\text{Time})$.

Table 7.1
Summary Statistics For Corn

County	# Farms	Yield Increase Per Year (bu/acre)	Yield Average (bu/acre)	Standard Deviation (bu/acre)	Coefficient of Variation (%)
Allegan	4	1.230	103.1	13.330	0.129
Antrim	1	-0.024	79.8	21.361	0.268
Barry	2	1.730	104.8	15.019	0.143
Berrien	1	3.060	138.0	23.685	0.172
Branch	4	1.600	106.1	17.817	0.168
Calhoun	3	1.770	116.5	15.724	0.135
Clinton	3	0.907	96.6	19.383	0.201
Eaton	1	1.530	99.5	14.048	0.141
Emmet	1	1.720	88.3	26.560	0.301
Gratiot	1	1.960	130.4	17.506	0.134
Hillsdale	4	1.130	98.9	14.162	0.143
Ingham	4	2.090	98.9	16.982	0.172
Ionia	4	1.430	100.7	19.624	0.195
Isabella	2	0.818	99.5	21.677	0.218
Jackson	5	2.070	111.4	15.461	0.139
Kalamazoo	2	0.885	77.7	16.304	0.210
Kent	2	0.975	96.0	12.476	0.130
Lapeer	2	2.340	117.3	17.553	0.150
Lenawee	3	1.920	101.2	12.586	0.124
Livingston	2	1.320	102.9	15.530	0.151
Mason	2	0.950	77.4	15.763	0.204
Mecosta	1	-0.163	64.1	18.749	0.292
Monroe	1	1.920	131.9	19.625	0.149
Montcalm	4	2.910	114.2	17.814	0.156
Muskegon	2	1.040	86.5	15.972	0.185
Oakland	1	0.761	86.6	19.208	0.222
Oceana	1	2.220	108.3	18.681	0.172
Oscola	1	0.310	66.8	14.790	0.221
Sanilac	7	1.770	109.3	16.088	0.147
Shiawassee	3	1.820	109.2	17.927	0.164
St.Clair	2	0.572	102.5	13.665	0.133
St.Joseph	2	2.650	132.0	20.341	0.154
Tuscola	3	1.770	129.2	15.893	0.123
VanBuren	3	0.505	82.6	15.435	0.187
Washtenaw	3	1.430	96.4	17.050	0.177
AVERAGE *	87	1.546	103.353	16.709	0.165

* calculated using a weighted average

Table 7.2
Summary Statistics For Wheat

County	# Farms	Yield Increase Per Year (bu/acre)	Yield Average (bu/acre)	Standard Deviation (bu/acre)	Coefficient of Variation (%)
Barry	2	-0.260	43.6	9.340	0.214
Branch	1	0.567	49.1	6.375	0.130
Calhoun	2	0.490	52.6	5.786	0.110
Clinton	1	0.440	55.3	5.788	0.105
Eaton	1	0.310	40.2	9.141	0.227
Gratiot	1	0.390	52.1	11.506	0.221
Ionia	2	0.288	51.2	8.959	0.175
Jackson	4	0.373	48.7	7.567	0.155
Kalamazoo	2	0.490	46.3	6.922	0.150
Mason	1	0.149	44.6	5.540	0.124
Mecosta	1	0.081	32.1	11.482	0.358
Missaukee	1	0.872	31.6	11.337	0.359
Monroe	1	1.010	72.0	7.220	0.100
Montcalm	4	0.606	46.3	9.361	0.202
Muskegon	2	-0.222	46.0	7.557	0.164
Newaygo	1	0.603	49.2	5.901	0.120
Sanilac	1	0.637	57.2	10.902	0.191
Shiawassee	1	0.360	42.1	9.343	0.222
St.Joseph	2	-0.228	38.9	8.393	0.216
Tuscola	2	0.282	52.8	11.204	0.212
VanBuren	1	-0.781	28.6	8.957	0.313
Washtenaw	1	-0.001	34.7	9.117	0.263
AVERAGE *	35	0.292	46.6	8.476	0.190

* calculated using a weighted average

Table 7.3
Summary Statistics For Soybeans

County	# Farms	Yield Increase Per Year (bu/acre)	Yield Average (bu/acre)	Standard Deviation (bu/acre)	Coefficient of Variation (%)
Branch	1	-0.480	23.2	6.237	0.269
Gratiot	1	0.266	33.9	4.870	0.144
Hillsdale	1	0.190	29.1	9.189	0.316
Shiawassee	1	1.570	35.2	5.318	0.151
St.Clair	2	0.532	29.5	6.732	0.228
St.Joseph	1	0.706	26.3	7.698	0.293
AVERAGE *	7	0.474	29.5	6.682	0.233

* calculated using a weighted average

Enterprise Budgets

Enterprise budgets present the gross revenue and variable costs on a per unit basis for a specific production activity. The primary purpose of an enterprise budget is to calculate the gross margin, i.e., the difference between income and selected cash expenses. In determining the gross margin of an enterprise, it is directly affected by the production per unit, output price received, and the quantity and costs of inputs needed to create the output. Knowledge of the variability of factors comprising the enterprise budget will give the decision-maker added insight into the variation of gross margin per unit for planning purposes. The extent of this research study will consider only the production risk of the enterprise budget.

Taking the parameters estimated for variability of corn yield in Berrien County from Table 7.1, a hypothetical enterprise budget for corn is constructed in Table 7.4. Table 7.4 shows the effect of the estimated variation in corn yield on the gross margin of the corn enterprise. The expected value column is the most likely yield times the price. One standard deviation in yield (23.685 bu.) will increase or decrease the gross margin +$52.46 or -$52.46, respectively given these price and cost assumptions. Yields two standard deviations from the expected yield average would vary the gross margin +/- $104.92 accordingly. The $P[X \leq G.M.]$ gives the cumulative probability that the gross margin could be less than the predicted value. The probabilities are based on the Normal distribution. Notice that the probability of the gross margin being in the range $91.85 - 301.69 is 95.4 percent ((.977 - .023) * 100). There is a 2.3 percent possibility that the gross margin could be below $91.85 and a 2.3 percent possibility that

Table 7.4
Corn Enterprise Budget - Berrien County

			-2 S.D.	-1 S.D.	Exp Val	+1 S.D.	+2 S.D.
		Yield	91	114	138	162	185
Income	Yield	Price					
Corn	138	2.75	249.23	314.37	379.50	444.63	509.77
Std Dev	23.685						
Variable Costs							
Seed			21.00	21.00	21.00	21.00	21.00
Fertilizer			44.70	44.70	44.70	44.70	44.70
Chemicals			15.20	15.20	15.20	15.20	15.20
Repairs-mach			18.00	18.00	18.00	18.00	18.00
Gas,Fuel,Oil			10.00	10.00	10.00	10.00	10.00
Drying			27.19	34.29	41.40	48.51	55.61
Utilities			2.27	2.86	3.45	4.04	4.63
Marketing			.91	1.14	1.38	1.62	1.85
Trucking			18.13	22.86	27.60	32.34	37.07
Total Cash Expense			157.39	170.06	182.73	195.40	208.07
Gross Margin			91.85	144.31	196.77	249.23	301.69
Prob[random variable] < G.M.			.023	.159	.500	.841	.977

the gross margin could be above $301.69. Figure 7.1 gives a pictorial representation of the probability distribution of the gross margin and Figure 7.2 gives a pictorial representation of the cumulative probability distribution for gross margin. Having an estimation of the nature and size of variation in gross margin due to yield risk for an enterprise can be a useful tool for a farm manager making enterprise selection decisions and developing strategies for enterprise improvement.

The enterprise budget is also useful in comparing farms when considering variations in gross margin. Consider for example (see Table 7.5) an enterprise budget constructed using yield variation parameters estimated for Shiawassee County. Assuming the costs of production are relatively the same in producing for an above average yield, the gross margins could be compared between counties. Berrien County has a higher average yield, but also a larger standard deviation in yield where Shiawassee County has both a lower average yield and standard deviation in yield. Because of the higher average yield, the expected gross margin of Berrien County is $63.79 larger than Shiawassee County. Shiawassee County has an advantage in that the gross margin will decrease only $39.71 for each standard deviation in yield where Berrien County will deviate $52.46. But as is the nature of risk, the upside potential for Shiawassee County is not as attractive as for Berrien County. After studying the downside risk and the range of possible outcomes in gross margin, Berrien County shows a better contribution from the corn enterprise than Shiawassee County. Berrien County has a 95.4 percent possibility of returning as little as $91.85 to as much as $301.69 above variable costs as contrasted to a $53.56 to $212.39 range

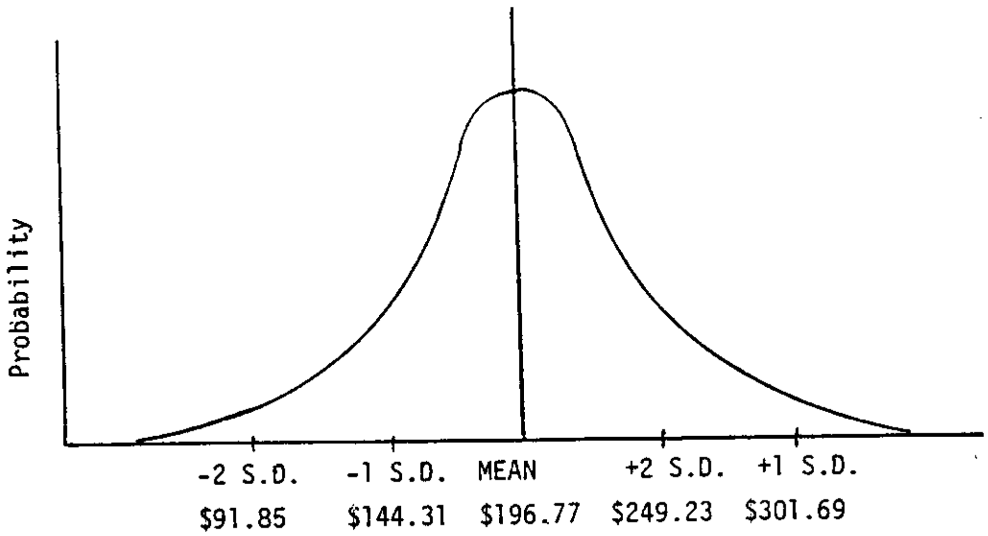


Figure 7.1
Probability Distribution Of Gross Margin Corn Produced
For Berrien County

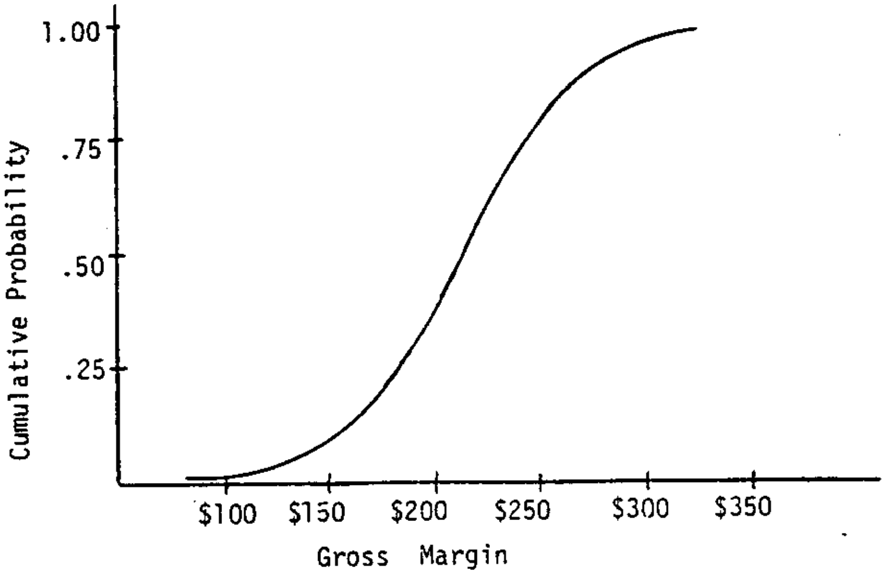


Figure 7.2
Cumulative Probability Distribution Of Gross Margin Corn Produced
For Berrien County

Table 7.5
Corn Enterprise Budget - Shiawassee County

			-2 S.D.	-1 S.D.	Exp Val	+1 S.D.	+2 S.D.
		Yield	73	91	109	127	145
Income	Yield	Price					
Corn	109.2	2.75	201.70	251.00	300.30	349.60	398.90
Std Dev	17.927						
Variable Costs							
Seed			21.00	21.00	21.00	21.00	21.00
Fertilizer			44.70	44.70	44.70	44.70	44.70
Chemicals			15.20	15.20	15.20	15.20	15.20
Repairs-mach			18.00	18.00	18.00	18.00	18.00
Gas,Fuel,Oil			10.00	10.00	10.00	10.00	10.00
Drying			22.00	27.38	32.76	38.14	43.52
Utilities			1.83	2.28	2.73	3.18	3.63
Marketing			.73	.91	1.09	1.27	1.45
Trucking			14.67	18.25	21.84	25.43	29.01
Total Cash Expense			148.14	157.73	167.32	176.91	186.50
Gross Margin			53.56	93.27	132.98	172.69	212.39
Prob[random variable] < G.M.			.023	.159	.500	.841	.977

- corn price is an average breakeven figure for non-partisipation in the set-aside program for Shiawassee County. Considering the downside risk of a given level of gross margin to cover fixed costs and family living, Berrien County has a smaller possibility of being under that level of margin.

Having an understanding of variation in the yield of a crop enterprise can provide information useful for financial planning and developing strategies to reduce risk in meeting financial objectives. Many questions can be raised from studying the impact of the estimated variation in corn yield on enterprise returns for the two counties. To illustrate , is the difference in yield averages between counties attributed to soil type differences, variations in management approaches, or both? Does the magnitude of yield variation (standard deviation) depend on soil type, tillage practices, management decisions, or a combination of all three? What makes yield variation increase over time? Whether these questions can be answered by further research or not, the fact remains that if the size and nature of the variation can be estimated from available data, much information can be gained for farm planning purposes.

The same approach of determining gross margin variation can also be used in evaluating wheat and soybean enterprises. Tables 7.6 and 7.7 give an estimation of the variation in gross margin for wheat grown in Clinton and Sanilac counties, respectively. From Table 7.2 we find that the average wheat yield estimated for Clinton county is 55.3 bushels/year, and has a standard deviation of 5.788 bushels as presented in Table 7.6. The expected value of the gross margin is $107.36 and varies by $21.36 for each standard deviation in yield. The average yield of 57.2 bushels/year for Sanilac county is almost the same as Clinton county, but the variation in yield is almost double at 10.902

Table 7.6
Wheat Enterprise Budget - Clinton County

			-2 S.D.	-1 S.D.	Exp Val	+1 S.D.	+2 S.D.
		Yield	44	50	55	61	67
Income	Yield	Price					
Wheat	55.3	3.90	170.52	193.10	215.67	238.24	260.82
Std Dev	5.788						
Variable Costs							
Seed			12.00	12.00	12.00	12.00	12.00
Fertilizer			59.10	59.10	59.10	59.10	59.10
Chemicals			1.00	1.00	1.00	1.00	1.00
Repairs-mach			16.00	16.00	16.00	16.00	16.00
Gas,Fuel,Oil			7.10	7.10	7.10	7.10	7.10
Utilities			1.50	1.50	1.50	1.50	1.50
Marketing			.44	.50	.55	.61	.67
Trucking			8.74	9.90	11.06	12.22	13.38
Total Cash Expense			105.88	107.10	108.31	109.53	110.74
Gross Margin			64.64	86.00	107.36	128.71	150.07
Prob[random variable] < G.M.			.023	.159	.500	.841	.977

Table 7.7
Wheat Enterprise Budget - Sanilac County

			-2 S.D.	-1 S.D.	Exp Val	+1 S.D.	+2 S.D.
		Yield	35	46	57	68	79
Income	Yield	Price					
Wheat	57.2	3.90	138.04	180.56	223.08	265.60	308.12
Std Dev	10.902						
Variable Costs							
Seed			12.00	12.00	12.00	12.00	12.00
Fertilizer			59.10	59.10	59.10	59.10	59.10
Chemicals			1.00	1.00	1.00	1.00	1.00
Repairs-mach			16.00	16.00	16.00	16.00	16.00
Gas,Fuel,Oil			7.10	7.10	7.10	7.10	7.10
Utilities			1.50	1.50	1.50	1.50	1.50
Marketing			.35	.46	.57	.68	.79
Trucking			7.08	9.26	11.44	13.62	15.80
Total Cash Expense			104.13	106.42	108.71	111.00	113.29
Gross Margin			33.91	74.14	114.37	154.60	194.82
Prob[random variable] < G.M.			.023	.159	.500	.841	.977

- wheat price is an average breakeven figure for non-partisipation in the set-aside program bushels for each standard deviation (see Table 7.7). Enterprise budgets for soybeans can be found in Tables 7.8 and 7.9. The budgets were constructed from yield and variation estimates for Shiawassee and Gratiot Counties (see Table 7.3).

Whole-Farm Budgets

Whole-farm budgets are a combination of the enterprise budgets for the farm business and the fixed costs or other non-enterprise related expenses. The whole-farm budget gives the expected net farm cash flow prediction for the year. Incorporating the production variability for relevant enterprises will give the decision-maker an appreciation for the variability of net cash flow due to variation in yields.

The Agricultural Risk Management Simulator (ARMS), a microcomputer program that evaluates strategies for managing risk, will be used to evaluate the variability of net cash flow. ARMS was designed to be used by farmers, lenders, and farm management advisors as a tool for whole farm planning and for evaluating opportunities to manage risk through the purchase of crop insurance and/or the use of forward contracting. ARMS will be used by this research study to evaluate the variability of net cash flow for the whole-farm budget and using the variability of net cash flow as a basis of comparison for different management strategies.

One reason for considering whole-farm budgeting is to ascertain net cash flow for the farm business. Incorporating the variation of yields into the whole-farm budget provides a probabilistic distribution of different net cash flows. Table 7.10 presents information for a case farm, which is assumed to be a cash crop farm of 500 acres growing corn, wheat, and soybeans in rotation and participates in the set-aside

Table 7.8
Soybean Enterprise Budget - Shiawassee County

			-2 S.D.	-1 S.D.	Exp Val	+1 S.D.	+2 S.D.
		Yield	25	30	35	41	46
Income	Yield	Price					
Soybeans	35.2	6.00	147.38	179.29	211.20	243.11	275.02
Std Dev	5.318						
Variable Costs							
Seed			12.00	12.00	12.00	12.00	12.00
Fertilizer			15.50	15.50	15.50	15.50	15.50
Chemicals			17.90	17.90	17.90	17.90	17.90
Repairs-mach			16.00	16.00	16.00	16.00	16.00
Gas,Fuel,Oil			9.00	9.00	9.00	9.00	9.00
Utilities			1.50	1.50	1.50	1.50	1.50
Marketing			.37	.45	.53	.61	.69
Trucking			4.91	5.98	7.04	8.10	9.17
Total Cash Expense			77.18	78.32	79.47	80.61	81.75
Gross Margin			70.20	100.97	131.73	162.50	193.26
Prob[random variable] < G.M.			.023	.159	.500	.841	.977

Table 7.9
Soybean Enterprise Budget - Gratiot County

			-2 S.D.	-1 S.D.	Exp Val	+1 S.D.	+2 S.D.
		Yield	24	29	34	39	44
Income	Yield	Price					
Soybeans	33.9	6.00	144.96	174.18	203.40	232.62	261.84
Std Dev	4.87						
Variable Costs							
Seed			12.00	12.00	12.00	12.00	12.00
Fertilizer			15.50	15.50	15.50	15.50	15.50
Chemicals			17.90	17.90	17.90	17.90	17.90
Repairs-mach			16.00	16.00	16.00	16.00	16.00
Gas,Fuel,Oil			9.00	9.00	9.00	9.00	9.00
Utilities			1.50	1.50	1.50	1.50	1.50
Marketing			.36	.44	.51	.58	.65
Trucking			4.83	5.81	6.78	7.75	8.73
Total Cash Expense			77.09	78.14	79.19	80.24	81.28
Gross Margin			67.87	96.04	124.21	152.38	180.56
Prob[random variable] < G.M.			.023	.159	.500	.841	.977

Table 7.10

CROP PRODUCTION COSTS				
Crop Name	CORN	WHEAT	SOYBEANS	SETASIDE
Production Unit	bu.	bu.	bu.	bu.
Preharvest Variable Costs ($/ac)	108.90	96.70	71.90	0.00
Harvest Costs:[2] Per Acre Component ($/ac)	4.50	4.00	4.00	0.00
Harvest Costs:[3] Per Unit Component ($/unit)	0.53	0.21	0.21	0.00

FARM SIZE AND OVERHEAD EXPENSES

Total Tillable Acres: 500
Total Overhead Expenses: 50000.00

FEED GRAIN PROGRAM ASSUMPTIONS

Crop to plnt	Base acres	ASCS yield	Def pmt	Total pmt	Loan rate
Corn	350	120	$1.16	$38,976(280A)	$1.77/bu
Wheat	75	40	$1.23	$2,656(54A)	$3.00/bu

CROP MIX

	planted	set-aside
Corn	280	70
Wheat	54	21
Soybeans	75	--

---

[2] Labor, Fuel, and Parts.

