A SYNTHESIS OF STOCHASTIC LINEAR PROGRAMMING

by

Charles Samuel Matheny

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in candidacy for the degree of

MASTER OF SCIENCE

in

Statistics

APPROVED:______________________________
Professor Whitney Johnson

____
Dr. Boyd Harshbarger

____
Professor Brian W. Conolly

____
Dr. R. H. Meyers

March, 1967

Blacksburg, Virginia

TABLE OF CONTENTS

I. INTRODUCTION

1.1 Discovery and development of linear programming

1.1.1 Early operations research

It was once said that war often speeds and gives emphasis to research that peacetime could never afford. Linear programming, although not a direct outgrowth of a war was made necessary by the need for rapid military decisions. Wood and Geisler (33) stated in 1951 that:

> "It was once possible for a Supreme Commander to plan operations personally. As the planning problem expanded in space, time, and general complexity, however, the inherent limitations in the capacity of any one man were encountered. Military histories are filled with instances of commanders who failed because they bogged down in details, not because they could not have eventually mastered the details, but because they could not master all the relevant details in the time available for decision."

The one-man planning of military operations continued until about 1860 when the use of a General Staff emerged. This was the early approach to operations planning, permitting the subdivision of the planning process. By this point in time much work had been done on linear equations by Fourier (11) and Gauss(14). Dantzig (6) points out that Fourier "may have been aware of its potential" as early as 1823. It is very doubtful that, even if the complete theory of linear programming were known at this early date, it could have been applied to military problems since a) their great complexity would have required large computers and b) they had not been defined as a set of linear inequalities.

So the military continued to further subdivide the planning process and theory continued to progress in the general direction of linear programming. In 1939, Kantorovich (17) of the U.S.S.R. advanced proposals that were very close to linear programming, but oddly, in a society that depends on central planning, his proposals were ignored.

World War II put a tremendous strain on the staff planning concept. It was during this war that a Central Monitor was placed on the process. Dantzig (6) stated that:

> "The entire program was started off with a war plan in which were contained the wartime objectives. From this plan, by successive stages, the wartime program specifying unit deployment to combat theaters, training requirements of flying personnel and technical personnel, supply and maintenance, etc., was computed. To obtain consistent programming the ordering of the steps in the schedule was so arranged that the flow of information from echelon to echelon was only in one direction, and the timing of information availability was such that the portion of the program prepared at each step did not depend on any following step. Even with the most careful scheduling, it took about seven months to complete the process."

After the war the control function was further consolidated and it became quite obvious that the application of scientific computational techniques would be in order. Several events had taken place by 1947 which brought together the knowledge and capability to tackle such problems; a) the design and construction of large scale computers, such as the Harvard Mark I and the ENIAC and b) the expression, in terms of linear inequalities, of allocation problems complete with objective functions such as the diet problem and Leontief's input-output model without the usual objective function and c) a growing interest in utility functions which as will be seen had an even greater impact on stochastic linear programming.

1.1.2 Project SCOOP

These developments led to the formation in 1947 of a group of scientists who, under the sponsorship of the United States Air Force, were charged with the investigation of the Scientific Computation of Optimum Programs. Primary participants in this effort were G. B. Dantzig, M. Wood, John Norton and Murray Geisler. From their work resulted the simplex method for solving linear programs. Due to Air Force sponsorship, the first real applications were to Air Force problems. Several of these applications were(6):

> "(a) contract bidding, (b) balanced aircraft, crew training, and wing deployment schedules, (c) scheduling of maintenance overhaul cycles, (d) personnel assignment, and (e) airlift routing problems."

It was only after Project SCOOP had demonstrated the usefulness of the technique and its application to many problems, that economists began to use it in their work. Much of their past work led up to the discovery of the simplex method but none seemed inclined to attempt to develop a practical algorithm. Several economists such as Von Neumann approached the linear programming problem with appropriate linear formulations but resulted only with an elegant mathematical theorem. When the technique was developed, T. C. Koopmans organized a conference on linear programming held in Chicago on June 20-24, 1949, and attended by many of the most well-known economists and mathematicians of the day. This began a period of intensive study of applications of linear programming to economic problems and to "theory of the firm".

1.2 Extension to stochastic linear programming

1.2.1 Dantzig's first paper

This state of affairs continued for several years with literally hundreds of papers being written on linear programming. All of these papers had one point in common; they assumed deterministic costs, coefficients and resources. In 1955, Dantzig (5) published a paper which for the first time considered the possibility of uncertainty in the determination of the costs. This problem had undoubtly been apparent before this time but the increased efficiency brought about by the use of linear programming with the assumption of deterministic coefficients had been sufficiently great as to make this further refinement unnecessary. Dantzig's approach seemed in retrospect to be somewhat over simplified and to have missed the point which is basic to both linear programming and statistics, that is, that the final objective is a decision and a decision is subjective.

1.2.2 Early Problem Formulations

Dantzig, in his 1955 paper (5), examined the situation of random costs, where the problem formulation is

minimize $c = px$

subject to $Ax = b$

where p, x and b are vectors

and A is a matrix of coefficients.

Then he stated that to minimize the expected cost we merely minimize

$$E(c) = E(p)x$$

$$= \bar{p}x$$

$$\text{where } \bar{p}_j = E(p_j)$$

which will indeed minimize the expected cost. Note that nothing at all has been said about the variance of the p_j or about confidence in the value of c computed in this manner.

