STOCHASTIC LINEAR PROGRAMMING

by

Paul van Moeseke

A Thesis Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
MASTER OF SCIENCE

Major Subject: Mathematical Economics

Signatures have been redacted for privacy

Iowa State University
Of Science and Technology
Ames, Iowa

1960

QA265
M64x
c.2

1126-21

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	MODELS WITH A VECTOR OF STOCHASTIC RESTRICTIONS OR INPUTS	10
	A. Certainty - Equivalence Theorems	10
	B. Orthogonal-Channel Chains	13
III.	MODELS WITH STOCHASTIC INPUT COEFFICIENTS	22
	A. Passive Approach	22
	B. Active Approach	28
IV.	GENERAL MODELS	36
	A. Approximation by a Quotient Distribution for Normal Variates	36
	B. Replacing Variates by Their Boundaries	40
	C. Selection of the Maximum Lower-Confidence Limit: a New Model	49
V.	AGRICULTURAL MODELS WITH STOCHASTIC YIELDS AND PRICES	59
	A. Method of Hybrid Activities	59
	B. Selection of the Maximum Lower-Confidence Limit: A Practical New Model	64
VI.	NON-RANDOM VARIATION	75
VII.	NON-LINEAR STOCHASTIC PROGRAMMING	79
VIII.	CONCLUSIONS WITH SPECIAL REFERENCE TO STATE PROGRAMMING	81
IX.	REFERENCES	85
X.	ACKNOWLEDGMENTS	90

T14342

I. INTRODUCTION

If one selects from a set the activities (relevant to a specific enterprise) which are to be carried out over some planned period, and then assigns definite, non-negative levels to each of them, one has a program. This assignment will be optimum if the combination of activity levels is such that an objective is maximized within the restrictions imposed by the conditions appropriate to the enterprise. The objective to be maximized is a function of the activity levels: the shape of the function partly determines how relative weights are attached to the various activities. Thus an entrepreneur or farmer has to consider an objective function F of activity levels x_i

1.1 $$F = F(x_1, x_2, \ldots, x_n) \qquad (x_i \geqslant 0;\ i=1,2,\ldots,n)$$

but is restricted by limited resources c_i: the way in which each activity uses up a resource is, of course, characteristic of both resource and activity. The restrictions may, therefore, be written:

1.2
$$\begin{aligned} \Psi_1 &= \Psi_1(x_1, x_2, \ldots, x_n) \leq c_1 \\ \Psi_2 &= \Psi_2(x_1, x_2, \ldots, x_n) \leq c_2 \\ &\ldots \qquad\qquad \ldots \qquad\qquad \ldots \\ \Psi_m &= \Psi_m(x_1, x_2, \ldots, x_n) \leq c_m \end{aligned}$$

(For simplicity's sake we consider all c_i as maxima.)
Set 1.2 obviously characterizes the business concern or farm as

a distinct entity or "body" in the m-dimensional restriction space.

A program is commonly called <u>linear</u> when it is, in fact, doubly linear, i.e., both equation 1.1 and set 1.2 are approximated linearly for relatively small changes in the x's. If set 1.2 is linear we assume constant returns to scale, a fact which has been found to be statistically plausible in a number of instances (53, p. 55). Another important assumption is additivity: the total amount of resources used by several activities must be equal to the sum of the resources used by each individual enterprise (no interaction), cfr. Heady and Candler (25, p. 17). The notation becomes:

$$1.3 \qquad P = a_1x_1 + a_2x_2 + \ldots + a_nx_n \qquad (x_i \geq 0;\ i=1,2,\ldots,n)$$

subject to

$$1.4 \qquad \begin{aligned} b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n &\leq c_1 \\ b_{21}x_1 + b_{22}x_2 + \ldots + b_{2n}x_n &\leq c_2 \\ \ldots \qquad\qquad \ldots &\quad \ldots \\ b_{m1}x_1 + b_{m2}x_2 + \ldots + b_{mm}x_n &\leq c_m \ . \end{aligned}$$

In matrix notation we shall write:

maximize

$$1.5 \qquad P = a'x \qquad (x \geq 0)$$

subject to the linear inequalities

$$1.6 \qquad Bx \leq c;$$

where a and x are column vectors of n elements, B an m x n

matrix, and c a column vector of m elements, while ' denotes transposition.

In economic terms a may be net prices, x outputs, B the "technical matrix" of resource or input requirements per unit level of output, and c the amounts of available resources. P then represents total net profit.

A linear-programming problem can be looked at from two formally different angles. First, total profit 1.5 derived from outputs may be maximized under conditions 1.6 which make sure that these outputs do not use up more resources than the available amounts (vector c), as explained. Second, the same problem may be expressed as one of minimizing the value of the inputs (resources) but now under conditions assuring that value going into an activity cannot be less than its net price. This second problem is the dual or "reflected image" of the first one. In matrix notation we now have to minimize

1.5b $\quad P^{o} = c'x^{o} \qquad (x^{o} \geqslant 0)$

subject to

1.6b $\quad B'x^{o} \geqslant a,$

where P^{o} is a scalar representing the total value of inputs and x^{o} is an m dimensional column vector of resource values; B' is the transpose of the original technical matrix from the first problem. It has been proved (31, Ch. XIX) that the linear program and its dual have either a common solution or no

solution at all (duality theorem). We shall here continue the argument in terms of the original problem.

Inequalities 1.4 or 1.6 can be transformed into equalities by allowing free disposal or non-use of every resource, and, therefore, including as many slack variables as there are resources. The i^{th} line of conditions 1.4 now reads

1.7 $$b_{i1}x_1+b_{i2}x_2 + \ldots + b_{in}x_n+0y_1+0y_2+\ldots+1y_i+\ldots+0y_m=c_i,$$

where the inclusion of y_i allows for disposal of y_i units of the i^{th} resource. In matrix notation this means that we have to define an enlarged column vector of n+m activity (and disposal) levels

1.8 $$z = \{x,y\} ,$$

an enlarged column vector of n+m prices (profits from non-use are zero)

1.9 $$p = \{a, o\} ,$$

and an enlarged m x (m+n) matrix

1.10 $$D = [B,I],$$

where I is the unit matrix of rank m.

Now we maximize the new profit function

1.11 $$G = p'z \qquad (z \geqslant 0)$$

under the conditions

1.12 $$Dz = c,$$

which we can restate, by splitting D into its column-vector components $d^{(s)}$ as

$$1.13 \qquad \sum_{s=1}^{n} z_s d^{(s)} = c,$$

where z_s is the s^{th} element of the above-defined z vector and c denotes again the m-dimensional resource vector. To each z vector there corresponds a point $z = (z_1, z_2, \ldots, z_n)$. Given c and $d^{(s)}$ $(s=1,2,\ldots,n)$ in the m-dimensional restriction space, the points $z = (z_1, z_2, \ldots, z_n)$ which are non-negative and satisfy 1.13 comprise a convex set Γ in the activity space (2, pp. 554-556). Evidently, the m(relevant) restrictions map the programming problem into a positive, convex hyper-polyhedron in this space; each restriction defines a facet. We need the following properties of convex sets:

(1) a linear function 1.11 defined for point z in Γ has a maximum at an extreme point of Γ;

(2) point $z = (z_1, z_2, \ldots, z_n)$ is an extreme of Γ if and only if the $d^{(s)}$ with positive weights p form a linearly independent set among all $d^{(s)}$. So, by virtue of the non-negativity condition in 1.11, only m weights are non-zero.

As a result, the optimum solution may be found by selecting all possible square m x m submatrices $D^{(k)}$ $(k=1,2,\ldots,K; K=C_m^{m+n})$ from the enlarged matrix D in 1.10, as stated by Dorfman (15, pp. 31 sqq). The solution vector maximizing G must be one of the K solution vectors $z^{(k)}$, of m elements, solved

from

1.14 $$D^{(k)}z^{(k)} = c,$$

provided D has rank m (non-degeneracy condition). This is the selection method which makes up the framework for our theoretical considerations.

Our linear-programming problem depends on three kinds of parameters: the p vector, the c vector or "right-hand side variables" and the technical coefficients of B. If all these parameters are fixed (ordinary or deterministic programming problem), this means that we are only considering one point in the parameter space S; this point has probability one. If, however, some or all of these parameters are subject to variation, one has to consider a joint probability distribution (54, p. 200).

1.15 $$P = P(p_1, p_2, \ldots, p_n, c_1, c_2, \ldots, c_m, b_{11}, b_{12}, \ldots, b_{mn}; \theta_1, \theta_2, \ldots),$$

defined over S and characterized by parameters θ (means, variances, covariances, etc.) In the case of <u>subjective risk</u>, where this distribution is known with certainty or probability one, this means that an individual anticipates that the chances for p_1 to belong to the interval $p_1 + d_{p1}$ and, simultaneously, for p_2 to belong to the interval $p_2 + dp_2$, and likewise for the other p's and for the c's and b's are (51, 52)

1.16 $$P(p_1, p_2, \ldots, p_n, c_1, c_2, \ldots, c_m, b_{11}, b_{12}, \ldots, b_{mn}; \theta_1, \theta_2, \ldots) \cdot \prod_j dp_j \cdot \prod_i dc_i \cdot \prod_{ij} db_{ij}$$

$(j=1,2,\ldots,n;\ i=1,2,\ldots,m;\ ij=11,12,\ldots,mn)$.

A situation of <u>subjective uncertainty</u> arises where there exists an a priori distribution of the probability distributions themselves; in practice we shall, however, restrict ourselves to the assumption that the distribution 1.15 is given and, in most cases, that it may be approximated by the normal.

From 1.15 one may, at least in principle, derive

1.17 $$Q = Q(G^*;\ w_1, w_2, \ldots;\ u_{11}, u_{12} \ldots),$$

the distribution of the profits G^* resulting from optimum programs; this distribution is characterized by parameter sets w and u where the w's are given by "nature", the u's by human decision. In agriculture for instance, the w's may be characteristic of the productivity of a certain soil type or region while the u_{ij} may represent the proportion of resource i reserved for activity j. Assuming that the entrepreneur has a preference functional

1.18 $$h = h[Q(G^*; w_1, w_2, \ldots;\ u_{11}, u_{12}, \ldots)]$$

one may maximize h with respect to the alternative sets of w available, which is the passive approach, or with respect to the u's, which is the active approach. Active models which have been presented include Tintner's method and our practical agricultural model VB.

The table of contents shows that most methods developed up to now are partly stochastic, i.e., in expression 1.15 either the net profits, or the resources, or the matrix coefficients are random. Only models listed sub IV cope with the joint distribution of all parameters.

Tinbergen says somewhere that it is rational and time-saving to mention in the introduction what is new in the essay at hand. According to this rule we point out that the original research presented here includes:

(1) A general probabilistic model with all relevant parameters subject to stochastic variation without the assumption of independence;

(2) a practical model for agriculture, including yield and price variation, with a new simplex criterion and computation of a stochastic program for agriculture in the State of Iowa;

(3) a more efficient procedure for Talacko's "boundary" method. In addition, Heady and Candler's method of programming with stochastic yields has been formalized and a new general formula presented for Tintner's active method.

A final remark concerning the use of symbols. In each of the subsequent outlines the notation of the respective author has been observed. This procedure is certainly questionable from the angle of notational consistency. However, since this survey is intended mainly as a guidance to those

interested in the methods under review, it is understood that they should be able to refer to the original for further details, without being puzzled once more by a new switch in conventions. We have not been able to create uniformity in the notations of different writers within the time limits set for the completion of this dissertation. Nuance and the complexity of various concepts have even compelled us to add some more symbols for the new methods presented here.

II. MODELS WITH A VECTOR OF STOCHASTIC RESTRICTIONS OR INPUTS

We start the discussion with models where all parameters are deterministic, except for the right-hand side vector of system 1.6. Taking the step from 1.6 to 1.14, as explained, it appears that $z^{(k)}$ is a transformation of c, hence transformation channels applied to a random input vector should also be grouped under this heading. Before considering confidence intervals we shall concentrate on the expected value of the objective function under these conditions.

A. Certainty-Equivalence Theorems

Madansky (38) considers a linear-programming problem: minimize

2.1 $$C(b,x) = c'x + f'y$$

with respect to x and y, under the restrictions

2.2 $$Ax + By = b \qquad (x,y \geqslant 0),$$

where the symbols denote:
C(b,x), a scalar, the objective function;
c and f, vectors of weights attached to the activities, with n_1 and n_2 components resp.;
x and y, activity vectors of n_1 and n_2 components resp.;
A and B, technical-coefficients matrices of dimensions $n_1 \times m$ and $n_2 \times m$ resp.;

b, the restriction victor of m components. Only the "right-hand side" or b vector is random, with known expectation Eb.

Since in practice one often tackles the problem by simply replacing b by Eb and then optimizing C(b,x) by the customary techniques, the author's contributions focus on the bounds (37) for the value of the objective function for the approximate solution, obtained after replacing b by Eb. He concentrates on the conditions of equality between the expected value of the optimum solution and the value of the optimum approximate solution.

It has been stated by Dantzig (14) that one cannot, in general, obtain an optimal solution by substituting Eb for b. Ferguson and Dantzig (17), in a study on the allocation of aircraft to routes under uncertain demand, obtained

2.3 $\min_x$ E C(b,x) = \$1,524,000

as compared to

2.4 $\min_x C(Eb,x)$ = \$1,000,000.

The paper (37) under consideration shows that the direction of this inequality may be generalized. Denoting by $\bar{x}(Eb)$ the decision which minimized $C[Eb,\bar{x}(Eb)]$, it follows that

2.5 $EC[b,\bar{x}(Eb)] \geqslant \min_x EC(b,x)$.

It is further demonstrated that

2.6 $\min_x EC(b,x) \geqslant E\min_x C(b,x)$,

while Vajda (58) proves:

2.7 $\quad E \min_x C(b,x) \geqslant \min_x C(Eb,x).$

The argument thus establishes the theorem

2.8 $\quad EC[b,\bar{x}(Eb)] \geqslant \min_x EC(b,x) \geqslant E\min_x C(b,x) \geqslant \min_x C(Eb,x),$

implying the generalization of the above outcome: 2.3 $\geqslant$ 2.4.

