CHANCE-CONSTRAINED PROGRAMMING
WITH 0-1 OR BOUNDED DECISION VARIABLES

by

Frederick S. Hillier

TECHNICAL REPORT NO. 92

August 26, 1966

Supported by the Army, Navy, Air Force, and NASA under
Contract Nonr-225(53)(NR-042-002)
with the Office of Naval Research*

Gerald J. Lieberman, Project Director

*This research was partially supported by the Office of Naval Research under Contract Nonr-225(89)(NR-047-061).

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

CHANCE-CONSTRAINED PROGRAMMING
WITH 0-1 OR BOUNDED DECISION VARIABLES

by

Frederick S. Hillier

1. Introduction

To introduce chance-constrained programming, consider the linear programming model,

$$\max x_0 = cx ,$$

subject to
$$Ax \leq b ,$$
$$x \geq 0 ,$$

where c and x are n-vectors, b is an m-vector, and A is an $m \times n$ matrix. Now suppose that some or all of the elements of A, b, and c are random variables rather than constants. Several approaches to this problem of linear programming under risk have been developed.[1] One of these, called "stochastic linear programming," is primarily concerned with the probability distribution of $\max x_0$.[2] A second approach, originally called "linear programming under uncertainty," deals with a special case of the problem by reducing it to an ordinary linear programming problem.[3] The "chance-constrained programming" approach reformulates the

[1] See Näslund [27] for a survey and comprehensive set of references.

[2] See Tintner [34] for the original presentation of this approach.

[3] This special case is one in which each random variable has only a finite number of possible values, and the particular value it actually takes on will become known before certain of the decision variables must be assigned values. See Dantzig [11] for the original presentation of this approach.

problem as:

$$\text{optimize} \quad f(c, x) \ ,$$

subject to $$\text{Prob}\,\{Ax \leq b\} \geq \alpha \ ,$$

$$x \geq 0 \ ,$$

where α is an m-dimensional column vector whose elements lie between 0 and 1. Thus, a nonnegative solution x is feasible if and only if

$$\text{Prob}\left\{\sum_{j=1}^{n} a_{ij}x_j \leq b_i\right\} \geq \alpha_i \quad \text{for } i = 1, \ldots, m \ ,$$

so that the complementary probability, $1 - \alpha_i$, represents the allowable risk that the random variables will take on values such that $\sum_{j=1}^{n} a_{ij}x_j > b_i$. If $a_{i1}, \ldots, a_{in}, b_i$ are all constants rather than random variables for a particular value of i, then α_i becomes irrelevant and the i^{th} constraint can remain in the form, $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$. The objective function $f(c, x)$ often is taken to be the mathematical expectation of cx, $E(cx) = \sum_{j=1}^{n} E(c_j)x_j$, although other criteria also may be used.[4] If certain of the random variables will be observed before certain elements of x must be specified, the problem may be formulated in terms of choosing a decision rule, $x = \psi\,(A, b, c)$, instead of specifying all elements of x directly. In this case, the function ψ normally would be restricted to a specified class of functions (e.g., the class of linear functions) but the parameters of ψ may be decision variables.

[4] See Charnes and Cooper [5] for an analysis of alternative criteria.

Chance-constrained programming was formulated originally by Charnes, Cooper, and Symonds [7] and Charnes and Cooper [4], and has since been further developed and applied by Charnes and Cooper [5, 6], Charnes, Cooper and Thompson [8, 9], Ben-Israel [3], Kataoka [21], Kirby [23], Naslund [26], Naslund and Whinston [28], Sinhal [31], Thiel [32], Van De Panne and Popp [35], and Miller and Wagner [25]. The point of departure of this paper from this previous work is three-fold. First, several linear inequalities will be introduced that permit the approximate solution and analysis of chance-constrained programming problems with either zero-order or linear decision rules as ordinary linear programming problems. Second, the case where some or all of the elements of x are 0-1 (yes-no) variables rather than continuous variables also is considered, and both exact and approximate solution procedures are presented. Third, since linear decision rules are not meaningful with 0-1 variables, another method of making "second-stage decisions" is developed for this case.

The original motivation for this work came from an earlier paper by the author [19], which was the award winner in the TIMS-ONR Program on "Capital Budgeting of Interrelated Projects." In Chapter 6 of this paper, a capital budgeting problem under risk was formulated as a chance-constrained programming problem with 0-1 decision variables, and a simple linear inequality was introduced that permitted its reduction to an ordinary linear programming problem. It then became evident that this approach could be greatly extended in a more general context, which is done here.

2. Formulation

It is assumed here that the decision variables are either continuous variables with known bounds or discrete variables restricted to two values (taken to be 0 or 1)[5] as when a yes-or-no decision must be made. It may be assumed without loss of generality that the bounded continuous variables lie between 0 and 1, since this can always be effected by the appropriate change of scale and translation of the coefficients of the respective variables. For concreteness, it is assumed that the objective function is $E(cx)$.[6] Therefore, the chance-constrained programming model to be considered here is

$$\max E(cx) = \sum_{j=1}^{n} E(c_j)x_j \; ,$$

subject to

$$\text{Prob}\,[Ax \leq b] \geq \alpha \; ,$$

$$0 \leq x_j \leq 1 \quad \text{for} \quad j \in C \; ,$$

$$x_j = 0 \text{ or } 1 \text{ for } \quad j \in D \; ,$$

where $C \cap D = \varphi$ and $C \cup D = \{1, \ldots , n\}$.

---

[5] As is well-known, a general integer variable restricted to the values, 0, 1, ... , N, can also be reduced to this case by replacing the variable by $\sum_{k=1}^{N} y_k$, where the y_k are 0-1 variables.

[6] However, certain other objective functions also could be handled within the framework of the following analysis. One suggested by Kataoka [21] is: maximize y, subject to Prob $\{cx \leq y\} = \beta$. This constraint can be rewritten in the standard form as Prob $\{-cx + y \leq 0\} \geq 1 - \beta$ without altering the resulting optimal solution (provided that c has a continuous probability distribution). Another such objective function is: minimize $\text{Var}(cx)$, which can be replaced by: maximize y, subject to $y + \sqrt{\text{Var}(cx)} \leq 0$. It will be seen subsequently that this constraint is in an acceptable form.

Each of the elements of A, b and c is permitted to be either a constant or a random variable, and the random variables are permitted to be statistically dependent.[7] However, it is assumed that the joint probability distribution of the random variables is not disturbed by the choice of x. For the moment, a zero-order decision rule is assumed, so that x is chosen without observing any of the random variables. However, other decision rules will be considered in the concluding sections.

The first step in solving this chance-constrained programming problem is to reduce it to a deterministic equivalent form. Consider a typical constraint,

$$\text{Prob}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i \leq 0\right\} \geq \alpha_i .$$

Assume that the expected values and covariance matrix of $a_{i1}, \ldots,$ a_{in}, b_i are known. Denote them by $E(a_{i1}), \ldots, E(a_{in}), E(b_i),$ and by V_i, respectively. Further assume that the functional form of the probability distribution of $\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right)$ is known, and that the fractiles of this distribution are completely determined by its mean and variance. For example, if $a_{i1}, \ldots, a_{in}, b_i$ have a multivariate normal distribution then $\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right)$ has a normal distribution for any x. If $C = \varphi$, then $\sum_{j=1}^{n} a_{ij}x_j$ has a chi-square distribution or a Poisson distribution if $a_{i1}, \ldots, a_{in}$ have independent chi-square

[7]However, if the random variables in different constraints are strongly dependent, so that the probabilities of satisfying the respective inequality constraints are strongly dependent, then another formulation imposing a lower bound on a single probability that all of the inequality constraints are satisfied simultaneously may be more suitable, albeit less tractable. Miller and Wagner [25] have analyzed such a formulation for a special case.

distributions or Poisson distributions, respectively. If the individual random variables have arbitrary distributions, then $\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right)$ may still be approximately normal by some version of the Central Limit Theorem, which holds under fairly weak conditions for independent random variables and under rather strong conditions for dependent random variables. A survey of the various sets of conditions under which the Central Limit Theorem holds is given by this author elsewhere [19, Sect. 4.2]. Whatever the distribution of $\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right)$ happens to be, let $F(\cdot)$ denote the cumulative distribution function of

$$\frac{\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right) - E\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}}{\sqrt{\operatorname{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}}} .$$

Given β, $0 \leq \beta \leq 1$, define K_β by the relationship,

$$F(K_\beta) = \beta .$$

Thus, by proceeding in the usual way,[8] the deterministic equivalent form of the constraint becomes

$$E\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\} + K_{\alpha_i}\sqrt{\operatorname{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} \leq 0 ,$$

which reduces to

[8] For example, see Cooper and Charnes [6] or Kataoka [21]. Also see Hillier and Lieberman [20] for a detailed expository treatment.

$$\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i}\sqrt{[x^t, -1]V_i \begin{bmatrix} x \\ -1 \end{bmatrix}} \leq E(b_i) \; .$$ [9]

The problem now is to reduce this deterministic equivalent form further to a more tractable form. The objective will be to linearize the constraints so that linear programming and integer linear programming algorithms can be used. The basic approach is suggested by the following obvious result.

Fundamental Lemma: Assume that $g_1(x) \leq g_2(x) \leq g_3(x)$ for all admissible x. Consider a solution x which is feasible if and only if $g_2(x) \leq k$, (i.e., x satisfies all other conditions for feasibility).

(i) If $g_3(x) \leq k$, then x is feasible.

(ii) If x is feasible, then $g_1(x) \leq k$.

Thus, $g_2(x) \leq k$ will represent some exact deterministic equivalent form of the constraint, whereas $g_3(x) \leq k$ and $g_1(x) \leq k$ will represent linear constraints that are uniformly tighter and uniformly looser, respectively. These linear approximations will be introduced in the next section, and procedures for obtaining both exact and approximately optimal solutions (initially with zero-order decision rules and

---

[9] If the functional form of the distribution $\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right)$ is not known, so K_{α_i} is not known, then the one-sided Chebyshev Inequality yields $\sqrt{\frac{\alpha_i}{1-\alpha_i}}$ as an upper bound on K_{α_i}. Hence, this bound may be used here when K_{α_i} is not known in order to guarantee that $\text{Prob}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\} \geq \alpha_i$. (However, it should be noted that the bound is based on the worst possible distribution and therefore will usually overestimate K_α greatly, so that it would tend to yield constraints that are considerably tighter than necessary.)

then with other decision rules) will then be described in the succeeding sections.

