1987

THE MEAN AND VARIANCE OF THE MEAN-VARIANCE DECISION RULE

by

James A. Chalfant, Robert N. Collender, and Shankar Subramanian

February, 1987

Decision-making

AAEA paper, 1987

The Mean and Variance of the Mean-Variance Decision Rule

The sampling properties of the mean-variance decision vector are studied. We show that, when the parameters of yield distributions are uncertain, risk in estimation causes bias in these decisions and large variances. Especially with small samples, decisions based on estimated parameters can be very poor estimates of optimal behavior.

THE MEAN AND VARIANCE OF THE MEAN-VARIANCE DECISION RULE

1. Introduction

Since Freund (1956) examined the land allocation problem using quadratic programming, agricultural economists have made extensive use of mean-variance (E-V) analysis to study the choice among uncertain alternatives. Along with the allocation of land, examples include participation in agricultural commodity programs, hedging, and the adoption of new technologies. The widely accepted expected utility approach produces a linear E-V objective when either utility is quadratic or the returns from each alternative are jointly normal.

Since quadratic utility implies increasing risk aversion, it is the latter which serves as the justification for E-V. While some papers have examined the normality of returns (e.g., Day, Buccola) and others have considered alternatives to normality (e.g., Collender and Zilberman), the fact that the decision-maker must estimate the underlying distributions is rarely addressed. Yet, this estimation risk has dramatic implications for the usefulness of E-V analysis and for the actual expected utility to be derived from applying the E-V approach to decision making. This is especially true when there is a limited amount of information available for estimating the distributions in question and holds whether or not the normality assumption is appropriate.

In this paper, we illustrate this problem by examining the sampling properties of the mean-variance decision vector for the land allocation problem. An unbiased decision vector is derived using these results. We also examine the sampling variance of these decision vectors, finding that variation around the correct decision is surprisingly large. The results show that E-V analysis should be undertaken only with strong caveats in many cases.

2. The Setup of the Model

The setup for the land allocation problem is identical to that used in the expected utility moment-generating function approach in papers by Collender and Zilberman and Collender and Chalfant. The problem is to allocate L acres of land to K crops, where returns per acre $\underline{x}$ are distributed as $N_K(\underline{\mu}, \Sigma)$. We assume an exponential utility function

$$u(\pi) = -\exp(-r\pi)$$

where r is the Pratt-Arrow measure of risk aversion and π denotes profits:

$$\pi = \sum_{i=1}^{K} \ell_i \, x_i .$$

We assume that per acre returns are net of production costs, and we treat the technologies as predetermined and consider only the acreage decision.

The first-order conditions for maximizing expected utility involve the derivatives of the moment-generating function with respect to each $t_i = -r\,\ell_i$. They are

$$\frac{M_1}{M} = \frac{M_i}{M}, \qquad i = 2, \ldots, K.$$

These conditions are then equivalent to

$$M^{-1} A \nabla M = \underline{0},$$

∇M being the K-vector of derivatives of M and A being a (K - 1) x K matrix of the form

$$A = \left[\underline{i} \;\vdots\; -I_{K-1} \right]$$

where $\underline{i}$ is used throughout the paper to denote a vector of ones. For a multivariate normal m.g.f. with $\underline{t} = -r\ \underline{\ell}$, this condition gives us

$$M^{-1} A \cdot M[\underline{\mu} - r\ \Sigma\ \underline{\ell}] = \underline{0}$$

or

$$A \cdot \Sigma\ \underline{\ell} = \frac{1}{r} A\ \underline{\mu}.$$

Note that A Σ is not a square matrix so it cannot be inverted to solve for $\underline{\ell}$. It is only (K - 1) x K, and we need one more restriction on $\underline{\ell}$ so we add that the farm size is L ($\underline{i}_K' \underline{\ell} = L$). Then, the system of K restrictions on $\hat{\underline{\ell}}$ for maximization of expected utility can be solved:

$$\hat{\underline{\ell}} = \begin{bmatrix} A\ \hat{\Sigma} \\ \underline{i}_K' \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{r} A\ \hat{\underline{x}} \\ L \end{bmatrix}$$

where $\hat{\Sigma}$ estimates Σ and $\hat{\underline{x}}$ estimates $\underline{\mu}$.