Now the following results are important in stochastic linear programming. For the last two members of 2.8 we have

2.9 $\quad E \min_x C(b,x) = \min_x C(Eb,x)$

if, and only if, $\min_x C(b,x)$ is a linear function of b, cfr. Savage (42, p. 265). If C(b,x) is a linear function of b, equality also holds between the second and third members of 2.8, and one can finally equate the second and fourth members

2.10 $\quad \min_x EC(b,x) = \min_x C(Eb,x).$

In ordinary linear programming this outcome is a formal account of the intuitively-acceptable idea that the optimum value of the objective-expectation is derived by the usual techniques applicable to a deterministic model, simply by substituting Eb for the stochastic b vector.

For practical purposes this statement is not immediately helpful because one is concerned with confidence limits rather than with expected values of the objective function.

In a two-stage program considered by Dantzig (14), first, a decision x is tried, whereafter the random b's are observed; at the second stage the activity vector y is used to compensate for "inaccuracies" in the first decision, assuming a feasible (x,y) for each (b,x). There one cannot in general replace the "right-hand side" by Eb: Madansky (38) calls it "the here-and-now problem". For this case, the essay goes into the computation of the upper and lower bounds on $\min_x$ EC(b,x), namely EC[b,$\bar{x}$(Eb)] and $\min_x$C(Eb,x), a matter which need not be pursued here.

B. Orthogonal-Channel Chains

In a recent publication (1, pp. 70-108) which is an exposition of principles regarding the formalization of management control functions, Adam describes a chain of orthogonal-transformation channels, the input messages of which are subject to "background noise" ("weisses Rauschen"), i.e., they are subject to a normal distribution. The technique may, therefore, be classified with the programming methods for a stochastic "right-hand side" vector. In the following lines the elements of this vector belong to a set with supposedly the same distribution. One may notice that the subsequent analysis is an interpretation, from the standpoint of information theory, of some of Linder's ideas (34, 35).

Exemplifying this idea by an agricultural experiment, let us denote by:

(n), the yield of a plot treated with N (nitrogen);
(p), the yield of a plot treated with P (phosphorus);
(k), the yield of a plot treated with K (potassium);
(1), the yield of the control plot.

Now we are interested in normalized output messages of the form

2.11 $$N^1 = \frac{(n)-(1)}{\sqrt{2}}$$

which give information about the normalized nitrogen effect. If we also have plots with 2 and 3 fertilizers combined, denote by:
$(1 \cap 1 \cap 1)$, the yield of the control plot;
$(n \cap p)$, the yield of a plot treated with N and P;
$(n \cap p \cap k)$, the yield of a plot treated with N, P, and K;
etc.

The normalized messages become:

$$N^1 = -(1 \cap 1 \cap 1)+(n \cap 1 \cap 1)-(1 \cap p \cap 1)-(1 \cap 1 \cap k)+(n \cap p \cap 1) +(n \cap 1 \cap k)-(1 \cap p \cap k)+(n \cap p \cap k)$$

2.12 $$N^1 \cap P^1 = (1 \cap 1 \cap 1)-(n \cap 1 \cap 1)-(1 \cap p \cap 1)+(1 \cap 1 \cap k)+ (n \cap p \cap 1)-(n \cap 1 \cap k)-(1 \cap p \cap k)+(n \cap p \cap k)$$

$$N^1 \cap P^1 \cap K^1 = -(1 \cap 1 \cap 1)+(n \cap 1 \cap 1)+(1 \cap p \cap 1)+(1 \cap 1 \cap k) -(n \cap p \cap 1)-(n \cap 1 \cap k)-(1 \cap p \cap k)+(n \cap p \cap k).$$

etc.

The interpretation of the signs in 2.12 follows the same logic as in the simplest case 2.11. One should bear in mind that expressions of the kind $N^1 \cap P^1$ are only intended to

single out the additional effect of the interaction among several components.*

A complete picture of the orthogonal-channel operation is given by the transformation

$$2.13 \qquad \begin{bmatrix} I \\ N^1 \\ P^1 \\ K^1 \\ N^1 \cap P^1 \\ N^1 \cap K^1 \\ P^1 \cap K^1 \\ N^1 \cap P^1 \cap K^1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 \end{bmatrix} \frac{1}{\sqrt{8}} \cdot \begin{bmatrix} 1 \cap 1 \cap 1 \\ n \cap 1 \cap 1 \\ 1 \cap p \cap 1 \\ 1 \cap 1 \cap k \\ n \cap p \cap 1 \\ n \cap 1 \cap k \\ 1 \cap p \cap k \\ n \cap p \cap k \end{bmatrix} .$$

or, in abbreviated notation,

$$2.14 \qquad y = Ax.$$

In order to have an idea about the "background noise" of the inputs we estimate their variance by performing more experiments with the same treatments, e.g. by building into the above channel a repetition of each experiment. We could, of course, repeat each experiment a different number of times, but let us stick to the simplest illustration of the principle.

---

*The results may be established by formal logic, viz., through deriving the matrix for conjunctive channels. The author shows that conjunctions of channels corresponds to the Kronecker matrix product just like a serial chain corresponds to the ordinary matrix product. Both products result in orthogonal matrices provided the components are so.

The system thus enlarged is shown in Table 1. The numerical coefficients have been introduced for the purpose of normalization.

Suppose, in the general case, that we dispose of m inputs (yield figures in the example) for k information targets (the elements of vector y, or the 8 first elements of the left-hand side vector in the enlarged system). One then obtains (m-k) informations S_i about the distribution of the inputs as the reader may verify from table 1. By the additivity theorem for independent chi-square distributions (26, p. 138), the mean S^2 of the S_i^2 may be used to estimate the variance of the stochastic inputs.

It is important that the transformation channel A in 2.14 be orthogonal; so, premultiplying both members by their transposes and remembering that $A'A=A^{-1}A=I$, we get

2.15 $\quad y'y = x'x.$

In the table we see that the right-hand side vector y may be partitioned into two parts z and s. The z elements stand for the k information targets, the s elements for the S_i (lower part of the left-hand side vector in Table 1); then 2.15 becomes

2.16 $\quad y'y = \begin{pmatrix} z \\ s \end{pmatrix}' \begin{pmatrix} z \\ s \end{pmatrix} = z'z + s's = x'x.$

Since

2.17 $\quad s's = (m-k)S^2,$

Table 1. Enlarged transformation channel for repeated experiments

Outputs y		Channel A	Inputs x (with repetitions*)
I		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1
N^1		-1 -1 1 1 -1 -1 -1 -1 1 1 1 1 1 -1 1 1	1*
p^1		-1 -1 -1 -1 1 1 -1 -1 1 1 -1 -1 1 1 1 1	n
K^1		-1 -1 -1 -1 -1 -1 1 1 -1 -1 1 1 1 1 1 1	n*
$N^1 \cap p^1$		1 1 -1 -1 -1 -1 1 1 1 1 -1 -1 -1 -1 1 1	p
$N^1 \cap K^1$		1 1 -1 -1 1 1 -1 -1 -1 -1 1 1 -1 -1 1 1	p*
$p^1 \cap K^1$		1 1 1 1 -1 -1 -1 -1 -1 -1 -1 -1 1 1 1 1	k
$N^1 \cap p^1 \cap K^1$		-1 -1 1 1 1 1 1 1 -1 -1 -1 -1 -1 -1 1 1	k*
.....	$= \frac{1}{\sqrt{16}}$	.....	
S_1		$\sqrt{8}$ $-\sqrt{8}$	$n \cap p$
S_2		$\sqrt{8}$ $-\sqrt{8}$	$(n \cap p)^*$
S_3		$\sqrt{8}$ $-\sqrt{8}$	$n \cap k$
S_4		$\sqrt{8}$ $-\sqrt{8}$	$(n \cap k)^*$
S_5		$\sqrt{8}$ $-\sqrt{8}$	$p \cap k$
S_6		$\sqrt{8}$ $-\sqrt{8}$	$(p \cap k)^*$
S_7		$\sqrt{8}$ $-\sqrt{8}$	$n \cap p \cap k$
S_8		$\sqrt{8}$ $-\sqrt{8}$	$(n \cap p \cap k)^*$

one derives immediately

2.18 $$S^2 = \frac{1}{m-k}(x'x - z'z),$$

so that one need not compute the s partition of the y vector.

The effect or target z_i--in the example the effect of a fertilizer treatment--is significant for

2.19 $$|z_i| > t(m-k;\eta).\ S$$

where t is Student's t with (m-k) degrees of freedom and $(1-\eta)$ is the adopted confidence level.

Quality control may be viewed along the same lines if the sequence of technical operations is symbolized by a chain of orthogonal channels. For a tolerance η, products will be accepted as long as the target qualities z_i, grouped in a vector z, satisfy

2.20 $$-e.t(m-k;\eta)\sqrt{\frac{s's}{m-k}} \leq z-z_o \leq e.t(m-k;\eta)\sqrt{\frac{s's}{m-k}},$$

where e is a vector all elements of which are one, and z_o the vector of means z_{io} of the z_i.

Application to cybernetics

An interesting aspect of those concepts is their applicability to cybernetics. Since any number of ideas, i.e., combinations of elements belonging to a finite "alphabet", may be registered in a digital code (a text composed of k characters allows for 2^k messages in a binary code), the use of computers

with preset courses (Programmsteuerung) has now brought synchronization of technical processing and management control within reach. Numerous processes (channels) characterized by sequences of inputs and outputs of "messages" can be reproduced by homogeneous affine transformation systems or matrices with real elements. The programming problems involved are essentially stochastic in the inputs; the description of their randomness, however, presents us with special problems. The kind of information required in management control is often broader, but less specified, than the usual sort of metrically graduated data (metric graduation scale, symbol $\leftrightarrow$) as prices, weights etc. The graduation scale may indeed be of merely topological definition, i.e., classificatoric as in ordinary language (diversity scale $\neq$) or comparative as in the subject-matter of non-parametric statistics (rank scale $<$).

The stochastic approach to these concepts consists in specifying their state of indetermination by the entropy H, an analogon to the variance in a metrically graduated universe. Consider a language V as a finite scheme of elements (symbols) V_i, classified only by diversity, (Khinchin, 28)

$$2.21 \qquad \begin{bmatrix} V_1 \neq V_2 \neq \ldots V_k \\ p_1 \quad p_2 \quad\quad p_k \end{bmatrix} \quad ; \quad p_i \geq 0 \; ; \qquad \sum_i p_i = 1 \; ;$$

where notation 2.21 means that every symbol V_i has a probability p_i of occurring in a message, the total probability adding up to one for any message. This language may, in principle,

be used to single out an empirical class $Y \subset X$ (for $\overline{Y} \cap Y = E$) among the set of imaginable worlds

2.22 $$X = \bigcup_{j(1)} \cdots \bigcup_{j(n)} \bigcap_{i=1}^{n} X_{ij(i)} \text{ , where } X_{ij(i)} \cap X_{ik(i)} = \begin{cases} X_{ij(i)} & \text{for } j=k \\ E & \text{for } j \neq k \end{cases} .$$

The indetermination of the language is measured by the entropy in the Boltzmann-Shannon sense

2.23 $$H = -\sum_i p_i \log p_i .$$

Automatic management control is now brought about by the following cybernetic principle. Let us resume the discussion of the above orthogonal-channel chains and assume a normal distribution of the target outputs, described this time in terms of entropy

2.24 $$P(|z_i - z_{io}| \leq u\sqrt{H}) = \Phi(u) , \quad \Phi(u) = \frac{1}{2\sqrt{\pi}} \int_{-u}^{+u} e^{-\frac{x^2}{2}} dx .$$

A filter criterion analogous to 2.20 is stored in the computer: upon violation a signal is fed back to the planning unit which thus learns to modify the set-up of the experiment under consideration.

Adam further introduces a stochastic evaluation measure for the possibility of automatic control. The "channel" (process) dispatches inputs coded in "language V" in order to yield

messages in a "language W" (the V and W messages correspond to x and y in 2.14). Call "measure of control" (Bestimmtheits-mass) the magnitude

2.25 $$B(W|V) = R(W|V) : H(W),$$

which is defined in terms of the rate of transmission

2.26 $$R(W|V) = H(W) - H(W|V).$$

It follows that $\max |B| = 1$. A process is amenable to automatic control provided

2.27 $$|B(z_i \mid V_1 \cup V_2 \cup \ldots)| > 1 - \varepsilon \qquad (i=1,2,\ldots),$$

where ε is a small positive number.

III. MODELS WITH STOCHASTIC INPUT COEFFICIENTS

Making some simplifying assumptions, an approximate distribution for the maximizing solution has been derived by Tintner (54, 56). The passive and active programming concepts have been presented in the introduction. It will be easier to make a clear distinction after the passive method has been developed.

A. Passive Approach

We remember from 1.14 that we can solve the equation

$$3.1 \qquad D^{(k)}z^{(k)} = c$$

for the k^{th} selection of activities, viz., the elements of $z^{(k)}$. Now consider the general problem of estimating the distribution of the solutions in the case of this non-homogeneous linear system when the coefficients are random.

Upon differentiation of 3.1 one has

$$3.2 \qquad dD^{(k)}.z^{(k)} + D^{(k)}.dz^{(k)} = dc,$$

where d is the differential operator. The mean value of $z^{(k)}$, as solved from 3.1, is

$$3.3 \qquad z^{(k)} = D^{(k)-1}.c \ ,$$

where D and c are taken at the means of their respective elements. Deriving the vector of deviations $dz^{(k)}$ from 3.2 we

get the linear expression

3.4 $$dz^{(k)} = -D^{(k)-1} \cdot dD^{(k)} \cdot z^{(k)} + D^{(k)-1} \cdot dc .$$

It is assumed, for simplicity's sake, that only the elements of $D^{(k)}$ are random, so the right-hand side term vanishes:

3.5 $$dz^{(k)} = -D^{(k)-1} \cdot dD^{(k)} \cdot z^{(k)} .$$

It will not be possible, however, to derive the exact distributions (22), so approximations around the means will be used, retaining only the linear terms of Taylor expansions.

The elements of $D^{(k)}$ are now supposed to be <u>independently and normally distributed</u> with known means and variances (known, e.g. by observation through time). Under these conditions the elements of the deviation vector $dz^{(k)}$, associated with the solution vector $z^{(k)}$, are also subject to a normal distribution. The approximation is supported by a theorem given by Cramer (12), which shows that the distribution of a function of a large number of independent random variables which possesses continuous first and second derivatives will tend to normal.

