The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments

Author(s): Paul A. Samuelson

Source: *The Review of Economic Studies*, Oct., 1970, Vol. 37, No. 4 (Oct., 1970), pp. 537-542

Published by: Oxford University Press

Stable URL: https://www.jstor.org/stable/2296483

The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments[1]

I

James Tobin [7, 8], Harry Markowitz [3, 4], and many other writers have made valuable contributions to the problem of optimal risk decisions by emphasizing analyses of means and variances. These writers have realized that the results can be only approximate, but have also realized that approximate and computable results are better than none.

Recently, Karl Borch [1] and Martin Feldstein [2] have re-emphasized the lack of generality of mean-variance analysis and evoked a reply from Tobin [9]. None of the writers in this symposium refer to a paper of mine (Samuelson, [6]) which suggested that most of the interesting propositions of risk theory can be proved for the general case with no approximations being involved. This same paper pointed out all the realms of applicability of mean-variance analysis and also its realms of non-applicability.

There is no need here to redescribe these arguments. But I think it important to re-emphasize an aspect of the mean-variance model that seems not to have received sufficient attention in the recent controversy, namely the usefulness of mean and variance in situations involving less and less risk—what I call " compact " probabilities. The present paper states and proves the two general theorems involved. In a sense, therefore, it provides a defence of mean-variance analysis—in my judgement the most weighty defence yet given. (In economics, the relevant probability distributions are *not* nearly Gaussian, and quadratic utility in the large leads to well-known absurdities). But since I improve on mean-variance analysis and show its exact limitations—along with those for any r-moment model—the paper can also be regarded as a critique of the mean-variance approach. In any case, the theorems here provide valuable insight into the properties of the general case. I should add that their general content has long been sensed as true by most experts in this field, even though I am unable to cite publications that quite cover this ground.

II

The Tobin-Markowitz analysis of risk-taking in terms of mean and variance alone is rigorously applicable only in the restrictive cases where the statistical distributions are normally Gaussian or where the utility-function to be maximized is quadratic. In only a limited number of cases will the central limit law be applicable so that an approximation to normality of distribution becomes tenable; and it is well known that quadratic utility has anomalous properties in the large—such as reduced absolute and relative risk-taking as wealth increases, to say nothing of ultimate satiation.

[1] Aid from the National Science Foundation is gratefully acknowledged, and from my M.I.T. students and co-researchers: Robert C. Merton, from whose conversations I have again benefited, and Dr. Stanley Fischer (now of the University of Chicago) whose 1969 M.I.T. doctoral dissertation, *Essays on Assets and Contingent Commodities* contains independently-derived results on compact distributions.

However, a defence for mean-variance analysis can be given (Samuelson, [6; p. 8]) along other lines—namely, when riskiness is " limited ", the quadratic solution " approximates " the true general solution. The present note states the underlying approximation theorems, and shows the exact limit of their accuracy.

III

Let the return from investing \$1 in each of " securities " 1, 2, ..., n be respectively the random variable $(X_1, \ldots, X_n)$, subject to the joint probability distribution

$$\text{prob}\,\{X_1 \leqq x_1 \text{ and } X_2 \leqq x_2 \text{ and } \ldots X_n \leqq x_n\} = F(x_1, \ldots, x_n). \qquad \ldots(1)$$

Let initial wealth W be set (by dimensional-unit choice) at unity and $(w_1, w_2, \ldots, w_n)$ be the fractions of wealth invested in each security, where $\sum_1^n w_j = 1$. Then the investment outcome is the random variable $\sum_1^n w_j X_j$, and if $U[W]$ is the decision maker's concave utility, he is postulated as acting to choose $[w_j]$ to maximize expected utility, namely

$$\max_{\{w_i\}} \bar{U}[w_1, \ldots, w_n] = \int_0^\infty \cdots \int_0^\infty U\left[\sum_1^n w_j X_j\right] dF(X_1, \ldots, X_n). \qquad \ldots(2)$$

In general, the solution to this problem will *not* be the same as the solution to the quadratic case

$$\max_{\{w_i\}} \int_0^\infty \cdots \int_0^\infty \left\{U[1] + U'[1]\left[\sum_1^n w_j X_j - 1\right] + \tfrac{1}{2}U''[1]\left[\sum_1^n w_j X_j - 1\right]^2\right\} dF(X_1, \ldots, X_n). \qquad \ldots(3)$$

