Stochastic linear programming

Charles Blanchard Millham

Iowa State University
MILLHAM, Charles Blanchard, 1936—
STOCHASTIC LINEAR PROGRAMMING.

Iowa State University of Science and Technology
Ph.D., 1962
Mathematics

University Microfilms, Inc., Ann Arbor, Michigan
STOCHASTIC LINEAR PROGRAMMING

by

Charles Blanchard Millham

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Economics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa
1962
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Previous Results in Stochastic Programming</td>
<td>23</td>
</tr>
<tr>
<td>The Passive Approach in Stochastic Linear Programming</td>
<td>54</td>
</tr>
<tr>
<td>The Active Approach in Stochastic Programming</td>
<td>78</td>
</tr>
<tr>
<td>Concluding Comments</td>
<td>95</td>
</tr>
<tr>
<td>Summary</td>
<td>98</td>
</tr>
<tr>
<td>Literature Cited</td>
<td>100</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>104</td>
</tr>
</tbody>
</table>
INTRODUCTION

The maximization (or minimization) of a linear function, called the objective function, subject to a set of linear constraints is called linear programming. A minimization problem, for example, may have the form given by

minimize the objective function

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to

$$x_j \geq 0, \ j = 1, \ldots, n$$

and

$$a_{11}x_1 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + \cdots + a_{mn}x_n = b_m$$

where the $$c_j, j = 1, \ldots, n$$, $$b_i, i = 1, \ldots, m$$, and $$a_{ij}, i = 1, \ldots, m, j = 1, \ldots, n$$, are constants. We shall further assume that $$m < n$$, and that we have an underdetermined system of equations forming the constraints.

An example of such a (maximization) programming problem is the activity analysis problem.

A manufacturer has at his disposal fixed amounts of a number of different resources, such as labor, raw material, and equipment, which can be combined to produce any one of several different commodities or combinations of commodities.
He is assumed to know how much of resource $i$ is needed in the production of one unit of output $j$ and how much profit he makes for each unit of commodity $j$ produced. He will then desire to produce that commodity or combination of commodities that will result in the greatest profit.

To establish this as a linear programming problem we make the following definitions.

Let $m =$ the number of resources available, $n =$ the number of commodities of which production is contemplated, $a_{ij} =$ the number of units of resource $i$ required to produce one unit of commodity $j$, $b_i =$ the number of units of resource $i$ available, $c_j =$ profit per unit of commodity $j$ produced, and let $x_j$ be the level of activity (the amount produced) of the $j$th commodity.

The $a_{ij}$ are sometimes called input-output coefficients.

The total amount of resource $i$ that is used is given by

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n.$$ Since this may not exceed the amount of resource $i$ available, we have, for all $i$, that

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leq b_i.$$ Since profit is given by profit per unit times number of units we wish to maximize $c_1x_1 + c_2x_2 + \cdots + c_nx_n$ subject to

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leq b_i$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m.$$ Since, further, it makes no sense to contemplate produc-
tion of negative quantities of a commodity, we add the further restriction that \( x_j \geq 0 \), for all \( j \).

Such a problem is much more compactly expressed as follows.

Let

\[
\begin{pmatrix}
 b_1 \\
 b_2 \\
 \vdots \\
 b_m 
\end{pmatrix},
\begin{pmatrix}
 c_1 \\
 c_2 \\
 \vdots \\
 c_n 
\end{pmatrix}
\]

and

\[
\begin{pmatrix}
 x_1 \\
 x_2 \\
 \vdots \\
 x_n 
\end{pmatrix}
\]

be column vectors and let

\[
A = \begin{pmatrix}
 a_{11} & \cdots & a_{1n} \\
 \vdots & \ddots & \vdots \\
 a_{m1} & \cdots & a_{mn} 
\end{pmatrix}
\]

be an \( m \times n \) matrix. The problem may then be expressed as

\[
\text{maximize } c'x
\]

subject to

\[
Ax \leq b \text{ and } x \geq 0,
\]

where \( c' \) indicates the transpose of \( c \).

To briefly present another example, now of a minimization problem, let us suppose that we are given the nutritional value of several different foods, the cost of each food, and the minimum daily requirement of each nutrient and that we are asked to find the minimum-cost diet which still fulfills the minimum daily requirements of nutrition. Again, we make the following definitions: let \( m = \) number of nutrients, \( n = \)
number of foods, $a_{ij}$ = number of milligrams (say) of the $i$th nutrient in one ounce of the $j$th food, $b_i$ = the minimum number of milligrams of the $i$th nutrient required, $c_j$ = the cost per ounce of the $j$th food, and $x_j$ = number of ounces of the $j$th food to be purchased. The programming problem may then be phrased as

$$\text{minimize } c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \geq b_2$$
$$\vdots$$
$$a_{mn}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_n$$

$$x_j \geq 0, \ j = 1, \cdots, n.$$
Definition 1.2: A set $S$ of points $x$ is said to be convex if and only if every convex combination of points $x$ in $S$ is again a point of $S$.

Definition 1.3: A point $k$ of a set $S$ is said to be an extreme point of $S$ if and only if $k$ cannot be expressed as a convex combination of any other two distinct points of $S$.

It will be remembered that the two examples of programming problems presented earlier had their constraints expressed in the form of inequalities. Before proceeding further it is necessary to demonstrate how these may be transformed to equalities, as the discussion which follows is valid only for constraints expressed as equalities.

Suppose we have a maximization (deterministic) programming problem expressed as

\[
\text{maximize } c_1 x_1 + c_2 x_2 + \cdots + c_n x_n
\]

subject to
\[
\begin{align*}
a_{11} x_1 + \cdots + a_{1n} x_n &\leq b_1 \\
a_{21} x_1 + \cdots + a_{2n} x_n &\leq b_2 \\
\quad &\vdots \\
a_{m1} x_1 + \cdots + a_{mn} x_n &\leq b_m
\end{align*}
\]

$x_j \geq 0$, $j = 1, \ldots, n$.

These inequalities may be transformed to equalities by

\[
i = 1, \ldots, n, \text{ and } \sum_{i=1}^{n} a_i = 1.
\]
the device of adding to each inequality $i$ a slack or disposal variable $x_{n+1} \geq 0$. Thus, we may write the above as

$$\text{maximize } c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

subject to

$$a_{11} x_1 + \cdots + a_{1n} x_n + x_{n+1} = b_1$$
$$a_{21} x_1 + \cdots + a_{2n} x_n + x_{n+2} = b_2$$
$$\vdots$$
$$a_{m1} x_1 + \cdots + a_{mn} x_n + x_{n+m} = b_m$$

$$x_j \geq 0, j = 1, \ldots, m+n.$$ 

The matrix form is then clearly

$$\text{max } d^T z$$

subject to $Bz = e$

$$z \geq 0, \text{ where } z = (x_1, x_2, \ldots, x_{n+1}, x_{n+m})^T + (z_1, z_2, \ldots, z_n, z_{n+1}, \ldots, z_{n+m})^T,$$

$$B = \begin{pmatrix}
    a_{11} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\
    a_{21} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\
    \vdots & & \vdots & & \vdots & & \vdots \\
    a_{m1} & \cdots & a_{mn} & 0 & 0 & \cdots & 0 & 1
\end{pmatrix}$$

$$d = \begin{pmatrix}
    c_1 \\
    \vdots \\
    \vdots \\
    c_n \\
    0 \\
    \vdots \\
    0
\end{pmatrix}, \quad e = \begin{pmatrix}
    b_1 \\
    b_2 \\
    \vdots \\
    b_m \\
    0 \\
    \vdots \\
    0
\end{pmatrix}.$$
The transformation of the constraint inequalities in a minimization problem is accomplished similarly, except that the nonnegative slack variables are subtracted from the inequalities.

Until otherwise indicated, we shall immediately assume that any linear programming problems under consideration have been placed in the equality form or representation. We may now proceed with more preliminary results.

Definition 1.4: Let a deterministic linear programming problem be given by

maximize $d'z$

subject to $Bz = e$, $z \geq 0$.

Then a feasible solution is a vector $z$ satisfying the constraints.

The following theorems are now stated, mostly without proof, as they are well-known and may be found in most books on linear programming, for example Gass (15).

Theorem 1.1: The set of all feasible solutions to the linear programming program forms a convex set.

Proof: Suppose $z^{(1)}$ and $z^{(2)}$ both satisfy $Bz^{(1)} = e$, $Bz^{(2)} = e$, $z^{(1)} \geq 0$, $z^{(2)} \geq 0$. Let $z = \alpha z^{(1)} + (1-\alpha)z^{(2)}$, $\alpha \geq 0$. Then $Bz = B[\alpha z^{(1)} + (1-\alpha)z^{(2)}] = \alpha Bz^{(1)} + (1-\alpha)Bz^{(2)} = \alpha e + (1-\alpha)e = e$. This proves the theorem, since we have shown that a convex combination of two feasible solutions is
also a feasible solution. It is clear that $z \geq 0$.

Theorem 1.2: The objective function assumes its maximum at an extreme point of the convex set generated by the set of feasible solutions to the programming problem. If it assumes its maximum at two distinct extreme points of the set, then it assumes its maximum at any convex combination of those two points.

This theorem, of course, states that if we wish to solve a deterministic linear programming problem, we need only to examine the extreme points of the set of feasible solutions.

Let us now express the columns of the matrix $B$ as vectors $P_1, P_2, \cdots, P_{n+m}$, where

$$
P_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad P_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots, P_{n+1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad P_{n+m} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}
$$

and so forth. A feasible solution is then a vector $z \geq 0$ such that

$$
z_1 P_1 + z_2 P_2 + \cdots + z_n P_n + \cdots + z_{m+n} P_{n+m} = e.
$$

we now state the following crucial theorem.

Theorem 1.3: Let $k$ be the convex set of feasible solutions to the programming problem. If $z = (z_1, z_2, \cdots, z_{m+n})$ is an extreme point of $k$ then the vectors $P_j$ associated with positive
z_j form a linearly independent set. From this it is clear that at most m of the z_j are positive. Conversely, associated with every extreme point of K is a set of m linearly independent vectors from the set P_1, P_2, ..., P_{m+n}.

We may thus conclude this discussion with the following observations: If a deterministic linear programming problem has a solution, it is assumed at an extreme point of the set K of feasible solutions. Hence we need examine only extreme points. Further, every extreme point of K has m linearly independent vectors of the set of m+n vectors formed by the columns of the matrix B associated with it. Thus, to solve a deterministic problem, we need merely consider all \( \binom{m+n}{m} \) possible combinations of column vectors arranged to form a matrix. If these vectors are linearly independent, the positive z^0 in the representation \( z_1P_1 + z_2P_2 + \cdots + z_{m+n}P_{m+n} = e \) form an extreme point to be examined for optimality of the objective function. If the chosen vectors are not linearly independent we discard the selection. This technique is called the complete description method, or method of selections.

Duality in Deterministic Programs

The following definitions and theorems are well known indeed, but we rely on them heavily in what follows, and they are therefore stated here.
Definition 1.4: A general (deterministic) linear programming problem consists of an m x n matrix \((a_{ij}) = a\), an m-tuple \(b' = (b_1, \ldots, b_m)\) and an n-tuple \(c' = (c_1, \ldots, c_n)\), where \(b\) and \(c\) are column vectors. The primal problem is then defined as the following: let \(X\) be the set of all n-tuples \(x = (x_1, \ldots, x_n)\) such that

1) \(x_j \geq 0, j = 1, \ldots, n\) and

2) \(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1, i = 1, \ldots, m\).

Find those \(x\) in \(X\) such that \(c_1x_1 + c_2x_2 + \cdots + c_nx_n\) is a maximum.

The dual problem is then defined as the following: let \(Y\) be the set of all m-tuples \(y = (y_1, y_2, \ldots, y_m)\) such that

1') \(y_i \geq 0, i = 1, \ldots, m\) and

2') \(a_{11}y_1 + a_{21}y_2 + \cdots + a_{n1}y_n \geq c_1, j = 1, \ldots, n\).

Find those \(y\) in \(Y\) for which the linear form \(b_1y_1 + b_2y_2 + \cdots + b_my_m\) is a minimum.

We now state, for future reference, the well-known duality theorem.

Theorem 1.4: If \(X\) and \(Y\) are both non-empty (i.e. if feasible solutions exist to both problems) then

\[
\sum_{i=1}^{m} b_iy_1 \geq \sum_{j=1}^{n} c_jx_j.
\]

If \(\sum_{i=1}^{m} b_iy_1\) achieves a minimum, then \(\sum_{j=1}^{n} c_jx_j\) achieves a maxi-
mum, and conversely. Furthermore, if $x^0$ is such that
\[ \sum_{j=1}^{n} c_j x_j^0 \geq \sum_{j=1}^{n} c_j x_j \text{ for all } x = (x_1, \ldots, x_n) \text{ and } y^0 = (y_1^0, \ldots, y_m^0) \text{ is such that } \sum_{i=1}^{m} b_i y_i^0 \leq \sum_{i=1}^{m} b_i y_i \text{ for all other } y \text{ in } Y, \text{ then } \sum_{j=1}^{n} c_j x_j^0 = \sum_{i=1}^{m} b_i y_i^0.

Again, a proof may be found in any good book on linear programming, e.g. Gass (15).

**Topological Preliminaries**

As most of the results in this work have been arrived at by means of the definition of a particular topological space and the application of topological methods, it seems well advised to include in the introduction some elementary topological concepts.

**Definition 1.5:** A topological space is a collection $S$ of undefined objects called "points" together with a collection of subsets of $S$ called "open sets" which satisfy the following axioms:

1. Every open set is a set of points.
2. The empty set is an open set.
3. For every point $p$ in $S$, there exists at least one open set $V$ containing $p$.
4. The union of any collection of open sets is an open set.
5. The intersection of any finite number of open sets is an open set.

Definition 1.6: A set $S$ of points is said to be a metric set if and only if there is associated with $S$ a mapping $\rho: S \times S \rightarrow \mathbb{R}$ (where $\mathbb{R}$ is the real line together with its usual topology and $S \times S$ is, of course, the Cartesian product of $S$ with itself) which has the following properties:

Let $x$, $y$, and $z$ be points of $S$. Then

1. $\rho(x, y) \geq 0$.
2. $\rho(x, y) = 0$ if and only if $x = y$.
3. $\rho(x, y) = \rho(y, x)$.
4. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.

Property 4 is called the "triangle inequality". The mapping $\rho$ is called the metric for the space $S$.

Definition 1.7: Let $S$ be a metric set. Then, with every point $p$ of $S$ and every real number $r > 0$ we associate a subset $N_r(p)$ of $S$ called the spherical neighborhood of $p$ with radius $r$. A point $q$ of $S$ is in $N_r(p)$ if and only if $\rho(p, q) < r$.

Definition 1.8: A metric set $S$ is said to be a metric space if and only if the topology of $S$ is that which is generated by the collection of all spherical neighborhoods in $S$. This topology for $S$ is said to have been induced by the metric $\rho$ defined on $S$. 
Definition 1.9: Let $S$ be a set and let $\sigma$ be a collection of subsets of $S$. Then $\sigma$ is said to generate the collection $\mathcal{T}$ of subsets of $S$ defined as follows: a subset $K$ of $S$ is an element of $\mathcal{T}$ if and only if $K$ is the union of a collection of elements of $\sigma$. The collection $\sigma$ is said to be a basis for the collection $\mathcal{T}$ which it generates.

Definition 1.10: A collection $\sigma$ of neighborhoods of a point $p$ in a space $S$ is said to be a basis at $p$ if and only if, given any neighborhood $U$ of $p$ there exists a $V$ belonging to $\sigma$ such that $V \subseteq U$.

Definition 1.11: A space $S$ is said to be first countable if and only if there exists a countable basis at $p$, for every $p$ in $S$.

Theorem 1.5: Every metric space is first countable.
Proof: The collection $N_{1/n}(p)$ of neighborhoods of radius $1/n$ of $p$ clearly form a countable basis at $p$.

Definition 1.12: Let $S$ and $T$ be spaces and $f: S \to T$ a mapping of $S$ into $T$. Then $f$ is said to be continuous at a point $s$ of $S$ if and only if, given any open subset $G$ of $T$ such that $s$ belongs to $f^{-1}(G)$, there exists an open set $V$ contained in $S$ such that $s$ belongs to $V \subseteq f^{-1}(G)$.

Definition 1.13: A mapping $f$ of $S$ into $T$ is said to be continuous on $S$ if and only if $f$ is continuous at every point of
The following theorem is stated without proof.

**Theorem 1.6:** Let $S$, $T$ and $W$ be spaces, and let $f: S \to T$ and $g: T \to W$ be continuous mappings. Then the composition mapping $h: S \to W$ defined by $h(x) = g(f(x))$ for every $x$ in $S$ is continuous.

**Definition 1.14:** A sequence is a function having as domain the positive integers.

**Theorem 1.7:** Let $S$ and $T$ be spaces and $f: S \to T$ a continuous mapping. Let $\{x_n\}$ be a sequence of points in $S$ which converges to a point $x$ in $S$. Then the sequence $\{f(x_n)\}$ of points in $T$ converges to $f(x)$.

A last few definitions and theorems conclude the section on topological preliminaries.

**Definition 1.14:** Let $\{X_\alpha\}_{\alpha \in \Delta}$ be a collection of sets designated by an index set $\Delta$. Then the Cartesian product $\prod_{\alpha \in \Delta} X_\alpha$ of the sets of this collection is defined to be the set of all mappings $f: \Delta \to \bigcup_{\alpha \in \Delta} X_\alpha$ such that $f(\alpha)$ is an element of $X_\alpha$ for all $\alpha$ in $\Delta$.

