PARAMETRIC EVALUATION AND MEAN-STANDARD DEVIATION ANALYSIS IN STOCHASTIC PROGRAMMING MODELS*, **

By HIROYUKI ITAMI***

I. *Introduction*

In every decision making situation, a decision has various implications in many ways. One of the virtues of management science efforts is that they make these implications clear and explicit to some extent by tracing the effects of a decision using formal models and logical reasoning. In this paper we will try to investigate three kinds of implications arising from the use of some stochastic programming models in the decision making situation under risk and from the decision based on these models. Almost all through this paper we shall be mainly interested in the situation where a decision maker is confronted with a problem involving a stochastic objective function, rather than a problem with stochastic constraints. Several models of this type are presented and analyzed extensively.

Our main interest will be in investigating three kinds of implications of stochastic programming models and their uses.

First, we shall see the implications that particular values for key parameters in these models, which a decision maker has to predetermine by some judgment, have in terms of the effects of the change of these values on the objective function in each model. Since the determination of values of parameters in these decision models are often considered to be in the realm of 'decision maker's subjective judgment', information from this parametric evaluation will be very helpful for a decision maker. Although some efforts have been previously undertaken in this area, e.g., Agnew, et al. [2], the author believes that this paper is the first extensive study of parametric evaluation in stochastic programming models containing a discussion of explicit computational methods.

The second kind of implications of stochastic programming models we shall see in this paper is the relationship between several stochastic programming models and mean-standard deviation analysis which has been developed in portfolio selection. It is shown that there is a very close relationship between these two different approaches.

Thirdly, we will see the utility implications of stochastic programming models presented in this paper. It is shown that most of the common stochastic programming models

* This report was prepared as part of the activities of the Management Sciences Research Group at Carnegie-Mellon University's Graduate School of Industrial Administration.

** The author would like to acknowledge very helpful comments from Professors K.J. Cohen and Y. Ijiri, Graduate School of Industrial Administration, Carnegie-Mellon University, on the earlier draft of this paper.

*** Lecturer (*Kôshi*) of Management Science.

presented in this paper are not compatible with the risk-averter's utility function and its expected value maximization.

In the final part of this paper, we will attempt to impute the value of key parameter(s) in each stochastic programming model and some other utility maximization models from the actual decision made by a decision maker, assuming that the decision is based on one of the models presented in this paper. This is to know the implications of a decision by referring to some particular decision model with a particular parameter value implied by a decision maker's decision. In a way, this imputation will lead to finding a decision maker's attitude toward risk by imputed values for parameters in several decision models.

Since we will present in Chapter II all the stochastic programming models which we will investigate, some readers might think it worthwhile to read this paper model-wise, that is, picking up sections in each chapter pertinent to a particular model, rather than following the order of chapters.

II. *Several Stochastic Programming Formulations and Parametric Evaluations*

Stochastic programming is a generalization of linear programming to the case where some or all of the parameters of the model are random, not deterministic. We consider the follwoing LP model as the starting point to the generalizations to the stochastic programming models.

$$\begin{aligned} \max\quad & cx \\ \text{s.t.}\quad & Ax \leq b \\ & x \geq 0, \end{aligned}$$

where $c \cdots 1 \times n$ vector
$x \cdots n \times 1$ vector
$b \cdots m \times 1$ vector
$A \cdots m \times n$ matrix.

(We do not indicate explicitly a transpose of a matrix when it is clear from the context.)

There are various ways of handling the randomness in A, b, c. We consider those lines of attack which A. Charnes and W.W. Cooper first developed and did extensive research upon, that is, the chance-constrained programming (CCP) approach.

In this approach we transform the original problems into so-called 'deterministic equivalents' of stochastic programming models, which are often non-linear programming problems. We will make a brief survey of various deterministic equivalents in the CCP approach and then investigate the information that dual evaluators (Lagrangean multipliers) in these models can give. Dual evaluation in non-linear programming has almost the same properties as in linear programming. Using these dual evaluators in non-linear programming, we can obtain valuable information about the effect of parametric changes in CCP models. This information may give the decision maker very fruitful insight into both the structure of the model and his own attitude towards risk which is supposedly expressed in the particular values he selected for parameters in the model, but which may be only implicit in his mind.

II. 1 General Model

The most general model of CCP models assumes randomness in all parameters, i.e., A, b, c, and has the following formulation.

$$\max \quad f(c, x) \qquad (2.\ 1)$$
$$\text{s.t.} \quad Pr(Ax \leq b) \geq \alpha \qquad (2.\ 2)$$
$$x \geq 0 \qquad (2.\ 3)$$

where α is a $m \times 1$ vector.
Here f is some criterion function to be maximized and (2.2) are chance constraints, which means that the i-th constraint of the basic LP model above has to be satisfied with the probability of α_i. The decision maker has to specify the form of the objective function, f.

The most common approach to deriving the deterministic equivalent for this model is to assume that c and each row of A are independently and normally distributed vectors. The normality assumption will be maintained throughout this paper. Although we can incorporate the randomness of b in the following derivation with a few additional assumptions, we will refrain from incorporating the randomness of b only because of the notational complexity it requires. We will use the following notations in the discussions below.

$\mu \cdots \mathscr{E}(c)$, $1 \times n$ vector $\qquad \mathscr{E} \cdots$expectation operator

$V \cdots \mathscr{E}(c-\mu)(c-\mu)$

$p^i \cdots \mathscr{E}(a^i)$, $a^i \cdots$the i-th row of A

$\sum_i \cdots \mathscr{E}(a^i - p^i)(a^i - p^i)$

$$F(q) \cdots \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{q} exp(-t^2/2)dt$$

We also assume that $\alpha_i > .5$ for all i.

We can rewrite the i-th chance-constraint in the following successive way.

$$Pr(a^i x \leq b_i) \geq \alpha_i$$
$$Pr((a^i x - p^i x)/\sqrt{x \textstyle\sum_i x} \leq (b_i - p^i x)/\sqrt{x \textstyle\sum_i x}) \geq \alpha_i$$
$$F((b_i - p^i x)/\sqrt{x \textstyle\sum_i x}) \geq \alpha_i$$
$$(b_i - p^i x)/\sqrt{x \textstyle\sum_i x} \geq F^{-1}(\alpha_i)$$

Putting $\qquad \beta_i = F^{-1}(\alpha_i)$

$$p^i x + \beta_i \sqrt{x \textstyle\sum_i x} \leq b_i$$

Therefore the deterministic equivalent to this general model is the following non-linear programming problem.

$$\max \quad f(c, x) \qquad (2.\ 1)$$
$$\text{s.t.} \quad p^i x + \beta_i \sqrt{x \textstyle\sum_i x} \leq b_i \qquad i=1, \cdots, m \qquad (2.\ 4)$$
$$x \geq 0 \qquad (2.\ 3)$$

Since $\sqrt{x \sum_i x}$ is a convex function of x (see Kataoka [15]) and $\beta_i > 0$ from the assumption that $\alpha_i > .5$, the left-hand-side of (2.4) is a convex function. This implies that the feasible region of the non-linear programming problem above is a convex set.