If the actual parameters were used, $\hat{\underline{\ell}}$ would be the optimal decision (call it $\underline{\ell}^*$) in the sense of maximizing expected utility. Estimation risk (Bawa, Brown, and Klein) exists if estimates are used and $\hat{\underline{\ell}}$ is calculated using what Pope and Ziemer (1984) called the "plug in" method--sample estimates plugged in for population parameters with no adjustments for the estimation risk. The decision will be suboptimal if $\hat{\ell}$ differs from $\underline{\ell}^*$ and $EU(\pi|\hat{\ell}) < EU(\pi|\ell^*)$.

The plug-in method is the standard practice, making $\hat{\underline{\ell}}$ random, as it is a function of past realizations of returns, through $\hat{\mu}$ and $\hat{\Sigma}$. Work on estimation risk in the finance literature has shown that the efficient set of portfolios is unaffected by the problem, though decisions are, and Bayesian decision-making techniques have been explored. When that approach is not followed and the plug-in method is used without prior beliefs about μ or Σ, the sampling behavior of $\hat{\ell}$ is of interest in determing the usefulness of the E-V approach.

We assume that data are available on the K different returns per acre observed over n periods and are collected in a K x n matrix X. Column t of X is a draw, at time t, from $N_K(\mu, \Sigma)$, and we assume timewise independence in these draws. Our estimates $\hat{\underline{x}}$ and $\hat{\Sigma}$ are obtained as

$$\hat{\underline{x}} = \frac{1}{n} X\, i_n$$

and

$$\hat{\Sigma} = (n - 1)^{-1} (X - \hat{\underline{x}}\, \underline{i}_n') (X - \hat{\underline{x}}\, \underline{i}_n')'.$$

If we let $Z = X - \underline{\mu}\, \underline{i}_n'$ be deviations from population means so that the columns of Z, $Z_{\cdot i}$ are independent draws from $N_K(0, \Sigma)$, we can then write

$$\hat{\underline{x}} = \frac{1}{n} Z\underline{i}_n + \underline{\mu}.$$

Also, our estimator for the variance matrix, Σ, can be expressed as

$$\hat{\Sigma} = (n - 1)^{-1} (X - \hat{\underline{x}}\, \underline{i}_n') (X - \hat{\underline{x}}\, \underline{i}_n')' = (n - 1)^{-1} X\, P_n\, X'$$

$$= (n - 1)^{-1} Z\, P_n\, Z' = (n - 1)^{-1} VV'$$

where $P_n = I - i_n(i_n' i_n)^{-1} i_n' = I - \frac{1}{n} i_n i_n'$ is a symmetric, idempotent matrix with rank n - 1. Only deviations from the population mean contribute to the variance estimator. Then, the expression for $\hat{\ell}$ becomes

$$(1) \qquad \hat{\ell} = \begin{bmatrix} (n - 1)^{-1} \ A \ Z \ P_n Z' \\ i_K' \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{r} A(\mu + \frac{1}{n} Zi_n) \\ L \end{bmatrix}.$$

The inverse in (1) exists and is equal to the partitioned expression

$$A'[AVV'A']^{-1} \ (n - 1) \left[I_{K-1} \ \vdots \ -(n - 1)^{-1} \ AVV'e_1 \right] + \left[\underline{\underline{0}} \ \vdots \ e_1 \right]$$

where the 0 matrix is K x (K - 1) and e_1 denotes the first elementary vector of length n.