3. Useful Inequalities

Consider again the typical constraint, $\text{Prob}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i \leq 0\right\} \geq \alpha_i$, and its deterministic equivalent form given in the preceding section. Assume initially that $a_{i1}, \ldots, a_{in}, b_i$ are mutually independent, so that

$$\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\} = \sum_{j=1}^{n} \text{Var}(a_{ij})x_j^2 + \text{Var}(b_i).$$

Define $\sigma_i^2 = \sum_{j=1}^{n} \text{Var}(a_{ij}) + \text{Var}(b_i)$, (for $i = 1, \ldots, m$), and

$\sigma_{ij}^2 = \text{Var}(a_{ij})$, (for $i = 1, \ldots, m$; $j = 1, \ldots, n$).

<u>Theorem 1</u>: Assume that $0 \leq x_j \leq 1$ for $j \in C$ and $x_j = 0$ or 1 for $j \in D$, and that $a_{i1}, \ldots, a_{in}, b_i$ are mutually independent. Then

$$\text{(i)} \quad \sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} \leq \sum_{j\in C}\sqrt{\sigma_1^2 - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2} + \sum_{j\in D} \sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2}\, x_j$$

$$+ \sum_{j\in D}\sqrt{\sigma_i^2 - \sigma_{ij}^2} - (n-1)\sigma_i .$$

$$\stackrel{\text{def}^n}{=\!=} R_i(x)$$

$$\text{(ii)} \quad \sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} = R_i(x) \qquad \text{if } \sum_{j=1}^{n} x_j = n \text{ or if}$$

$$\sum_{\substack{j=1\\j\neq k}}^{n} x_j = n - 1 \qquad \text{for any } k = 1, \ldots, n.$$

(iii) $R_i(x) \leq h(x)$ for any function $h(x)$ of the form,

$$h(x) = \sum_{j \in C} f_j(x_j) + \sum_{j \in D} d_j x_j + a_0,$$

such that

$$\sqrt{\operatorname{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} \leq h(x)$$

for all admissible x and

$$\sqrt{\operatorname{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} = h(x)$$

if $\sum_{j=1}^{n} x_j = n$.

Proof: To verify Part (i), notice that

$$\sqrt{\operatorname{Var} \sum_{j=1}^{n} a_{ij}x_j - b_i} = \sqrt{\sum_{j=1}^{n} \sigma_{ij}^2 x_j^2 + \operatorname{Var}(b_i)}$$

$$\equiv \sigma_i - \sum_{j=1}^{n} \left[\sqrt{\sigma_i^2 - \sum_{k=1}^{j-1} \sigma_{ik}^2(1 - x_k^2)} - \sqrt{\sigma_i^2 - \sum_{k=1}^{j} \sigma_{ik}^2(1 - x_k^2)}\right]$$

$$\leq \sigma_i - \sum_{j=1}^{n} \left[\sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2(1 - x_j^2)}\right], \text{ (since } \sqrt{y} \text{ is concave)},$$

$$= -(n-1)\sigma_i + \sum_{j \in C} \sqrt{\sigma_i^2 - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2} + \sum_{j \in D} \left[x_j\sigma_i + (1 - x_j)\sqrt{\sigma_i^2 - \sigma_{ij}^2}\right],$$

(since $x_j = 0$ or 1 for $j \in D$),

which is the desired result (after recombining terms).

Part (ii) is evident by inspection.

To verify Part (iii), notice that $R_i(x) = h(x) = \sigma_i$ when $\sum_{j=1}^{n} x_j = n$, and that $h(x)$ is a sum of separable functions of the individual variables. Therefore, it follows from Part (ii) that

$$f_j(1) - f_j(x_j) \le \sqrt{\sigma_i^2 - \sigma_{ij}^2 + \sigma_{ij}^2(1)} - \sqrt{\sigma_i^2\sigma_{ij}^2 + \sigma_{ij}^2 x_j^2} \quad \text{for} \quad 0 \le x_j \le 1 \quad \text{and} \quad j \epsilon C ,$$

$$d_j \le \sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2} , \quad \text{for} \quad j \epsilon D ,$$

so that

$$\sigma_i - h(x) \le \sigma_i - R_i(x)$$

for all admissible x, so that

$$h(x) \ge R_i(x) .$$

This completes the proof of Theorem 1.

<u>Corollary 1 to Theorem 1:</u> In addition to the assumptions of Theorem 1, assume that $C = \Phi$ and $\sum_{j=1}^{n} x_j < n_1 < n - 1$. Let J_1 be a subset of $\{1, \ldots , n\}$ containing exactly $(n - n_1)$ elements such that

$$J_1 = \{j | \sigma_{ij}^2 < \sigma_{ik}^2 \quad \text{for all} \quad k \notin J_1\} ,$$

and define

$$a_i = \sigma_i - \sum_{j \epsilon J_1} \left[\sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2} \right] - \sqrt{\sigma_i^2 - \sum_{j \epsilon J_1} \sigma_{ij}^2} .$$

(i) $$\sqrt{\mathrm{Var}\left\{ \sum_{j=1}^{n} a_{ij}x_j - b_i \right\}} \le R_i(x) - a_i .$$

(ii) $$\sqrt{\mathrm{Var}\left\{ \sum_{j=1}^{n} a_{ij}x_j - b_i \right\}} = R_i(x) - a_i \quad \text{if} \quad \sum_{j=1}^{n} x_j = n_1 \quad \text{and}$$

$$\sum_{j \epsilon J_1} x_j = 0 .$$

Proof: Under these assumptions, the two functions reduce to

$$\sqrt{\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} = \sqrt{\sigma_i^2 - \sum_{j=1}^{n} (1 - x_j)\sigma_{ij}^2} ,$$

$$R_i(x) = \sigma_i - \sum_{j=1}^{n} (1 - x_j)\left[\sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2}\right] .$$

Let $$S = \left\{x \middle| x_j = 0 \text{ or } 1, \ \sum_{j=1}^{n} x_j \leq n_1\right\} ,$$

$$g(x) = R_i(x) - \sqrt{\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} ,$$

$$d = \min_{x \in S} g(x) ,$$

so that

$$\sqrt{\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x - b_i\right\}} \leq R_i(x) - d .$$

Since $\sqrt{y}$ is a concave function, it is evident that $g(x)$ must be minimized by some solution x such that $\sum_{j=1}^{n} x_j = n_1$.
Therefore, it only remains to show that, among the set of solutions such that $\sum_{j=1}^{n} x_j = n_1$, $g(x)$ is minimized by $x^{(1)}$, where $x_j^{(1)} = \begin{cases} 0, & \text{if } j \in J_1 \\ 1, & \text{if } j \notin J_1 \end{cases}$, since this will show that $a_i = d$.

Let J_2 be an arbitrary subset of $\{1, \ldots , n\}$ that contains exactly $(n - n_1)$ elements, and let $x^{(2)}$ be the solution such that

$$x_j^{(2)} = \begin{cases} 0, & \text{if } j \in J_2 \\ 1, & \text{if } j \notin J_2 \end{cases} .$$

Let $J_3 = J_1 - J_1 \cap J_2$ and $J_4 = J_2 - J_1 \cap J_2$, let n_3 be the number of elements in J_3 (or J_4), and let $h_3(1), \ldots , h_3(n_3)$ and $h_4(1), \ldots , h_4(n_4)$ be the elements of J_3 and J_4, respectively.

Finally, define

$$j' = \begin{cases} j, & \text{if } \ j \notin J_3 \quad . \\ h_4(k), & \text{if } \ j = h_3(k) \ \text{ for some } \ k = 1, \ldots , n_3 \ . \end{cases}$$

Therefore,

$$\begin{aligned}
& g(x^{(2)}) - g(x^{(1)}) \\
&= \left\{ \sigma_i - \sum_{j=1}^{n} (1-x_j^{(2)}) \left(\sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2} \right) - \sqrt{\sigma_i^2 - \sum_{j=1}^{n} (1-x_j^{(2)}) \sigma_{ij}^2} \right\} \\
&\quad - \left\{ \sigma_i - \sum_{j=1}^{n} (1-x_j^{(1)}) \left(\sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2} \right) - \sqrt{\sigma_i^2 - \sum_{j=1}^{n} (1-x_j^{(1)}) \sigma_{ij}^2} \right\} \\
&= \sqrt{\sigma_i^2 - \sum_{j \epsilon J_1} \sigma_{ij}^2} - \sqrt{\sigma_i^2 - \sum_{j \epsilon J_1} \sigma_{ij}^2 - \sum_{k=1}^{n_3} (\sigma^2_{i,h_4(k)} - \sigma^2_{i,h_3(k)})} \\
&\quad - \sum_{j \epsilon J_3} \left[\sqrt{\sigma_1^2 - \sigma_{ij}^2} - \sqrt{\sigma_i^2 - \sigma_{ij}^2 - (\sigma_{ij'}^2 - \sigma_{ij}^2)} \right] \\
&= \sum_{j=1}^{n} \left\{ \begin{array}{l} \sqrt{\sigma_i^2 - \sum_{j \epsilon J_1} \sigma_{ij}^2 - \sum_{k=1}^{j-1} (\sigma_{ij'}^2 - \sigma_{ij}^2)} - \sqrt{\sigma_i^2 - \sum_{j \epsilon J_1} \sigma_{ij}^2 - \sum_{k=1}^{j} (\sigma_{ij'}^2 - \sigma_{ij}^2)} \\ - \left[\sqrt{\sigma_i^2 - \sigma_{ij}^2} - \sqrt{\sigma_{ij}^2 - \sigma_{ij}^2 - (\sigma_{ij'}^2 - \sigma_{ij}^2)} \right] \end{array} \right\}
\end{aligned}$$

≥ 0 ,

since $\sqrt{y}$ is a concave function so that each term in the summation is nonnegative. This completes the proof.