If the decision maker specifies some concave objective function, the general model of CCP reduces to the so-called concave programming model (a concave maximizing objective function with a convex constraint set) for which various computational methods like gradient methods are developed under the assumption of the differentiability of all the functions involved. For a survey of these concave programming methods, see Hadley [11] and

Wolfe [23].

In this model, the decision maker has to decide (or specify) two things before actually selecting the optimal decision. They are the form of f and the values of risk levels, α. It is very helpful for a decision maker if he can estimate the implications of various values of α's, especially in terms of the effect of parametric changes of α's on the value of the maximand. It is possible to do this by using the dual evaluators in the non-linear programming problem (2.1), (2.4), (2.3).

In solving this non-linear programming problem, we will get the Lagrangean multipliers, u, which is a $1\times m$ vector, and each component of which corresponds to each constraint of (2.4). In many concave programming methods, we can get u as a byproduct of the computation to get the optimal solution x^0. If we do not have this byproduct in a particular algorithm, but if we get the optimal solution x^0, we can get the value of u by solving the following LP problem, which is a dual problem to (2.1)–(2.3) with x^0 inserted.[1] The duality theorem in non-linear programming assures us that this method gives the right values of u (See Wolfe [24]).

$$\min \quad ub \tag{2. 5}$$

$$\text{s.t.} \quad \partial f(c, x)/\partial x_j \geq \sum_{i=1}^{m} u_i(p_j^i + \beta_i \textstyle\sum_{ij} x/\sqrt{x\sum_i x}) \qquad j=1, \cdots, n \tag{2. 6}$$

$$u \geq 0 \tag{2. 7}$$

where $\sum_{ij}$ is the j-th column of $\sum_i$.

If we consider the optimal value of f as a function of b in this model as b's change, u is equal to $\partial f/\partial b$ under the following assumption.

i) Tight constraints at x^0 remain tight in the new optimal solution after infinitesimal change of b's.

ii) If $x_j^0=0$, $x_j=0$ in the new optimal solution after infinitesimal change of b's.

For a detailed explanation on these conditions in the general non-linear programming models, see Abadie [1], especially his preface.

Under the similar assumptions as above, dual evaluations as follows are possible.

$$\partial f/\partial \beta_i = -\sqrt{x\textstyle\sum_i x}\, u_i$$

$$\partial f/\partial \alpha_i = -\sqrt{x\textstyle\sum_i x}\, u_i/\varphi(\beta_i) \tag{2. 6$'$}$$

where $$\varphi(t)=\frac{1}{\sqrt{2\pi}}\exp(-t^2/2).$$

For the general discussion of the derivation of parametric evaluations using Lagrangean multipliers, see Naslund and Whinston [18]. There is also some discussion on the parametric evaluation in the CCP model in Agnew at el. [2].

(2.6)$'$ is the evaluation of risk levels in terms of the effect to the optimal objective function value. The decision maker may adjust the risk levels after getting this information, so that he is satisfied with the trade-off between f and α which is implied by his selection of a particular value for α_i.

Now we shall turn our discussion into more specific models which have been developed so far in the context of the CCP approach. In these discussions, we are going to assume that randomness occurs only in c in order to focus attention on decision making situations with stochastic benefit, in this case, the objective function value, cx. This assumption will be maintained throughout the rest of this paper.

[1] We assume that f is differentiable.

II. 2 E-model

If we consider the maximization of the expected value of the objective function, cx, as an appropriate criterion, we have the following simple LP problem:

$$\begin{aligned} \max \quad & \mathscr{E}(cx)=\mu x \\ \text{s.t.} \quad & Ax\leq b \\ & x\geq 0 \end{aligned}$$

We call this model E-model, and ordinary dual evaluations in LP are available regarding the effect of changes of b's on $\mathscr{E}(cx)$.

II. 3 V-model

When the legitimate criterion is to minimize the variance of the stochastic return (which we might consider as a measure of risk), the deterministic equivalent of CCP is reduced to the following quadratic programming problem. This model is called V-model.

$$\begin{aligned} \min \quad & \text{Var}\ (cx)=xVx \\ \text{s.t.} \quad & Ax\geq b \\ & x\geq 0 \end{aligned}$$

Since V is a positive definite martix (or a positive semi-definite martix in some special cases), various quadratic programming methods are available for solving this problem.

In quadratic programming, dual evaluation is also possible. The dual evaluator will tell us the rate of change of the risk (variance of the objective, cx) according to the infinitesimal increase of b's.

II. 4 α-fractile model (α-model)

In this model, the decision maker wants to maximize the lower bound of cx above which he can be sure that the objective function value will fall with probability α. α is predetermined by the decision maker, and we call this lower bound the α-fractile. This model is called the α-fractile model or, more simply, the α-model. The formulation in probabilistic terms is the following.

$$\begin{aligned} \max \quad & f \\ \text{s.t.} \quad & Pr(cx\geq f)=\alpha \\ & Ax\leq b \\ & x\geq 0 \end{aligned}$$

The deterministic equivalent for this model is the following non-linear programming problem.

$$\max \quad f=\mu x-q\sqrt{xVx} \qquad (2.\ 8)$$
$$\text{s.t.} \quad Ax\leq b \qquad (2.\ 9)$$
$$x\geq 0, \qquad (2.\ 10)$$

where $q=F^{-1}(\alpha)$

If we assume $\alpha>.5$, which is very reasonable, then $q>0$ and f becomes a concave function.

There are several algorithms which solve (2.8)–(2.10). One of them is Kataoka's method (Kataoka [15]) which is specially designed for (2.8)–(2.10) His method uses Wolfe's quadratic programming algorithm as its main component. Another one is Geoffrion's method (Geoffrion [13]), using the bi-criterion method which he has developed for more general mathematical programs.

Lagrangean multipliers of (2.9) in the problem (2.8)–(2.10) give us $\partial f/\partial b=u$. Another evaluation which might interest the decision maker is the change of the α-fractile itself in response to the change of the value of α, that is,

$$\partial f/\partial\alpha=-\sqrt{xVx}/\varphi(q)$$

According to Geoffrion [13], his method for solving (2.8)–(2.10) can also give valuable information about the trade-off between the α-fractile and the expected value of the objective function (μx) as a byproduct of the computation of the optimal solution. This trade-off curve plots the relationship between α-fractiles and expected values of the objective function corresponding to the optimal solution (x's) as α varies.