Dantzig then explored a technique for solving the special case of the linear programming problem used by Ferguson and Dantzig (10) in solving the aircraft allocation problem. The problem as stated by Dantzig was to:

$$\text{minimize } c = \sum_i \sum_j c_{ij} x_{ij} + \sum_j v_j$$

$$\text{subject to } \sum_j x_{ij} = a_i$$

$$\text{and } \sum_i b_{ij} x_{ij} = \mu_j$$

where x_{ij} is the amount of the i^{th} resource assigned to the j^{th} destination,

b_{ij} is the number of units of demand at destination j which can be satisfied by one unit of resource i,

c_{ij} is the cost of moving one unit of the i^{th} resource to the j^{th} destination,

$$\text{and } d_j = \mu_j + v_j - s_j$$

where d_j is the demand,

v_j is the shortage,

and s_j is excess.

He then proceeded to prove that the expected value of the objective function is convex, which was necessary in order that it have a unique and readily computable minimum. This led to an approximation technique in which the continuous objective function is replaced by a step function. This technique involved only the expected value and therefore made the implicit assumption that the variances are equal. The technique would not, therefore, be generally useful.

II VARIOUS FORMULATIONS OF THE PROBLEM

In discussing the various formulations of stochastic linear programming, it should be noted that all of these formulations revolve around one or a combination of random variates; the cost vector, the matrix of coefficients and the resources or right hand side. In the majority of papers reviewed, the possibility of a random coefficient matrix was ignored. This will also be done here, except to mention that a paper by Barbar (1) gives an excellent discussion of the matter. In that paper, the situation is treated as a problem in solution stability and the probability that a given solution will remain optimal under variation of the coefficient matrix is determined. This same problem can be approached with perturbation techniques(24).

2.1 Dantzig's two stage problem

In his first paper on stochastic linear programming (5) and later in his book on linear programming (6), Dantzig formulated a method of solution which Madansky (20) later dubbed as the "Here and Now" method. It was:

$$\text{maximize } z = c'x + f'y$$

$$\text{subject to } Ax \leq b^*$$

$$\text{and } Ax + By \leq b$$

where c' is the vector of "profits" on x

and f' is the vector of penalties for constraint violation.

This is a two stage problem with a random resource vector, where, in the first stage, the vector b^* is unknown and must be estimated and the x vector chosen. Then in the second stage, the b vector becomes known and the 'slack' vector, y, must be chosen to minimize any error which occurred in the choice of x. This method was supposed to represent the actual decision process which management experiences and reduces to the problem of determining a method for finding the "best" estimate of the resource vector, b. Here "best" may not be "best" in the usual sense. It may well be that the "best" estimate is one in which $P(Ax \geq b) \leq \varepsilon$ for small values of ε.

2.2 Elmaghraby's allocation problem

In 1959 S. E. A. Elmaghraby published a paper (8) dealing with a linear programming problem having a cost vector which was distributed as a discrete random variable. The method of solution presented was to separate each variable into several new variables. Each of these new variables would then be used to represent only a portion of the density function of the variable for which they were substituted. This approach was quite similar to Dantzig's approach to the allocation problem which, notably, was not listed in the bibliography of the paper.

Then, in 1960, Elmaghraby (9) published a paper which was a condensation and extension of his Ph. D. thesis (7) of 1958. In this paper he addresses the allocation (transportation type) problem which he distinguishes from the programming problem. The problem is formulated as:

$$\text{minimize } c = f(x_{ij})$$

$$\text{subject to } \quad g_i(x_{ij}) \leq a_i \qquad x_{ij} \geq 0$$

where the $g_i(x_{ij})$ are of the form

$$\sum_j x_{ij} \leq a_i$$

and the $f(x_{ij})$ are continuous density functions.

Elmaghraby then proved the existence of an optimum by virtue of the convexity of the objective function and constraints and then established the necessary conditions for a feasible x to be the optimal x.

2.3 Freund's risk programming

In 1956, Freund published a paper (12) on risk programming. In this paper he considered the problem from the viewpoint of a statistician, and in so doing generated one of the best of the early papers on stochastic linear programming. He treated the problem:

$$\text{maximize} \quad c'x = r$$

$$\text{subject to} \quad Ax \leq b, \quad x \geq o$$

$$\text{where } c_i : N\{\mu_i, \sigma_i^2\}$$

Then the net revenue, r, is distributed

$$r : N(\mu' x, x' Vx)$$

where $\mu = E(c)$
V = matrix of variances and covariances

He then introduced a utility function, $y(r) = 1 - e^{-\beta r}$, which he assumed will describe the behavior of the decision maker. There is no justification for that particular function's being used except that it would seem to represent a "conservative" decision maker. It is a strong point in that it involves the notion of utility. He illustrated the point by showing other utility functions which might represent a "gambler" and so forth.

Using the specified utility function, Freund computed its expected value

$$E(u) = \int_{-\infty}^{\infty} (1-e^{-\beta r})e^{-(r-u)^2/2\sigma^2}dr$$

and showed that maximization of the above is accomplished if

$$E(u^*) = u-\frac{\beta}{2}\sigma^2$$

is maximized or, in terms of the original problem variables;

$$E(u^*) = c'x-\frac{\beta}{2}xVx$$

subject, of course, to

$$Ax \leq b$$
$$x \geq 0$$

Notice that in this formulation of the problem, one can choose to maximize any choice of mean and variance simply by the appropriate choice of the constant β. Thus a larger β would tend to be more conservative, introducing, less "risk" into the final solution. This was similar to the "coefficient of optimism" introduced by Luce and Raiffa(19).