Unfortunately, the assumption of independence remains unrealistic, especially in agriculture: deviations of coefficients in the different columns of the technical-coefficients matrix are closely related to--though not wholly determined by--yield variation, and are thereby largely accounted for by the same meteorological conditions. This will be the case

especially for crops which cover the field during the same period of the year. The author advises, therefore, to subject the original variables to an orthogonal transformation in order to make them independent and then to work in the space of the independent random variables. Complications arise, however, because the transformation affects the constraints; this difficulty will be dealt with in a forthcoming publication by Hartley.* We have presented a solution to this problem, without transformation, in the general model IV C.

Writing out 3.5 one has

$$3.6 \qquad \begin{bmatrix} dz_1 \\ dz_2 \\ \cdots \\ dz_m \end{bmatrix} = - \begin{bmatrix} d_{11} & d_{12} \cdots d_{1m} \\ d_{21} & d_{22} \cdots d_{2m} \\ \cdots & \cdots\cdots \\ d_{m1} & d_{m2} \cdots d_{mm} \end{bmatrix}^{(k)} \begin{bmatrix} u_{11} & u_{12} \cdots u_{1m} \\ u_{21} & u_{22} \cdots u_{2m} \\ \cdots & \cdots\cdots \\ u_{m1} & u_{m2} \cdots u_{mm} \end{bmatrix}^{(k)-1} \begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_m \end{bmatrix}^{(k)} .$$

Comparing 3.5 to 3.6 the reader will notice that, for reasons of computational convenience, the variations of the d_{ij} have been replaced by standardized normal variates $u_{ij}=(d^*_{ij}-d_{ij})/s_{ij}$, where the numerator is the deviation of the observed value d^*_{ij} from the mean d_{ij}. The variates u_{ij} have mean zero and variance one. It should be noticed that one first needs to solve the ordinary, deterministic problem in order to obtain solution 3.3 for the z's and plug it into 3.6.

---

*Hartley, Herman O., Iowa State University of Science and Technology, Statistical Laboratory, Ames, Iowa. Forthcoming paper on non-linear programming. Private communication. 1960.

The article at hand exemplifies the computation of a program for an Iowa farm, taking into account only two crops (corn, flax) and two resource restrictions (land, capital) because of the extensive calculations. We shall try here to continue the argument in terms of the general case.

Writing out system 3.6 we get

$$3.7 \qquad dz_i^{(k)} = -\sum_{l=1}^{m}\sum_{j=1}^{m} d_{ij}^{(k)} u_{jl}^{(k)} z_l^{(k)} \qquad (i=1,2,\ldots,m),$$

for the deviation of the i^{th} element of the $z^{(k)}$ vector. For the k^{th} selection of activities the profit function takes the form

$$3.8 \qquad \underline{G}^{(k)} = \sum_{l=1}^{m} p_l^{(k)(k)} z_l \quad ,$$

where $p_1^{(k)}$ is the net profit derived from the unit level of the 1^{th} activity in the k^{th} selection. If the z's are subject to variation, $\underline{G}^{(k)}$ in 3.8 is only the expected value of $G^{(k)}$. Leaving room for variation we could write instead

$$3.9 \qquad G^{(k)} = \sum_{l=1}^{m} p_l^{(k)} (z_l^{(k)} \pm dz_l^{(k)})$$

or substituting the value of the $dz_1^{(k)}$ given in 3.7

$$3.10 \qquad G^{(k)} = \sum_{l=1}^{m} p_l^{(k)} z_l^{(k)} \mp \sum_{i=1}^{m}\sum_{l=1}^{m}\sum_{j=1}^{m} d_{ij}^{(k)} u_{jl}^{(k)} p_l^{(k)} z_l^{(k)} \quad ,$$

which is a linear form in the m^2 different u_{jl}, and may, therefore, be written

3.11 $$G^{(k)} = \sum_{1=1}^{m} p_1^{(k)} z_1^{(k)} + a_1^* u_{11}^{(k)} + a_2^* u_{12}^{(k)} + \ldots + a_{m^2-1}^* u_{m.m-1}^{(k)} + a_{m^2}^* u_{mm}^{(k)} ,$$

where the a^*'s are some constants, viz., products of the d_{ij}, p_1 and z_1.

The distribution of $G^{(k)}$ is now approximated numerically by assigning the discrete levels 0, ± 1, and ± 2 to the u's. Since the latter are independent and normal with mean 0 and variance 1, we are able to estimate the probability of all possible combinations for the three levels for all the u's in 3.9. In this fashion we have roughly described the joint probability distribution of the m^2 u's, and thus of $G^{(k)}$, in function of the parameter space S.

Owing to the non-negativity condition imposed upon the z's (real or disposal activities cannot be included at negative levels), we have to discard all points in S, corresponding to joint values of the u's such that

3.12 $$z_i^{(k)} - dz_i^{(k)} < 0 \qquad [(k) = (1),(2),\ldots,(K)],$$

i.e., cfr. 3.7,

3.13 $$z_i^{(k)} > \sum_{1=1}^{m} \sum_{j=1}^{m} d_{ij}^{(k)} u_{j1}^{(k)} z_1^{(k)} .$$

The subregion of S where 3.13 is fulfilled is called $S^{(k)}$.

We have to apply the same procedure to each of the $K = C_m^{m+n}$ possible selections of m activities. We compute all the corresponding objective functions $G^{(1)}$, $G^{(2)}$,..., $G^{(K)}$ at all feasible points as described by 3.13 .

So one obtains the overlapping, "feasible" subregions $S^{(k)}$ of S. The parameter space may also be divided into K subregions $T^{(k)}$ with the property that in $T^{(k)}$

3.14 $$G^{(k)} = \max_i G^{(i)}.$$

By their definition these subregions are non-overlapping. Define $U^{(k)}$ as

3.15 $$U^{(k)} = S^{(k)} \cap T^{(k)} ,$$

meaning the subregion where selection k is both feasible and optimum.

What we are interested in is the distribution of the maximum solution G^*, which coincides with $G^{(1)}$ in region $U^{(1)}$, with $G^{(2)}$ in region $U^{(2)}$,..., with $G^{(k)}$ in region $U^{(k)}$, hence

3.16 $$G^* = (G^{(k)} \mid U^{(k)}) \qquad (k=1,2,\ldots,K).$$

Since we have computed all the $G^{(k)}$ at each point of the parameter space (in practice only at selected points) we have a

numerical approximation to the distribution of G^*, if at each point considered we take the largest $G^{(k)}$.

For a large, "life-size" matrix, the computations require powerful electronic machines, especially if one wants a more refined numerical analysis, thereby considering more than three values of the standardized normal variates.

Let

3.17 $\quad Q(G^*; w_1, w_2, \ldots)$

denote the distribution obtained for G^*, characterized by parameters w_i, and assume that the entrepreneur has a (subjective) preference functional

3.18 $\quad h[Q(G^*; w_1, w_2, \ldots)]$.

He will choose the operational conditions which maximize 3.18. In the case of a farmer, who has to decide between different plots of land or different regions of possible settlement, the choice will fall upon the region with the set of relevant distribution characteristics such that 3.18 is maximum. In our case these characteristics are the means and variances of the input coefficients.

B. Active Approach

The foregoing approach is passive in that one simply computes the distribution of optimum solutions resulting from a <u>given situation</u> (characterized by parameter variation). A

decision is required only at the stage of the preference function, as to which distribution, i.e., which given situation, e.g. region, will be favored. It is desirable to decide at an earlier stage and to relate the stochastic choice criterion to the use of resources. This is the active approach, as described by Tintner in Econometrica (forthcoming). We shall here use his notations but proceed in a somewhat different fashion, thereby generalizing a criterion as to which activities are included in the objective function.

Denote by

3.19 $$p = a'x$$

the linear objective function to be maximized where a and x are vectors of net profits and activity levels respectively. The linear inequalities or constraints now take the form

3.20 $$b_{ij}x_j \leq c_i u_{ij} \qquad (i=1,2,\ldots,m;\ j=1,2,\ldots,n)\ (x_j \geq 0)$$

3.21 $$\sum_{j=1}^{n} u_{ij} = 1 \qquad (i=1,2,\ldots,m)$$

3.22 $$0 \leq u_{ij} \leq 1 ,$$

where the symbols represent:

b_{ij}, a coefficient of the technical matrix B;

c_i , a restriction or amount of resource available;

u_{ij}, the proportion of resource c_i reserved for activity x_j;

The u_{ij} are decision variables* and we denote their matrix by

3.23 $$U = [u_{ij}] .$$

One assumes, as before, a known joint probability distribution

3.24 $$P = P(a,B,c),$$

we hypothesize e.g. that the parameters are normally distributed and their means and variances known by observation. In addition, a priori decisions affect the totality of the resources. With the same numerical method as in the passive case one can derive an approximation to the probability distribution of the anticipated net profits p

3.25 $$R(p; U)$$

which will now depend on the choice of the decision variables u. If a preference functional

3.26 $$f[R(p;U)]$$

is defined we can maximize it with respect to the u_{ij}.

The author illustrates with an application to Iowa agriculture which we shall briefly describe because the article has not yet been published; the data are mentioned in detail

---

*They should not be confounded with the standardized variates mentioned sub the passive approach. The latter are again used in the numerical approximation, but are not further mentioned.

by Babbar (6, 7). Denote by:

u_{11} (resp. u_{21}), the proportion of land (resp. capital) reserved for corn;

u_{12} (resp. u_{22}), the proportion of land (resp. capital) reserved for flax.

Resources : 148 acres of land and \$1800 of net disposable capital.

Prices : \$1.56 per bu. corn and \$3.81 per bu. flax.

One has to find the distribution of

$$3.27 \qquad p = 1.56x_1 + 3.81\,x_2,$$

under the conditions:

$$3.28 \qquad \begin{aligned} &b_{11}\,x_1 \leq 148\,u_{11}\ ;\ b_{21}\,x_1 \leq 1800\,u_{21}\ ; \\ &b_{12}\,x_2 \leq 148\,u_{12}\ ;\ b_{22}\,x_2 \leq 1800\,u_{22}\ ; \end{aligned}$$

and

$$3.29 \qquad \begin{aligned} &u_{11} + u_{12} = 1\ ; \\ &u_{21} + u_{22} = 1\ ; \qquad (0 \leq u_{ij} \leq 1)\ . \end{aligned}$$

In deriving the R(p; U), I have found the following values for p:

$$p = \frac{1.56 \times 148u_{11}}{b_{11}} + \frac{3.81 \times 1800u_{12}}{b_{12}} \quad \text{if} \quad \frac{148u_{11}}{b_{11}} < \frac{1800u_{21}}{b_{21}};$$

$$\frac{148u_{12}}{b_{12}} < \frac{1800u_{22}}{b_{22}}\ ;$$

(continued)

3.30

$$p = \frac{1.56 \times 148u_{11}}{b_{11}} + \frac{3.81 \times 1800u_{22}}{b_{22}} \text{ if } \frac{148u_{11}}{b_{11}} < \frac{1800u_{21}}{b_{21}};$$

$$\frac{148u_{12}}{b_{12}} > \frac{1800u_{22}}{b_{22}};$$

$$p = \frac{1.56 \times 148u_{21}}{b_{21}} + \frac{3.81 \times 1800u_{12}}{b_{12}} \text{ if } \frac{148u_{11}}{b_{11}} > \frac{1800u_{21}}{b_{21}};$$

$$\frac{148u_{12}}{b_{12}} < \frac{1800u_{22}}{b_{22}};$$

$$p = \frac{1.56 \times 148u_{21}}{b_{21}} + \frac{3.81 \times 1800u_{22}}{b_{22}} \text{ if } \frac{148u_{11}}{b_{11}} > \frac{1800u_{21}}{b_{21}};$$

$$\frac{148u_{12}}{b_{12}} > \frac{1800u_{22}}{b_{22}}.$$

From the first line of conditions 3.30 we see, for instance, that the amount of corn produced is restricted to $148u_{11}/b_{11}$ bushels because the most restrictive resource which, by decisions u_{11} and u_{21} has been assigned to corn is land, given the acreage and capital requirements b_{11} and b_{21} per bushel.

In order to generalize the set of equations and inequalities 3.30 we could introduce the following formula for the objective function p which depends on n activities and m restrictions:

3.31 $$p = \sum_{i=1}^{n} \frac{p_i c_i u_{k(i)i}}{b_{k(i)i}}$$

where we get the $u_{k(i)i}$ from the condition

$$3.32 \qquad \frac{c_i u_{k(i)i}}{b_{k(i)i}} = \min_j \left[\frac{c_i u_{ji}}{b_{ji}}\right] (j=1,2,\ldots,m;\ i=1,2,\ldots,n).$$

The subscript k(i)i denotes that k is, in general, a different ordinal number for each i and thus is itself a function of i, defined by 3.32.

Example. For the above-quoted application to Iowa agriculture .023 and .092 acres of land were required per bu. corn and flax respectively, so $b_{11} = .023$ and $b_{12} = .092$. The respective capital requirements in \$/bu. were $b_{21} = .318$ and $b_{22} = .970$. As for the decision variables, suppose the farmer reserves $u_{11} = \frac{1}{2}$ and $u_{12} = \frac{1}{2}$ of the land for corn and flax respectively, while the capital will be available to these two crops in the proportions $u_{21} = 3/4$ and $u_{22} = 1/4$. The inequalities of 3.30 become in this particular case:

$$3.33 \qquad \frac{148\ u_{11}}{b_{11}} = 3230 < \frac{1800\ u_{21}}{b_{21}} = 4250;\ \frac{148 u_{12}}{b_{12}} = 804 >$$

$$> \frac{1800\ u_{22}}{b_{22}} = 465.$$

For subscript i = 1 it is clear from 3.33 that

$$3.34 \qquad \min_j \left[\frac{c_i u_{ji}}{b_{ji}}\right] = \frac{148\ u_{11}}{b_{11}}$$

while for subscript i = 2 one has

$$3.35 \qquad \min_j \left[\frac{c_i u_{ji}}{b_{ji}} \right] = \frac{1800\, u_{22}}{b_{22}} \quad .$$

Hence, in this case we see that in 3.31 k=k(i) is equal to 1 for i=1 and is equal to 2 for i=2, so 3.31 becomes

$$3.36 \qquad p = \frac{p_1 c_1 u_{11}}{b_{11}} + \frac{p_2 c_2 u_{22}}{b_{22}}$$

which is, of course, in agreement with 3.30. In other words, formulae 3.31 and 3.32 mean nothing else than that the u_{ij} are ex ante decisions: their ex post realization depends on 3.32. What is decided on beforehand, however, is which activities will be in the plan; viz., any i^{th} activity for which $u_{ji} > 0$ (j=1,2,...,m).