II. 5 Probability maximum model (P-model)

As the name suggests, the criterion function of this model is the maximization of the probability that the objective function value (cx) exceeds a certain lower limit, θ. In the mathematical formulation,

$$\begin{aligned} \max\quad & P_\theta=Pr(cx\geq\theta)\\ \text{s.t.}\quad & Ax\leq b\\ & x\geq 0. \end{aligned}$$

The deterministic equivalent for this version is the following. (See Kataoka [16])

$$\max\quad f(x)=(\mu x-\theta)/\sqrt{xVx} \tag{2.11}$$

$$\text{s.t.}\quad Ax\leq b \tag{2.12}$$

$$x\geq 0. \tag{2.13}$$

Although we are assuming normality for the c vector, Roy [20] got the same problem as (2.11)–(2.13) using the Tchevychev approximation for $Pr(cx>\theta)$ without the normality assumption for c. In Roy's formulation, θ is called the disaster level and $Pr\ (cx\leq\theta)$ is considered to be the probability of survival.

Since $f(x)$ in the deterministic equivalent is not concave, but quasiconcave, we cannot use any ordinary algorithm for concave programming. Kataoka [16] and Geoffrion [13], however, have independently developed special algorithms for (2.11)–(2.13). Geoffrion's method uses the same bi-criterion method as in the α-model.

Several parametric evaluations may give us some useful information on the structure of the problem in this model. Some of them are

$$\partial f/\partial\theta=-1/\sqrt{xVx},\quad \partial f/\partial b=u$$

$$\partial P_\theta/\partial\theta=-\varphi((\theta-\mu x)/\sqrt{xVx})/\sqrt{xVx}$$

$$\partial P_\theta/\partial b=u\varphi((\theta-\mu x)/\sqrt{xVx}).$$

Since the decision maker has to determine the disaster level, θ, in advance, $\partial P_\theta/\partial\theta$ will give him very important ideas on his own selection of θ. He might want to change θ after knowing $\partial P_\theta/\partial\theta$, which is implied by a particular value for θ. In Geoffrion's method of computation, we can get the trade-off curve between $\mathscr{E}(cx)$ and P_θ as in the case of the α-model. This trade-off curve will help the decision maker reach a subjective judgment concerning the appropriate value of θ.

II. 6 Probability of loss model (PL-model)

This model maximizes the expected value of cx, subject to a probabilistic constraint on the loss which might occur. Its formulation is

$$\begin{aligned} \max \quad & \mu x \\ \text{s.t.} \quad & Pr(cx \leq \pi) \leq \gamma \\ & Ax \leq b \\ & x \geq 0. \end{aligned}$$

This probabilistic constraint means that the probability that the objective function value falls short of some predetermined level, π, has to be smaller than or equal to γ, which is also predetermined.

This model seems to be appropriate in many practical business situations, like portfolio selection of common stocks. The probabilistic constraint in this model is shown to be equivalent to Baumol's confidence limit restriction in mean-standard deviation analysis. (See p. 72 in Section III. 2.5 below.)

The deterministic equivalent for this model is

$$\max \quad f(x) = \mu x \tag{2. 14}$$

$$\text{s.t.} \quad \mu x - r\sqrt{xVx} \geq \pi \tag{2. 15}$$

$$Ax \leq b \tag{2. 16}$$

$$x \geq 0 \tag{2. 17}$$

where $r = -F^{-1}(\gamma)$.

Except in some extraordinary cases, we can assume $\gamma < .5$ and therefore $r > 0$. Positive r leads to a convex constraint set in this deterministic equivalent and we can use ordinary concave programming methods in solving this deterministic equivalent numerically.

Dual evaluations and parametric evaluations give the decision maker valuable insight into the problem in this case, too. For example,

$\partial f/\partial \pi = v$ $\quad v$ is a Lagrangean multiplier for (2.15)

$\partial f/\partial b = u$

$\partial f/\partial \gamma = v\sqrt{xVx}/\varphi(-r).$

$\partial f/\partial \pi$ and $\partial f/\partial \gamma$ might have special importance to the decision maker in selecting appropriate values of π and γ.

III. *Mean-Standard Deviation Analysis and Stochastic Programming Models*

In this chapter we shall see that the mean-standard deviation analysis developed mainly in the context of portfolio selection has some interesting correspondence with the special cases of stochastic programming models (or their deterministic equivalents) presented in the previous chapter.

III. 1 Mean-standard deviation analysis.

Markowitz first proposed in his pioneering paper, 'Portfolio Selection', the selection of portfolios of common stocks based on the mean and the standard deviation (or variance) of the stochastic return of the portfolios, and developed the concept of the mean-standard deviation efficient frontier (*E-S* frontier) of the available set of alternative portfolios. The efficient frontier means that there is no other feasible portfolio which has both higher mean return (E) and smaller standard deviation of the return (S) than those portfolios on the efficient frontier. If we accept the standard deviation as a measure of the risk of the portfolio, it is very reasonable to select the portfolio only from those on the E-S frontier, since every portfolio on the efficient frontier minimizes the risk for some level of the expected return, and maximizes the expected return for some level of the risk. This mean-standard deviation analysis is also justified as expected utility maximization at least in two cases. One case is when the utility function of the decision maker is quadratic, regardless of the types of the probability distribution of the stochastic return. The other is the case when the probability distribution of the stochastic return of the portfolio is a normal distribution, regardless of what utility function the decision maker has in his mind, except that it is concave. (See Tobin [20], Feldstein [9], Tobin [21].)