<u>Corollary 2 to Theorem 1</u>: Under the assumptions of Theorem 1, the following statements hold.

(i) Any solution x that satisfies the set of constraints,

$$\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i} R_i(x) \leq E(b_i), \quad \text{for } i = 1, \ldots, m,$$

$$0 \leq x_j \leq 1, \quad \text{for } j \in C,$$

$$x_j = 0 \text{ or } 1, \quad \text{for } j \in D,$$

necessarily is a feasible solution.

(ii) If the additional assumptions of Corollary 1 hold, then (i) will still hold after replacing $R_i(x)$ by $[R_i(x) - a_i]$ for $i = 1, \ldots, m$.

(iii) If, for each $i = 1, \ldots, m$, $K_{\alpha_i} \geq 0$ and each nonlinear term $\sqrt{\sigma_i^2 - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2}$ in $R_i(x)$ is approximated by a piecewise-linear function that coincides with $\sqrt{\sigma_i^2 - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2}$ only at $x_j = 0$, $x_j = 1$, and at the points where the slope of the piecewise-linear function changes, then both (i) and (ii) will still hold. Furthermore, each of these piecewise-linear functions necessarily is convex.

(iv) Any feasible solution x such that $\sum_{j=1}^{n} x_j = n$ or $\sum_{\substack{j=1 \\ j \neq k}}^{n} x_j = n - 1$ for any $k = 1, \ldots, n$ necessarily satisfies the set of constraints in (i). Furthermore, if $x_k = 0$ also, then this statement still must hold after introducing the piecewise-linear functions described in (iii).

Proof: Given the Fundamental Lemma, all of these statements are an immediate consequence of Theorem 1. The convexity of the piecewise-linear functions described in (iii) is demonstrated simply by noting that

$$\frac{d^2}{dx_j}\left\{-\sqrt{\sigma_i^2 - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2}\right\} = \frac{(\sigma_i^2 - \sigma_{ij}^2)\sigma_{ij}^2}{\left(\sigma_i^2 - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2\right)^{3/2}} \geq 0 .$$

To gain some insight into the goodness of the approximation introduced by Theorem 1, consider as an example a case where $n = 5$, $E(a_{ij}) = 10$ and $\text{Var}\ (a_{ij}) = 10$ for all $j = 1, \ldots, 5$, $E(B_i) = 50$ and $\text{Var}\ (b_i) = 50$, $K_{\alpha_i} = 2$, and $C = \Phi$. For this case, the following numerical results are obtained.

$\sum_{j=1}^{n} x_j$	$\sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}}$	$R_i(x)$	$\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i}\sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}}$	$\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i} R_i(x)$
5	10	10	70	70
4	9.487	9.487	58.974	58.974
3	8.944	8.974	47.888	47.948
2	8.367	8.461	36.734	36.922
1	7.746	7.948	25.492	25.896
0	7.071	7.435	14.142	14.870

Thus, the approximation introduced by Theorem 1 is excellent here for the larger values of $\sum_{j=1}^{n} x_j$, which is where accuracy tends to be important.

Whereas the above results provide uniformly tighter constraints, Theorem 2 below will provide uniformly looser constraints.

<u>Theorem 2</u>: Assume that $0 \leq x_{j'} \leq 1$ for $j \in C$ and $x_j = 0$ or 1 for $j \in D$, and that $a_{(i)} \ldots a_{in}$, b_i are mutually independent. Then

(i) there exists a unique real constant v_i,

$$\text{Var}\ (b_i) + \max_{j \in \{1,\ldots,n\}} \left\{\sigma_{ij}^2\right\} \leq v_i \leq \sigma_i^2 , \text{ such that}$$

$$\sqrt{\text{Var}\ (b_i)} + \sum_{j=1}^{n}\left[\sqrt{v_i} - \sqrt{v_i - \sigma_{ij}^2}\right] = \sigma_i .$$

(ii) if $y \geq v_i$ (i.e., if $\sqrt{\text{Var}(b_i)} + \sum_{j=1}^{n}\left[\sqrt{y} - \sqrt{y - \sigma_{ij}^2}\right] \leq \sigma_i$),

then

$$h(x, y) \underline{\underline{\text{def}}} \sum_{j\in C}\left[\sqrt{y-\sigma_{ij}^2+\sigma_{ij}^2 x_j^2} - \sqrt{y-\sigma_{ij}^2}\right] + \sum_{j\in D}\left[\sqrt{y} - \sqrt{y-\sigma_{ij}^2}\right] x_j + \sqrt{\text{Var}(b_i)}$$
$$\leq \sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} .$$

(iii) $h(x, y) = \sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}}$ if $\sum_{j=1}^{n} x_j = 0$ or if

$y = v_i$ and $\sum_{j=1}^{n} x_j = n$.

Proof: To prove (i), let

$$g(y) = \sqrt{\text{Var}(b_i)} + \sum_{j=1}^{n}\left[\sqrt{y} - \sqrt{y-\sigma_{ij}^2}\right] \quad \text{for } y \geq \max_{j\in\{1,\ldots,n\}}\left\{\sigma_{ij}^2\right\} ,$$

and note that $g(y)$ must be a strictly monotone decreasing continuous function since $\sqrt{y}$ is a strictly concave function. Therefore, by appealing to the Intermediate Value Theorem, it is only necessary to show that

$$g(\text{Var}(b_i) + \max_{j\in\{1,\ldots,n\}}\left\{\sigma_{ij}^2\right\}) \geq \sigma_i \geq g(\sigma_i^2) .$$

However, this becomes evident by assuming (without loss of generality) that $\sigma_{i1}^2 = \max_{j\in\{1,\ldots,n\}}\left\{\sigma_{ij}^2\right\}$ and then expressing σ_i as

$$\sigma_i = \sqrt{\text{Var}(b_i)} + \sum_{j=1}^{n}\left[\sqrt{\text{Var}(b_i) + \sum_{k=1}^{j}\sigma_{ik}^2} - \sqrt{\text{Var}(b_i) + \sum_{k=1}^{j}\sigma_{ik}^2 - \sigma_{ij}^2}\right],$$

since

$$\sqrt{\mathrm{Var}\ (b_i)+\sigma_{i1}^2} - \sqrt{\mathrm{Var}\ (b_i)+\sigma_{i1}^2 - \sigma_{ij}^2}$$

$$\geq \sqrt{\mathrm{Var}\ (b_i)+ \sum_{k=1}^{j} \sigma_{ij}^2} - \sqrt{\mathrm{Var}\ (b_i)+ \sum_{k=1}^{j} \sigma_{ik}^2 - \sigma_{ij}^2}$$

$$\geq \sigma_i - \sqrt{\sigma_i^2 - \sigma_{ij}^2} \ .$$

To prove (ii), note that, for fixed x, $h(x, y)$ is a strictly monotone decreasing function of y. Thus, it is sufficient to prove (ii) for $y = v_i$. It will now be assumed that $\sigma_{in}^2 = \max_{j\in\{1,\ldots,n\}} \left\{\sigma_{ij}^2\right\}$. First consider the case where

$$v_i \geq \max_{j\in\{1,\ldots,n\}} \left\{\mathrm{Var}\ (b_i) + \sum_{k=1}^{j-1} \sigma_{ik}^2 x_k^2 + \sigma_{ij}^2\right\}$$

$$= \mathrm{Var}\ (b_i) + \sum_{k=1}^{n-1} \sigma_{ik}^2 x_k^2 + \sigma_{in}^2 \ .$$

Then

$$h(x, v_i) \leq \sqrt{\mathrm{Var}(b_i)} + \sum_{j=1}^{n} \left[\sqrt{\mathrm{Var}(b_i)+ \sum_{k=1}^{j} \sigma_{ik}^2 x_k^2} - \sqrt{\mathrm{Var}(b_i) - \sum_{k=1}^{j-1} \sigma_{ik}^2 x_k^2}\right]$$

$$= \sqrt{\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} \ .$$

Now consider the complementary case where

$$v_i \leq \min_{j\in\{1,\ldots,n\}} \left\{\mathrm{Var}\ (b_i) + \sum_{j=1}^{n} \sigma_{ij}^2 - \sum_{k=1}^{j-1} \sigma_{ik}^2 (1-x_k^2)\right\}$$

$$= \min_{j\in\{1,\ldots,n\}} \left\{\mathrm{Var}\ (b_i) + \sum_{k=1}^{j-1} \sigma_{ik}^2 x_k^2 + \sum_{k=j}^{n} \sigma_{ik}^2\right\}$$

$$= \mathrm{Var}\ (b_i) + \sum_{k=1}^{n-1} \sigma_{ik}^2 x_k^2 + \sigma_{in}^2 \ .$$

Then

$$h(x,v_i) = \sigma_i - \sum_{j\in C}\left[\sqrt{v_i} - \sqrt{v_i-\sigma_{ij}^2(1-x_j^2)}\right] - \sum_{j\in D}\left[\sqrt{v_i} - \sqrt{v_i-\sigma_{ij}^2}\right](1-x_j)$$

$$\leq \sigma_i - \sum_{j=1}^{n}\left[\sqrt{\mathrm{Var}(b_i)+\sum_{k=1}^{n}\sigma_{ik}^2-\sum_{k=1}^{j-1}\sigma_{ik}^2(1-x_k^2)} - \sqrt{\mathrm{Var}(b_i)+\sum_{k=1}^{n}\sigma_{ik}^2-\sum_{k=1}^{j}\sigma_{ik}^2(1-x_k^2)}\right]$$

$$= \sqrt{\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j-b_i\right\}} \ ,$$

which completes the proof of (ii).

Part (iii) is evident by inspection.

<u>Corollary to Theorem 2</u>: Under the assumptions of Theorem 2, the following statements hold.