But now let us suppose that $F(\cdot, \ldots, \cdot)$ belongs to a family of " compact " or " small-risk " distributions, defined so that as some specified parameter goes to zero, all our distributions converge to a sure outcome. An appropriate family would be

$$F(x_1, \ldots, x_n) = P\left(\frac{x_1 - \mu - \sigma^2 a_1}{\sigma\sigma_1}, \frac{x_2 - \mu - \sigma^2 a_2}{\sigma\sigma_2}, \ldots, \frac{x_n - \mu - \sigma^2 a_n}{\sigma\sigma_n}\right). \qquad \ldots(4)$$

Here, the variables have been defined so that, as the parameter $\sigma \to 0$, it becomes ever more certain that the outcome for $(X_1, \ldots, X_n) = (\mu, \mu, \ldots, \mu)$, where μ might be 1 + the " safe " rate of interest. Fig. 1 illustrates for the one-dimensional case, $n = 1$, the meaning of such a " compact family ". As $\sigma \to 0$ all the probability piles up at μ. Note that no normality of distributions is involved—originally or (in any non-trivial sense) asymptotically.

By convention, $P(y_1, \ldots, y_n)$ is defined to have the properties

$$E[Y_i] = \int_0^\infty \cdots \int_0^\infty Y_i dP(Y_1, \ldots, Y_n) = 0,$$

$$E[Y_i^2] = \int_0^\infty \cdots \int_0^\infty Y_i^2 dP(Y_1, \ldots, Y_n) = 1, \qquad \ldots(5)$$

$$E[Y_i Y_j] = \int_0^\infty \cdots \int_0^\infty Y_i Y_j dP(Y_1, \ldots, Y_n) = r_{ij},$$

where $[r_{ij}]$ is a symmetric positive definite correlation matrix. Similarly higher moments $E[Y_i^{k_i} Y_j^{k_j} \ldots]$, for k's integers, can be defined. It follows then that

$$E[X_i] = \mu + \sigma^2 a_i, \quad E[X_i - E[X_i]]^2 = \sigma_i^2 \sigma^2, \text{ etc.} \qquad \ldots(6)$$

An explanation may be needed to motivate our putting the σ parameter in the numerator of P's arguments. If $[\mu + \sigma^2 a_i]$ had been replaced by the simple constants $[\mu_i]$, and if they were not all equal, then the securities with the largest μ_i would dominate the rest as $\sigma \to 0$. All other w's would either go to zero, or if borrowing and selling short were freely permitted—so that the w's need not be non-negative—infinitely profitable arbitrage would be possible. This bizarre, infinite case is of trivial interest. Hence, $\mu_i(\sigma)$ must as $\sigma \to 0$ approach a common limit.

Anyone familiar with Wiener's Brownian motion will identify σ with the square root of time, $\sqrt{t}$: in the numerator σ^2 appears because means grow linearly with time in Brownian motion; in the denominator σ appears because standard deviations grow like $\sqrt{t}$, with variance growing linearly. Dimensionally $a_i\sigma^2$ has the same dimension, namely dollars, as does X_i and $\sigma\sigma_i$; in Brownian terms a_i would be " dollars/time ". The inequality of instantaneous mean gains in Brownian motion is indicated by inequalities among the a_i, not among the μ_i. Although I have couched this heuristic explanation in terms of Brownian motion in time, the concept of a compact family is an independent and completely general one. Fortunately, it gains in importance because it does throw light on the reasons why enormous " quadratic " simplicities occur in continuous-time models, as the cited Fischer thesis make clear and as is evident in Merton [5].

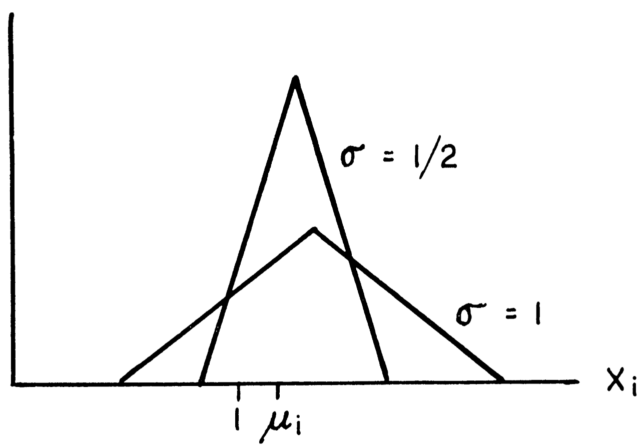

FIGURE 1
Example of family of compact probability densities.