The E-S efficient frontier in the portfolio selection model with stochastic return cx may be traced out by solving the following non-linear programming model, changing S^0 parametrically from 0 to ∞.

$$\max \quad \mu x \tag{3. 1}$$
$$\text{s.t.} \quad \sqrt{xVx} \leq S^0 \tag{3. 2}$$
$$ex=1 \tag{3. 3}$$
$$x \geq 0, \tag{3. 4}$$

where e is a $1 \times n$ vector all of whose elements are unities. Here x_i is considered to be the proportion of the total fund which will be invested in the i-th financial asset, hence the sum of x_i's has to be unity. The resulting efficient frontier on the E-S plane (expected return, E, on the vertical axis and standard deviation, S, on the horizontal axis) is a continous curve made of segments of hyperboli. This curve typically will have some kinks. (See Markowitz [17] p. 153) Computationally it is easier to solve the following quadratic programming model, changing E^0 parametrically, in order to get the E-S frontier.

$$\begin{aligned} \min \quad & xVx \\ \text{s.t.} \quad & \mu x \geq E^0 \\ & ex=1 \\ & x \geq 0 \end{aligned}$$

Applying the same principle as in the mean-standard deviation analysis in portfolio selection, we can have the E-S frontier in the stochastic programming models in the previous chapter. This frontier will be traced out by solving the following non-linear programming model, which we could consider as a generalization of the model (3.1)–(3.4). This is a concave programming problem, since $\sqrt{xVx}$ is convex. We call it problem I.

$$\max \quad E=\mu x \tag{3. 5}$$
$$\text{s.t.} \quad \sqrt{xVx} \leq S^0 \tag{3. 6}$$
$$Ax \leq b \tag{3. 7}$$
$$x \geq 0 \tag{3. 8}$$

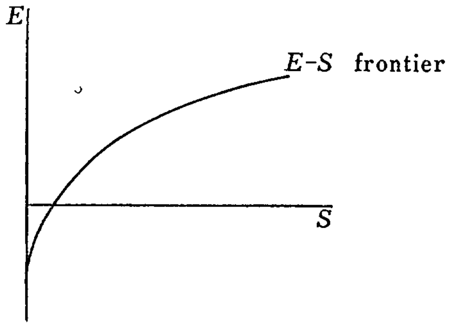
As a computational device, we can solve the following quadratic programming problem, changing E^0 parametrically, and get the mean-variance efficient frontier and then transform it into the mean-standard deviation frontier.

$$\min \quad xVx \tag{3. 9}$$
$$\text{s.t.} \quad \mu x \geq E^0 \tag{3. 10}$$
$$Ax \leq b \tag{3. 11}$$

$$x \geq 0 \tag{3. 12}$$

The *E-S* efficient frontier from the stochastic programming problem[2] is not necessarily made of segments of hyperboli on the *E-S* plane, but it is concave, if we consider this frontier as $E=E(S)$. This can be seen by a simple application of Lemma 1 in Gale [12]. This *E-S* frontier is depicted in Figure 1. If the *E-S* frontier cuts the *S*-axis, it is on the non-negative part of *S*-axis, since $\sqrt{xVx} \geq 0$. And we might have the leftward end point of this frontier on the *E*-axis, either on the positive part or non-positive part, when some riskless assets (like government bonds and cash) exist.

FIG. 1.



If we denote the slope of this frontier as dE/dS, its value at a point (E^0, S^0) is equal to the Lagrangean multiplier, k, to (3.6) in problem I, i.e.,

$$k=dE/dS|_{S=S^0}, \tag{3. 13}$$

except at several points where the *E-S* frontier may have kinks. At points of kinks, however,

$$dE/dS|_{-} \leq k \leq dE/dS|_{+} \tag{3. 14}$$

For the reasonings behind (3.13) and (3.14), Balinski and Baumol [3] have more general discussion on the properties of dual evaluators.

III.2 The relationship with stochastic programming models

When Markowitz proposed the *E-S* frontier in portfolio selection, he did not prescribe which point on the frontier the decision maker should select. He left this task to the decision maker's own subjective judgment. The decision maker has to have some decision model, either formal or intuitive, in order to select a single point on the *E-S* frontier from many points on the frontier. This should depend upon his attitude towards risk, or upon

[2] When we talk about the *E-S* frontier in the rest of this paper, it always means the *E-S* frontier from (3.5)-(3.8), not from (3.1)-(3.4).

his utility function for returns (i.e., for income or wealth), etc.

In the remaining portion of this chapter we shall see that all the stochastic programming models developed in the previous chapter, except the general model, have their optimal solutions on the *E-S* frontier. They, therefore, can serve as the mechanism for selecting a single point from the *E-S* frontier.[3] Thus stochastic programming models can be considered as further aides to the decision maker in the mean-standard deviation analysis, complimentary or supplementary to the *E-S* frontier.

III. 2.1 *E*-model

Since the *E-S* frontier is the north-west boundary of the feasible region (i.e., the shaded area in Figure 2) on the *E-S* plane, the E-model is easily seen to seek the highest point on the frontier. This is point *A* in Figure 2. Any other point in the feasible region (including the frontier) has smaller $E(=\mu x)$, and thus cannot be the optimal solution of *E*-model This case is rather trivial.

III. 2.2 V-model

The optimal solution of the V-model is point *B* in Figure 2, since V-model's objective is the minimization of the variance of *cx*, i.e., the minimization of the standard deviation, irrespective of the value of *E*. Point *B* is the most leftward point of the feasible region. This could be below the *S*-axis in some cases and above or on the *S*-axis in other cases.

FIG. 2.

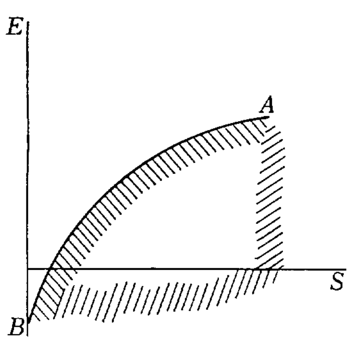
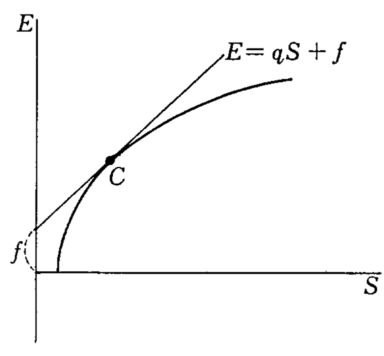


FIG. 3.



II. 2.3 α-model

Let us denote the optimal solution for the α-model x^0, and $\sqrt{x^0 V x^0} = S^0$, $\mu x^0 = E^0$. We assume $\alpha > .5$, hence $q > 0$.

[3] The author found, after completing this research, that Pyle and Turnovsky [19] took the criterion functions of the α-model, the P-model and the PL-model as the mechanisms for doing this and discussed the connection between these criteria and the expected utility approach. In this paper they show that these criteria 'lead to optimization of expressions involving the mean and standard deviation.' Here we further show algebraically that the optimization of these criteria *necessarily* leads to the optimal solutions on the *E-S* frontier, thus strengthening the validity of their contention of using these criteria in the mean-standard deviation analysis. That is, we establish a close relationship between stochastic programming models and the mean-standard deviation analysis. Notice that for some expressions involving the mean and the standard deviation, their optimizations do not necessarily lead to points on the *E-S* frontier.