**Theorem 1.7:** Let $\{X_\alpha\}_{\alpha \in \Delta}$ be a collection of spaces indexed by $\Delta$. Denote by $\sigma$ the collection of all subsets of $\prod_{\alpha \in \Delta} X_\alpha$ that are of the form $\prod_{\alpha \in \Sigma} Y_\alpha$ where, for some finite subset $\Sigma$ of $\Delta$, $Y_\beta$ is
an open subset of $X_\beta$ for every $\beta \in \Sigma$, and $Y_\alpha = X_\alpha$ for all $\alpha$ in $\Delta - \Sigma$. Then $\sigma$ is a basis for a topology for the product set $\prod_\alpha X_\alpha$.

Definition 1.15: Let $\{X_\alpha\}_{\alpha \in \Delta}$ be a collection of spaces indexed by $\Delta$. The product space $\prod_\alpha X_\alpha$ is then the set $\prod_\alpha X_\alpha$ with the topology having as a basis the collection $\sigma$.

Stochastic Linear Programming

A linear programming problem is said to be stochastic, or is called a risk programming problem, if one or more of the parameters in the problem is known only as to its statistical distribution. It is, we believe, the case that almost any conceivable linear programming problem derived from a practical need will necessarily be of a stochastic nature. In production planning, for example, future demand and prices of output are known only through forecasts which are necessarily subject to error, and the prices of inputs, except where contracted ahead, will also be subject to deviation from their most probable value. Even the input-output coefficients, due to the vagaries of mechanical contrivances and the nature of their human operators, may be subject to deviation from an expected norm. The application of linear programming in production planning, allocation, transportation, etc., still is accomplished in many circles by the expedient of utilizing the
means or medians of compiled relevant data and handling the program as though it were completely deterministic. While undoubtedly this is better than nothing and a planning or inquiring organization may be content merely to be in the same ballpark, so to speak, with the desired objectives, especially if they are to be ascertained experimentally, we feel that a much more complete and informative analysis of any situation involving linear programming can be obtained by using a technique, i.e., stochastic programming, which does not obscure the statistical nature of the data.

Stochastic programming as a study is still in its infancy. It was born no earlier than the middle 1950's, and, as with any new field, the workers in the area have created some rather divergent approaches to the field. We shall be concerned with only one of these many basic approaches, that of Professor Tintner (36, 37, 39, 41), and while the others are described briefly in the next chapter, no attempt is made here to concern the reader with them.

Let \( A = (a_{ij}) \) be an \( m \times n \) matrix whose elements are random variables with known probability density functions. Let

\[
\begin{pmatrix}
    b_1 \\
    \vdots \\
    b_m
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
    c_1 \\
    \vdots \\
    c_n
\end{pmatrix}
\]

be column vectors whose elements are also random variables with known distributions, and let us further assume that the elements of \( A, b \) and \( c \) are mutually
and jointly independent. Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ be a column vector.

Then a maximization passive stochastic linear program is defined by

$$\text{maximize } c'x \text{ subject to }$$

$$Ax \leq b$$

$$x \geq 0.$$  

We transform the constraints to equalities, to get

$$\text{maximize } d'z$$

$$Bz = e, \ z \geq 0,$$  

where

$$z = \begin{pmatrix} x_1 \\ \vdots \\ x_m \\ x_{m+1} \\ \vdots \\ x_{n+m} \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix},$$

$$d = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \ e = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \text{ and } B = \begin{pmatrix} a_{11} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Using the method of selections, we form all $\binom{m+n}{m}$ combinations of the columns of $B$ to get $m \times m$ matrices, one for each
selection. Discarding all those whose determinants are identically zero, i.e., all those whose columns are not linearly independent, we proceed to find the objective function for each selection. Since there are only a finite number of columns in the matrix B, there are at most a finite, although perhaps a large, number of steps.

Recall now that the elements of A, b, and c are random variables. Hence a parameter space is formed which consists of all possible variations in the parameters \( a_{ij}, b_i, c_j \).

Let the objective function be designated as \( F \). Then for each of the \( \frac{(m+n)!}{m!n!} \) selections we get values for the variables \( z_1, \ldots, z_{m+n} \) and a corresponding objective function. Index the selections by \( k \); \( k = 1, \ldots, \frac{(m+n)!}{m!n!} \), thus designating the values of the variables for the \( k \)th selection as \( z_1^{(k)}, z_2^{(k)}, \ldots, z_{m+n}^{(k)} \), and designating the corresponding objective function by \( F^{(k)} \).

For each selection \( k \) find the regions in the parameter space where the selection is feasible, i.e., where the \( z_1^{(k)} \geq 0 \) and where the selection is optimal, i.e., where \( F^{(k)} > F^{(\overline{k})} \), where \( \overline{k} \) is any other selection. The intersection of these regions of feasibility and those of optimality then are what is sought, and the probability density function of the objective function may be approximated.

This is the so-called "passive" approach. The "active"
approach, in which the entrepreneur, analyst or programmer optimizes with respect to "decision variables", is the following: for \( i, j = 1, 2 \), let \( a_{ij}, b_i \) and \( c_j \) be random variables with known probability density functions, and define "decision variables" \( u_{ij} \) which may assume values between 0 and 1. Then an active programming representation may be given by

\[
\begin{align*}
\text{maximize } & c_1x_1 + c_2x_2 \\
\text{subject to } & a_{11}x_1 \leq b_1u_{11} \\
& a_{12}x_2 \leq b_1u_{12} \\
& a_{21}x_1 \leq b_2u_{21} \\
& a_{22}x_2 \leq b_2u_{22} \\
& x_1 \geq 0, \ x_2 \geq 0. \\
& u_{11} + u_{12} = 1, \ u_{21} + u_{22} = 1.
\end{align*}
\]

For each distinct value of \( u_{11} \) and \( u_{22} \), a different program is determined. The problem determined by each value of \( u_{11} \) and \( u_{12} \) may be solved using the selections technique, and the analyst may optimize then with respect to the decision variables. As a detailed discussion of active stochastic programming is forthcoming later, more is not said at this time.

Intuitively, it should be quite apparent what the term "parameter space" has reference to. We would, however, like to formalize its mathematical structure somewhat so as to endow it with some properties that will be highly desirable later.
It is easy to show that the real line $\mathbb{R}$ with its usual metric $\rho(x,y) = \sqrt{x^2 - y^2}$ for any two points $x, y$ in $\mathbb{R}$ forms a topological space with a topology induced by $\rho$ that is first countable. We shall soon show that the topological product of $\mathbb{R}$ with itself any finite number of times is also first countable. But first it is in order to provide some motivation for this seemingly impertinent digression.

Consider a set of $n$ continuously distributed and mutually independent random variables $y_1, y_2, \ldots, y_n$ having each a probability density function $f_i(y_i), i = 1, \ldots, n$. These are real-valued functions of a real variable. These random variables have, moreover, a joint probability density function $g(y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} f_i(y_i)$ having as domain the cartesian product of the real line $\mathbb{R}$ with itself $n$ times. Let us index $n$ copies of the real line, and express this domain as $\prod_{i=1}^{n} \mathbb{R}_i$.

It should now be apparent that the parameter space mentioned earlier is, if we have an $m \times n$ matrix and an $m$-vector and an $n$-vector defining the elements of our stochastic programming problem, precisely the cartesian product of $\mathbb{R}$ with itself $(m \times n) + m + n$ times. It at least satisfies one's intuition on the subject and does, we believe, justify the following definition.

Definition 1.16: Let a stochastic programming problem be defined by an $m$-vector, an $n$-vector, and an $m \times n$ matrix, each
of whose elements are random variables with known and mutually independent probability density functions. We then define the parameter space associated with the problem to be the cartesian product of the real line $\mathbb{R}$ with itself $(m \times n) + m + n$ times.

We thus find the parameter space described by an $(m \times n) + m + n$-dimensional rectangular coordinate system, and it is in order to consider the topological structure of the parameter space. Let us, for the sake of simplicity, denote this parameter space by $S$.

**Theorem 1.8:** Let $\sigma$ denote the collection of all subsets of $S$ having the form $\cap_{i=1}^{(m \times n) + m + n} \delta_i$, where $\delta_i$ is an element of a basis for $R_i$. Then $\sigma$ forms a basis for $S$.

**Proof:** The proof follows from Theorem 1.7 and the fact that $S$ is the cartesian product of a finite number of spaces.

The following theorem forms a crucial underpinning for the remainder of this work.

**Theorem 1.9:** $S$ is first countable.

**Proof:** We must show that given any point $a$ in $S$, we can find a countable basis at $a$. Let $a$ be specified by the ordered $n$-tuple $(y_1, \ldots, y_{m \times n + m + n})$, $y_i \in R_i$, and consider the $(m \times n) + m + n$ projection mappings $f_i(a) = y_i$. Since $R_i$ is first countable, we can find for all $i$ a countable basis $u_{y_{i,k}}$, $k$ a
positive integer, of $y_i$. The collection $\bigcap_{i=1}^{\infty} I_n k, k = 1, 2, \ldots$, then forms a basis at $a$, since it clearly satisfies the requisite definition and it is countable, since countable collections of countable sets are countable.

Definition 1.17: A topological space $T$ is called an I space if and only if it is first countable and each of its points is a closed set.

Theorem 1.10: Every metric space $T$ is an I space.

As a corollary to the above theorem we have the following result.

Theorem 1.11: The parameter space $S$ is an I space.

The next theorem concludes this chapter.

Theorem 1.12: Let $T$ be a subset of an I space $S$. Then a point $p$ of $S$ is a limit point of $T$ if and only if there exists a sequence of distinct points of $T$ that converges to $p$. 
PREVIOUS RESULTS IN STOCHASTIC PROGRAMMING

New as it is, stochastic programming has captured the interest of perhaps a dozen well-known scholars. Each seems to have his own approach to the study of linear programming problems with one or more variable parameters, and several have found stochastic programming formulations for problems that had been classically dealt with through the use of entirely different techniques.

It is our intention to present in this chapter some illustrations of other work done in this general area of stochastic programming. The contents of this chapter represent, to the best of our knowledge, a quite complete survey of the literature on the theory of stochastic programming, and are included partially in order that the reader may contrast the approach used in this work, which is due to Professor Tintner (38, 39, 41, 43), with those of other workers in the field.

Danzig (9) considers linear programming problems in which allocations in a first "stage" are made to meet the uncertain demands with "known" distributions of a second stage. He gives the following examples, among others.

1. A nutrition expert must advise on a minimum cost diet without knowledge of the prices and costs involved, but with the assumption of the knowledge of a distribution of possible
prices. Let $x_j$ be the quantity of food $j$, $p_j$ be its price, and let $a_{ij}$ be the quantity of the $i^{th}$ nutrient contained in a unit quantity of food $j$. Let $b_i$ be the minimum good-health requirement of nutrient $i$. The stochastic programming problem is then to minimize $E(c) = \sum_j \bar{p}_j x_j$

subject to

$$\sum_{j=1}^n a_{ij} x_j \geq b_i, \quad x_j \geq 0 \quad \text{and} \quad \bar{p}_j = E(p_j).$$

2. A factory has on hand 100 items which may be shipped to an outlet at a cost of $\$1$ each, to meet an uncertain demand $d_2$. If demand should exceed supply, it is assumed necessary to meet unsatisfied demand by purchases on the local market at $\$2$ each.

Let $x_{11} =$ number shipped from the factory; let $x_{12} =$ number stored at the factory, $x_{21} =$ number bought on the local market, $x_{22} =$ excess of supply over demand, $d_2 =$ unknown demand uniformly distributed between 70 and 80, and let $c =$ total costs.

The system must then satisfy

$$100 = x_{11} + x_{12},$$

$$d_2 = x_{11} + x_{21} - x_{22},$$

$$c = x_{11} + 2x_{21},$$

with

$$x_{1j} \geq 0, \quad 1, \quad j = 1, \quad 2.$$
assigning the resources to the destinations $j$, and the costs (or lost revenues) incurred due to failure of the amounts $u_1$, $u_2$, $\cdots$, $u_n$ assigned to meet the unknown demands $d_1$, $d_2$, $\cdots$, $d_n$.

The author sets up the following two-stage structure:

For the first stage, 
\[ \sum_{j=1}^{n} x_{1j} = a_1, \quad x_{1j} \geq 0, \quad j = 1, \cdots, n, \]
and
\[ \sum_{j=1}^{m} b_{1j} x_{1j} = u_j, \]
where $x_{1j}$, again, represents the amount of resource $i$ assigned to destination $j$, and $b_{1j}$ represents the number of units of demand at destination $j$ that can be satisfied by one unit of resource $i$.

For the second stage, 
\[ d_j = u_j + v_j - s_j, \quad (j = 1, \cdots, n), \]
where $v_j$ is the excess of demand over supply and $s_j$ is the excess of supply over demand.

The total cost function is then
\[ c = \sum_{j=1}^{n} c_{1j} x_{1j} + \sum_{j=1}^{m} a_j v_j. \]

The objective is to minimize expected cost.

Let $\mathcal{O}_j(u_j | d_j) = \text{minimum costs at destination } j$. Then
\[ \mathcal{O}_j(u_j | d_j) = \begin{cases} a_j(d_j - u_j) & \text{if } d_j \geq u_j \\ 0 & \text{if } d_j < u_j. \end{cases} \]

The author shows that
\[ \mathbb{E}(c) = \sum_{j=1}^{n} c_{1j} x_{1j} + \sum_{j=1}^{m} a_j \mathcal{O}_j(u_j), \]
where
\[ \phi_j'(u_j) = \frac{\Sigma \phi_j(u_j \phi_j)}{\phi_j(u_j \phi_j)} = a_j \int_{x=u_j}^{\infty} (x-u_j) \sigma(x) dx, \]

and \( p(d_j) \) is the probability density function of \( d_j \). He shows that the \( \phi_j(u_j) \) are convex functions and from this result arrives at an approximation scheme for minimum costs.

Ferguson and Danzig (12) in a later paper apply these results to the problem of allocating aircraft routes under the assumption of uncertain demand.

Another worker in the area of stochastic programming, Freund (14) considers a programming problem of the following form.

Maximize \( s'x \) subject to \( x \geq 0, Tx \leq v \), where \( s, x \) and \( v \) are column vectors and \( T \) is a matrix. In this example, the objective function \( s'x \) represents net revenue.

Assuming the elements of \( T \) to be fixed and that the net revenue of each output \( i \) is normally distributed: \( s_i \sim N(v_i, \sigma_i^2) \) with covariances between the net revenue of two processes defined as \( \sigma_{ij} \), the net revenue, \( r \), will also be normally distributed: \( r \sim N(v'x, x' \Sigma x) \), where \( \Sigma \) is the variance-covariance matrix for the \( s_i \).

The author then assumes a utility function of the form \( y(r) = 1 - e^{-ar} \) and, treating the problem as one of maximizing expected utility

\[ E(u) = \int_{-\infty}^{\infty} (1 - e^{-ar})(e^{-\frac{(r-u)^2}{2\sigma^2}}) dr, \]
he demonstrates that this can be accomplished by maximizing
\[ E(u) = s'x - \frac{1}{2} x' \leq x \text{ subject to } Tx \leq v, \ x \geq 0 \] which is a problem in quadratic programming.

Madansky (27) considers a one-stage problem similar to that described by Danzig (9) but written in a somewhat different form. He states a problem as: minimize \( c'x + b'y \) (with respect to \( x \) and \( y \)) subject to \( Ax + By = b, \ x \geq 0, \ y \geq 0, \) where \( A \) and \( B \) are known \( m \times n_1 \) and \( m \times n_2 \) matrices respectively, and \( b, c, \) and \( f \) are \( m, n_1, \) and \( n_2 \)-dimensional column vectors.

Denote by \( c(b, x) \) the function \( \min_y (c'x + f'y) \), for a given \( x \). The author then proves two results: first, that \( \min_x c(b, x) \) is a convex function of \( b \), and secondly, that \( \min_x c(b, x) \) is a continuous function of \( b \).

Suppose now that \( b \) is a vector whose elements are stochastic. Let \( \overline{x}(Eb) \) be that vector which minimizes \( c(Eb, x) \), \( \overline{x} \) be that vector \( x \) which minimizes \( Ec(b, x) \), and let \( \overline{x}(b) \) be that vector \( x \) which minimizes \( c(b, x) \). The author then demonstrates that \( Ec[b, \overline{x}(Eb)] \geq \min_x Ec(b, x) \geq E \min_x c(b, x) \geq \min_x c(Eb, x) \), and points out that a sufficient condition that \( \min_x Ec(b, x) = E \min_x c(b, x) \) is for \( c(b, x) \) to be linear in \( b \). On the other hand, \( Ec(b, \overline{x}(Eb)) \) and \( \min_x c(Eb, x) \) are easily-computed upper and lower bounds for both \( \min_x Ec(b, x) \) and \( E \min_x c(b, x) \).

The remainder of his paper is taken up with the presentation of a technique to sharpen these bounds, and an applica-
tion of his results to Danzig's two-stage example (9).

Other authors and workers have also considered the problem of finding bounds for the objective function of a stochastic linear program. Another school of thought on this subject is exemplified by the work of Talacko (35) who presents the following approach.