Actually our family can be defined more generally than in terms of $[x_i - \mu - \sigma^2 a_i]/(\sigma\sigma_i)$. All that is required—and it *is* required if $[w_i(\sigma)]$, the optimal portfolio proportion as a function of the parameter σ, is to approach a unique and smooth limit $[w_i(0)]$—is that for the standardized variables $Z_i = X_i - \mu$

$$\lim_{\sigma\to 0} \frac{E[Z_i]}{E[Z_i^2]} = \frac{A}{B}, \quad \lim_{\sigma\to 0} \frac{E[Z_i^r]}{E[Z_i^2]} = \sigma^{r-2}C_r \quad (r = 3, 4, \ldots). \qquad \ldots(7)$$

The P family defined above does have this property, as can now be verified for the one-dimensional case, $X_1 = X$,

$$\begin{aligned} E[X] &= \int_0^\infty X dP\left(\frac{X-\mu-\sigma^2 a}{\sigma\sigma_1}\right) = \mu + \sigma^2 a, \\ E[(X-\mu-\sigma^2 a)^2] &= \sigma^2\sigma_1^2, \\ E[Z] &= \int_0^\infty (X-\mu) dP\left(\frac{X-\mu-\sigma^2 a}{\sigma\sigma_1}\right) = \sigma^2 a, \\ E[Z^2] &= E[(X-\mu-\sigma^2 a)^2] + (\sigma^2 a)^2, \qquad \ldots(8) \\ &= \sigma_1^2\sigma^2 + \sigma^4 a^2 = \sigma^2(\sigma_1^2 + \sigma^2 a^2), \\ E[Z^3] &= E[\{(X-\mu-\sigma^2 a) + \sigma^2 a\}^3], \\ &= \sigma^3\sigma_1^3\mu_3 + \sigma\psi_3, \end{aligned}$$

where μ_3 is the third moment around the mean of $P(y)$ and $\sigma\psi_3$ involves powers of σ higher than 3. Similarly, it can be shown that

$$E[Z^4] = \sigma^4\sigma_1^4\mu_4+\sigma\psi_4, \ldots, E[Z^r] = \sigma^r\sigma_1^r\mu_r+\sigma\psi_r, \quad r>2. \qquad \ldots(9)$$

An example of an admissible compact family that cannot quite be written in the $P(\cdot)$ form is

$$\text{prob}\,\{X_1 = 1+\sigma\} = \text{prob}\,\{X_1 = (1+\sigma)^{-1}\} = \tfrac{1}{2}, \qquad \ldots(10)$$

Here,
$$E[X-1] = \frac{\sigma^2}{2} + \text{higher powers of } \sigma,$$

$$E[(X-1)^2] = 2\sigma^2 + \text{higher powers of } \sigma.$$

But this does satisfy our needed asymptotic conditions as defined in (7).

We can now state our fundamental approximation theorem

Theorem 1. *The solution to the general problem,* $[w_i(\sigma)]$, *does, as* $\sigma\to 0$ *have the property that* $[w_i(0)]$ *is the exact solution to the quadratic problem*

$$\max_{\{w_i^*\}} \int_0^\infty \ldots \int_0^\infty \left\{U[\mu]+U'[\mu]\left[\sum_1^n w_j^* X_j-\mu\right] + \frac{U''(\mu)}{2} \cdot \left[\sum_1^n w_j^* X_j-\mu\right]^2\right\} dP\left(\frac{X_1-\mu-\sigma^2 a_1}{\sigma\sigma_1}, \ldots, \frac{X_n-\mu-\sigma^2 a_n}{\sigma\sigma_n}\right),$$

i.e.,
$$\lim_{\sigma\to 0} w_i^*(\sigma) = w_i^*(0) = w_i(0) = \lim_{\sigma\to 0} w_i(\sigma).$$

But it is definitely *not* the case that the higher approximation implied by $w_i'(0) = w_i^{*\prime}(0)$ will hold. Theorem 2 will show that one must use cubic-utility 3-moment theory to achieve this higher degree of approximation, and that in general (r-moment, rth degree) utility theory must be used in order to get agreement between $[w_i(0), w_i'(0), \ldots, w_i^{[r-2]}(0)]$ and $[w_i^*(0), w_i^{*\prime}(0), \ldots, w_i^{*[r-2]}(0)]$.