Suppose there is a feasible combination of E and S ($E=\mu x$, $S=\sqrt{xVx}$) such that $E\geq E^0$, $S\leq S^0$. Since $q>0$,

$$E-qS\geq E^0-qS^0$$
$$\therefore \quad \mu x-q\sqrt{xVx}\geq \mu x^0-q\sqrt{x^0Vx^0}$$

The last inequality contradicts the optimality assumption of x^0. Hence there could not be feasible combination of E and S such that $E\geq E^0$ and $S\leq S^0$. This means x^0, hence (E^0, S^0) is on the E-S frontier.

In the E-S plane, the α-model seeks to get to the highest line with a predetermined slope, q, that is, a line with the greatest intercept. This will be obtained by the tangency point, C, in Figure 3. As in that figure, the objective function of the α-model is easily rewritten as

$$\max \quad f=E-qS$$

thus

$$\max \quad f \text{ in } E=qS+f \tag{3. 15}$$

III. 2.4 P-model

The objective function in this case is

$$\max \quad f=(\mu x-\theta)/\sqrt{xVx},$$

which is, in E and S notation,

$$\max \quad f \quad \text{in } E=fS+\theta$$

In the E-S plane the maximization of f is equivalent to the maximization of the slope of lines through (0, θ). This will lead to the tangency point, D, of these lines with the E-S frontier. Therefore, P-model picks up a point on the E-S frontier as its optimal solution. The similar proof by contradiction as in the case of α-model is also possible and easy to do in this case. Point D is depicted on the E-S frontier in Figure 4.

FIG. 4

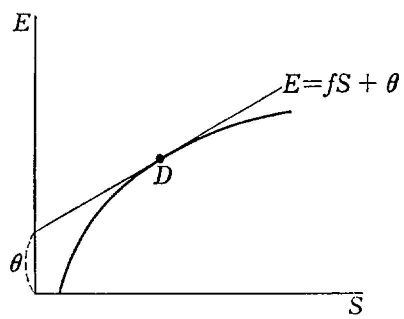
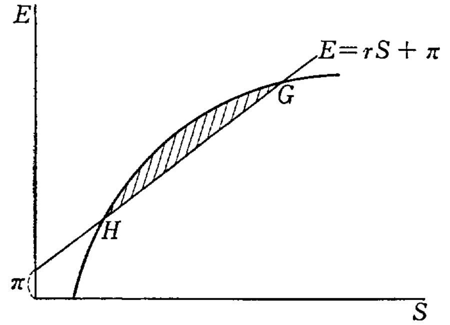


FIG. 5



III. 2.5 PL-model

Let x^0 be the optimal solution for the deterministic equivalent of PL-model, (2.14)–(2.17), and $E^0=\mu x^0$, $S^0=\sqrt{x^0Vx^0}$.

Since the point on the E-S frontier with $S=S^0$ is the solution of max $\{\mu x|Ax\leq b, \sqrt{xVx}=S^0, x\geq 0\}$, we have only to prove that x^0 is the optimal solution to this latter problem in order to prove that (E^0, S^0) is on the E-S frontier. Let us call this latter problem Problem (1) and call (2.14)-(2.17) Problem (2).

Let $\bar{x}$ be the optimal solution of Problem (1). x^0 is a feasible solution to both Problem (1) and (2) by construction. Therefore,

$$\mu x^0 - rS^0 \geq \pi, \tag{3. 17}$$

and since $\bar{x}$ is optimal for Problem (1),

$$\mu\bar{x} \geq \mu x^0. \tag{3. 18}$$

From (3.17) and (3.18)

$$\mu\bar{x} \geq \pi + rS^0$$

But $\quad S^0 = \sqrt{\bar{x}V\bar{x}}.$

Therefore

$$\mu\bar{x} \geq \pi + r\sqrt{\bar{x}V\bar{x}}$$

$$\therefore \quad \mu\bar{x} - r\sqrt{\bar{x}V\bar{x}} \geq \pi \tag{3. 19}$$

This means that $\bar{x}$ is a feasible solution for Problem (2). Thus, from the optimality of x^0 to Problem (2),

$$\mu x^0 \geq \mu\bar{x}. \tag{3. 20}$$

From (3.18) and (3.20)

$$\mu x^0 = \mu\bar{x}.$$

Therefore, if the optimal solution to Problem (1) is unique,

$$\bar{x} = x^0.$$

Even if the optimal solution to Problem (1) is not unique, x^0 has been shown to be the optimal solution to Problem (1). This implies that (E^0, S^0) is on the *E-S* frontier.

In the *E-S* plane, the probabilistic constraint of PL-model (or its deterministic equivalent (2.15)) is

$$E \geq rS + \pi,$$

and makes the feasible region narrower than without this constraint. The narrowed feasible region is the shaded part in Figure 5, and the *E-S* frontier is also shortened to the arc *GH*. The restriction (2.15) does the same thing as Baumol's confidence limit restriction (see Baumol [4]) and the arc *GH* is the same thing as what Baumol calls the *E-L* efficient frontier.

From this shortened frontier, the PL-model picks up point *G* according to the expected value ($\mathscr{E}(cx)$) maximization. *G* is the higher of the two points of intersection of the *E-S* frontier and the line

$$E = rS + \pi. \tag{3. 21}$$

IV. *Utility Implications of Stochastic Programming Models and Mean-Standard Deviation Analysis*

IV. 1 Some utility models

In the previous chapter, we saw that several stochastic programming formulations with normality assumptions that were discussed in Chapter II lead to optimal solutions on the *E-S* efficient frontier. In Chapter III we also mentioned that the quadratic utility function, combined with the expected utility maximization principle by von Neuman and

Morgenstern, leads to the selection of a point on the *E-S* frontier as the optimal solution. In this case, too, the decision maker will select different *E-S* combinations depending on the parameter values of his quadratic utility function.

Another utility function also leads to the optimal *E-S* combination which will be on the *E-S* frontier, if we assume expected utility maximization and normality of the random return in the model (see Freund [10]). In his model, Freund suggests the exponential utility function,

$$U(R)=1-exp(-aR)$$

where R is the random variable (return, etc.) and a is a positive parameter, implying the risk-aversion of the decision maker.

Taking the expected value of this utility function and assuming a normal distribution for R, Freund derived the following objective function to be maximized (see Freund [10]).

$$f=E-aS^2/2 \qquad (5.\ 1)$$

where $E=\mathscr{E}(R)$, $S=\sqrt{\mathscr{E}(R-E(R))^2}$. (5.1) clearly shows that the maximization of this function will lead to the selection of some point of the *E-S* frontier.[4]

As a short summary of what we have shown in the previous chapter and also the firs part of this section, we can say that the *E-S* frontier could be considered as a set of optimal solutions of various decision models, some of which are stochastic programming-type models with normality assumptions, the others being the expected utility maximization models using a quadratic or an exponential utility function.