Let a programming problem be given as that of maximizing $c'x$ subject to $Ax \leq b$ and $x \geq 0$, where $A$ is an $m \times n$ matrix, $b$ is an $m$-tuple, and $c$ is an $n$-tuple. Assume, moreover, that $a_{ij}$, $b_i$, and $c_j$, the elements of $A$, $b$ and $c$ respectively, are statistics $S$: $a_{ij}^-$, $b_i^-$, $c_j^-$, with known intervals of interest $a_{ij}^- \leq a_{ij} \leq a_{ij}^+$, $b_i^- \leq b_i \leq b_i^+$, and $c_j^- \leq c_j \leq c_j^+$, and let these intervals have associated probabilities $p(b_i^-)$, $p(c_j^-)$ and $p(a_{ij}^-)$.

Represent by $S_i^-$ and $S_i^+$ the lower and upper bounds for these statistics. The author then proceeds to find bounds for the objective function $F(x) = \sum_{j=1}^{n} c_j x_j$ in this manner.

Let $W$ be the set of all $x$ satisfying $Ax \leq b$, $x \geq 0$, for the various possible values of $a_{ij}$, $b_i$, and $c_j$ over the previously described intervals of interest. He then proves two lemmas: first, that there exist vectors $u$ and $u^*$ in $W$ such that $F(u^*) \leq F(x) \leq F(u)$, and secondly, that if $W$ corresponds to any fixed set of intermediate values of the $a_{ij}$ and $c_i$, then $W^- \subset W \subset W^+$, where $W^-$ and $W^+$ correspond to the lower and upper bounds, respectively, of the intervals of interest.

On the basis of these two lemmas, he proves the following
Theorem: Let \( \max F(x) \) correspond to a problem involving any intermediate values of the statistics \( S_i \). Let \( \max^+ F^+(x) \) denote \( \max F^+(x) \) over \( W^+ \), and let \( \max^- F^-(x) \) be the maximum of \( F^-(x) \) over \( W^- \). Then \( \max^+ F^+(x) \leq \max F(x) \leq \max^- F^-(x) \).

The upper and lower limits of the objective function are then found, we see, by solving two programming problems:

\[
\max \sum_{j=1}^{n} c_i x_i \quad \text{subject to} \quad \sum_{j=1}^{n} a_{ij} x_i \leq b_i, \quad i = 1, \ldots, m, \quad \text{and} \quad x \geq 0, \max \sum_{j=1}^{n} c_i^+ x_i \quad \text{subject to} \quad \sum_{j=1}^{n} a_{ij}^+ x_i \leq b_i^+, \quad i = 1, \ldots, m
\]

and \( x \geq 0 \).

He then closes by giving several numerical examples illustrating his technique.

Another interesting work is that done by Elmaghraby (11), who has extended the allocation problem for the deterministic case to a situation involving uncertain demand with a continuous probability density function.

The problem of allocation under uncertain demand can be represented as follows:

Minimize \( c = f(X_{ij}), \quad i = 1, \ldots, m, \quad j = 1, \ldots, n, \quad \text{subject to} \quad g_i(X_{ij}) \leq a_{ij} \quad \text{and} \quad X_{ij} \geq 0, \) where \( f \) is a continuous function of the \( X_{ij} \), and the \( g_i \) are of the form \( \sum_{j} X_{ij} \leq a_i \).

This may be reformulated in the following language of production and inventory.

Let \( A_1, \ldots, A_m \) be a set of facilities with limited capacities \( a_1, \ldots, a_m \). These facilities are to be utilized
in the production of goods $P_1, \ldots, P_n$. Let $f_j(y)$ represent the probability density function of the demand for the $j$th product, and let $F_j(y)$ be its cumulative density function.

Let $r_{ij}$ be the productivity of resource $i$ when producing item $j$; let $c_{ij}$ be the cost of manufacturing a unit of product $j$ using activity $i$; $c_{0j}$, the storage cost per time unit of one unit of item $j$; $c_{uj}$, the cost of understorage per unit of time of one unit of item $j$; $S_j$, the total amount produced per time unit of item $j$, and $x_{ij}$, the amount of resource $i$ devoted to product $j$ per time unit.

It is clear that $\sum_{i} r_{ij} x_{ij} = S_j, j = 1, \ldots, m, i = 1, \ldots, n$, and $\sum_{j} x_{ij} \leq a_i$.

The problem is then to minimize the expected total cost of production

$$C = \sum_{i} \sum_{j} c_{ij} r_{ij} x_{ij} + \sum_{j} c_{0j} \int_0^{S_j} (S_j - y)f_j(y)dy$$

$$+ \sum_{j} c_{uj} \int_{S_j}^{\infty} (y - S_j)f_j(y)dy.$$

The author then proves as a lemma that the functions $c(X)$ and $h_1(X)$ are convex in $X$, where $h_1(X) = a_1 - \sum_{j} x_{ij} \geq 0$.

His main theorem is this:

For a feasible solution (allocation matrix) $X^0$ to be optimal, it is necessary that

a) if the inequality $h_1(X^0) > 0$ holds for some $i$, then the partial derivative of $c$ with respect to each component of
$X^0$, which is positive and is also involved in the equality must vanish;

b) if the equality $h_1(X^0) = 0$ holds for some $i$, then the partial derivative of $c$ with respect to all the components of $X^0$ which are positive and are also involved in the equality must have the same value $v_i$, where $v_i \geq 0$.

These two conditions together with the following two are sufficient for optimality.

c) In case (a), the partial derivative of $c$ with respect to each component of $X^0$ which is equal to zero and is involved in the equality must be non-negative;

d) In case (b), the partial derivative of $c$ with respect to all the components of $X^0$ which are equal to zero and are involved in the same equality must be greater than or equal to $v_i$.

Having proved this theorem and giving conditions necessary and sufficient for optimality, the author presents a sequence of decisions in chart form which, taken together, have the effect of an algorithm. The basic ideas guiding the construction of the chart may be summarized in the following two statements:

1. For any fully-utilized capacity, it is the difference among the derivatives of the allocations in the same row, and not their absolute values, that are important, while for non-fully utilized capacities, it is the magnitude of the alge-
braically smallest derivative that matters.

2. A change in allocation in any one cell affects the allocation in all the cells in the same column, and at least one cell in the same row for fully-utilized capacities.

The author closes with a numerical example illustrating the application of his theory to the aircraft allocation problem discussed by Danzig and Ferguson (12), together with a comparison of their results with his.

Simon (34) has published some results in which he concerns himself with uncertainty in the case of a dynamic program with a quadratic function. He has been able to show that the problem for the case of uncertainty can be reduced to the problem for the case of certainty merely by replacing, in the computation for the optimal "first period", the future values of the variables by their expectations. Toward this end, he considers the three kinds of information a decision-maker might possess about the future values of certain variables relevant to his decision. These may be listed as follows:

a) He might know these future values with certainty;

b) He might know their unconditional expected values;

c) He might know the joint probability distributions of the variables over the whole sequence of future time periods.

The initial task of the decision-maker is to determine his course of action for the first time period. Having done this, and on the basis of the new information now available to
him, he can choose a course of action for the second time period, and so on.

The author then discusses two different approaches to be considered in light of the available data.

The first of these is the "certainty-equivalent" method. If the decision-maker knows (c), he can compute the expected values of (b) and, if he so chooses, behave as if these expected values were (unknown) certain values, as in (a). This transforms the problem into one of dynamic programming under certainty. At the end of the first period, he has new initial conditions and a new joint probability distribution (c) from which he can obtain new expected values (b), and so on. The author notes that the only prerequisite to application of this technique is knowledge of (b).

The so-called "general programming" method, on the other hand, requires a knowledge of (c). Knowing (c), he can determine an optimal action for the first period, carry out this action, and replace for the second period using this same technique.

The author mentions that, by definition, if the planner has only (c) as information, no planning procedure can yield a higher expected value of the criterion function than the general programming method, or yield a lower value if the planner is minimizing.

Simon's main result is then the following: the certainty-
equivalent method will lead to exactly the same prescription of action as the general programming model and that the former, requiring only knowledge ($b$), represents a solution to the problem of dynamic programming under uncertainty with a quadratic criterion function.

His proof does not require independence of the stochastic variable, but does assume that the criterion function is a quadratic form in the variables and their derivatives.

Another, still different, approach to the area of stochastic programming is described by Vajda (46), who presents several interesting results covering situations in which only the coefficients of the objective function are stochastic, and secondly, where the right-hand sides of the constraints are stochastic as well.

He describes his problem in the following form:

minimize $c'x$ subject to

$$Ax = b, \ x \geq 0.$$ 

Let the solution to this deterministic problem be $x_0$ and let $c'x_0 = M$.

Consider now the problem described by the following:

minimize $c'x$ subject to $Ax = b + \beta, \ x \geq 0$, where $\beta$ is a random vector taking on the form $b_t$ with probability $p_t$, and $E\beta = 0$.

Let the solution of the problem whose constraints are $Ax = b + b_t, \ x \geq 0$, be $x_{0t}$, and denote $c'x_{0t}$ by $M_t$. He then proves, using the duality theorems, that $EM_t = \Sigma p_t M_t \geq M$. 
In the second type of problem mentioned above, he discusses a problem that is represented by the following:

\[
\text{minimize } E(c'x + f'y_t), \quad \text{where } Ax + y_t = b + b_t, \quad Eb_t = 0, \quad \text{and } x \geq 0 \text{ form the constraints, and } f \text{ and } y_t \text{ are column vectors.}
\]

If the elements of \( y_t \) are not assumed non-negative, then the objective function can be written as the following:

\[
\text{minimize } E[c'x + f'(b + b_t - Ax)] = E[c'x + f'(b - Ax)]
\]

with no constraints other than \( x \geq 0 \).

If the minimum is denoted by \( M_t \), he then demonstrates that \( M_t \leq M \). This result remains valid if, for example, it is required that \( y_t \geq 0 \).

His final result is a demonstration that, if \( y_t \geq 0 \), then, dependent on \( f \), the minimum \( \bar{M} \) might be larger or smaller than \( M \).

Votaw (49) presents an approach which diverges rather thoroughly from those of other workers in the field. He begins, conventionally enough, by defining statistical programming as linear programming in which the information about one or more of the parameters involved is statistical in nature. He presents as examples the transportation problem and the personnel-assignment problem, which he finds to be equivalent mathematically except that the variables must assume only the values 0 or 1 in the latter.

The transportation problem has the following form. Let
$c_{ij}$ represent the cost of shipping a unit of the product from origin $i$ to destination $j$, and let $x_{ij}$ be the amount thus shipped. Find an $n \times n$ matrix $(x_{ij})$ of real numbers for which

$$\sum_{i,j=1}^{n} c_{ij} x_{ij}$$

is a minimum, subject to

$$\sum_{i=1}^{n} x_{ij} = \sum_{j=1}^{n} x_{ij} = 1,$$

$$x_{ij} \geq 0.$$ 

It is now apparent that the admissible program matrices $(x_{ij})$ are just the $n!$ permutation matrices of order $n \times n$.

Let $(j_1^i, \ldots, j_n^i)$ denote any permutation of $(1, \ldots, n)$. Find that permutation such that the objective function is a minimum.

The problem becomes statistical if one considers the following.

Let $\mathbb{W}^*$ be an $n^2$-dimensional Euclidean space with points $(w_{11}, \ldots, w_{nn})$. Let $H(w_{11}, \ldots, w_{nn})$ be the probability density function associated with $\mathbb{W}^*$, and let $(w_{11}, \ldots, w_{nn})$ be a random $n^2$-tuple whose distribution is $H$. Assume that the $c_{ij}$ are parameters of $H$ whose numerical values are unknown. Assume also that an observation $(w_{11}', \ldots, w_{nn}')$ of $(w_{11}, \ldots, w_{nn})$ can be obtained. Having obtained this observation, the programmer selects a permutation of $(1, \ldots, n)$. Statistical programming then occurs when the analyst partitions the sample space into $n!$ mutually exclusive and exhaustive subsets and establishes a 1-1 correspondence between these subsets and the $n!$ permutations of $(1, \ldots, n)$.

A natural kind of statistical programming is then that in
which one considers estimates of the n! permutation sums and selects that permutation which corresponds to the smallest sum. This is called "programming by estimation".

On the other hand, let \((J^*_1, \ldots, J^*_n)\) be a purely random permutation such that for any chosen permutation, the probability that \((J^*_1, \ldots, J^*_n) = (j_1, \ldots, j_n)\) is \(1/n!\), where \((j_1, \ldots, j_n)\) is the chosen permutation. Let \(c_1J^*_1 + \cdots + c_nJ^*_n\) be the resultant permutation sum and let \(P(s) = \Pr(S \leq s), -\infty < s < \infty\). \(P(s)\) is then a discrete distribution having no more than \(n!\) saltuses. The selection of a value of \((J^*_1, \ldots, J^*_n)\) will be called "purely random programming".

For each \(i, j\), let \(y^*_{ij} = w_{ij} - c_{ij}\). Represent the sample space of the \(y^*_{ij}\) by \(Y\). Let \((y^*_{11}, \ldots, y^*_{nn})\) be any point of \(Y\) and let the distribution of \((y^*_{11}, \ldots, y^*_{nn})\) be \(K(y^*_{11}, \ldots, y^*_{nn})\). Let \((J'_1, \ldots, J'_n)\) be a permutation to be selected under programming by estimation with an observed value of \((w_{11}, \ldots, w_{nn})\). Let \(z = c_1J'_1 + \cdots + c_nJ'_n\) and let \(G(s)\) be the cumulative distribution function of \(z\), i.e., \(G(s) = \Pr(z \leq s), -\infty \leq s \leq \infty\). The author then proves that \(G(s)\) has the same saltus points as \(P(s)\), and that if \(K(y^*_{11}, \ldots, y^*_{nn})\) is continuous and completely symmetric in its variables, then \(G(s) \leq P(s)\) for \(s\) in a certain specified interval.

Babbar (1, 2) has applied some results regarding the distribution of the solution of a set of simultaneous linear equations to a linear programming model.
His linear equations have the form \((B + b)x = (Q + e)\), where \(B\) is an \(m \times m\) non-singular matrix whose elements are known constants, \(b\) is an \(m \times m\) matrix of random errors \(b_{ij}\) such that \(\mathbb{E}(b_{ij}) = 0, \, i, \, j = 1, \ldots, m,\) and \(\mathbb{E}(b_{ij}^2) = \sigma_{ij}^2\). The matrix \((B + b)\) is also assumed non-singular. \(Q\) is an \(m\)-dimensional column vector whose elements \(q_i\) are known constants, and \(e\) is an \(m\)-dimensional column vector of random errors \(e_i\), with \(\mathbb{E}(e_i) = 0\) and \(\mathbb{E}(e_i^2) = \gamma_i^2, \, i = 1, \ldots, m\).

To derive approximate distributions for the variables \(x_1, \ldots, x_m\), he establishes the following notational conventions.

Let \(|B|\) denote the determinant of the matrix \(B\), and let \(|B + b|\) denote the determinant of the matrix \((B_{ij} + b_{ij})\); let \(\varphi_{ij}\) denote the cofactor of the element \(B_{ij}\) in \(|B|\), and let \(|D^k|\) denote the determinant of \((B_{ij})\) with its \(k^{th}\) column replaced by column \(q_i\). Let \(\varphi^k_{ij}\) denote the cofactor of the element in the \(i^{th}\) row and \(j^{th}\) column of \(|D^k|\), and finally, let \(|D^k + d^k|\) denote the determinant of the matrix \((B_{ij} + b_{ij})\) when its \(k^{th}\) column is replaced by the column vector \((q_i + e_i)\).

Then, the solution of the system \((B + b)x = (Q + e)\) is given by

\[
x_k = \frac{|D^k + d^k|}{|B + b|} = \frac{N(x_k)}{D(x_k)}, \quad k = 1, \ldots, m,
\]

and the author addresses himself to the question of finding the distribution of the function
\[ y = \sum_{r=1}^{m} (c_r + c_{-r}) x_r = \frac{1}{|B + b|} \sum_{r=1}^{m} (c_r + c_{-r}) \sigma \text{d}^r + \text{d}^r = \frac{N(y)}{\nu(y)}. \]

He observes that since \( N(x_k) \), \( D(x_k) \), \( N(y) \) and \( D(y) \) are linear functions of normally distributed variables, they themselves will be normally distributed, and the problem becomes one of finding the distribution of the ratios

\[ \frac{N(x_k)}{D(x_n)} \text{ and } \frac{N(y)}{D(y)}. \]

Using a result due to Geary (16), he shows that the probability density of \( x_k \) is given by

\[
f(x_n) = \frac{1}{\sqrt{2\pi}} \left[ \frac{\delta_k^2 - \delta_n \sigma_{Bk}}{\sigma_k^2} + x_k \left( \frac{\delta_n \sigma_{Bk}^2 - \delta_{Bk}}{\sigma_k^2} \right) \right] \cdot \frac{1}{\left[ \sigma_k^2 - 2\sigma_{Bk} x_n + \sigma_{Bk}^2 \right]^{\frac{3}{2}}} \cdot \exp \left\{ -\frac{1}{2} \left[ \frac{(x_n \delta_k - \sigma_k^2)^2}{\sigma_k^2} - 2\delta_{Bk} x_k + \sigma_{Bk}^2 \right] \right\},
\]

where

\( E \left[ N(x_k) \right] = \delta_k \) and \( \text{var} \left[ N(x_k) \right] = \sigma_k^2 \), \( E \left[ N(y) \right] = \delta_y \), and

\( \text{var} \left[ N(y) \right] = \sigma_y^2 \). \( E \left[ D(x) \right] = E(D(y)) = \beta \), \( \text{var} \left[ D(x) \right] = \text{var} \left[ D(y) \right] = \sigma_B^2 \).