Finally, Freund introduced an unproven but workable algorithm to solve the resulting non-linear (note the squared term in the objective function) programming problem. This result was based on an earlier paper of Hildreth(16). Hildreth's paper will not be covered here as the author feels that more sophisticated techniques are now available (cf. Williams(31)).

The algorithm which Elmaghraby developed to solve the problem bears some resemblence to that of Freund's and, although he references many of those works upon which Freund bases his method, he does not reference Freund's paper.

These results of Freund's were further extended in 1959, in an unpublished thesis by Rein (23), to allow a ready choice of the variable β by plotting isograms of conservatism thus giving a basis to the subjective evaluation of this constant.

2.4 Madansky's "Here and Now" problem

Madansky, (20) who worked with Dantzig in later development of the two stage problem, tagged the two stage problem the "Here and Now" method. In this problem, a decision must be made in the first stage (here and now), knowing only the distribution of the resource vector. This is in contrast to the "Wait and See" approach in which the decision maker "waits" and observes the random vector and then solves the problem. This latter method is how Madansky characterized the approach of Freund and Tintner(26).

In his first paper, using this approach, (20) Madansky merely derived certain inequalities which show the conditions under which the "Here and Now" results approach the "Wait and See" results and further, under what conditions the use of the expected value of the resource vector will yield good results.

This result is developed by letting the objective function of the stochastic linear programming problem represented by $C(b,x)$ and proving this to be a convex, continuous function of b.

Now, note that the situation in which the expected value of the objective function is to be minimized by using an observed value of the vector b, is to be contrasted with that of using the expected value of b. If $\bar{x}(E(b))$ is used to represent the value of x which minimizes the objective function involving the expected value of b, $C(E(b),x)$, then consider the function of b, $C(b,\bar{x}(E(b)))$. The values which can be assumed by varying b over its allowable range can be determined by noting that since $A\bar{x}(E(b)) \geq E(b)$, then for this inequality to hold with probability 1, i.e. $Pr(A\bar{x}(E(b)) \geq b) = 1$, b must be restricted to values greater than $E(b)$. Therefore, for values of b over this restricted range, it follows that $C(b,\bar{x}(E(b))) \geq$ minimum $E(C(b,x))$.

Then let $\bar{x}$ be the x which minimizes $EC(b,x)$ and let $\bar{x}(b)$ be that x which minimizes $C(b,x)$. In the former case b is assumed to be an as yet unobserved random vector ("Here and Now") and in the latter case, b has already been observed ("Wait and See").

Next, note that in the former case, in order that $Pr(A\bar{x} \geq b) = 1$, $\bar{x}$ must be chosen to be much larger or at least as large as $\bar{x}(b)$ in the "Wait and See" situation. Thus, it follows that $C(b,\bar{x}) \geq C(b,\bar{x}(b))$ and hence $EC(b,\bar{x}) \geq EC(b,\bar{x}(b))$ and also minimum $EC(b,x) \geq E$ minimum $C(b,x)$ or, in other words, the minimum value of the objective function of the "Here and Now" problem will always be at least as great as the minimum value of the "Wait and See" problem.

The above proof is based primarily upon the work of Madansky, but this author added those elements of the proof referencing $P_r(Ax \geq b) = 1$. The question of considering the effect on the objective of allowing the case $P(Ax \geq b) \leq 1$ to occur (cf. Charnes and Cooper (3,4)) was not mentioned by Madansky. This case should be considered since it will, in general, allow "better" solutions to the problem. To observe this, it is only necessary to compare that value of the vector x which minimizes

C(b,x) subject to $Pr(Ax > b) = 1$, to the value of the vector x, say x which minimizes C(b,x) subject to $Pr(Ax \leq b) \leq \varepsilon$ for small ε. That is, in the latter case, the constraints may be violated with probability ε. Thus the value of minimum $C(b,x) > C(b,x)$. A "more optimal" solution was obtained by allowing a certain amount of risk.

Finally, returning to the above proof, Madansky invokes the convexity and continuity of the function C(b,x) with respect to b along with Jensen's inequality and shows that

$$E \min_x C(b,x) > \min_x C(E(b),x)$$

and hence

$$E(C(b,\bar{x}(Eb))) \geq \min_x E(C(b,x)) \geq E(\min_x C(b,x))$$
$$\geq \min_x C(E(b),x)$$

One notes that bounds have now been established on the value of the objective function of the "Here and Now" problem, $\min_x EC(b,x)$, which are $E(C(b,\bar{x}(Eb)))$ and $\min_x C(E(b),x)$ and are computable. Note that these are also bounds on the expected value of the "Wait and See" problem, E min C(b,x) as well. The use of the above inequalities therefore yields a method of determining whether the use of the expected value of the random resource vector will result in a close approximation to the "Here and Now" problem. And, in fact, Madansky not only refines the above inequalities but also shows that if $C(b,x) = C_1(b,x) + C_2(b)$ then the "certainty" and "uncertainty" problems are equivalent with b = Eb.