[3] Drying and Hauling.

program for feed grains. Crop rotation consists of wheat after soybeans and corn after wheat. Enterprise budget figures for corn, wheat, and soybeans are taken from Tables 7.4, 7.6, and 7.8, respectively. Using the yield parameters from Tables 7.4, 7.6, and 7.8 for corn, wheat, and soybeans, ARMS estimates a yield distribution with approximately the same parameters (see Table 7.11). The estimated deficiency payments are given as a per acre yield for set-aside land. Table 7.12 summarizes the price distributions. The minimum price for corn and wheat is set by the government program loan rate with upward movement depicted by subjective probabilities. The price variation for soybeans is a subjective forecast. Histograms for the price distribution assumptions for corn, wheat, and soybeans are given in Figures 7.3, 7.4, and 7.5, respectively. Figure 7.6 depicts the percentage of the maximum set-aside payment that could be received. The set-aside payment is the difference between the target price and the average market price. The correlation between the set-aside payments and prices for corn and wheat are summarized in Table 7.13. Correlations are also given for yield with yield.[x] The crop mix is given in Table 7.14.

Once all the necessary assumptions are made, ARMS simulates random draws from the given assumptions and then summarizes the information in terms of net cash flow. Table 7.15 gives a cumulative distribution of net cash flow for the case farm example. Figure 7.7 depicts the same cumulative distribution in graphical form. The cumulative probability given in Figure 7.7 is the likelihood that the actual net cash flow will be less than the given amount.

---

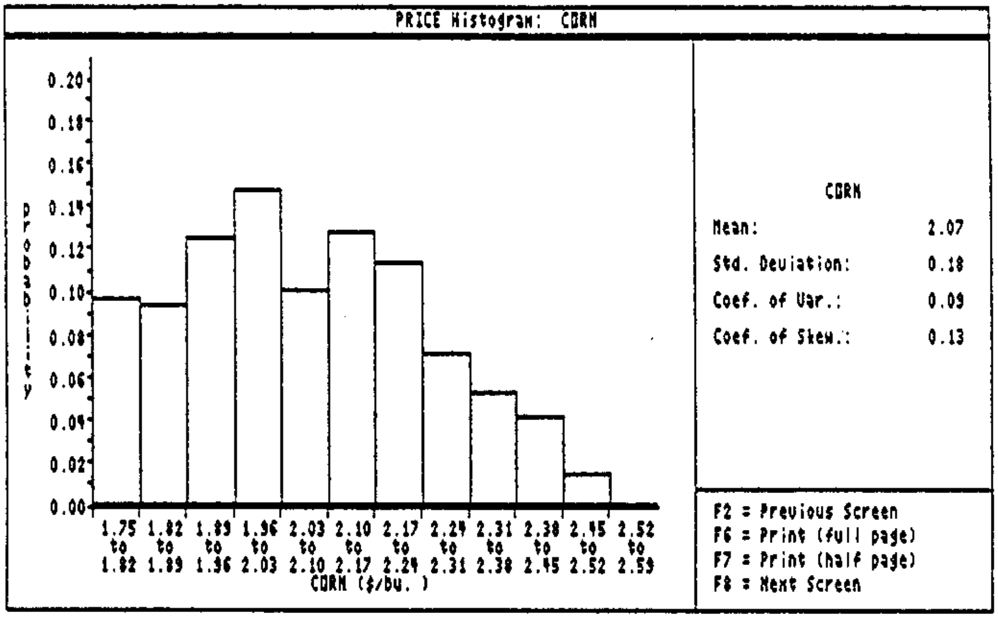
[x] Taken from unpublished research done by J. Roy Black, MSU, 1984.

Table 7.11

YIELD CUMULATIVE DISTRIBUTIONS				
	CORN	WHEAT	SOYBEANS	SETASIDE
Minimum	57.69	36.32	15.77	457.00
1st Percentile	84.72	39.87	21.88	457.00
5th Percentile	99.60	45.46	26.28	457.00
10th Percentile	110.16	47.76	28.07	457.00
20th Percentile	119.50	51.03	30.77	457.00
40th Percentile	132.21	54.65	33.71	457.00
50th Percentile	138.51	55.78	35.26	457.00
60th Percentile	143.49	57.28	36.17	457.00
80th Percentile	157.91	60.41	39.94	457.00
90th Percentile	169.66	63.34	42.19	457.00
95th Percentile	180.51	65.01	44.23	457.00
99th Percentile	187.19	69.32	47.69	457.00
Maximum	197.32	71.72	53.60	457.00
Mean	138.54	55.74	35.16	457.00
Std. Deviation	23.66	6.10	5.67	0.00
Coef. of Var.	0.17	0.11	0.16	0.00
Coef. of Skew.	0.00	-0.02	-0.06	0.00

Table 7.12

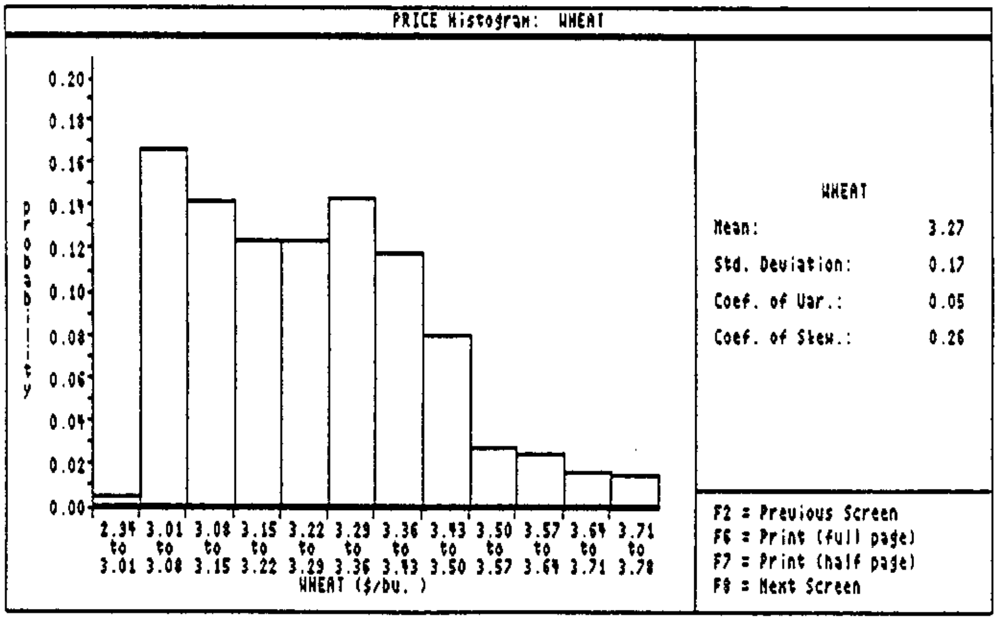
PRICE CUMULATIVE DISTRIBUTIONS				
	CORN	WHEAT	SOYBEANS	SETASIDE
Minimum	1.77	3.00	6.00	0.50
1st Percentile	1.77	3.02	6.02	0.51
5th Percentile	1.80	3.05	6.05	0.59
10th Percentile	1.82	3.07	6.08	0.65
20th Percentile	1.90	3.09	6.15	0.72
40th Percentile	1.98	3.19	6.36	0.81
50th Percentile	2.06	3.25	6.46	0.82
60th Percentile	2.12	3.30	6.56	0.87
80th Percentile	2.23	3.42	6.73	0.95
90th Percentile	2.32	3.49	6.87	0.97
95th Percentile	2.40	3.59	6.95	0.98
99th Percentile	2.46	3.73	7.12	0.99
Maximum	2.49	3.74	7.23	1.00
Mean	2.07	3.27	6.46	0.82
Std. Deviation	0.18	0.17	0.30	0.12
Coef. of Var.	0.09	0.05	0.05	0.15
Coef. of Skew.	0.13	0.25	0.04	-0.02



```
CORN
Mean:                  2.07
Std. Deviation:        0.18
Coef. of Var.:         0.09
Coef. of Skew.:        0.13

F2 = Previous Screen
F6 = Print (full page)
F7 = Print (half page)
F8 = Next Screen
```

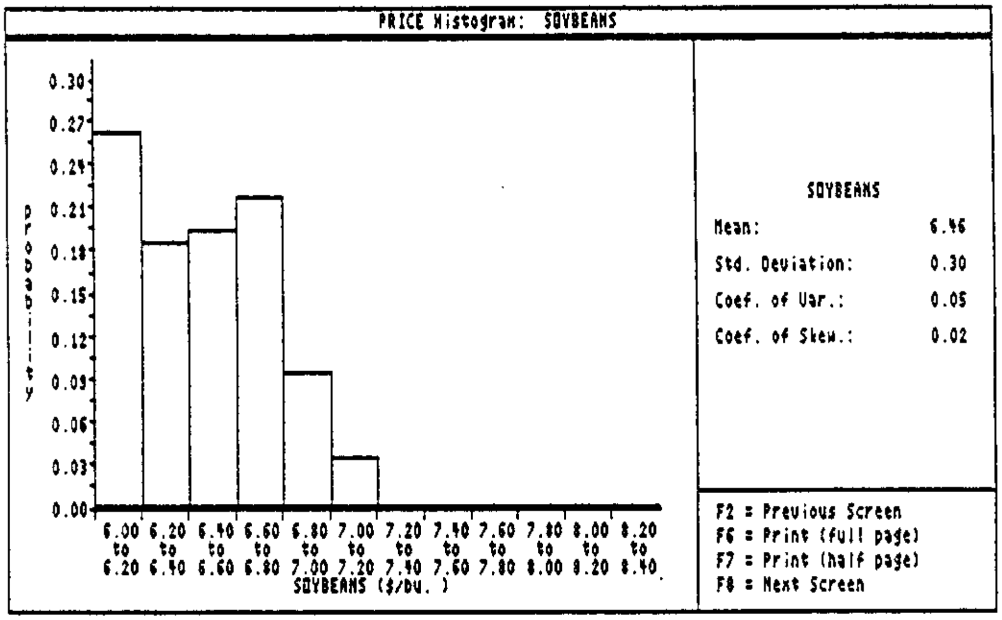
Figure 7.3
Price Histogram - Corn



```
WHEAT
Mean:                  3.27
Std. Deviation:        0.17
Coef. of Var.:         0.05
Coef. of Skew.:        0.26

F2 = Previous Screen
F6 = Print (full page)
F7 = Print (half page)
F8 = Next Screen
```

Figure 7.4
Price Histogram - Wheat


PRICE Histogram: SOYBEANS

SOYBEANS

Mean: 6.46

Std. Deviation: 0.30

Coef. of Var.: 0.05

Coef. of Skew.: 0.02

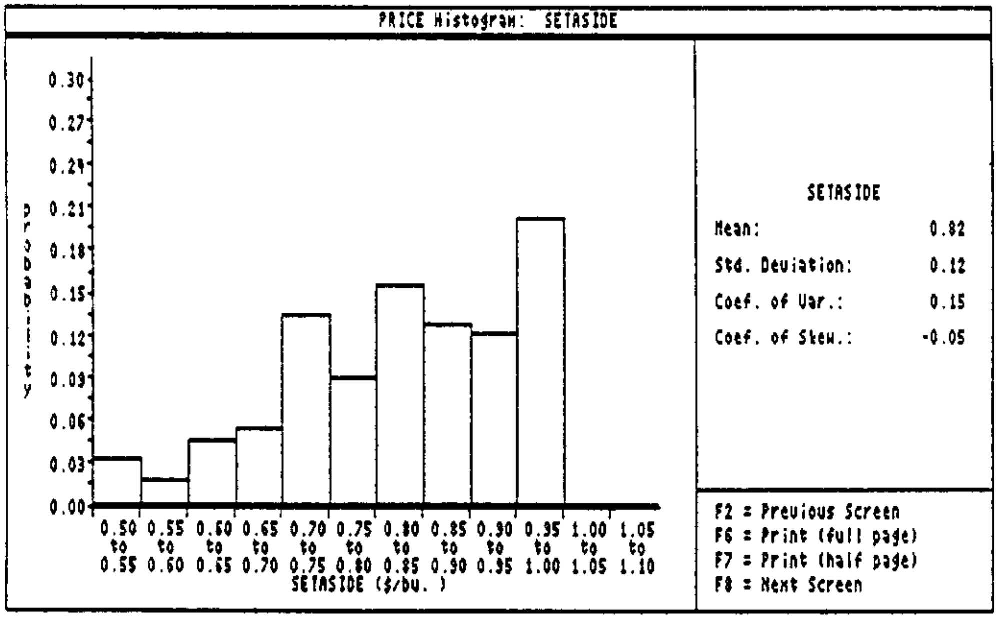
F2 = Previous Screen

F6 = Print (full page)

F7 = Print (half page)

F8 = Next Screen

Figure 7.5
Price Histogram - Soybeans


PRICE Histogram: SETASIDE

SETASIDE

Mean: 0.82

Std. Deviation: 0.12

Coef. of Var.: 0.15

Coef. of Skew.: -0.05

F2 = Previous Screen

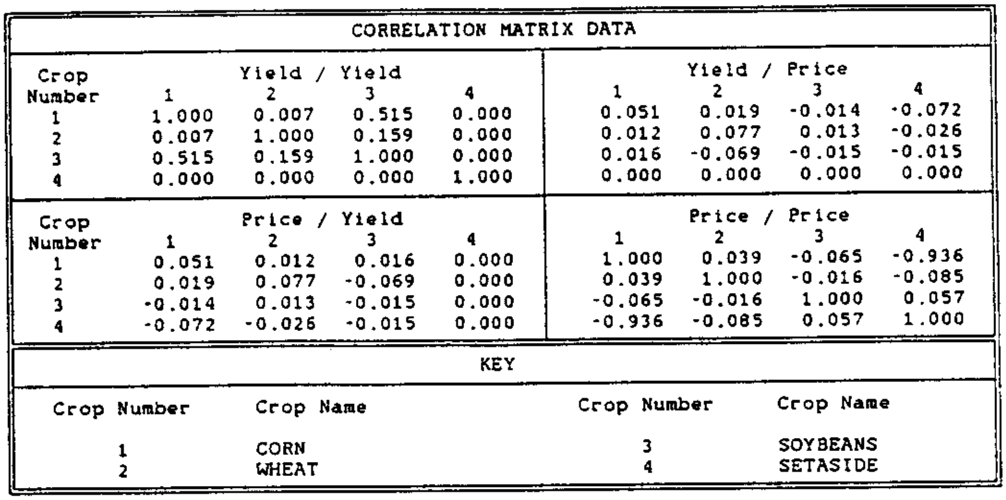
F6 = Print (full page)

F7 = Print (half page)

F8 = Next Screen

Figure 7.6
Price Histogram - Setaside

Table 7.13



CORRELATION MATRIX DATA

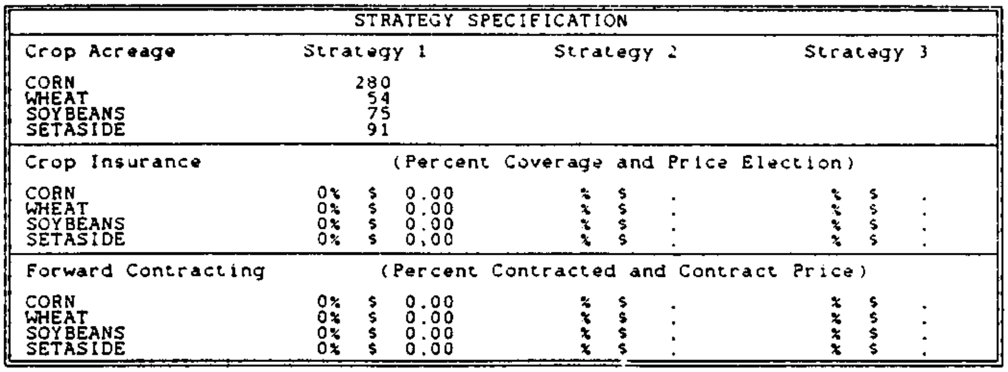
Crop Number	Yield / Yield 1	2	3	4	Yield / Price 1	2	3	4
1	1.000	0.007	0.515	0.000	0.051	0.019	-0.014	-0.072
2	0.007	1.000	0.159	0.000	0.012	0.077	0.013	-0.026
3	0.515	0.159	1.000	0.000	0.016	-0.069	-0.015	-0.015
4	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000

Crop Number	Price / Yield 1	2	3	4	Price / Price 1	2	3	4
1	0.051	0.012	0.016	0.000	1.000	0.039	-0.065	-0.936
2	0.019	0.077	-0.069	0.000	0.039	1.000	-0.016	-0.085
3	-0.014	0.013	-0.015	0.000	-0.065	-0.016	1.000	0.057
4	-0.072	-0.026	-0.015	0.000	-0.936	-0.085	0.057	1.000

KEY

Crop Number	Crop Name	Crop Number	Crop Name
1	CORN	3	SOYBEANS
2	WHEAT	4	SETASIDE

Table 7.14

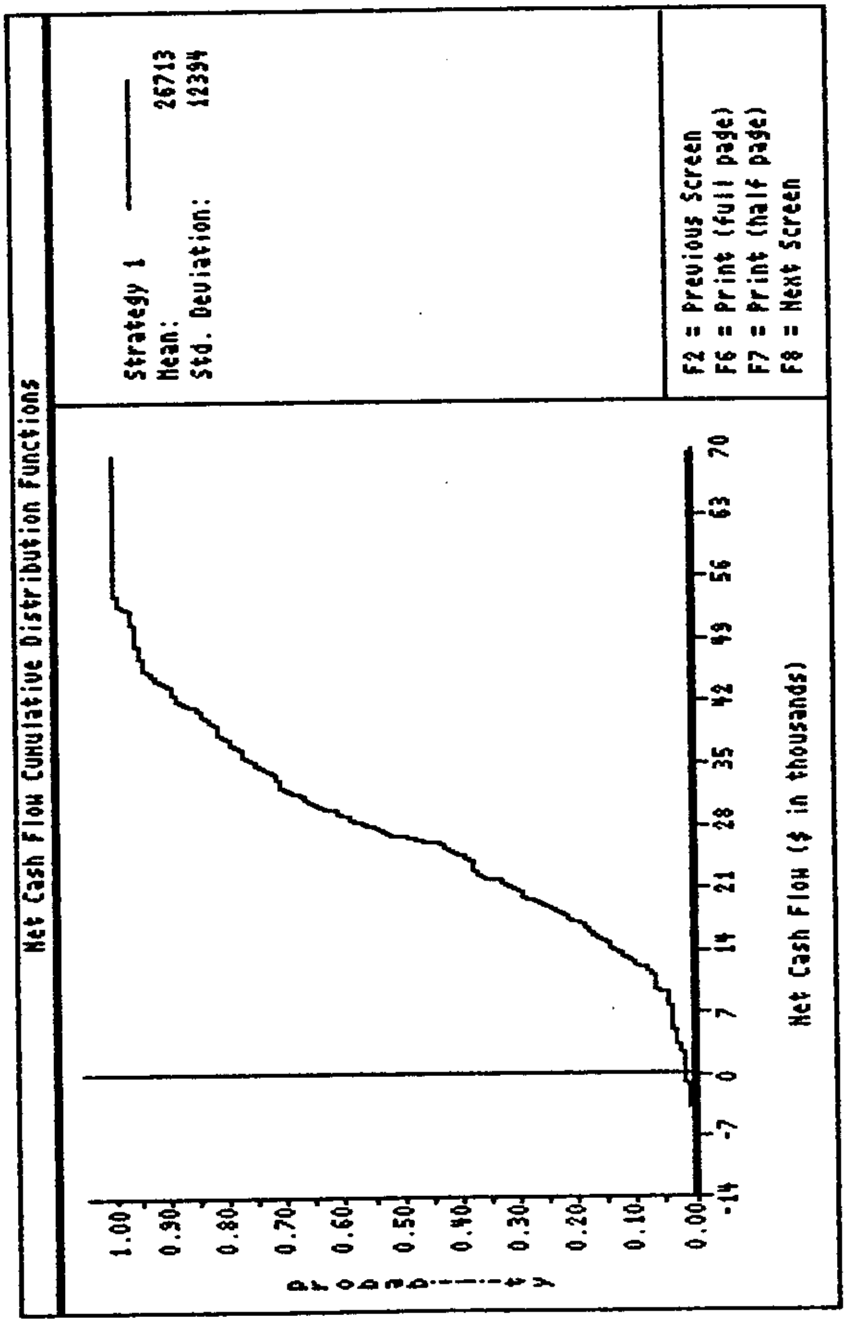


STRATEGY SPECIFICATION

Crop Acreage	Strategy 1	Strategy 2	Strategy 3
CORN	280		
WHEAT	54		
SOYBEANS	75		
SETASIDE	91		
Crop Insurance	(Percent Coverage and Price Election)		
CORN	0% $ 0.00	% $.	% $.
WHEAT	0% $ 0.00	% $.	% $.
SOYBEANS	0% $ 0.00	% $.	% $.
SETASIDE	0% $ 0.00	% $.	% $.
Forward Contracting	(Percent Contracted and Contract Price)		
CORN	0% $ 0.00	% $.	% $.
WHEAT	0% $ 0.00	% $.	% $.
SOYBEANS	0% $ 0.00	% $.	% $.
SETASIDE	0% $ 0.00	% $.	% $.