(i) If $u_i \leq v_i$ for all $i = 1, \ldots, m$, then any feasible solution x necessarily satisfies the set of constraints,

$$\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i} h(x, u_i) \leq E(b_i) \ , \quad \text{for} \quad i = 1, \ldots, m \ ,$$

$$0 \leq x_j \leq 1 \ , \quad \text{for} \quad j \in C \ ,$$

$$x_j = 0 \quad \text{or} \quad 1 \ , \quad \text{for} \quad j \in D \ .$$

(ii) $h(x,\sigma_i^2) = R_i(x) + \sqrt{\mathrm{Var}(b_i)} + (n-1)\sigma_i - \sum_{j=1}^{n}\sqrt{\sigma_i^2-\sigma_{ij}^2}$, for $i = 1,\ldots,m$.

(iii) If, for each $i = 1, \ldots, m$, $K_{\alpha_i} \geq 0$ and each nonlinear term $\sqrt{u_i - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2}$ in $h(x, u_i)$ is approximated by a convex piecewise-linear function that never exceeds $\sqrt{u_i - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2}$, then (i) will still hold.

(iv) Any solution x that satisfies the set of constraints in (i) necessarily is a feasible solution if $\sum_{j=1}^{n} x_j = 0$, or if $u_i = v_i$ (for $i = 1, \ldots, m$) and $\sum_{j=1}^{n} x_j = n$.

Although no explicit solution for the v_i has been given, they can be readily determined (within a specified error) by standard numerical methods since $h(x, y)$ is a strictly monotone decreasing function of y.

To illustrate the application of Theorem 2, consider again the numerical example introduced following Theorem 1. For this case, $\sqrt{v_i} - \sqrt{v_i - \sigma_{ij}^2} = 0.5858$ (so $v_i = 77.94$), which yields the following results.

$\sum_{j=1}^{n} x_j$	$h(x, v_i)$	$\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i} h(x, v_i)$
5	10	70
4	9.414	58.828
3	8.828	47.656
2	8.243	36.486
1	7.657	25.314
0	7.071	14.142

Comparison with the values of $\sqrt{\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}}$ and $\left[\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i}\sqrt{\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}}\right]$ given earlier indicate that this is an excellent "uniformly looser" linear approximation.

Now consider the case where $a_{ij}, \ldots, a_{in}, b_i$ are not mutually independent, so that

$$\mathrm{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\} = \left[x^T, -1\right] V_i \begin{bmatrix} x \\ -1 \end{bmatrix}$$

$$= \sum_{j=1}^{n} \mathrm{Var}(a_{ij})x_j^2 + \mathrm{Var}(b_i) + \sum_{\substack{k=1 \\ k\neq j}}^{n} \sum_{j=1}^{n} \mathrm{Cov}(a_{ij},a_{ik})x_j x_k - \sum_{j=1}^{n} \mathrm{Cov}(a_{ij},b_i)x_j .$$

<u>Lemma 1</u>: Define

$$\alpha_{ij} = K_{\alpha_i}^2 \mathrm{Var}(a_{ij}) - \left[E(a_{ij})\right]^2, \quad \text{for} \quad j = 1, \ldots, n,$$

$$\beta_{ijk} = K_{\alpha_i}^2 \mathrm{Cov}(a_{ij},a_{ik}) - E(a_{ik})D(a_{ik}), \quad \text{for} \quad j,k = 1, \ldots n \ (k\neq j),$$

$$\gamma_{ij} = 2E(b_i)E(a_j) - K^2_{\alpha_i} \text{ Cov } (a_{ij}, b_i) \ , \quad \text{for} \quad j = 1, \dots , n \ ,$$

$$r_i = \left[E(b_i)\right]^2 - K^2_{\alpha_i} \text{ Var } (b_i) \ .$$

Then, if $K_{\alpha_i} \geq 0$, the constraint,

$$\sum_{j=1}^{n} E(a_{ij})x_j + K_{\alpha_i} \sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} \leq E(b_i) \ ,$$

is equivalent to the pair of constraints,

$$\sum_{j=1}^{n} \alpha_{ij}x_j^2 + \sum_{\substack{k=1 \\ k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk}x_jx_k + \sum_{j=1}^{n} \gamma_{ij}x_j \leq r_i \ ,$$

$$\sum_{j=1}^{n} E(a_{ij})x_j \leq E(b_i) \ .$$

<u>Proof</u>: Rewrite the original constraint in the equivalent form,

$$K_{\alpha_i} \sqrt{\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}} \leq E(b_i) - \sum_{j=1}^{n} E(a_{ij})x_j \ .$$

This constraint is unaltered by squaring both sides,

$$K^2_{\alpha_i} \text{ Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\} \leq \left[E(b_i) - \sum_{j=1}^{n} E(a_{ij})x_j\right]^2 ,$$

provided that $E(b_i) - \sum_{j=1}^{n} E(a_{ij})x_j \geq 0$, i.e.,

$$\sum_{j=1}^{n} E(a_{ij})x_j \leq E(b_i) \ .$$

Substituting in the expression for $\text{Var}\left\{\sum_{j=1}^{n} a_{ij}x_j - b_i\right\}$ which was given preceding the Lemma, and then rearranging terms yields the desired result.

Theorem 3: Assume that $0 \le x_j \le 1$ for $j \in C$ and $x_j = 0$ or 1 for $j \in D$.

(i) $$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k \le \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} - \sum_{j=1}^{n} \left[\sum_{\substack{k=1\\k\neq j}}^{n} \left(\beta_{ijk} + \min\left\{\beta_{ijk}, 0\right\}\right) \right] (1-x_j)$$

$$\underset{=}{\text{def}} \; {}^{n}U_i(x) \; .$$

(ii) $$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k = U_i(x) \quad \text{if} \quad \sum_{j=1}^{n} x_j = n \; , \quad \text{or if}$$

$\beta_{ijk} \le 0$ for all $j, k = 1, \ldots, n (k \neq j)$ and $\sum_{\substack{j=1\\j\neq k}}^{n} x_j = n - 1$

for any $k = 1, \ldots, n$, or if $\beta_{ijk} \ge 0$ for all

$j, k = 1, \ldots, n (k \neq j)$ and $\sum_{j=1}^{n} x_j = 0$.

(iii) If $\beta_{ijk} \le 0$ for all $j, k = 1, \ldots, n (k \neq j)$,

then $U_i(x) \le h(x)$ for any function

$h(x)$ of the form,

$$h(x) = a_o + \sum_{j=1}^{n} d_j x_j \; ,$$

such that

$$\sum_{\substack{j=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k \le h(x)$$

for all admissible x and

$$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k = h(x)$$

if $\sum_{j=1}^{n} x_j = n$.

Proof: To prove (i), observe that

$$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk}x_jx_k = \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk}\,[1 - (1-x_k) - (1-x_j)x_k] .$$

Therefore, since $\beta_{ijk} = \beta_{ikj}$, so that $\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk}x_k = \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk}x_j$,

it follows that

$$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n}\beta_{ijk}x_jx_k = \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n}\beta_{ijk} - \sum_{j=1}^{n}\left[\sum_{\substack{k=1\\k\neq j}}^{n}\beta_{ijk}\right](1-x_j) - \sum_{j=1}^{n}\left[\sum_{\substack{k=1\\k\neq j}}^{n}\beta_{ijk}x_k\right](1-x_j)$$

$$\leq \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} - \sum_{j=1}^{n}\left[\sum_{\substack{k=1\\k\neq j}}^{n} (\beta_{ijk} + \min\,\{\beta_{ijk}, 0\})\right](1-x_j) ,$$

as was to be shown.

Part (ii) is evident by inspection, and (iii) follows immediately from (ii).

Corollary 1 to Theorem 3: In addition to the assumptions of Theorem 3, assume that $n_0 \leq \sum_{j=1}^{n} x_j < n_i$. Let $(\beta_{ijp_j}(1), \cdots , \beta_{ijp_j}(n-1))$ be a permutation of $\beta_{ij1}, \cdots , \beta_{ij(j-1)}, \beta_{ij(j+1)}, \cdots , \beta_{ijn}$ such that $\beta_{ijp_j}(1) \geq \beta_{ijp_j}(2) \geq \beta_{ijp_j}(n-1)$. Define

$$e_{ij} = \sum_{\substack{k=1\\k\neq j}}^{n}\beta_{ijk} + \sum_{k=n-n_0+1}^{n-1}\max\left\{\beta_{ijp_j}(k), 0\right\} + \sum_{k=n-n_1}^{n-1}\min\left\{\beta_{ijp_j}(k), 0\right\}, \text{ for } j=1,\ldots,n.$$

(i) $\sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk}x_jx_k \leq \sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk} - \sum_{j=1}^{n} e_{ij}(1-x_j)$.

(ii) Define $\Delta_{ij} = e_{ij} - \left[\sum_{\substack{k=1\\k\neq j}}^{n} (\beta_{ijk} + \min\{\beta_{ijk}, 0\})\right]$, for $j = 1,\ldots,n$, and let s_i be the sum of the $(n-n_1)$ smallest elements of $(\Delta_{i1},\ldots,\Delta_{in})$.

Then $\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk}x_jx_k \leq U_i(x) - s_i$.

Proof: The key step is to note that

$$\sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk}x_k = \sum_{\substack{k=1\\k\neq j}}^{n} \max\{\beta_{ijk}, 0\}\, x_k + \sum_{\substack{k=1\\k\neq j}}^{n} \min\{\beta_{ijk}, 0\}\, x_k$$

$$\geq \sum_{k=n-n_0+1}^{n-1} \max\left\{\beta_{ijp_j(k)}, 0\right\} + \sum_{k=n-n_1}^{n-1} \min\left\{\beta_{ijp_j(k)}, 0\right\} .$$

Part (i) then follows immediately by using the expression obtained in the proof of Theorem 3,

$$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk}x_jx_k = \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} - \left[\sum_{j=1}^{n} \sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk}\right](1-x_j) - \sum_{j=1}^{n} \left[\sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk}x_k\right](1-x_j)$$

$$\leq \sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk} - \sum_{j=1}^{n} e_{ij}(1-x_j) .$$

After noting that $\Delta_{ij} \geq 0$ for all j and that $\sum_{j=1}^{n} (1-x_j) \geq (n-n_1)$, Part (ii) then follows from Part (i).