Let us now come back to stochastic considerations. Suppose, as before, that only the elements of the B matrix are subject to variation. Again we shall use standardized variates for the distribution of the b's, and assign them only the values 0, ± 1, ±2. If the entrepreneur has allocated to each activity its part of the resources, i.e., determined by decision the value of the n different $u_{k(i)i}$ in 3.31, we may, as above, numerically compute discrete points of the distribution of p by calculating its value for all possible combinations of the n standardized normal variates, each assuming 3 values. We discard values for which

$$3.37 \qquad x_i = \frac{e_i u_{k(i)i}}{b_{k(i)i} \pm \beta_{k(i)i}} < 0 \quad ,$$

where $b_{k(i)i}$ represents the expectation of the coefficient and $\beta_{k(i)i}$ the related standardized normal variate.

One may, however, be interested in approximating the mathematical expectation of p for all possible decisions.

$$3.38 \qquad Ep = \int p.R(p; U)\, dp$$

where the integral is to be taken over the whole range of variation of p. The numerical approximation is carried out by computing p for the different values of the standard variates as before, and repeating this for all combinations of some discrete values of the u's in 3.31. Each of them can be assigned, e.g., values 0, 1/4, 1/2, 3/4, 1. Only combinations of levels obeying 3.21 are acceptable.

Likewise, somebody who does not see his business primarily as a long-run profit source (in which case he maximizes Ep) but looks at it from the conservative standpoint may compute, e.g., the lower .05 confidence level $p_{.05}$ over the whole variation range etc. Those u's are selected which maximize Ep, respectively $p_{.05}$.

IV. GENERAL MODELS

We recall the linear-programming problem described in the introduction:

maximize

4.1 $P = a'x$

subject to the linear restraints

4.2 $Bx \leq c$

4.3 $x \geq 0$,

where a and x are column vectors of n components, B an m x n matrix, and c a column vector of m elements, while ' denotes transposition. Now all or several elements of a, B, and c may be of a probabilistic nature and interdependent. Three models dealing with this situation are presented in the following paragraphs.

A. Approximation by a Quotient Distribution for Normal Variates

Babbar (6, 7) expresses the solutions as a quotient of two determinants, which he keeps linear in the deviations by dropping higher-order terms. He then applies Geary's results (22) on the frequency distribution of the quotient of two normal variates.

Consider the distribution of an objective function

4.4 $y = (C + c)' X$

under the restraints

4.5 $(B + b) X = (Q + \varepsilon)$.

The latter are again written as an equality, since this is, for any chosen selection including the optimal combination of activity levels, the way in which X is solved.

C and X are n dimensional column vectors, B is an m x n technical-coefficients matrix, and Q is an m dimensional column vector. The lower-case letters are vectors and matrices of deviation; they have appropriate dimensions.

Denote by

$|D^k|$, the determinant of B after replacing its k^{th} column by Q;
$|D^k + d^k|$, idem for replacement by $(Q + \varepsilon)$;

The objective function y, after substitution of the solution of 4.5,

4.6 $$x_k = \frac{|D^k + d^k|}{|B + b|} \qquad (k=1,2,\ldots, m),$$

becomes

4.7 $$y = \frac{1}{|B+b|} \sum_{r=1}^{m} (C_r + c_r) \cdot |D^r + d^r| \cdot$$

It will be noticed that expressions 4.6 and 4.7 may be written as quotients in functional notation:

$$4.8 \qquad x_k = \frac{N(x_k)}{D(x)}$$

$$4.9 \qquad y = \frac{N(y)}{D(y)} \quad .$$

The author uses the following approximation: the functions $N(x_k)$, $D(x)$, $N(y)$, and $D(y)$ are, of course, obtained by expanding the determinants in 4.6 and 4.7. Ignoring all products of errors of the second and higher orders in this expansion, the expressions obtained for the above functions are linear in the (supposedly) normally distributed errors and, in turn, subject to a normal distribution.

We can now apply Geary's theorem referred to before: If N and D are normally-distributed variables with E(N) and E(D) as the mean values, σ_N^2, σ_D^2 as variances and σ_{ND} as covariance, and $Z = N/D$ is the quotient, then the expression

$$4.10 \qquad t = \frac{E(D)Z - E(N)}{(\sigma_N^2 - 2\sigma_{ND}Z + \sigma_D^2 Z^2)^{\frac{1}{2}}}$$

is approximately normally distributed with mean zero and variance one, provided

$$4.11 \qquad E(D) > 3\sigma_D,$$

i.e., D should have a coefficient of variation lower than 33%.

Confidence limits for Z (and thus for the x_k and for y) are established as follows. If the probability coefficient is a fraction, .95 say, the standard normal tables give a positive number γ such that the probability of the absolute value of the standard normal variate being greater than γ is $(1-\alpha)$. One has

$$4.12 \qquad P\left[\left|\frac{E(D)Z - E(N)}{(\sigma_N^2 - 2\sigma_{ND}Z - \sigma_D^2 Z^2)^{\frac{1}{2}}}\right| \leq \gamma\right] = \alpha,$$

i.e.,

$$4.13 \qquad P[\{E(D)^2 - \gamma^2\sigma_D^2\}Z^2 - 2\{E(D).E(N) - \gamma^2\sigma_{ND}\}Z + \{E(N)^2 - \gamma^2\sigma_N^2\} \leq 0] = \alpha,$$

where P stands for a probability statement. The left-hand side of 4.13 is of the second degree in Z; putting

$$4.14 \qquad \{E(D)^2 - \gamma^2\sigma_D^2\}Z^2 - 2\{E(D).E(N) - \gamma^2\sigma_{ND}\}Z + \{E(N)^2 - \gamma^2\sigma_N^2\} = 0$$

and solving for Z, one obtains the two limits between which Z will lie with 100α% probability.

This result can now be applied to find confidence intervals for the x_k in 4.8 and for y in 4.9 directly, without the help of Pearson's bivariate tables (41).

Fieller's procedure (18) concerning the cumulative distribution is also mentioned in this paper. He computes the

chance of obtaining a value of the variable Z = N/D not less than v, viz.,

$$4.15 \qquad [1 - F(v)] = \int_{h}^{\infty}\int_{k}^{\infty} N(\varrho)\, dx\, dy + \int_{-h}^{\infty}\int_{-k}^{\infty} N(\varrho)\, dx\, dy,$$

where $N(\varrho)$ is the bivariate normal distribution (of x and y) with zero means, both variances equal to unity and correlation coefficient ϱ. The value of h is the coefficient of variation for D; and k is computed as follows:

$$4.16 \qquad k = \frac{E(N) - vE(D)}{[\sigma_N^2 - 2\sigma_{ND} + \sigma_D^2 v^2]^{\frac{1}{2}}}.$$

Some complications arise if either or both of h and k are negative. The most interesting result for practical purposes is formula 4.13 because the confidence level is what we are especially interested in.

The author has computed a practical application to agriculture.

B. Replacing Variates by Their Boundaries

There is a possibility of deriving the boundaries of solution deviations, from given confidence boundaries of the coefficients. This has been worked out by Talacko (44). There is, however, a more efficient procedure for solving the problem along these lines, which we shall present in a comment on Talacko's method.

The approach is basically non-parametric.

First, consider the maximization problem of our usual, constrained objective function

4.17 $F(x) = \max c'x$

$A x \leq b, \qquad (x \geq 0)$

where c and x are n dimensional column vectors, A an m x n matrix and b an m dimensional column vector. A and c are deterministic, and the interval of $b_i^- \leq b_i \leq b_i^+$ is known with a given probability $p(b_i)$.

Taking only the b_i^- and b_i^+ boundaries would enable us to find $2^m \times C_m^{m+n}$ values F_o by trivial solution, viz., 2^m values in each of the C_m^{m+n} possible selections (and not only a total of 2^m, as the author asserts), under appropriate non-degeneracy assumptions. The case reduces to consideration of only two solutions:

4.18 $Ax \leq b^-$ with $F_o^-(x) = \max c'x$

$Ax \leq b^+$ with $F_o^+(x) = \max c'x$

The result is then generalized for stochastic elements of A and c as well.

All, or some, of the technical coefficients a_{ij} of A, the restrictions b, and the weights c attached to the activities by the objective function are now random variables.

4.19 $S_i : a_{ij}\ ,\ b_i\ ,\ c_j\ .$

Si lies in the interval $(S_i^-\ ,\ S_i^+)$ defined by its lower and

upper boundaries, with a probability $p(S_i)$. These random variables must satisfy the constraints

$$
\begin{aligned}
& 0 \leq a_{ij}^{-} \leq a_{ij} \leq a_{ij}^{+} \\
4.20 \qquad & 0 \leq b_i^{-} \leq b_i \leq b_i^{+} \\
& c_j^{-} \leq c_j \leq c_j^{+} ,
\end{aligned}
$$

the first two of which are very stringent and reduce considerably the applicability range of the model. Why they are necessary will become clear when the solutions are presented.

The argument is developed as follows. Denoting by W the set of all x satisfying relations 4.17 for various a_{ij} and b_j conforming to 4.18, the author then proves

Lemma I:

There exist vectors u and u* such that $u \in W$ and $u^* \in W$, and for all other $x \in W$,

$$4.21 \qquad F(u^*) \leq F(x) \leq F(u).$$

Let W^+ and W^- stand for the vectors x satisfying respectively

$$4.22 \qquad \sum_{j=1}^{n} a_{ij}^{-} x_j \leq b_i^{+}$$

and $\qquad (i = 1,2,\ldots,m)$

$$4.23 \qquad \sum_{j=1}^{n} a_{ij}^{+} x_j \leq b_i^{-} .$$

One may then demonstrate

Lemma II:

If W corresponds to any fixed set of intermediate values of the a_{ij} and c_i within ranges 4.20, there exists an inequality

4.24 $$W^- \leq W \leq W^+$$

and obviously

4.25 $$F^-(x) \leq F(x) \leq F^+(x)$$

for

4.26 $$F^+(x) = \sum_{j=1}^{n} c_j^+ x_j \quad \text{and} \quad F^-(x) = \sum_{j=1}^{n} c_j^- x_j \qquad (x \geq 0).$$

Theorem Let max F(x) correspond to the problem involving any intermediate values for the S_i. Denoting by $\max^+ F^+(x)$ and $\max^- F^-(x)$ the maximum values $F^+(x)$ over set W^+ and $F^-(x)$ over set W^- respectively, we have

$$\max^+ F^+(x) \geq \max F(x) \geq \max^- F^-(x).$$

Solution The upper and lower limits of the optimum functional can thus be found by solving only two problems:

4.27 $$F_o^- = \max \sum_{j=1}^{n} c_j^- x_j \quad \text{for} \quad \sum_{j=1}^{n} a_{ij}^+ x_j \leq b_j^-$$

and $(i=1,2,\ldots,m)$

4.28 $$F_o^+ = \max \sum_{j=1}^{n} c_j^+ x_j \quad \text{for} \quad \sum_{j=1}^{n} a_{ij}^- x_j \leq b_j^+ \ .$$

Inspection of conditions 4.27 and 4.28 shows the sense of non-negativity conditions in 4.20; they may be dropped in the simpler case, with only b stochastic, as is easily verified.

If N different S_i are involved, the probability $p(F_o)$ that F_o lies in the interval (F_o^-, F^+) must satisfy

$$4.29 \qquad \prod_{i=1}^{N} p(S_i) \leq p(F_o).$$

An example is given for the model with only the b_i stochastic and subject to a Poisson distribution. Estimating $\sigma_i = \sqrt{b_i}$ and fixing the boundaries at the .95 confidence limits, conditions 4.27 and 4.28 become:

$$4.30 \qquad \begin{aligned} F_o^- &= \max \sum_{j=1}^{n} c_j x_j \text{ for } \sum_{j=1}^{n} a_{ij} x_j \leq b_i - 2\sqrt{b_i} \\ F_o^+ &= \max \sum_{j=1}^{n} c_j x_j \text{ for } \sum_{j=1}^{n} a_{ij} x_j \leq b_i + 2\sqrt{b_i} \quad . \end{aligned}$$

Further numerical examples illustrate the generalized approach.

Increasing the method's efficiency

From system

$$4.31 \qquad Ax = b \ ,$$

where A is an m x m matrix with elements a_{ij} and x, b are column vectors of m components, we solve the i^{th} element of the x vector as

$$4.32 \qquad x_i = \sum_{j=1}^{m} a^{ij} b_j$$

with the a^{ij} being the elements of the inverse matrix A^{-1}.

For a model where only the a's or input coefficients of the technical matrix are random*, the author proves, in summary, the intuitively acceptable theorem

$$4.33 \qquad \bar{x}_i = \sum_{j=1}^{m} a^{+ij} b_j,$$

where $\bar{x}_i$ denotes the lower boundary of x_i and a^{+ij} is an element of the inverse matrix of A where, before inversion, every element a_{ij} has been taken at its upper boundary $a^{+}ij$.

Suppose that we let the "boundaries" coincide with the $(1-\alpha)$ confidence limits. Since the probability that all a_{ij} are at their higher .95 confidence level (say) is extremely small, the "efficiency" of the model is low. We shall try to improve upon this procedure by determining the confidence interval for all the variables at such a level that their combined occurrence has the probability of $(1-\alpha)$.

For the sake of notational simplicity, denote the m^2 elements of the A matrix, viz., the original a_{ij}, as $a_1, a_2, \ldots, a_i, \ldots, a_{m^2-1}, a_{m^2}$. The probability p^o that any of these elements, with means $\underline{a}_i$, exceeds its upper confidence limit a_i^+ is, in case they are independent (or orthogonalized) (40, p. 29):

---

*We shall later give a normalization method which incorporates variability of the profit and resource vectors into the technical matrix. We focus therefore upon the latter's variability.

$$4.34 \qquad p^o = p^o\,(a_1 > a_1^+ \text{ or } a_2 > a_2^+ \text{ or } \ldots \text{ or } a_{m2} > a_{m2}^+)$$

$$= \sum_{i=1}^{m^2} P(a_i > a_i^+) - \sum_{i,j} P(a_i > a_i^+ \,,\; a_j > a_j^+) +$$

$$\mp \sum_{i,j,k} P(a_i > a_i^+ \,,\; a_j > a_j^+ \,,\; a_k > a_k^+) - \ldots \pm$$

$$+ P(a_1 > a_1^+,\; a_2 > a_2^+,\; \ldots,\; a_{m2} > a_{m2}^+) \quad .$$

Only a few right-hand side terms of 4.39 need actually be computed, because the higher-order terms rapidly tend to zero.