Table 7.15

NET CASH FLOW CUMULATIVE DISTRIBUTIONS ($/year)			
	Strategy 1		
Minimum	-12005		
1st Percentile	-2941		
5th Percentile	7719		
10th Percentile	11965		
25th Percentile	18611		
40th Percentile	24300		
50th Percentile	26552		
60th Percentile	28782		
75th Percentile	34772		
90th Percentile	43333		
95th Percentile	46826		
99th Percentile	54439		
Maximum	63126		
Mean	26713		
Std. Deviation	12393		



Net Cash Flow Cumulative Distribution Functions

Strategy 1

Mean: 26713

Std. Deviation: 12394

Figure 7.7
Net Cash Flow Cumulative Distribution Function For Case Example Farm

Notice that the probability of net cash flow being negative is approximately 2 percent. The mean or average net cash flow is estimated to be $26,713 with a standard deviation of $12,393.

To illustrate the net cash flow distribution as a measure of comparison for analyzing the whole-farm, the yield for corn is lowered by 8 bushels and the standard deviation decreased by approximately 5 bushels. Table 7.16 gives the new estimates for the yield distribution of corn and Table 7.17 summarizes the corresponding net cash flow distribution (see Figure 7.8 for a graphical representation of the cumulative adjusted net cash flow distribution). With a lower corn yield average and smaller standard deviation, the probability of net cash flow being negative is less than 1 percent. The average net cash flow is less at $23,047, but varies only $9,718 for each standard deviation. Notice also that the upside potential is not as appealing as compared to the higher yield and standard deviation.

In summary, having information on the variation of production for the farm business gives a basis for building planning tools that give a more realistic approach to decision making.

7.2 Long-Range Planning

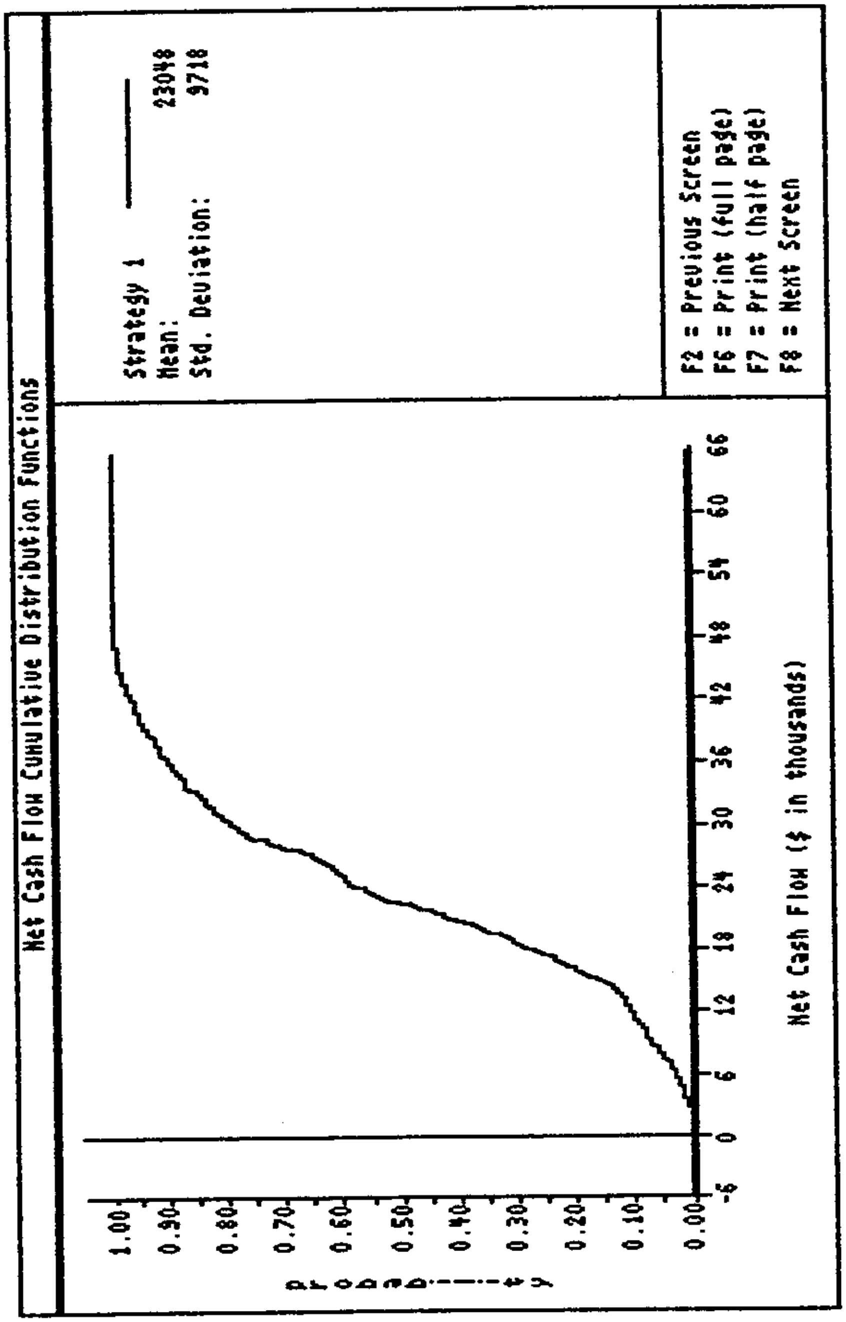
When considering a planning horizon of more than 2 years, the effect of growth in yield average should be included. This may or may not have a noticeable effect on gross margin as production expenses vary from year to year. Estimates for growth in yield average over time can be found in Tables 7.1, 7.2, and 7.3 for corn, wheat, and soybeans, respectively. Other notable observations that apply to long-range planning relate to the relationship of average yield with yield growth

Table 7.16
Yield Cumulative Distributions

YIELD CUMULATIVE DISTRIBUTIONS				
	CORN	WHEAT	SOYBEANS	SETASIDE
Minimum	75.99	36.69	20.95	457.00
1st Percentile	91.95	43.38	24.64	457.00
5th Percentile	101.43	46.90	28.03	457.00
10th Percentile	109.23	48.28	29.75	457.00
20th Percentile	116.26	50.01	31.52	457.00
40th Percentile	124.64	53.97	34.32	457.00
50th Percentile	128.05	55.30	35.78	457.00
60th Percentile	133.37	56.98	36.91	457.00
80th Percentile	146.48	60.77	39.74	457.00
90th Percentile	152.68	63.16	41.80	457.00
95th Percentile	163.63	65.33	43.98	457.00
99th Percentile	177.31	69.27	46.37	457.00
Maximum	182.92	71.99	48.13	457.00
Mean	130.50	55.51	35.68	457.00
Std. Deviation	18.44	5.82	4.82	0.00
Coef. of Var.	0.14	0.10	0.14	0.00
Coef. of Skew.	0.40	0.11	-0.06	0.00

Table 7.17
Net Cash Flow Cumulative Distributions

NET CASH FLOW CUMULATIVE DISTRIBUTIONS ($/year)			
	Strategy 1		
Minimum	-412		
1st Percentile	2624		
5th Percentile	6922		
10th Percentile	10735		
25th Percentile	17145		
40th Percentile	20353		
50th Percentile	22314		
60th Percentile	24810		
75th Percentile	28570		
90th Percentile	35740		
95th Percentile	40109		
99th Percentile	47005		
Maximum	55724		
Mean	23047		
Std. Deviation	9718		



Strategy 1
Mean: 23048
Std. Deviation: 9718

Figure 7.8
Net Cash Flow Cumulative Distribution Functions
For Adjusted Case Example

and with yield variation. Figures 7.9-7.11 indicate that high yield averages are commonly associated with high growth in yield averages. Also, it can be seen from Figures 7.12-7.14 that the size of yield variation seems to be independent of yield average. These observations seem logical in that high yield averages would be characteristic of progressive producers and that variation is random for all producers.

7.3 Forward Pricing

Having an estimation of yield average and yield variation, forward pricing becomes more of a calculated planning tool. Combining the variation of price with the variation in yield gives a more realistic approach to farm planning. With this information available, various strategies can be evaluated.

Using the scenario in Section 7.1 with the same assumptions, ARMS is used to evaluate possible forward pricing decisions. Table 7.18 summarizes three example strategies. Strategy 1 has no forward pricing, Strategy 2 forward prices 50 percent of the corn production at $2.00/bushel, and Strategy 3 forward contracts 50 percent of corn production at $2.25/bushel. The corresponding net cash flow distributions for the three strategies is given in Table 7.19 (see Figure 7.15 for a graphical representation). The net cash distribution for Strategy 2 shows the results of contracting corn below the expected mean price (see Figure 7.3 and Table 7.12). Strategy 3 shows the net cash distribution for contracting corn above the predicted mean price.

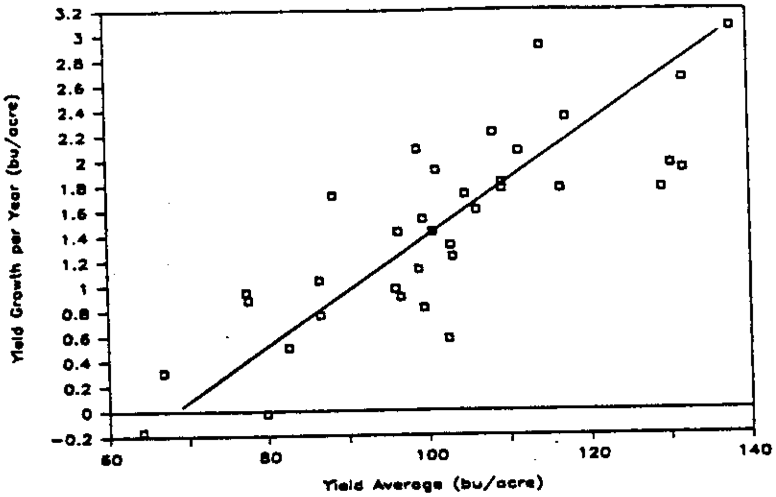


Figure 7.9
Yield Growth vs Yield Average - Corn

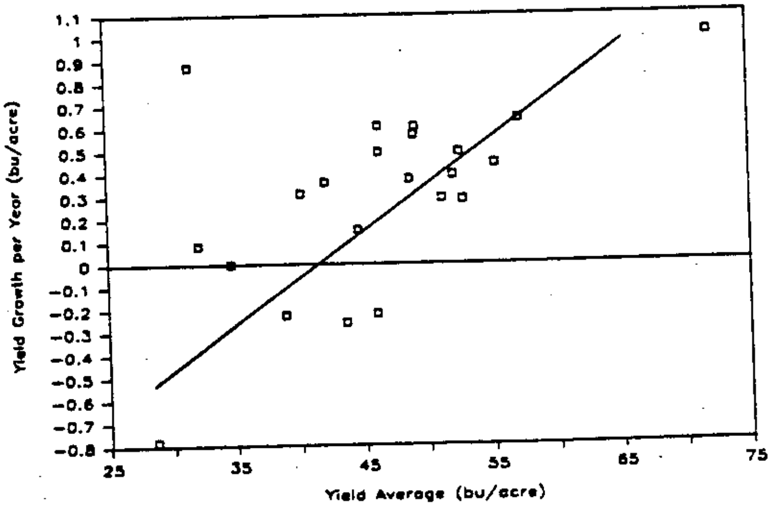


Figure 7.10
Yield Growth vs Yield Average - Wheat

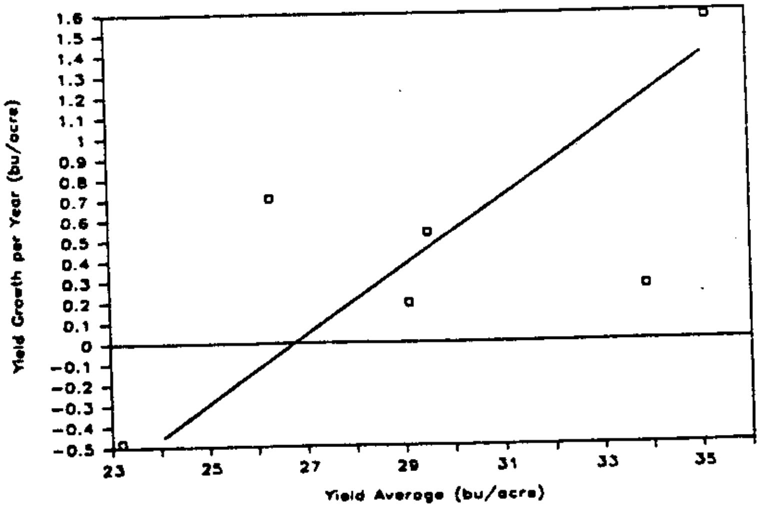


Figure 7.11
Yield Growth vs Yield Average - Soybeans

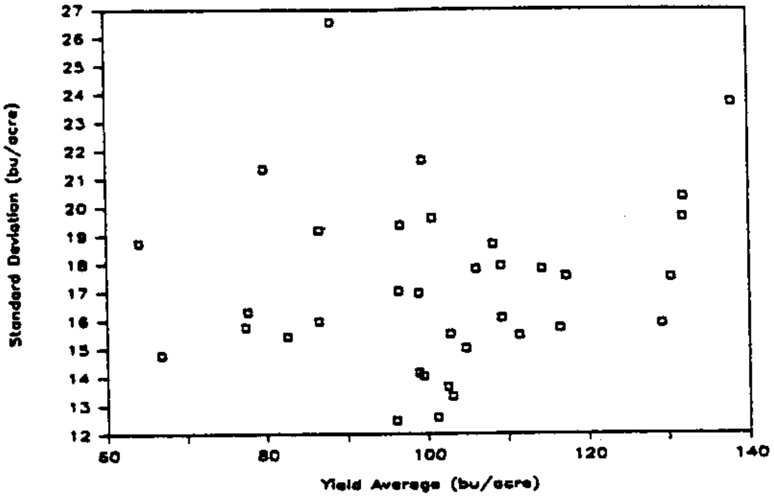


Figure 7.12
Standard Deviation vs Yield Average - Corn

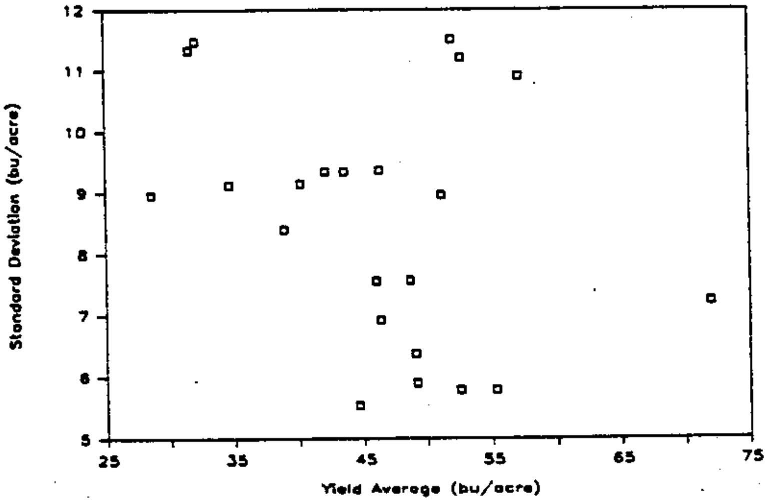


Figure 7.13
Standard Deviation vs Yield Average - Wheat

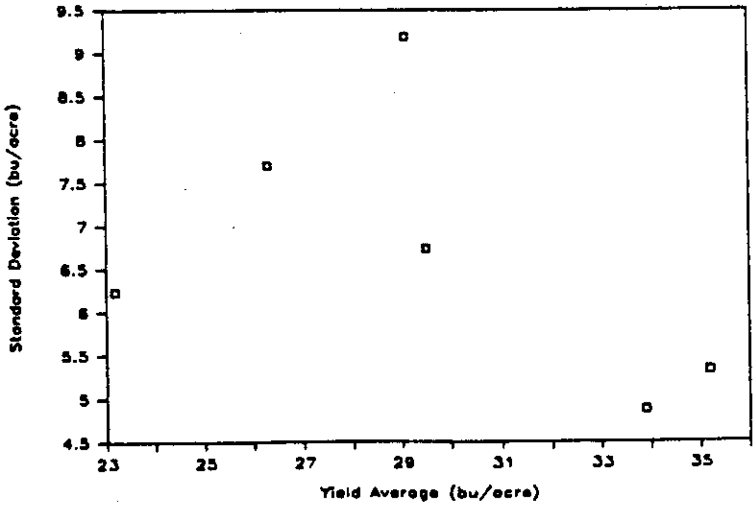


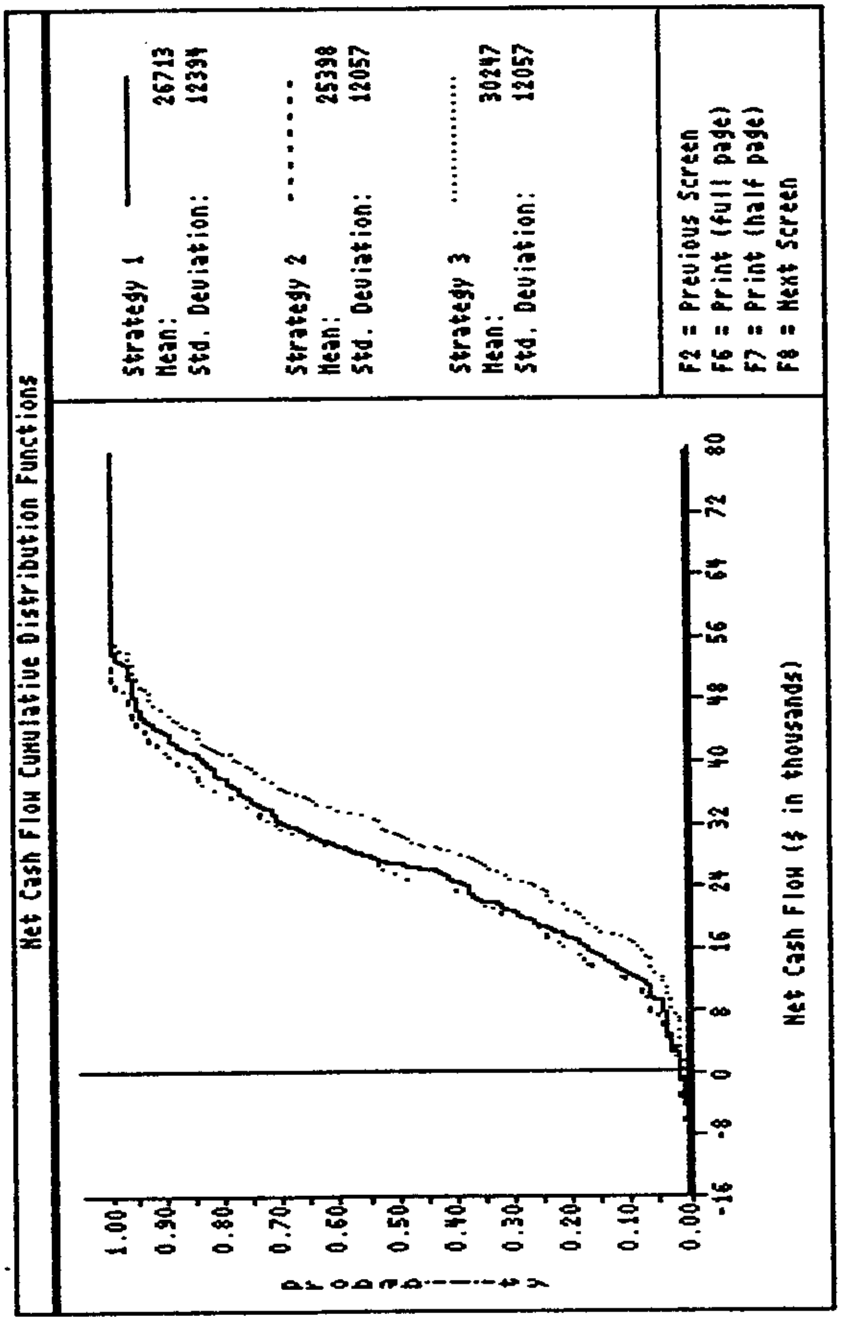
Figure 7.14
Standard Deviation vs Yield Average - Soybeans

Table 7.18
Strategy Specification

STRATEGY SPECIFICATION			
Crop Acreage	Strategy 1	Strategy 2	Strategy 3
CORN	280	280	280
WHEAT	54	54	54
SOYBEANS	75	75	75
SETASIDE	91	91	91
Crop Insurance	(Percent Coverage and Price Election)		
CORN	0% $ 0.00	0% $ 0.00	0% $ 0.00
WHEAT	0% $ 0.00	0% $ 0.00	0% $ 0.00
SOYBEANS	0% $ 0.00	0% $ 0.00	0% $ 0.00
SETASIDE	0% $ 0.00	0% $ 0.00	0% $ 0.00
Forward Contracting	(Percent Contracted and Contract Price)		
CORN	0% $ 0.00	50% $ 2.00	50% $ 2.25
WHEAT	0% $ 0.00	0% $ 0.00	0% $ 0.00
SOYBEANS	0% $ 0.00	0% $ 0.00	0% $ 0.00
SETASIDE	0% $ 0.00	0% $ 0.00	0% $ 0.00

Table 7.19
Net Cash Flow Cumulative Distributions

NET CASH FLOW CUMULATIVE DISTRIBUTIONS ($/year)			
	Strategy 1	Strategy 2	Strategy 3
Minimum	-12005	-14466	-9617
1st Percentile	-2941	-5366	-517
5th Percentile	7719	5932	10781
10th Percentile	11965	11119	15968
25th Percentile	18611	18022	22871
40th Percentile	24300	22782	27631
50th Percentile	26552	25098	29946
60th Percentile	28782	28395	33243
75th Percentile	34772	33628	38477
90th Percentile	43333	40779	45628
95th Percentile	46826	44649	49497
99th Percentile	54439	50079	54928
Maximum	63126	58064	62913
Mean	26713	25397	30246
Std. Deviation	12393	12057	12057



Strategy 1 ———
Mean: 26713
Std. Deviation: 12394

Strategy 2 - - - - - - - -
Mean: 25398
Std. Deviation: 12057

Strategy 3 ....
Mean: 30247
Std. Deviation: 12057

Figure 7.15
Net Cash Flow Cumulative Distribution Functions
For Three Pricing Strategies

7.4 Crop Insurance

The conclusions found in Chapters 5 and 6 have considerable application to crop insurance for the formulation of rates concerning corn, wheat, and soybeans grown in Michigan. Actual premium rates determined by the American Association of Crop Insurers can be compared with breakeven rates to evaluate appropriateness of crop insurance purchase for farmers with corn, wheat, and soybeans in Michigan. Hillsdale County was selected to illustrate this economic evaluation process.