The solution vector, $\hat{\ell}$, can therefore be written as

$$\hat{\ell} = \left\{ \left[A'(AVV'A')^{-1} \ (n - 1) + \underline{\underline{0}} \quad -A'(AVV'A')^{-1} \ AVV'e_1 + e_1 \right] \right\}_K \ _K\begin{bmatrix} \frac{1}{r} A\mu + \frac{1}{rn} \quad AZi_n \\ L \end{bmatrix}$$

or

$$\hat{\ell} = A'(AVV'A')^{-1} \cdot \frac{n - 1}{r} A(\mu + \frac{1}{n} Zi_n) + Le_1 - LA'(AVV'A')^{-1} AVV'e_1.$$

By a series of substitutions, this express reduces to

$$\underline{\hat{\ell}} = A'(AVV'A')^{-1} \ \frac{n - 1}{r} A(\underline{\mu} + \frac{1}{n} Z\underline{i}_n) + Le_1 - LA'\alpha - LA'(AVV'A)^{-1} AVu.$$

where $\alpha = (A\Sigma A')^{-1} A\Sigma_{\cdot 1}$, $\Sigma_{\cdot 1}$ denotes column one of Σ, and $AVu = AV(V'e_1) - AVV'A'\alpha$.

These substitutions make the proofs more convenient but are otherwise of no importance. The expected value of $\hat{\ell}$ consists of three terms:

$$E(\hat{\ell}) = A' \ E(TT')^{-1} \ A \ \frac{n - 1}{r} \ \underline{\mu} - LA'\alpha + Le_1,$$

and

$$\hat{\ell} - E(\hat{\ell}) = \frac{n - 1}{r} A' \left[(TT')^{-1} - E(TT')^{-1} \right] \ A(\underline{\mu} + \frac{1}{n} Zi_n)$$

$$+ \frac{n - 1}{r} A' \ E(TT')^{-1} \ A \frac{1}{n} Zi_n - LA'(TT')^{-1} \ AVu.$$

The variance matrix for $\hat{\ell}$, $V(\hat{\ell}) = E[\hat{\ell} - E(\hat{\ell})]\,[\hat{\ell} - E(\hat{\ell})]'$, consists of expectations of the products of each term above and its transpose. All cross-products disappear by independence of each term from the others. The random terms are T, Zi_n, and u. Since u can be shown to be independent of both T = AV and Zi_n, the cross products involving u will have a zero expectation. The same is true of those involving the first term since the expectation over T of $[(TT')^{-1} - E(TT')^{-1}]$ is zero.

Our expression for the variance matrix for $\hat{\ell}$ becomes

$$V(\hat{\ell}) = \left(\frac{n-1}{r}\right)^2 A' E_T\left\{\left[(TT')^{-1} - E(TT')^{-1}\right] A(\mu\mu' + \frac{1}{n}\Sigma) A' \left[(TT')^{-1} - E(TT')^{-1}\right]\right\} A$$

$$+ \left(\frac{n-1}{r}\right)^2 A'E(TT')^{-1} A \frac{1}{n} \Sigma A'E(TT')^{-1} A$$

$$+ L^2 A'E_T\left\{(TT')^{-1} T \bar{\Sigma} I_{n-1} T'(TT')^{-1}\right\} A$$

$$= \left(\frac{n-1}{r}\right)^2 A' E_T\left\{\left[(TT')^{-1} - E(TT')^{-1}\right] A(\mu\mu' + \frac{1}{n}\Sigma) A' \left[(TT')^{-1} - E(TT')^{-1}\right]\right\} A$$

$$+ \frac{1}{n}\left(\frac{n-1}{r}\right)^2 A'E(TT')^{-1} A\Sigma A'E(TT')^{-1} A + \bar{\Sigma} L^2 A'E(TT')^{-1} A.$$

where

$$\bar{\Sigma} = \Sigma_{11} - \Sigma_{\cdot 1}' A'(A\Sigma A')^{-1} A\Sigma_{\cdot 1}.$$

The last term simplifies to

$$\bar{\Sigma} L^2 A' \cdot \frac{(A\Sigma A')^{-1}}{n - K - 1} A = \frac{\bar{\Sigma} L^2}{n - K - 1} A'(A\Sigma A')^{-1} A.$$