Theorem 2. *The solution to the general problem above is related asymptotically to that of the r-moment problem*

$$\max_{\{w_i^{**}\}} \int_0^\infty \ldots \int_0^\infty \left\{\sum_0^r U^{[j]}(\mu)\frac{\left[\sum_1^n w_i X_i-\mu\right]^j}{j!}\right\} dP\left[\frac{X_1-\mu-\sigma^2 a_1}{\sigma\sigma_1}, \ldots, \frac{X_n-\mu-\sigma^2 a_n}{\sigma\sigma_n}\right]$$

by the high-contact equivalences

$$w_i(0) = w_i^{**}(0),\ w_i'(0) = w_i^{**\prime}(0), \ldots, w_i^{[r-2]}(0) = w_i^{**[r-2]}(0).$$

To prove the theorems most rapidly, note that if U possesses an exact Taylor's expansion, $[w_i(\sigma)] \equiv [w_i^{**}(\sigma)]$ is for $r = \infty$ trivially identical. Hence, $w_i(\sigma)$ is formally defined from the power series

$$\bar{U}(\sigma) = \max_{\{w_i\}} \sum_0^\infty U^{[j]}(\mu)\,\frac{E\left[\left(\sum_1^n w_i Z_i\right)^j\right]}{j!}. \qquad \ldots(11)$$

The first-order conditions for a regular optimum are

$$\frac{\partial\bar{U}}{\partial w_1} = \frac{\partial\bar{U}}{\partial w_2} = \ldots = \frac{\partial\bar{U}}{\partial w_n}, \qquad \ldots(12)$$

where each of these is a power series in all the w's, with coefficients that depend on the moments $E[Z_1^{k_1}Z_2^{k_2}\ldots]$. However, the moments in the infinite series that are involved if we truncate the series at $r-1<\infty$ will be seen to be only the " lower " moments, in which $k_1+k_2+\ldots \leqq r$. Also straightforward but tedious formal algebra shows that in the infinite power series for $w_i(\sigma)$, the coefficients of the terms of order σ^{r-2} or less do depend only on the coefficients of the $\bar{U}(\sigma)$ power series only up to the rth degree terms and rth moments. Hence, the indicated agreement of derivatives in Theorem 2 does hold. Theorem 1 is of course only a special, but important, sub-case of Theorem 2.

There are a number of obvious corollaries of the theorems. Thus, the well-known Tobin separation theorem, which is valid for quadratic utilities, will necessarily be asymptotically valid in the sense of becoming true as $\sigma \to 0$. I.e. $w_i(0)/w_j(0)$ for assets held along with cash will be independent of the form of $U[\cdot]$.

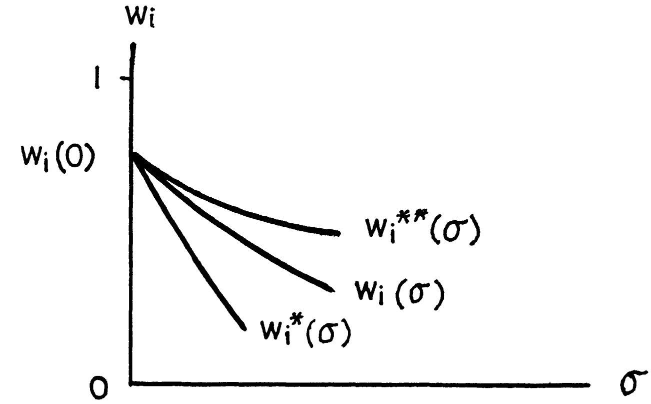

FIGURE 2

Although the quadratic mean-variance solution, $w_i{}^*(0)$, approaches the same intercept as the true solution, $w_i(0)$ the 3-moment cubic solution, $w_i{}^{**}(0)$, has higher contact with the true solution.