The AACI provides producers with Multi-Peril Crop Insurance (MPCI) at three levels of production and three levels of price giving the grower nine different choices. The Breakeven Farmer Rate (BEFR) is the average indemnity or the breakeven premium rate for the producer (see Appendix F-1).

Table 7.20 summarizes the premium rates taken from the MPCI Rate and Transitional Yield Factor Table for corn, wheat, and soybeans and the Breakeven Farmer Rates (BEFR) calculated from average or above average values of standard deviation. Figures 7.16-7.24 give XY plots comparing the premium rates for two insurance options with the breakeven farmer rates for all three levels of protection.

Comparing the premium rates charged by a multi-peril crop insurance policy and the breakeven rates calculated for the producer, assuming the yield distribution is normally distributed, shows significant disparity for the example chosen. In general, low average yield producers should be charged more and high average yield producers charged less. The rate structure should seemingly be shifted so the benefit of any doubt be

Table 7.20
Insurance Rate Compared To BEFR

Corn

		50% COVERAGE			65% COVERAGE			75% COVERAGE		
YIELD	C. Var	H+F/in	H+F/out	BEFR	H+F/in	H+F/out	BEFR	H+F/in	H+F/out	BEFR
70	0.25	4.7	4	0.420	6.5	5.5	1.41	11.9	10.1	2.78
85	0.22	3.9	3.3	0.170	5.4	4.6	0.30	9.8	8.3	1.87
114	0.16	2.9	2.5	0.007	4	3.4	0.12	7.2	6.1	0.54
130	0.13	2.6	2.2	0.001	3.6	3	0.04	6.5	5.5	0.27

Wheat

		50% COVERAGE			65% COVERAGE			75% COVERAGE		
YIELD	C. Var	H+F/in	H+F/out	BEFR	H+F/in	H+F/out	BEFR	H+F/in	H+F/out	BEFR
30	0.30	2.7	2.2	1.180	3.6	3.1	2.77	6.6	5.7	4.53
40	0.24	1.7	1.4	0.320	2.4	2	1.18	4.3	3.7	2.46
50	0.19	1.3	1	0.040	1.8	1.5	0.37	3.2	2.7	1.11
60	0.13	1.1	0.8	0.001	1.5	1.3	0.02	2.8	2.4	0.18

Soybeans

		50% COVERAGE			65% COVERAGE			75% COVERAGE		
YIELD	C. Var	H+F/in	H+F/out	BEFR	H+F/in	H+F/out	BEFR	H+F/in	H+F/out	BEFR
25	0.30	4.3	3.6	1.180	6	5	2.770	10.8	9.1	4.53
30	0.22	3.6	3	0.170	4.9	4.1	0.800	9	7.6	1.87
35	0.14	3.1	2.5	0.001	4.2	3.6	0.040	7.7	6.6	0.27
40	0.10	2.7	2.1	0.000	3.7	3.2	0.000	6.7	5.7	0.02

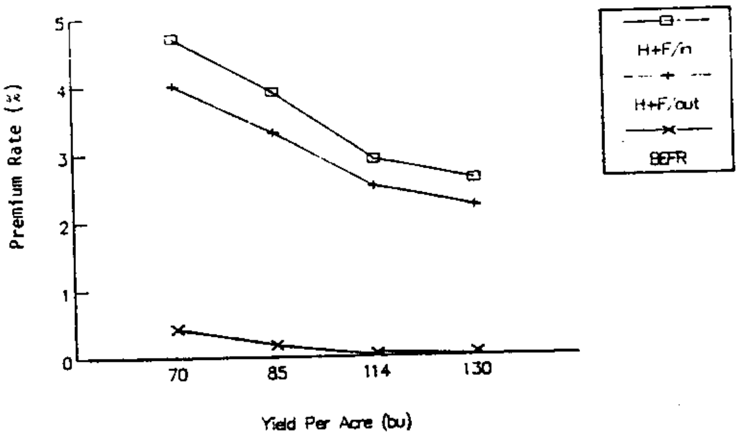


Figure 7.16
Crop Insurance Rates Compared To Break Even Farmer Rates
For Corn At 50% Coverage

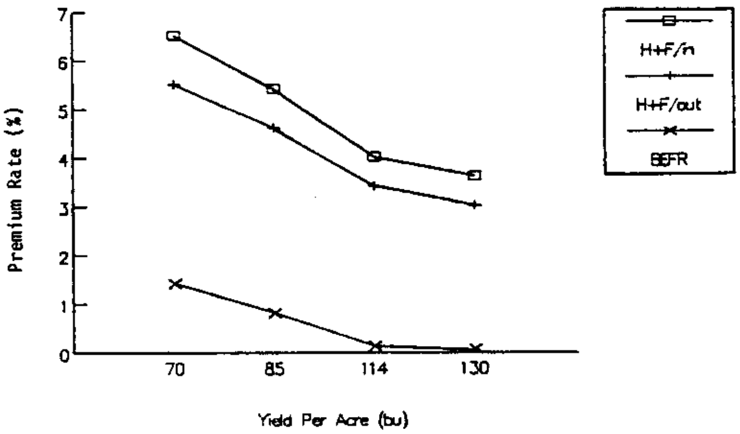


Figure 7.17
Crop Insurance Rates Compared To Break Even Farmer Rates
For Corn At 65% Coverage

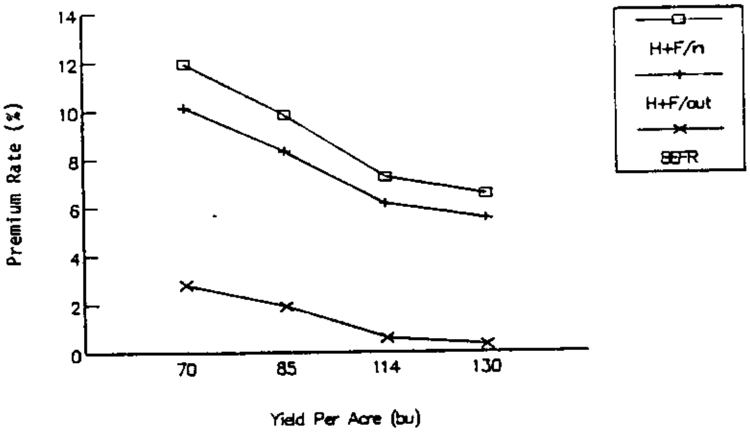


Figure 7.18
Crop Insurance Rates Compared To Break Even Farmer Rates
For Corn At 75% Coverage

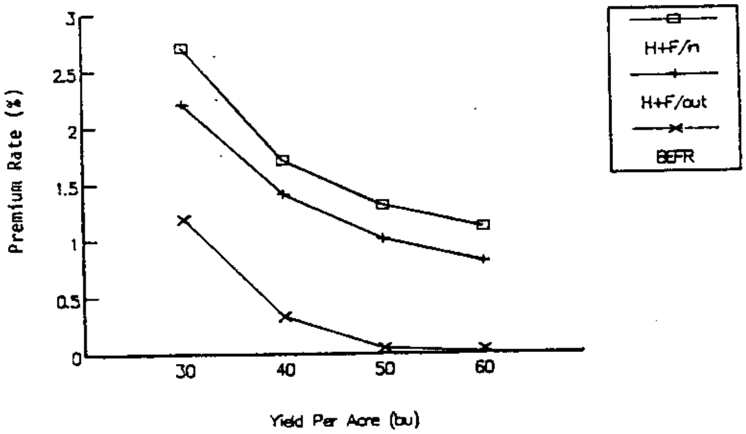


Figure 7.19
Crop Insurance Rates Compared To Break Even Farmer Rates
For Wheat At 50% Coverage

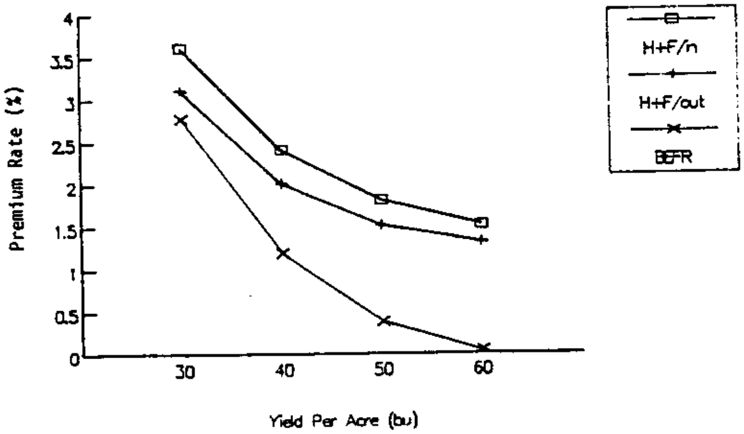


Figure 7.20
Crop Insurance Rates Compared To Break Even Farmer Rates
For Wheat At 65% Coverage

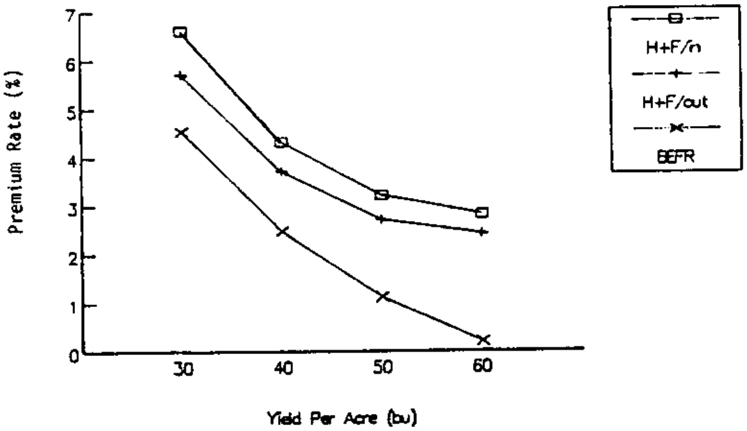


Figure 7.21
Crop Insurance Rates Compared To Break Even Farmer Rates
For Wheat At 75% Coverage

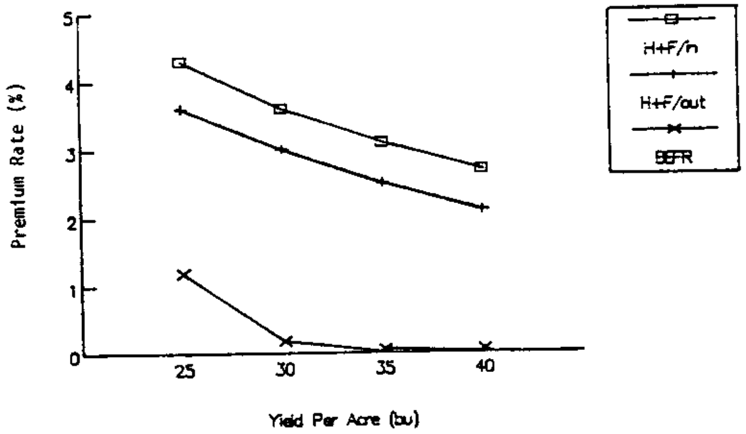


Figure 7.22
Crop Insurance Rates Compared To Break Even Farmer Rates
For Soybeans At 50% Coverage

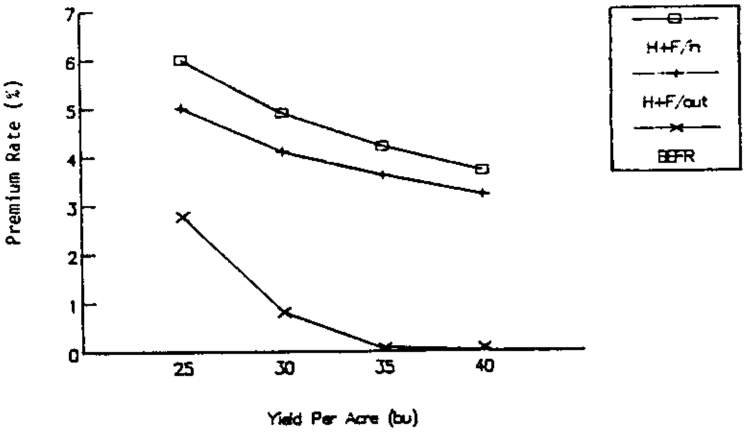


Figure 7.23
Crop Insurance Rates Compared To Break Even Farmer Rates
For Soybeans At 65% Coverage

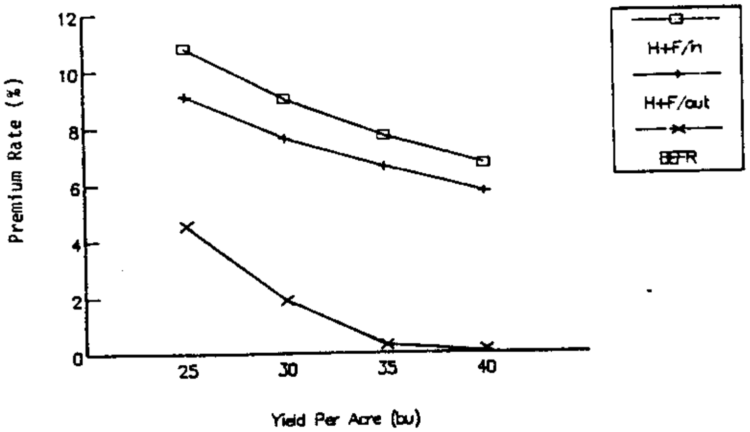
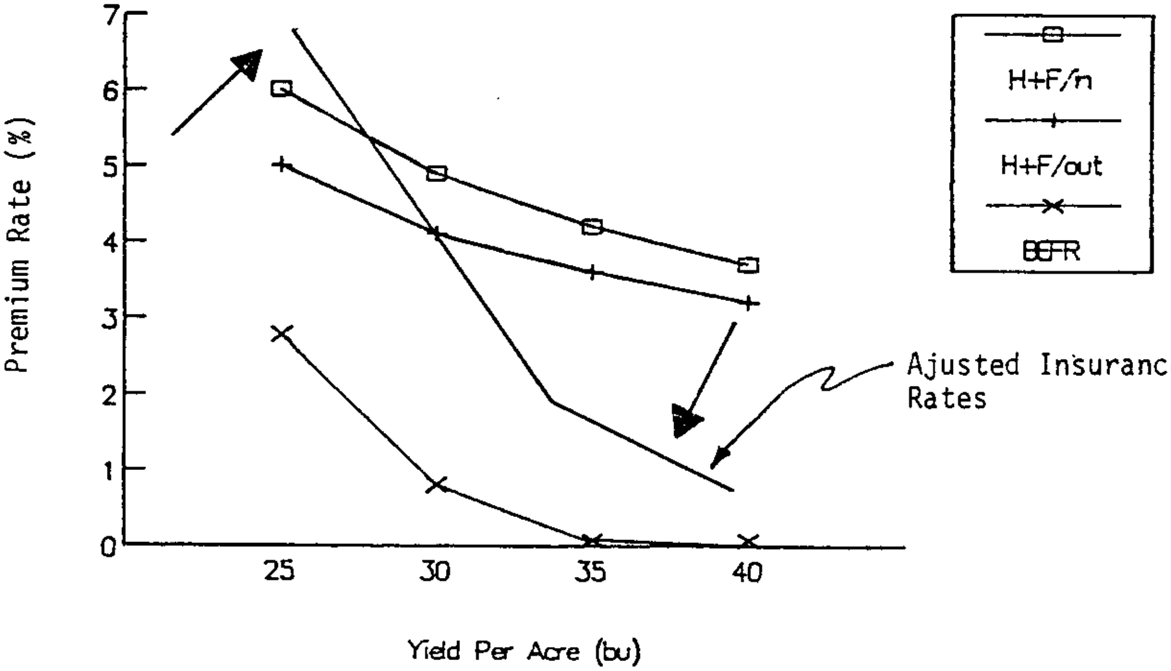


Figure 7.24
Crop Insurance Rates Compared To Break Even Farmer Rates
For Soybeans At 75% Coverage

given to high yield producers as they exhibit the less risk (see Figure 7.25).

Several points must be evaluated to see why this suggestion is of importance. Consider Figures 7.26-7.28. The coefficient of variation (standard deviation/yield average) is plotted against yield average. As a matter of simple mathematics, the low producers have almost double the percent variation in mean yields as high producers. This result has a direct impact on the ability of multi-peril crop insurance to provide economical risk reduction for producer depending on their level of average production. Multi-peril crop insurance provides coverage at 50 percent, 65 percent, and 75 percent of the producers average production. Figure 7.29 gives the coefficient of variation as interpolated from Figure 7.26 for various levels of production. It should be noted that as the level of production increases, the crop insurance levels move farther out on the tail of the distribution. Figure 7.29 shows that at low average production, the coverage levels of 50 percent, 65 percent, and 75 percent suit the producers distribution to offer viable liability coverage. As the coefficient of variation decreases, the coverage levels move farther out on the tail of the distribution giving the high average producer minimal coverage at the 75 percent level.