The second term is

$$\frac{1}{n}\left(\frac{n-1}{r}\right)^2 A' \frac{(A\Sigma A')^{-1}}{n-K-1} A\Sigma A' \frac{(A\Sigma A')^{-1}}{n-K-1} A$$

$$= \left(\frac{n-1}{r}\right)^2 \frac{1}{n(n-K-1)^2} A'(A\Sigma A')^{-1} A.$$

The first term is a bit more involved. It is necessary to use results of Shaman (1980) and extensive algebraic manipulations to show that the first term is

$$\left(\frac{n-1}{r}\right)^2 \frac{1}{(n-K-1)(n-K)^2} A'(A\Sigma A')^{-1} A\mu\mu' A'(A\Sigma A')^{-1} A$$

$$+ \left(\frac{n-1}{r}\right)^2 \frac{1}{n(n-K-1)(n-K)^2} A'(A\Sigma A')^{-1} A.$$

If we collect all three parts of $V(\hat{\ell})$, we have

$$V(\hat{\ell}) = \left(\frac{n-1}{r}\right)^2 \frac{1}{(n-K-1)(n-K)^2} A'(A\Sigma A')^{-1} A\mu\mu' A(A\Sigma A')^{-1} A$$

$$+ \left(\frac{n-1}{r}\right)^2 \frac{1}{n(n-K-1)} \left[\frac{1}{(n-K)^2} + \frac{1}{(n-K-1)}\right] A'(A\Sigma A')^{-1} A$$

$$+ \frac{\overline{\Sigma L}^2}{(n-K-1)} A'(A\Sigma A')^{-1} A.$$

An expression for the bias in $\hat{\ell}$ can be obtained by subtraction:

$$\text{Bias}\ (\hat{\ell}) = E(\hat{\ell}) - \ell^* = \left[\left(\frac{n-1}{n-K-1}\right) - 1\right] A'(A\Sigma A')^{-1} A \cdot \frac{1}{r} \underline{\mu}.$$

An unbiased estimator, call it $\tilde{\ell}$, can be then obtained by rescaling $\hat{\Sigma}$ by the term $(n - 1)/(n - K - 1)$. It is easy to show that

$$V(\tilde{\ell}) = \left(\frac{n - K - 1}{r}\right)^2 \frac{1}{(n - K - 1)\,(n - K)^2} A'(A\Sigma A')^{-1} A\mu\mu' A'(A\Sigma A')^{-1} A$$

$$+ \left(\frac{n - K - 1}{r}\right)^2 \frac{1}{n\,(n - K - 1)} \left[\frac{1}{(n - K)^2} + \frac{1}{(n - K - 1)}\right] A'(A\Sigma A')^{-1} A$$

$$+ \frac{\overline{\Sigma L}^2}{(n - K - 1)} A'(A\Sigma A')^{-1} A.$$

Clearly, this implies a smaller variance matrix for $\tilde{\ell}$ in the sense that $V(\hat{\ell}) - V(\tilde{\ell})$ is positive semidefinite.

To summarize the results so far, we have considered the effect on mean-variance decisions of estimation risk. With returns following a multivariate normal $N_K(\underline{\mu}, \Sigma)$ and μ and Σ unknown, the decision vector $\hat{\ell}$ obtained using sample estimates is biased as an estimator of the unknown optimum ℓ^*. It also has greater variation than the unbiased vector we derived, $\tilde{\ell}$, making the latter an improved rule for mean-variance decisions. Of necessity, both $EU(\pi|\hat{\ell})$ and $EU(\pi|\tilde{\ell})$ are less than $EU(\pi|\ell^*)$ so estimation risk must reduce average welfare.

3. Examples

The importance of these results can be seen from plugging in the values from representative applications. A range of coefficients of absolute risk aversion (10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}) similar to those in Collender and Chalfant (1986b) was used with sample sizes of 5, 10, and 30 to see how estimation risk might affect the reliability of mean-variance allocations.

First, we examined the problem of allocating \$10,000 to securities of Chrysler, New York Shipping, and Bulova--the problem studied by Frankfurter et al. (1971). The means and variances in rates of return for those securities were taken from their article. They used

$$\underline{\mu} = (16.64,\ 6.64,\ 21.35)$$

and

$$\Sigma = \begin{bmatrix} 2102 & -115 & 1115 \\ -115 & 1664 & -37 \\ 1115 & -37 & 2223 \end{bmatrix}.$$

Table 1 shows the optimal allocation, $\hat{\underline{\ell}}$, and the variance of $\hat{\underline{\ell}}$ for selected combinations of risk aversion and sample sizes.