For computational convenience we replace the a_i by standardized normal variates.

$$4.35 \qquad y_i = \frac{a_i - \underline{a}_i}{s_i} \quad ;$$

in this formula s_i estimates the standard deviation of a_i. The probability that each y_i exceeds at the same time some (standardized) deviation v_i may be derived from their joint, supposedly normal distribution as a product of integrals

$$4.36 \qquad \prod_{i=1}^{m^2} \int_{v_i}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{\frac{-y_i^2}{2}}\, dy_i \quad .$$

Calculate p^o from 4.34. We want to determine which is the standardized deviation v, the same for every y_i, so that the joint probability 4.36 of being exceeded is equal to p^o. This gives the equation

$$4.37 \qquad \prod_{i=1}^{m^2} \int_{v}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{\frac{-y_i^2}{2}}\, dy_i = p^o$$

i.e.

$$4.38 \qquad \int_{v}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy = \frac{m^2}{\sqrt{p^o}}$$

In the normal tables one finds the v corresponding to this value of the integral. From 4.35 we finally derive

$$4.39 \qquad (a_i^+)^o = \underline{a}_i + v.s_i$$

$$(a_i^-)^o = \underline{a}_i - v.s_i \ ,$$

where the new notations $(a_i^+)^o$, $(a^-)^o$ denote that these estimates will, in general, be different from their counterparts a_i^+, a_i^- . We may introduce

$$4.40 \qquad 100 \frac{a_i^+ - (a_i^+)^o}{s_i} \%$$

as a measure of the increase in "efficiency" in terms of the standard error of a_i.

Example If there were only two* a_i, say

$\underline{a}_1 = 50$, with $s_1 = 10$;

$\underline{a}_2 = 40$, with $s_2 = 15$;

the upper .95 confidence limits ($\alpha = .05$) would be:

---

*Since A is a square matrix there should be at least four a_i but this does not affect the argument, which is purely expository.

$$a_1^+ = \underline{a}_1 + 1.96\ \sigma_{a_1} = 69.6\ ;$$

4.41

$$a_2^+ = \underline{a}_2 + 1.96\ \sigma_{a_2} = 69.4\ ;$$

which are the boundaries Talacko would consider. However, the chances that they exceed these boundaries together are only $(.025)^2 = .0006$. Instead we consider the probability that either one exceeds these same boundaries; from 4.34:

4.42 $$P^o = (.025 + .025) - (.025 \times .025) = .049.$$

The individual, standardized deviations of a_1 and a_2 having the probability P^o of occurring together would be equal to v as found from equation

4.43 $$\frac{1}{\sqrt{2\pi}} \int_v^\infty e^{\frac{-y^2}{2}}\, dy = \sqrt{.049} = .22$$

i.e.,

4.44 $$v = .77\ .$$

Fixing the confidence levels for the individual a_i at .95 is, therefore, <u>equivalent</u> to assigning them, when considered jointly, an upper boundary

4.45 $$(a_i^+)^o = \underline{a}_i + .77\ s_i\ .$$

For $(a_1^+)^o$ we get 57.7, which is an increase in "efficiency" of

4.46 $$\frac{100\ (69.6 - 57.7)\%}{10} = 119\%.$$

In formulae 4.27 and 4.28 we compute F_o^- and F_o^+ as before, but substitute $(a_i^+)^o$ for a_i^+ . We could as well write $(a_{ij}^+)^o$, remembering that we replaced the a_{ij} by a_i for notational simplicity.

It should again be borne in mind that all these considerations concern every single system among the C_m^{m+n} systems of equations considered by the selection method, as pointed out before.

C. Selection of the Maximum Lower-Confidence Limit: a New Model

The model presented below is again general in that it takes up a linear-programming problem where all parameters (prices, resources, and technical or input coefficients) are random; they may also be correlated. The distinctive feature of this method is a direct computation of the variance of the objective function, which is simpler, yet just as accurate as in the other models. It will be shown that the same degree of approximation as in Tintner's and Babbar's models is used, in as far as linearity in the error terms is concerned in the development of a matrix with stochastic elements. No other assumptions are, however, necessary.

From the original, known or observed, variance of the parameters, the variance of the objective function is derived, theoretically for all the $K = C_m^{m+n}$ solutions obtained

by the selection method, but in practice only for a (sufficiently large) number of solutions found in the vicinity of the non-stochastic optimum vertex of the convex hyperpolyhedron. If the entrepreneur takes risk into account he will not be primarily interested in the expected (or long-run) optimum solution. In the most usual, or conservative, case he will prefer the program with the highest lower $(1-\alpha)$ confidence threshold for the objective function g, i.e., g^o is preferred provided

$$4.47 \qquad g^o - vs_{g^o} = \max_k (g_k - vs_{g_k}) \qquad (k=1,2,\ldots,K),$$

where s_{g^o} is a standard-deviation estimator. Assuming that the g_k are distributed in a not too anormal fashion, v corresponds to a chosen $(1-\alpha)$ lower confidence level evaluated from tabulated areas under the normal probability curve (3, 40, 55):

$$4.48 \qquad \alpha = \tfrac{1}{2} - \int_0^v \Phi(t)\, dt.$$

Our goal will thus be reached if we succeed in computing s_{g_k} for any k.

It has been shown in the introduction that the maximization of a linear function under restraints given by linear inequalities can be solved, under certain conditions, by computing the objective function for each of the K solutions obtained by the selection method. In each of these cases our stochastic

problem reduces, therefore, to finding the distribution of

4.49 $g = p'\underline{z}$ $(\underline{z} \geqslant 0)$

subject to

4.50 $\underline{B}\,\underline{z} = c$

where the symbols denote:
g, a scalar, the profit or objective function;
p, a column vector of m net prices;
$\underline{z}$, a column vector of m activity levels;
$\underline{B}$, an m x m matrix of technical coefficients;
c, a column vector of m amounts of resources;
(k subscripts for g and z have been dropped for notational simplicity).

Supposing that the elements of p, $\underline{B}$, c are random variables, we want to derive confidence limits of $\underline{z}$ and g for each of the K solutions or, geometrically speaking, for each of the vertices of the convex hyperpolyhedron resulting from a mapping of the operational restrictions into the space defined by the activities.

Normalization for fixed c and p vectors The statistical information is supposed to provide us with T, e.g. yearly, observations p_t, c_t, and $\underline{B}_t$, where t takes the values 1,2,...,T. Expressing element $\underline{b}_{ijt}$ of $\underline{B}_t$ per unit of p_j and per unit of c_i, one gets

$$4.51 \qquad b_{ijt} = \frac{\bar{b}ijt}{p_t \, c_t} ,$$

where b_{ijt} is a normalized element. In matrix notation:

$$4.52 \qquad B_t = C_t^{-1} \, \bar{B}_t \, P_t^{-1} ,$$

where B_t is the normalized matrix (i.e., one of C_m^{m+n} such matrices) for the t^{th} period and C_t, P_t are diagonal matrices with the elements of c_t, p_t along the diagonal such that

$$4.53 \qquad c_t = C_t \{1\}$$
$$p_t = P_t \{1\}$$

if $\{1\}$ is the column vector with all elements equal to one. The solution vector corresponding to a normalized matrix is denoted by z (without bar).

In economic terms p, z, B, and c stand for net profits, activity levels, input coefficients, and amounts of resources, respectively. After this operation the interpretation of b_{ijt} is simply the input per unit of resource c_i and of net income p_j derived from the unit level of activity z_j in year t. The original problem has now been converted by transformation 4.52 into an equivalent one <u>where only B is subject to variation</u>, with deterministic c and p vectors. The c_{it}, p_{jt} are measured in decimal units as in the expository example below; or, even more conveniently, in the expected trend values of the c_i, p_j in the year for which the program is designed: this avoids reconversion of the final solution.

Example Agriculture for the State of Iowa, in the years mentioned, was characterized by capital requirements per acre; the b_{ijt}, and available capital resources c_{it} given in Table 2.

Table 2. Capital inputs (in \$/acre), net profits (in \$/acre), and capital available (in $\$10^6$) for Iowa agriculture[a]

Year	Capital inputs			Net profits			Available capital
	Corn	Oats	Soybeans	Corn	Oats	Soybeans	
1935	9.85	6.41	7.41	14.84	2.86	6.06	138.5
1936	7.55	5.18	6.32	11.95	4.41	8.37	111.3

1952	22.10	12.79	23.95	59.00	9.96	34.10	351.4

[a]Computed from (57b).

Table 3 shows elements of $\underline{B}_t P_t^{-1}$, while the second step, $C_t^{-1}(\underline{B}_t P_t^{-1})$, has been taken in Table 4.

Table 3. Standardization of inputs (Table 1) per unit of net profits (\$10)

Year	Capital inputs			Net profits			Available capital
	Corn	Oats	Soybeans	Corn	Oats	Soybeans	
1935	6.63	22.40	12.25	10	10	10	138.5
1936	6.32	11.70	7.85	10	10	10	111.3

1952	3.75	12.84	7.20	10	10	10	351.4

Table 4. Standardization of inputs per unit of net profits ($10) and per unit of available capital (10^8)

Year	Capital inputs			Net profits			Available capital
	Corn	Oats	Soybeans	Corn	Oats	Soybeans	
1935	4.80	16.24	8.86	10	10	10	100 million
1936	5.69	10.52	6.80	10	10	10	100 million

1952	1.07	3.66	2.05	10	10	10	100 million

Computation of the variances for the solution vectors x and the objective functions g We now have a model where only the technical matrix B is subject to random variation. Denote

$$4.54 \qquad B = \frac{1}{T} \sum_{t=1}^{T} B_t$$

as a matrix whose elements b_{ij} estimate the means of the b_{ijt}. From the system of constraint equations

$$4.55 \qquad Bz = c$$

we obtain the mean value of the solution vector

$$4.56 \qquad z = B^{-1}c.$$

Defining β_t and ξ_t as the deviations in period t:

$$4.57 \qquad \beta_t = B_t - B$$

4.58 $$\zeta_t = B_t^{-1}c - B^{-1}c ,$$

we obtain the following restrictions in the t^{th} period:

4.59 $$(B + \beta_t)(z + \zeta_t) = c .$$

Upon subtraction of 4.55 and neglecting difference terms of order higher than one*, 4.59 reduces to (16, pp.161 sqq.)

4.60 $$\zeta_t = -B^{-1}(\beta_t z),$$

from which expression we derive for the k^{th} element of vector ζ_t,

4.61 $$\zeta_{kt} = \sum_{i=1}^{m} \sum_{j=1}^{m} b^{ki} \beta_{ijt} z_j ,$$

where the b^{ki}, β_{ijt} are elements of B^{-1} and β_t respectively. If the β_{ij} are independently distributed, their cross moments vanish and the variance of the k^{th} element of vector z is estimated by

---

*This approximation is formally the same as retaining only the linear terms in a Taylor development as used in Tintner's method quoted above:

$$F(B+\beta, z+\zeta) = F(B,z) + \sum_{n=1}^{\infty} \frac{1}{n!} \left(\beta \frac{\partial}{\partial B} + \zeta \frac{\partial}{\partial z}\right)^n F(B,z),$$

where $\left(\frac{\partial}{\partial B} + \frac{\partial}{\partial z}\right)^n$ represents the operator for the n^{th} partial derivatives. Since $F(B,z) = Bz$, we get

$$(B+\beta)(z+\zeta) = Bz + \beta z + B\zeta$$

if we drop the non-linear terms.

$$4.62 \qquad s^2_{z_k} = E\,(z_k - Bz_k)^2 = \frac{1}{T-1} \sum_{t=1}^{T} \zeta^2_{kt} \,.$$

Upon substitution of ζ_{kt} by its value in 4.61 we may equate 4.62 to

$$4.63 \qquad s^2_{z_k} = \sum_{i=1}^{m} \sum_{j=1}^{m} (b^{ki})^2 \, z_j^2 \, \sigma^2_{ij},$$

a function of the variance of the b_{ij}, viz.,

$$4.64 \qquad \sigma^2_{ij} = \frac{1}{T-1} \sum_{t=1}^{T} \beta^2_{ijt} \,.$$

Generalizing 4.63 to the case with non-orthogonal b_{ij}, one has to include the cross moments of the β_{ij}; substitution of ζ_{kt} from 4.61 in 4.62 gives by induction

$$4.65 \qquad s^2_{z_k} = \left[[z'z'\ldots z'] \begin{pmatrix} B^{k1} & & & \\ & B^{k2} & & \\ & & \ddots & \\ & & & B^{km} \end{pmatrix} \right] V \left[[z'z'\ldots z'] \begin{pmatrix} B^{k1} & & & \\ & B^{k2} & & \\ & & \ddots & \\ & & & B^{km} \end{pmatrix} \right]' ,$$

i.e.

$$4.66 \qquad [z'z'\ldots z'] \begin{pmatrix} B^{k1} & & & \\ & B^{k2} & & \\ & & \ddots & \\ & & & B^{km} \end{pmatrix} V \begin{pmatrix} B^{k1} & & & \\ & B^{k2} & & \\ & & \ddots & \\ & & & B^{km} \end{pmatrix} [z'z'\ldots z']' \,,$$

where the symbols represent:
$[z'z'\ldots z']$, a row vector of m^2 elements, generated by m repetitions of the transpose of z; B^{k1} the product of scalar b^{ki} with the unit matrix I; B^{ki} is thus a diagonal matrix of rank m, and itself an element of a

diagonal matrix of rank m^2; V, the $m^2 \times m^2$ variance-covariance matrix of all b_{ij}, the typical element of which is, cfr. 4.64,

$$4.67 \qquad \sigma_{ij.kl} = \frac{1}{T-1} \sum_{t=1}^{T} \beta_{ijt} \cdot \beta_{klt} \cdot$$

It may be noticed that, at this stage, the method constitutes an approach to the general problem of solving a system of linear equations with parameters subject to a probability distribution.