In a later paper, (21) Madansky considered the solution of the problem

$$\text{minimize } c'x + f'y$$

$$\text{subject to } Ax + By \geq b$$

for random b. He approached the solution with three methods; the expected value method, the "fat method" and the "slack method". The expected value method was discussed above. In the "fat method", the values are chosen to be sufficiently pessimistic that $P(Ax \geq b) = 1$.

In the "slack method", $P(Ax \geq b) < 1$ is allowed, by choosing some y to minimize $cx + fy$, subject to $Ax + By = b$.

The vector f is a penalty placed on the "inaccuracies", that is, it is a utility function placed upon the violation of a constraint.

2.5 Tintner's approaches

Gerhard Tinter was writing papers (25) on risk programming long before such a name was recognized. One of these was a rather general exposition which developed a general theory of production where the objective was to maximize profits from production subject to uncertainty of prices and certain production transformation equations. Tintner proposed the use of Lagrange multipliers to solve these equations then assumed the use of charactistic functions to determine the distribution of the profit. Finally, he introduced a risk preferential function to determine an individuals' evaluation of profits and risk.

Although the approach was impractical it, nevertheless, contained some essential points which were missed by many later authors.

First, he was interested in the distribution of the profit. This was either overlooked or simply not noted in many papers twenty years later. Whether a stochastic resource vector or a cost vector is being considered, the value of the objective function does indeed have a distribution.

Secondly, he presented (perhaps not for the first time) the idea of a risk preferential function. This function is a measure of the subjectivity of the decision maker. This point, while carefully concealed in most statistical decisions by the use of "standard" critical points, was, instead, emphasized in Tintner's paper.

In 1960, Tintner (26) described the "passive" and "active" approaches to stochastic linear programming. The "passive" approach was to determine the distribution of the objective function and base decisions upon this distribution. This might be accomplished by solving the problem for all values of the random variables and then computing the distribution of the objective function. As Tintner pointed out, this would be rather laborious.

He suggested, instead, that the problem

$$\text{maximize } p = a'x$$

$$\text{subject to } Bx \leq c, \quad x \geq 0$$

where $P(a,B,c)$ is a probability distribution,

be transformed to

$$\text{maximize } p = a'x$$

$$\text{subject to } b_{ij}x_j = c_i u_{ij} \quad \text{all } i \text{ and } j$$

$$\text{where all } 0 \leq u_{ij} \leq 1 \; ; \; \sum_j^n u_{ij} = 1$$

and where each u_{ij} denotes the amount of resource c_i to be allocated to activity x_j. This allocation of resources to be used in the problem solution he called the "active" approach. A probability function $R(p,U)$ would be derived from $P(a,B,c)$ which would depend on the u_{ij}, which he states would be rather difficult to derive in practice. Tintner again introduced the preferential functional $f = f(R(p,U))$ where this function was to be maximized with respect to the elements of U. He pointed out that a special case is $E(p) = \int p\,dR(p,U)$, where the integral is to be taken over the whole range of p. Tintner suggested that dynamic programming might be used to solve this problem.

While this was an interesting approach, how the distribution of p is to be derived was not clear. If it is assumed that a and c are normally distributed then x will also be distributed normally. But since no assumptions can be made the independence of the elements of c under transformation, x cannot be assumed independent. Therefore, the vector product $a'x$ has an unknown distribution. The distribution of $a'x$ for independent a and independent x was noted by Miller (22) to be a modified Bessel function of the second kind. There appears to be very limited possibility of extending this result.

2.6 Charnes and Cooper's "certainty equivalents"

Charnes, in conjunction with others, wrote several papers (23) in the late Fifties and early Sixties on what they refered to as "chance constrained programming". The problem was one of

$$\text{maximizing } f(c,x)$$
$$\text{subject to } P(Ax \leq b) \geq \alpha$$
$$\text{for small values of the vector } \alpha .$$

The approach taken was based on the availability of large amounts of historical data from which the desired probability estimates could be computed. The approach then paralleled the "expected value" method.

In 1963, Charnes and Cooper published a paper (4) with the word "satisfice" in the title. Here again the idea of "satisfying" an individual is brought forth, emphasizing the subjective nature of the choice of the vector, α .

Assuming that some vector, α , could be chosen, Charnes and Cooper presented three models, the "E model", the "V model" and the "P model". These were the maximization of the expected value, the minimization of the variance, and the maximization of certain probability levels of the objective.

The "E model" was simply

$$\text{maximize } E(c'x)$$

$$\text{subject to } P(Ax \leq b) \geq \alpha$$

where Charnes, et al proposed the use of a decision rule matrix D, such that

$$x = Db$$

and D is to be determined. They then proceeded in the following manner:

$$E(c'x) = E(c'Db)$$

$$= (E(c))' \ D \ (E(b))$$

or, setting $\mu_c' = (E(c))'$ and $\mu_b' = (E(b))'$, then $E(c'x) = \mu_c'D \, \mu_b$. The objective function is now deterministic.