The probability density function of \( y \) is given by

\[
f(y) = \frac{1}{\sqrt{2\pi}} \left[ \frac{\beta \sigma_N^2(y) - \delta_y \sigma_{BN}(y)}{\sigma_y^2} + y \left( \frac{\delta_y \sigma_{BN}^2 - \delta_{BN}(y)}{\sigma_y^2} \right) \right] \cdot \frac{1}{\left[ \sigma_y^2 - 2\sigma_{BN}(y) y + \sigma_{BN}^2 \right]^{\frac{3}{2}}} \cdot \exp \left\{ -\frac{1}{2} \left[ \frac{(y \delta_y - \sigma_y^2)^2}{\sigma_y^2} - 2\sigma_{BN}(y) y + \sigma_{BN}^2 \right] \right\},
\]
The author then discusses confidence intervals for the ratios, cumulative distributions, etc., and closes with a numerical application of these techniques to linear programming.

Charnes and Cooper (7) consider a situation in which a company wishes to determine its ship chartering policy in the light of its already-available fleet.

Suppose that charters for ships fall into two classes: spot or short term, and five-year or long-term, and that these are mutually exclusive and exhaustive classes. Suppose further that the demands for shipping capacity are uncertain in that they are known, from period to period, only by statistical distribution with known means and variances, and that the charter rates for each period are similarly known, that is, by distribution. It is required that shipping capacity demands be met at least to the levels of preassigned probabilities.

Assume that there exists a stipulated horizon, or planning interval and a maximal allowance, stated probabilistically, for the number of ship charters which will be in effect on the terminal data. The problem is then to find a sequence of decision rules that are linear in the demand variates and are
also optimal in that they produce minimum total charter costs while honoring the chance constraints.

The problem may be formulated as follows:

Let $D_i$, the amount of shipping demand for period $i$, be normally distributed with mean $\bar{D}_i$ and variance $\sigma_i^2$. Let $L_j$ and $S_j$ be, respectively, the amount of long-term and spot chartering to be undertaken at the beginning of period $j$ ($i, j = 1, \ldots, N$, where $N$ fixes the horizon). Represent the long-term and spot charter rates respectively as $\ell_i(D_1, \ldots, D_i, \ldots, D_N)$ and $s_i(D_1, \ldots, D_N)$ to indicate that the rates $\ell_i$ and $s_i$ may depend, as desired, on past, present, and future demand.

Let $\beta_i$ represent the minimum probability specified for honoring a constraint in period $i$. The problem may then be stated as

$$\text{minimize } E\left\{ \frac{\sum_{j=1}^{N} L_j + S_j}{\sum_{j=1}^{N} L_j + S_j} \right\} \text{ subject to }$$

$$\Pr \left\{ \sum_{j=i}^{N} L_{i-j+1} + S_i \geq D_i \right\} \geq \beta_i, \ i = 1, \ldots, N, \text{ and }$$

$$\Pr \left\{ \sum_{j=i}^{N} L_{N-j+1} \leq L \right\} \geq \beta_{N+1}, \text{ where the last constraint requires that unexpired charters not exceed a level } L.$$

Since it is the case that in the $i$th period only $D_1, \ldots, D_{i-1}$ are known, any specification of $L_i$ must represent only these $D_1, \ldots, D_{i-1}$, and therefore the most general linear
decision rule expressing \( L_i \) as a function of these \( D_k \) is

\[
L_i = \sum_{k=1}^{\frac{1-1}{L_i}} y_{ik}d_k + \gamma_i, \quad i = 1, \ldots, N, \quad \text{where} \quad d_k = D_k - \bar{D}_k.
\]

The above representation then becomes

\[
E \left\{ \sum_{j=1}^{\frac{1-1}{l}} \sum_{k=1}^{\frac{1-1}{l}} y_{jk}d_k + y_k \right\} + \delta_j s_j \right\} \\
= \sum_{j=1}^{\frac{1-1}{l}} \sum_{k=1}^{\frac{1-1}{l}} \left[ (\mu_k^j + \mu_j^i y_{jk} + \mu_j^i y_{jk}) \right] + \delta_j s_j \right\}
\]

subject to

\[
\Pr \left\{ \sum_{k=1}^{\frac{1-1}{l}} y_{ik}d_k + \gamma_i \right\} \geq \delta_i \}
\]

\[
\Pr \left\{ \sum_{k=1}^{\frac{1-1}{l}} y_{ik}d_k + \gamma_i \leq \delta_i \right\} \geq \beta_i
\]

\[
\Pr \left\{ \sum_{k=1}^{\frac{1-1}{l}} y_{ik}d_k + \gamma_i \leq 0 \right\} \geq a_i.
\]

The bar, as in \( \bar{y}_j \), indicates, of course, the mean.

Vajda presents in his book (46) an approach discussed by Talacko (35) and Beale (3). An interesting extension of their work that he considers, however, is this.

Consider a problem defined by:

\[
\text{Minimize} \quad \sum_{i=1}^{n} c_i x_i + \sum_{t=1}^{m} \sum_{j=1}^{n} f_j x_{n+j}, t = \sum_{i=1}^{n} c_i x_i + \sum_{t=1}^{m} \sum_{j=1}^{n} f_j (b_j + b_{jt} - \sum_{i=1}^{n} a_{ij} x_i) \\
\text{subject to} \quad \sum_{i=1}^{n} a_{ij} x_i + x_{n+j}, t = b_j + b_{jt}, \quad j = 1, \ldots, m, \quad \text{where}
\]
the \( b_{jt} \) are random errors, the \( x_{n+j, t} \) are slack variables, \( f_j \) is the cost of a deficiency in the slack variables, and \( p_t \) is the probability with which the random errors \( b_{jt} \) occur. If no information is available about the \( b_{jt} \), not even their means, the problem is modified to read

\[
\min_{x} \max_{t} \left[ \sum_{i=1}^{n} c_i x_i + \sum_{j=1}^{m} f_j (b_j + b_{jt} - \sum_{i=1}^{n} a_{ij} x_i) \right]
\]

subject to

\[
\sum_{i=1}^{n} a_{ij} x_i \leq b_j + b_{jt+1}, \quad j = 1, \ldots, m \text{ and } x_i \geq 0 \text{ for } i = 1, \ldots, n.
\]

Since it is known that \( \min_{x} \max_{t} \geq \max_{x} \min_{t} x \) for any function of \( x \) and \( t \), we may define the above as a programming problem.

Let \( \max_{t} \left[ \sum_{i=1}^{n} c_i x_i + \sum_{j=1}^{m} f_j (b_j + b_{jt} - \sum_{i=1}^{n} a_{ij} x_i) \right] = V. \)

We then seek values of \( x_i \) which minimize \( V \) subject to

\[
\sum_{i=1}^{n} a_{ij} x_i \leq b_j + b_{jt}, \quad j = 1, \ldots, m, \text{ and } \sum_{i=1}^{n} c_i x_i + \sum_{j=1}^{m} f_j (b_j + b_{jt} - \sum_{i=1}^{n} a_{ij} x_i) \leq V \text{ for all } t, \quad x_i \geq 0, \quad i = 1, \ldots, n.
\]

Write the last set of constraints as

\[
\sum_{i=1}^{n} c_i x_i + \sum_{j=1}^{m} f_j (b_j + b_{jt} - \sum_{i=1}^{n} a_{ij} x_i) + d_t = V, \quad d_t \geq 0
\]

for all \( t \). Now select a value of \( t \), say \( t = s \), and subtract, in the last set, the constraints for \( s \) from all others, to obtain finally the problem
Sasieni (33) has performed an industrial analysis which is of interest to us because it was later transformed into a programming problem by Dreyfus (10). The problem was concerned with the manufacture of rubber tires in which two "bladders" are used on a machine to produce two tires, one tire on each bladder simultaneously. Bladders can fail during the course of manufacture, and this is discovered when the faulty tire produced fails to pass inspection. When a bladder fails, cost is incurred due to the necessity of stripping the machine and replacing the bladder, the cost of the bladder, and the lost production time. However, once a machine has been stripped, replacing the second bladder is easily done, and the only extra cost incurred is the price of the second new bladder.

It has been found that the probability with which a bladder fails is a function of the number of tires made on it. The following policies have appeared to be optimal with regard to cost minimization:

a) Replace bladders which have been used in the manufacture of a predetermined number of tires.

\[
\begin{align*}
\text{minimize} & \quad \sum_{i} \left( c_{i} - \sum_{j} f_{j} a_{i,j} x_{i} + \sum_{j} f_{j} (b_{j} + b_{j,n}) + d_{s} \right) \\
\text{subject to} & \quad \sum_{i} a_{i,j} x_{i} \leq b_{j} + b_{j,t}, \quad j = 1, \ldots, m, \quad x_{i} \geq 0, \\
& \quad i = 1, \ldots, n, \quad \sum_{j} f_{j} (b_{j,t} - b_{j,s}) + d_{t} - d_{s} = 0 \text{ for all } t, \text{ and} \\
& \quad d_{t} \geq 0 \text{ for all } t, \text{ including } t = s.
\end{align*}
\]
b) When a machine has been stripped to replace one blad-
der, replace the other also if it has produced more
than a given number of tires.

The "age" of a bladder will be defined to be the number
of tires produced on it.

The author then shows how the ages required for policies
a) and b) can be determined so as to minimize the average
costs of replacing bladders per useful tire, which are the
cost of stripping the machine, etc. Let \( p_i \) = the probability
that a bladder of age \( i \) will fail on the next tire. Let the
state of a pair of bladders on the machine be characterized
by their ages, \( (i, j) \), and let \( i_j^{PRS} \) be the probability that,
on producing a pair of tires, the bladders pass from ages
\( (i, j) \) to ages \( (r, s) \), i.e., the Markov transition probabili-
ties. Let \( h_{ij} \) be the probability of ultimately finding the
system in state \( (i, j) \), let \( m \) be the predetermined age in
policy, \( n(n \leq m) \) be the predetermined age in policy b, \( c_1 \) be
the cost of purchasing a bladder, \( c_2 \) the cost of the scrap
produced if a bladder fails, \( c_3 \) be the labor cost in stripping
the machine, and let \( c_4 \) be the cost in lost production due to
such a bladder failure.

The probabilities \( i_j^{PRS} \) can be computed from production
records and the transition matrix found. Thus, the matrix
will have the following states with non-zero probabilities:
\( (i+1, j+1) \), \( (i+1, 0) \), \( (0, j+1) \) and \( (0, 0) \), where 0

indicates replacement of a bladder.

This matrix reaches a steady state, if the values m, n are consistently used, then the proportion of machines in each of the possible states (i, j) become independent of time. Let \( h_{ij} \) represent these proportions. Then the matrix \((ijpr_{rs})\), the expected number of single and double replacements, and the cost associated with any pair of ages \((m, n)\) may be computed.

Let \( P = (ijpr_{rs}) \) and \( H = (h_{ij}) \). Then \( H = PH \), or \((P-I)H = 0\), where \( I \) is the identity matrix. We thus have a homogeneous set of equations which reduce to

\[
\begin{align*}
    h_{00} &= \sum_{rs} rsP_{00} h_{rs}, \quad h_{0j} = \sum_{i, j-1} P_{ij} h_{i, j-1}, \\
    h_{ij} &= i-1, j-1 P_{ij} h_{i-1, j-1}, \quad i, j \neq 0.
\end{align*}
\]

The cost of purchasing bladders, per cycle, is \( c_1 \left[ 2h_{00} + \sum_{j=1}^{m} h_{0j} \right] \); that of scrap is \( c_2 \sum_{i, j} h_{ij} (p_i + p_j) \); that of lost production time is \( c_3 \sum_{i, j} h_{ij} (p_i + p_j - p_i p_j) \), while that of stripping the machine is \( c_3 \sum_{i=0}^{m} h_{0i} \).

The problem is to choose \((m, n)\) so as to minimize the sum of the costs.

Dreyfus (10) has restated the same problem as one in dynamic programming.

Let \( f_N(i, j) = \) total expected cost of producing \( N \) additional tires where bladder 1 has already produced \( i \) tires and
bladder 2 has produced j tires. The expected cost per tire for the production of N tires is then \( f_N(0, 0)/N \).

Let \( p_i \) be defined as before. Then the problem may be stated as the following recurrence relation:

\[
f_N(i, j) = \min \left\{ \begin{array}{l}
(1 - p_i)(1 - p_j)f_{N-2}(i+1, j+1) + p_ip_j \left[ (2c_1 + 2c_2 + c_3 + c_4 + f_N(0, 0)) + p_i(1 - p_j)\min\left[ c_1 + c_2 + c_3 + c_4 + f_{N-1}(0, j+1), 2c_1 + c_2 + c_3 + c_4 + f_{N-1}(i+1, 0) \right] + p_j(1 - p_i)\min\left[ c_1 + c_2 + c_3 + c_4 + f_{N-1}(i, 0), 2c_1 + c_2 + c_3 + c_4 + f_{N-1}(0, 0) \right] \right\}
\]

He also has a numerical example.

Geisler and Karr (17) include in their work a statement as to the formulation of the following problem as programming under uncertainty.

The military services make much use of supply tables with the following characteristics: each supply table consists of a bundle of spare parts selected in advance of use to meet supply needs of vehicles, aircraft, ships, etc.; the table is the sole source of supply during the period it is in use, and there is always some limitation on size.

It is desired to design supply tables under the following assumptions:

(a) The criterion for selecting the quantities and kinds of parts to go into a table is to minimize the expected number of shortages weighted by the "criticality" of the part.
(b) The demand for a part is independent of the amount stocked.

(c) Substitution among items in the table to meet demand is not allowed.

(d) It is possible to stock fractional parts of each item instead of only integral units.

Let \( p_i(x) \) be the probability density function for a demand of \( x \) units of the \( i^{th} \) part during the period that the table is to be used. Let \( S_i \) be the quantity of the \( i^{th} \) part placed in the table, where clearly \( S_i \geq 0 \), and let \( c_i \) be the "criticality factor" for the \( i^{th} \) part. Let \( m \) be the number of parts eligible for placing in the table, and let \( w_i \) be the unit size of the \( i^{th} \) part (in weight, volume, or whatever).

Then \( w = \sum_{i=1}^{m} w_i S_i \) represents the total size of the table if \( S_i \) units of the \( i^{th} \) part are placed in the table, and

\[
\sum_{i=1}^{m} c_i \int_{S_i}^{\infty} (x - S_i)p_i(x)\,dx
\]

gives the expected total number of shortages over the \( m \) different parts considered, each shortage weighted by its criticality.

The problem is then to minimize

\[
\sum_{i=1}^{m} c_i \int_{S_i}^{\infty} (x - S_i)p_i(x)\,dx
\]

subject to \( w = \sum_{i=1}^{m} w_i S_i \).

Radner (32) develops an elementary stochastic programming problem in which the random variables are discretely distributed.
Consider a firm with two activities, say production and promotion, dealing with only one product. Let a denote the amount of money allocated to production and let \( x_a \) be the resulting quantity produced, while \( b \) can denote the amount of money devoted to promotion and \( y_b \) the resulting demand generated. If both the product and generated demand are perishable and if the price of the commodity is one, then the profit will be \( \min (x_a, y_b) - (a + b) \).

Suppose now that the firm is uncertain about the true values of \( x \) and \( y \) and that accurate prediction is costly. The firm could then pay for accurate estimation (the case of full information) or it could rely on the probability distribution of \( x \) and \( y \) (the case of routine operation) and choose \( a \) and \( b \) so as to maximize expected profit. A third alternative (the decentralized case) may occur where decisions about \( a \) are made by one person, and those about \( b \) are made by another, all according to some predetermined decision rule.

Let \( x \) and \( y \) be statistically independent, and let each possess a finite number of discrete values with probabilities \( p(x,y) \). Let \( u(a,b; x,y) = \min f_n(a,b; x,y) \), where \( f_n \) is linear in \( a \) and \( b \) for all \( n, x, y \). Let \( r = R(x, y) \) be the information upon which action on \( a \) is based and let \( s = S(x, y) \) be the information on which action \( b \) is based. Let \( A \) denote any function of \( r \) (a decision rule for \( a \)) and let \( B \) be a corresponding function of \( s \). Let \( z \) be any function of \( x \)
and $y$. The author then claims that $(A, B)$ is a solution of
\[
\max E u(A[R(x, y)], B[S(x, y)]) \quad x, y
\]
subject to $A(r), B(s)$ nonnegative if and only if there is a $z$ such that
$(z, A, B)$ is a solution of $\max E z(x, y)$ (by selection of $A, B$
and $z$) subject to $z(x, y) \leq f_n(A[R(x, y)], B[S(x, y)]) \quad x, y$
for all $n, x, y$, where $E$ designates, as usual, the expectation
operator.

Since $E z(x, y) = \sum_{x,y} p(x, y) z(x, y)$ is linear in $z(x, y)$
and the constraints are linear in $z(x, y), A(r)$ and $B(s)$, the
latter problem is one in linear programming.

This paper closes with a numerical example.

Martin (30) discusses the "Expected-Return-Variance-of-
Return" theory of Markowitz (29) in the selection of invest­
ment portfolios. It is of interest to us because of its ready-
adaptability to a programming model.