The covariances of the elements of the x vector are computed analogously to 4.66:

$$4.68 \qquad s_{z_k z_l} = \frac{1}{T-1} \sum_{t=1}^{T} \zeta_{kt} \zeta_{lt}$$

$$= [z'z' \ldots z'] \begin{pmatrix} B^{k1} & & & \\ & B^{k2} & & \\ & & \ddots & \\ & & & B^{km} \end{pmatrix} V \begin{pmatrix} B^{l1} & & & \\ & B^{l2} & & \\ & & \ddots & \\ & & & B^{lm} \end{pmatrix} [z'z' \ldots z']'.$$

The profit function $g = \sum_i p_i z_i$ is equal to $\sum_i z_i$ as a result of the normalization of the activities per unit of net profits. The variance of this linear form is expressed by the quadratic function

$$4.69 \qquad s_g^2 = \sum_{i=1}^{m} z_i^2 s_{z_i}^2 + 2 \sum_{i<j} \sum z_i z_j s_{z_i z_j} = z' \Omega z .$$

If Ω represents the discriminant of the quadratic form: the formula for a typical element of Ω is 4.68.

We now recall formula 4.48 which provides an estimate of the $(1-\alpha)$ lower confidence level for the net income derived from any selection of activities, and formula 4.47 which serves as a criterion for finding the highest lower confidence level. In practice it will in all probability be sufficient to compare the lower confidence levels only for selections in the immediate vicinity of the deterministic optimum, i.e., the optimum one reaches through the usual methods (simplex). The assumption of normality for the distribution of the g's is supported by Cramer's theorem already quoted in the third chapter.

V. AGRICULTURAL MODELS WITH STOCHASTIC YIELDS AND PRICES

A considerable simplification of general models is permissible in agriculture, because yields and prices are by far the foremost sources of income variation (24, 25). The former source is meteorological in nature whereas the latter is partly due to the inelastic production pattern of the farming business. Resources, on the other hand, are generally subject to entrepreneurial decision, and inputs--apart from a variation component linked with yields--occupy an intermediary position. Taking only the above two kinds of variation into account may, therefore, be thought of as a minor loss of accuracy, in exchange for a considerable gain in simplicity.

A. Method of Hybrid Activities

One may conceive of a programming model where an observed variance is typical for a particular crop and then set out, at each stage of the program, to detect the combination of activities which associate a minimum increase in income variance with a given increase in income. This model has been taken up by Heady and Candler (25), at least as far as the purely computational techniques are concerned. They do not, however, present the theoretical background of the matter. Here we shall proceed in the inverse manner and try to establish a simple

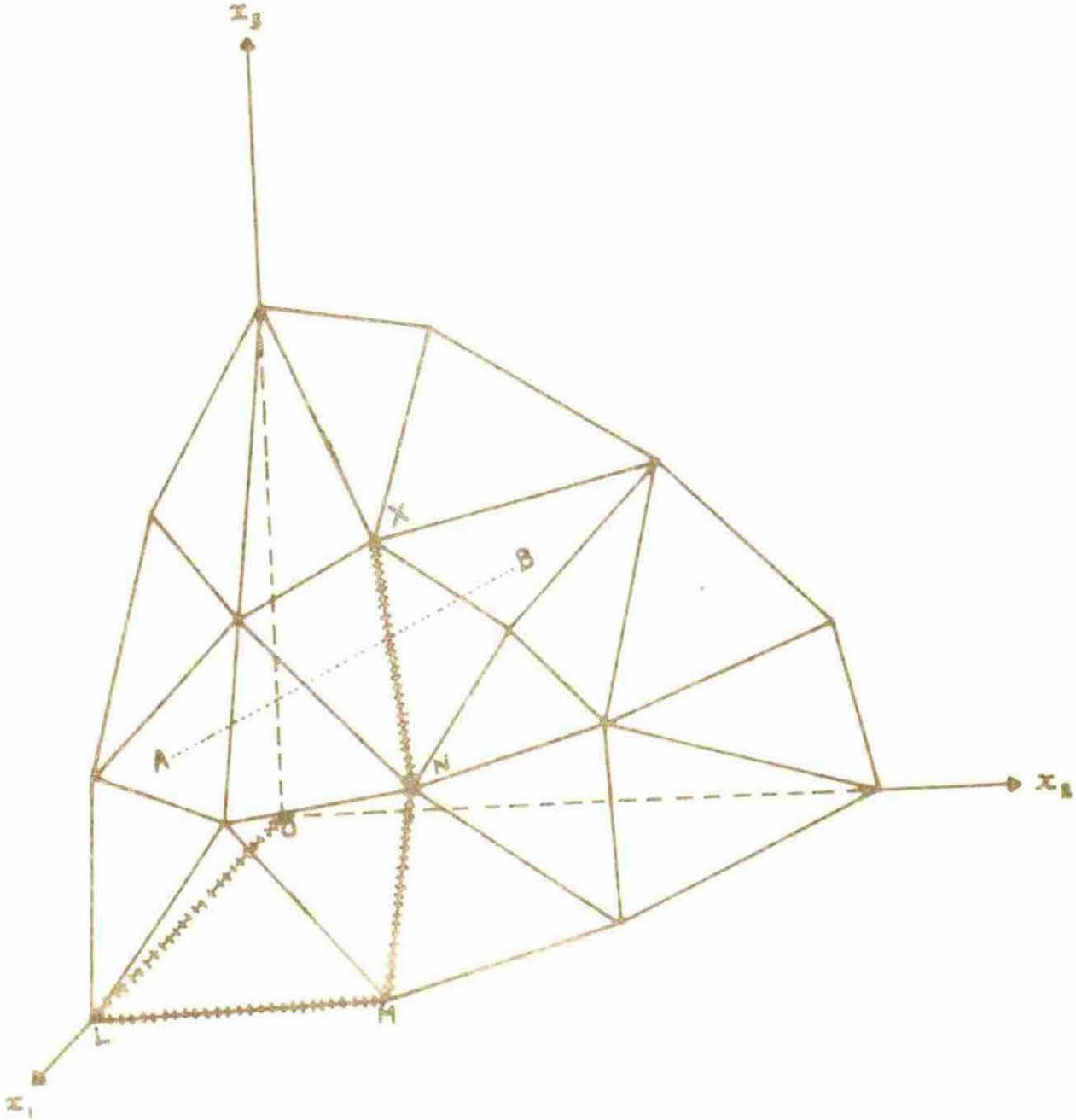
geometric and algebraic framework, dropping the details of computation which are thoroughly described in the reference.

Consider the general linear-programming problem. Without loss of generality we may, for simplicity's sake, suppose that the activities have been normalized per unit net profit, so that the objective function becomes

$$5.1 \qquad g = \{1\}' x = \sum_i x_i$$

where $\{1\}$ is a column vector of m unit elements and x a column vector of m activity levels x_i which have been normalized per unit of net profit. In the case of a unique solution, the optimum has to coincide with a vertex of the hyperpolyhedron bordered by the restrictions in a space defined over the activities. The simplex solution, therefore, only considers vertices, and shifts from one to another in the direction having the greatest profit gradient. This is done because one may, in general, expect that for a hyperpolyhedron with a large number of facets the number of jumps (simplex tableaus or cycles) required to reach the optimum X (Fig. 1) will be smallest if, at each vertex, this gradient is selected.

For the problem at hand the shift direction is chosen according to a different criterion: instead of following the border lines with the greatest profit gradient we select the internal line with minimum variance gradient per unit of profit added and follow it until we push through another border



Pig. 1. Convex polyhedron with 3 activities (x_1,x_2,x_3) and 20 relevant restrictions

hyperplane.* Typical simplex moves are, therefore, $\overline{OL}$, $\overline{LM}$, $\overline{MN}$, and $\overline{NX}$. A typical move for the method under consideration is $\overline{AB}$. If the x's are measured in units of net profit, the non-stochastic optimum X is the common point of the convex hull and a tangent plane which cuts all the axes at 45°. The optimum under risk, however, will not, in general, belong to the extremal of the hull: it will most likely be a point like B, i.e., a point on the base of a hypercone, defined by a positive convex linear combination

$$5.2 \qquad \vec{OB} = \sum_i y_i \vec{OE}_i$$

for

$$5.3 \qquad \sum_i y_i = 1, \qquad (y_i \geqslant 0)$$

where the E_i belong to the extremal and the y_i are a non-trivial set.

Starting from a feasible solution we have to find the set of y_i, i.e., the activity mix, with minimum variance increase per unit of profit added. Call V the variance-covariance matrix of the m activities or x's expressed in profit units according to the above convention. The variance of the mix is given by a form, quadratic in the y's:

---

*Related topics are discussed by Markowitz (39, pp. 129-153).

$$5.4 \qquad Q = \begin{bmatrix} y_1x_1 & y_2x_2 \cdots y_mx_m \end{bmatrix} \; V \; \begin{bmatrix} y_1x_1 \\ y_2x_2 \\ \cdots \\ y_mx_m \end{bmatrix} ,$$

i.e.,

$$5.5 \qquad Q = (yx)' \, V \, (yx) ,$$

where (yx) is a column vector. The procedure consists of trying extremal values of 5.5 under restraint 5.3. One may solve one y_i from 5.3 and substitute into 5.5 or use a Lagrange multiplier λ and then find a critical point for

$$5.6 \qquad \Phi = (yx)' \, V \, (yx) + \lambda \left(\sum_{i=1}^{m} y_i - 1 \right)$$

with respect to every y_i. One obtains a system,

$$\frac{\delta \Phi}{\delta y_i} = \frac{\delta Q}{\delta y_i} + \lambda = 0 \qquad (i=1,2,\ldots,m)$$

$$5.7 \qquad \sum_{i=1}^{m} y_i = 1 ,$$

of m+1 equations in m+1 variables, viz., the y_i and λ.

Expanding Q into

$$Q = \sum_{i=1}^{m} x_i^2 \, s_{x_i}^2 y_i^2 + 2 \sum_{i<j} \sum x_i y_i s_{x_i x_j} x_j y_j$$

where the s's are elements of V, system 5.7 can be made explicit:

5.8
$$\frac{\delta \Phi}{\delta y_i} = 2 \sum_{j=1}^{m} x_i \sigma_{ij} x_j y_j + \lambda = 0 \qquad (i=1,2,\ldots,m)$$
$$\sum_{i=1}^{m} y_i = 1 .$$

Solving for the y's, a "trial hybrid activity" is obtained, meaning that the new activity which is tried in the simplex tableau is not one of the original activities, as usual, but a compound (100 y_1 % of the first enterprise, 100 y_2 % of the second etc.) intended to increase income with a minimum increment in variance.

B. Selection of the Maximum Lower-Confidence Limit: A Practical New Model

We shall conclude this inquiry into stochastic linear programming by presenting a new and computationally practicable model which permits the farmer to cope with risk in his decision in precisely the same way as expressed by formulae 4.47 and 4.48 of the previous model IV C. However, the sources of variation will now be restricted to the items which are of a real concern to the agricultural entrepreneur: yield (weather) and prices as has already been mentioned.

Let us recall again from the introduction that the maximization of a linear function under restraints given by linear inequalities can be solved under certain conditions by computing the objective function for each of the $K = C_m^{m+n}$ solutions obtained by the selection method. In each of these

cases our stochastic problem reduces, therefore, to finding the distribution of

5.9 $\quad g = p' x \qquad (x \geqslant 0)$

subject to

5.10 $\quad Bx = c$

where the symbols denote:
g, a scalar, the profit or objective function;
p, a column vector of m net prices;
x, a column vector of m activity levels;
B, an m x m matrix of technical coefficients;
c, a column vector of m amounts of resources.
In addition, denote by $[\alpha_{ij}]$ the input matrix per unit of activity for which the variation is given; since yields are recorded per acre, one obtains direct figures for the variation per acre too, so this would be the activity unit in agriculture. In an example quoted below the average corn yield was 43.8 bu./acre and the average capital requirement $13.00/acre. If b_{ij} represents the capital input per bushel of corn one would have: $b_{ij} = 13.00/43.8 = .297$ and the corresponding α_{ij} is, of course, 13.00. Hence we see that the input coefficients b_{ij} are inversely proportionate to the yields f_j since, by definition, these b_{ij} represent inputs per unit of output, bushels, say:

5.11 $$b_{ij} = \frac{\alpha_{ij}}{f_j} .$$

If only yields are random, all the b_{ij} belonging to the j^{th} column have one single source of variation, viz., f_j, and, therefore, the same coefficient of variation.

Price variability is incorporated into the model by applying the first part of the normalization procedure outlined under section IV C (Table 3); 5.11 would then become

5.12 $$\hat{b}_{ij} = \frac{\alpha_{ij}}{f_j p_j} ,$$

with the same outcome as before: the b_{ij} of column j have again the same coefficient of variation and the (normalized) price vector becomes fixed. No such normalization has been carried out in the subsequent pages since the treatment is identical and the examples are only expository.

Define a diagonal matrix F from

5.13 $$f = F \{1\} ,$$

where f is the vector of the f_j and $\{1\}$ the column vector with only unit elements; now 5.11 becomes

5.14 $$B [\alpha_{ij}] F^{-1}$$

Substituting (5.14) into (5.10), we get

5.15 $$[\alpha_{ij}] F^{-1} x = c ,$$

with solutions

5.16 $$x = P[\alpha_{ij}]^{-1} c = P[\alpha^{ij}] c \quad .$$

Each element x_i of x can now be expressed as

5.17 $$x_i = f_i \sum_{j=1}^{m} \alpha^{ij} c_j \qquad (i=1,2,\ldots,m) \ .$$

Putting

5.18 $$\alpha_i^* = p_i \sum_{j=1}^{m} \alpha^{ij} c_j \ ,$$

the profit function may be written

5.19 $$g = \sum_{i=1}^{m} \alpha_i^* f_i \ ,$$

and under our assumption of fixed resources, has a variance

5.20 $$E(g - Eg)^2 = \alpha^{*\prime} \Phi \alpha^* \ ,$$

where Φ is the variance-covariance matrix of the f_i and α^* is a column vector with elements α_i^* .

Example The following data were computed from (24).