In this sense, the mean-standard deviation analysis can be said to be a comprehensive approach to decision problems under risk, embracing several specific decision models within its general framework.

IV. 2 *E-S* indifference curves

In picking up a point from the *E-S* frontier as the optimal *E-S* combination, we essentially use the *E-S* indifference curve in the *E-S* plane. There could be many types of indifference curves, depending on what decision model we use.

If we denote the maximand in each model discussed before as f, a list of the maximands in those models is as follows.

model	maximand
E-model	$f=E$
V-model	$f=-V=-S^2$
P-model	$f=(E-\theta)/S$
α-model	$f=E-qS$
PL-model	$f=E$
quadratic utility model[5]	$f=E-a(E^2+S^2)$
Freund's model	$f=E-aS^2/2$

From these maximands, it is very easy to derive the corresponding indifference curves on the *E-S* plane. Since we have been assuming that the *E-S* frontier is computed subject

[4] For the ease of explanation, we assume that R in Freund's model (and quadratic utility model) is essentially cx in stochastic programming models and we have the side condition $Ax\leq b$, $x\geq 0$, also in both models. In this way, we can talk about the same constrained *E-S* frontier, all through this paper without explicitly saying so.

[5] The quadratic utility function is assumed to be

$$U(R)=R-aR^2.$$

to the side conditions, $Ax \leq b$, $x \geq 0$, we can determine the optimal solutions to each model using this E-S frontier and the derived indifference curves, a list of which is as follows.

Model	Indifference Curve
E-model	$E=f$
V-model	$S=-f$
α-model	$E=qS+f$
P-model	$E=fS+\theta$
PL-model	$E=f$
quadratic utility model	$(E-1/2a)^2+S^2=-f/a+1/4a^2$
Freund's model	$E=a/2S^2+f$

In each family of indifference curves, f is the parameter which distinguishes one curve of the indifference curve family from the other curve of the same family.

It can be easily seen that all the indifference curves for the various stochastic programming models are linear in the E-S plane, some of which are horizontal, some vertical, some sloping.

The indifference curve for the quadratic utility model is the circle with a center on the E-axis with $E=1/2a$. Since we want to maximize f, we in effect want to make these indifference circles as small as possible, but at the same time compatible with the E-S frontier. The optimal solution for this model, therefore, is the point of tangency between the indifference circle and the E-S frontier, depicted as I in Figure 6.

FIG. 6

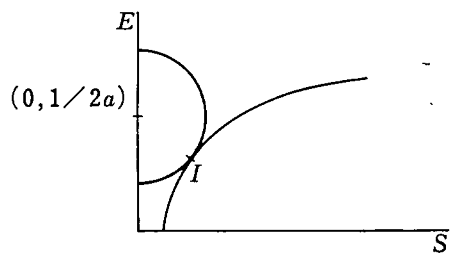
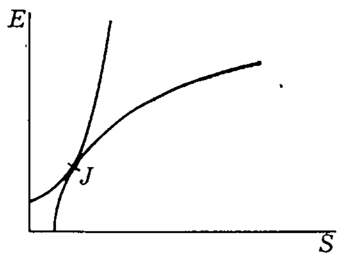


FIG. 7



In Freund's model, the indifference curves are a family of paraboli which are symmetrical with respect to the E-axis and with a slope aS. The optimal point in this model is the point of tangency between the E-S frontier and the parabola with as great an intercept with the E-axis as possible, the point J in Figure 7.

IV. 3 Utility implications of stochastic programming models

Since we have not connected the expected utility maximization and stochastic programming formulations in any explicit way so far, let us now try to see what meaningful information can be derived from this effort. This effort means that we try to find the forms or properties of utility functions which are implied if we suppose that each stochastic programming model is an application of the expected utility maximization principle with a suitable utility function.[6] We shall now try to identify the utility function which is thus supposed to underlie each stochastic programming model. If identification of this function is not possible, then we will try to find out some of the properties of this utility function. As an essential assumption in this section we suppose that the expected utility maximization principle is consistent with each of the stochastic programming formulations in Chapter III. When the normality of c in the objective function is not assumed, K. Borch has shown by a very ingenious example of a special probability distribution that PL-model is, at least in one case, not compatible with the axioms of the expected utility maximization principle. (See Borch [5], pp. 42) We maintain, however, the normality assumption of c throughout this paper.

For the E-model, the maximand is just $E=\mathscr{E}(cx)$. If we suppose that

$$\mathscr{E}(U(cx))=E,$$

then U has to be either cx or else a positive linear transformation of cx.[7]

For the V-model,

$$\mathscr{E}(U(cx))=-S^2,$$

and therefore

$$U(cx)=-(cx-E)^2.$$

Thus, in this case, the underlying utility function is a special type of the quadratic utility function.

For the α-model, P-model and PL-model it is very difficult to identify the underlying utility functions, although it is rather easy to find an important property of these functions with important implications. If we consider the maximands in the deterministic equivalents of the α-model and the P-model as expected values of some utility functions, we have

for the α-model: $\mathscr{E}(U(cx))=E-qS$

for the P-model: $\mathscr{E}(U(cx))=(E-\theta)/S$.

Although it is very difficult to work backward from these expected values to the underlying utility functions, we can at least deduce an important implication. That is, as we have seen in the previous section, the indifference curves in the E-S plane in these three cases are linear. As Tobin has shown in his pioneering work (Tobin [21]), an indifference curve has to be strictly convex in the E-S plane for a utility function of strict concavity' that is, a risk-averter's utility function, when normality is assumed for the random benefit, in this case, cx.

From this result we can at least say that the underlying utility functions for the α-model, P-model and PL-model are not utility functions of a risk-averter, since their indifference curves in the E-S plane are not strictly convex.

V. *Imputation of Key Parameters of Various Decision Models*

It seems reasonable to assume that behind every decision made by a decision maker there lies some decision model, formal or intuitive. In the context of the discussions on decision under risk that we have developed so far, we will now try to impute the value of

[6] For another different approach to see the connection between the expected utility maximization and these stochastic programming models, see Pyle and Turnovsky [9].

[7] Because a utility function is unique only up to a positive linear transformation.

the key parameter in various decision models implied in the decision maker's decision under risk, i.e., a combination of E and S he decided to select.