Suppose an investor has either statistical information
or a probabalistic degree of belief regarding the expected
rate of return and the variance of return on a set of $n$ possi­
bile investments. Let these expected returns be $u_1, u_2, \cdots,$
$u_n$, and let the variances and covariances of return among the
investment alternatives be $\sigma_{ij}, i, j = 1, \cdots, n$. Let $x_i$
designate the fraction of available funds invested in alter­
native $i$. Then the problem may be formulated mathematically
as
\[
\sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0, \quad E = \sum_{i=1}^{n} x_i u_i, \quad V = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij}.
\]
The investor may then either stipulate $E$ and minimize $V$, or stipulate $V$ and maximize $E$. The author presents a numerical example.

Charnes, Cooper and Symonds (8) develop another programming model with chance constraints in a paper on planning the production of heating oil to meet stochastically-determined demand. It is pointed out initially that the planner might maximize expected profit, or minimize expected total cost while still supplying whatever demand happens to emerge.

Under the assumptions that the probability density functions of sales are known and that all relevant variables are statistically dependent, the authors express the constraints in the fashion described below.

For $j = 1, \ldots, N$, let $\Pr\{I_0 + \sum_{i=1}^{j} R_i \geq \sum_{i=1}^{j} S_i + I_{\text{min}} \geq a_j\}$ represent a sales constraint, where $a_j$ is a "confidence coefficient" prescribed for the $j^{\text{th}}$ production scheduling interval. $I_0$ is the initial inventory stated in terms of $M$ barrels of oil, $R_j \geq 0$ is the production rate (stated in terms of $M$ barrels of oil per day) to be scheduled in period $j$. $S_j \geq 0$ is the anticipated sales for period $j$ (stated in terms of $M$ barrels per day), $I_{\text{min}}$ is a minimum inventory level to be maintained, again stated in the same terms.

A set of storage constraints may be stated as

$$\Pr\{I_0 + \sum_{\ell=1}^{j} R_{\ell} \leq \sum_{\ell=1}^{j} S_{\ell} + I_{\text{max}} \geq u_j\},$$
where $I_{\text{max}}$ is the maximum inventory level allowed by storage capacity.

Let $S_j$ be a random variable with a known probability density function $f_j(S_j)$, $j = 1, \ldots, N$. Assume $u_j = 1, \ldots, N$. Then the problem may be stated as

$$\max E\pi = \max_{R,D} \int \pi(S, R) f(S) dS,$$

where

$$\pi(S, R) = \frac{1}{N} \sum_{j=1}^{N} \left[ p_j S_j - (c_j + T_j) R_j - k_j \bar{I}_j \right],$$

$R$ and $D$ are the domains of refinery and sales variations to be considered, $T_j$ is transport cost of taking produced crude oils to the refinery, $c_j$ represents average variable cost in dollars per barrel at period $j = 1, \ldots, N$, $k_j$ is inventory carrying cost in dollars per barrel per day for period $j$, $\bar{I}_0$ is average inventory in period $j$, and $p_j$ is the price per barrel expected to prevail in period $j$.

The work done by Professor Tintner (38, 39, 41, 43) is the foundation for the efforts contained in this work. In (41) he introduces the passive stochastic programming formulation as described in the Introduction applying it to an example of agricultural production planning. In (38) he presents the active approach to stochastic programming, again applying it to agricultural production planning and contrasting the results obtained by this method with those arrived at by means of the passive technique applied to the same data.
His theoretical framework has been the basis of other work, notably that in production planning done by Van Moeseke (47), and in planned growth models for a national economy (43).

While not concerned with stochastic programming, Gass (15) has done some work on programming with parametric objective functions which has some relationship with the active case discussed in this dissertation. Since his procedures and results are entirely based on the simplex method, with which we are not only not concerned but for which we can find no application here, we merely state his assumptions.

Let 0 and 0 be finite numbers, and consider 0 ≤ λ ≤ 0. Then the parametric programming problem of Gass is to find a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ which minimizes $\sum_{j=1}^{n} (d_j + \lambda d_j^i)x_j$

subject to $\sum_{j=1}^{m} a_{ij}x_j = b_i$, $i = 1, \ldots, m$, $x_j \geq 0$, where $d_j$, $d_j^i$ and $b_i$ are given constants.

He then gives an algorithm, based on the Simplex method, for finding such a vector and shows that a new basis, introduced by the Simplex method, yields a minimum for at least one value of $\lambda$. 
THE PASSIVE APPROACH IN STOCHASTIC LINEAR PROGRAMMING

It is our purpose in this chapter to discuss in detail the theoretical aspect and the characteristics of the passive approach. We will be able to demonstrate, for example, a weak duality theorem for the passive case which is of some interest theoretically. It is also the case, however, that a practical demonstration of our results would in itself entail a research problem not within the scope of this paper. This will be pointed out when it arises.

Let $A$ be an $m \times n$ matrix with elements $a_{ij}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Let $b$ be an $n$-dimensional column vector with elements $b_j$, let $c$ be an $m$-dimensional column vector with elements $c_i$, and let $x$ be an $n$-dimensional column vector with elements $x_j$. Let the elements $a_{ij}$, $b_j$, and $c_i$ be normally and independently distributed with known means and variances. Then a (maximization) stochastic programming problem may be stated as follows:

$$\text{maximize } b'x \text{ subject to } Ax \leq c \text{ and } x \geq 0.$$  

The dual of this problem may be stated as: minimize $c'y$ subject to $A'y \geq b$, $y \geq 0$, where $y$ is an $m$-dimensional column vector and $c'$ and $A'$ are the transposes of $c$ and $A$ respectively.

Each of these problems may be solved using the complete description technique, or the selections method.
A contemplation of the duality theorems might now lead one to conclude that if a particular region is feasible for the primal problem, then it must also be a region of feasibility for the dual, and conversely. This is quite a reasonable conjecture, and it is even the case that such a relationship is often encountered in empirical work done with the passive approach. It is, however, the case that such a relationship does not in general hold, and we present an example to demonstrate this. The example also serves to demonstrate the use of the selections technique.

Let us first consider a simple minimization problem given by:

\[
\text{minimize } c_1 y_1 + c_2 y_2 \text{ subject to}
\]

\[
a_{11} y_1 + a_{21} y_2 \geq b_1, \quad y_1 \geq 0, \quad y_2 \geq 0.
\]

\[
a_{12} y_1 + a_{22} y_2 \geq b_2
\]

Following the procedure outlined by the selections technique, we assign slack variables \(y_3\) and \(y_4\) to the inequalities to obtain the constraint system described by

\[
a_{11} y_1 + a_{21} y_2 - y_3 = b_1, \quad a_{12} y_1 + a_{22} y_2 - y_4 = b_2, \quad y_1 \geq 0, \quad y_2 \geq 0, \quad y_3 \geq 0, \quad y_4 \geq 0.
\]

The problem may now be expressed as:

\[
\text{minimize } c_1 y_1 + c_2 y_2 \text{ subject to}
\]

\[
\begin{pmatrix}
a_{11} & a_{21} & -1 & 0 \\
a_{12} & a_{22} & 0 & -1
\end{pmatrix}
\begin{pmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4
\end{pmatrix}
= \begin{pmatrix}
b_1 \\
b_2
\end{pmatrix},
\]
\[ y_1 \geq 0, \ y_2 \geq 0, \ y_3 \geq 0, \ y_4 \geq 0. \]

The rank of the coefficient matrix is 2. Hence we proceed to select two columns of the matrix in turn, and also the corresponding 2 variables, to form \( \binom{4}{2} = 6 \) sets of two equations in two unknowns.

For selection 1, we let \( y_3 = y_4 = 0 \) and select the first two columns. This gives us the system

\[
\begin{align*}
    a_{11}y_1^{(1)} + a_{21}y_2^{(1)} &= b_1 \\
    a_{12}y_1^{(1)} + a_{22}y_2^{(1)} &= b_2,
\end{align*}
\]

or

\[
y_1^{(1)} = \frac{\begin{vmatrix} b_1 & a_{21} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}} = \frac{b_1a_{22} - b_2a_{21}}{a_{11}a_{22} - a_{21}a_{12}}.
\]

and

\[
y_2^{(1)} = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{12} & b_2 \end{vmatrix}}{a_{11}a_{22} - a_{21}a_{12}} = \frac{a_{11}b_2 - b_1a_{12}}{a_{11}a_{22} - a_{21}a_{12}}.
\]

This selection will be feasible in all regions of the parameter space \( S \) where the numerator and denominator carry the same algebraic sign, for both variables \( y_1^{(1)} \) and \( y_2^{(1)} \).

Continuing the process, for the second selection we let \( y_1 = y_4 = 0 \) and select the second and third columns of the coefficient matrix. We then have the system
\[
\begin{pmatrix}
  a_{21} & -1 \\
  a_{22} & 0
\end{pmatrix}
\begin{pmatrix}
  y_2 \\
  y_3
\end{pmatrix}
= \begin{pmatrix}
  b_1 \\
  b_2
\end{pmatrix},
\]

or

\[a_{21}y_2 - y_3 = b_1 \quad \text{and} \quad a_{22}y_3 = b_2\]

This gives

\[y_3^{(2)} = \frac{b_2}{a_{22}} \text{ and } y_2^{(2)} = \frac{b_1}{a_{21}} + \frac{b_2}{a_{22}a_{21}},\]

and the selection is feasible if

\[\frac{b_2}{a_{22}} \geq 0 \text{ and if } \frac{b_1}{a_{21}} + \frac{b_2}{a_{22}a_{21}} \geq 0.\]

We let \(y_1 = y_3 = 0\) for the third selection, and choose the second and fourth columns. The resulting system of equations is

\[a_{21}y_2 = b_1 \quad \text{and} \quad a_{22}y_2 - y_4 = b_2,\]

giving

\[y_2^{(3)} = \frac{b_1}{a_{21}} \text{ and } y_4^{(3)} = \frac{a_{22}b_1}{a_{21}} - b_2,\]

a selection which is feasible if

\[\frac{b_1}{a_{21}} \geq 0 \text{ and if } \frac{a_{22}b_1}{a_{21}} \geq b_2.\]

Letting \(y_2 = y_3 = 0\) for the fourth selection and choosing the first and fourth columns of the matrix, the system of equations is
\[ a_{11} y_1 = b_1 \]
\[ a_{12} y_1 - y_4 = b_2, \]

from which
\[ y_1^{(4)} = \frac{b_1}{a_{11}} \] and \[ y_4^{(4)} = \frac{a_{12} b_1}{a_{11}} - b_2. \]

with the associated regions of feasibility determined by the requirements that
\[ \frac{b_1}{a_{11}} \geq 0 \] and \[ \frac{a_{12} b_1}{a_{11}} \geq b_2. \]

The fifth selection is found by taking the first and third columns and letting \( y_2 = y_4 = 0. \) This gives
\[ a_{11} y_1 - y_3 = b_1 \]
\[ a_{12} y_1 = b_2. \]

The associated regions of feasibility are
\[ \frac{b_2}{a_{12}} \geq 0 \] and \[ \frac{a_1 b_2}{a_{12}} - b_1 \geq 0. \]

The final selection is determined by taking the third and fourth columns. We find \( y_3 = -b_1 \) and \( y_4 = -b_2, \) with the regions of feasibility being \( b_1 \leq 0, \) \( b_2 \leq 0. \)

We have thus isolated all the regions of feasibility associated with this minimization problem. Let us now consider the dual of the problem. The dual is easily found to be:

\[
\text{maximize } b_1 x_1 + b_2 x_2 \text{ subject to } \]
\[ a_{11} x_1 + a_{12} x_2 \leq c_1 \]
\[ a_{21}x_1 + a_{22}x_2 \leq c_2 \]
\[ x_1 \geq 0, \; x_2 \geq 0. \]

Letting \( x_3 \) and \( x_4 \) be slack variables, the constraints become
\[ a_{11}x_1 + a_{12}x_2 + x_3 = c_1 \]
\[ a_{21}x_1 + a_{22}x_2 + x_4 = c_2 \]
\[ x_1 \geq 0, \; x_2 \geq 0, \; x_3 \geq 0, \; x_4 \geq 0, \]
or
\[
\begin{pmatrix}
  a_{11} & a_{12} & 1 & 0 \\
  a_{21} & a_{22} & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2 \\
  x_3 \\
  x_4
\end{pmatrix}
= \begin{pmatrix}
  c_1 \\
  c_2
\end{pmatrix}, \; x_j \geq 0, \; j = 1, \ldots, 4.
\]

The matrix of constraint coefficients for this, the dual problem, also has rank 2. Hence again there are \( \binom{4}{2} = 6 \) possible combinations of columns to consider and examine for independence. For the first selection, we may let \( x_3 = x_4 = 0 \) and choose the first and second columns. This gives us the system of equations
\[
\begin{pmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
= \begin{pmatrix}
  c_1 \\
  c_2
\end{pmatrix},
\]
and our solutions are easily found to be
\[
x_1^{(1)} = \frac{c_1a_{22} - c_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}} \quad \text{and} \quad x_2^{(1)} = \frac{a_{11}c_2 - a_{21}c_1}{a_{11}a_{22} - a_{21}a_{12}}.
\]
The selection is feasible if the numerator and denominator
have the same sign for both $x_1^{(1)}$ and $x_2^{(1)}$.

It is not necessary to describe any more selections or regions of feasibility for the maximization problem, since a comparison of the regions of feasibility just found with those described earlier for the dual minimization problem reveals that there are none for the minimization problem which match with the ones for the one selection found for the maximization problem. Thus it is not true that a region of feasibility for a primal stochastic problem will also be a region of feasibility for the dual.

It is apparent, however, that there must be regions which are feasible for both primal and dual, that is, there will be regions which will be subsets of the regions of feasibility for both primal and dual. To see that such non-empty subsets will always exist, consider regions of feasibility for the first selection from both the primal and dual in the above example. Let the set

$$\left\{ 0 \leq \frac{b_1 a_{22} - b_2 a_{21}}{a_{11} a_{22} - a_{21} a_{12}} \right\}$$

designate that region in the parameter space where $y_1^{(1)}$ is feasible in the minimization problem, and let

$$\left\{ 0 \leq \frac{c_1 a_{22} - c_2 a_{12}}{a_{11} a_{22} - a_{21} a_{12}} \right\}$$

be that region where $x_1^{(1)}$ is feasible in the maximization problem. It is clear, since these regions do not involve the
same parameters, that, if we let $\emptyset$ designate the empty set, then

$$\left\{ 0 \leq \frac{c_1a_{22} - c_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}} \right\} \cap \left\{ 0 \leq \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}} \right\} \neq \emptyset.$$  

The non-empty intersection is clearly a region of feasibility for both the maximization and minimization problems. That such an intersection of regions of feasibility for the two problems must, in fact, always be non-empty is apparent from the fact that the right-hand side of the constraint equations in the primal (after the selection is performed) is not the right-hand side of the constraint equations for the dual. Thus, all selections in the maximization problem will be feasible for all values of $b_1$ and $b_2$. Since a given selection in the minimization problem will be feasible for only certain values of the $b_1$ and $b_2$, the intersection is guaranteed.

Another problem is made apparent by examination of the solutions for the variables in this last example, and that is that the variables are not defined everywhere in the parameter space. Thus, examining each selection for the minimization problem in turn, we note that the first selection has no meaning in any subset of the parameter space where $a_{11}a_{22} - a_{21}a_{12} = 0$; the second is meaningless anywhere where $a_{22} = 0$, the third, if $a_{21} = 0$, the fourth, if $a_{11} = 0$, and so on. Taking an expectation, for example, over such a region or any part of it would be equivalent to integrating a function of a single
variable over a point of discontinuity of the function, and would be meaningless. Let us therefore make a definition suitable for our purposes.

Definition 3.1: Let \( L \) be a linear programming problem of the passive stochastic type, and let \( V \) be an area in the associated parameter space. Then \( V \) will be defined to be a critical region for the problem \( L \) if any variable in \( L \) expressed as a function of the stochastic parameters of \( L \) is undefined in \( V \).

We shall, in later results, always abstract away from such critical regions in any proofs to be given.

Another conjecture which has been prompted by results arising from applications of stochastic programming is that regions of optimality for the maximization problem are the regions of feasibility for the minimization problem, and conversely. This situation has occurred often enough to warrant some interest, and we present a simple example to demonstrate that this is not the case in general.

Consider a maximization problem defined by the following.

Maximize \( b_1 x_1 + b_2 x_2 \) subject to \( a_{11} x_1 + a_{12} x_2 \leq c_1 \), \( x_1 \geq 0, \ x_2 \geq 0 \).

Letting \( x_3 \) be a slack variable, the constraints may be expressed as \( a_{11} x_1 + a_{12} x_2 + x_3 = c_1 \), \( x_1 \geq 0, \ x_2 \geq 0, \ x_3 \geq 0 \).

The rank of the constraint coefficient matrix \( (a_{11} \ a_{12} \ 1) \) is one. For the first selection we let \( x_2 = x_3 = 0 \). The constraint becomes
\[ x^{(1)}_1 = \frac{c_1}{a_{11}}, \]

which is feasible if \( c_1 \geq 0, a_{11} > 0 \) or if \( c_1 \leq 0 \) and \( a_{11} < 0 \).

In the second selection we let \( x_1 = x_3 = 0 \), giving

\[ x^{(2)}_2 = \frac{c_1}{a_{12}}, \]

which is feasible if \( c_1 \geq 0 \) and \( a_{12} > 0 \), or if \( c_1 \leq 0 \) and \( a_{12} < 0 \).