This fact is significant when considering the functional form of the distribution for rate formulation. The thickness of the tail (negative side) is important to the rates for above average producers (coverage levels are on the negative end of the distribution). The area between -2 standard deviations and the mean is important in making rates



Ajusted Insurance Rates

Figure 7.25
Example Of Suggested Rate Adjustment

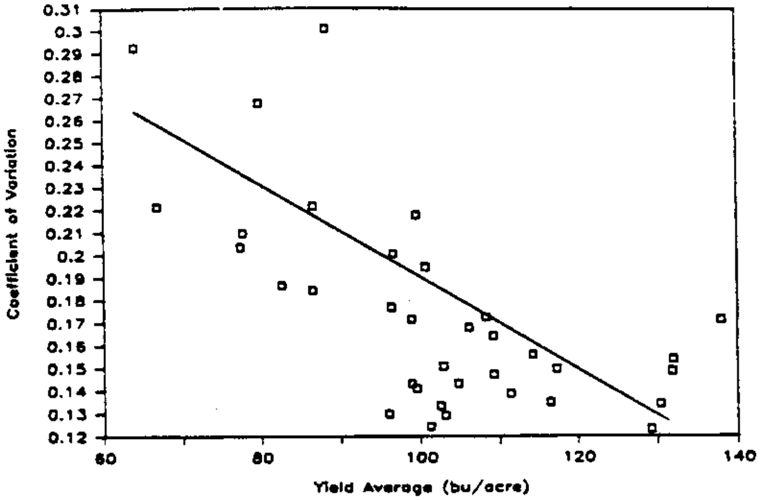


Figure 7.26
Coefficient of Variation vs Yield Average - Corn

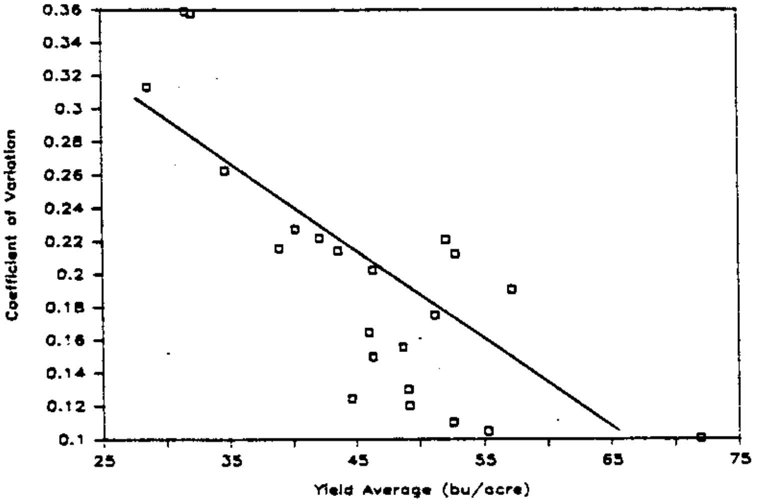


Figure 7.27
Coefficient of Variation vs Yield Average - Wheat

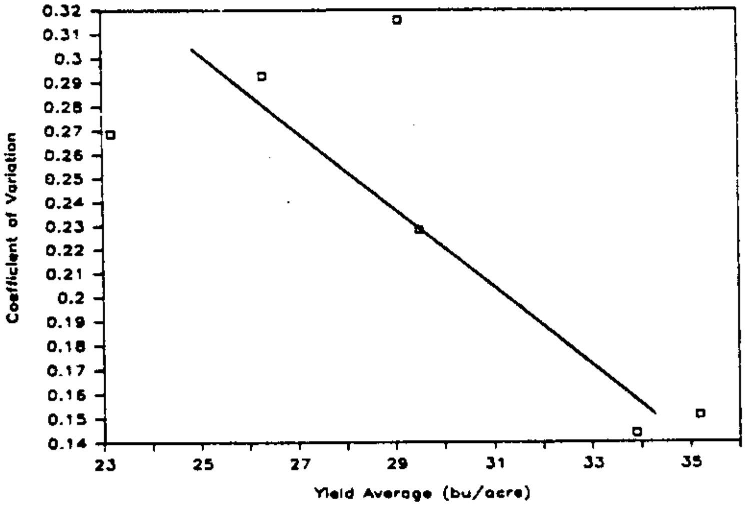


Figure 7.28
Coefficient of Variation vs Yield Average - Soybeans

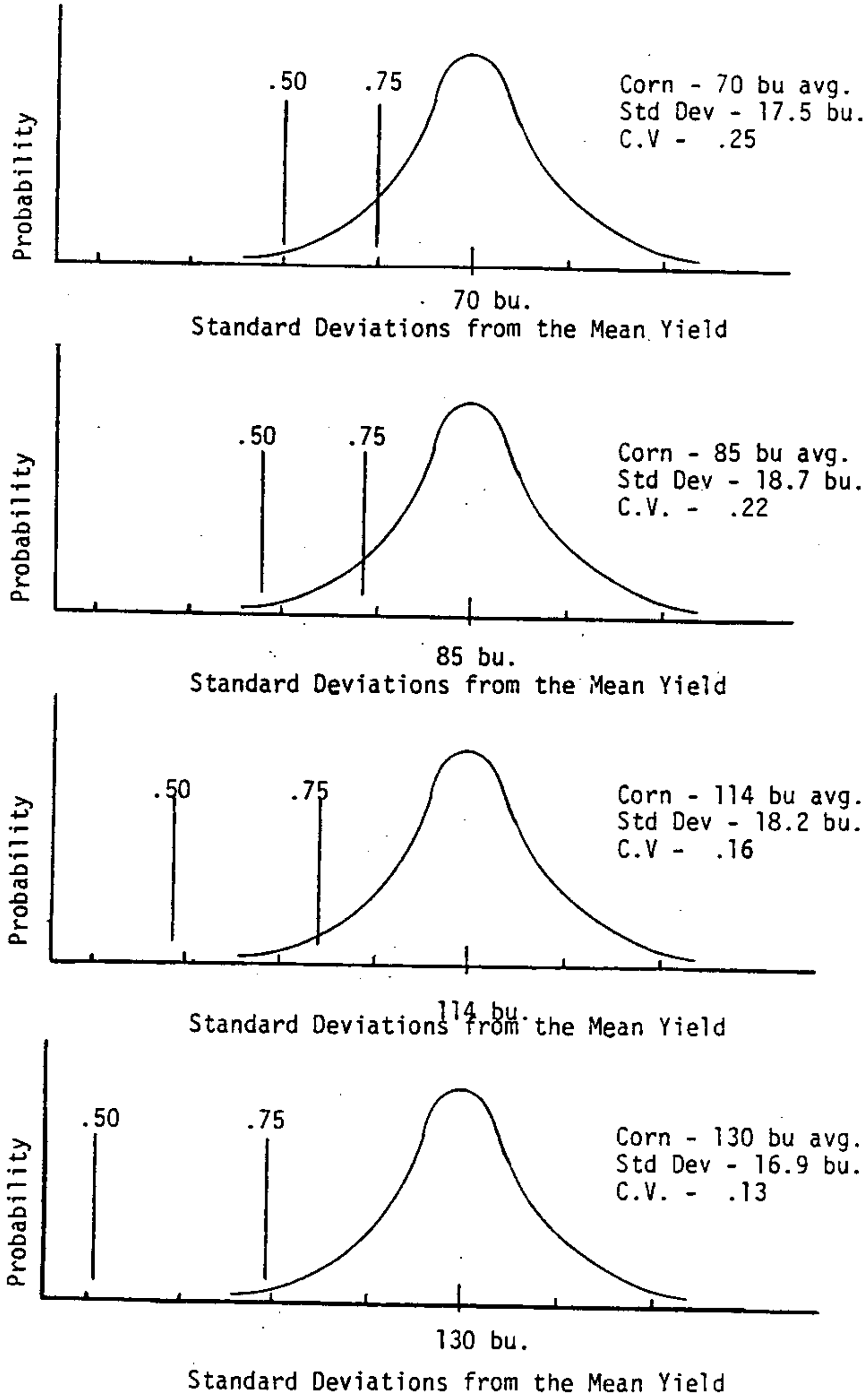


Figure 7.29
Relationship Between Coefficient Of Variation
And Level Of Yield Coverage For Crop Insurance

for below average producers. The results of Chapter 6 then become important for the formulation of rates for low average yield producers.

Another significant factor in rate formulation is the effect of growth in yield average. Recall from Section 7.2, graphs 7.9-7.11 depicting growth in yield average plotted against yield average. These graphs add an important addition to what has already been discussed. The plots strongly suggest that the high producers are getting better, faster and the low producers are slowly improving or getting worse. Interpretation of the negative side of these graphs is challenging, but a few explanations might be possible. The regression line might be misleading caused by one eratic data point and a visual inspection of the scattergram could change the opinion, or that particular county might not be suited for that crop, or possibly the business is on the road to dissolving and management is getting progressively worse each year. This question would be interesting for further research. Regardless of the background, the presence of negative growth means that over time the crop insurance coverage levels are moving up the negative side of the distribution because the coefficient of variation is increasing and, therefore, premium rates should be increased to keep in step with risk. For above average producers, because average yields are increasing, the insurance coverage levels will move farther out on the negative tail of the distribution. The premium rates will need to be lowered over time in order to attract better-than-average producers to utilization of the MPCI risk-management strategy.

7.5 Summary

This investigation was initially prompted by multiple computer runs using ARMS to evaluate the competitiveness of crop insurance for reducing farmers downside risk. The numbers were not at all appealing for above average producers which would seem to be the target market for crop insurance.

Significant factors in formulating insurance rates are all effected by the producers average yield. Average yield effects the coefficient of variation (percentage of production that is variation). It also appears to be strongly associated to the growth in average production over time. This single factor then has a significant impact in rate formulation and should be given careful consideration.

There needs to be an adjustment to the distribution that is used for rate formulation. It would seem that the insurance rates should be higher for low average producers and lower for high average producers. Results from fitting distributions to individual farms in Chapter 6 could serve as a reference. The main concern would be that the negative side of the distribution would be heavier as compared to the normal in the area from the mean to -2 standard deviations and that the tail to the left of the -2 standard deviations not be much heavier than the normal distribution.

Considerable attention should also be given to incorporating the growth in yield average over time especially for the low average producer. It would seem logical that insurance not be available for yield averages under a specified percentage of the county average yield unless the premium rate is adequately adjusted.

CHAPTER 8

SUMMARY AND CONCLUSIONS

8.1 Objective Of Study

The main objective of this research project has been to estimate the parameters (mean, variance) of continuous probability distributions that describe the variation in yield for corn, wheat, and soybeans grown in Michigan. The focus of the estimation process was taken from three different perspectives; 1) each crop as a whole, 2) soil groupings for each crop, and 3) each farm of a selected soil group for corn.

The second objective of this study was to develop the methodology for estimating continous probability distributions and evaluate their accuracy in modeling yield variation. Since tests for determining normality are low in power, several other continuous probability distributions were evaluated for goodness-of-fit.

The last concern of this study was to incorporate the use of continuous probability distributions into decision-making models that would be useful for risk management.

8.2 Data Sources

The data for this research project were taken from actual farm records of Michigan farmers enrolled in Telfarm. Telfarm is a self-supported record-keeping service provided by the Cooperative Extension Service of the State of Michigan at Michigan State University. After error-checking time-series data (last year-1983) for individual farms enrolled in Telfarm, 87 farms were available for analysis of corn, 35 farms for wheat and because of the relative newness of soybeans to Michigan agriculture, only 7 farms were available for soybeans.

8.3 Analysis And Results

Regression analysis was used on initial yield data to determine the presence of upward trend in yields. The detrended yield data appeared to be symmetrically distributed around the fitted line. This observation brought forth the hypothesis that the crop yield distributions could be represented by the Normal distribution. It was also hypothesized that the crop yields could be characterized as independently distributed having constant variance.

All data sets for all three crops were evaluated with detrended yield data so all data points could be evaluated on an equivalent technological basis. The Durbin-Watson test was used, along with the scattergrams, to evaluate the yield independence assumption. The constant variance assumption was tested using a procedure outlined by Stewart and Wallis along with the scattergram of the residuals from the regression analysis. The assumption of Normality for each farm was evaluated by the three statistics for skewness, kurtosis, and the Shapiro-Wilks W-test. Also, the Shapiro-Wilks W-tests for each farm (each crop) were grouped together to test the normality assumption for each crop.

The null hypothesis that yield variation for corn, wheat, and soybeans could be described by a Normal distribution, independently distributed with constant variance was not rejected. Very little evidence was present for serial correlation for the three crops to disprove independence of yield variation. Even with significant levels as high as 15 percent, little evidence was found to reject the constant variance assumption. Visual inspection of the scattergrams of the absolute value of regression residuals was also used as this test is low in power. The results for skewness, kurtosis, W-statistics, and the grouped W-statistics do not provide substantial evidence to disprove the normality assumption when considering all farms for each crop.

Farms were then classified by soil groupings for corn and wheat to determine whether there was any evidence to reject the hypothesis of independent and normally distributed yields with constant variance when farms were evaluated by soil groupings. The assumption of yield independence was supported for all soil groups for both corn and wheat. The constant variance assumption was found to hold for all soil groups and both crops except on loamy soils for corn. The normality assumption was supported for all soil groups for both crops except wheat grown on loamy soils.

The normality assumption was also tested for individual farm level data. Corn yield distributions for farms classified by wet loamy soils were analyzed to determine if another probability distribution would better describe the yield variation. It was found that farms with symmetrical data values around the mean are best described by the Normal distribution. Farms with fairly symmetrical data values dispersed around the mean, but having significant amounts of negative kurtosis, could be more accurately described by the Beta probability distribution. Data sets with negative outliers were found to be best fit with the Extreme Value Type A probability distribution. Because these statistical tests are low in power, the results must be used carefully.

8.4 Applications

These research results were applied to farm management decision-making for farm planning. The applications considered were short-term planning, long-range planning, forward pricing, and purchase of multiple peril crop insurance.

The distribution of yield variation and its parameters (mean, variance) were included in an enterprise budget to show the effect of production risk on gross margin for making short-term decisions. The distribution parameters were also used with the Agricultural Risk Management Simulator (ARMS), a whole-farm planning tool for evaluating the management of risk through crop insurance and/or forward contracting.

Parameters estimated for the yield distributions were also found to have significant application to long-range planning. When considering a planning horizon of more than two years, the effect of growth in yield average should be included. It was also noted that high yield averages are commonly associated with high growth in yield averages and that the size of yield variation seems to be independent of yield average.

The estimate of distribution parameters were also used for forward pricing decision-making. Parameters were used in the Agricultural Risk Management Simulator to evaluate various forward contracting scenarios.

Lastly, the assumption of normally distributed yields and distribution parameters were used to evaluate crop insurance premium rate competitiveness with a break-even farmer premium rate. Break-even rates were compiled for a case example and compared to the actual insurance premium rates. It was found that, in general, producers having low average yield levels were charged too little and that high average yield producers were over charged.

8.5 Future Research

Telfarm data is available (20-40 years) for four other soil groups for corn and four soil groups for wheat. Fitting distributions for individual farms on these other soil management groups would give a much broader base of information to evaluate the functional form of the continuous probability distribution at the farm level. This would seem of significant importance for evaluating risk management instruments as multiple-peril crop insurance.

Other noteable observations for this study that might be of topics of further research would be; 1) the presence of negative yield growth over time, 2) the presence of increasing/decreasing variance in yield average, and 3) high average yields being associated with high growth in yield averages.

APPENDICES

APPENDIX A - 1

Individual Farm Statistics For
Corn, Wheat, and Soybeans

TABLE A-1 Corn

FARMNUM	a ( )[1]	b ( )[1]	R ( )[2]	DW	Y M	α_2 ( )[2]	CV	STD DEV	$\sqrt{B_1}$	β_2	W-test
030001	-106.0 (33.0)	2.70 (.451)	.657 (.000)	2.20	120.1	.152 (.039)* ▲	.10	12.215	-.491	-.888	.93197
030022	-32.9 (44.9)	1.44 (.627)	.249 (.035)	1.88	82.6	.082 (.123)	.16	13.398	-.068	1.572*	.96815
030023	177.6 (44.2)	-1.04 (.609)	.155 (.105)	2.11	92.8	.116 (.082)	.14	13.259	-.263	-.646	.96074
030173	-29.2 (39.1)	1.76 (.534)	.363 (.004)	1.93	116.8	.032 (.215)	.12	14.447	.048	.972	.96517
050001	81.8 (63.6)	-.024 (.866)	.000 (.978)	1.95	79.8	.174 (.037)* ▲	.27	21.361	-.518	-.695	.92958
080031	-55.7 (38.6)	1.67 (.534)	.365 (.006)	1.56	81.5	.153 (.048)* ▶	.15	12.611	-.373	1.176*	.94843
080392	-19.2 (50.4)	1.79 (.694)	.271 (.019)	1.51	128.0	.015 (.301)	.14	17.426	-.521	-.837	.93112
110225	-116.2 (64.3)	3.06 (.877)	.403 (.003)	.897*	138.0	.042 (.192)	.17	23.685	-.581	.086	.92469
120001	-51.6 (68.6)	1.86 (.954)	.192 (.068)	2.02	99.4	.003 (.409)	.22	21.360	.063	-.644	.98573
120004	-55.4 (48.3)	2.00 (.659)	.325 (.007)	2.08	110.5	.095 (.086)	.16	17.848	-.014	-.805	.95022
120396	-14.3 (12.0)	1.45 (.185)	.595 (.000)	2.34	106.4	.005 (.319)	.14	14.707	-.056	-.832	.96674
120399	17.2 (46.9)	1.09 (.641)	.132 (.105)	1.71	107.9	.003 (.403)	.16	17.351	.247	.394	.95076
130035	-43.9 (12.9)	2.04 (.204)	.714 (.000)	1.62	125.5	.064 (.052)	.13	15.835	.225	-.413	.97095
130038	-13.7 (46.7)	1.43 (.638)	.209 (.037)	1.21*	105.2	.000 (.484)	.16	17.279	.269	.714	.98107
130203	-34.4 (39.2)	1.84 (.538)	.395 (.003)	2.14	118.8	.135 (.054)	.12	14.057	.140	-.790	.96375
190004	14.8 (51.1)	.942 (.711)	.099 (.203)	1.84	91.2	.205 (.029)* ▶	.17	15.914	-.443	-.907	.94191

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression
DW - Durbin-Watson test for serial correlation YM - Estimated yield mean
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estiimated standard deviation $\sqrt{B_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE A-1 Corn - continued

FARMNUM	a ( )[1]	b ( )[1]	R ( )[2]	DW	Y M	α_2 ( )[2]	CV	STD DEV	$\sqrt{b_1}$	β_2	W-test
190027	17.7 (47.5)	1.00 (.649)	.112 (.138)	2.17	101.3	.002 (.419)	.17	17.575	-.419	.609	.97341
190042	32.7 (69.6)	.778 (.955)	.035 (.426)	1.29	97.4	.058 (.151)	.25	24.660	-1.001*	.548	.90798
230044	-28.1 (11.5)	1.53 (.182)	.645 (.000)	1.80	99.5	.106 (.018)* ▲	.14	14.048	.507	-.370	.95375
240048	-53.1 (24.1)	1.72 (.382)	.373 (.000)	1.49	88.3	.016 (.229)	.30	26.560	1.112*	.796	.89233*
290018	-28.9 (15.2)	1.96 (.242)	.633 (.000)	1.62	130.4	.003 (.350)	.13	17.506	-1.740*	6.463*	.88982*
300011	-31.4 (11.5)	1.76 (.181)	.702 (.000)	1.64	114.7	.043 (.092)	.12	14.074	-.375	.387	.98530
300074	-5.3 (39.2)	1.29 (.539)	.243 (.027)	2.21	101.0	.007 (.355)	.13	13.536	-.397	.046	.97457
300101	64.9 (34.4)	.224 (.470)	.011 (.639)	3.01*	83.6	.062 (.137) ▲	.15	12.732	.015	.294	.99027
300174	-7.8 (49.1)	1.25 (.679)	.166 (.082)	1.92	96.2	.013 (.316)	.17	16.305	-.480	-.603	.95294
330040	-15.0 (34.8)	1.45 (.479)	.351 (.007)	1.44	105.8	.000 (.497)	.12	12.257	-.601	-1.001	.90513
330144	-63.7 (58.6)	2.12 (.806)	.278 (.017)	1.98	110.5	.000 (.460)	.18	20.251	.104	-.268	.94496
330222	-67.8 (44.0)	1.72 (.614)	.329 (.013)	1.91	69.8	.077 (.132) ▲	.19	13.118	.610	.114	.95358
330225	-142.5 (66.4)	3.07 (.918)	.412 (.004)	1.67	109.6	.114 (.085) ▼	.20	22.300	-.695	.314	.94207
340073	32.5 (59.2)	.842 (.803)	.057 (.308)	1.42	102.4	.018 (.288)	.20	20.176	-.259	-.679	.95916
340093	-51.4 (60.3)	1.87 (.820)	.225 (.034)	1.90	104.4	.142 (.050)	.20	21.618	.186	-.791	.97115
340097	27.5 (43.8)	.909 (.606)	.116 (.152)	1.42	103.0	.000 (.462)	.14	14.560	.599	-.044	.95894
340560	-77.9 (68.9)	2.11 (.954)	.223 (.041)	1.51	92.9	.014 (.311)	.24	22.140	-.579	-.912	.92537

---

[1] Standard Error [2] Significance level * - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression
DW - Durbin-Watson test for serial correlation YM - Estimated yield mean
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estimated standard deviation $\sqrt{b_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE A-1 Corn - continued

FARMNUM	a ( )[1]	b ( )[1]	R ( )[2]	DW	Y_M	α_2 ( )[2]	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-test
370145	-20.4 (50.9)	1.36 (.695)	.167 (.065)	1.23*	92.6	.051 (.162)	.21	18.821	.429	-.838	.94280
370146	83.6 (71.0)	.276 (.977)	.004 (.780)	2.13	106.3	.051 (.167)	.23	24.533	.273	-.474	.95653
380062	-66.3 (45.5)	2.18 (.622)	.394 (.002)	1.70	115.2	.030 (.225)	.15	16.825	.608	.137	.95228
380080	-60.2 (40.5)	1.97 (.554)	.399 (.002)	1.42	103.2	.034 (.211)	.15	14.984	-.272	.330	.97632
380485	-42.4 (46.0)	1.80 (.628)	.301 (.010)	2.39	106.9	.030 (.223)	.16	16.997	.031	-.967	.96930
380486	-62.7 (36.0)	2.13 (.493)	.509 (.000)	2.70	114.4	.103 (.083) ▼	.12	13.237	-.409	.492	.97370
380493	-71.2 (41.3)	2.27 (.564)	.460 (.001)	2.19	117.5	.060 (.141) ▼	.13	15.264	-.676	-.152	.93286
390002	16.9 (9.9)	.650 (.160)	.308 (.000)	1.12*	70.3	.092 (.029)* ▲	.16	11.418	.024	-.490	.98077
390095	-8.0 (18.0)	1.12 (.288)	.291 (.000)	1.65	85.1	.005 (.323)	.25	21.189	.266	-.490	.97026
410106	58.1 (35.1)	.309 (.483)	.022 (.530)	2.56	83.4	.173 (.033)* ▲	.15	12.125	-.940*	1.025	.89556*
410114	-26.2 (37.1)	1.64 (.510)	.365 (.005)	1.55	108.6	.033 (.220)	.12	12.826	-.841*	.928	.94254
440175	-66.5 (44.7)	2.35 (.623)	.470 (.002)	2.23	123.8	.007 (.368)	.11	13.549	-.535	-.782	.92217
440182	-82.5 (58.3)	2.32 (.797)	.309 (.009)	2.52	110.7	.160 (.035)* ▲	.19	21.557	.499	-.169	.94024
460007	45.2 (29.2)	.608 (.403)	.117 (.151)	1.89	95.7	.152 (.049)* ▲	.10	9.972	-.302	.487	.96792
460056	-113.1 (35.3)	2.55 (.476)	.627 (.000)	1.92	98.8	.036 (.217)	.12	11.446	-.045	-.211	.96578
460444	-107.3 (45.2)	2.60 (.615)	.499 (.000)	2.61	109.0	.009 (.343)	.15	16.341	-.842*	.711	.94060
470006	24.0 (14.4)	.725 (.231)	.214 (.003)	2.07	84.2	.034 (.133) ▲	.20	16.862	-.337	-.017	.97733