In particular a decision maker is assumed to select a point on the E-S frontier, subjectively or objectively, as the most suitable E-S combination to him. If we further assume that he is utilizing a particular decision model, explicitly or implicitly, we can impute parameter values in this decision model he is supposed to be using from the information contained in that selected point. Thus we can give him important information about the attitude toward risk implied by his selected point. This information is, of course, if-then type information, like 'If he is assumed to be using the α-model in his selection, the selected point implies that the value of the key parameter in the α-model, α, must be such and such.' In these imputations, we also assume that we have complete information about the normal distribution of c, as the decision maker is assumed to have.

Now let us denote the selected optimal solution as x^0 and the optimal E and S as E^0 and S^0, where

$$E^0=\mu x^0,\ S^0=\sqrt{x^0 V x^0}$$

V.1 Slope of the E-S frontier

As we shall see shortly, the slope of the E-S frontier plays an important role in imputing parameters to various decision models. We shall now discuss how to obtain this slope in computational terms.

In Chapter III, we saw that we can obtain the E-S frontier by solving (3.5)–(3.8) parametrically or, as a computational substitute, we can more easily solve the quadratic programming problem (3.9)–(3.12).

The slope of the E-S frontier, denoted as $dE\,/\,dS$, is easily obtained from the program (3.5)–(3.8) as the Lagrangean multiplier for (3.6), k, for a particular value of S, S^0, when we use the problem formulation (3.5)–(3.8) to derive the E-S frontier.

In fact, we have a sequence of values for k, each of which corresponds to some value of S. From this sequence, which will typically be supplied as byproduct of the optimization computation, we can determine $dE/dS|_{S=S^0}$ for any S. When the sequence of values for k is not directly available from the optimization computation, we can get $dE/dS|_{S=S^0}$ by solving the following linear programming problem for u and k.

$$\min\ ub+S^0k \tag{6. 1}$$
$$\text{s.t.}\ uA+x^0Vk/S^0\geq\mu \tag{6. 2}$$
$$u,\ k\geq 0 \tag{6. 3}$$

This is the dual problem of (3.5)–(3.8) with the optimal x^0 substituted, since, by assumption, we know x^0 from the decision maker. We can consider the optimal solution, k, in (6.1)-(6.3) as the Lagrangean multiplier for (3.6) in the concave programming problem (3.5)–(3.8) with $S=S^0$. For the reasoning behind this method, see p. 66 in Chapter II above.

When we use the parametric quadratic programming model (3.9)–(3.12) for the derivation of the E-S frontier, we usually get the sequence of values for the Lagrangean multiplier of (3.10), i.e., $dS^2\,/\,dE$, as a byproduct of the quadratic programming computation. In that case,

$$dS^2/dE=2SdS/dE \tag{6. 4}$$

Therefore except at point of kinks on the E-S frontier where we cannot define $dE\,/\,dS$

uniquely,

$$dE/dS=2S/(dS^2/dE) \qquad (6.5)$$

In these ways we can get the slope of the *E-S* frontier, $dE/dS|_{S=S^0}$ through various computational methods.

If the decision maker happens to select a kinky point on the *E-S* frontier as his optimal decision, dE/dS cannot be defined uniquely and the linear programming problem (6.1)–(6.3) will not have a unique solution. For further discussions, see p. 70 in Chapter III above.

In the following we assume that the decision maker has selected a regular (not kinky) point, and we let k^0 denote the slope of the *E-S* frontier at $S=S^0$.

V. 2 Imputation of parameters

If we know the value of the slope of the *E-S* frontier at $E=E^0$ and $S=S^0$ using one of the methods discussed in the previous section, we can impute the values of the key parameters of decision models by using the various relationships between those models and mean-standard deviation analysis discussed in Chapters III and IV. Since the E-model and the V-model of stochastic programming do not contain any parameters in them, we will omit these two models from the following discussion.

V. 2.1 α-model

If we assume that the decision maker is using the α-model as his underlying decision model under risk, we can then impute the values of the key parameters, q and α, from his selection of (E^0, S^0) from the *E-S* frontier in the following way.

As we discussed in Chapters III and IV, the optimal solution for the α-model is given by the tangency point in Figure 3 in Section III.2.3. At this point

$$k^0=q,$$

or

$$\alpha=F(k^0).$$

Therefore, assuming the α-model for the decision model, we can impute the risk tolerance, α, for the α-fractile in the decision maker's attitude toward risk.

V. 2.2 P-model

When we assume that the decision maker has selected the point (E^0, S^0) using the P-model as his implicit decision model, the following imputation of the parameter value, θ, that he is supposed to be using is possible.

Since what we are maximizing in this model is the slope of the line through (0, θ), we have the following relationship from (3.16) and the fact that the maximization is attained at the point of tangency, D, in Figure 4,

$$E^0=k^0S^0+\theta,$$

substituting k^0 for f, the maximand.

From this,

$$\theta=E^0-k^0S^0.$$

Therefore, in the P-model, the lower limit on the objective or disaster level, θ, can be imputed.

V. 2.3 PL-model

If the decision maker uses the PL-model, implicitly or explicitly, in making his selection from the E-S frontier, we can derive the following linear relationship between two parameters in the model, namely, r and π, from (3.21).

$$E^0 = rS^0 + \pi \qquad (6.8)$$

In this case we can only use the fact that the line (3.21) goes through (E^0, S^0). The slope of the E-S frontier at (E^0, S^0) is not relevant here.

Since we have only one relationship for two parameters, we cannot uniquely impute both r and π. If we set, however, one of these two parameter values, we can of course impute the value for the other parameter. For example if we set π, the minimum required level for the objective, cx, we can impute the value for γ from the decision maker's selection, (E^0, S^0) as follows,

$$\gamma = F((E^0 - \pi)/S^0).$$

This will give us the decision maker's implicit subjective judgement for the percentage of the time that he will tolerate the underattainment of the required level, π. Conversely, of course, we can impute π by setting γ.

V. 2.4 Quadratic utility model

In this model, we can obtain the slope of indifference curves by totally differentiating the expected utility function, $\mathscr{E}(U) = E - a(E^2 + S^2)$.

$$0 = dE - a(2EdE + 2SdS)$$
$$dE/dS = 2aS/(1 - 2aE).$$

Since the optimal solution for this model is the point of tangency between the expected utility function and the E-S frontier (see Figure 6),

$$dE/dS = k^0 \qquad \text{at} \quad E = E^0,\ S = S^0$$
$$2aS^0/(1 - 2aE^0) = k^0.$$

Rearranging for a,

$$a = k^0/(2k^0E^0 + S^0).$$

Thus the only parameter of the quadratic utility function, a, can be imputed and we can completely specify the quadratic utility function for this decision maker by knowing his selection of a point of the E-S frontier, assuming that his decision is implicitly or explicitly based on the maximization of the expected value of the quadratic utility function.