Finally, we let \( x_1 = x_2 = 0 \), and we find that \( x^{(3)}_3 = c_1 \), a selection which is feasible if \( c_1 \geq 0 \).

The dual to this problem is

\[
\begin{align*}
\text{minimize } c^y_1 & \quad \text{subject to} \\
(a_{11}) y_1 & \geq (b_1), \quad y_1 \geq 0.
\end{align*}
\]

As usual, we define \( y_2 \) and \( y_3 \) to be slack variables obtaining the problem

\[
\begin{align*}
\text{minimize } c^y_1 & \quad \text{subject to} \\
\begin{pmatrix} a_{11} & -1 & 0 \\ a_{12} & 0 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} & = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad y_1 \geq 0, \ y_2 \geq 0, \ y_3 \geq 0.
\end{align*}
\]

The rank of the coefficient matrix for this system of equations is 2. Hence we select in \( \binom{3}{2} = 3 \) ways two of the three columns, and the associated variables, to examine for linear independence.

For the first selection we let \( y_3 = 0 \) and select the
first two columns. This results in the system

\[ a_{11}y_1 - y_2 = b_1 \]

\[ a_{12}y_1 = b_2, \]

giving

\[ y_1^{(1)} = \frac{b_2}{a_{12}} \text{ and } y_2^{(1)} = a_{11}(\frac{b_2}{a_{12}}) - b_1, \]

a selection which is feasible in regions where \( b_2 \geq 0, a_{12} > 0 \)
or \( b_2 \leq 0, a_{12} < 0, \) and

\[ a_{11}(\frac{b_2}{a_{12}}) \geq b_1. \]

We make the second selection by choosing the first and third columns of the coefficient matrix and by letting \( y_2 = 0. \) The resulting system is

\[ a_{11}y_1 = b_1 \]

\[ a_{12}y_1 - y_3 = b_2, \]

giving

\[ y_1 = \frac{b_1}{a_{11}} \text{ and } y_3 = a_{12}(\frac{b_1}{a_{11}}) - b_2. \]

The associated regions of feasibility are clearly \( b_1 \geq 0, \)
\( a_{11} > 0 \) or \( b_1 \leq 0, a_{11} < 0, \) and

\[ a_{12}(\frac{b_1}{a_{11}}) \geq b_2. \]

The third selection, choosing the last two columns in the matrix, gives \( y_2 = -b_1 \) and \( y_3 = -b_2, \) and the regions of feasi-
bility are \( b_1 \leq 0, b_2 \leq 0 \).

Examining each of these selections for optimality, we see that the first selection will be optimal in regions of the parameter space where

\[
\frac{c_1 b_2}{a_{12}} < \frac{c_1 b_1}{a_{11}} < 0,
\]

that the second selection will be optimal where

\[
\frac{c_1 b_1}{a_{11}} < \frac{c_1 b_2}{a_{12}} < 0,
\]

and that the third selection will be optimal where

\[
0 < \frac{c_1 b_1}{a_{11}} \quad \text{and} \quad 0 < \frac{c_1 b_2}{a_{12}}.
\]

Let us complete our example by choosing one of these regions of optimality and showing that it coincides with none of the regions of feasibility found for the maximization problem. Along these lines, we consider the regions of optimality for the second selection

\[
\frac{c_1 b_1}{a_{11}} < \frac{c_1 b_2}{a_{12}} < 0.
\]

The regions of feasibility for the first selection for the maximization problem were found to be \( c_1 \geq 0 \) and \( a_{11} > 0 \) or \( c_1 \leq 0, a_{11} < 0 \), a set of restrictions completely independent of \( b_1, b_2, \) and \( a_{12} \). Thus, if we consider the region \( c_1 \geq 0, a_{11} > 0 \), whether or not the selection for the dual that we are considering is optimal or not depends on the values of \( b_1, b_2, \)
and $a_{12}$. We must, for example, have $b_1 \leq 0$, a restriction not present in the feasibility requirements. Thus this minimization selection can be optimal only along half of the $b_1$ axis, while the region of feasibility for the first maximization selection contains the entire $b_1$ axis. This is sufficient to demonstrate the failure of equivalence between this particular region of optimality for the dual and this particular region of feasibility for the primal. It is obvious that the same line of comparison between this region of optimality for the dual and the remaining two regions of feasibility for the primal will yield the same conclusion. This completes the demonstration of the counterexample.

We thus far have nothing except negative results on the subject of dual relationships in passive stochastic linear programming and one might begin to wonder whether any such relationship can be salvaged. There is, in fact, a weak duality theorem, which we now state and prove.

Theorem 3.1: Let $F^{(k)}(x)$ be the objective function determined by the $k$th selection in the solution of an arbitrary maximization stochastic linear programming problem, and let $U^{(k)}_F$ be that region in the parameter space throughout which $F^{(k)}(x)$ is both optimal and feasible. Let $a$ be any point in the interior of $U^{(k)}_F$. Then there exists a selection, say the $l$th selection, for the dual problem with corresponding objective function
\[G(x)\text{ such that } G(x) = F(x)\text{ and a neighborhood } N(x)\]

about \(a\) throughout which \(G(x)\) is both feasible and optimal, and such that \(G(x) = F(x)\) for all \(x\) in \(N(x)\).

Proof: By the duality theorem, since \(a\) is a fixed point in the parameter space and since we have assumed that the primal has a solution at \(a\), the dual must also have a solution there. Let this solution be given by the \(F\)th selection. We then have that \(G(x) = F(x)\). This completes the first part of the proof. We note in passing that there may be more than one selection in the dual for which this holds.

For the second part of the proof, suppose there is no neighborhood of \(a\) throughout which \(G(x)\) is optimal even though it is optimal at \(x = a\). Then, since the \(F\)th selection is feasible for the primal and \(F(x)\) is optimal and a solution to the primal throughout \(N\), the dual must likewise have a solution throughout \(U\). Suppose, then, that the \(F\)th selection for the dual is optimal at some points of an arbitrary neighborhood, say \(N(x)\) of \(a, N(x) \subset U\). We then have that while \(G(x) - G(x) < 0, G(x) - G(x) \geq 0\) at some points \(x\) in \(N(x)\).

If there are only a finite number \(x_1, x_2, \ldots x_q\) of such points \(x\), we may define \(f(x_i) = \rho(x, x_i)\), \(i = 1, \ldots, q\), and let \(e' = \min_{i} f(x_i)\). Then \(N(x)/2(x)\) is a neighborhood of \(a\) throughout which \(G(x)\) is optimal as well as feasible.
Suppose, then, that there are an infinite number of points $x_\lambda$, $\lambda \in \Delta$, where $\Delta$ is an uncountable or countable index set, for which $G(\ell)(x_\lambda) < G(\ell')(x_\lambda)$. Since $S = \bigcap_{i=1}^{\max n+m+n} \mathbb{R}_i$, the parameter space, is metric and first countable, there exists at $a$ a countable basis $\{N_{1/n}(a)\}$ contained in $S$, and in each $N_{1/n}(a)$ there exists at least one point $x_\lambda$.

Define the "doughnut" neighborhood $D_{1/n}(a)$ of $a$ as $D_{1/n}(a) = N_{1/n}(a) - N_{1/n+1}(a)$. Then either there are infinitely many of these neighborhoods $D_{1/n}(a)$ containing points $x_\lambda$, or there are only finitely many. If there are only finitely many then there exists an integer $N > 0$ such that for all $n' > N$, $D_{1/n'}(a)$ contains no points $x_\lambda$, in which case either $N_{1/n'}(a)$ contains no points $x_\lambda$ for all $n' > N$ and $N_{1/n'}(a)$ forms a neighborhood of $a$ throughout $G(\ell')(x) = F(\ell')(x)$, or there are infinitely many of the points $x_\lambda$ which belong to $N_{1/n'}(a)$ for all $n' > N$. These points form a set $T$ contained in the parameter space $S$, and $a$ is a limit point of $T$. Hence, by Theorem 1.12, there is a sequence of distinct points of $T$ converging to $a$. Denote this sequence of points by $\{x_{\lambda n}\}$.

Since the $G(\ell)(x)$ are continuous at all points in the parameter space not in a critical region, we have that the sequence $\{G(\ell)(x_{\lambda n'})\}$ converges to $G(\ell)(a)$ for all $\ell'$, and in particular $\ell'$. But $G(\ell')(a) - G(\ell)(a) < 0$ and $G(\ell')(x_{\lambda n'}) - G(\ell)(x_{\lambda n'}) \geq 0$ for all $n$. Thus we have that the sequence $W(x_{\lambda n'}) =$
If \( g' \) of \( \varphi' \), \( g(\varphi') \) does not converge to \( W(a) \), which contradicts the continuity of \( W(x) \) and therefore of \( G(\varphi')(x) \) and \( G(\varphi)(x) \). For this case, the problem is solved.

If there are infinitely many of the neighborhoods \( D_{1/n} a \) containing points \( x_\lambda \), we may use the axiom of choice on each such \( D_{1/n} a \) to select out of each one of these points \( x_\lambda \), and we shall call the point \( x_\lambda \), chosen out of the neighborhood \( D_{1/n} a \), \( x_{\lambda n} \). The points \( \{x_{\lambda n}\} \) thus form a sequence which converges to \( \alpha \), and we finish the problem as demonstrated for the case in which only finitely many of the neighborhoods \( D_{1/n} a \) contain points \( x_\lambda \). This proves the theorem.

We note that if \( a \) is a point on the boundary of the region of optimality and feasibility for a given selection, we cannot then maintain the truth of the above theorem.

**Decision Rules for Optimization**

We have heretofore in this work concerned ourselves with only a particular decision rule for deciding when a given selection was optimal. That is, we have stated that a selection \( \bar{k} \) was optimal for a maximization passive stochastic problem if the associated objective function \( f(\bar{k}) \) was greater than the objective function \( f(k) \) associated with another selection \( k \).

While this is certainly the only sensible rule possible
for a deterministic linear program, the flexibility of a stochastic program is such that an analyst might choose any one of several decision rules for deciding when a selection was "optimal". Which rule he happened to choose would be determined by situations and circumstances extraneous to the data of the problem itself. To illustrate and give meaning to these admittedly fuzzy statements, let us consider some examples.

Suppose we regard the objective function in a maximization passive stochastic programming problem as representing profit, where the variables are outputs. The constraints then represent the input restriction under which the entrepreneur is forced to operate.

Let us now conceive of an entrepreneur who feels that he must obtain a certain level of profit, say R, and that any profit greater than R, while certainly desirable, is of secondary importance. Such a situation could arise, for example, if the entrepreneur needed to achieve the level R of profit in the current planning period in order to survive at the beginning of the next planning period. In such a case it seems rational and reasonable that the entrepreneur would choose that selection whose associated objective function, in regions of feasibility, achieved values greater than or equal to R with probability p, say, in preference to a selection whose associated objective function took on values greater than R
with probability $p' < p$. Optimal regions for a selection would then be those regions throughout which the associated objective function assumed at least a value of $R$, and that selection would be chosen as "optimal" whose regions of optimality, intersected with those of feasibility, possessed the greatest probability measure. Such a "maximin" strategy or rule - maximum probability of minimum profit level allowed - would seem to be of considerable interest.

Another decision rule which suggests itself is this: choose that selection to be optimal whose associated objective function has the greatest expected value over the regions of feasibility for the selection. As an example, consider the maximization problem discussed in the first part of this chapter.

Maximize $b_1 x_1 + b_2 x_2$ subject to

\[
\begin{align*}
    a_{11} x_1 + a_{12} x_2 &\leq c_1 \\
    a_{21} x_1 + a_{22} x_2 &\leq c_2 \\
    x_1 &\geq 0, x_2 \geq 0.
\end{align*}
\]

Letting $x_3, x_4$ be slack variables, we found the system of constraints to be

\[
\begin{align*}
    a_{11} x_1 + a_{12} x_2 + x_3 &= c_1 \\
    a_{21} x_1 + a_{22} x_2 + x_4 &= c_2 \\
    x_1 &\geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0,
\end{align*}
\]
or
The first selection gave

\[
\begin{pmatrix}
    a_{11} & a_{12} & 1 & 0 \\
    a_{21} & a_{22} & 0 & 1
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    x_3 \\
    x_4
\end{pmatrix}
= \begin{pmatrix}
    c_1 \\
    c_2
\end{pmatrix}, \quad x_j \geq 0, \quad j = 1, \ldots, 4.
\]

The first selection gave

\[
x_1^{(1)} = \frac{a_{22}c_1 - c_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}} \quad \text{and} \quad x_2^{(1)} = \frac{a_{11}c_2 - a_{21}c_1}{a_{11}a_{22} - a_{21}a_{12}},
\]

and the associated objective function is easily found to be

\[
F^{(1)} = b_1 \left[ \frac{c_1 a_{22} - c_2 a_{12}}{a_{11} a_{22} - a_{21} a_{12}} \right] + b_2 \left[ \frac{a_{11} c_2 - a_{21} c_1}{a_{11} a_{22} - a_{21} a_{12}} \right].
\]

Let \( C_1 \) be the region of feasibility for the first selection.

For the second selection, we can let \( x_2 = x_4 = 0 \) and select the first and third columns of the matrix. This gives a system with equations:

\[
\begin{align*}
    a_{11} x_1 + x_3 &= c_1 \\
    a_{21} x_1 &= c_2
\end{align*}
\]
or

\[
x_1^{(2)} = \frac{c_2}{a_{21}} \quad \text{and} \quad x_3 = c_1 - \frac{a_{11} c_2}{a_{21}}.
\]

The objective function is then

\[
F^{(2)} = \frac{b_1 c_2}{a_{21}},
\]

and we again designate the region of feasibility for this
selection by $C_2$.

The third selection can be found by choosing the first and fourth columns, and by letting $x_2 = x_3 = 0$. We then have

$$a_{11}x_1 = c_1$$

$$a_{21}x_1 + x_4 = c_2$$

or

$$x_1^{(3)} = \frac{c_1}{a_{11}} \quad \text{and} \quad x_4^{(3)} = c_2 - \frac{a_{21}c_1}{a_{11}}.$$
The objective function is
\[ f(5) = \frac{b^2c^2}{a_{22}}, \]
and the feasible region is \( C_5 \).

For the sixth selection, we choose the last two columns and find \( x_3^{(6)} = c_1 \) and \( x_4^{(6)} = c_2 \). \( f(6) = 0 \).

We should like to mention, as a digression, that if a particular region of feasibility contains a critical region, it becomes apparent that it is not a connected region, for it is clear that it can be expressed as the union of two or more separated sets.

The analyst or programmer may now do the following: if the variates \( b_i, c_j, \) and \( a_{ij} \) have probability density functions \( f_{b_i}(b_i), f_{c_j}(c_j) \) and \( f_{a_{ij}}(a_{ij}), i = 1, \ldots, 2, \ j = 1, \ldots, 2, \) and are mutually independent, they have a joint probability distribution defined by \( f(b_1, b_2, c_1, c_2, a_{11}, \ldots, a_{22}) = \prod_{i=1}^{2} \prod_{j=1}^{2} f_{b_i} f_{c_j} f_{a_{ij}} \).

Define the conditional expectation
\[ E_{C_k}(F(k)) \]
of the objective function for the \( k \)th selection over the critical region \( C_k \) to be
\[ E_{C_k}(F(k)) = \int_{C_k} \int_{C_k} \int_{C_k} \int_{C_k} \int_{C_k} \prod_{i=1}^{2} \prod_{j=1}^{2} f_{a_{ij}}(a_{ij}) f_{b_i}(b_i) f_{c_j}(c_j). \]
\( w(k) = a_{11}a_{12} \cdots a_{22}b_{1}b_{2}b_{1}c_{1}c_{2} \). Then choose that \( k' \) such that \( E_{c_k}(F(k')) = \max_k E_{c_k}(F(k)) \) as the optimal selection.

Thus, for example, we would find \( E_{c_1}F(1) \) and \( E_{c_3}F(3) \) to be given by

\[
E_{c_1}F(1) = \sum_{a_{11}} \sum_{a_{12}} \cdots \sum_{a_{22}} \sum_{b_1} \sum_{b_2} \sum_{c_1} \sum_{c_2} \prod_{i} \prod_{j} f_{a_{ij}}(a_{ij}) f_{b_1}(b_1) f_{c_j}(c_j) \cdot
\left[
\frac{c_1 a_{22} - c_2 a_{12}}{a_{11}a_{22} - a_{21}a_{12}} + \frac{a_{11}c_2 - a_{21}c_1}{a_{11}a_{22} - a_{21}a_{12}}
\right] \cdot
da_{11} \cdots \, da_{22} \, db_1 \, db_2 \, dc_1 \, dc_2
\]

and

\[
E_{c_3}F(3) = \sum_{a_{11}} \sum_{a_{12}} \cdots \sum_{a_{22}} \sum_{b_1} \sum_{b_2} \sum_{c_1} \sum_{c_2} \prod_{i} \prod_{j} f_{a_{ij}}(a_{ij}) f_{b_1}(b_1) f_{c_j}(c_j) \left[ \frac{b_1 c_1}{a_{11}} \right] \cdot
da_{11} \cdots \, dc_2.
\]

It is, unfortunately, the case that an attempt to apply this decision rule empirically soon gives the worker to understand that it is a most difficult, perplexing, and unpleasant undertaking, due to the fact that the boundaries of the regions of feasibility over which one must integrate are defined by messy combinations of the same parameters that appear in the integrand. The problem is, moreover, compounded in complexity as the dimensions of the program increase. There is, to our knowledge, no general statistical approximation which could be applied to an entire class of such problems, and even
attempts to find such an approximation for a particularly simple special case involved a statistical analysis not within the scope of this paper.