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression

DW - Durbin-Watson test for serial correlation Y_M - Estimated yield mean

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE A-1 Corn - continued

FARMNUM	a ( )^1	b ( )^1	R ( )^2	DW	Y M	α_2 ( )^2	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-test
470119	-37.3 (38.4)	1.91 (.524)	.412 (.002)	1.95	121.7	.001 (.440)	.12	14.198	-.500	-.436	.94565
530013	-36.6 (13.3)	1.49 (.209)	.594 (.000)	1.59	86.0	.001 (.426)	.17	14.438	-.129	-.306	.98534
530067	34.8 (46.2)	.410 (.631)	.021 (.524)	2.70	68.8	.139 (.047)* ▲	.25	17.088	.347	-.388	.97417
540001	77.6 (50.8)	-.163 (.694)	.003 (.816)	1.31	64.1	.000 (.463)	.29	18.749	-.302	-1.224*	.92071
580219	-27.6 (54.3)	1.92 (.745)	.270 (.019)	2.12	131.9	.032 (.221)	.15	19.625	-1.878*	5.607*	.84178*
590010	-27.0 (12.1)	1.27 (.188)	.546 (.000)	1.76	79.0	.066 (.054)	.18	14.468	.202	-.499	.96551
590020	-247.1 (48.5)	4.70 (.667)	.758 (.000)	2.23	141.1	.117 (.081)	.11	15.185	-.197	-.424	.95554
590093	-145.3 (52.8)	3.23 (.717)	.529 (.000)	1.54	122.7	.000 (.484)	.15	18.276	.295	.102	.98168
590133	-87.5 (63.1)	2.42 (.862)	.294 (.011)	1.39	113.9	.007 (.351)	.20	23.326	-.363	-1.409*	.90563*
610043	-43.9 (52.9)	1.65 (.718)	.228 (.033)	2.87	93.7	.009 (.341)	.20	18.538	-.242	-.349	.97127
610051	43.7 (38.1)	.429 (.516)	.038 (.418)	2.34	79.3	.154 (.048)* ▼	.17	13.406	-.586	.240	.96470
630033	23.4 (52.0)	.761 (.710)	.056 (.297)	2.32	86.6	.185 (.025)* ▶	.22	19.208	-.741	.284	.94578
640001	-74.2 (54.1)	2.22 (.744)	.332 (.008)	.691*	108.3	.116 (.069) ▼	.17	18.681	-.468	.345	.97494
670034	41.0 (40.4)	.310 (.553)	.017 (.582)	2.76	66.8	.056 (.155)	.22	14.790	1.049*	.590	.88564*
740077	-92.0 (58.2)	2.64 (.792)	.395 (.004)	1.77	127.4	.033 (.226)	.16	19.963	-.493	.137	.95872
740087	63.5 (35.2)	.318 (.482)	.023 (.518)	2.16	89.9	.024 (.256)	.14	12.720	.115	-.668	.96974
740099	-85.1 (43.7)	2.62 (.596)	.505 (.000)	1.49	132.8	.043 (.182)	.12	16.137	-.416	-.841	.94640

---

[1] Standard Error [2] Significance level * - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression
DW - Durbin-Watson test for serial correlation YM - Estimated yield mean
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE A-1 Corn - continued

FARMNUM	a ( )[1]	b ( )[1]	R ( )[2]	DW	YM	α_2 ( )[2]	CV	STD DEV	$\sqrt{b_1}$	β_2	W-test
740127	-77.3 (35.7)	2.08 (.487)	.489 (.000)	2.51	95.3	.071 (.121)	.14	13.187	-.744	1.411*	.95171
740155	12.9 (31.1)	1.01 (.430)	.245 (.031)	2.18	95.0	.061 (.152)	.11	9.998	-.722	.199	.94268
740167	-108.2 (64.8)	2.07 (.897)	.360 (.007)	2.36	119.5	.053 (.170)	.18	21.172	-.413	-.539	.94081
740169	-29.8 (54.3)	1.62 (.744)	.209 (.042)	2.39	105.3	.026 (.247)	.18	19.437	-.852*	.810	.94227
760086	25.7 (55.7)	.766 (.768)	.055 (.322)	1.08*	88.5	.009 (.347)	.22	19.075	-.148	-.913	.96415
760106	-18.3 (51.1)	1.77 (.704)	.271 (.022)	1.80	128.6	.047 (.184)	.23	17.985	-.650	.819	.94814
760540	-129.0 (47.6)	2.92 (.659)	.521 (.000)	1.38	110.5	.007 (.362)	.15	16.721	-.185	-1.083	.94452
770005	-5291.5 (1122.2)	1.46 (.306)	.547 (.000)	1.93	112.3	.039 (.195)	.12	13.843	.015	-.767	.96299
770626	118.2 (45.2)	-.316 (.631)	.015 (.623)	1.41	92.6	.069 (.145)	.15	13.486	-.518	-.363	.93648
780007	-49.3 (16.3)	1.93 (.257)	.602 (.000)	2.01	111.0	.000 (.438)	.17	18.485	-.352	-.281	.96665
780014	-126.9 (60.1)	3.37 (.820)	.470 (.001)	2.38	153.0	.030 (.225)	.15	22.196	.525	.359	.97666
790115	-89.6 (41.3)	2.86 (.567)	.586 (.000)	2.14	148.1	.007 (.361)	.10	14.945	-.536	.211	.96542
790278	.291 (49.0)	1.50 (.665)	.220 (.037)	2.48	125.0	.069 (.130)	.14	17.173	-.023	-1.284*	.94449
790726	34.8 (42.1)	.959 (.575)	.127 (.112)	1.64	114.5	.011 (.325)	.14	15.561	.486	.600	.94107
800173	146.0 (47.0)	-1.05 (.656)	.138 (.129)	2.21	61.9	.046 (.196)	.23	14.011	-.450	.429	.97045
800309	115.7 (38.2)	-.344 (.522)	.022 (.518)	2.17	87.1	.055 (.152)	.16	14.120	-.845*	.938	.93357
800314	-142.8 (55.5)	2.91 (.758)	.464 (.001)	1.58	98.8	.048 (.182)	.18	18.174	.729	1.969*	.95724

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression

DW - Durbin-Watson test for serial correlation YM - Estimated yield mean

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estiimated standard deviation $\sqrt{b_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE A-1 Corn - continued

FARMNUM	a ( )[1]	b ( )[1]	R ( )[2]	DW	Y M	α_2 ( )[2]	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-test
810001	-3.25 (35.7)	1.10 (.488)	.221 (.036)	2.69	88.5	.016 (.297)	.15	13.148	-.528	.033	.95935
810038	-34.2 (30.8)	1.52 (.418)	.423 (.002)	2.42	92.0	.076 (.118)	.11	10.494	.117	.078	.98388
810516	-31.3 (79.7)	1.68 (1.07)	.126 (.136)	2.33	108.8	.030 (.238)	.25	27.509	-.294	-.791	.96046

---

[1] Standard Error [2] Significance level • - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression
DW - Durbin-Watson test for serial correlation YM - Estimated yield mean
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE A-1 Wheat

FARMNUM	a ( )[1]	b ( )[1]	R ( )[2]	DW	Y M	α_2 ( )[2]	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-test
080031	44.0 (32.1)	-.321 (.446)	.000 (.943)	2.60	41.4	.078 (.130) ▲	.25	10.270	.562	1.301*	.97104
080392	62.2 (24.3)	-.199 (.334)	.019 (.559)	2.50	45.8	.122 (.065)	.18	8.402	-.051	-.692	.97143
120396	2.03 (5.23)	.567 (.082)	.543 (.000)	2.09	49.1	.050 (.076)	.13	6.375	.118	-.090	.97660
130035	3.36 (5.29)	.574 (.083)	.555 (.000)	.937*	51.0	.008 (.287)	.13	6.437	.752	.848	.96097
130203	20.4 (14.3)	.405 (.196)	.191 (.054)	2.97*	54.1	.009 (.340)	.09	5.135	-.019	.028	.96413
190027	18.8 (15.9)	.440 (.210)	.184 (.059)	2.05	55.3	.000 (.451)	.10	5.788	-1.049*	.713	.91425
230044	13.8 (7.8)	.310 (.120)	.155 (.013)	2.41	40.2	.000 (.483)	.23	9.141	1.140*	2.613*	.93506*
290018	20.4 (10.5)	.390 (.170)	.130 (.026)	2.18	52.1	.064 (.062)	.22	11.506	-2.426*	10.012*	.81719*
340073	44.2 (26.4)	.005 (.361)	.000 (.989)	2.70	44.7	.202 (.020)* ▶	.22	9.784	.069	.844	.96439
340093	10.6 (24.3)	.570 (.320)	.151 (.100)	2.16	58.2	.001 (.442)	.14	8.134	.368	-.893	.94057
380080	20.4 (20.3)	.293 (.278)	.058 (.305)	2.36	44.8	.009 (.345)	.17	7.432	.605	-.375	.94573
380485	7.1 (18.9)	.516 (.258)	.181 (.061)	1.36	49.9	.082 (.110) ▲	.14	6.953	.038	.118	.96005
380486	7.9 (17.1)	.553 (.234)	.236 (.030)	2.28	53.9	.000 (.499)	.12	6.288	-.210	-.852	.93978
380493	45.1 (25.9)	.129 (.354)	.000 (.971)	1.72	46.2	.018 (.276)	.21	9.596	-1.458*	2.290*	.85966*
390002	7.8 (7.7)	.436 (.129)	.256 (.002)	1.74	41.0	.031 (.152)	.19	7.624	.093	1.812	.97449
390095	6.4 (5.1)	.543 (.080)	.534 (.000)	2.06	51.5	.002 (.381)	.12	6.219	-.413	-.319	.95934

---

[1] Standard Error [2] Significance level * - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression
DW - Durbin-Watson test for serial correlation YM - Estimated yield mean
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE A-1 Wheat - continued

FARMNUM	a ( )^1	b ( )^1	R ( )^2	DW	Y M	α_i ( )^2	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-test
530067	32.2 (15.6)	.149 (.212)	.026 (.490)	1.94	44.6	.017 (.286)	.12	5.540	-.090	-1.062	.95111
540001	25.3 (31.1)	.081 (.425)	.002 (.850)	1.95	32.1	.028 (.237)	.36	11.482	-1.332*	1.567*	.83500*
570321	-39.9 (32.8)	.872 (.451)	.171 (.069)	1.40	31.6	.071 (.126)	.36	11.337	.237	-.320	.96712
580219	-12.2 (21.1)	1.01 (.290)	.413 (.003)	1.60	72.0	.134 (.060) ▲	.10	7.220	.979	.087	.87899*
590010	13.4 (6.7)	.379 (.104)	.262 (.001)	1.89	44.9	.023 (.173)	.17	7.422	.416	-.281	.96580
590020	-10.2 (27.8)	.716 (.383)	.162 (.078)	1.99	48.4	.007 (.361)	.20	9.622	-1.215*	1.258*	.88614*
590093	3.03 (29.0)	.548 (.396)	.091 (.183)	1.98	48.5	.000 (.450)	.22	10.718	-.351	-.274	.97736
590133	-21.6 (28.8)	.781 (.389)	.191 (.061)	1.59	43.2	.075 (.127)	.22	9.680	-.077	-.498	.97086
610043	105.5 (24.5)	-.789 (.334)	.247 (.030)	1.90	39.9	.000 (.490)	.21	8.329	.496	-.130	.96449
610051	23.2 (18.6)	.345 (.253)	.093 (.190)	1.45	52.0	.002 (.412)	.13	6.785	-.216	-.294	.94052
620432	.366 (5.55)	.603 (.087)	.580 (.000)	1.95	49.2	.000 (.459)	.12	5.901	.063	1.086	.98755
740169	4.4 (30.4)	.637 (.417)	.114 (.145)	2.22	57.2	.094 (.093)	.19	10.902	-.373	.314	.98220
760086	12.5 (29.0)	.360 (.402)	.047 (.383)	2.11	42.1	.072 (.139)	.22	9.343	.447	.402	.97153
780001	49.6 (25.0)	-.160 (.346)	.012 (.649)	2.35	36.5	.073 (.131) ▼	.23	8.294	-.192	-1.016	.95789
780014	65.7 (25.5)	-.295 (.353)	.039 (.416)	2.23	41.2	.085 (.112) ▲	.21	8.491	-.250	.167	.97710
790724	20.5 (41.5)	.327 (.577)	.019 (.578)	2.19	47.1	.153 (.053)	.27	12.925	-1.567*	3.336*	.87403*
790726	38.7 (28.1)	.236 (.379)	.022 (.541)	1.52	58.5	.044 (.192)	.16	9.483	.317	.851	.93643

^{1} Standard Error ^{2} Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression

DW - Durbin-Watson test for serial correlation YM - Estimated yield mean

α_i - test for constant variance CV - coefficient of variation

STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE A-1 Wheat - continued

FARMNUM	a ( )^1	b ( )^1	R ( )^2	DW	Y M	α_2 ( )^2	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-test
800173	91.1 (30.0)	-.781 (.419)	.178 (.081)	1.90	28.6	.032 (.238)	.31	8.957	.157	.414	.97505
810038	34.9 (26.7)	-.001 (.363)	.000 (.997)	1.97	34.7	.153 (.043)* ▲	.26	9.117	.374	-.549	.93383

TABLE A-1 Soybeans

FARMNUM	a ( )^1	b ( )^1	R ( )^2	DW	Y M	α_2 ( )^2	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-test
120291	60.2 (42.5)	-.480 (.597)	.066 (.442)	2.96	23.2	.063 (.228)	.27	6.237	.190	1.129*	.95747
290018	12.9 (12.4)	.266 (.179)	.103 (.155)	2.23	33.9	.149 (.041)* ▶	.14	4.870	-.162	-1.041*	.96528
300101	13.3 (43.6)	.190 (.572)	.009 (.745)	2.02	29.1	.049 (.221)	.32	9.189	.177	-.021	.94752
760015	-95.4 (27.3)	1.57 (.350)	.638 (.001)	2.05	35.2	.211 (.056) ▼	.15	5.318	.447	-1.127*	.92178
770060	-52.1 (58.8)	1.09 (.834)	.176 (.227)	2.83	29.7	.012 (.380)	.24	7.145	.553	-.077	.96782
770626	31.5 (32.2)	-.027 (.437)	.000 (.951)	1.92	29.3	.014 (.349)	.22	6.318	-.932*	1.605*	.89414
780339	-26.5 (46.9)	.706 (.675)	.098 (.320)	2.54	26.3	.032 (.288)	.29	7.698	.469	-.526	.94717

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

a - Y - intercept b - slope (bu/acre) R - R^2 of Regression

DW - Durbin-Watson test for serial correlation YM - Estimated yield mean

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

APPENDIX A - 2

Shapiro - Wilks Group Test For
Corn, Wheat, and Soybeans

TABLE A-2

SHAPIRO-WILKS GROUPED TEST FOR CORN

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	21	.93197	-1.038594
2	18	.96815	.6227713
3	18	.96074	.2341094
4	21	.96517	.2620959
5	19	.92958	-.953444
6	19	.94843	-.3485942
7	20	.93112	-.9870243
8	20	.92469	-1.164591
9	18	.98573	2.086835
10	21	.95022	-.4243431
11	42	.96674	-.3747778
12	21	.95076	-.4031577
13	42	.97095	-8.133841E-02
14	21	.98107	1.409297
15	20	.96375	.2523799
16	18	.94191	-.5070687
17	21	.97341	.7730184
18	20	.90798	-1.569954
19	41	.95375	-1.065522
20	36	.89233*	-2.788057
21	40	.88982*	-2.987758
22	42	.98597	1.476631
23	20	.97457	.9177957
24	21	.99027	2.641272
25	19	.95294	-.1737485
26	19	.90513	-1.550583
27	20	.94496	-.5473041
28	18	.95358	-8.039856E-02
29	18	.94207	-.5017762
30	20	.95916	2.613974E-02
31	20	.97115	.6820565
32	19	.95894	8.489656E-02
33	19	.92537	-1.068149
34	21	.9428	-.6956696
35	20	.95653	-9.288645E-02
36	21	.95228	-.3423433
37	21	.97632	.9908023
38	21	.9693	.5016895
39	20	.9737	.8550544
40	21	.93286	-1.012361
41	39	.98077	.8744488
42	39	.97026	-4.554558E-02
43	20	.89556*	-1.831712
44	20	.94254	-.6310034
45	18	.92217	-1.076297
46	21	.94024	-.7817731
47	19	.96792	.5479765
48	19	.96578	.4273544
49	20	.9406	-.6957931
50	38	.97733	.5623665
51	21	.94565	-.5955353
52	37	.98534	1.470386
53	21	.97417	.8275876
54	20	.92071	-1.267834
55	20	.84178*	-2.733463
56	40	.96551	-.3885017
57	18	.95554	9.026527E-04
58	20	.98168	1.525998
59	21	.90563*	-1.702747
60	20	.97127	.6898642
61	19	.9647	.3691821
62	21	.94578	-.5908532
63	20	.97494	.9451103
64	20	.88564 *	-2.022527
65	19	.95872	.0748024
66	20	.96974	.5926113
67	21	.9464	-.568378
68	21	.95171	-.3653579
69	19	.94268	-.5520301
70	19	.94081	-.614152
71	20	.94227	-.6401396
72	19	.96415	.3402071
73	19	.94814	-.3593445
74	20	.94452	-.5627708
75	21	.96299	.1463523
76	18	.93648	-.6792417
77	39	.96665	-.2895036
78	21	.97666	1.017918
79	20	.96542	.3414993
80	20	.94449	-.5638208
81	21	.94107	-.7542338
82	18	.97045	.7611185
83	21	.93357	-.9912066
84	19	.95724	8.186817E-03
85	20	.95935	3.501606E-02
86	20	.98388	1.761841
87	19	.96046	.1560588
			-14.76958

* - significant at the 5% level.

Z = -1.583465

α = .057

TABLE A-2

SHAPIRO-WILKS GROUPED TEST FOR WHEAT

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	18	.97104	.7982721
2	20	.97143	.7003183
3	42	.9766	.384738
4	40	.96097	-.6551953
5	20	.96413	.2723088
6	20	.91425	-1.426054
7	39	.93506*	-1.73741
8	38	.81719*	-4.138614
9	21	.96439	.2199092
10	19	.94057	-.621995
11	20	.94573	-.5199619
12	20	.96005	6.805992E-02
13	20	.93978	-.7225971
14	21	.85966*	-2.551613
15	35	.97449	.3779249
16	42	.95934	-.8134561
17	20	.95111	-.3182278
18	20	.835 *	-2.829352
19	20	.96712	.436533
20	19	.87899	-2.057284
21	39	.9658	-.3432498
22	20	.88614*	-2.013246
23	21	.97736	1.074972
24	19	.97086	.7268873
25	19	.96449	.358069
26	20	.94052	-.698423
27	36	.98755	1.82912
28	20	.9822	1.579147
29	18	.97153	.8296848
30	19	.95789	3.717089E-02
31	19	.9771	1.172536
32	18	.87403*	-2.060686
33	19	.93643	-.7529159
34	18	.97505	1.071929
35	20	.93383	-.9076824
			-13.23038

* - significant at the 5% level

Z = -2.236343

$\hat{\alpha}$ = .013

SHAPIRO-WILKS GROUPED TEST FOR SOYBEANS

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	11	.95747	.5556972
2	21	.96528	.2681155
3	14	.94752	1.145172E-02
4	13	.92178	-.6212085
5	10	.96782	1.060582
6	13	.89414	-1.183999
7	12	.94717	.1382389
			.228878

Z = .0865077

$\hat{\alpha}$ = .531

APPENDIX B

Individual Farm Statistics by Soil Group
For Corn and Wheat

TABLE B-1

Parameters For Corn -- Soil Group M1 -- Clayey soils

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	Y = A + B TIME			RESIDUALS							
				A[1] ( )	B[1] ( )	DW	α_2[2] ( )	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
110225	63-83	72	20	-116.2 (64.3)	3.060 (.877)	.90*	.042 (.192)	-47.0	40.3	.17	23.685	-.581	.086	.92469
410106	63-83	83	20	58.1 (35.1)	.309 (.483)	2.56	.173 (.033)* ▲	-28.1	18.0	.15	12.125	-.940*	1.025	.89056*
460007	63-83	80,81	19	45.2 (29.2)	.608 (.403)	1.89	.152 (.049)* ▲	-21.4	20.9	.10	9.972	-.302	.487	.96792
530013	42-83	47,48,50 78,83	37	-36.6 (13.3)	1.490 (.209)	1.59	.001 (.426)	-30.6	31.3	.17	14.438	-.129	-.306	.98534
590010	42-83	52,57	40	-27.0 (12.1)	1.270 (.188)	1.76	.066 (.054) ▶	-31.4	26.3	.18	14.468	.202	-.499	.96551
610043	63-83	65	20	-43.9 (52.9)	1.650 (.718)	2.87*	.009 (.341)	-37.1	34.4	.20	18.538	-.242	-.349	.97127
610051	63-83	67,69	19	43.7 (38.1)	.429 (.516)	2.34	.154 (.048)* ▼	-29.3	21.1	.17	13.406	-.586	.240	.96470
800309	63-83	0	21	115.7 (38.2)	-.344 (.522)	2.17	.055 (.152)	-34.2	21.6	.16	14.120	-.845*	.938	.93357