Table 5. Input coefficients, yields, and prices for crops in N. Central Iowa, 1917-48[a]

Item	Activity	
	Corn	Flax
Land coefficient (acres/bu.)	.022831	.095238
Capital coefficient ($/bu.)	.296804	.876190
Yield (bu./acre)	43.8	10.5
Prices ($/bu.)	1.56	3.81

[a]These data are very similar to Tintner's example (54, p. 217).

A typical farm supposedly had available 148 acres of land and $1800 of capital. The variance of output is given per acre (Table 6) and the $[\alpha_{ij}]$, therefore, is set up in inputs per acre, according to 5.14 (Table 7).

Table 6. Variance-covariance matrix of yields in N. Central Iowa (bu./acre), 1917-1948

	Corn	Flax
Corn	129.6	17.8
Flax		19.7

Table 7. $[\alpha_{ij}]$ matrix for Table 5; inputs per acre

	Corn	Flax
Land (acres)	1	1
Capital ($)	13.00	9.20

There are only $C_2^{2+2} = 6$ solutions so we can solve by the selection method, estimating the standard deviation of g, viz., s_g for each solution according to formula 5.20. The preferred solution is 4.47 with v = 1.96 (say) corresponding to a confidence level of 97.5% as read from 4.48.

The second solution yields maximum expected income, but the first is to be preferred on the .975 confidence level. The break-even confidence level $\underline{\alpha}$ between selections 1 and 2

Table 8. Solutions, profits (g), estimated standard deviation of profits (s_g), and .975 lower confidence level ($g-1.96\,s_g$)

Selection no.	Activity level (acres)		Profits		
	Corn	Flax	Mean g	s_g	$g-1.96s_g$
			($)	($)	($)
1	115.37	32.63	9188	2302	4677
2	138.46	--	9461	2459	4641
3	not feasible				
4	not feasible				
5	--	148.00	5920	2503	1020
6	no real activities				

may be found from the system of equations:

5.21 $$g_1 - v\,s_{g_1} = g_2 - v\,s_{g_2}$$

5.22 $$\alpha = \tfrac{1}{2} - \int_0^v \Phi(t)\,dt\ .$$

One finds v = 1.73, corresponding to $\underline{\alpha} = .0418$.

Setting up the technical matrix directly in terms of the α_{ij}

It saves time to set up the $[\alpha_{ij}]$ immediately from the data, since the latter refer always to varying outputs per unit of activity, say, acres; and it will, therefore, be the variance per acre which is computed. Transformation 5.14 is then unnecessary: the outputs f_j^* are now defined as

5.23 $$f_j^* = \frac{f_j}{\bar{f}_j} ,$$

where the $\bar{f}_j$ are mean yields.

Therefore,

5.24 $$E\, f_j^* = 1 ,$$

i.e.,

5.25 $$F = I,$$

where I is the unit matrix. Table 9 shows a technical matrix interms of the α_{ij}.

Table 9. Inputs per acre and global resources available for agriculture, state of Iowa, 1928-52[a]

Resource	Amount available	Inputs			
		Corn	Oats	Soybeans	Wheat
Land[b]	17 255 360	1	1	1	1
Capital[c]	20 468 882	1.30	.96	1.08	1.02
March labor[d]	4 426 227	0	.777	0	0
April "	19 335 841	1.239	.973	.8036	0
May "	26 166 667	2.31	0	1.9926	0
June "	15 583 339	1.3755	0	1.189	0
July "	28 186 609	1.1235	2.485	.92	4.4625
August "	16 042 415	0	2.765	0	.9525
Sept. "	3 068 371	.21	0	.2378	2.085
Oct. "	19 184 500	1.554	0	3.075	0
Nov. "	22 578 993	2.142	0	0	0

[a]Computed from Agricultural Statistics, U. S. Dept. Agr., and from (6, 24).

[b]In acres.

[c]In $10.

[d]In man-hours.

From 5.23 the elements of the new variance-covariance matrix can be computed in terms of the original ones:

$$5.36 \qquad \sigma_{f_i^* f_j^*} = \frac{1}{\bar{f}_i \bar{f}_j} \sigma_{f_i f_j} ,$$

as shown in Table 10.

Table 10. Variance-covariance table for crop yields, state of Iowa, 1928-52[a]

	In bu./acre				In mean yields/acre[b]			
	Wheat	Corn	Oats	Soyb.	Wheat	Corn	Oats	Soyb.
Wheat	13.522	13.306	12.859	3.327	.03887	.01594	.01988	.00947
Corn		130.706	41.951	29.271		.06524	.02767	.03473
Oats			48.039	10.634			.03994	.01628
Soybeans				9.099				.02566

[a]Computed from Agricultural Statistics, U. S. Dept. Agr.

[b]Average yields were: 18.65, 44.76, 34.68, and 18.83 bu./acre in the order mentioned.

A modified simplex criterion

Testing all possible solutions, obtained by the selection method, leads to an almost prohibitive computational burden in the above State program of Iowa agriculture. Therefore, we base ourselves upon the ordinary simplex algorithm, and introduce a new criterion vector.

Each simplex iteration yields a plan with m activities, some of them possibly at the zero level. The ordinary criteri-

on replaces activity x_k of the k^{th} selection by the new activity $\hat{x}_k$, provided (15)

$$5.27 \qquad (p_{\hat{x}_k} - \pi'_{(k)}p) = \max_j (p_{\hat{x}_j} - \pi'_{(j)}\, p), \quad (j=1,2,\ldots,m)$$

where the $p_{\hat{x}_j}$ are prices or weights of new activities $\hat{x}_j$ being tried, and $\pi_{(j)}$ is a column vector of marginal rates of substitution between the new $\hat{x}_j$ and each of the x_j of the old activity vector. Since 5.27 measures the marginal profit increase due to the introduction of one unit of $\hat{x}_j$, the objective function g_k, corresponding to the k^{th} selection, grows to

$$5.28 \qquad g_k = p'x + \hat{x}_k\,(p_{\hat{x}_k} - \pi'_{(k)}p)$$

after $\hat{x}_k$ has been brought in at its maximum level.

In formulae 5.17 - 5.19 we defined vector α^* as a convenient transformation of x. At each stage of the simplex tableau a new activity $\hat{x}_k$ replaces the k^{th} element of x. So, when α^* corresponds to

$$5.29 \qquad [x_1, x_2, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_m]'$$

we define $\alpha^*_{(k)}$ as the same transformation, but now of the vector

$$5.30 \qquad [x_1, x_2, \ldots, x_{k-1}, \hat{x}_k, x_{k+1}, \ldots, x_m]' .$$

Assuming a preference expression

$$5.31 \qquad g^o - v\, s_{g^o} = \max_k (g_k - v\, s_{g_k})$$

as before, we have to develop a new criterion based on 5.27, at the same time taking account of the lower confidence level expressed by the term v s_g.

Analogously to 5.20:

5.32 $$s^2_{g_1} = \alpha^{*'}_{(1)} \Phi \alpha^*_{(1)} ,$$

and the new criterion expression becomes

5.33 $$\max_1 \left\{ p'x + \hat{x}_1 (p_{\hat{x}_1} - \pi' p) - v\sqrt{\alpha^{*'}_{(1)} \Phi \alpha^*_{(1)}} \right\} \quad (1=1,2,\ldots,m),$$

where v again follows from 4.48. Subscripts 1 replace k because, with the new criterion, the incoming activity may be different from the one which would be selected according to the usual criterion.

Stochastic state program for Iowa agriculture

The state program for Iowa agriculture, constructed with the help of the data of tables 9 and 10 is set up in such a fashion that, for each crop, the mean acreage actually used over the period considered was included in the optimum program; it proves to be the stochastic optimum as well, even on the .975 probability level. For stochastic comparison with some other high-profit programs, the reader may refer to Table 11 below. The figures have been obtained by computing a simplex table, but it was criterion 5.33 which determined at each stage

which new activity had to come in. Theoretically, expression 5.33 has to be computed at each stage for each possible new activity which might be tried but in practice most possibilities can be discarded at sight.

Table 11. Stochastic comparison of some high-profit state programs for Iowa agriculture

Program no.	Activities included (in 1000 acres) Corn	Oats	Soyb.	Wheat	Expected net profit (in $1000)	Lower .975 profit level (in $1000)
1	10541	5697	912	306	303,540	164,705
2	10541	5802	912	-	300,691	162,184
3	10541	5661	-	410	292,559	156,804
4	10541	5802	-	-	288,744	153,439
5	10541	-	912	-	267,039	136,135
6	10541	-	-	410	259,725	131,441
7	10541	-	-	-	255,092	127,384

VI. NON-RANDOM VARIATION

These models are not concerned with stochastic programming proper: the parameters under consideration are not supposed to vary at random but are deliberately fixed at several levels. The technically interesting aspects of such methods are the short-cut algorithms (consisting of modifications of the simplex routine) which make it possible to reach a new optimum solution without having to execute a complete simplex iteration. We shall only mention these techniques, because the subject is essentially beyond the scope of this dissertation.

Heady and Candler (25) present simplex modifications for non-stochastic variation in prices, resources, and input coefficients; the algorithms proposed for each case are extensively treated in the above reference.

However, in the case of one resource or constraint c_i subject to continuous variation Hartley and Loftsgard (23) have presented a modification of the simplex algorithm, related to Candler's solution (10) but, in practice, superior when programming on high speed computers. In presenting the essentials of the procedure, we shall assume the reader to be familiar with the simplex routine. Consider the variation of one resource, say, capital. Starting with some initial amount of capital suppose that we reach the corresponding optimum

program in the T^{th} simplex iteration and denote by: a_{ij}^{t}, the element of the i^{th} row and j^{th} column in the t^{th} simplex iteration tableau, the iteration count starting at 0; b_i^t, the element of the i^{th} row in the resource column, with the same meaning attached to the superscript t; $(b_i^t \geqslant 0)$ B, the matrix which transforms the columns a_{ij}^{o} to a_{ij}^{T} for every j.

If the capital slack column is in the k^{th} position, we write down the equalities:

6.1 $$a_{ik}^{T} = B\, a_{ik}^{o}$$

6.2 $$b_i^T = B\, b_i^o$$

Assume now that capital is increased in the original tableau by an amount ΔC, so that it grows to

6.3 $$b_1^o + \Delta C$$

if capital occupies the first row in the resource column. Since the capital slack column of the first section has 1 as its first element while the other elements are 0, any member of the new resource column may as well be written

6.4 $$b_i^o + \Delta C \cdot a_{ik}^o \,.$$

Carrying this column through the same T simplex iterations again, it becomes

6.5 $b_i^{T'} = E(b_i^{\circ} + \Delta C.a_{ik}^{\circ})$

i.e.

6.6 $b_i^{T'} = b_i^{T} + \Delta C.a_{ik}^{T}$,

meaning that the resource column of section T' may be obtained directly by adding ΔC times the slack column of section T to the resource column of this section.

Since the $b_i^{T'}$ column represents a "feasible solution" as long as

6.7 $b_i^{T} + \Delta C.a_{ik}^{T} \geqslant 0$ (for all i) , the authors go on to distinguish between two cases.

1. All elements of a_{ik}^{T} are $\geqslant 0$.

Expression 6.7 is always $\geqslant 0$ so we have feasible solutions for any increment ΔC, and since all $Z_j - C_j$ * are $\geqslant 0$, they are all optimum. In other words the set of activities already included in tableau T remains unaltered and their levels, as well as the profit, merely show a linear increase with ΔC: the problem is solved.

2. At least one element of a_{ik}^{T} is < 0.

The non-negativity condition for programming solutions requires for 6.7

6.8 $\Delta C \leq \min \frac{b_i^{T}}{|a_{ik}^{T}|}$

---

*Common indication for the criterion row of a simplex iteration.

where the minimum is valid for negative divisors a_{ik}^{T} only. We compute $b_1^{T'}$ as indicated in 6.6 and continue the standard simplwx routine until all $Z_j - C_j$ are ≥ 0, as usual. If there are still negative a_{ik} left, one repeats the procedure outlined in this paragraph.

The authors, in their afore-mentioned paper, have worked out along these lines an example by Candler (10). The method is important for our purpose in that it provides a simple criterion of the maximum variation one single resource can undergo without affecting the set of activities included in the optimum solution.

Apart from price and resource variation the input coefficients of the "technical matrix" may also be subject to variation. For programming in agricultural enterprises, Heady and Candler (25) have given a computational method for yield variation, thus changing by the same rate the coefficients of the column vector relevant for that activity. What they call "variation of coefficients" is, in fact, variation of an activity column. Their computational method has been treated in the reference.

VII. NON-LINEAR STOCHASTIC PROGRAMMING

This subject matter is manifestly beyond the scope of this paper. We shall, however, at least mention two articles published in Econometrica.

The first, by Freund (20), deals with the introduction of risk into a programming model. It is illustrated by an application to agriculture. Only net revenues of unit levels of output are considered subject to variation. The author assumes an exponential utility function of a rather arbitrary type and the manner in which he derives a quadratic objective function from the exponent seems doubtful. These considerations hold for the assumption of a normal distribution for the net revenues; in the case of a discrete distribution, however, he also arrives at a quadratic function. The maximization is carried out with the help of the results obtained by Kuhn and Tucker (32), who proved that the quadratic programming problem:

maximize

7.1 $\quad s'x - x' \Sigma x,$

subject to

7.2 $\quad Tx \leq v \ ;$

is equivalent to the minimax problem:

maximize with respect to x, and minimize with respect to u

7.3 $$\Phi(x, u) = x' \Sigma x - s'x + u'(v-Tx) \qquad (u \geq 0) .$$

The symbols stand for:

s, the vector of net profits;

x, the activity vector;

Σ, the variance-covariance matrix of the x's at unit level;

T, the "technical" matrix;

v, the vector of restrictions;

u, a vector containing as many elements as T rows.

The second article, by Simon (43), is of a theoretical nature. For a quadratic criterion function he reduces the case of dynamic programming under uncertainty to the planning problem under certainty. The set of (unconditional) expected values of the random variables is taken as a set of "certainty equivalents" and they replace, in the computation of the optimal action for the first period, the "certain" future values. A similar result (certainty-equivalence method) had been established by Theil (45) for static planning. These considerations are, however, of a purely-mathematical interest: they are intuitively acceptable and do not give any information as to the distribution of the expected outcomes, which is, of course, our real point of interest.

VIII. CONCLUSIONS WITH SPECIAL REFERENCE TO STATE PROGRAMMING

Obviously, the methods developed at this stage are also relevant to the purely mathematical problem of deriving the distribution of solutions for a system of linear equations whose coefficients have probability distributions.