A second and simpler example is the allocation of land (640 acres) to three crops with independent and identically distributed yields. We took net profits to be equal to \$1.00 and nonstochastic for each and used a mean of 1,714 pounds and standard deviation of 600 pounds similar to Day's calculations for cotton yields at 45 pounds of nitrogen per acre. Of course, it is optimal in this case to plant equal areas to each crop. Mean return is unaffected but variance is lowest with that decision. Table 2 shows how sample size for estimating $\underline{\mu}$ and $\underline{\Sigma}$, under various levels of risk aversion, affects the reliability of using mean-variance for allocating land.

The results demonstrate that estimation risk, operating through uncertainty about $\underline{\mu}$ and $\underline{\Sigma}$, can be devastating for our estimates of the optimal decision. Quite often, an interval of two (or less) standard deviations around the optimum includes the corner solution.

Table 1
Calculations Using Data from Frankfurter et. al.

n=5 r=1e-6

910.525 -9385.972 18475.44

2.233106e9	-5.734925e8	-1.659514e
-5.734923e8	1.232097e9	-6.586048e
-1.659514e9	-6.586046e8	2.318218e

n=5 r=0.00001

2653.976 3530.018 3816.006

64883520	-16800068	-48083448
-16800062	34827248	-18027185
-48083452	-18027182	66110632

n=5 r=0.0001

2828.32 4821.617 2350.062

43201288	-11233145	-31968144
-11233141	22854550	-11621411
-31968146	-11621408	43589552

n=5 r=0.001

2845.754 4950.777 2203.468

42984468	-11177475	-31806992
-11177472	22734824	-11557353
-31806994	-11557349	43364344

n=10 r=1e-6

2121.254 -416.5346 8295.281

145448592	-37326268	-108122320
-37326256	80440952	-43114704
-108122328	-43114692	151237008

n=10 r=0.00001

2775.048 4426.96 2797.939

8546563	-2217453	-6329108
-2217453	4555456	-2338003
-6329109	-2338003	8667112

n=10 r=0.0001

2840.428 4911.31 2248.26

7177542	-1866365	-5311177
-1866365	3796601	-1930237
-5311177	-1930236	7241413

n=10 r=0.001

2846.965 4959.747 2193.287

7163852	-1862854	-5300997
-1862854	3789013	-1926159
-5300998	-1926158	7227156

n=30 r=1e-6

2307.52 963.379 6729.100

25169106	-6465057	-18704050
-6465055	13877749	-7412694
-18704050	-7412692	26116744

n=30 r=0.00001

2793.675 4564.953 2641.371

1888324	-490233.0	-1398091
-490233	1004403.	-514171
-1398091	-514170.6	1912262

n=30 r=0.0001

2842.29 4925.11 2232.599

1655516.	-430484.8	-1225031.
-430484.6	875670.2	-445185.5
-1225031.	-445185.4	1670216.

n=30 r=0.001

2847.152 4961.126 2191.72

1653188	-429887.3	-1223300.
-429887	874382.8	-444495.7
-1223301	-444495.5	1667796.

Table 2
Calculations Using Day's Data

n=5 r=1e-6

213.3340 213.3340 213.3330

7498431	-3749216	-3749215
-3749216	7498431	-3749215
-3749215	-3749215	7498430

n=5 r=0.00001

213.3334 213.3334 213.3333

165096.3	-82548.1	-82548.1
-82548.2	165096.3	-82548.1
-82548.1	-82548.1	165096.2

n=5 r=0.0001

213.3332 213.3334 213.3333

91762.94	-45881.47	-45881.46
-45881.48	91762.93	-45881.46
-45881.46	-45881.46	91762.92