Corollary 2 to Theorem 3:

(i) Any solution x that satisfies the set of constraints,

$$\sum_{j=1}^{n} \alpha_{ij}x_j^2 + U_i(x) + \sum_{j=1}^{n} \gamma_{ij}x_j \leq r_i , \quad \text{for } i = 1, \ldots , m ,$$

$$\sum_{j=1}^{n} E(a_{ij})x_j \leq E(b_i), \text{for } i = 1, \ldots , m ,$$

$$0 \leq x_j \leq 1 , \quad \text{for } j \in C ,$$

$$x_j = 0 \text{ or } 1, \text{for } j \in D ,$$

necessarily is a feasible solution.

(ii) If the additional assumption of Corollary 1 holds, then (i) will still hold after replacing $U_i(x)$ by either $\left[\sum_{\substack{k=1\\k\neq j}}^{n}\sum_{j=1}^{n}\beta_{ijk}-\sum_{j=1}^{n}e_{ij}(1-x_j)\right]$ or $[U_i(x) - s_i]$ for $i = 1, \ldots, m$.

(iii) If $C = \Phi$, then both (i) and (ii) will still hold after replacing x_j^2 by x_j for $j = 1, \ldots, n$.

(iv) Assume that $C \neq \Phi$ and that, for $i = 1, \ldots, m$, $\alpha_{ij} \geq 0$ for each $j \in C$. Suppose that, for each $j \in D$, x_j^2 is replaced in (i) by x_j, and that, for each $j \in C$, x_j^2 is approximated in (i) by a piecewise-linear function that coincides with x_j^2 only at $x_j = 0$, $x_j = 1$, and at the points where the slope of the piecewise-linear function changes (so that this piecewise-linear function necessarily is convex). Then both (i) and (ii) will still hold.

(v) A feasible solution x necessarily satisfies the set of constraints in (i) if (1) $\sum_{j=1}^{n} x_j = n$, or (2) $\beta_{ijk} \leq 0$ for all $j, k = 1, \ldots, n(k\neq j)$ and $\sum_{\substack{j=1\\j\neq k}}^{n} x_j = n-1$ for any $k = 1, \ldots, n$, or (3) $\beta_{ijk} \leq 0$ for all $j, k = 1, \ldots, n(k\neq j)$ and $\sum_{j=1}^{n} x_j = 0$. Furthermore, if $x_k = 0$ also in condition (2), then this entire statement still must hold after making the changes described in (iv).

<u>Proof:</u> Given the Fundamental Lemma and Lemma 1, these statements are an immediate consequence of Theorem 3 and Corollary 1.

<u>Theorem 4</u>: Assume that $0 \leq x_j \leq 1$ for $j \in C$, $x_j = 0$ or 1 for $j \in D$, and $n_0 \leq \sum_{j=1}^{n} x_j \leq n_1$. Define $p_j(k)$ as in Corollary 1 to Theorem 3, and let

$$\gamma'_{ij} = \sum_{k=1}^{n_0-1} \beta_{ijp_j(k)} + \sum_{k=n_0}^{n_1} \max\left\{\beta_{ijp_j(k)}, 0\right\}, \quad \text{for } j = 1, \ldots, n .$$

Then

$$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k < \sum_{j=1}^{n} \gamma'_{ij} x_j .$$

Proof: Note that $\sum\limits_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk} x_k \leq \gamma'_{ij} x_j$, so that

$$\begin{aligned}\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k &= \sum_{j=1}^{n} \left[\sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk} x_k\right] x_j \\ &\leq \sum_{j=1}^{n} \gamma'_{ij} x_j ,\end{aligned}$$

as was to be shown.

Corollary to Theorem 4: After imposing the assumptions of Theorem 4, Statement (i) and its extensions in Statements (iii) and (iv) of Corollary 2 to Theorem 3 will still hold after replacing $U_i(x)$ by $\sum\limits_{j=1}^{n} \gamma'_{ij} x_j$.

The decision whether to use Theorem 3 or Theorem 4 to obtain a linear upper bound on $\sum\limits_{\substack{k=1\\k\neq j}}^{n} \sum\limits_{j=1}^{n} \beta_{ijk} x_j x_k$ depends largely on the anticipated values of $\sum\limits_{j=1}^{n} x_j$. The bounds provided by Theorem 3 and Corollary 1 to Theorem 3 tend to be relatively tight if $\sum\limits_{j=1}^{n} x_j$ is relatively close to n. However, if the interesting feasible solutions tend to yield values of $\sum\limits_{j=1}^{n} x_j$ that are relatively small with respect to n, then the bound provided by Theorem 4 may be better, especially if $(n_1 - n_0)$ is not large and the β_{ijk} are not too variable. To illustrate, consider once again the numerical example introduced after Theorem 1, and impose the additional restriction that $2 \leq \sum\limits_{j=1}^{n} x_j \leq 3$. Thus, $\alpha_{ij} = -60$, $\beta_{ijk} = -100$, $\gamma_{ij} = 1000$, $r_i = 2300$, $e_{ij} = -700$, $s_i = -200$, and $\gamma'_{ij} = -100$ for all j, k, which yields the following results.

$\sum_{j=1}^{n} x_j$	$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k$	$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} - e_{ij}(1-x_j)$	$U_i(x)-s_i$	$\sum_{j=1}^{n} \gamma'_{ij} x_j$	$r_i - \sum_{j=1}^{n} \alpha_{ij} x_j^2 - \sum_{j=1}^{n} \gamma_{ij} x_j$
2	- 200	+ 100	+ 200	- 200	+ 420
3	- 600	- 600	- 600	- 300	- 540

Whereas both Theorems 3 and 4 provided upper bounds on $\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k$, Theorem 5 below will provide a linear lower bound on this function.

Theorem 5: Assume that $0 \leq x_j \leq 1$ for $j \in C$, $x_j = 0$ or 1 for $j \in D$, and $n_0 \leq \sum_{j=1}^{n} x_j \leq n_1$. Define $p_j(k)$ as in Corollary 1 to Theorem 3, let

$$N_j = \begin{cases} n_1, & \text{if } j \in C \\ n_1 - 1, & \text{if } j \in D, \end{cases}$$ and let

$$q_{ij} = \sum_{k=1}^{n_0 - 1} \beta_{ijp_j(n-k)} + \sum_{k=n_0}^{N_j} \min \left\{ \beta_{ijp_j(n-k)}, 0 \right\}, \quad \text{for } j = 1, \ldots, n \text{ .}$$

Then

(i) $$\sum_{j=1}^{n} q_{ij} x_j \leq \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k \text{ .}$$

(ii) $$\sum_{j=1}^{n} q_{ij} x_j = \sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k \quad \text{if} \quad \sum_{j=1}^{n} x_j = 0 \text{ .}$$

Proof: Note that

$$\left[\sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk} x_k \right] x_j \geq q_{ij} x_j \text{ ,} \quad \text{for } j = 1, \ldots, n \text{ ,}$$

so that

$$\sum_{\substack{k=1\\k\neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k = \sum_{j=1}^{n} \left[\sum_{\substack{k=1\\k\neq j}}^{n} \beta_{ijk} x_k \right] x_j \geq \sum_{j=1}^{n} q_j x_j \text{ .}$$

This verifies (i), and (ii) is obvious by inspection.

Corollary to Theorem 5:

(i) If $t_{ij} \leq q_{ij}$ for all i, j $(i = 1, \ldots, m;\ j = 1, \ldots, n)$, then any feasible solution x such that $n_0 \leq \sum_{j=1}^{n} x_j \leq n_1$ necessarily satisfies the set of constraints,

$$\sum_{j=1}^{n} \alpha_{ij} x_j^2 + \sum_{j=1}^{n} (t_{ij} + \gamma_{ij}) x_j \leq r_i \ , \quad \text{for} \quad i = 1, \ldots, m \ ,$$

$$\sum_{j=1}^{n} E(a_{ij}) x_j \leq E(b_i) \ , \quad \text{for} \quad i = 1, \ldots, m \ ,$$

$$0 \leq x_j \leq 1 \ , \qquad \text{for} \quad j \in C \ ,$$

$$x_j = 0 \text{ or } 1, \quad \text{for } j \in D \ .$$

(ii) Assume that $\alpha_{ij} \geq 0$ for $i = 1, \ldots, m$ and $j \in C$. Suppose that, for each $j \in D$, x_j^2 is replaced in (i) by x_j, and that, for each $j \in C$, x_j^2 is approximated in (i) by a convex piecewise-linear function that never exceeds x_j^2. Then (i) will still hold.

Referring again to the numerical example used to illustrate Theorems 3 and 4, $q_{ij} = -200$ for all j, so the following results would be obtained.

$\sum_{j=1}^{n} x_j$	$\sum_{\substack{k=1 \\ k \neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k$	$\sum_{j=1}^{n} q_{ij} x_j$	$r_i - \sum_{j=1}^{n} \alpha_{ij} x_i^2 - \sum_{j=1}^{n} \gamma_{ij} x_j$
2	- 200	- 400	+ 420
3	- 600	- 600	- 540

4. Exact Solution Procedures

To explore how to find an exact optimal solution to the chance-constrained programming problem formulated in Section 2, three exhaustive cases will be considered.

First, suppose that $D = \varphi$, so that all of the decision variables are continuous variables. Kataoka [21, pp. 194-5] has shown[10] that $\sqrt{[x^t, -1]v_i\begin{bmatrix} x \\ -1 \end{bmatrix}}$ is a convex function. Therefore, if $K_{\alpha_i} \geq 0$ for $i = 1, \ldots, m$, it is known that the deterministic equivalent form of the problem is an ordinary convex programming problem, for which there exists a number of algorithms. These include the ones developed by Rosen [30], Zoutendijk [38], Kelley [22], and Fiacco and McCormich [13, 14].