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)

DW - Durbin-Watson test for serial correlation

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE B-2

Parameters For Corn -- Soil Group M5 -- Sandy soils

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A^1 ( )	B^1 ( )	DW	α_2 ( )^2	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
030173	63-83	0	21	-29.2 (39.1)	1.760 (.534)	1.93	.032 (.215)	-29.5	34.8	.120	14.447	.048	.972	.96517
370145	63-83	0	21	-20.4 (50.9)	1.360 (.695)	1.23	.051 (.162)	-28.0	33.2	.210	18.821	.429	-.838	.94280
540001	63-83	73	20	77.6 (50.8)	-.163 (.694)	1.31	.000 (.463)	-36.0	23.8	.290	18.749	-.302	-1.224	.92071
640001	63-83	83	20	-74.2 (54.1)	2.220 (.744)	.69*	.116 (.069) ▼	-42.8	31.7	.170	18.681	-.468	.345	.97494

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-3

Parameters For Corn -- Soil Group M3 -- Loamy soils

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A[1] ( )	B[1] ( )	DW	α_2[2] ( )	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
030023	63-83	64,82,83	18	177.6 (44.2)	-1.040 (.609)	2.11	.116 (.082) ▶	-23.3	23.2	.14	13.259	-.263	-.646	.96074
080031	63-83	81,83	19	-55.7 (38.6)	1.670 (.534)	1.56	.153 (.048)* ▶	-27.4	27.4	.15	12.611	-.373	1.176*	.94843
080392	63-83	83	20	-19.2 (50.4)	1.790 (.694)	1.51	.015 (.301)	-35.6	22.1	.14	17.426	-.521	-.837	.93112
120004	63-83	0	21	-55.4 (48.3)	2.000 (.659)	2.08	.095 (.086) ▶	-29.5	29.4	.16	17.848	-.014	-.805	.95022
120399	63-83	0	21	17.2 (46.9)	1.090 (.641)	1.71	.003 (.403)	-38.0	33.0	.16	17.351	.247	.394	.95076
190004	63-83	78,82,83	18	14.8 (51.1)	.942 (.711)	1.84	.205 (.029)* ▶	-28.4	23.1	.17	15.914	-.443	-.907	.94191
230044	42-83	73	41	-28.1 (11.5)	1.530 (.182)	1.80	.106 (.018)* ▲	-26.7	29.5	.14	14.048	.507	-.370	.95375
290018	42-83	82,83	40	-28.9 (15.2)	1.960 (.242)	1.62	.003 (.350)	-72.5	31.0	.13	17.506	-1.740*	6.463*	.88982*
300011	42-83	0	42	-31.4 (11.5)	1.760 (.181)	1.64	.043 (.092) ▶	-37.4	32.4	.12	14.074	-.375	.387	.98597
300101	63-83	0	21	64.9 (34.4)	.224 (.470)	3.01*	.062 (.137) ▲	-26.0	27.7	.15	12.732	.015	.294	.99027

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)

DW - Durbin-Watson test for serial correlation

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE B-3 continued

Parameters for Corn -- Soil Group M3 -- Loamy soils

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	Y = A + B TIME			RESIDUALS							
				A[1] ( )	B[1] ( )	DW	α_2 ( )[2]	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
330144	63-83	83	20	-63.7 (58.6)	2.120 (.806)	1.98	.000 (.460)	-34.4	44.2	.18	20.251	.104	-.268	.94496
330222	63-83	81,82,82	18	-67.8 (44.0)	1.720 (.614)	1.91	.077 (.132) ▲	-19.5	27.8	.19	13.118	.610	.114	.95358
330225	63-83	74,78,83	18	-142.5 (66.4)	3.070 (.918)	1.67	.114 (.085) ▼	-46.3	38.5	.20	22.300	-.695	.314	.94207
340073	63-83	63	20	32.5 (59.2)	.842 (.803)	1.42	.018 (.288)	-42.2	28.5	.20	20.176	-.259	-.679	.95916
340093	63-83	67	20	-51.4 (60.3)	1.870 (.820)	1.90	.142 (.050)* ▶	-34.4	41.4	.20	21.618	.186	-.791	.97115
340097	63-83	81,82	19	27.5 (43.8)	.909 (.606)	1.42	.000 (.462)	-20.6	32.9	.14	14.560	.599	-.044	.95894
380062	63-83	0	21	-66.3 (45.5)	2.180 (.622)	1.70	.030 (.225)	-30.3	35.1	.15	16.825	.608	.137	.95228
380080	63-83	0	21	-66.3 (45.5)	2.180 (.622)	1.70	.030 (.225)	-30.3	35.1	.15	16.825	.608	.137	.95228
380485	63-83	0	21	-42.4 (46.0)	1.800 (.628)	2.39	.030 (.223)	-27.5	31.0	.16	16.997	.031	-.967	.96930
380486	63-83	76	20	-62.7 (36.0)	2.130 (.493)	2.70	.103 (.083) ▼	-29.2	24.9	.12	13.237	-.409	.492	.97370

---

[1] Standard Error [2] Significance level * - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-3 continued

Parameters for Corn -- Soil Group M3 -- Loamy soils

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A[1] ( )	B[1] ( )	DW	α_2 ( )[2]	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
380493	63-83	0	21	-71.2 (41.3)	2.270 (.564)	2.19	.060 (.141) ▼	-33.0	26.6	.13	15.264	-.676	-.152	.93286
410114	63-83	83	20	-26.2 (37.1)	1.640 (.510)	1.55	.033 (.220)	-33.8	17.7	.12	12.826	-.841*	.928	.94254
440175	63-83	80,82,83	18	-66.5 (44.7)	2.350 (.623)	2.23	.007 (.368)	-26.2	16.3	.11	13.549	-.535	-.782	.92217
440182	63-83	0	21	-82.5 (58.3)	2.320 (.797)	2.52	.160 (.035)* ▲	-27.4	50.2	.19	24.557	.499	-.169	.94024
470006	42-83	71,73,74 81	38	24.0 (14.4)	.725 (.231)	2.07	.034 (.133) ▲	-41.2	31.3	.20	16.862	-.337	-.017	.97733
470119	63-83	0	21	-37.3 (38.4)	1.910 (.524)	1.95	.001 (.440)	-31.2	24.3	.12	14.198	-.500	-.436	.94565
630033	63-83	0	21	23.4 (52.0)	.761 (.710)	2.32	.185 (.025)* ▶	-47.5	24.7	.22	19.208	-.741	.284	.94578
740127	63-83	0	21	-77.3 (35.7)	2.080 (.487)	2.51	.071 (.121) ▶	-36.0	22.1	.14	13.187	-.744	1.411*	.95171
740155	63-83	82,83	19	12.9 (31.1)	1.010 (.430)	2.18	.061 (.152)	-22.0	16.0	.11	9.998	-.722	.199	.94268
780007	42-83	43,78,79	39	-49.3 (16.3)	1.930 (.257)	2.01	.000 (.438)	-40.1	34.9	.17	18.485	-.352	-.281	.96665

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-3 continued

Parameters For Corn -- Soil Group M3 -- Loamy soils

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A[1] ( )	B[1] ( )	DW	α_2[2] ( )	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
790278	63-83	65	20	.3 (49.0)	1.500 (.665)	2.48	.069 (.130) ▶	-26.4	26.4	.14	17.173	-.023	-1.284*	.94449
790726	63-83	0	21	34.8 (42.1)	.959 (.575)	1.64	.011 (.325)	-26.2	36.0	.14	15.561	.486	.600	.94107
810001	63-83	75	20	-3.3 (35.7)	1.100 (.488)	2.69	.016 (.297)	-28.5	23.8	.15	13.148	-.528	.033	.95935
810038	63-83	63	20	-34.2 (30.8)	1.520 (.418)	2.42	.076 (.118) ▶	-19.3	23.2	.11	10.494	.117	.078	.98388
810516	63-83	66,67	19	-31.3 (79.7)	1.680 (1.07)	2.33	.030 (.238)	-53.6	39.8	.25	27.509	-.294	-.791	.96046

---

[1] Standard Error [2] Significance level * - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-4

Parameters For Corn -- Soil Group M4 -- Wet Loamy soils

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A[1] ( )	B[1] ( )	DW	α_2[2] ( )	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
190027	63-83	0	21	17.7 (47.5)	1.000 (.649)	2.17	.002 (.419)	-43.1	34.7	.17	17.575	-.419	.609	.97341
190042	63-83	81	20	32.7 (69.7)	.778 (.955)	1.29	.058 (.151)	-62.7	29.9	.25	24.660	-1.001*	.548	.90798
340560	63-83	82,83	19	-77.9 (68.9)	2.110 (.954)	1.51	.014 (.311)	-43.3	30.2	.24	22.140	-.579	-.912	.92537
370146	63-83	83	20	83.6 (71.0)	.276 (.977)	2.13	.051 (.167)	-38.6	46.9	.23	24.533	.273	-.474	.95653
580219	63-83	79	20	-27.6 (54.3)	1.920 (.745)	2.12	.032 (.221)	-64.6	24.9	.15	19.625	-1.878*	5.607*	.84178*
740077	63-83	64,77	19	-92.0 (58.2)	2.640 (.792)	1.77	.033 (.226)	-40.1	33.7	.16	19.963	-.493	.137	.95872
740087	63-83	79	20	63.5 (35.2)	.318 (.482)	2.16	.024 (.256)	-19.4	24.7	.14	12.720	.115	-.668	.96974
740099	63-83	0	21	-85.1 (43.7)	2.620 (.596)	1.49	.043 (.182)	-31.6	21.6	.12	16.137	-.416	-.841	.94640
740167	63-83	81,83	19	-108.2 (64.8)	2.070 (.897)	2.36	.053 (.170)	-45.9	35.3	.18	21.172	-.413	-.539	.94081
740169	63-83	80	20	-29.8 (54.3)	1.620 (.744)	2.39	.026 (.247)	-50.2	25.8	.18	19.437	-.852	.810	.94227

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-4 continued

Parameters For Corn -- Soil Group M4 -- Wet Loamy soils

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A[1] ( )	B[1] ( )	DW	(α_2)[2]	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
760086	63-83	76,83	19	25.7 (55.7)	.766 (.768)	1.08*	.009 (.347)	-31.7	32.9	.22	19.075	-.148	-.913	.96415
760540	63-83	83	20	-129.0 (47.6)	2.920 (.659)	1.38	.007 (.362)	-30.8	23.6	.15	16.721	-.185	-1.083	.94452
770005	63-83	0	21	-5291.5 (1122.2)	1.460 (.306)	1.93	.039 (.195)	-24.7	24.1	.12	13.843	.015	-.767	.96299
770626	63-83	81,82,83	18	118.2 (45.2)	-.316 (.631)	1.41	.069 (.145)	-28.3	18.9	.15	13.486	-.518	-.363	.93648
790115	63-83	79	20	-89.6 (41.3)	2.860 (.567)	2.14	.007 (.361)	-36.1	23.7	.10	14.945	-.536	.212	.96542
800173	63-83	81,82,83	18	146.0 (38.2)	-1.050 (.522)	2.21	.046 (.152)	-31.5	25.4	.23	14.011	-.450	.429	.97045

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance
CV - coefficient of variation
STD DEV - Estiimated standard deviation
$\sqrt{\beta_1}$ - Skewness
β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-5

Parameters For Corn -- Soil Group M7 -- Loamy soils underlain by sand & gravel

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A ^{1} ( )	B ^{1} ( )	DW	α_2 ( ) ^{2}	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
030001	63-83	0	21	-106.0 (33.0)	2.700 (.451)	2.20	.152 (.039)* ▲	-25.0	15.5	.10	12.215	-.491	-.888	.93197
030022	63-83	63,83	18	-32.9 (44.9)	1.440 (.627)	1.88	.082 (.123) ▶	-31.5	30.3	.16	13.398	-.068	1.572*	.96815
120396	42-83	0	42	-14.3 (12.0)	1.450 (.185)	2.34	.005 (.319)	-29.0	27.2	.14	14.707	-.056	-.832	.96674
130035	42-83	0	42	-43.9 (12.9)	2.040 (.204)	1.62	.064 (.052) ▶	-26.7	39.8	.13	15.835	.225	-.413	.97095
130038	63-83	0	21	-13.7 (46.7)	1.430 (.638)	1.21*	.000 (.484)	-35.2	41.9	.16	17.279	.269	.714	.98107
130203	63-83	80	20	-34.4 (39.2)	1.840 (.538)	2.14	.135 (.054) ▶	-22.2	25.6	.12	14.057	.140	-.790	.96375
300074	63-83	83	20	-5.3 (39.2)	1.290 (.539)	2.21	.007 (.355)	-30.6	21.9	.13	13.536	-.397	.046	.97457
300174	63-83	81,82	19	-7.8 (49.1)	1.250 (.679)	1.92	.013 (.316)	-31.9	22.5	.17	16.305	-.480	-.603	.95294
330040	63-83	76,81	19	-15.0 (34.8)	1.450 (.479)	1.44	.000 (.497)	-22.7	15.7	.12	12.257	-.601	-1.001	.90513
390002	42-83	74,81,83	39	16.9 (9.9)	.650 (.160)	1.12*	.092 (.029)* ▲	-27.0	23.2	.16	11.418	.024	-.041	.98077

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-5 continued

Parameters For Corn -- Soil Group M7 -- Loamy soils underlain by sand & gravel

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	Y = A + B TIME A_1 ( )	 B_1 ( )	 DW	RESIDUALS α_2 ^{2} ( )	 LOW	 HIGH	 CV	 STD DEV	 $\sqrt{\beta_1}$	 β_2	 W-TEST
390095	42-83	64,78,81	39	-8.0 (18.0)	1.120 (.288)	1.65	.005 (.323)	-37.2	53.0	.25	21.189	.266	-.490	.97026
460444	63-85	68	20	-107.3 (45.2)	2.600 (.615)	2.61	.009 (.343)	-40.6	24.8	.15	16.341	-.842*	.711	.94060
590020	63-83	64,79,83	18	-247.1 (48.5)	4.700 (.667)	2.23	.117 (.081) ▶	-30.5	22.8	.11	15.185	-.197	-.424	.95554
590093	63-83	64	20	-145.3 (52.8)	3.230 (.717)	1.54	.000 (.484)	-32.1	40.3	.15	18.276	.295	.102	.98168
590133	63-83	0	21	-87.5 (63.1)	2.420 (.862)	1.39	.007 (.351)	-38.3	32.4	.20	23.326	-.363	-1.409*	.90563*
760106	63-83	76,81	19	-18.3 (51.1)	1.770 (.704)	1.80	.047 (.184)	-42.6	34.3	.23	17.985	-.650	.819	.94814
780014	63-83	0	21	-126.9 (60.1)	3.370 (.820)	2.38	.030 (.225)	-39.9	51.8	.15	22.196	.525	.359	.97666
800314	63-83	64,82	19	-142.8 (55.5)	2.910 (.758)	1.58	.048 (.182)	-33.6	48.7	.18	18.174	.729	1.969*	.95724

( ) Standard Error ^{2} Significance level   * - Significant at 5% level

▲ - Increasing variance   ▼ - Decreasing variance   ▶ - Constant variance

A - Y - intercept   B - slope (bu/acre)

DW - Durbin-Watson test for serial correlation

α_2 - test for constant variance   CV - coefficient of variation

STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE B-6

Parameters For Wheat -- Soil Group M1 -- Clayey soils

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A [1] ( )	B [1] ( )	DW	α_2 [2] ( )	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
590010	42-83	42,43,52	39	13.4 (6.7)	.379 (.104)	1.89	.023 (.173)	-12.2	17.5	.17	7.422	.416	-.281	.96580
610043	63-83	65,80	19	105.5 (24.5)	-.789 (.334)	1.90	.000 (.490)	-13.6	16.7	.21	8.329	.496	-.130	.96449
610051	63-83	69	20	23.2 (18.6)	.345 (.253)	1.45	.002 (.412)	-12.8	14.0	.13	6.785	-.216	-.294	.94052
620432	42-83	45,50,53 76,82,83	36	.4 (5.55)	.603 (.087)	1.95	.000 (.459)	-15.0	16.0	.12	5.901	.063	1.086	.98755

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)

DW - Durbin-Watson test for serial correlation

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE B-7

Parameters For Wheat -- Soil Group M4 -- Wet Loamy soils

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	Y = A + B TIME A[1] ( )	 B[1] ( )	 DW	RESIDUALS α_2[2] ( )	 LOW	 HIGH	 CV	 STD DEV	 $\sqrt{\beta_1}$	 β_2	 W-TEST
190027	63-83	78	20	18.8 (15.9)	.440 (.21)	2.05	.000 (.451)	-14.2	7.200	.10	5.788	-1.049 *	.713	.91425
580219	63-83	79,82	19	-12.2 (21.2)	1.010 (.29)	1.60	.134 (.060)	-8.9	14.100	.10	7.220	.979 *	.087	.87899 *
740169	63-83	82	20	4.4 (30.4)	.637 (.417)	2.22	▲ .094 (.093)	-25.8	19.900	.19	10.902	-.373	.314	.98220
760086	63-83	76,81,83	18	12.5 (29.0)	.360 (.402)	2.11	▶ .072 (.139)	-14.9	21.900	.22	9.343	.447	.402	.97153
800173	63-83	81,82,83	18	91.1 (30.0)	-.781 (.419)	1.90	▶ .032 (.238)	-16.7	19.300	.31	8.957	.157	.414	.97505

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)

DW - Durbin-Watson test for serial correlation

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estiimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE B-8

Parameters For Wheat -- Soil Group M3 -- Loamy soils

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	Y = A + B TIME A $(\ )^1$	 B $(\ )^1$	 DW	RESIDUALS α_2 $(\ )^2$	 LOW	 HIGH	 CV	 STD DEV	 $\sqrt{\beta_1}$	 β_2	 W-TEST
080031	63-82	77,81,83	18	44.0 (32.11)	-.322 (.446)	2.60	.078 (.130)	-19.5	25.5	.25	10.270	.562	1.301*	.97104
080392	63-82	83	20	62.2 (24.34)	-.199 (.334)	2.50	▲ .122 (.065)	-16.8	13.1	.18	8.402	-.051	-.692	.97143
230044	42-83	66,73,78	39	13.8 (7.6)	.310 (.12)	2.41	▶ .000 (.483)	-14.5	32.1	.23	9.141	1.140*	2.613*	.93506
290018	42-80	77,81,82, 83	38	20.4 (10.5)	.390 (.17)	2.18	.064 (.062)	-51.3	17.2	.22	11.506	-2.426*	10.012*	.81719*
340073	63-83	0	21	44.2 (26.4)	.005 (.361)	2.70	▶ .202 (.020)*	-22.7	19.2	.22	9.784	.069	.844	.96439
340093	63-83	65,66	19	10.6 (24.3)	.570 (.32)	2.16	▶ .001 (.442)	-11.6	14.2	.14	8.134	.368	-.893	.94057
380080	63-83	77	20	20.4 (20.3)	.293 (.278)	2.36	.009 (.345)	-11.0	15.3	.17	7.432	.605	-.375	.94573
380485	63-83	71	20	7.1 (18.9)	.516 (.258)	1.36	.082 (.110)	-11.8	15.6	.14	6.953	.038	.118	.96005
380486	63-83	76	20	7.9 (17.1)	.553 (.234)	2.28	▲ .000 (.499)	-10.3	9.4	.12	6.288	-.210	-.852	.93978
380493	63-83	0	21	45.1 (25.9)	.129 (.354)	1.72	.018* (.276)	-26.1	13.9	.21	9.596	-1.458*	2.290*	.85966*

---

[1] Standard Error [2] Significance level * - Significant at 5% level

▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)

DW - Durbin-Watson test for serial correlation

α_2 - test for constant variance CV - coefficient of variation

STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis

W-test - Shapiro-Wilks test for normality

TABLE 8-8 continued

Parameters For Wheat -- Soil Group M3 -- Loamy soils

FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	Y = A + B TIME A[1] ( )	B[1] ( )	DW	RESIDUALS α_2[2] ( )[1]	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
780001	63-83	80,83	19	49.6 (25.0)	-.160 (.346)	2.35	.073 (.131) ▼	-15.9	11.9	.23	8.294	-.192	-1.016	.95789
790726	63-83	65,67	19	38.7 (28.1)	.236 (.379)	1.52	.044 (.192)	-15.8	23.5	.16	9.483	.317	.851	.93643
810038	63-83	63	20	34.9 (26.7)	-.001 (.363)	1.97	.153 (.043)* ▲	-14.8	17.2	.26	9.117	.374	-.549	.93383

---

[1] Standard Error [2] Significance level * - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