n=5 r=0.001

213.3333 213.3334 213.3333

91029.6	-45514.8	-45514.8
-45514.8	91029.6	-45514.8
-45514.8	-45514.8	91029.6

n=10 r=1e-6

213.3335 213.3335 213.3335

482857.5	-241428.8	-241428.7
-241428.8	482857.5	-241428.7
-241428.7	-241428.7	482857.4

n=10 r=0.00001

213.3332 213.3334 213.3333

19847.24	-9923.62	-9923.62
-9923.62	19847.24	-9923.62
-9923.62	-9923.62	19847.24

n=10 r=0.0001

213.3332 213.3334 213.3333

15217.13	-7608.568	-7608.566
-7608.568	15217.13	-7608.565
-7608.566	-7608.566	15217.13

n=10 r=0.001

213.3332 213.3334 213.3333

15170.83	-7585.417	-7585.416
-7585.418	15170.83	-7585.415
-7585.416	-7585.416	15170.83

n=30 r=1e-6

213.3334 213.3332 213.3332

83035.0	-41517.5	-41517.5
-41517.5	83035.0	-41517.5
-41517.5	-41517.5	83035.0

n=30 r=0.00001

213.3332 213.3334 213.3333

4296.196	-2148.098	-2148.098
-2148.098	4296.196	-2148.097
-2148.098	-2148.098	4296.195

n=30 r=0.0001

213.3333 213.3334 213.3333

3508.807	-1754.404	-1754.403
-1754.404	3508.807	-1754.403
-1754.403	-1754.403	3508.807

n=30 r=0.001

213.3333 213.3334 213.3333

3500.934	-1750.467	-1750.467
-1750.467	3500.933	-1750.466
-1750.466	-1750.466	3500.933

4. Conclusions

The use of mean-variance and similar analyses is so widespread in agricultural economics as to be considered one of the main quantitative techniques. However, the problem of estimation risk is infrequently addressed, and it is common to see both positive and normative applications with very sparse data sets. The results in this paper are a first step toward assigning confidence levels or constructing interval estimates for the optimal decision, and they show that extreme care should be taken in some applications.

The results establish, under the normality assumption, that the usual "plug in" type decision vector is biased and, unless large amounts of data are available, its elements have large variances. An unbiased decision vector can be constructed, but it still suffers from the problem of sparse data.

Our applications were chosen to provide some representative cases, possibly erring on the side of large variances. Still, the fact that the optimal decision is estimated with so little precision is startling. While the magnitudes are specific to the application, the need for care in obtaining estimates of optimal decisions is likely to be less so.

REFERENCES

Bawa, Vijay S., Stephen J. Brown, and Roger W. Klein: Estimation Risk and Optimal Portfolio Choice. Amsterdam: North-Holland Publishing Co., 1979.

Buccola, Steven T.: "Testing for Nonnormality in Farm Net Returns," American Journal of Agricultural Economics, 68(1986), 334-343.

Collender, Robert N., and James A. Chalfant: "An Alternative Approach to Decisions Under Uncertainty Using the Empirical Moment-Generating Function," American Journal of Agricultural Economics, 68(1986a), 727-731.

________. "Sparse Data and Risk-Efficient Choice Under Uncertainty." University of California, Department of Agricultural and Resource Economics, Working Paper No. 403, Berkeley, April, 1986b.

Collender, Robert N., and David Zilberman: "Land Allocation Under Uncertainty for Alternative Specifications of Return Distributions," American Journal of Agricultural Economics, 67(1985), 779-786.

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

Frankfurter, G. M., H. E. Phillips, and J. P. Seagle: "Portfolio Selection: The Effects of Uncertain Means, Variances, and Covariances," Journal of Financial and Quantitative Analysis, 6(1971), 1251-1262.

Freund, R. J.: "The Introduction of Risk Into a Programming Model," Econometrica, 14(1956), 253-263.

Pope, Rulon D., and Rod F. Ziemer: "Stochastic Efficiency, Normality, and Sampling Errors in Agricultural Risk Analysis," American Journal of Agricultural Economics, 66(1984), 31-40.

Shaman, Paul: "The Inverted Complex Wishart Distribution and its Application to Spectral Estimation," Journal of Multivariate Analysis, 10(1980), 51-59.