Now consider the case where $C = \varphi$, so that all of the decision variables are constrained to be either 0 or 1. This case may be solved in a straight-forward manner as follows. First, replace the deterministic equivalent form of the set of constraints by a set of uniformly tighter linear constraints. If its assumptions are satisfied, such sets are identified by Parts (i) and (ii) of Corollary 2 to Theorem 1. Otherwise, use one of the sets identified by Part (iii) of Corollary 2 to Theorem 3 or the Corollary to Theorem 4. Then find a good suboptimal solution to the resulting integer linear programming problem, which may be done by using one of the suboptimal algorithms developed by the author [18]. This solution necessarily is feasible for the original problem. Next, replace the deterministic equivalent form of the original constraints by a set of uniformly looser linear constraints. Such a set may be obtained from the Corollary to Theorem 2, if its assumptions hold, or from the Corollary to Theorem 5. This yields an ordinary integer linear programming problem whose set of feasible solutions includes all solutions that are feasible for the original problem. The final step is to apply a

[10] Also see Sinhal [31] and Van de Panne and Popp [35, pp. 421-2] for related investigations.

slightly modified version of a bound-and-scan algorithm developed by the author [17] for the integer linear programming problem. Given a good feasible solution, this algorithm repeatedly finds successively better ones until an optimal solution is reached. Therefore, the one modification that is required is that a new "better feasible solution" should be discarded if it is not feasible for the original problem. The final solution yielded by this modified algorithm will then be the optimal solution to the original problem.

Finally, consider the general case where there exist both continuous and integer decision variables. This problem is considerably more difficult than the special cases discussed above. However, if $K_{\alpha_i} \geq 0$ for $i = 1, \ldots, m$, then an algorithm recently developed by Veinott [36] is applicable. Although its computational efficiency is untested at present, this algorithm in principle will converge to the optimal solution for this problem.

5. Approximate Solution Procedures

Despite the availability of the exact solution procedures described above, approximate procedures that expedite computation and sensitivity analysis also are of considerable practical interest. Such procedures will be described below. To clarify the exposition, the requirement that $x_j = 0$ or 1 rather than $0 \leq x_j \leq 1$ for $j \in D$ will be ignored until the latter part of the section.

A good "definitely feasible" solution may be obtained relatively easily by applying the results given in Section 3. The first step is to replace the deterministic equivalent form of the set of constraints by a set of uniformly tighter piecewise-linear constraints. This new set may be selected from any of those provided by Part (iii) of Corollary 2 to Theorem 1, Parts (iii) and (iv) of Corollary 2 to Theorem 3, and the Corollary to Theorem 4 for which the assumptions hold. The optimal solution to this new problem is the desired "definitely feasible" solution. It may be obtained by converting the problem to an ordinary linear programming problem by the well-known separable convex programming technique (e.g., see Hadley [16, Ch. 4]) and then solving it by a streamlined version of the simplex method. For example, Hadly [16, pp. 126-9] describes how to use the decomposition principle to simplify the computational procedure.

The above approach is a conservative one in that many "barely feasible" solutions are excluded from consideration. It isn't always desirable to be this conservative, especially since the α_i often represent only rough guidelines that were set on a subjective basis. An opposite approach would be to consider all feasible solutions plus some "barely infeasible" ones. This may be done by proceeding exactly as before except that the new set of constraints would be selected from those provided by Part (iii) of the Corollary to Theorem 2 or Part (ii) of the Corollary to Theorem 5. This will provide a solution which yields a value for the objective function that is at least as large as that for the optimal solution but which may not be quite feasible for the given values of the α_1.

Another approach that may be more satisfactory is to combine the above two. One way to do this is to search for the best feasible solution along the line segment between the "definitely feasible" solution and the "nearly feasible" solution described above. Another way is to search for the best feasible solution yielded by using a weighted average of the two sets of constraints. If these two sets differ only in the right-hand side of the respective constraints, then the solutions may be obtained easily by standard parametric programming procedures (e.g., see Hillier and Lieberman [20]).[11]

Still another approach is to use the solution obtained in any of the above ways to construct a better set of approximate linear constraints, which are then used to find the final solution. For example, let $x^* = [x_1^*, \ldots, x_n^*]^T$ be the "definitely feasible" solution described above. Then one may replace $\sum_{\substack{k=1 \\ k \neq j}}^{n} \sum_{j=1}^{n} \beta_{ijk} x_j x_k$ by $\sum_{j=1}^{n} \left[\sum_{\substack{k=1 \\ k \neq j}}^{n} \beta_{ijk} x_k^* \right] x_j$ in Lemma 1 and use the resulting pair of constraints to replace the corresponding original constraint for each $i = 1, \ldots, n$. Alternatively, the set of constraints described in Corollary 2 to Theorem 1 may be modified by replacing $R_i(x)$ by

$$\left\{ \sum_{j \in C} \sqrt{\bar{v}_i - \sigma_{ij}^2 + \sigma_{ij}^2 x_j^2} + \sum_{j \in D} \left[\sqrt{\bar{v}_i} - \sqrt{\bar{v}_i - \sigma_{ij}^2} \right] x_j + d_i \right\}$$

for each $i = 1, \ldots, m$, where $\bar{v}_i = \text{Var}\left\{ \sum_{j=1}^{n} a_{ij} x_j^* - b_i \right\}$ and d_i is

---

[11] For this case, an upper bound on the difference between the value of the objective function for the optimal solution and the selected solution may be obtained easily as follows. Let M be the number of functional constraints, and let $[y_1^*, \ldots, y_M^*]^T$ be the optimal dual solution corresponding to the selected primal solution. Let Δb_i be the difference between the right-hand side of the i^{th} functional constraint for the "nearly feasible" solution described above and that for the selected solution, where $i = 1, \ldots, M$. Then this upper bound is $\sum_{i=1}^{M} y_i^* \Delta b_i$. (See the section entitled "Bounding Procedure for Group 3 and 4 Variables" in [17] for the justification of this result.) In general, the value of the objective function for the "nearly feasible" solution also provides an upper bound on the corresponding value for the optimal solution.

the constant required to make this new function equal to $\sqrt{\operatorname{Var}\left\{\sum_{j=1}^{n} a_{ij} x_j^* - b_i\right\}}$ when $x = x^*$. For either case, one would then apply the separable convex programming technique and solve as before. It appears that the resulting solution should tend to be nearly optimal and either feasible or sufficiently close to feasible for most practical purposes.

All of the above approximate procedures reduce to solving a linear programming problem. This has important advantages over the exact solution procedures described in the preceding section. One is the relatively high efficiency with which a linear programming problem can be solved originally and then subjected to sensitivity analysis. Another advantage is the availability of linear duality theory for analyzing the problem. Thus the optimal dual solution, which is an automatic by-product of the ordinary linear programming calculations, is available for further guidance and for study of the original policy decisions made when constructing the chance-constrained programming model.

Now consider how to deal with variables that are restricted to be either 0 or 1. The linear programming solution procedures described above may assign fractional values to some of these variables, so that some modification is required. The simplest approach is to attempt to round such variables up or down in such a way as to obtain a feasible solution with a relatively large value of the objective function. Fortunately, according to a theorem due to Weingartner [37, pp. 35 ff.], the number of fractional variables in these linear programming solutions cannot exceed the number of functional constraints. Therefore, if m is not large, there can only be a relatively few variables that will need adjusting.

Furthermore, since the final solution needs to be feasible only for the original problem, and not for the approximating linear programming problem, it may be possible to increase the value of the objective function over that for the optimal linear programming solution.

A more systematic approach to the integer or mixed integer problem is to formulate an approximating linear programming problem as described above and then apply one of the available algorithms for the pure integer or mixed integer programming problem (see Balinski [1] and Beale [2] for a survey of these algorithms). However, it may require more computation time than can be justified, especially considering that the resulting solution need not be optimal for the original problem. A more efficient procedure is to instead apply a suboptimal integer programming algorithm, such as those developed by Reiter and Rice [29] and (if $C = \varphi$) by the author [18]. The Reiter-Rice algorithm also can be applied directly to the original problem in its deterministic equivalent form, which should tend to yield a slightly better solution with somewhat less efficiency.

6. Linear Decision Rules for Continuous Variables

Now consider the situation where the values taken on by certain of the random variables will become known before some of the decision variables must be assigned values. It is highly desirable to formulate and solve problems of this type in such a way that the ultimate decisions will be partially based on the new information that has become available. As indicated in Section 1, one way to do this is to formulate the problem in terms of choosing a decision rule,

$$x = \psi(A, b, c) .$$

where ψ is a vector-valued function to be determined.

This approach becomes quite tractable when ψ is restricted to the class of linear decision rules. In particular, let the components of ψ be

$$x_k = \sum_{i=1}^{m} \sum_{j=1}^{n} u_{ijk} a_{ij} + \sum_{i=1}^{m} v_{ik} b_i + w_k \,, \quad \text{for} \quad k = 1, \ldots , n \,,$$

where u_{ijk} and v_{ik} are decision variables, except that they are set equal to zero if $k \in D$ or if x_k must be assigned a value before observing a_{ij} and b_i, respectively; w_k is a decision variable if $k \in D$ or if x_k must be assigned a value before observing any of the a_{ij} and b_i, and it is an arbitrary constant otherwise.

Since a general linear decision rule may assign a value other than 0 or 1 to x_k, it has been necessary to reserve these rules for only the continuous decision variables. Therefore, if $k \in D$, then $x_k = w_k$ only so that it does not depend on the values taken on by the a_{ij} and b_i. An alternative approach that does permit deferring 0 - 1 decisions will be described in the next section.