It is to be noted now that of the two most recently-discussed decision rules for optimization, neither has in its favor the "dual" property that the optimal expected value, or probability measure, as the case may be, for the dual is equal to that of the primal. A decision rule which does possess an analogous such property is the following. Recall our initial discussion in which a selection \( k' \) was defined to be "optimal" for a maximization problem if the associated objective function \( F(k') \) was greater than the objective function associated with any other selection. If \( U_{k'} \) was that region in the parameter space \( S \) where \( F(k') > F(k) \) for all other \( k \), we proved that there is some subset of \( U_{k'} \) throughout which a selection \( l' \) for the dual is optimal and feasible if it is optimal and feasible at a point within this subset of \( U_{k'} \), and that \( F(k')(x) = G(l')(x) \) for all \( x \) in this same subset. We may now prove the following rather trivial result.

**Theorem 3.2:** Let \( U_{k'} \) be a region of both feasibility and optimality (in the present sense) for a selection \( k' \) for a maximization problem, let \( a \) be a point in \( U_{k'} \) such that \( F(k')(a) = G(l')(a) \), where \( G(l') \), as earlier, is the objective function associated with a selection \( l' \) for the dual, and let \( N_e(a) \) be that neighborhood of \( a \) throughout which, by Theorem
3.1. \( F(k')(x) = G(k')(x) \) for all \( x \) in \( N_\varepsilon(\alpha) \). Let the conditional expectation of \( F(k')(x) \) given that \( x \) is in \( N_\varepsilon(\alpha) \) be denoted by \( E_{N_\varepsilon(\alpha)}[F(k')(x)] \), and similarly for the conditional expectation \( E_{N_\varepsilon(\alpha)}[G(k')(x)] \) of \( G(k')(x) \) given that \( x \) is in \( N_\varepsilon(\alpha) \). Then \( E_{N_\varepsilon(\alpha)}[F(k')(x)] = E_{N_\varepsilon(\alpha)}[G(k')(x)] \).

Proof: The proof follows directly from the well-known statistical theorem that if \( h(x) = g(x) \) for all \( x \) in a domain \( D \), then \( E(h(x)) = E(g(x)) \), where the expectation is taken over \( D \).
THE ACTIVE APPROACH IN STOCHASTIC PROGRAMMING

The active approach to stochastic linear programming has been described briefly in the Introduction, and the reader will easily recall the characteristics which distinguish it from the passive problem just discussed. In this chapter we propose to consider an active stochastic programming problem and to restate it in a form that is amenable to finding the dual of the problem.

Two things should perhaps be pointed out at the outset of our discussion. First, the apparent compactness of the active approach is deceiving, and a matrix-vector formulation of an active representation proliferates in dimensionality to an extent that becomes totally unmanageable on standard-sized paper unless the number of constraints are restricted. Because of this, we restrict ourselves to problems having four constraints. Secondly, it will become evident that there is no unique matrix-vector representation of an active problem. We will consider only two such representations, one of which seems considerably more "natural" than the other, and we will be able to show that, in a restricted sense, and under a particular decision rule for optimality, the same solution is common to both representations.

Finally, our results are proved only for the case of four constraints. This may well be unnecessarily restrictive, but
the flexibility of the active approach is such that techniques for extending the results to problems of greater dimensionality seem difficult to come by. The seemingly obvious approach, double (mathematical) induction, was a failure.

Consider the active stochastic linear programming problem defined by

\[
\begin{align*}
\text{maximize} & \quad b_1 x_1 + b_2 x_2 \\
\text{subject to} & \quad a_{11} x_1 + a_{12} x_2 \leq c_1 u_{11} + c_1 u_{12} \\
& \quad a_{21} x_1 + a_{22} x_2 \leq c_2 u_{21} + c_2 u_{22} \\
& \quad x_1 \geq 0, x_2 \geq 0.
\end{align*}
\]

The \( u_{ij} \) are "decision variables" with respect to which the programmer optimizes the objective function, and are under the restriction that \( \sum_{j=1}^{2} u_{ij} = 1, i = 1, 2 \). An interpretation of the \( u_{ij} \) has been previously discussed.

We note that with the restriction \( \sum_{j=1}^{2} u_{ij} = 1, i = 1, 2 \), we may add the constraints having common values of \( i \), and the active problem becomes a passive problem. Thus, we may write

\[
\begin{align*}
& \quad a_{11} x_1 + a_{12} x_2 \leq c_1 u_{11} + c_1 u_{12} = c_1 \\
& \quad a_{21} x_1 + a_{22} x_2 \leq c_2 u_{21} + c_2 u_{22} = c_2 \\
& \quad x_1 \geq 0, x_2 \geq 0.
\end{align*}
\]

An example of an application of the active approach which also serves to provide a motivating interpretation of the parameters and variables involved is presented by Professor Tintner (38). Suppose we are concerned with production planning for a particular farm growing corn and flax, and are
confronted with the following data: the net price of corn is $b_1 = 1.56$ per unit, that of flax is $b_2 = 3.81$ per unit; the input coefficients for land in relation to corn and flax, respectively, are $a_{11}$ and $a_{12}$, while the input coefficients for capital are, again for corn and flax in that order, $a_{21}$ and $a_{22}$. We suppose that the $a_{ij}$ are normally and independently distributed with the following statistics: the mean of $a_{11}$ is 0.022740 and its standard deviation is 0.0065205; the mean of $a_{12}$ is 0.072440 and its standard deviation is 0.0256583; the mean of $a_{21}$ is 0.317720 and it has a standard deviation of 0.0853977, while the mean of $a_{22}$ is 0.969500 and the standard deviation is 0.4338177. We suppose that we have $c_1 = 148$ units of land and $c_2 = 1800$ units of capital available for the enterprise.

The programming problem for the passive case is easily formulated. It is

$$\text{maximize } 1.56 x_1 + 3.81 x_2 \text{ subject to}$$

$$a_{11} x_1 + a_{12} x_2 \leq 148$$

$$a_{21} x_1 + a_{22} x_2 \leq 1800$$

$$x_1 > 0, x_2 > 0.$$

$x_1$ and $x_2$ are, of course, the number of units of corn and flax, respectively, that are produced.

Let $u_{11}$ represent the proportion of land used to grow corn, and let $u_{12} = 1 - u_{11}$ represent the proportion of land used to grow flax. Let $u_{21}$ represent the proportion of
capital used to produce corn and \( u_{22} = 1 - u_{21} \) represent the proportion of capital devoted to growing flax. The active stochastic programming problem is then the following:

Find the probability distribution of \( p = 1.56 x_1 + 3.81 x_2 \) subject to the constraints

\[
\begin{align*}
  a_{11} x_1 &\leq 148 u_{11} & a_{21} x_1 &\leq 1800 u_{21} \\
  a_{12} x_2 &\leq 148 u_{12} & a_{22} x_2 &\leq 1800 u_{22} \\
  \sum_{j=1}^{2} u_{ij} & = 1, \ i = 1, 2, \ u_{ij} & \geq 0 \text{ for } i, j = 1, 2, x_i & \geq 0, \ i = 1, 2.
\end{align*}
\]

The \( a_{ij} \) are random variables with known distributions.

Professor Tintner then uses numerical methods to approximate the distribution of \( p \) and its mathematical expectation, varying the decision variables \( u_{ij} \) to maximize the expectation of \( p \).

This is the only theoretical work to be found involving parametric programming with stochastic coefficients, and it is upon this that we base ourselves.

There are a multitude of matrix and vector formulations which can be related to an active programming problem of the type described. Let us consider two formulations of the same general active program. As previously, suppose the problem is to

maximize \( b_1 x_1 + b_2 x_2 \) subject to

\[
\begin{align*}
  a_{11} x_1 &\leq c_{11} u_{11} & a_{21} x_1 &\leq c_{21} u_{21} \\
  a_{12} x_2 &\leq c_{12} u_{12} & a_{22} x_2 &\leq c_{22} u_{22}
\end{align*}
\]
82

\[ x_1 \geq 0, \ x_2 \geq 0, \quad \sum_{j=1}^{2} u_j = 1, \quad i = 1, 2. \]

To find our first formulation, let us re-write this program in the form

**maximize**

\[ \frac{b_1}{2} x_1 + \frac{b_2}{2} x_2 + \frac{b_1}{2} x_1 + \frac{b_2}{2} x_2 \]

**subject to**

\[
\begin{pmatrix}
  a_{11} & 0 & 0 & 0 \\
  0 & a_{12} & 0 & 0 \\
  0 & 0 & a_{21} & 0 \\
  0 & 0 & 0 & a_{22}
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2 \\
  x_1 \\
  x_2
\end{pmatrix}
\leq
\begin{pmatrix}
  c_{1u_{11}} \\
  c_{1u_{12}} \\
  c_{2u_{21}} \\
  c_{2u_{22}}
\end{pmatrix},
\]

\[ x_1 \geq 0, \ x_2 \geq 0. \]

Letting \( x_3, \ x_4, \ x_5, \) and \( x_6 \) be slack variables, the constraints may assume the form

\[
\begin{pmatrix}
  a_{11} & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
  0 & a_{12} & 0 & 0 & 0 & 1 & 0 & 0 \\
  0 & 0 & a_{21} & 0 & 0 & 0 & 1 & 0 \\
  0 & 0 & 0 & a_{22} & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2 \\
  x_1 \\
  x_2 \\
  x_3 \\
  x_4 \\
  x_5 \\
  x_6
\end{pmatrix}
\leq
\begin{pmatrix}
  c_{1u_{11}} \\
  c_{1u_{12}} \\
  c_{2u_{21}} \\
  c_{2u_{22}}
\end{pmatrix}, \quad x_1 \geq 0,
\]

\[ x_2 \geq 0, \ x_3 \geq 0, \ x_4 \geq 0, \ x_5 \geq 0, \ x_6 \geq 0. \]

The problem may be streamlined more by the following procedure: let \( z_1 = x_3 = x_1, \ z_2 = z_4 = x_2, \ z_5 = x_3, \ z_6 = x_4, \ z_7 = x_5, \) and \( z_8 = x_6. \) The problem then becomes
maximize
\[
\frac{b_1}{2} z_1 + \frac{b_2}{2} z_2 + \frac{b_1}{2} z_3 + \frac{b_2}{2} z_4
\]

subject to
\[
\begin{pmatrix}
  a_{11} & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
  0 & a_{12} & 0 & 0 & 0 & 1 & 0 & 0 \\
  0 & 0 & a_{21} & 0 & 0 & 0 & 1 & 0 \\
  0 & 0 & 0 & a_{22} & 0 & 0 & 0 & 1 \\
  1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
  0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\
\end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \\ z_8 \end{pmatrix} = \begin{pmatrix} c_{11} u_{11} \\ c_{12} u_{12} \\ c_{21} u_{21} \\ c_{22} u_{22} \end{pmatrix},
\]

\[z_1 \geq 0, \, i = 1, \ldots, 8.\]

The last two restrictions merely assure \(z_1 = z_3\) and \(z_2 = z_4\).

The coefficient matrix has 6 rows and 8 columns. It is easy to verify that it has rank 6 (consider the determinant formed by deleting the first and last columns) and hence we have \(\binom{8}{6} = 28\) combinations of columns to examine to locate the extreme points.

A second and much more natural formulation, and one which is considerably more advantageous from a computational aspect is the following.

Letting the objective function again be \(b_1 x_1 + b_2 x_2\), we write the constraints in the form
\[ \begin{pmatrix} a_{11} & 0 \\ 0 & a_{12} \\ a_{21} & 0 \\ 0 & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c_{1u_{11}} \\ c_{1u_{12}} \\ c_{2u_{21}} \\ c_{2u_{22}} \end{pmatrix} \]

\[ x_1 \geq 0, \ x_2 \geq 0, \ 2 \sum_{j=1}^{2} u_{1j} = 1, \ i = 1, 2. \]

Upon insertion of the slack variables \( x_3 \geq 0, \ x_4 \geq 0, \ x_5 \geq 0, \ x_6 \geq 0, \) we re-write the problem to read

maximize \( b_1 x_1 + b_2 x_2 \) subject to

\[ \begin{pmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{12} & 0 & 1 & 0 & 0 \\ a_{21} & 0 & 0 & 0 & 1 & 0 \\ 0 & a_{22} & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} c_{1u_{11}} \\ c_{1u_{12}} \\ c_{2u_{21}} \\ c_{2u_{22}} \end{pmatrix}, \ x_1 \geq 0, \ i = 1, \ldots, 6, \ 2 \sum_{j=1}^{2} u_{1j} = 1, \ i = 1, 2. \]

This coefficient matrix, with four rows and six columns, clearly has rank 4 and we have \( \binom{6}{4} = 15 \) selections to make and examine for extreme points.

We have now two representations of our original active problem. Let us denote these representations, respectively, by A and B. If one wishes, one may demonstrate by evaluating the 45 selections that both A and B have critical regions not
common to the other. Hence there will be points in the parameter space where $A$ is defined and $B$ is not, and vice versa. It is also conceivable that $A$ might assume its optimum in a critical region of $B$, and vice versa. Thus the two representations are not completely equivalent.

It is, however, the case that the objective functions for the two representations will be equal if evaluated at any point in the parameter space where $x_1 = z_1$ and $x_2 = z_2$, an obvious assertion. With this preliminary discussion, we should like to prove some results regarding relationships between the two representations.

**Theorem 4.1**: Let $a$ be a fixed point in the parameter space which is not in a critical region of either of the representations $A$ and $B$. Let $u_{11}$ and $u_{12}$ be fixed for both representations, and let $F_A(a)$ and $F_B(a)$ be the respective objective functions, with parameters fixed at the point $a$, to be maximized. Suppose that neither of the two convex regions formed by the restrictions for the representations $A$ and $B$ has an extreme point which is in a critical region for the other. Then $\max F_A(a) = \max F_B(a)$.

**Proof**: Let $\beta_{FA}$ be an extreme point at which $F_A(a)$ assumes its maximum value, and let $z_1\beta_{FA}$ and $z_2\beta_{FA}$ be the values of $z_1$ and $z_2$ associated with the point $\beta_{FA}$. Then $F_B(a)$ assumes the value $\max F_A(a)$ at any point in the parameter space where $x_1 =$
\[ z_{18}^{\mathcal{A}} \text{ and } x_2 = z_{25}^{\mathcal{A}}. \] We thus have that \( \max F_A(a) \geq \max F_B(a) \).

Similarly, we may show that \( \max F_B(a) \geq \max F_A(a) \). This proves the theorem.

Let us now, before proceeding further, introduce some new notation that will make our next efforts considerably more understandable. Suppose that the decision variables \( u_{11} \) and \( u_{12} \) can each assume at most a finite number of values between 0 and 1 inclusive, e.g., let \( u_{11} \) assume \( p_1 \) values and let \( u_{22} \) assume \( p_2 \) values. Let us index these values, letting \( \emptyset = 1, \ldots, p_1 \) and \( \theta = 1, \ldots, p_2 \). Suppose that there are \( \lambda_A \) selections to be made in representation \( \mathcal{A} \) and \( \lambda_B \) selections in representation \( \mathcal{B} \). Let \( F_A^{\emptyset, \theta}(a) \) be the objective function for representation \( \mathcal{A} \) evaluated at the point \( a \), with \( \emptyset \) and \( \theta \) indicating the values of \( u_{11} \) and \( u_{22} \) under consideration, while \( F_B^{\emptyset, \theta}(a) \) indicates the corresponding situation surrounding representation \( \mathcal{B} \). We may now demonstrate another result.

**Theorem 4.2:** Let the conditions described in the preceding theorem hold for all \( p_1 p_2 \) possible combinations of values for \( u_{11} \) and \( u_{22} \). Let the objective function for one representation, say \( \mathcal{A} \), assume, for a fixed point \( a \) in the parameter space, its maximum at \( \emptyset = \emptyset' \) and \( \theta = \theta' \). Then the objective function \( F_B^{\emptyset, \theta}(a) \) also assumes its maximum at \( \emptyset = \emptyset' \) and \( \theta = \theta' \).

**Proof:** Let \( F_A^{\emptyset', \theta'}(a) = \max_{\emptyset, \theta} F_A^{\emptyset, \theta}(a) \). Then, by the preceding
theorem. \( F_B^g \theta' (a) = F_A^g \theta' (a) \). If there exists \( \overline{\theta} \), \( \overline{\theta} \) such that

\( F_B^\overline{\theta} (a) > F_B^\theta' (a) \), we then have, again by the preceding theorem, that \( F_A^\overline{\theta} (a) = F_B^\overline{\theta} (a) > F_B^\theta' (a) = F_A^\theta' (a) \), contradicting our assumption that \( F_A^\theta' (a) = \max_{\theta, \theta'} F_A^\theta (a) \), and thus proving the theorem.

Another result that one would reasonably expect to hold is that if \( a \) is not a limit point of a critical region for either A or B, and that if \( F_A^\theta' (a) = F_B^\theta' (a) \), then there exists a neighborhood \( N_{\epsilon} (a) \) about \( a \) for which \( F_A^\theta' (x) = F_B^\theta' (x) \) for all \( x \) in \( N_{\epsilon} (a) \). It is the case that all methods of proof break down on this result: the conditions simply aren't strong enough. It may be true but for the moment it must remain an open question.