Problems of small errors in the solutions of linear equations with coefficients subject to errors have been considered by several writers not yet referred to, e.g. Lonseth (36), Hotelling (27), and Turing (57a). These considerations are of the rounding-error type. Confidence regions for systems with coefficients having a known variance-covariance matrix have been considered by Box and Hunter (9), who extended an argument by Fieller (18): the method does not work, however, if the system is "poorly conditioned".*

As for economics, the extension of the methods discussed to such related problems as game theory (60), Leontief systems (33), and allocation of resources (31) is straightforward; the

---

*From the system of non-homogeneous equations $a_0+Ax = \delta$, derive the quadric $\delta'\delta = (Ax+a_0)' (Ax+a_0) = \sum_i \delta_i^2$, where a_0 and δ are vectors with constant elements, A is a matrix, and x a vector of unknowns. For fixed values of $z=\sum_i \delta_i^2$ the quadric defines contours (ellipsoids) which may be referred to their centre as origin and their principal axes as coordinates:

$$z = \sum_i \lambda_i X_i^2 ,$$

where the λ_i, X_i are latent roots and vectors of $A'A$. If some λ_i are relatively small, the conditioning surface is very attenuated in this direction and large deviations of the x's cause but a negligible effect on $\sum_i \delta_i^2$.

above topics have been linked together (2), so the stochastic approach applies to all of them. We would like to attract attention to a less cited topic, viz., state programming. One may think of the subject as the development of a plan for maximizing aggregate income in one sector, e.g. agriculture, of a given state: a stochastic appraisal for Iowa agriculture viewed as a state program has been worked out in the last section.

In a broader sense state programming consists of maximizing some welfare function (59), under constraints approximated by an econometric model of the nation's economy (21), e.g. the 27-equations model of the Central Planning Bureau at The Hague (46, 48, 49, 50) or the Klein-Goldberger model for the United States (19, 30). Consider such a model

$$8.1 \qquad Ay + Bx = u, \qquad (Eu = 0)$$

where y and x are vectors of endogenous and exogenous variables respectively. In the Tinbergen terminology the endogenous variables stand for targets and irrelevant variables, the exogenous variables for instruments and data. One may, therefore, partition 8.1 as follows:

$$8.2 \qquad \begin{bmatrix} A_T & A_{IR} \end{bmatrix} \begin{bmatrix} y_T \\ y_{IR} \end{bmatrix} + \begin{bmatrix} B_{IN} & B_D \end{bmatrix} \begin{bmatrix} x_{IN} \\ x_D \end{bmatrix} = u ,$$

where the significance of the subscripts is obvious. In the

Tinbergen approach the y_T are fixed; the x_{IN} are then solved in terms of the y_T, or target levels, and the expected values of the other vectors, so the scheme of government measures is, in principle, determined.

In the more recent approach a welfare function

8.3 $$w(x_{IN}, y_T) = p' x_{IN} + p^{o'} y_T$$

is maximized under constraints 8.1. The weight vectors p and p^o comprise the so-called barter terms, determined as follows:

Rewrite 8.3 as

8.4 $$w(z) = \pi' z$$

i.e.,

8.5 $$w(z) = \sum_i \pi_i z_i,$$

where z and π are the compound vectors $\{x_{IN}, y_T\}$ and $\{p, p^o\}$ respectively. Let z_i represent the target "balance of payments surplus", measured in 100 million guilders, and z_j the target "increase in the rate of employment", measured in %. If the pressure groups, politicians, and advisors vote to set a 100 million surplus on the balance of payments against (or estimate it as much of an increase in welfare as) an increase in the rate of employment of .20%, then $\pi_i = 1.0$ and $\pi_j = .20$ etc.

State planning may thus be tackled by formal linear-programming models. Our most important conclusion is that the

general stochastic methods developed above now enable us to introduce a quantified risk concept into political measures. Since the statistical techniques used in building model 8.1, such as maximum likelihood, limited information, and double-stage least squares (4, 5, 8, 47, 53), provide us with an estimate of the variances and covariances of the coefficients of A and B, the foregoing methods, developed for stochastic matrices, apply directly. Appreciation of risk and uncertainty has been brought within the reach of welfare policy.

As for planning on the firm's scale, several of the above methods are promising, especially the active approach presented; the very expensive calculations are, however, a major drawback.

The new, complete model coping with variation from all sources, profits, resources, and input coefficients, is as yet of primarily theoretical interest. It may be developed into a computationally feasible algorithm.

In agriculture, however, we have succeeded in constructing a workable model, the computations of which require hardly more clerical work than the ordinary simplex routine. The risks involved in agriculture's foremost hazards, yield variation due to unpredictable meteorological effects, and price variation due to the unwieldiness of short-run farming decisions have thus been incorporated in an active approach to the optimum program, at any desired confidence level.

IX. REFERENCES

1. Adam, A. Messen und regeln in der Betriebswirtschaft. Würzburg, Germany, Physica-Verlag. 1959.

2. Allen, R. G. D. Mathematical economics. 2nd ed. London, England, Macmillan and Co., Ltd. 1959.

3. Anderson, R. L. and Bancroft, T. A. Statistical theory in research. New York, N. Y., McGraw-Hill Book Company, Inc. 1952.

4. Anderson, T. W. and Rubin, H. Estimation of the parameters of a single equation in a complete system of stochastic equations. The Annals of Mathematical Statistics 20:46-63. 1949.

5. _____ and _____. The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations. The Annals of Mathematical Statistics 21:570-582. 1950.

6. Babbar, Madan Mohan. Statistical approach in planning production programs for interdependent activities. Unpublished Ph.D. Thesis. Ames, Ia., Library, Iowa State University of Science and Technology. 1953.

7. _____. Distributions of solutions of a set of linear equations. Journal of the American Statistical Association 5:854-869. 1955.

8. Basmann, R. L. The computation of generalized classical estimates of coefficients in a structural equation. Econometrica 27:72-81.

9. Box, G. E. P. and Hunter, J. S. A confidence region for the solution of a set of simultaneous equations with an application to experimental designs. Biometrika 41:190-199. 1953.

10. Candler, Wilfred V. A modified simplex solution for linear programming with variable capital restrictions. Journal of Farm Economics 38:940-955. 1956.

11. Charnes, W., Cooper, W. W., and Henderson, A. An introduction to linear programming. New York, N. Y., John Wiley and Sons, Inc. 1953.

12. Cramer, Harald. Mathematical methods of statistics. Princeton, Princeton University Press. 1946.

13. Dantzig, George B. Maximization of a linear function of variables subject to linear inequalities. In Koopmans, Tjalling C., ed. Activity analysis of production and allocation. pp. 339-347. New York, N. Y., John Wiley and Sons, Inc. 1951.

14. _____. Linear programming under uncertainty. Management Science 1:197-206. 1955.

15. Dorfman, R. Application of linear programming to the theory of the firm. Berkeley, Calif., University of California Press. 1951.

16. Dwyer, P. S. Linear computations. New York, N. Y., John Wiley and Sons, Inc. 1951.

17. Ferguson, A. R. and Dantzig, G. B. The allocation of aircraft to routes--an example of linear programming under uncertain demand. Management Science 3: 45-73. 1956.

18. Fieller, E. C. The distribution of the index in a normal bivariat population. Biometrika 24:428-440. 1932.

19. Fox, Karl A. Econometric analysis for public policy. Ames, Ia., Iowa State College Press. 1958.

20. Freund, Rudolf T. The introduction of risk into a programming model. Econometrica 24:253-263. 1956.

21. Frisch, R. Statistical confluence analysis by means of complete regression systems. Oslo, Norway, Klara Civiltryckeri. 1934.

22. Geary, R. C. The frequency distribution of the quotient of two normal variables. Journal of the Royal Society 93:442-446. 1930.

23. Hartley, Herman O. and Loftsgard, Lowel D. Linear programming with variable restraints. Iowa State College Journal of Science 33:161-172. 1958.

24. Heady, Earl O., Kehrberg, Earl W., and Jebe, Emil H. Economic instability and choices involving income and risk in primary or crop production. Iowa Agr. Exp. Sta. Research Bul. 404. 1954.

25. Heady, Earl O. and Candler, Wilfred. Linear programming methods. Ames, Ia., The Iowa State College Press. 1958.

26. Hoel, Paul G. Introduction to mathematical statistics. New York, N. Y., John Wiley and Sons, Inc. 1947.

27. Hotelling, Harold. Some new methods in matrix calculations. Annals of Mathematical Statistics 14:1-33. 1943.

28. Khinchin, A. I. The entropy concept in probability theory. Uspekhi Matematicheskich Nauk 8, No. 3: 3-20. 1953.

29. Klein, Lawrence R. A textbook of econometrics. Evanston, Ill., Row, Peterson and Company. 1953.

30. _____ and Goldberger, A. S. An econometric model of the United States, 1929-1952. Amsterdam, Holland, North-Holland Publishing Co. 1955.

31. Koopmans, Tjalling C., ed. Activity analysis of production and allocation. New York, N. Y., John Wiley and Sons, Inc. 1951.

32. Kuhn, H. W. and Tucker, A. W. Nonlinear programming. Second Berkeley Symposium on Mathematical Statistics and Probability. Proc. 5:481-492. 1950.

33. Leontief, W. W. Structure of the American economy, 1919-1929. 2nd ed. New York, N. Y., Oxford University Press. 1951.

34. Linder, A. Statistische Methoden. Mathematische Reihe. Band 3. Basel, Switzerland, Verlag Birkhäuser. 1951.

35. _____. Planen und auswerten von Versuchen. Reihe der experimentellen Biologie. Band 13. Basel, Switzerland, Verlag Birkhäuser. 1953.

36. Lonseth, A. T. Systems of linear equations with coefficients subject to error. Annals of Mathematical Statistics 13:332-337. 1942.

37. Madansky, Albert. Bounds on the expectation of a convex function of a multivariate random variable. Annals of Mathematical Statistics 30:733-746. 1959.

38. _____. Inequalities for stochastic linear programming problems. Management Science 6:197-204. 1960.

39. Markowitz, Harry M. Portfolio selection. Cowles Foundation Monograph No. 16. New York, N. Y., John Wiley and Sons, Inc. 1959.

40. Mood, Alexander McFarlane. Introduction to the theory of statistics. New York, N. Y., McGraw-Hill Book Company, Inc. 1950.

41. Pearson, Karl. Tables for statisticians and biometricians, Part II. London, England, Cambridge University Press. 1924.

42. Savage, L. J. The foundations of statistics. New York, N. Y., John Wiley and Sons, Inc. 1959.

43. Simon, Herbert A. Dynamic programming under uncertainty with a quadratic criterion function. Econometrica 24:74-81. 1956.

44. Talacko, J. On stochastic linear inequalities. Trabajos de Estadistica 10, No. 2:89-112. 1959.

45. Theil, H. Econometric models and welfare maximisation. Weltwirtschaftliches Archiv 72:60-83. 1954.

46. _____. Economic forecasts and policy. Amsterdam, Holland, North-Holland Publishing Co. 1958.

47. _____ and Kloek, T. The statistics of systems of simultaneous economic relationships. Statistica Neerlandica 13, No. 1. 1959.

48. Tinbergen, J. An econometric approach to business cycle problems. Paris, France, Hermann et Cie. 1937.

49. _____. On the theory of economic policy. 2nd ed. Amsterdam, Holland, North-Holland Publishing Co. 1955.

50. _____. Economic policy, principles and design. Amsterdam, Holland, North-Holland Publishing Co. 1956.

51. Tintner, G. The theory of choice under subjective risk and uncertainty. Econometrica 9:298-303. 1941.

52. _____. The pure theory of production under technological risk and uncertainty. Econometrica 9:305-312. 1941.

53. _____. Econometrics. New York, N. Y., John Wiley and Sons, Inc. 1952.

54. Tintner, G. Stochastic linear programming with applications to agricultural economics. The Second Symposium on Linear Programming. Proc. 1:197-228. 1955.

55. _____. Mathematics and statistics for economists. New York, N. Y., Rinehart and Co., Inc. 1957.

56. _____. La teoria dei giochi, la programmazione lineare e l'analisi delle interdipendenze strutturali. Milan, Italy, L'industria. 1957.

57a. Turing, A. M. Rounding off errors in matrix processes. Quarterly Journal of Mechanics and Applied Mathematics 1:287-308. 1947.

57b. U. S. Dept. Agr. Agricultural statistics, 1935-1952.

58. Vajda, S. Inequalities in stochastic linear programming. Bulletin of the International Statistical Institute 36:357-363. 1958.

59. Van Eijk, C. J. and Sandee, J. Quantitative determination of an optimum economic policy. Econometrica 27: 1-13. 1959.

60. Von Neumann, J. and Morgenstern, O. Theory of games and economic behavior. 2nd ed. Princeton, Princeton University Press. 1947.

X. ACKNOWLEDGMENTS

The author wishes to express his sincere gratitude to Dr. Gerhard Tintner who suggested the topic of this dissertation and carefully read through the manuscript. He also offered continuous encouragement and assistance during the course of the underlying investigations. Special thanks are due to Dr. Herman O. Hartley and Dr. Carl E. Langenhop for making several criticisms concerning the accuracy of the argument, and accepting membership of the committee.

Dr. Karl A. Fox, Head of the Economics Department, and Dr. Erik Thorbecke appreciably stimulated my interest in state programming within the framework of the Economic-Policy Seminar (Ford Foundation).

This research has been made possible by the grant of a C.R.B. fellowship, Belgian American Educational Foundation: I feel deeply indebted to Mr. Perrin C. Galpin, President, Clark E. Stillman, Secretary in the United States, and Jacques van der Belen, Secretary in Belgium.

The author is much obliged to his Excellency Mgr. Honore Van Wayenbergh, Rector Magnificus of Louvain University for his personal interest, and to Messrs. F. Jurion and Prof. J. Lebrun, respectively Director-General and Secretary-General of the Institut National pour l'Etude Agronomique du Congo, Colonial Office, for their cooperative attitude.

I also feel indebted to my friends: Morton A. Stelling (Iowa State University), who kindly read the manuscript and eliminated most non-idiomatic expressions; and to Messrs. Roger Mauldon (Perth, Australia) and Loren Ihnen for their help and advice.