TABLE B-9

Parameters For Wheat -- Soil Group M7 -- Loamy soils underlain by sand & gravel

				Y = A + B TIME			RESIDUALS							
FARM NUMBER	TIME SERIES	MISSING YEARS	TOTAL CASES	A ^{1} ( )	B ^{1} ( )	DW	α_2 ^{2} ( )	LOW	HIGH	CV	STD DEV	$\sqrt{\beta_1}$	β_2	W-TEST
120396	42-83	0	42	2.0 (5.23)	.567 (.082)	2.09	.050 (.0760) ▶	-13.1	14.0	.13	6.375	.118	-.090	.97660
130035	42-83	65,70	40	3.4 (5.29)	.574 (.083)	.94*	.008 (.287)	-11.5	19.7	.13	6.437	.752	.848	.96097
130203	63-83	80	20	20.5 (14.34)	.405 (.196)	2.97*	.009 (.340)	-10.0	10.3	.09	5.135	-.019	.028	.96413
390002	42-76	0	35	7.8 (7.7)	.436 (.129)	1.74	.031 (.152)	-20.5	21.4	.19	7.624	.093	1.812*	.97449
390095	42-83	0	42	6.4 (5.1)	.543 (.080)	2.06	.002 (.381)	-13.4	11.4	.12	6.219	-.413	-.319	.95934
590020	63-83	83	20	-10.2 (27.8)	.716 (.383)	1.99	.007 (.361)	-25.2	10.2	.20	9.622	-1.215*	1.258*	.88614*
590093	63-83	0	21	3.0 (29.0)	.548 (.396)	1.98	.000 (.450)	-22.2	19.5	.22	10.718	-.351	-.274	.97736
590133	63-83	64,68	19	-21.6 (28.8)	.781 (.389)	1.59	.075 (.127) ▶	-17.3	16.7	.22	9.680	-.077	-.498	.97086
780014	63-83	81,82	19	65.7 (25.5)	-.295 (.353)	2.23	.085 (.112) ▲	-19.0	14.3	.21	8.491	-.250	.167	.97710
790724	63-83	78,82,83	18	20.5 (41.5)	.327 (.577)	2.19	.153 (.053) ▶	-37.8	16.4	.27	12.925	-1.567*	3.336*	.87403*

---

[1] Standard Error [2] Significance level * - Significant at 5% level
▲ - Increasing variance ▼ - Decreasing variance ▶ - Constant variance

A - Y - intercept B - slope (bu/acre)
DW - Durbin-Watson test for serial correlation
α_2 - test for constant variance CV - coefficient of variation
STD DEV - Estimated standard deviation $\sqrt{\beta_1}$ - Skewness β_2 - Kurtosis
W-test - Shapiro-Wilks test for normality

APPENDIX C

Shapiro - Wilks Group Test by Soil Group
For Corn and Wheat

TABLE C-1

SHAPIRO-WILKS GROUPED TEST FOR CORN - SG M1

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	20	.92469	-1.164591
2	20	.89556*	-1.831712
3	19	.96792	.5479765
4	37	.98534	1.470386
5	40	.96551	-.3885017
6	20	.97127	.6898642
7	19	.9647	.3691821
8	21	.93357	-.9912066
			-1.298602

* - significant at the 5% level

$Z = -.459125$

$\hat{\alpha} = .326$

TABLE C-2

SHAPIRO-WILKS GROUPED TEST FOR CORN SG M3

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	18	.96074	.2341094
2	19	.94843	-.3485942
3	20	.93112	-.9870243
4	21	.95022	-.4243431
5	21	.95076	-.4031577
6	18	.94191	-.5070687
7	41	.95375	-1.065522
8	40	.88982*	-2.987758
9	42	.98597	1.476631
10	21	.99027	2.641272
11	20	.94496	-.5473041
12	18	.95358	-8.039856E-02
13	18	.94207	-.5017762
14	20	.95916	2.613974E-02
15	20	.97115	.6820565
16	19	.95894	8.489656E-02
17	21	.95228	-.3423433
18	21	.97632	.9908023
19	21	.9693	.5016895
20	20	.9737	.8550544
21	21	.93286	-1.012361
22	20	.94254	-.6310034
23	18	.92217	-1.076297
24	21	.94024	-.7817731
25	38	.97733	.5623665
26	21	.94565	-.5955353
27	21	.94578	-.5908532
28	21	.95171	-.3653579
29	19	.94268	-.5520301
30	39	.96665	-.2895036
31	20	.94449	-.5638208
32	21	.94107	-.7542338
33	20	.95935	3.501606E-02
34	20	.98388	1.761841
35	19	.96046	.1560588
			-5.400125

* - significant at the 5% level

$Z = -.9127878$

$\hat{\alpha} = .181$

TABLE C-3
SHAPIRO-WILKS GROUPED TEST FOR CORN - SG M4

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	21	.97341	.7730184
2	20	.90798	-1.569954
3	19	.92537	-1.068149
4	20	.95653	-9.288645E-02
5	20	.84178*	-2.733463
6	19	.95872	.0748024
7	20	.96974	.5926113
8	21	.9464	-.568378
9	19	.94081	-.614152
10	20	.94227	-.6401396
11	19	.96415	.3402071
12	20	.94452	-.5627708
13	21	.96299	.1463523
14	18	.93648	-.6792417
15	20	.96542	.3414993
16	18	.97045	.7611185
			-5.499525

* - significant at the 5% level

$Z = -1.374881$

$\hat{\alpha} = .085$

TABLE C-4
SHAPIRO-WILKS GROUPED TEST FOR CORN SG M7

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	21	.93197	-1.038594
2	18	.96815	.6227713
3	42	.96674	-.3747778
4	42	.97095	-8.133841E-02
5	21	.98107	1.409297
6	20	.96375	.2523799
7	20	.97457	.9177957
8	19	.95294	-.1737485
9	19	.90513	-1.550583
10	39	.98077	.8744488
11	39	.97026	-4.554558E-02
12	20	.9406	-.6957931
13	18	.95554	9.026527E-04
14	20	.98168	1.525998
15	21	.90563	-1.702747
16	19	.94814	-.3593445
17	21	.97666	1.017918
18	19	.95724	8.186817E-03
			.6072252

$Z = .1431244$

$\hat{\alpha} = .555$

TABLE C-5
SHAPIRO-WILKS GROUPED TEST FOR WHEATM3

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	18	.97104	.7982721
2	20	.97143	.7003183
3	39	.93506	-1.73741
4	38	.81719 *	-4.138614
5	21	.96439	.2199092
6	19	.94057	-.621995
7	20	.94573	-.5199619
8	20	.96005	6.805992E-02
9	20	.93978	-.7225971
10	21	.85966 *	-2.551613
11	19	.95789	3.717089E-02
12	19	.93643	-.7529159
13	20	.93383	-.9076824
			-10.12906

* - significant at the 5% level

$Z = -2.809296$

$\hat{\alpha} = .0026$

TABLE C-6
SHAPIRO-WILKS GROUPED TEST FOR WHEATM7

CASE NUMBER	# OF OBS.	W-TEST	G(I) VALUE
1	42	.9766	.384738
2	40	.96097	-.6551953
3	20	.96413	.2723088
4	35	.97449	.3779249
5	42	.95934	-.8134561
6	20	.88614 *	-2.013246
7	21	.97736	1.074972
8	19	.97086	.7268873
9	19	.9771	1.172536
10	18	.87403 *	-2.060686
			-1.533217

* - significant at the 5% level

$Z = -.4848456$

$\hat{\alpha} = .315$

APPENDIX D

Equivalent Growth Test by Soil Group
For Corn and Wheat

TABLE D - 1

F - tests for corn

Soil Group M1

$$F_{M1} = \frac{50,899.91 - 44,661.65/8}{44,661.65/180} = 25.14$$

$F_{(8,180)}$ 1.94

Soil Group M3

$$F_{M3} = \frac{719,692.5 - 208,271.3/35}{208,271.3/728} = 51.07$$

$F_{(35,278)}$ 1.46

Soil Group M4

$$F_{M4} = \frac{129,579.19 - 102,375.77/16}{102,375.77/283} = 62.06$$

$F_{(16,283)}$ 1.67

Soil Group M5

$$F_{M5} = \frac{26,951.71 - 24,568.82/4}{24,568.82/74} = 1.79$$

$F_{(4,74)}$ 2.45

Soil Group M7

$$F_{M7} = \frac{305,903.07 - 115,737.5/18}{115,737.5/402} = 36.70$$

$F_{(18,402)}$ 1.57

TABLE D - 2

F - tests for wheat

Soil Group M1

$$F_{M1} = \frac{52,287.10 - 5,434.27/4}{5,434.27/106} = 228.47$$

$$F_{(4,106)} \quad 2.45$$

Soil Group M3

$$F_{M3} = \frac{93,488.67 - 23,310.38/13}{23,310.38/268} = 62.06$$

$$F_{(13,268)} \quad 1.67$$

Soil Group M4

$$F_{M4} = \frac{32,283.97 - 6,681.27/5}{6,681.27/85} = 65.14$$

$$F_{(5,85)} \quad 2.29$$

Soil Group M7

$$F_{M7} = \frac{110,871.32 - 17,225.91/10}{17,225.91/256} = 139.17$$

$$F_{(10,256)} \quad 1.83$$

APPENDIX E

Equivalent Variance Test by Soil Group For Corn and Wheat

TABLE E-1

```
---------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M1 FOR CORN
---------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          99.44     18
          2          147.02    19
          3          179.72    18
          4          199.37    20
          5          208.46    36
          6          209.32    39
          7          343.66    19
          8          560.98    19
---------------------------------------------------------
Q1 STATISTIC = 19.62116       MEAN OF VARIANCE     = 243.4963
Q  STATISTIC = 19.61365       STD DEV OF VARIANCE = 48.29898
---------------------------------------------------------
```

$\chi^2_{7,.05} = 14.06$

TABLE E-2

```
---------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M3 FOR CORN
---------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          99.96001  18
          2          110.1241  19
          3          159.0373  18
          4          162.1038  20
          5          164.5063  19
          6          172.0819  17
          7          172.8699  19
          8          173.897   20
          9          175.2182  19
          10         175.8011  17
          11         183.5754  17
          12         197.3463  40
          13         198.0775  41
          14         201.5832  20
          15         211.9936  18
          16         224.5203  20
          17         232.9897  20
          18         242.1447  20
          19         253.2554  17
          20         283.0807  20
          21         284.327   37
          22         288.898   20
          23         294.9119  19
          24         301.0572  20
          25         303.6655  19
          26         306.4601  39
          27         318.5511  20
          28         341.6953  38
          29         368.9473  20
          30         407.071   19
          31         410.103   19
          32         467.3379  19
          33         497.29    17
          34         603.0462  20
          35         756.7451  18
---------------------------------------------------------
Q1 STATISTIC = 69.72656       MEAN OF VARIANCE     = 278.4078
Q  STATISTIC = 69.7085        STD DEV OF VARIANCE = 23.25701
---------------------------------------------------------
```

$\chi^2_{34,.05} = 55.75$

TABLE E-3

```
---------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M4 FOR CORN
---------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          161.7984  19
          2          181.8722  17
          3          191.6287  20
          4          196.3081  17
          5          223.353   19
          6          260.4028  20
          7          279.5919  19
          8          308.8807  20
          9          363.8557  18
          10         377.797   19
          11         385.1406  19
          12         398.5213  18
          13         448.2537  18
          14         490.1796  18
          15         601.8683  19
          16         608.1156  19
---------------------------------------------------------
Q1 STATISTIC = 24.69873      MEAN OF VARIANCE    = 342.348
Q  STATISTIC = 24.69408      STD DEV OF VARIANCE = 34.57819
---------------------------------------------------------
```

$X^2_{15,.05} = 24.99$

TABLE E-4

```
---------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M5 FOR CORN
---------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          208.72    20
          2          348.98    19
          3          351.53    19
          4          354.23    20
---------------------------------------------------------
Q1 STATISTIC = 1.848816      MEAN OF VARIANCE    = 315.865
Q  STATISTIC = 1.848687      STD DEV OF VARIANCE = 30.94401
---------------------------------------------------------
```

$X^2_{3,.05} = 7.81$

TABLE E-5

```
------------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M7 FOR CORN
------------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          130.3707  38
          2          149.2062  20
          3          150.234   18
          4          179.5064  17
          5          183.2233  19
          6          197.5993  19
          7          216.2958  41
          8          230.5842  17
          9          250.7472  41
         10          265.853   18
         11          267.0283  19
         12          298.5638  20
         13          323.4602  18
         14          330.2943  18
         15          334.0122  19
         16          448.9737  38
         17          492.6624  20
         18          544.1023  20
------------------------------------------------------------
Q1 STATISTIC = 35.88745      MEAN OF VARIANCE    = 277.3732
Q  STATISTIC = 35.87985      STD DEV OF VARIANCE = 27.29567
------------------------------------------------------------
```

$\chi^2_{17,.05} = 27.58$

TABLE E-6

```
------------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M1FOR WHEAT
------------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          34.82     38
          2          46.04     18
          3          55.09     19
          4          69.37     35
------------------------------------------------------------
Q1 STATISTIC = 4.434876      MEAN OF VARIANCE    = 51.33
Q  STATISTIC = 4.432509      STD DEV OF VARIANCE = 6.325264
------------------------------------------------------------
```

$\chi^2_{3,.05} = 7.81$

TABLE E-7

```
----------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M3FOR WHEAT
----------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          39.54     19
          2          48.34     19
          3          55.23     19
          4          66.16001  18
          5          68.79     18
          6          70.59     19
          7          83.12     19
          8          83.56     38
          9          89.93     18
          10         92.08     20
          11         95.73001  20
          12         105.47    17
          13         132.39    37
----------------------------------------------------------
Q1 STATISTIC = 14.48023       MEAN OF VARIANCE     = 79.30231
Q  STATISTIC = 14.47032       STD DEV OF VARIANCE = 6.688396
----------------------------------------------------------
```

$\chi^2_{12,.05} = 21.02$

TABLE E-8

```
----------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M4FOR WHEAT
----------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          33.5      19
          2          52.13     18
          3          87.29     17
          4          89.57     17
          5          118.85    19
----------------------------------------------------------
Q1 STATISTIC = 8.702392       MEAN OF VARIANCE     = 76.26801
Q  STATISTIC = 8.692053       STD DEV OF VARIANCE = 13.45124
----------------------------------------------------------
```

$\chi^2_{4,.05} = 9.48$

TABLE E-9

```
----------------------------------------------------------
TEST FOR EQUAL VARIANCES ON SOIL GROUP M7FOR WHEAT
----------------------------------------------------------
         FARM       VARIANCE  D.F.
          1          26.37     19
          2          38.68     41
          3          40.64     41
          4          41.43     39
          5          58.13     34
          6          72.1      18
          7          92.58     19
          8          93.7      18
          9          114.88    20
          10         167.06    17
----------------------------------------------------------
Q1 STATISTIC = 35.45374       MEAN OF VARIANCE     = 74.557
Q  STATISTIC = 35.42049       STD DEV OF VARIANCE = 13.04102
----------------------------------------------------------
```

$\chi^2_{9,.05} = 16.91$

APPENDIX F

Computation of the Breakeven Farmer Rate

- long version
- computer program (Basic listing)

BREAK EVEN FARMER RATE
THE LONG VERSION

Soybeans 50% coverage -- 13 bu.

E(Y) = 25 bu SD(Y) = 7.5 bu CV = .30 Price -$5

Probability	Coverage - Yield = Indemnity
(.0011)	(13 - 2)(5) = .0605
(.0015)	(13 - 4)(5) = .0675
(.0031)	(13 - 6)(5) = .1085
(.0062)	(13 - 8)(5) = .155
(.0109)	(13 - 10)(5) = .1635
(.032)	(13 - 12)(5) = .16

Premium=Rate*Coverage*Price

$$\text{Rate} = \frac{\text{Premium}}{\text{Coverage*Price}}$$

Average Indemnity or Premium = .715/(13)(5) =.011 x 100 =1.1% Rate

The average indemnity is the breakeven premium payment for the producer. Each payout is weighted by the probability of its occurrence. The total of these weighted payouts represents the average payout from the insurance pool. To find the breakeven rate for the producer, the breakeven premium (average indemnity payout) is divided by coverage * price. This is then the break even farmer rate or BEFR. This rate should be the same or sufficiently close to the BEFR calculations adapted from Botts & Boles. Checking with Table 7.20, the Botts & Boles calculation is 1.18% vs 1.10% as calculated above.

BREAK EVEN FARMER RATE
COMPUTER VERSION

BASIC PROGRAM LISTING

```
100 CLS
120 INPUT "E(Y) ";EY
140 INPUT "S(Y) ";SY
160 INPUT "COVERAGE ";COV
180 '
190 Y6=COV*EY
200 V1=(EY-YG)/SY
220 V2=(1/SQR(2*3.1415926#))*EXP(-.5*V1*V1)
240 T=1/(1+.33267*V1)
260 V3=1-V2*(.4361836*T-.1201676*T*T+.937289*T*T*T)
280 ALC=(1-V3)*(Y6-EY)+V2*SY
300 BEFR=(ALC/Y6)*100
320 PRINT YG,V1,V2,T,V3,ALC,BEFR
340 END
```

OUTPUT

```
E(Y) ? 70
S(Y) ? 17.5
COVERAGE ? 0.75
 52.5          1             .2419707      .7503733      .8413522
 1.458151      2.777431
```

BIBLIOGRAPHY

Antle, John M. "Incorporating Risk In Production Analysis." American Journal of Agricultural Economics 65(1983):1099-1106.

Botts, Ralph R., and James N. Boles. "Use Of Normal-Curve Theory In Crop Insurance Ratemaking." Journal of Farm Economics 40(1968):733-740.

Conover, W. J. Practical Nonparametric Statistics. 2nd Edition, New York:Wiley & Sons, 1980.

Cooperative Extension Service. "Soil Association Map Of Michigan." East Lansing, Michigan:Michigan State University, December 1981, Extension Bulletin E-1550.

Crop Insure - Farm Protection Computer Model Demonstrating Risk Management. American Association of Crop Insurers, Michigan State University Extension Service, 1988.

Day, Richard H. "Probability Distributions Of Field Crop Yields." Journal Farm Economics 47(1965):713-741.

Durbin J., and G. S. Watson. "Testing For Serial Correlation In Least Squares Regression." Biometrica 37(1950):409-428.

Glejser, H. "A New Test For Heteroscedasticity." Journal of the American Statistical Association 64(1969):316-323.

Harsh, Stephen B., Larry J. Connor, and Gerald D. Schwab. Managing The Farm Business. New Jersey:Prentice-Hall, 1981.

Holt, John. "Remarks About Research On Risk Management Strategies For Agricultural Production Firms." Proceedings of a seminar sponsored by Southern Regional Research Project S-180, San Antonio, TX, March 28-30, 1983.

Holt, John, and Kim B. Anderson. "Teaching Decision Making Under Risk And Uncertainty To Farmers." Paper presented at the Summer Meetings of the American Agricultural Economics Association, July 31-August 3, 1977.

Ikerd, John, and Kim Anderson. "Teaching Risk Rated Management Strategies To Farmers And Ranchers." Paper presented at Workshop For Extension Specialists On Marketing, Risk, And Financial Management, Minneapolis, MN, April 2-4, 1984.

Johnston, John J. Econometric Methods. 3rd Edition, New York:McGraw-Hill, 1984.

Jolly, Robert W. "Risk Management In Agricultural Production." American Journal of Agricultural Economics 65(1983):1107-1113.

King, Robert P. "Crop Insurance Research Needs." Proceedings of a seminar sponsored by Southern Regional Research Project S-180, San Antonio, TX, March 28-30, 1983.

King, Robert P., Fred J. Benson, and J. Roy Black. Agricultural Risk Management Simulator - User Manual. University of Minnesota Extension Service, 1987.

Knight, Frank H. Risk, Uncertainty, And Profit. Boston:Riverside Press, 1921.

Law, Averill M., and W. David Kelton. Simulation Modeling And Analysis. New York:McGraw-Hill, 1982.

Lee, Warren F., Michael D. Boehlje, Aaron G. Nelson, and William G. Murray. Agricultural Finance. 7th Edition, Iowa State University Press, 1980.

Mapp, Harry P. "Crop Insurance Research Needs: Comments." Proceedings of a seminar sponsored by Southern Regional Research Project S-180, San Antonia, TX, March 28-30, 1983.

Nelson, A. Gene, George L. Casler, and Odell L. Walker. Making Farm Decisions In A Risky World: A Guidebook. Oregon State University Extension Service, 1978.

Park, R. E. "Estimation With Heteroscedastic Error Terms." Econometrica 34(1966):888.

Pindyck, Robert S., and Daniel L. Rubinfeld. Econometric Models And Economic Forecasts. 2nd Edition, New York:McGraw-Hill, 1981.

Pope, C. Arden, III., and Earl O. Heady. "The Effects Of Technological Progress And Weather On Corn Belt Crop Yields." Report by the Center for Agricultural and Rural Development, Iowa State University, October 1982.

Schwab, Gerald D., Sherrill B. Nott, Myron P. Kelsey, and Ting-Ing Ho." Michigan Crop And Livestock Estimated 1983 Budgets." Report by Department of Agricultural Economics, Michigan State University, January 1983.

Shapiro, S. S., and M. B. Wilk. "An Analysis Of Variance Test For Normality (complete samples)." Biometrica 52(1965):591-611.

Shapiro, S. S., M. B. Wilk, and H. J. Chen. "A Comparative Study Of Various Tests For Normality." Journal of the American Statistical Association 63(1968):1343-1372.

Stephens, M. A. "EDF Statistics For Goodness Of Fit And Some Comparisons." Journal of the American Statistical Association 69(1974):730-737.

Stewart, Mark B., and Kenneth F. Wallis. Introductory Econometrics, 2nd Edition, New York:Halsted Press, 1981.