To solve the chance-constrained programming problem with the indicated linear decision rule, notice that

$$\begin{aligned}\sum_{k=1}^{n} c_k x_k &= \sum_{k=1}^{n} c_k \left[\sum_{i=1}^{m} \sum_{j=1}^{n} u_{ijk} a_{ij} + \sum_{i=1}^{m} v_{ik} b_i + w_k \right] \\ &= \sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{n} c_k a_{ij} u_{ijk} + \sum_{k=1}^{n} \sum_{i=1}^{m} c_k b_i v_{ik} + \sum_{k=1}^{n} c_k w_k \,,\end{aligned}$$

and, similarly,

$$\sum_{k=1}^{n} a_{tk} x_k - b_t = \sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{n} a_{tk} a_{ij} u_{ijk} + \sum_{k=1}^{n} \sum_{i=1}^{m} a_{tk} b_i v_{ik} + \sum_{k=1}^{n} a_{tk} w_k - b_t \,.$$

Therefore, if the c_k are statistically independent of the a_{ij} and b_i, and if the restriction that decision variables lie between 0 and 1 is ignored, then the original model reduces to

$$\max \left\{ \sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{n} E(c_k)E(a_{ij})u_{ijk} + \sum_{k=1}^{n} \sum_{i=1}^{m} E(c_k)E(b_i)v_{ik} + \sum_{k=1}^{n} E(c_k)w_k \right\} ,$$

subject to

$$\text{Prob}\left\{ \sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{n} a_{tk}a_{ij}u_{ijk} + \sum_{k=1}^{n} \sum_{i=1}^{m} a_{tk}b_i v_{ik} + \sum_{k=1}^{n} a_{tk}w_k - b_t \le 0 \right\} \ge \alpha_t ,$$

$$\text{for} \quad t = 1, \ldots , m$$

$$w_k = 0 \text{ or } 1, \text{ for } k \in D .$$

Note that the decision variables here play the same role that the x_k played before with the zero-order decision rule. Similarly, the corresponding $(a_{tk}a_{ij})$, $(a_{tk}b_j)$, and a_{tk} now play the same role as the a_{tk} before. Therefore, except for two added complications, one may proceed exactly as before to solve for the decision variables. These complications are (1) it is now more difficult to determine the expectation, variance, and covariance terms, and (2) the requirement that the decision variables must lie between 0 and 1 still remains to be taken into account. The remainder of the section will be devoted to discussing these two complications.

To illustrate the former complication, consider how one would find the expectation, variance, and covariance terms involving only $(a_{tk}a_{ij})$. Assume that the elements of A and b are mutually independent, that the expected value and variance of these elements are known, and that $u_{ikk} = 0$ for all i and k. Then

$$E(a_{tk}a_{ij}u_{ijk}) = u_{ijk}E(a_{tk})E(a_{ij})$$

and

$$\begin{aligned}\mathrm{Var}(a_{tk}a_{ij}u_{ijk}) &= u_{ijk}^2\left[E\{(a_{tk}a_{ij})^2\} - E(a_{tk}a_{ij})^2\right] \\ &= u_{ijk}^2\left[E(a_{tk}^2)E(a_{ij}^2) - E(a_{tk})^2E(a_{ij})^2\right] \\ &= u_{ijk}^2\left[\left(\mathrm{Var}(a_{tk})+E(a_{tk})^2\right)\left(\mathrm{Var}(a_{ij})+E(a_{ij})^2\right)-E(a_{tk})^2E(a_{ij})^2\right] .\end{aligned}$$

If $k \neq k'$ and either $i \neq i'$ or $j \neq j'$, then $\mathrm{Cov}\{a_{tk}a_{ij}u_{ijk}, a_{tk'}a_{i'j'}u_{i'j'k'}\} = 0$. However, if $k = k'$ instead, then

$$\begin{aligned}&\mathrm{Cov}\{a_{tk}a_{ij}u_{ijk}, a_{tk}a_{i'j'}u_{i'j'k}\} \\ &= u_{ijk}u_{i'j'k}E\{(a_{tk}a_{ij} - E(a_{tk}a_{ij}))(a_{tk}a_{i'j'} - E(a_{tk}a_{i'j'}))\} \\ &= u_{ijk}u_{i'j'k}[E(a_{tk}^2a_{ij}a_{i'j'}) - E(a_{tk}a_{ij})E(a_{tk}a_{i'j'})] \\ &= u_{ijk}u_{i'j'k}E(a_{ij})E(a_{i'j'})[E(a_{tk}^2) - E(a_{tk})^2] \\ &= u_{ijk}u_{i'j'k}E(a_{ij})E(a_{i'j'}) \mathrm{Var}(a_{tk}) .\end{aligned}$$

If $k \neq k'$, but $i = i'$ and $j = j'$, it follows similarly that

$$\mathrm{Cov}\{a_{tk}a_{ij}u_{ijk}, a_{tk'}a_{ij}u_{ijk'}\} = u_{ijk}u_{ijk'}E(a_{tk})E(a_{tk'}) \mathrm{Var}(a_{ij}) .$$

(The corresponding expressions involving $a_{tk}b_i$ and a_{tk} are obtained in the same way.) These results have been based on the assumption of independent random variables. If this assumption does not hold, it then becomes necessary to use the joint probability distribution of the elements of A and b in order to calculate these expectation, variance, and covariance terms.

Now consider the former requirement that bounds of 0 and 1 be imposed upon the decision variables. If the objective is to obtain an optimal decision rule by the solution procedures described in Section 4, then these lower and upper bound constraints play no essential role and can be omitted safely. However, this is not the situation if an approximate decision rule is being sought by means of the inequalities presented in Section 3, since these inequalities are based critically on the variables lying between 0 and 1. Therefore, it is necessary for this case that the current decision variables be so constrained, although the solution procedure would not require that the replaced variables - the x_k - lie between these bounds. However, unless appropriate adjustments are made, arbitrarily imposing such constraints on the u_{ijk} and v_{ik} may eliminate interesting decision rules from consideration. These adjustments are discussed below.

One may essentially insure that all of the interesting values of the u_{ijk} and v_{ik} are nonnegative merely by assigning sufficiently small (possibly negative) values to each w_k that is not a decision variable. A scaling factor may then be used to essentially insure that these interesting values do not exceed one. In other words, each a_{ij} (and b_i) and the corresponding u_{ijk} (and v_{ik}) would be multiplied and divided, respectively, by a sufficiently large constant (not necessarily the same for all a_{ij} and b_i). This scales down the u_{ijk} and v_{ik} without changing the problem. After sufficient translation and change of scale of the decision variables, the lower and upper bound constraints may be imposed without seriously reducing the set of feasible solutions.

This process of translating and scaling the decision variables may be conducted somewhat by trial and error. For example, a trial approximate solution might be obtained after a modest amount of scaling. If many of the variables in this solution equal one, it would then be evident that additional scaling is required to reduce the distortion caused by adding the upper bound constraints. On the other hand, one should be careful not to scale down the variables too far, since this would cause the difference between the number of decision variables and the sum of the variables to be relatively large. The approximations introduced by Theorems 1 and 3 in Section 3 can deteriorate seriously if this difference becomes too large. If a large difference is unavoidable, it may be best to use the approximation introduced by Theorem 4 instead. It may also be very desirable to use the methods described in Section 5 for improving upon an initial approximate solution.

7. Two-stage Decision Rules

The linear decision rule approach described in the preceding section provides a convenient way of permitting the deferment of decisions represented by continuous decision variables. However, it does not apply to discrete decision variables. This section develops an alternative approach which, although less precise and flexible, does apply to both continuous and discrete variables. It is motivated by the two-stage formulation of linear programming under uncertainty with discrete probability distributions that was developed by Dantzig [11] and others (see Naslund [27] for references to related work). What is presented here is essentially an extension to include continuous distributions in a chance-constrained programming format.

Suppose that certain of the decisions must be made immediately (stage 1) and the remainder can be deferred until some later point in time (stage 2) when the values taken on by certain of the random variables will be known. Let $x = [{y \atop z}]$, where the elements of y and z are the stage 1 and stage 2 decision variables, respectively, and let $c = [c_Y, c_Z]$ be the corresponding partitioning of c. Let n_Y and $n_Z = n - n_Y$ be the number of elements of y and z, respectively. Similarly, partition the set of functional constraints,

$$\text{Prob}\,\{Ax \leq b\} \geq \alpha ,$$

into those constraints (if any) involving only the stage 1 variables,

$$\text{Prob}\,\{A_Y y \leq b_Y\} \geq \alpha_Y ,$$

and the others,

$$\text{Prob}\,\{A_{YZ} y + A_Z z \leq b_Z\} \geq \alpha_Z .$$

Therefore, the original formulation of the chance-constrained programming problem can be written as

$$\max E\{C_Y y + C_Z z\} ,$$

subject to

$$\begin{aligned} &\text{Prob}\,\{A_Y y \leq b_Y\} \geq \alpha_Y , \\ &\text{Prob}\{A_{YZ} y + A_Z z \leq b_Z\} \geq \alpha_Z , \\ &0 \leq x_j \leq 1, \quad \text{for} \quad j \in C , \\ &x_j = 0 \text{ or } 1, \quad \text{for} \quad j \in D . \end{aligned}$$

Let M be the array,

$$M = \begin{bmatrix} c_Y & c_Z & 0 \\ A_Y & 0 & b_Y \\ A_{YZ} & A_Z & b_Z \end{bmatrix} ,$$

so that M has a multivariate probability distribution. Let R be the random vector whose elements are the random variables whose value will be known when z must be specified, and let n_r be the number of elements of R. Let S be the range space of R, so that S is the set in n_r-dimensional Euclidean space consisting of the values R can take on. Suppose that S is partitioned into n_s subsets, $S_1, S_2, \ldots, S_{n_s}$, such that $S_i \cap S_j = \varphi$ for all $i \neq j$ and $S_1 \cup S_2 \ldots \cup S_{n_s} = S$. Let

$$p_i = \text{Prob}\,\{R \in S_i\}\ , \quad \text{for} \quad i = 1, \ldots, n_s\ .$$

For each $i = 1, \ldots, n_s$, let $M^{(i)}$ be the array of random variables,

$$M^{(i)} = \begin{bmatrix} c_Y^{(i)} & c_Z^{(i)} & 0 \\ A_Y^{(i)} & 0 & b_Y^{(i)} \\ A_{YZ}^{(i)} & A_Z^{(i)} & b_Z^{(i)} \end{bmatrix}$$

such that the joint distribution of $M^{(i)}$ coincides with the conditional joint distribution of M given that $R \in S_i$.