V. 2.5 Freund's model

If the decision maker uses Freund's model as his decision model, we can specify the exponential utility function used in his model from the decision maker's own selection, (E^0, S^0).

As we have discussed in the second section in Chapter IV, Freund's model has indifference paraboli with a slope aS on the E-S plane, and picks up the optimal solution as the tangency point of this indifference curve and the E-S frontier. Thus if (E^0, S^0) is the optimal solution for this model,

$$k^0 = aS^0,$$
$$\therefore \quad a = k^0/S^0.$$

a is the only parameter in Freund's model, which he calls a parameter of 'risk aversion'.

Complete specification of the underlying utility function is thus possible from knowledge of the decision made and the assumption that Freund's model has been utilized.

VI. *Concluding Remarks*

In this paper we have investigated several types of implications of using some stochastic programming models in decisions under risk. In particular, we have found that parametric evaluation using dual evaluators can give some helpful information in determining appropriate values of key parameters in the models. We have also found that each stochastic programming model presented in this paper (except the general model) has its optimal solution on the *E-S* frontier. Furthermore, we have examined indifference curves for stochastic programming models and some utility models, and we have seen that a risk-averter could not use the E-model, α-model, P-model and PL-model because they implicitly contradict the concavity property of his utility function.

Since mean-standard deviation analysis is rather popular in investment decision analysis, we have analyzed the implications that subjective selection on the *E-S* frontier has, in terms of its relationship with various decision models. Since we can impute the value of key parameters in each decision model presented in this paper, this approach could be used as a framework for empirical research on investors' or other decision makers' attitude toward risk.

Several implications that we have seen so far in stochastic programming models could give a decision maker valuable information in deciding many things when using these stochastic programming models, including of course the final decision on the course of action he should take.

He has to, first of all, decide what particular model to use in his case. At this stage, the implications that the particular model has in terms of the way it selects a point from the *E-S* frontier or the utility implication of the particular model may give him some hint and insight.

Next, he has to determine, in most cases, the specific value for the key parameter(s) to use in his selected model (e.g. the value for α in the α-model.) Parametric evaluation in Chapter II will give him some insights into the trade-off structure in the model.

When he has had the optimal solution from his selected model with a particular parameter value he has determined, he can impute the values of key parameters in other models which are implied by his optimal solution from his selected model. These imputed values of key parameters of other models provide additional information which may induce him to reconsider the tentative decision he has already made.

In these ways we can provide several types of dialogues between formal models and the decision maker himself. This man-model interaction is one of the most essential factors in any successfully implemented management science project.

Thus, the purpose of this paper as a whole might be summarized as an effort to give more aids to a decision maker under risk-taking situation in a more practical manner, not by inventing new models or new tools, but by letting him know many implications arising from the uses of stochastic programming models. Although it is certainly important to invent and devise new and better models and tools in management science, it is no less important to search for more practical and comprehensive ways of utilizing models and tools that have been developed so far. (August 1971)

REFERENCES

1. Abadie, J. (ed.), *Nonlinear Programming*, North-Holland, Amsterdam, 1967.
2. Agnew, N.H., R.A. Agnew, J. Rasmussen and K.R. Smith, 'An application of chance-constrained programming to portfolio selection in a casualty insurance firm,' *Man. Sci.*, 15 (10), June 1969, B512-B521.
3. Balinski, M. L. and W.J. Baumol, 'The dual in nonlinear programming and its economic interpretation,' *Review of Econ. Studies*, 35 (3), July 1968, 237-256.
4. Baumol, W.J., 'An expected gain-confidence limit criterion for portfolio selection,' *Man. Sci.*, 10 (1), October 1963, 174-182.
5. Borch, K., *The Economics of Uncertainty*, Princeton Univ. Press, Princeton, N.J., 1968.
6. _______, 'A note on uncertainty and indifference curves,' *Review of Econ. Studies*, 36 (1), January 1969, 1-4.
7. Charnes, A. and W.W. Cooper, 'Chance constraints and normal deviates,' *JASA*, 57 (297), March 1962, 134-148.
8. _______ and _______, 'Deterministic equivalents for optimizing and satisficing under chance constraints,' *Oper. Res.*, 11 (2), January-February 1963, 18-39.
9. Feldstein, M.S., 'Mean-variance analysis in the theory of liquidity preference and portfolio selection,' *Review of Econ. Studies*, 36 (1), January 1969, 5-12.
10. Freund, R.J., 'The introduction of risk into a programming model,' *Econometrica*, 24 (3), July 1956, 253-263.
11. Hadley, G., *Nonlinear and Dynamic Programming*, Addison-Wesley, Reading, Mass., 1964.
12. Gale, D., 'A geometric duality theorem with economic applications,' *Review of Econ. Studies*, 35 (1), January 1968, 18-24.
13. Geoffrion, A.M., 'Stochastic programming with aspiration or fractile criteria,' *Man. Sci.*, 13 (9), May 1967, 672-679.
14. _______, 'Solving bi-criterion mathematical programs,' *Oper. Res.*, 15 (1), January-February 1967, 38-54.
15. Kataoka, S., 'A stochastic programming model,' *Econometrica*, 31 (1-2), January-April 1963, 181-196.
16. _______, 'Stochastic programming——maximum probability model,' *Hitotsubashi Journal of Arts and Sciences*, September 1967, 51-59.
17. Markowitz, H., *Portfolio Selection*, Wiley, New York, 1958.
18. Naslund, B. and A. Whinston, 'A model of multi-period investment under uncertainty,' *Man. Sci.*, 8 (2), January 1962, 184-200.
19. Pyle, D.H. and S.J. Turnovsky, 'Safety-first and expected utility maximization in mean-standard deviation portfolio analysis,' *Review of Econ. and Stat.*, 52 (1), February 1970, 75-81.
20. Roy, A.D., 'Safety-first and the holding of assets,' *Econometrica*, 20 (3), July 1952,

431-449.

21. Tobin, J., 'Liquidity preference as behavior towards risk,' *Review of Econ. Studies*, 25 (2) February 1958, 65-86.
22. _______, 'Comment on Borch and Feldstein,' *Review of Econ. Studies*, 36 (1), January 1969, 13-14.
23. Wolfe, P., 'Methods in nonlinear programming,' in Abadie [1].
24. Wolfe, P., 'A duality theorem in nonlinear programming,' *Quarterly of Applied Mathematics*, 19 (3), 1961, 239-244.