The constraints still involve a random vector.

$$\text{Let } \hat{b} = b - \mu_b$$

$$\text{and } a_i' = (a_{i1}, a_{i2}, \ldots, a_{in})$$

then assume that

$$(\hat{b}_i - a_i' D\hat{b}) \ : \ N(\mu_{bi} - a_i' \ D \, \mu_b, \ E\{\hat{b}_i - a_i' D\hat{b}\}^2)$$

thus,

$$P(a_i' D\hat{b}-\hat{b}_i \leq 0) = P(\hat{b}_i - a_i' D\hat{b} \geq 0)$$

$$= P(\hat{b}_i - a_i' D\hat{b} \geq -\mu_{bi} + a_i' D \mu_b)$$

$$= P\left(\frac{\hat{b}_i - a_i' D\hat{b}}{\sqrt{E\{\hat{b}_i - a_i' D\hat{b}\}^2}} \geq \frac{-\mu_{bi} + a_i' D\mu_b}{\sqrt{E\{\hat{b}_i - a_i' D\hat{b}\}^2}}\right)$$

Let z_i equal the first term above so that z_i : N(0,1). Then,

$$P\left(z_i \geq \frac{-\mu_{bi} + a_i' D\mu_b}{\sqrt{E\{\hat{b}_i - a_i' D\hat{b}\}^2}}\right) \geq \alpha_i$$

for the i^{th} constraint, or

$$F_i\left(\frac{(-\mu_{bi} + a_i' D\mu_b)}{\sqrt{E\{\hat{b}_i - a_i' D\hat{b}\}^2}}\right)$$

and hence

$$\frac{-\mu_{bi} + a_i' D\mu_b}{\sqrt{E\{\hat{b}_i - a_i' D\hat{b}\}^2}} \leq F_i^{-1}(\alpha_i) = -K_{\alpha_i}$$

where α_i will be assumed to be chosen such that K_{α_i} 0, all i.

These equations lead to

$$\mu_{bi} - a_i' D\mu_b \geq v_i \geq K_{\alpha_i} \sqrt{E\{\hat{b}_i - a_i' D\hat{b}\}^2} \geq 0$$

which leads to the deterministic convex programming problem:

minimize $-\mu_c' D \mu_b$

subject to $\mu_i(D) - v_i \geq 0$

$$\text{and } -K^2_{\alpha_i}\sigma^2_i(D) \quad K^2_{\alpha_i}\ \mu^2_i(D)+v^2_i \geq 0$$

$$\text{and } v_i \geq 0$$

$$\text{where } \sigma^2_i(D) \equiv E(a_i' \, Db-b_i)^2$$

$$\mu^2_i(D) \equiv (\mu_{bi}-a_i' \, D \, \mu_b)^2$$

Charnes and Cooper pointed out that this encompasses the results of Tintner, and this is easily seen if the elements u_{ij} are divided into the b_{ij} thus:

$$\frac{b_{ij}}{u_{ij}} \quad x_{ij} = c_i$$

where, now, the matrix D can be seen to be a function of the original coefficient matrix and of Tintner's "decision variables'.

To obtain the minimum variance about a chosen objective, Charnes and Cooper stated that it is only necessary to minimize $E(c'Db-z^0)^2$ subject to the above constraints; where z^0 is to be the chosen objective value.

The "P model" does not compare with any model presented by any other author and will not be discussed here.

2.7 Other formulations

Other authors have also treated the stochastic linear programming problem in recent years. Notable among these are Wets (30), Van Slyke and Wets (27), Walkup and Wets (29), Geoffrion (15) and Williams (31). The problem attacked by Wets, Van Slyke and Walkup was the original two stage problem of Dantzig. Their methods differ primarily in the use of measure theory in a manner that merely results in over complicating the problem.

Williams, on the other hand, approached the problem involving only a stochastic resource vector and in which "salvage" and "penalty" costs are used to place a premium on constraint violation, similar to the penalty vector of Madansky's. He then derived an approximation formula to the solution of this problem and showed that it satisfied Madansky's inequalities.

It is interesting to note that even though published in 1965 and 1966, none of these authors referenced any other papers on stochastic linear programming except their own and those of Dantzig's and Madansky's. An exception to this is Geoffrion who, writing in 1966, referenced Katoaka (18), Dantzig (5,6), Charnes and others, but did not reference Rein or Freund. Geoffrion described an "E model", a "P model and a "fractile model" of which the latter will bear closer examination. The former two are similar to those of Charnes and Cooper and need not be discussed.

The fractile model is an interesting problem, which is

$$\text{maximize } px$$
$$\text{subject to } Ax \leq b$$
$$\text{where } \quad p : N(\mu, V)$$

This problem was originally discussed by Kataoka (18), but Geoffrion's, who referenced Kataoka, will be cited here. Geoffrion stated that it can be shown that

$$F_\alpha (x) = \mu ' x + \Phi^{-1}(\alpha) \sqrt{x' V x}$$

where $F_\alpha (x)$ is the α fractile of px and is to be maximized. This function bears a great deal of resemblence to the objective function proposed by Freund. Indeed, Geoffrion noted that a solution to the problem can be obtained by the graph of the (E, σ) tradeoff curve which, although not referenced by Geoffrion, is the approach used by Rein (23) in his unpublished thesis. The comparison becomes even more complete by noting that since $\Phi^{-1}(\alpha)$ represents the inverse standard normal function, that values of α less than 0.5 result in negative $\Phi^{-1}(\alpha)$ and hence a relationship between the normal curve and Freund's value β, is established.