Returning now to stress the stochastic nature of the active programming problem, let us assume that a formulation has been chosen. We can offer no reason why we should not choose the one offering the least resistance to computation, specifically representation B, although we must admit the weakness of our reason for this choice.

We should now like to suggest the following as a procedure for evaluating the active stochastic program. Modifying our notation with the understanding that we shall henceforth refer only to representation B, we allow a subscript in the
objective function to index the selections and let $F_k$, $k = 1, \ldots, \lambda$, denote the objective function associated with the $k^{th}$ selection. To optimize with respect to any decision rule desired, assign all $p_1p_2$ possible combinations of values to the decision variables $u_{11}$ and $u_{22}$, choosing, for each combination, that selection which results in an "optimal" objective function. Taking this set of optimal objective functions, one for each of the $p_1p_2$ combinations of values for the $u_{11}$ and $u_{22}$, we optimize with respect to the index numbers $\phi$ and $\theta$ to find an over-all optimal selection and combination of values for the $u_{ij}$. More formally, we find $\max \max \max F_{k\theta}$. It is clear that the three operators $\max, \max$ and $\max$ will commute. We also note the fact that for given values of $u_{11}$ and $u_{22}$, the function $F_{k\theta}$ is a continuous function of the random variables anywhere outside of the critical region. We now prove another result.

**Theorem 4.3:** Let $\phi'$, $\theta'$ and $k'$ indicate a particular value for $\phi$, $\theta$ and $k$ respectively. Let $F_{k'}\theta'$ be the corresponding objective function and let $\alpha$ be a point in the interior of a region of feasibility of $F_{k'}\theta'$ such that $F_{k'}\theta' (\alpha) > F_{k\theta} (\alpha)$ for all other $\phi$, $\theta$ and $k$. Then there is some neighborhood of $\alpha$ such that, for points $x$ in this neighborhood,

$$F_{k'}\theta' (x) > F_{k\theta} (x),$$

for all other $\phi$, $\theta$ and $k$. 
Proof: Suppose there is no neighborhood of \( a \) for which this is true. We then have that, while \( F_k^{\prime}{\varphi}^\prime(a) > F_k{\varphi}(a) \), \( F_k^{\prime}{\varphi}^\prime(x_\lambda) \leq F_k{\varphi}(x_\lambda) \) for infinitely many points \( x_\lambda \) contained in an \( \epsilon \)-neighborhood \( N_\epsilon(a) \), for every \( \epsilon \).

Let \( N_{1/n}(a) \) be a spherical neighborhood of radius \( 1/n \) of \( a \), and, as before, define \( D_{1/n}(a) = N_{1/n}(a) - N_{1/n+1}(a) \). Then, should there be only finitely many of the \( D_{1/n}(a) \) containing points \( x_\lambda \), we treat this proof as in the proof of Theorem 3.1, and the problem is solved. If there are infinitely many of the \( D_{1/n}(a) \) containing points \( x_\lambda \), we use the axiom of choice to select one point \( x_\lambda \) from each \( D_{1/n}(a) \), and call this point \( x_{\lambda n} \). The \( x_{\lambda n} \) form a sequence \( \{x_{\lambda n}\} \) converging to \( a \), but the sequence \( H(x_{\lambda n}) = F_k^{\prime}{\varphi}^\prime(x_{\lambda n}) - F_k{\varphi}(x_{\lambda n}) \) does not converge to \( H(a) \), since \( H(a) > 0 \) and \( H(x_{\lambda n}) \leq 0 \) for all \( n \). This implies that \( H(x) \) is not a continuous function, and thus contradicts the statement that \( F_k^{\prime}{\varphi}^\prime \) and \( F_k{\varphi} \) are continuous. This proves the theorem.

We have now dwelt at some length on the active approach to stochastic linear programming and have now to consider one of its more interesting aspects, which is the dual. The dual of an active representation has the characteristic that the decision variables appear, after the dual transformation, in the objective function. For example, let the primal (maximizing) problem be given by
maximize \( b_1 x_1 + b_2 x_2 \) subject to
\[
\begin{pmatrix}
a_{11} & 0 \\
0 & a_{12} \\
a_{21} & 0 \\
0 & a_{22}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
\leq
\begin{pmatrix}
c_{1u_{11}} \\
c_{1u_{12}} \\
c_{2u_{21}} \\
c_{2u_{22}}
\end{pmatrix},
x_1 \geq 0, x_2 \geq 0.
\]

The dual is then
\[
\text{minimize } c_{1u_{11}}y_1 + c_{1u_{12}}y_2 + c_{2u_{21}}y_3 + c_{2u_{22}}y_4 \text{ subject to }
\begin{pmatrix}
a_{11} & a_{21} & 0 \\
0 & a_{12} & 0 \\
a_{21} & 0 & a_{22}
\end{pmatrix}
\begin{pmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4
\end{pmatrix}
\geq
\begin{pmatrix}
a_1 \\
a_2
\end{pmatrix},
y_j \geq 0, j = 1, \ldots, 4.
\]

Letting \( y_5 \geq 0 \) and \( y_6 \geq 0 \) be slack variables, the constraints may be written as
\[
\begin{pmatrix}
a_{11} & a_{21} & 0 & -1 & 0 \\
0 & a_{12} & 0 & a_{22} & 0 & -1
\end{pmatrix}
\begin{pmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4 \\
y_5 \\
y_6
\end{pmatrix}
= \begin{pmatrix}
b_1 \\
b_2
\end{pmatrix},
y_j \geq 0, j = 1, \ldots, 6.
\]

Again, the rank of the matrix is 2, and since there are six columns, we have \( \binom{6}{2} = 15 \) selections to make and examine for linear independence.

Another theorem is now in order.
Theorem 4.4. Let $\varphi'\theta'$ be a selection for the primal and let $a$ be a point in the parameter space at which $F_k^{\varphi'\theta'}$ is feasible and also $F_k^{\varphi'\theta'}(a) > F_k^{\varphi\theta}(a)$ for all other $\varphi, \theta$ and $k$. Then there is a neighborhood $N_\varepsilon(a)$ about $a$ and a selection $h'$ for the dual and values $\varphi'', \theta''$ of $\varphi$ and $\theta$ such that $G_n^{\varphi''\theta''}(x) = F_k^{\varphi'\theta'}(x)$ for all $x$ in $N_\varepsilon(a)$, where $G$ is the objective function for the dual.

Proof: By the duality theorem, if $a$ is a fixed point in the parameter space, we know that if $\varphi'' = \varphi'$ and $\theta'' = \theta'$, then there is a selection $h'$ for the dual such that $G_n^{\varphi'\theta'}(a) = F_k^{\varphi'\theta'}(a)$. The values $\varphi', \theta'$ and $h'$ then suffice to make the theorem true at the point $a$. The proof that these same values also make the theorem true throughout some neighborhood $N_\varepsilon(a)$ of $a$ follows from the proof of Theorem 3.1 for the passive case, and is hence omitted.

One might also like to believe that the selection $h'$ for the dual is also the optimal dual selection at the point $a$. That is, one might conjecture that $G_n^{\varphi'\theta'} < G_n^{\varphi\theta}$ for all possible values of $\varphi, \theta$ and $h$. This is not the case, and we present a theorem to attest to its falsity.

Theorem 4.5: Let $a$ be a point which is not a limit point of a critical region and let $F_k^{\varphi'\theta'}$ be the objective function associated with the numbers $\varphi', \theta'$ and $k'$ such that $F_k^{\varphi'\theta'}(x) > F_k^{\varphi\theta}(x)$
for all \( x \) in some neighborhood of \( a \), for all other \( \varnothing, \Theta \) and \( k \).

Let \( G_{h'}^{\varnothing, \Theta'} \) be the objective function associated with a selection \( h' \) of the dual such that \( G_{h'}^{\varnothing, \Theta'} (x) = F_{k'}^{\varnothing, \Theta'} (x) \) throughout a neighborhood of \( a \), as in Theorem 4.4. Then there exists a selection \( h'' \) and values \( \varnothing'' \) and \( \Theta'' \) such that

\[
G_{h''}^{\varnothing'' \Theta''} (x) < G_{h'}^{\varnothing, \Theta'} (x)
\]

throughout this same neighborhood of \( a \). That is, \( G_{h'}^{\varnothing, \Theta'} \) is not optimal in this neighborhood.

Proof: Since \( F_{k'}^{\varnothing, \Theta'} \) is optimal throughout this neighborhood of \( a \), it follows that \( F_{k'}^{\varnothing'' \Theta''} (a) < F_{k'}^{\varnothing, \Theta'} (a) \) where \( \varnothing'' \) and \( \Theta'' \) are any other values of \( \varnothing \) and \( \Theta \), and hence that \( G_{h'}^{\varnothing'' \Theta''} (a) < G_{h'}^{\varnothing, \Theta'} (a) \).

As before, it is easy to show that if this is true, then

\[
G_{h'}^{\varnothing, \Theta'} (x) < G_{k'}^{\varnothing'' \Theta''} (x)
\]

for all \( x \) in this same neighborhood. This proves the theorem.

It is, in fact, the case that if one selects any point \( a \) which is not a limit point of a critical region and chooses a selection, say \( k' \), for the primal, then chooses those \( \varnothing'', \Theta'' \) such that \( F_{k'}^{\varnothing'' \Theta''} (a) < F_{k'}^{\varnothing, \Theta} (a) \) for all \( \varnothing, \Theta \), the optimal selection \( G_{h}^{\varnothing \Theta} \) with values \( \varnothing \) and \( \Theta \) for \( \varnothing \) and \( \Theta \) will be less than or equal to \( G_{h'}^{\varnothing, \Theta'} = F_{k'}^{\varnothing, \Theta'} \) throughout a neighborhood of \( a \).

Thus the results that one might expect to exist regarding duality in the active case are not at all strong.

We would now like to examine any relationship that may
exist between the passive and active versions of a particular programming problem.

It is apparent that a relationship can be demonstrated for a passive 2 by 2 programming problem and the corresponding active representation. For example, consider the passive constraints given by

\[
\begin{align*}
  a_{11}x_1 + a_{12}x_2 & \leq c_1 \\
  a_{21}x_1 + a_{22}x_2 & \leq c_2 \\
  x_1 & \geq 0, \ x_2 \geq 0
\end{align*}
\]

and the active constraints given by

\[
\begin{align*}
  a_{11}x_1 & \leq c_1u_{11} & a_{21}x_1 & \leq c_2u_{21} \\
  a_{12}x_2 & \leq c_1u_{12} & a_{22}x_2 & \leq c_2u_{22} \\
  x_1 & \geq 0, \ x_2 \geq 0, \ u_{11} + u_{12} & = 1 \\
  & & u_{21} + u_{22} & = 1
\end{align*}
\]

Two systems of inequalities are said to be equivalent if a solution for one system is also a solution for the other.

It is clear that any solution for the set of active constraints, for any values of the \(u_{ij}\) satisfying the restriction \(\sum_j u_{ij} = 1, \ i = 1, 2\), will also satisfy the passive constraints. To see this we have only to add the active inequalities to obtain

\[
\begin{align*}
  a_{11}x_1 + a_{12}x_2 & \leq c_1u_{11} + c_1u_{12} = c_1 \\
  a_{21}x_1 + a_{22}x_2 & \leq c_2u_{21} + c_2u_{22} = c_2,
\end{align*}
\]

which are the constraints for the passive case. Thus, any feasible solution for the active case must clearly be a feasi-
ble solution for the passive case. On the other hand, we can find a solution for the passive case which is not contained in the set of solutions for the active, specifically, after transformation to equalities, the solution

\[
x_1 = \frac{c_1 a_{12}}{a_{11} a_{22}} = \frac{c_1 a_{22} - c_2 a_{12}}{a_{11} a_{22} - a_{21} a_{12}}
\]

and

\[
x_2 = \frac{a_{11} c_2 - a_{21} c_1}{a_{11} a_{22} - a_{21} a_{12}}
\]

This demonstrates the following result.

**Theorem 4.6:** Let \( a \) be any point in the parameter space which is not a limit point of a critical region for either the active or passive representations. Let \( A \) represent the maximum value of the set of objective functions for the active case when all values of the decision variables are considered, and let \( P \) represent the maximum of the set of objective functions for the passive case. Then \( A \leq P \).

We must admit, however, discovery of the fact that this result has long been known to Professors Tintner and J. K. Sengupta of Iowa State University, Ames, Iowa, (private communication, 1962) and we can, therefore, not claim originality. We include it for completeness.
CONCLUDING COMMENTS

It may well be the case that we have more future research problems implicitly contained in this work than we have solved.

First, a numerical approximation is badly needed for evaluating the conditional expectations taken over regions of feasibility. For example, let us consider a very elementary stochastic or risk programming problem

\[
\text{maximize } c^T x + Cg \text{ subject to } \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \ x_1 \geq 0, \ x_2 \geq 0,
\]

where we will assume \( c_1 \sim \mathcal{N}(4,1), c_2 \sim \mathcal{N}(5,1), b_1 \sim \mathcal{N}(4,2), b_2 \sim \mathcal{N}(6,4), \) for the sake of explication. The coefficient matrix is assumed deterministic.

Adding slack variables \( x_3 \geq 0 \) and \( x_4 \geq 0 \), the system of constraints becomes

\[
\begin{pmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \ x_1 \geq 0, \ i = 1, \ldots, 4.
\]

The first selection may be the first two columns and the first two variables. We then have that

\[
x_1^{(1)} = \frac{3b_1 - b_2}{2}, \quad x_2^{(1)} = \frac{2b_2 - 4b_1}{2},
\]

which is feasible if \( b_1 \geq 1/3 \ b_2, b_2 \geq 2b_1 \). The objective
function becomes \( F(1) = \frac{3}{2} c_1 b_1 - \frac{1}{2} c_1 b_2 + c_2 b_2 - 2c_2 b_1 \), and the mathematical expectation \( E(F(1)) \) is given by

\[
E(F(1)) = \left( \frac{1}{2\pi} \right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{0} \int_{2b_1}^{3b_1} db_2 e^{-(c_1 - 4)^2} \cdot e^{-(c_2 - 5)^2} e^{-(b_1 - 4)^2/2} e^{-(b_2 - 6)^2/4} e^{3c_1 b_1 - \frac{b_1}{2} c_1 b_2 + c_2 b_2 - 2c_2 b_1}
\]

The problem is obvious: the appearance of the same variables in the limits of integration that appear in the integrand render this intractable, and the problem is magnified greatly if the coefficient matrix is allowed random elements.

Secondly, we feel that there is much more left to be done in regard to the relationship between the active and passive approaches, and we feel that the way to accomplish this, and many other results, is to define a measure-theoretic structure for the parameter space. We can conjecture that a proper definition of a measure space, or probability measure defined on the parameter space, would be a powerful tool allowing a much more critical analysis of any stochastic or risk programming problem. At this time, such an effort goes far beyond the scope of this work.

Lastly, it is a well-known fact that associated with any solvable deterministic linear programming problem is a certain two-person zero-sum game. We are frankly curious as to the
game associated with an active stochastic programming problem, and we conjecture that the reduction from program to game will go through without a hitch anywhere outside of a critical region. We think the solution to such a game, using any appropriate decision rule, would be of interest.
In this work we have considered a linear programming problem in which the coefficients in the constraints and the objective function are random variables. This is known as stochastic or risk programming. After defining a topology for the parameter space, which, if there are $m \times n + m + n$ coefficients in the problem, consists of the topological product $\bigotimes_{i=1}^{m+n} R_i$ where $R$ is the real line, we were able to demonstrate a weak duality theorem for two types of stochastic programming problems. The first is the so-called passive problem, a representative of which class is

$$\max c_1 x_1 + c_2 x_2 \text{ subject to }$$

$$a_{11} x_1 + a_{12} x_2 \leq b_1$$
$$a_{21} x_1 + a_{22} x_2 \leq b_2$$
$$x_1 \geq 0, \ x_2 \geq 0,$$

while the second type is the active problem, exemplified by

$$\max b_1 x_1 + b_2 x_2 \text{ subject to }$$

$$a_{11} x_1 \leq c_1 u_{11}$$
$$a_{12} x_2 \leq c_1 u_{12}$$
$$a_{21} x_1 \leq c_2 u_{21}$$
$$a_{22} x_2 \leq c_2 u_{22}$$
$$x_1 \geq 0, \ x_2 \geq 0, \ u_{ij} \geq 0, \ i, j = 1, 2, \ \sum_{j=1}^{2} u_{ij} = 1, \ i=1,2.$$
The $u_{ij}$ are "decision variables" which assume a finite number of discrete values and with respect to which the analyst optimizes his desired goals. Some decision rules for such optimization were given, and it was pointed out that these are meaningful if the problem is a stochastic or risk program, and have no interpretation in the deterministic (non-stochastic) case.

Other results demonstrated included a theorem to the effect that the maximum for a passive problem will be greater than or equal to the maximum for the active case, and a theorem stating that the values of the decision variables that optimize the objective function for the active maximization case will not optimize the objective function for the dual minimization case.
LITERATURE CITED


ACKNOWLEDGMENTS

I would like to express my deepest gratitude and appreciation to Professor Gerhard Tintner, who has served as my advisor and major professor throughout most of my time in graduate school and who suggested the topic and directed the research for this dissertation.

I would also like to thank Dr. J. K. Sengupta for his assistance and helpful comments during the preparation of this manuscript, Mr. Ray Guenther for comments which aided in the organization of this work, and Mrs. Mary Haensel, typist.