The exposition here has been heuristic, but can, subject to some hypotheses on the utility U and probability P functions, be made quite rigorous. The kind of formal algebra involved can be illustrated by the important case of cash versus one risky-asset (say a stock). Call cash X_0 and call the stock $X_1 = X$. Then the outcome for initial unit wealth is $(1-w)+wX = w(X-1)+1 = wZ+1$, where $w = w_1 = 1-w_0$, and our general problem becomes

$$\bar{U}(w) = \max_{\{w\}} \int_0^\infty U[wZ+1]dP\left(\frac{Z-\sigma^2 a}{\sigma}\right),$$

$$\bar{U}'(w) = 0 = \int_0^\infty ZU'[wZ+1]dP \qquad \ldots(13)$$

$$= U'[1]\int_0^\infty ZdP + U''[1]\frac{w}{1}\int_0^\infty Z^2dP + U'''[1]\frac{w^2}{2!}\int_0^\infty Z^3dP + w^3R_3,$$

where dP is short for $dP[\sigma^{-1}(Z-\sigma^2 a)]$ and w^3R_3 involves higher powers of w than w^2.

If we truncate the series before the $U'''[1]$ term, we have the Tobin-Markowitz mean-variance approximation, with solution

$$w^*(0) = \lim_{\sigma\to 0} - \frac{a\sigma^2 U'[1]}{\sigma^2(\sigma_1^2 + a^2\sigma^2)U''[1]} = \frac{a}{\sigma_1^2}\left(\frac{U'[1]}{-U''[1]}\right), \qquad \ldots(14)$$

a not-surprising result. If from the above expression for $w^*(\sigma)$ we calculate $w^{*\prime}(0)$, we will not get the same result as if we had carried one more term in our truncated infinite expansion. Call this last value $w^{**}(\sigma)$ and define it as the root of the equation

$$0 = U'[1](a\sigma^2)+U''[1]w\sigma^2(\sigma_1^2+a^2\sigma^2)+U'''(1)\frac{w^2}{2}\sigma^3(\sigma_1^3\mu_3+\sigma\psi_3), \qquad ...(15)$$

where $\sigma^3\sigma\psi_3$ involves powers of σ beyond σ^3 and where the third moment of $P(y)$ has been earlier defined in (8). Then ignoring $\sigma^4\psi_3$, as we may, we find that $w^{**\prime}(0)$ most definitely *is* affected by the value of $U'''[1]\sigma_1^3\mu_3$. Hence, the quadratic approximation is not locally of " high " contact. Fig. 2 illustrates the phenomenon. Notice how higher-than second moments do improve the solution. But it also needs emphasizing that near to $\sigma = 0$, when " risk is quite limited ", the mean-variance result is a very good approximation. When the heat of the controversy dissipates, that I think will be generally agreed on.[1]

Massachusetts Institute of Technology PAUL A. SAMUELSON

First version received January 1970; *final version received March* 1970

REFERENCES

[1] Borch, K. " A Note on Uncertainty and Indifference Curves ", *Review of Economic Studies*, **36** (1969).

[2] Feldstein, M. S. " Mean-Variance Analysis in the Theory of Liquidity Preference and Portfolio Selection ", *Review of Economic Studies*, **36** (1969).

[3] Markowitz, H. " Portfolio Selection ", *The Journal of Finance*, **7** (1952).

[4] Markowitz, H. *Portfolio Selection: Efficient Diversification of Investments* (New York, John Wiley & Sons, 1959).

[5] Merton, R. C. " Lifetime Portfolio Selection under Uncertainty: the Continuous-Time Case ", *Review of Economics and Statistics*, **51** (1969).

[6] Samuelson, P. A. " General Proof that Diversification Pays ", *Journal of Financial and Quantitative Analysis*, **2** (1967).

[7] Tobin, J. " Liquidity Preference as Behaviour Towards Risk ", *Review of Economic Studies*, **25** (1958).

[8] Tobin, J. " The Theory of Portfolio Selection ", in F. H. Hahn and F. P. R. Brechling (eds.), *The Theory of Interest Rates* (London, Macmillan, 1965).

[9] Tobin, J. " Comment on Borch and Feldstein ", *Review of Economic Studies*, **36** (1969).

1 See Markowitz [4] p. 121 for an argument that can be related closely to that here.