The observation that the value of $\Phi^{-1}(\alpha)$ represented the utility of the expected value relative to the variance of the objective function was also made by Kataoka, but independently of Freund. The primary result of his paper was an algorithm for solution of this problem utilizing a method of Wolfe's(32). Geoffrion's result was also an algorithm based upon a similar technique.

III TOWARD A UNIFIED APPROACH

After studying these sometimes fragmentary approaches and results and noting the lack of communication and coordination which results in the re-discovery of earlier results, the question of which approach and result is correct, becomes intriguing. Charnes and Cooper stated in 1959 (3) that:

> "The problem of stochastic (chance constrained) programming involves difficulties of an order incommensurate to that of "certainty" programming. These difficulties stem fundamentally from the probabilistic constraints, which experience (let alone theory) has made clear, are NOT adequately represented as some have done by applying the expectation operator to the stochastic form."

Yet, after making this statement, Charnes and Cooper continued to apply the expectation operator to the stochastic form(3,4).

Any optimization involving risk must also involve a subjective determination of what is the "best" risk. The notion of "best" risk may vary from one situation to another depending on the utility value associated with risk. The gambler will allow, even desire, large risk while a banker will allow almost none at all. Freund, introduced a "risk coefficient" to allow the control of risk in the stochastic linear programming problem. In statistics, the control of risk is allowed by the use of "critical points", a 99% point yields less risk than does the use of a 90% point. Therefore, when Kataoka and Geoffrion developed their objective function, they introduced the standard normal curve in such a fashion that it replaced the "risk coefficient" used by Freund. Thus, the amount of risk being taken could be quantitized in a statistical manner.

It would seem, then, that a complete synthesis of stochastic linear programming has been achieved by Kataoka and Geoffrion. Indeed, Kataoka has presented a formulation involving a stochastic resource vector and cost vector as well as an objective function in which risk has been quantitized. But, further analysis of Kataoka's formulation reveals that in order to show the solution vector as deterministic, he used a device, first proposed by Charnes and Cooper. This was the fact that solution space described by the set of linear equations with a deterministic resource vector is a convex cone. To allow for a stochastic resource vector, this convex cone was simply enlarged. Then, that portion of the solution space which would result in a constraint being violated with probability greater than α is disallowed, i.e., the solution space is restricted to "admissible" values of the solution vector. Therefore, using this device, the solution vector is now deterministic and hence its product with the stochastic cost vector will give the result shown by Kataoka.

The above can be shown to be a false picture by the following consideration. Suppose a value for the stochastic resource vector is chosen from the appropriate distribution and the (now deterministic) linear program is solved. The result would be a solution vector x. Repeat this process a large number of times and the values of these optimal x's would form their own distribution. Hence, it can be seen that they are indeed a function of the stochastic resource vector and are themselves stochastic. Suppose this were carried a step further so that for each solution vector computed, a value of the cost vector was observed. The value of the objective function would depend on the product of two stochastic vectors and, therefore, would itself be stochastic. This last distribution is the one assumed by Kataoka to be normal but, in fact, the objective function would be distributed according to a much more complex distribution, as observed earlier. This is true since not one, but both vectors in the objective function are stochastic.

The fact that the solution vector was itself stochastic was later recognized by Charnes and Cooper. Their derivation of a set of "certainty equivalent" constraints makes this fact clear. But in the derivation of an objective function, they again returned to that which they had earlier referred to as an "inadequate approach" and took expected values. Therefore, using an approach similar to that taken by Kataoka (18), but taking into consideration the stochastic nature of the solution vector, it would seem reasonable to combine Charnes and Cooper's "E model" and "V model" in the following manner.

Let

$$z = c'x$$

be the objective function, subject to

$$Ax \leq b$$

Then assume

$$b : N(\mu_b, \Sigma_b)$$

$$c : N(\mu_c, \Sigma_c)$$

and let the objective be to maximize

$$F_{\alpha}\ (c'x) = F_{\alpha}\ (z)$$

This leads in a manner similar to Kataoka's derivation to

$$\text{minimize} \ - \ \mu_c' D \ \mu_b + G^{-1}(\alpha)V(D)$$

where $G^{-1}(\alpha)$ is the standardized function and

$$V(D) = E(c'Db - \mu_c' D \ \mu_b)^2$$

subject to the same constraints derived by Charnes and Cooper. The fact that G(z) is unknown presents only a momentary problem.

First, observe that since G is standardized, the values of primary interest will be negative. This can also be arrived at by roughly the same reasoning covered by Freund. Further, though the nature of $G^{-1}(\alpha)$ is unknown, it can readily be assumed that practical values will not be infinite, that is, there will be reasonable values of $G^{-1}(\alpha)$ which will be finite. If this were not true, the distribution F(z) would not integrate to unity.

Next, substitute a parameter, $-\beta$ for $G^{-1}(\alpha)$. This parameter can then be used with a suitable algorithm for solving convex programming problems to compute a range of solutions (this range will be continuous) which can be used to partially determine the shape of the function F(z). It is possible, in theory at least, to develop an estimate of the α involved and chose a suitable value of β, from the set of β 's, for the desired α level.

IV CONCLUSION

Tracing the development of stochastic linear programming from its inception in 1955 to the present has resulted in an interesting array of problem formulations and solutions. First attempts at solution involved taking expected values and returning the problem to the familiar linear programming problem. This approach was valuable if for no other reason than it illustrated the fact that uncertainty costs the decision maker.

Then it seemed as though everyone in the field created his own private "uncertainty" problem and started trying to solve it. There were one stage, two stage, transportation, infinite horizon, here and now, wait and see, and passive and active problems. Notions of utility crept in and thought was given to the inviolability of constraints. Slowly, sometimes by independent paths, it became apparent that this was a statistical problem, there were distributions involved and perhaps statistical techniques should be used to solve it.

It is now clear that the results in this paper will obtain a solution to the truly stochastic linear programming problem, although the distribution of the objective function may never be analytically determined when both the resource and cost vectors are stochastic. This formulation is an amalgamation, a synthesis of the work of Freund, Rein, Kataoka, Charnes and Cooper and embodies and encompasses the two stage problem; Madansky's fat, slack, and expected value problem; Tintner's active approach; and, of course, Charnes and Cooper's uncertainty problem. Each of these cases can be obtained by proper choice of parameters. This, then, is indeed a unified approach.

A lot remains to be done. Methods for solving nonlinear programming problems are not complete. The nature of the distribution of the objective function should be explored for various distributions and combinations of the stochastic variables. Theory should be expanded to include the simultaneous consideration of stochastic matrix elements along with the other variables. Finally, everyone should agree on one formulation of the problem, hopefully, this one.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation to Mr. Whitney L. Johnson for his many useful suggestions which were invaluable in the development of this thesis.

The author also wishes to thank Dr. Conolly, Dr. Meyers, Dr. Jensen and Dr. Harshbarger for their interest and suggestions. Much appreciation is also due John Philpot for his useful comments.

Finally, the author wishes to thank his wife, Mary, for her patience.

BIBLIOGRAPHY

Literature Cited

1. Barbar, M. M. "Distributions of Solutions of a Set of Linear Equations". Journal of the American Statistical Association. Vol. 5 (1955) pp. 854 -869.

2. Charnes, A., Cooper, W. W., and G. H. Symonds. "Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil". Management Science. Vol. 4 (1958) pp. 235-263.

3. Charnes, A. and W. W. Cooper. "Chance Constrained Programming". Management Science. Vol. 6 (1959/60) pp. 73-79.

4. Charnes, A. and W. W. Cooper. "Deterministic Equivalents for Optimizing and Satisficing Under Chance Constraints". Operations Research. Vol. 11 (1963) pp. 18-39.

5. Dantzig, G. B. "Linear Programming Under Uncertainty". Management Science. Vol. 1 (1955) pp. 197-206.

6. Dantzig, G. B. Linear Programming and Extensions. RAND Report R-366-PR (1963) pp. 12-20.

7. Elmaghraby, S. E. A. Programming Under Uncertainty. Cornell University Press (1958) Xeroxed by University Microfilms Inc., Ann Arbor, Michigan (1964)

8. Elmaghraby, S. E. A. "An Approach to Linear Programming Under Uncertainty". Operations Research. Vol. 7 (1959) pp. 208-216.

9. Elmaghraby, S. E. A. "Allocation Under Uncertainty When the Demand has Continuous D. F.". Management Science. Vol. 6 (1960) pp. 270-294.

10. Ferguson, A. R. and Dantzig, G. B. "The Allocation of Aircraft to Routes: An Example of Linear Programming Under Uncertain Demand". Management Science. Vol. 3 (1956) pp. 45-73.

11. Fourier, Jean B. J. "Solution d'une Question Particuliere du Calcul des Inegalites". (1826) Original Not Seen.

12. Freund, R. J. "The Introduction of Risk into a Programming Model". Econometrica. Vol. 24 (1956) pp. 253-263.

14. Gauss, Karl F. Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, Supplementum, Vol. 4. Werke, Gottingen (1826) pp. 55-93. Original Not Seen.

15. Geoffrion, A. M. On Stochastic Linear Programming. Working Paper No. 103, Western Management Science Institute, University of California at Los Angeles. (1966).

16. Hildreth, C. "Point Estimates of Ordinates of Concave Functions". Journal of the American Statistical Association. Vol. 49 (1954) pp. 598-619.

17. Kantorovich, L. V. "Mathematical Methods in the Organization and Planning of Production". Management Science. Vol. 6, (1960) pp. 366-422.

18. Kataoka, S. "A Stochastic Programming Model". Econometrica. Vol. 31 (1963) pp. 181-196.

19. Luce, R. D. and H. Raiffa. Games and Decisions, Introduction and Critical Survey. John Wiley and Sons. New York (1957).

20. Madansky, A. "Inequalities for Stochastic Programming Problems". Management Science. Vol. 6 (1960) pp. 197-204.

21. Madansky, A. "Methods of Solution of Linear Programs Under Uncertainty". Operations Research. Vol. 10 (1962) pp. 463-471.

22. Miller, K. S. Multivariate Gaussian Distributions. SIAM Series in Statistics. John Wiley and Sons.

23. Rein, MacEason. Aspects of Risk Programming. Unpublished Masters Thesis. Virginia Polytechnic Institute (1959).

24. Saaty, T. L. "Coefficient Perturbation of a Constrained Extremum". Operations Research. Vol. 7 (1959) pp. 294-302.

25. Tintner, G. "The Pure Theory of Production under Technological Risk and Uncertainty". Econometrica. Vol. 9 (1941) pp. 305-312.

26. Tintner, G. "A Note on Stochastic Linear Programming". Econometrica. Vol. 28 (1960) pp. 490-495.

27. Van Slyke, R. and R. Wets. "Programming under Uncertainty and Stochastic Optimal Control". SIAM Journal on Control. Vol. 4 (1966) pp. 179-193.

28. Von Neumann, John. "A model of General Economic Equilibrium". Review of Economic Studies. Vol. 13 (1945) pp. 1-9.

29. Walkup, D. W. and R. Wets. Stochastic Programs with Recourse. Boeing Scientific Research Laboratory Document DI-82-0551. (1966).

30. Wets, R. "Programming Under Uncertainty: The Equivalent Convex problem". Journal SIAM on Applied Mathematics. Vol. 14 (1966) pp. 89-105.

31. Williams, A. C. "On Stochastic Linear Programming". Journal SIAM on Applied Mathematics. Vol. 13 (1965) pp. 927-940.

32. Wolfe, P. "The Simplex Method for Quadratic Programming". Econometrica. Vol. 27 (1959) pp. 382-398.

33. Wood, M. K. and M. A. Geisler. "Development of Dynamic Models for Program Planning". in T. C. Koopman's (ed) Activity Analysis of Production and Allocation. John Wiley and Sons (1951) pp. 189-192.

Literature Examined

1. Arrow, K. and L. Hurwicz. A Gradient Method for Approximating the Saddle Points. RAND Paper P-223 (1051).

2. Arrow, K. "Alternative Approaches to the Theory of Choice in Risk-taking Situations". Econometrica. Vol. 19 (1951) pp. 404-437.

3. Bellman, R. Dynamic Programming. Princeton University Press. (1957)

4. Charnes, A. and W. W. Cooper. "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints". Operations Research. Vol. 11 (1963) pp. 18-39.

5. Dantzig, G. B. "Application of the Simplex Method to the Transportation Problem". Chapter XXIII in T. C. Koopman's Activity Analysis of Production and Allocation. John Wiley and Sons, New York. (1951)

6. Derman, Cyrus. "On Sequential Decisions and Markov Chains". Management Science. Vol. 9 (1962/63) pp. 16-24.

7. Gale, D. Theory of Linear Economic Models. McGraw Hill. New York. (1960).

8. Holt, C. C., F. Modigliani and H. A. Simon. "Linear Decision Rule for Production and Employment Scheduling". Management Science. Vol. 2 (1955) pp. 1-30.

9. Kelley, J. E. "Parametric Programming and the Primal-Dual Algorithm". Operations Research. Vol. 7 (1959) pp. 327-334.

10. Mangasarian, O. L. "Nonlinear Programming Problems with Stochastic Objective Functions". Management Science. Vol. 10 (1963/64) pp. 353-359.

11. Milnow, J. "Games Against Nature" Chapter 4 in R. M. Thrall et. al. Decision Processes. John Wiley and Sons, New York.

12. Reiter, S. "Surrogates for Uncertain Decision Problems: Minimal Information for Decision Making". Econometrica. Vol. 25 (1957) pp. 339-345.

13. Rich, R. N. "A Statistical Basis for Approximation and Optimization". Annals of Mathematical Statistics. Vol. 37 (1966) pp. 59-65.

14. Sengupta, J. K., G. Tintner and C. Millham. "On Some Theorms of Stochastic Linear Programming with Applications". Management Science. Vol. 10 (1963) pp. 143-159.

15. Sengupta, J. K. and G. Tintner. "On the Stability of Solutions under Error in Stochastic Linear Programming". Metrica. Vol. 9. (1965) pp. 47-60.

16. Simon, H. A. "Dynamic Programming under Uncertainty with a Quadratic Criterion Function". Econometrica. Vol. 24 (1956) pp. 74-81.

17. Stigler, G. J. "The Development of Utility Theory". Journal of Political Economics. Vol. 58 (1950) pp. 307-327, 373-396.

18. Szwarc, Wlodzimierz. "The Transportation Problem with Stochastic Demand". Management Science. Vol. 11 (1964/65) pp. 33-50.

19. Theil, H. "A Note on Certainty Equivalence in Dynamic Planning". Econometrica. Vol. 25 (1957) pp. 346-349.

20. Tintner, G. "The Theory of Choice under Subjective Risk and Uncertainty". Econometrica. Vol. 9 (1941) pp. 298-304.

21. Weiss, Lionel. Statistical Decision Theory. McGraw Hill, New York. (1961)

22. Wolfe, Phillip and Dantzig, G. B. "Linear Programming in a Markov Chain". Operations Research. Vol. 10 (1962) pp. 463-471.

The vita has been removed from the scanned document

ABSTRACT

A Synthesis of Stochastic Linear Programming

by

Charles Samuel Matheny

With the growing application of statistics to many areas of science and engineering it was inevitable that the indeterministic nature of linear programming would be recognized. The first paper in the area of stochastic linear programming was written by one of the pioneers of linear programming, G. B. Dantzig. This paper led to increased interest in the field and was followed by the works of several prominent authors in linear programming.

At this point, there was considerable divergence in the problem formulations and in methods of solution. These various formulations are discussed and compared, illustrating the lack of communication among the authors.

Then, a problem formulation is developed which incorporates many of the formulations of the other authors and recognizes completely the stochastic nature of the problem elements.