Given the above development, the decision structure of the problem can now be refined considerably. Rather than having to choose z independently of R, the choice of z will be made conditional upon the identity of the subset of S which contains the value taken on by R. Thus, for each $i = 1, \ldots, n_s$, let $z^{(i)}$ be the value of z that will be chosen if $R \in S_i$. The resulting formulation of the problem is to determine $y, z^{(1)}, \ldots, z^{(n_s)}$ so as to

$$\max E(cx) = \sum_{j=1}^{n_Y} E(c_j)y_j + \sum_{i=1}^{n_s} p_i \sum_{j=n_Y+1}^{n} E(c_j^{(i)}) z_{j-n_Y}^{(i)}\ ,$$

subject to Prob $\{A_Y y \leq b_Y\} \geq \alpha_Y$

$$\text{Prob } \{A_{YZ}^{(i)} y + A_Z^{(i)} z^{(i)} \leq b_Z^{(i)}\} \geq \alpha_Z \ , \quad \text{for} \quad i = 1, \ldots , n_s$$

$$0 \leq y_j \leq 1, \quad \text{for} \quad j \in C \cap \{1, \ldots , n_Y\}$$

$$0 \leq z_{j-n_Y}^{(i)} \leq 1 \ , \quad \text{for} \quad j \in C \cap \{n_Y + 1, \ldots , n\} \ , \ i = 1, \ldots, n_s$$

$$y_j = 0 \text{ or } 1, \quad \text{for} \quad j \in D \cap \{1, \ldots , n_Y\}$$

$$z_{j-n_Y}^{(i)} = 0 \text{ or } 1, \quad \text{for} \quad j \in D \cap \{n_Y + 1, \ldots , n\}, \ i = 1, \ldots, n_s \ .$$

This is an ordinary chance-constrained programming problem with 0 - 1 and bounded decision variables, so it can still be solved by the procedures described in Sections 4 and 5. Furthermore, since it takes advantage of the information that becomes available between the first-stage and second-stage decisions, this refined formulation should yield significantly better decisions, provided that the S_i are chosen strategically.

The main consideration in choosing the S_i is that the points within a subset should be as similar as possible in their impact upon what the second-stage decisions should be, whereas points in different subsets should be as different as possible in this regard. For example, suppose that the relevant new information for the second-stage decisions is how favorable were the overall consequences of the first-stage decisions. One might then construct say five categories — very unfavorable, somewhat unfavorable, neutral, somewhat favorable, and very favorable. The possible values of R would be matched up with these categories and thereby assigned to the five subsets, $S_1, \ldots , S_5$. If the individual consequences of certain groups of the first-stage decisions are also particularly relevant, it might be desirable to use combinations of these categories and thereby obtain a larger number of subsets.

It is quite evident that this two-stage formulation could be generalized to a k-stage formulation. However, for most problems, it is doubtful that the benefits of doing so would justify the considerable increase in the complexity of setting up the problem.

REFERENCES

[1] Balinski, M. L., "Integer Programming: Methods, Uses, Computation," Management Science, Vol. 12 (1965), pp. 253-313.

[2] Beale, E. M. L., "Survey of Integer Programming," Operational Research Quarterly, Vol. 16 (1965), pp. 219-228.

[3] Ben-Israel, A., "On Some Problems of Mathematical Programming," Ph.D. Dissertation, Northwestern University, Evanston, Illinois, June, 1962.

[4] Charnes, A. and Cooper, W. W., "Chance-Constrained Programming," Management Science, Vol. 6 (1959), pp. 73-79.

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

[6] Charnes, A. and Cooper, W. W., "Chance Constraints and Normal Deviates," Journal of the American Statistical Association, Vol. 57 (1962), pp. 134-148.

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

[8] Charnes, A., Cooper, W. W., and Thompson, G. L., "Constrained Generalized Medians and Hypermedians as Deterministic Equivalents for Two-Stage Programs under Uncertainty," Management Science, Vol. 12 (1965), pp. 83-112.

[9] Charnes, A., Cooper, W. W., and Thompson, G. L., "Critical Path Analyses Via Chance-Constrained and Stochastic Programming," Operations Research, Vol. 12 (1964), pp. 460-470.

[10] Dantzig, George B., Linear Programming and Extensions, Princeton University Press, Princeton, N. J., 1963.

[11] Dantzig, George B., "Linear Programming under Uncertainty," Management Science, Vol. 1 (1955), pp. 197-206.

[12] Feller, William, An Introduction to Probability Theory and Its Applications, Vol. 1, 2nd edition, John Wiley and Sons, New York, 1957.

[13] Fiacco, Anthony V. and McCormick, Garth P., "The Sequential Unconstrained Minimization Technique for Nonlinear Programming, A Primal-Dual Method," Management Science, Vol. 10 (1964), pp. 360-366.

[14] Fiacco, Anthony V. and McCormick, Garth P., "Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming," Management Science, Vol. 10 (1964), pp.601-617.

[15] Graves, Robert L. and Wolfe, Phillips (eds.), Recent Advances in Mathematical Programming, McGraw-Hill, New York, 1963.

[16] Hadley, G., Nonlinear and Dynamic Programming, Addison-Wesley, Reading, Mass., 1961.

[17] Hillier, Frederick S., "An Optimal Bound-and-Scan Algorithm for Integer Linear Programming," Technical Report No. 3, Contract Nonr-225(89), Stanford University, August 19, 1966. Submitted to Operations Research.

[18] Hillier, Frederick S., "Efficient Suboptimal Algorithms for Integer Linear Programming with an Interior," Technical Report No. 2, Contract Nonr-225(89), Stanford University, August 12, 1966. Submitted to Operations Research.

[19] Hillier, Frederick S., "The Evaluation of Risky Interrelated Investments," Technical Report No. 73, Contract Nonr-225(53), Stanford University, July 24, 1964. To appear in Capital Budgeting of Interrelated Projects, Office of Naval Research Monograph on Mathematical Methods in Logistics (edited by Dorothy M. Gilford, A. Charnes, and W. W. Cooper), 1967.

[20] Hillier, Frederick S., and Lieberman, Gerald J., Introduction to Operations Research, Holden-Day, San Francisco, 1967.

[21] Kataoka, Shinji, "A Stochastic Programming Model," Econometrica, Vol. 31 (1963), pp. 181-196.

[22] Kelley, J. E., Jr., "The Cutting-Plane Method for Solving Convex Programs," Journal of the Society for Industrial and Applied Mathematics, Vol. 8 (1960), pp. 703-712.

[23] Kirby, M., "Generalized Inverses and Chance-Constrained Programming," Northwestern University Research Project: Temporal Planning and Management Decision under Risk and Uncertainty, Evanston, Ill., March, 1965.

[24] Loève, Michel, Probability Theory, 2nd edition, D. Van Nostrand, Princeton, New Jersey, 1960.

[25] Miller, Bruce L. and Wagner, Harvey M., "Chance-Constrained Programming with Joint Constraints," Operations Research, Vol. 13 (1965), pp. 930-945.

[26] Naslund, Bertil, "A Model of Capital Budgeting under Risk," The Journal of Business, Vol. XXXIX (1966), pp. 257-271.

[27] Naslund, Bertil, "Mathematical Programming under Risk," The Swedish Journal of Economics, 1965, pp. 240-255.

[28] Naslund, Bertil, and Whinston, Andrew, "A Model of Multi-Period Investment under Uncertainty," Management Science, Vol. 8 (1962), pp. 184-200.

[29] Reiter, Stanley and Rice, Donald B., "Discrete Optimizing Solution Procedures for Linear and Nonlinear Integer Programming Problems," Institute Paper No. 109, Institute for Research in the Behavioral, Economic, and Management Sciences, Purdue University, Lafayette, Indiana, May, 1965.

[30] Rosen, J. B., "The Gradient Projection Method for Nonlinear Programming — Part II: Nonlinear Constraints," Journal of the Society for Industrial and Applied Mathematics, Vol. 9 (1961), pp. 514-532.

[31] Sinhal, S. M., "Programming with Standard Errors in the Constraints and the Objective," p. 121 in [15].

[32] Thiel, H., "Some Reflections on Static Programming under Uncertainty," Weltwirtschaftliches Archiv., Vol. 87 (1961); reprinted as Publication No. 5 of the International Center for Management Science (The Netherlands School of Economics, Rotterdam).

[33] Thompson, G. L., Cooper, W. W., and Charnes, A., "Characterizations by Chance-Constrained Programming," pp. 113-120 in [15].

[34] Tintner, G., "Stochastic Linear Programming with Applications to Agricultural Economics," in H. A. Antosiewics (ed.)., Second Symposium on Linear Programming, National Bureau of Standards, Washington, 1955.

[35] Van De Panne, C. and Popp, W., "Minimum-Cost Cattle Feed under Probabilistic Protein Constraints," Management Science, Vol. 9, (1963), pp. 405-430.

[36] Veinott, Arthur F., Jr., "The Supporting Hyperplane Method for Unimodal Programming," Technical Report No. 10, NSF Grant GP-3739, Stanford University, March 10, 1966.

[37] Weingartner, H. Martin, Mathematical Programming and the Analysis of Capital Budgeting Problems, Prentice-Hall, Englewood Cliffs, N. J., 1963.

[38] Zoutendijk, G., Methods of Feasible Directions, Elsevier Publishing Company, Amsterdam, 1960.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Stanford University Department of Statistics Stanford, California	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Chance-Constrained Programming with 0-1 or Bounded Decision Variables

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (Last name, first name, initial)

Hillier, Frederick S.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 26, 1966	44	38

8a. CONTRACT OR GRANT NO. Contract Nonr-225(53)	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 92
b. PROJECT NO. NR-042-002	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) Nonr-225(89)(NR-047-061)
d.	

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Logistics and Mathematical Statistics Branch Office of Naval Research Washington, D. C. 20360

13. ABSTRACT

This paper considers the chance-constrained programming problem where the decision variables can be either bounded and continuous or restricted to be either zero or one, and where some or all of the elements of A, b, and c are random variables that may be statistically dependent. Both exact and approximate solution procedures are presented, where most of these are based on several linear inequalities that permit this problem to be approximated by a number of ordinary (integer or noninteger) linear programming problems. Either zero-order or linear decision rules are allowed for the continuous variables, and a general method of making "second-stage decisions" with either continuous or 0-1 variables is developed.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
operations research mathematical programming chance-constrained programming linear programming under uncertainty linear bounds						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________ ."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________ ."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________ ."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional