CHANCE-CONSTRAINED LINEAR PROGRAMMING WITH DISTRIBUTION-FREE CONSTRAINTS

Technical Report No. 59

SUBMITTED BY:

Frederick M. Allen
Research Associate

R. M. Braswell
Principal Investigator

APPROVED BY:

B. D. Sivazlian
Editor, Project THEMIS

Department of Industrial and Systems Engineering
The University of Florida
Gainesville, Florida 32601

December 1971

This research was performed under Project THEMIS, ARO-D Contract No. DAH C04 68C0002.

Security Classification

DOCUMENT CONTROL DATA - R&D

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

1. ORIGINATING ACTIVITY (Corporate author) University of Florida	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP Unclassified

3. REPORT TITLE

CHANCE-CONSTRAINED LINEAR PROGRAMMING WITH DISTRIBUTION-FREE CONSTRAINTS

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

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

ALLEN, Frederick M.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
December 1971	61	29

8a. CONTRACT OR GRANT NO. DAH C04 68C0002	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 59
b. PROJECT NO. 2T0 14501B81B	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. AVAILABILITY/LIMITATION NOTICES

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	ARO-Durham

13. ABSTRACT

~~This~~ report is concerned with methods of approximating the chance-constrained set $S = \{\underline{x} | \Pr[\underline{A}\ \underline{x} \le B] \le \alpha\}$ when the underlying distribution, $F(\cdot)$ of the random variate $(\overline{A}, B)$ is non-normal. The resulting sets are completely distribution-free in that no assumptions are made about the form of $F(\cdot)$ or any of its parameters.

The concept employed is the distribution-free tolerance region. This is a sample based region containing 100α percent of the population, at a confidence level, β. The elements of the distribution-free sets satisfy the chance-constraint, $\Pr[\underline{A}\ \underline{x} \le B] \le \alpha$ with a confidence of at least β. Furthermore, the sample size required to attain this level of confidence is readily available in tabular or graphical form. The superiority of the distribution-free approach over existing chance-constrained methods is demonstrated using simulated gamma variates.

DD FORM 1 JAN 64 1473

Security Classification

ABSTRACT

This report is concerned with methods of approximating the chance-constrained set $S = \{\underline{x} | Pr[\underline{A}\ \underline{x} \leq B] \geq \alpha\}$ when the underlying distribution, $F(\cdot)$ of the random variate $(\underline{A}, B)$ is non-normal. The resulting sets are completely distribution-free in that no assumptions are made about the form of $F(\cdot)$ or any of its parameters.

The concept employed is the distribution-free tolerance region. This is a sample based region containing 100α percent of the population, at a confidence level, β. The elements of the distribution-free sets satisfy the chance-constraint, $Pr[\underline{A}\ \underline{x} \leq B] \geq \alpha$ with a confidence of at least β. Furthermore, the sample size required to attain this level of confidence is readily available in tabular or graphical form. The superiority of the distribution-free approach over existing chance-constrained methods is demonstrated using simulated gamma variates.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Consider the linear programming problem of the form

maximize $Z = \underline{C}\,\underline{x}$

subject to

$$\underline{A}_i\,\underline{x} \leq B_i \qquad i = 1,\ldots, q \qquad (1.1)$$

$$\underline{x} \geq \underline{0}$$

$\underline{x}$ is an n-dimensional column vector, and $\underline{A}_i$ and C are n-dimensional row vectors. In real-world problems the elements of $\underline{C}$, $\underline{B}$ and $\underline{A}_i$ may be random variables and in such a case the above formulation (1.1) has no meaning. The random variable Z cannot be maximized and must be replaced with some deterministic function. The most widely used function is the expected value of Z, although other choices have been suggested in the literature [1,2,3]. This research is concerned only with random variation in the constraints. In particular, the chance-constrained formulation originally proposed by Charnes et al. [4] is considered. For a review of other possible reformulations of linear programming problems subject to random variation, the reader is referred to the survey paper by McQuillan [5].

Chance-Constraints

In chance-constrained programming it is not required that the constraints always be satisfied, but rather that they be satisfied with given probabilities. More precisely, the chance-constrained reformulation of (1.1) associates with each constraint a preassigned number α_i, $0 \leq \alpha_i \leq 1$, $i = 1,\ldots, q$ such that $\Pr[\underline{A}_i \underline{x} \leq B_i] \geq \alpha_i$, $i = 1,\ldots, q$. The corresponding feasible solution set is then given by

$$S = \{\underline{x} | \Pr[L_i(\underline{x}) \leq 0] \geq \alpha_i, \quad i = 1,\ldots, q; \ \underline{x} \geq 0\} \qquad (1.2)$$

where

$$L_i(\underline{x}) = \underline{A}_i \underline{x} - B_i \qquad i = 1,\ldots, q$$

It is desired to convert S into a form more amenable to existing mathematical programming techniques. The method of conversion suggested by Charnes [2] yields the equivalent form

$$S_Q = \{\underline{x} | E[L_i(\underline{x})] + K_{\alpha_i} \sigma[L_i(\underline{x})] \leq 0, \ i = 1,\ldots, q, \ \underline{x} \geq \underline{0}\} \qquad (1.3)$$

where $E[L_i(\underline{x})]$ and $\sigma[L_i(\underline{x})]$ denote the expected value and standard deviation of $L_i(\underline{x})$, respectively. K_{α_i} is the smallest number satisfying

$$\Pr[T_i(\underline{x}) \leq K_{\alpha_i}] \geq \alpha_i$$

where $T_i(\underline{x})$ is the standardized variate of $L_i(\underline{x})$. (K_{α_i} is often referred to as the quantile of order α_i.) When $K_{\alpha_i} \geq 0$, it can be shown [6] that S_Q is convex. In such a case any one of a number of convex programming algorithms could be used to solve the resulting problem.

The above approach, henceforth called the Quantile Method, is limited to a special class of distributions which are referred to as "stable" [7]. The common property of this class is that the distributions are completely specified by two parameters U and V, and the convolution of any K distributions $F[(x - U_1)/V_1]$, ..., $F[(x - U_K)/V_K]$ is again of the form $F[(x - U)/V]$. One such distribution belonging to this class is the normal, thus giving the Quantile Method some appeal. However, many times the elements of $\underline{A}_i$ and $\underline{B}$ are not normal. For example, the elements of $\underline{A}_i$ may represent rates which have to be non-negative. In such cases alternative approaches [8,9] have been proposed for obtaining convex solution sets which approximate the set S_Q. The most general procedure is given by Sinha [9], in that only the means, variances and covariances of the random variables need be specified. Using the Tchebysheff Extended Lemma [10], it is shown that S_Q contains the convex set

$$S_T = \{\underline{x} | E[L_i(\underline{x}) + \left(\frac{\alpha_i}{1 - \alpha_i}\right)^{1/2} \cdot \sigma[L_i(\underline{x})] \leq 0,$$

$$i = 1, \ldots, q; \ \underline{x} \geq 0\} \qquad (1.4)$$

This method of conversion shall henceforth be referred to as the Tchebysheff Method.

Motivation and Objective of this Research

Although the Tchebysheff Method makes possible the solution of chance-constrained programs under non-normal conditions, there still exists a reliance upon parameters of the underlying distribution. In real world situations, the values associated with these parameters are estimates derived from random samples. The accuracy of these estimates can be measured in terms of levels of significance or degrees of confidence, but there is no way of directly incorporating these measures into the set S_T. Thus, the effect of bad estimates upon the solution obtained using the Tchebysheff Method cannot be ascertained. A similar situation also occurs with the Quantile Method when normality assumptions are sample based.

The above discussion suggests a need for a more generalized theory and method of solving chance-constrained linear programming problems when random sampling is necessary. The most general would be a method which could be used regardless of the forms of the underlying distributions or any of their parameters. While such a requirement precludes the use of any classical statistical techniques, there exists a special class of so-called distribution-free techniques which are applicable in situations similar to the above.

The objective of this research is to develop methods for constructing a distribution-free set from a sample of size N such that for any $\underline{x}$ contained in this set, it can be asserted with a confidence level β, that a constraint will hold with a certain probability, α. The concept to be employed is that of a distribution-free tolerance region, similar to the one used by Allen and Braswell [11].

CHAPTER II

THEORY AND METHODS OF DISTRIBUTION-FREE TOLERANCE REGIONS

This chapter deals with the development of the theory and methods which serve as the statistical foundation for Chapter III. The chapter begins with a definition of a distribution-free tolerance region, then follows with a general procedure for constructing such a region.

Definition of a Distribution-free Tolerance Region

Let $\underline{Y} = (Y_1, \ldots, Y_n)$ be an n-dimensional random variable with a cumulative distribution function (c.d.f.) $H_{\underline{Y}}(\cdot)$. Let $O_N = (\underline{Y}_k,\ k = 1, \ldots, N)$ be a sample of size N drawn from a population with c.d.f. $H_{\underline{Y}}(\cdot)$. Let T be a region that lies in the sample space of $\underline{Y}$, and assume that the exact shape and size of T depends upon the observed values of O_N. Define the coverage, U, of the region, T, as the probability measure of T. Since T is random, U will also be random. Now if the corresponding c.d.f. of U is independent of $H_{\underline{Y}}(\cdot)$, and if for $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$

$$\Pr[U \geq \alpha] = \beta$$

then T is called a 100α percent distribution-free tolerance region at a probability level, β [12]. This concept was originally introduced by Shewhart [13] in 1931.

The above definition is interpreted by Fraser [14] as: "In repeated sampling the probability is β that the region T contains at least 100α percent of the population." Now for a particular experimental value of O_N, the corresponding region, T, may or may not contain at least α of the population. However, one can assert with a confidence of β that it does.

It should be noted that the term non-parametric has also been used to describe the above region [14,15]. As Noether [16] indicates, "this term has come to refer to methods that are valid in some sense or other under less restrictive assumptions than those of normality or another specific distribution type." The terms distribution-free and non-parametric are not always synonymous, however; for example, in testing statistical hypotheses, a non-parametric test is one which makes no hypothesis about the value of a parameter in a statistical density function, whereas a distribution-free test is one which makes no assumptions about the precise form of the sampled population [17].

Construction of a Distribution-free Tolerance Region

One-dimensional Case

The general method of constructing a distribution-free tolerance region is best introduced by considering the case of a one-dimensional random variable Y with continuous c.d.f. $H_Y(\cdot)$.

Let $[Y_{(1)},\ldots,Y_{(N)}]$ be the order statistics of a random sample from a population with continuous c.d.f. $H_Y(\cdot)$. In 1941 Wilks [18] showed that the symmetric interval $[Y_{(j)}, Y_{(N-j+1)}]$ could serve as a distribution-free tolerance region, and in 1943 Wald [23] derived similar results for any two order statistics. Robbins [19] subsequently showed that order statistics alone could be used to construct a distribution-free tolerance interval.

The Dirichlet distribution is used in the construction of a distribution-free tolerance interval (and region), so it is worthwhile to review the definition of this distribution and two of its properties.

Definition: Let $(A_1,\ldots,A_n)$ be an n-dimensional random variable with a probability density function (p.d.f.) of the form

$f_{\Lambda_1,\ldots,\Lambda_n}(\lambda,\ldots,\lambda_n) =$

$$\begin{cases} \dfrac{\Gamma(\nu_1+\ldots+\nu_{n+1})}{\Gamma(\nu)\ldots\Gamma(\nu_{n+1})} \lambda_1^{\nu_1-1} \ldots \lambda_n^{\nu_n-1} (1-\lambda_1-\ldots-\lambda_n)^{\nu_{n+1}-1}, & (\lambda_1,\ldots,\lambda_n) \subset S_n \\ 0 & \text{otherwise} \end{cases}$$

where S_n is the simplex $\{(\lambda_1,\ldots,\lambda_n) | \lambda_i > 0,\ i = 1,\ldots,n,\ \sum_{i=1}^{n} \lambda_i \le 1\}$ in R_n, ν_i, $i = 1,\ldots,n+1$ are real and positive, and $\Gamma(\cdot)$ denotes the gamma function.

A distribution having the above p.d.f. is called an n-variate Dirichlet distribution and is denoted by $D(\nu_1,\ldots,\nu_n;\ \nu_{n+1})$.

The two properties of a Dirichlet distribution that will be used in this section are the following [12].

Property 1: If $(\Lambda_1,\ldots,\Lambda_n)$ is distributed as the n-variate Dirichlet $D(\nu_1,\ldots,\nu_n;\ \nu_{n+1})$, then the marginal distribution of $(\Lambda_1,\ldots,\Lambda_k)$ $k < n$, is the k-variate Dirichlet $D(\nu_1,\ldots,\nu_k;\ \nu_{k+1}+\ldots+\nu_{n+1})$.

Property 2: If $(\Lambda_1,\ldots,\Lambda_n)$ is distributed as the n-variate Dirichlet distribution $D(\nu_1,\ldots,\nu_n;\ \nu_{n+1})$, then the sum $\Lambda_1+\ldots+\Lambda_n$ is distributed as a beta distribution $B(\nu_1+\ldots+\nu_n, \nu_{n+1})$.

Turning now to the construction of a distribution-free tolerance interval, consider the random intervals $(-\infty, Y_{(1)}]$, $(Y_{(1)}, Y_{(2)}]$, $\ldots$, $(Y_{(N)}, +\infty)$, and let U_i,

$i = 1,\ldots,N+1$ denote the corresponding coverages associated with these intervals. It can be shown [12] that the coverages $U_1,\ldots,U_N$ are random variables having the N-variate Dirichlet distribution $D(1,\ldots,1;1)$, which is completely symmetric in the variables. It follows from symmetry and Property 1 that any k coverages ($k < N$) have the k-variate Dirichlet distribution $D(1,\ldots,1;N-k+1)$, and from this and Property 2 it also follows that the sum of any k coverages has the beta distribution $B(k,N-k+1)$.

Now for any two order statistics $Y_{(k_1)}$, $Y_{(k_1+k_2)}$, the coverage U_{k_2} associated with the random interval $[Y_{(k_1)}, Y_{(k_1+k_2)}]$ is the sum of k_2 coverages and hence has the beta distribution $B(k_2,N-k_2+1)$. Since this holds for any distribution, then with

$$\Pr[U_{k_2} \geq \alpha] = \beta \tag{2.1}$$

for $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $[Y_{(k_1)}, Y_{(k_1+k_2)}]$ is a 100α percent distribution-free tolerance interval at probability level β.

Using K. Pearson's [20] notation for the incomplete beta function, (2.1) reduces to

$$I_{1-\alpha}(N-k_2+1,\ k_2) = \beta \tag{2.2}$$

Now for fixed α, β, k_2, there may exist no sample size N for which (2.2) holds exactly. However, since the left-hand side of (2.2) is a monotone increasing function of N, there exists a smallest integer N for which

$I_{1-\alpha}(N-k_2+1,\ k_2) \geq \beta$

For example, for $\alpha = .95$, $\beta = .99$, $k_2 = 128$, one could use the tables of the incomplete beta function [20] to find $N = 130$. It should be noted that Murphy [21] gives graphs of α as a function of N for fixed values of β and $m = N-k_2+1$ (number of intervals excluded). Somerville [15] extends Murphy's results in tabular form.

Scheffé and Tukey [22] extended the above results to the case where Y is discontinuously distributed by showing that the closed interval $[Y_{(k_1)}, Y_{(k_1+k_2)}]$ could serve as a 100α percent distribution-free tolerance interval at a probability of at least β, and the open interval $[Y_{(k_1)}, Y_{(k_1+k_2)}]$ at a probability level of at most β.

n-dimensional Case

Wald [23] extended the above method to the case of a continuously distributed n-dimensional random variable $\underline{Y}$. His resulting distribution-free tolerance region consisted of the union of rectangular regions in R_n. In this section a generalization of this method due to Tukey [24] is presented. Further generalizations, due to Fraser [14,25] and Kemperman [26], do not concern this research. The basic underlying notion in Tukey's method is that of a "statistically equivalent block" which is the multivariate analogue of the interval between two adjacent order statistics. To visualize the block construction, it is convenient to think of the random sample, O_N, as N points in R_n. Let $\phi_i(\underline{Y})$, i = 1,...,N be numerical valued functions with continuous c.d.f.'s. The exact choice of these functions will depend on the desired form of the tolerance region to be constructed. Suppose these functions are used to section R_n in the following manner:

First divide R_n into two complementary regions, θ_1 and $\bar{\theta}_1$, such that

$$\theta_1 = \{\underline{Y} | \phi_1(\underline{Y}) > W_1\}, \tag{2.3}$$

by means of the cut

$$\Omega_1 = \{\underline{Y} | \phi_1(\underline{Y}) = W_1\}$$

where

$$W_1 = \max_{k} \phi_1(\underline{Y}_k) \equiv \phi_1(\underline{Y}_{k_1})$$

which defines $\underline{Y}_{k_1}$.

Let $\bar{\theta}_1$ be divided into the two complementary regions θ_2 and $\bar{\theta}_2$, such that

$$\theta_2 = \{\underline{Y} | \phi_1(\underline{Y}) \leq W_1,\ \phi_2(\underline{Y}) > W_2\} \tag{2.4}$$

and $\bar{\theta}_2$ by means of the cut

$$\Omega_2 = \{\underline{Y} | \phi_1(\underline{Y}) < W_1,\ \phi_2(\underline{Y}) = W_2\}$$

where

$$W_2 = \max_{k \neq k_1} \phi_2(\underline{Y}_k) = \phi_2(\underline{Y}_{k_2})$$

Continue this procedure for the remaining sample points where, in general, $p \leq N$

$$\theta_p = \{\underline{y} | \phi_1(\underline{y}) < W_1, \ldots, \phi_{p-1}(\underline{y}) < W_{p-1}, \phi_p(\underline{y}) > W_p\} \quad (2.5)$$

and

$$\Omega_p = \{\underline{y} | \phi_1(\underline{y}) \leq W_1, \ldots, \phi_{p-1}(\underline{y}) < W_{p-1}, \phi_p(\underline{y}) = W_p\}$$

where

$$W_p = \max_{j \neq k_1, k_2, \ldots, k_{p-1}} \phi_p(\underline{Y}_k) = \phi_p(\underline{Y}_{k_p}) \quad (2.6)$$

The resulting regions $\theta_1, \ldots, \theta_n$, $\overline{\theta}_n$ are the statistically equivalent blocks mentioned above. In Reference [12] it is shown that the coverages $U_1, \ldots, U_N$ associated with the blocks $\theta_1, \ldots, \theta_N$ have the Dirichlet distribution $D(1, \ldots, 1; 1)$. Thus if U_m denotes the coverage of the sum of any m blocks, then U_m is distributed as a beta distribution $B(m, N-m+1)$.

Let U_{N-m+1} be the coverage of the region T_{N-m+1} formed by removing m blocks from R_n. If

$$\Pr[U_m \leq 1 - \alpha] = I_{1-\alpha}(m, N-m+1) = \beta$$

for some α, β $(0 \leq \alpha \leq 1, 0 \leq \beta \leq 1)$, then

$$\Pr[U_{N-m+1} \geq \alpha] = I_{1-\alpha}(m, N-m+1) = \beta$$

and T_{N-m+1} is a 100α percent distribution-free tolerance region at probability level β. It should be noted that if T_{N-m+1} is found by removing the first m blocks $\theta_1,\ldots,\theta_m$, then only the functions $\phi_i(\cdot)$, $i = 1,\ldots,m$ need to be specified. Furthermore, the graphs and tables for the one-dimensional case can be used to relate the parameters α, β, N, and m in the n-dimensional case.

For the case of discontinuous distributions, θ_i, $i = 1,\ldots,N+1$ are defined as above with the exception that ($<$) is replaced by ($\leq$) and ($>$) is replaced by ($\geq$). The resulting region becomes a 100α percent distribution-free tolerance region at a probability level of <u>at least</u> β. The theoretical justification for such a statement can be found in [27].

It should be noted that in dealing with discontinuous distributions, a situation might arise in which two or more sample points minimize a particular $\phi_i(\cdot)$. In such a case, the construction procedure is no longer unique, and one must specify in advance a rule for selecting among these alternative points. Tukey [27] suggested such a rule using the concept of lexicographical ordering. $(a_1,\ldots,a_n)$ is said to be less than $(b_1,\ldots,b_n)$ in the lexicographic sense if any of the following hold

1. $a_1 < b_1$

2. $a_1 = b_1$, and $a_2 < b_2$

$$\vdots$$

n. $a_i = b_i$, $i < n$, and $a_n < b_n$

By defining the functions $\Phi_1, \Phi_2, \ldots, \Phi_N$ as

$$\Phi_i(\cdot) = \{\phi_i(\cdot), \phi_{i+1}(\cdot), \ldots, \phi_N(\cdot)\}$$

a tie-breaking rule would be to select the sample point for which $\Phi_i(\cdot)$ is minimized in the lexicographical sense. For example, if r points minimize the function $\phi_1(\cdot)$, then find the r_1 points among these r points that minimize the function $\phi_2(\cdot)$. If $r_1 = 1$, then select the point which minimizes $\phi_2(\cdot)$; otherwise find the r_2 points among the r_1 points that minimize $\phi_3(\cdot)$. Continue the procedure until $r_i = 1$, $i < N$, or $r_N > 1$. In the latter case the method of constructing the sample blocks will be the same regardless of the point selected among the r_N points.

In conclusion, this chapter encompasses developments and refinements of the theory and methods of distribution-free tolerance regions specifically for application to chance-constrained linear programming. Chapter III merges this material with the theory and methods of linear programming to formulate new procedures for chance-constrained linear programming with distribution-free constraints.

CHAPTER III

DISTRIBUTION-FREE CONSTRAINT SETS

In this chapter methods are developed for constructing a distribution-free set $S(\alpha,\beta)$ such that for any $\underline{x} \subset S(\alpha,\beta)$ it can be asserted with a preassigned confidence, β, that a constraint will hold at least 100α percent of the time. The required number of samples is a direct function of the values assigned to α and β.

The Distribution-free Set

The meaning of a distribution-free set S is best explained by considering the chance-constraint

$$Pr[Ax \leq B] \geq \alpha \tag{3.1}$$

where A and B are random variables. Let $C_{(1)},\ldots,C_{(N)}$ be the order statistics of a sample of size N from the distribution of $C = B/A$. If $U_1,\ldots,U_{N+1}$ are the coverages associated with the random intervals $(-\infty,C_{(1)}]$, $[C_{(1)},C_{(2)}]$, ..., $[C_{(N)},+\infty)$ and U' is the sum of the coverages $U_2,\ldots,U_{N+1}$, there exists a β for which $Pr[U' \geq \alpha] = \beta$. The random interval $[C_{(1)},+\infty)$ is a 100α percent distribution-free interval at probability level β. Thus, if $c_{(1)}$

is the observed value of $C_{(1)}$, it can be asserted with a confidence, β, that $\Pr[C \geq c_{(1)}] \geq \alpha$. Then if $S(\alpha,\beta) = \{x | x \leq c_{(1)}\}$, for any $x \subset S(\alpha,\beta)$ it can be asserted with a confidence, β, that $\Pr[C \geq x] \geq \alpha$ or $\Pr[Ax \leq B] \geq \alpha$.

For the general case of q chance-constraints it is desired to find q distribution-free sets with the above property. That is, for any $\underline{x} \subset S^i$ it can be asserted with a confidence of at least β_i that $\Pr[\underline{A}_i\ \underline{x} \leq B_i] \geq \alpha_i$. Maximization of the objective function would then be over the intersection of these q sets. The next section describes the fundamental approach to be taken in this research for constructing a distribution-free set for a particular constraint. For convenience the superscript, i, is omitted, and the right-hand side, B, fixed at one. When B is random, the procedures which follow are applicable to the random vector $\underline{A}/B$.

Constructing a Distribution-free Set

The approach for constructing a distribution-free set can be described in two basic steps.

1. Construct a 100α percent distribution-free tolerance region with confidence, β, from samples of the elements of $\underline{A}$. Denote this region by $T(\alpha,\beta)$.

2. Determine the set $S(\alpha,\beta)$ such that for any $\underline{x} \subset S(\alpha,\beta)$, $\underline{A}\ \underline{x} \leq 1\ \forall \underline{A} \subset T(\alpha,\beta)$. The justification for taking these steps is that for any $\underline{x} \subset S(\alpha,\beta)$ the half-space

$\{\underline{A} | \underline{A}\ \underline{x} \leq 1\}$ contains a 100α percent distribution-free tolerance region with confidence β. Hence, for any $\underline{x} \subset S(\alpha,\beta)$ it can be asserted with a confidence of at least β that $\Pr[\underline{A}\ \underline{x} \leq 1] \geq \alpha$.

To illustrate the construction of a deterministic set $S(\alpha,\beta)$ from a distribution-free tolerance region, let such a region be constructed by removing a statistical block with the linear cutting function $\phi = \underline{A}\ \underline{y}$. The elements of $\underline{y}$ are assumed to be arbitrarily chosen and constant. This region is given by

$$T(\alpha,\beta) = \{\underline{A} | \underline{A}\ \underline{y}^* \leq 1\}$$

where

$$y_j^* = y_j/W \qquad j = 1,\ldots,n$$

and

$$W = \max_k \phi(\underline{A}_k) \qquad k = 1,\ldots,N$$

The desired set corresponding to this region would then be given by

$$S(\alpha,\beta) = \{\underline{x} | \underline{x} = \lambda\underline{y},\ 0 \leq \lambda \leq 1,\ \underline{x} \geq 0\}$$

Example 3.1 further illustrates this approach.

<u>Example 3.1</u>

Suppose it is desired to find a set $S(.5, .95)$ such that for any $\underline{x} = (x_1, x_2) \subset S(.5, .95)$ it can be asserted with a confidence of at least .95 that

$$Pr[A_1x_1 + A_2x_2 \leq 1] \geq .50$$

Choosing $y_1 = 1$ and $y_2 = 2$ (arbitrarily), the function $\phi = A_1 + 2A_2$ is used to generate a tolerance region $T_1(.5, .95)$. The required sample size is then the smallest integer value N satisfying the relationship

$$I_{.5}(1, N) \geq 1 - .95$$

which can be shown to be N = 5.

To illustrate, a random sample of size N = 5 was taken from a population of independently and identically distributed normal variates with means and variances of 3 and 1, respectively. The resulting sample values are $(A_{1,1}, A_{2,1}) = (3.485, 2.618)$; $(A_{1,2}, A_{2,2}) = (4.345, 1.398)$; $(A_{1,3}, A_{2,3}) = (.538, 1.534)$; $(A_{1,4}, A_{2,4}) = (3.043, .361)$; and $(A_{1,5}, A_{2,5}) = (2.084, 3.598)$. Then $W = \max_k(A_{1k} + 2A_{2k}) = 2.084 + 2(3.598) = 9.280$, and $T(.5, .95) = \{A_1, A_2 | .108 A_1 + .216 A_2 \leq 1\}$.

The corresponding distribution-free set is then

$$S(.5, .95) = \{x_1,x_2 | (x_1,x_2) = \lambda(.108, .216),$$
$$0 \leq \lambda \leq 1, (x_1,x_2) \geq 0\}$$

A scatter diagram of the original sample points is given in Figure 3.1, along with the tolerance region T(.5, .95). Figure 3.2 contains a graphical representation of the set S(.5, .95).

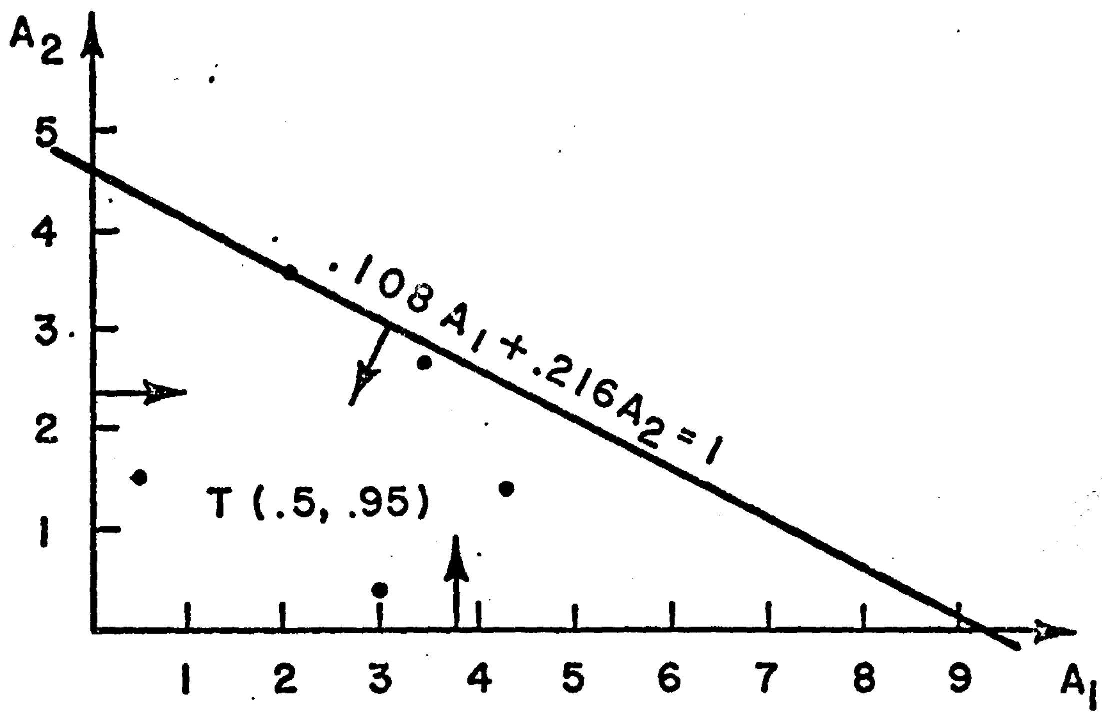


Figure 3.1. T(.5, .95) for Example 3.1.

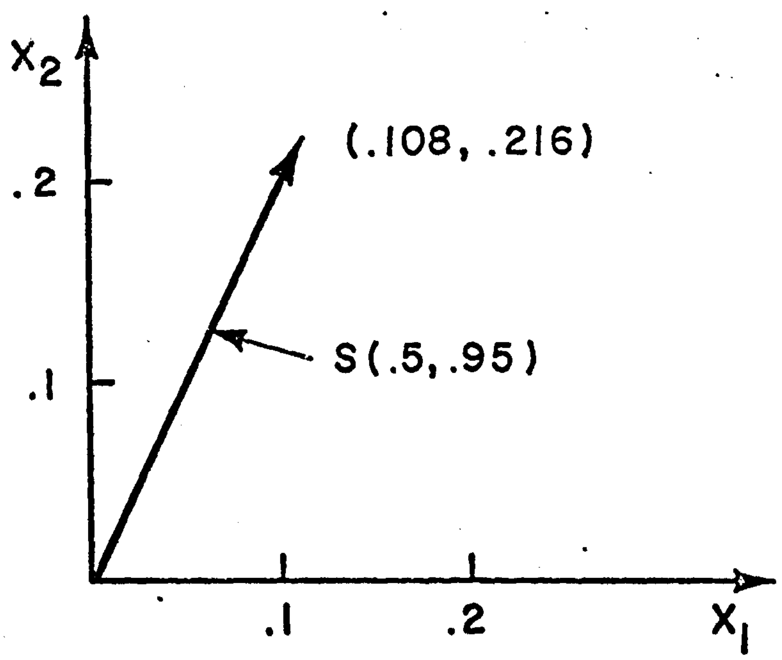


Figure 3.2. S(.5, .95) for Example 3.2.

The foregoing approach was used to illustrate the notion of constructing a deterministic set S from a distribution-free tolerance region T. The relative merit of such an approach is dubious when dealing with more than one constraint, since the choice of $\underline{x}$ is restricted to points along a vector in R_n . The same cutting function must be used for each constraint, otherwise the only choice for $\underline{x}$ would be the origin. The two methods which follow provide considerably more freedom in the choice of the shape of the distribution-free regions corresponding to each constraint.

A Distribution-free Linear Constraint Set

It is possible to represent a distribution-free set as a linear constraint set in the following manner. First construct a distribution-free tolerance region $T_L(\alpha,\beta)$ using a sequence of cutting functions of the form

$$\phi_j = A_j \qquad j = 1,\ldots,n$$

The resulting region would be given by

$$T_L(\alpha,\beta) = \{\underline{A} | \underline{A} \leq \underline{W}\} \tag{3.2}$$

where the elements of $\underline{W} = (W_1,\ldots,W_n)$ are determined from Eqs. (2.3) and (2.6) of the previous chapter. The desired set S_L is then given by

$$S_L(\alpha,\beta) = \{\underline{x} | \underline{W}\ \underline{x} \leq 1,\ \underline{x} \geq 0\} \tag{3.3}$$

as is evidenced by the following theorem.

Theorem 3.1

Let S_L and T_L be given by Eqs. (3.2) and (3.3), respectively. (For convenience the (α,β) designation is deleted.) Then a necessary and sufficient condition for $\underline{A}\,\underline{x} \leq 1 \;\forall \underline{A} \subset T_L$ is that $\underline{x} \subset S_L$.

Proof

(Sufficient)

Assume $\underline{x} \subset S_L$

Show

$$\underline{A}\,\underline{x} \leq 1 \;\forall\; \underline{A} \subset T_L \tag{3.4}$$

Rewrite Eq. (3.4) as

$$\underline{W}\,\underline{x} - 1 \leq (\underline{W} - \underline{A})\,\underline{x} \tag{3.5}$$

Now the r.h.s. of Eq. (3.5) is always greater than or equal to zero $\forall \underline{A} \subset T_L$, so the inequality will always hold provided the l.h.s. is non-positive.

Since by definition of the set S_L

$$\underline{W}\,\underline{x} - 1 \leq 0$$

the sufficiency part of the proof is complete. The necessary part of the proof follows since

$$\underline{A}\,\underline{x} \leq 1 \;\forall \underline{A} \subset T_L$$

and in particular, the relationship

$\underline{W}\,\underline{x} \leq 1$

must be satisfied.

Example 3.2

For comparative purposes the problem stated in Example 3.1 will be used, along with the same five sample values. However, before proceeding it is necessary to increase the sample size to N = 8, since this is the smallest number for which

$I_{.5}(2, N-1) \geq 1 - .95$

The additional simulated sample values are found to be (2.502, 2.972), (2.341, 2.143) and (3.456, 4.116).

Figure 3.3 contains a scatter diagram of the eight sample values along with the resulting tolerance region

$T_L(.5, .95) = \{A_1, A_2 | A_1 \leq 4.345,\ A_2 \leq 4.116\}$

The desired linear set (shown in Figure 3.4) is then given by

$S_L(.5, .95) = \{x_1, x_2 | 4.345\ x_1 + 4.116\ x_2 \leq 1,\ (x_1, x_2) \geq 0\}$

A major disadvantage of the above approach is that as the number of variables, n, increases, so does the number of required cuts, m. This in turn requires a larger sample size N for fixed levels of α and β. As is shown in Table 3.1,

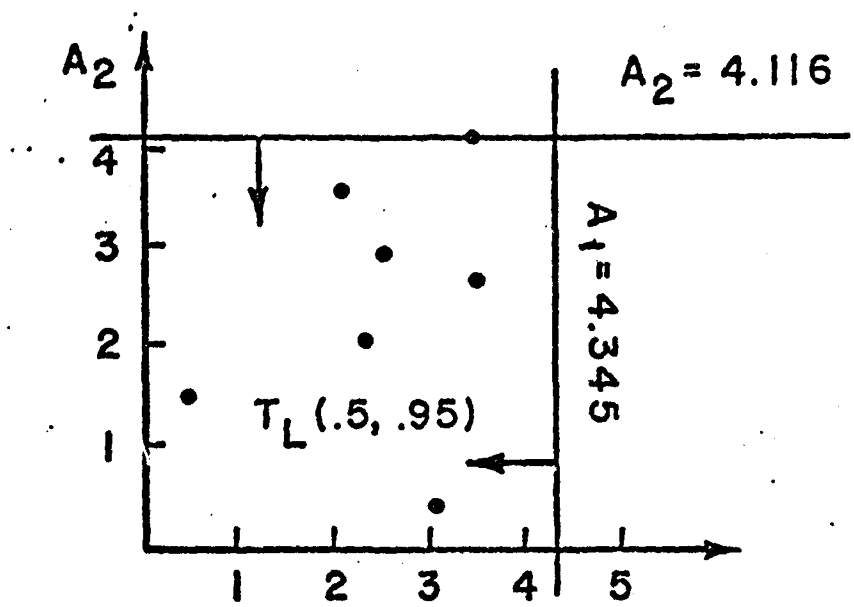


Figure 3.3. $T_L(.5, .95)$ for Example 3.2.

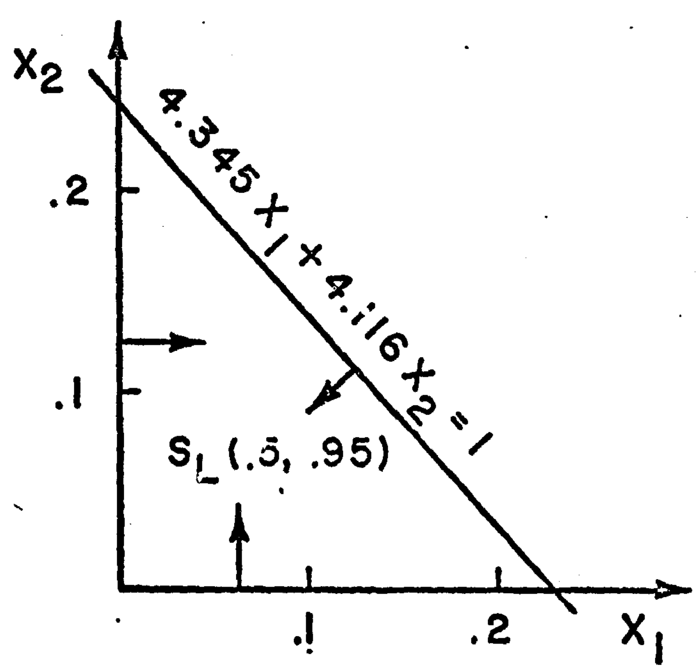


Figure 3.4. $S_L(.5, .95)$ for Example 3.2.

with $\alpha = .90$ and $\beta = .95$, the size of N for even modest values of m is quite large. For the case of limited or costly data, this restriction could be very significant.

Table 3.1

Values of m and N with $\alpha = .90$; $\beta = .95$

m	N
5	90
10	155
15	215
20	275
25	335
30	390
40	500
50	600

The next section shows how a spherical cutting function can be used to construct a convex distribution-free constraint set without the above restriction.

A Distribution-free Convex Constraint Set

Suppose a distribution-free tolerance region is constructed via the cutting function

$$\phi(\underline{A}) = |\underline{A} - \underline{d}| = \left(\sum_{j=1}^{n} (A_j - d_j)^2\right)^{1/2}$$

where $\underline{d}$ is a row vector of preassigned constants. The resulting tolerance region is given by

$$T_S(\alpha,\beta) = \{\underline{A} \mid |\underline{A} - \underline{d}| \leq \rho\} \tag{3.6}$$

where

$$\rho = \max_k |\underline{A}_k - \underline{d}|$$

This region is the surface and interior of an n-dimensional hypersphere centered at $\underline{d}$ with radius ρ. The corresponding distribution-free set is given by

$$S_S(\alpha,\beta) = \{\underline{x} \mid |\underline{x}| \leq (1 - \underline{d}\,\underline{x})/\rho,\ \underline{x} \geq 0\}$$

Theorem 3.2

Let T_S be given by Eq. (3.6). A necessary and sufficient condition for $\underline{A}\,\underline{x} \leq 1\ \forall \underline{A} \subset T_S$ is

$$|\underline{x}| \leq (1 - \underline{d}\,\underline{x})/\rho \tag{3.7}$$

Proof

(Sufficient)

Assume Eq. (3.7) holds. Show that $\underline{A}\,\underline{x} \leq 1\ \forall \underline{A} \subset T_S$.

$$\begin{aligned}
\underline{A}\,\underline{x} &= \underline{d}\,\underline{x} + (\underline{A} - \underline{d})\,\underline{x} \\
&\leq \underline{d}\,\underline{x} + |\underline{A} - \underline{d}|\,|\underline{x}| \\
&\leq \underline{d}\,\underline{x} + [(1 - \underline{d}\,\underline{x})/\rho]\,\rho \\
&\leq 1
\end{aligned}$$

(Necessary)

Assume $\underline{A}\ \underline{x} \leq 1\ \forall \underline{A} \subset T_S$. Show that Eq. (3.7) holds.

$$\underline{A}\ \underline{x} \leq 1$$

$$\underline{d}\ \underline{x} + (\underline{A} - \underline{d})\ \underline{x} \leq 1$$

$$(\underline{A} - \underline{d})\ \underline{x} \leq 1 - \underline{d}\ \underline{x}$$

$$|\underline{A} - \underline{d}|\ |\underline{x}| \cos(\underline{A} - \underline{d}, \underline{x}) \leq 1 - \underline{d}\ \underline{x} \tag{3.8}$$

Note that if Eq. (3.8) holds for points on the surface of the hypersphere defined by T_S, then this relationship also holds for all points contained in this hypersphere. Thus $|\underline{A} - \underline{d}|$ can be replaced with ρ in Eq. (3.8) to give

$$|\underline{x}| \cos(\underline{A} - \underline{d}, \underline{x}) \leq (1 - \underline{d}\ \underline{x})/\rho \tag{3.9}$$

For Eq. (3.9) to hold, it must hold for an $\underline{A}^*$ on the surface of T_S for which $\cos(\underline{A}^* - \underline{d}, \underline{x}) = 1$. Such a point exists and is given by $\underline{A}^* = [(\rho\underline{x})/|\underline{x}|] + \underline{d}$. Replacing $\cos(\underline{A} - \underline{d}, \underline{x})$ with $\cos(\underline{A}^* - \underline{d}, \underline{x}) = 1$ in Eq. (3.9) yields

$$|\underline{x}| \leq (1 - \underline{d}\ \underline{x})/\rho$$

and the end of the necessary part of the proof.

The convexity of the set S_S can be proven by letting $\underline{x}^* = \lambda\underline{x}_1 + (1 - \lambda)\ \underline{x}_2$ where $0 \leq \lambda \leq 1$ and $\underline{x}_1, \underline{x}_2 \subset S_S$. Then

$$|\underline{x}^*| - (1 - \underline{d}\ \underline{x}^*)/\rho =$$

$$|\lambda\underline{x}_1 + (1 - \lambda)\ \underline{x}_2| - [1 - (\underline{d}(\lambda\underline{x}_1 + (1 - \lambda)\ \underline{x}_2)]/\rho$$

$$\leq \lambda|\underline{x}_1| + (1 - \lambda)|\underline{x}_2| - [1 - \lambda \underline{d}\ \underline{x}_1 - (1 - \lambda)\ \underline{d}\ \underline{x}_2]/\rho$$

$$\leq \{\lambda(1 - \underline{d}\ \underline{x}_1) + (1 - \lambda)(1 - \underline{d}\ \underline{x}_2) - [1 - \lambda\ \underline{d}\ \underline{x}_1 - (1 - \lambda)\ \underline{d}\ \underline{x}_2]\}/\rho$$

$$\leq \lambda + (1 - \lambda) - 1$$

$$\leq 0$$

The relationship (3.7) can be described geometrically as the surface and interior of a sphere, elipsoid, paraboloid or one nappe of a hyperboloid depending on δ (i.e., $\delta = \rho$, <0, $= 0$ or >0) where $\delta = \rho - |\underline{d}|$. This is illustrated in Figures 3.5 through 3.8, where $\underline{d}$ varies and ρ remains fixed at .5. Also included in these figures are the corresponding tolerance regions described by Eq. (3.6).

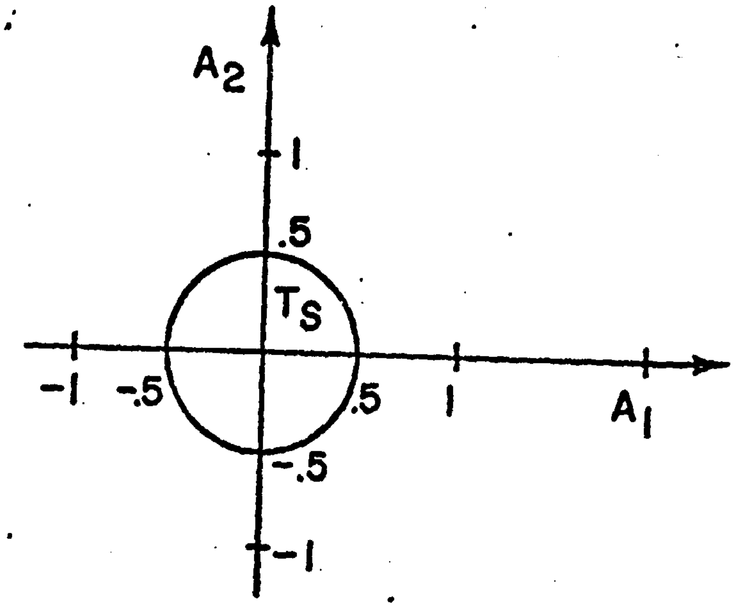
Example 3.3 illustrates the foregoing method with respect to the preceding examples.

Example 3.3

For convenience, the circular cutting function with $\underline{d} = (0,0)$ is considered. Since only one cut is required, the original five sample points are used to determine the value of $\rho = (A_{12}^2 + A_{22}^2)^{1/2} = 4.5$. Then the resulting tolerance region (shown in Figure 3.4) is given by

$$T_S(.5, .95) = \{(A_1, A_2) | (A_1^2 + A_2^2)^{1/2} \leq 4.5\}$$

and the corresponding distribution-free set (shown in Figure 3.5) is given by



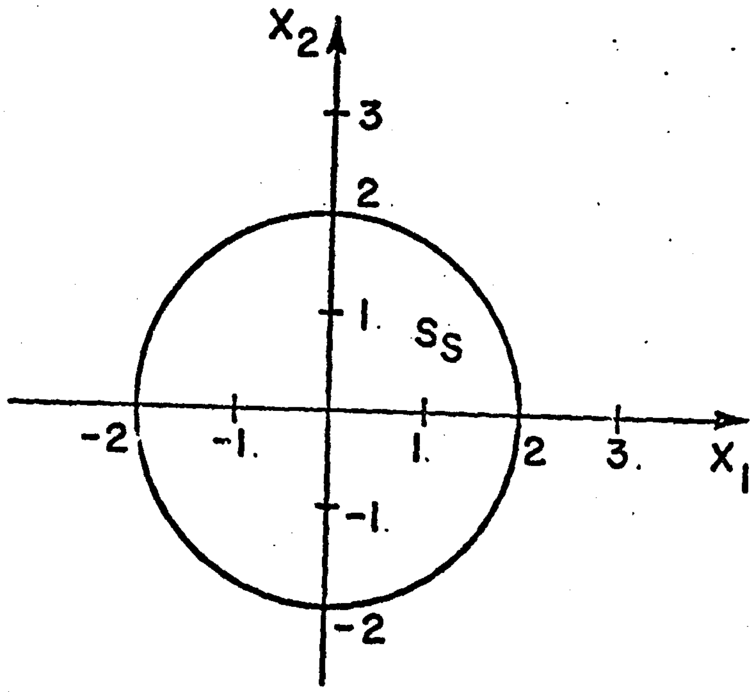


Figure 3.5. T_S and S_S with $\delta = \rho$.

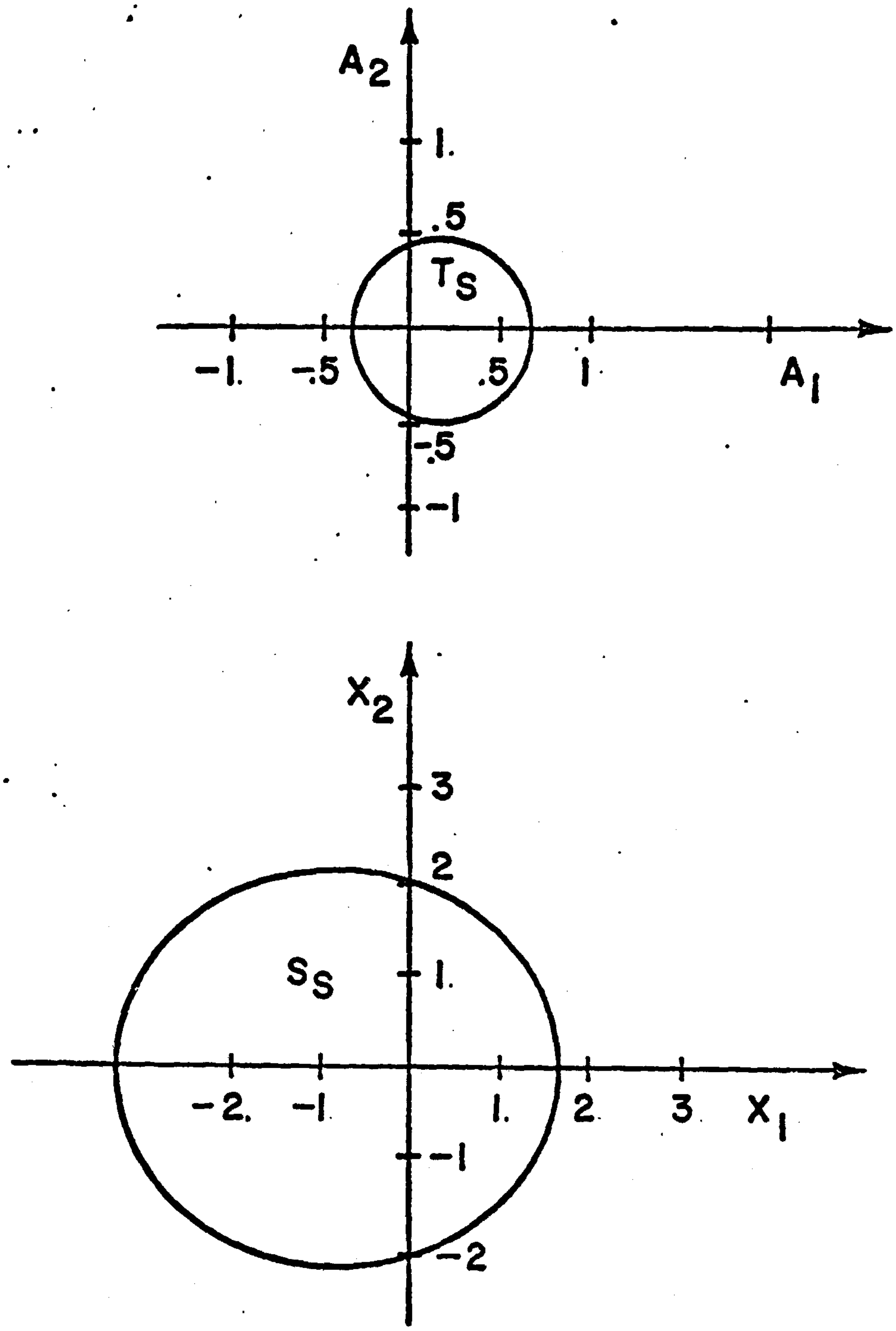


Figure 3.6. T_S and S_S with $\delta < 0$.

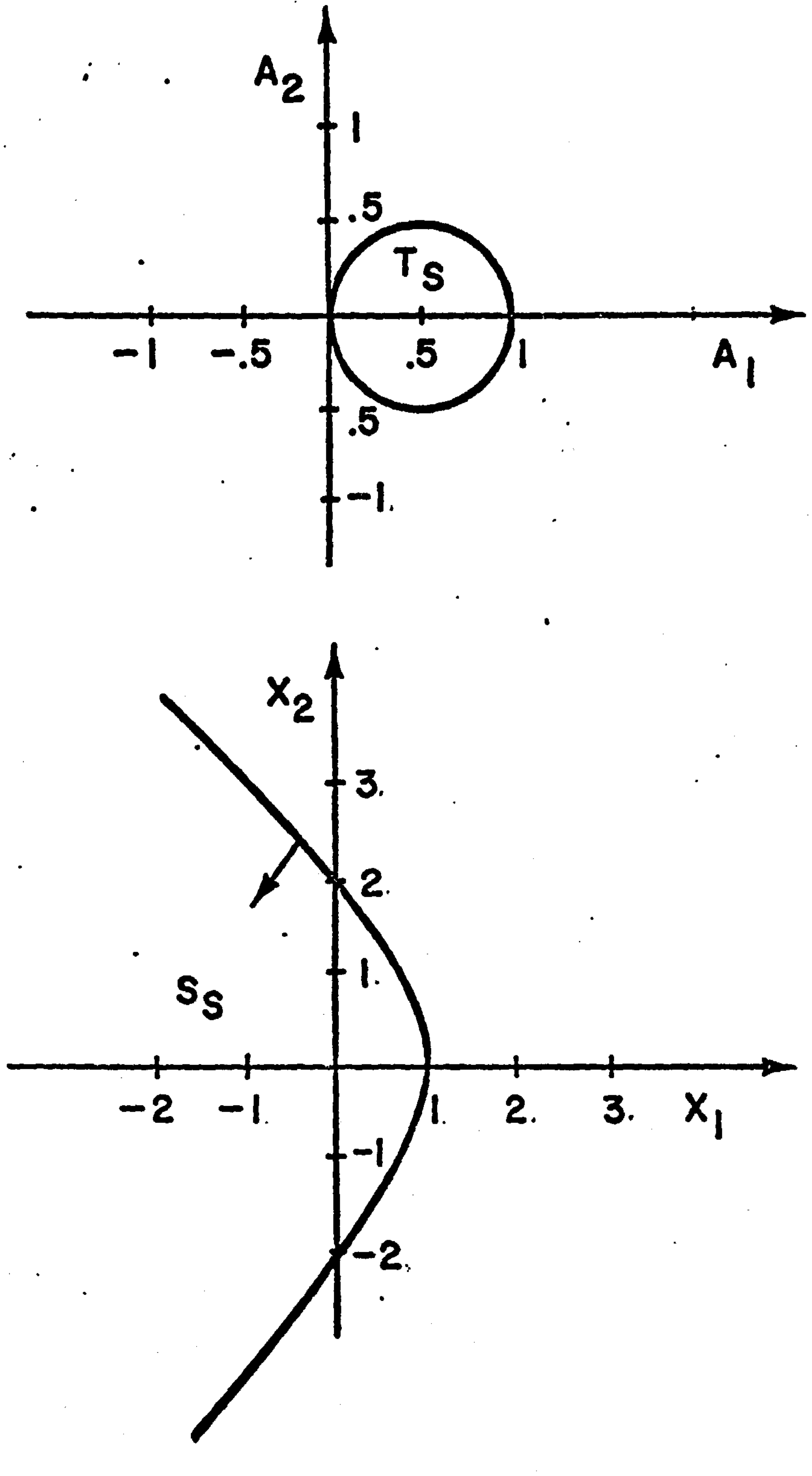


Figure 3.7. T_S and S_S with $\delta = 0$.

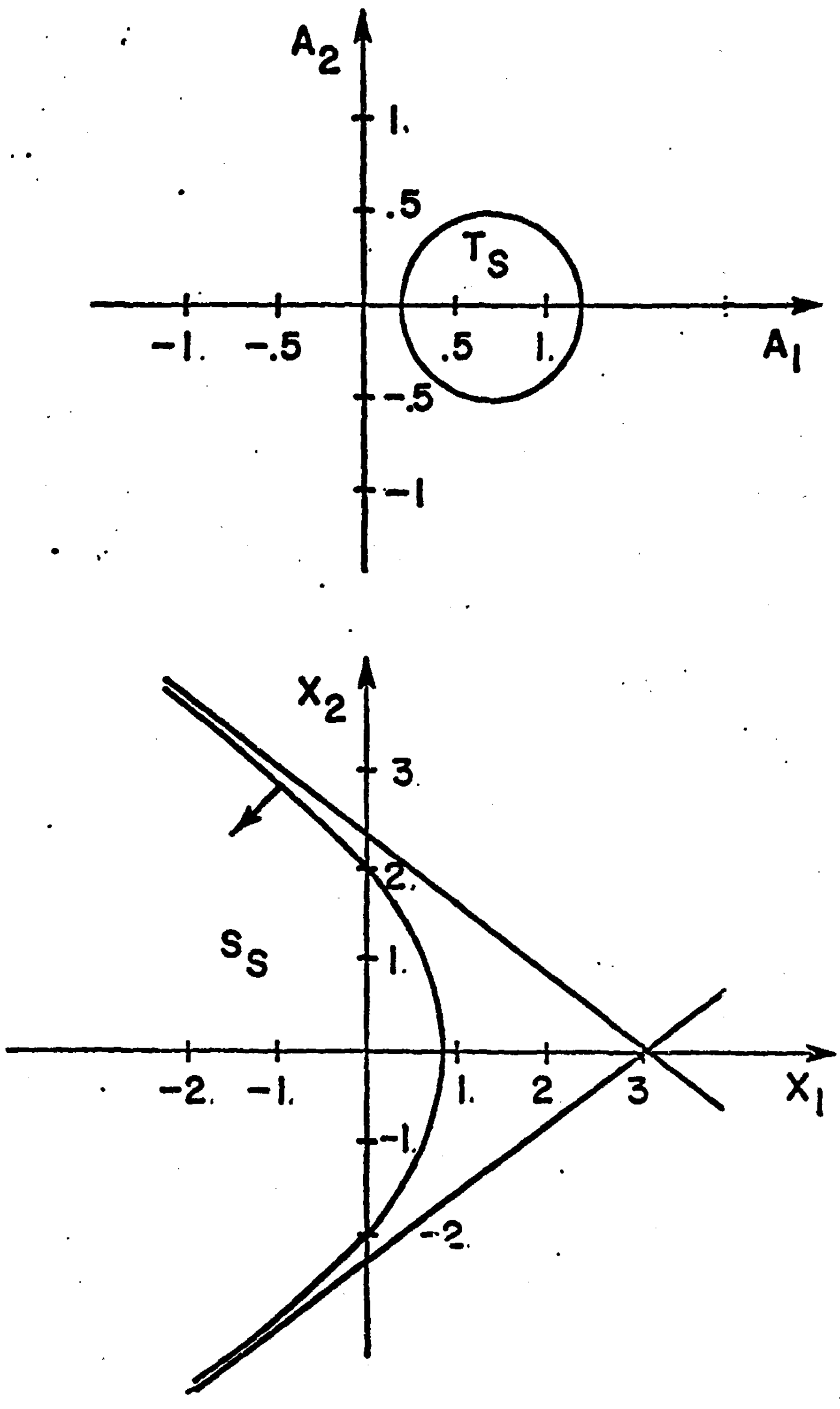


Figure 3.8. T_S and S_S with $\delta > \rho$.

$$S_S(.5, .95) = \{x_1, x_2 | (x_1^2 + x_2^2) \leq 1/4.5,\ (x_1, x_2) \geq (0,0)\}$$

The tolerance regions and distribution-free sets of the foregoing examples are shown together in Figures 3.9 and 3.10, respectively.

Expanding the Size of a Distribution-free Set

This section is concerned with the problem of expanding the size of a distribution-free set, $S(\alpha,\beta)$, after it has been constructed from a sample of size N. Such an expansion might be motivated by an undesirable value of the objective function obtained by maximizing over $S(\alpha,\beta)$. If this occurs in the use of the Quantile or Tchebysheff Methods, the sets can be expanded by reducing the preassigned probability level, α, for constraint satisfaction. This results in a smaller value of K_α or $(\alpha/1 - \alpha)$ and thus increases the size of the respective sets S_Q, S_T as described by Eqs. (1.3) and (1.4) in Chapter I. In the case of a distribution-free set the problem could be similarly resolved by reassigning lower levels of α and β and repeating the construction procedure with reduced sample sizes. If there are no samples available, then it may be possible to obtain a larger set by reducing the original tolerance region $T(\alpha,\beta)$ by taking additional cuts. The coverage of the resulting region is described in the following theorem. (The proof of theorem 3.3 is presented in the Appendix.)

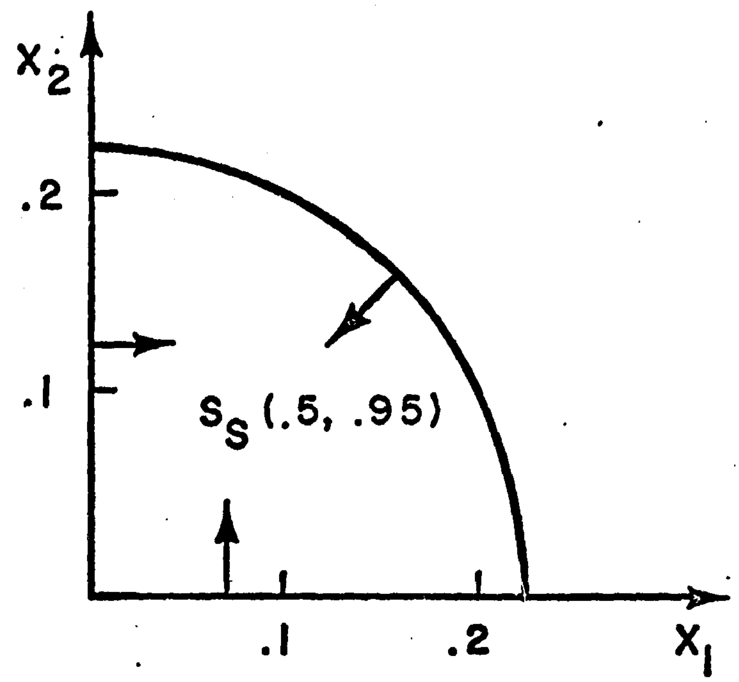


Figure 3.9. $S_S(.5, .95)$ for Example 3.3.

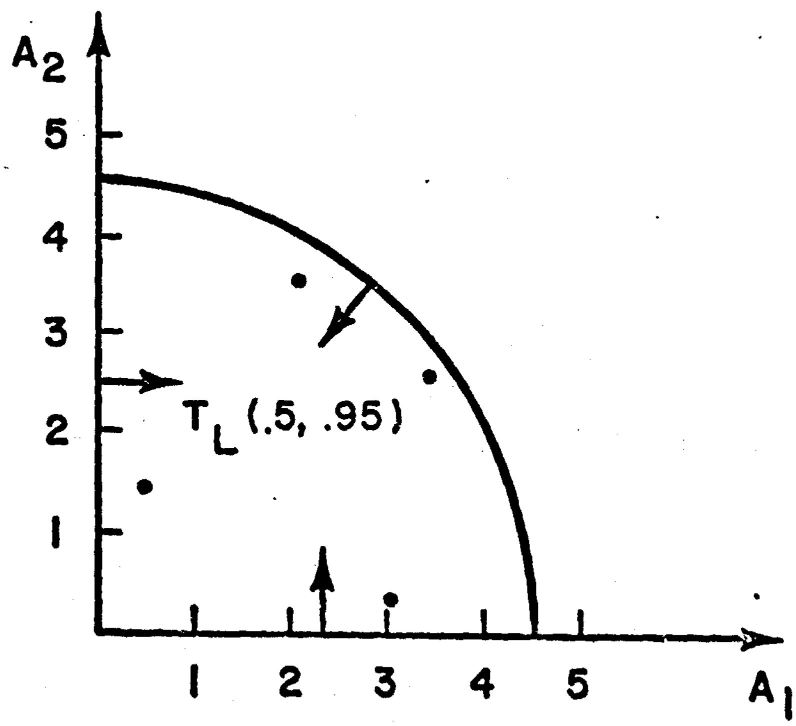


Figure 3.10. $T_L(.5, .95)$ for Example 3.3.

Theorem 3.3

Let U^m be the coverage of the region $T(\alpha,\beta)$ constructed from a sample of size N by removing m blocks. Let $U^{m+m'}$ be the coverage of the region $T'(\alpha',\beta')$ formed by removing m' additional blocks from $T(\alpha,\beta)$. Then

$$\Pr[U^{m+m'} \geq \alpha'] = 1 - I_{\alpha'/U^m}(N-m-m'+1,\ m') = \beta' \qquad (3.8)$$

From the above theorem it is seen that the confidence level, β', associated with the region $T(\alpha',\beta')$ is dependent upon the coverage, U^m, of the original region, $T(\alpha,\beta)$. Once the sample has been drawn, U^m is a fixed but unknown quantity. Thus, it is impossible to determine the value of β' for a given level of α'. However, relationship (3.8) can be used to approximate the coverage, $U^{m+m'}$, by replacing U^m with a suitable estimate. One such estimate is the original value of α, since it is known with a confidence of at least β that $U^m \geq \alpha$.

CHAPTER IV

EXPERIMENTATION AND COMPUTATIONAL RESULTS

This chapter includes the results of investigations into the performance of linear and spherical distribution-free constraint sets using simulated data from a non-normal distribution. The value of such investigations is two-fold. First, it provides a clearer understanding of the meaning and interrelationship of the parameters α and β. Second, it provides a means of comparing the relative merit of a distribution-free set versus one obtained using the Quantile or Tchebysheff Method in the absence of any knowledge of the underlying distribution.

Notation and Assumptions

Consider the single chance-constraint

$$\Pr\{\underline{A}\ \underline{x} \leq 1\} \geq \alpha,\ \underline{x} \geq 0 \qquad (4.1)$$

Let $S_S(\alpha,\beta)$ denote a convex distribution-free set as described in the previous chapter. All $\underline{x} \subset S_S(\alpha,\beta)$ will satisfy (4.1) with a confidence of at least β. Let $S_L(\alpha,\beta)$ denote a linear distribution-free set with the same property. $S_T(\alpha)$ and $S_Q(\alpha)$ denote sets obtained by the Tchebysheff and

Quantile Methods. The information required to construct the above sets is based upon sample data from independent and identically distributed gamma variates with parameters $\mu = 10$ and $\nu = 5$. These variates were generated on an IBM 360-65 computer using the FORTRAN program suggested in reference [28].

Construction of $S_S(\alpha,\beta)$, $S_L(\alpha,\beta)$, $S_T(\alpha)$ and $S_Q(\alpha)$

Consider the case of n = 2, and suppose it is desired to find a region in the positive quadrant of $\underline{x} = (x_1,x_2)$ such that any point in this region will satisfy the chance-constraint

$$Pr[A_1\ x_1 + A_1\ x_2 \leq 1] \geq .90$$

Such a region can be determined using a spherical cutting function with a confidence of .95 from a sample of size N = 29. For the purpose of generality this region is constructed using a spherical cutting function with $\underline{d} = (0,0)$. Table 4.1 contains the 29 simulated sample points $(A_{1k},A_{2k},$ $k=1,\ldots,29)$. The sixth sample value (.706, .734) yields the maximum value of $\rho = (.706)^2 + (.734)^2 = 1.037$. The resulting distribution-free set is

$$S_S(.90,\ .95) = \{x_1,x_2 \mid (x_1^2 + x_2^2)^{1/2} \leq (1/1.037)^{1/2}\}$$

Table 4.1

29 Simulated Samples of (A_1, A_2)

k	A_{1k}	A_{2k}	k	A_{1k}	A_{2k}	k	A_{1k}	A_{2k}
1	.493	.380	11	.288	.383	21	.492	.527
2	.773	.405	12	.615	.365	22	.623	.777
3	.490	.382	13	.293	.303	23	.252	.565
4	.384	.472	14	.358	.450	24	.405	.398
5	.277	.456	15	.651	.525	25	.718	.408
6	.706	.734	16	.685	.412	26	.484	.215
7	.635	.416	17	.421	.070	27	.890	.399
8	.446	.670	18	.489	.630	28	.317	.725
9	.172	.122	19	.650	.458	29	.232	.329
10	.366	.625	20	.419	.592	--	--	--

To illustrate empirically the meaning of $\alpha = .90$ and $\beta = .95$, the above procedure for constructing $S_S(.90, .95)$ was repeated for 99 additional sample sets of size $N = 29$. For each set, the surface point $\underline{x}^*$ for which $x_1 = x_2$ was selected and in 1,000 realizations of the random variables A_1, A_2, the number of times that the relationship $A_1x_1^* + A_2x_2^* \leq 1$ was satisfied was recorded, and denoted as ALPHA. In BETA = 96 of the 100 trials, the value of ALPHA was found to be greater or equal to 900.

Table 4.2 exhibits these values along with other observed values of BETA for various values of ALPHA. These observations can be compared with actual values of α and β with $N = 29$ and $m = 1$ in Table 4.3.

Table 4.2

Observed Values of ALPHA and BETA

ALPHA	BETA
800	99
850	99
900	96
950	76
960	70
970	59

Table 4.3

Actual Values of α and β with $N = 29$ and $m = 1$

α	β
.800	.998
.850	.992
.900	.953
.950	.775
.960	.697
.970	.585

A linear distribution-free set $S_L(.90, .95)$ would require additional sample points since two cuts are needed (as opposed to one in the spherical case). Rather than taking any more samples, the set $S_L(.90, .83)$ is constructed from the original sample of size 29. The maximum value of A_{1k} is given by $A_{1,27} = .890$. The maximum value of $A_{2,k}$ (after deleting $A_{2,27}$) is $A_{2,22} = .777$. The resulting linear set is given by

$$S_L(.90, .83) = \{x_1, x_2 | .890\ x_1 + .777\ x_2 \leq 1\}$$

To construct the set $S_T(.90)$ using the Tchebysheff Method, it is necessary to calculate sample means and variances from the 29 sample values of Table 4.1. The resulting set is given by

$$S_T(.90) = \{x_1, x_2 | .484\ x_1 + .455\ x_2 + 3.0(.033\ x_1^2 + .028\ x_2^2)^{1/2} \leq 0\}$$

Assuming (erroneously) A_1 and A_2 to be independent normal variates, the Quantile Method could be used to generate the set

$$S_Q(.90) = \{x_1, x_2 | .484\ x_1 + .455\ x_2 + 1.282(.033\ x_1^2 + .028\ x_2^2)^{1/2} \leq 0\}$$

Comparative Analysis

To illustrate geometrically the relative accuracy of the above sets with respect to the true set S = $\{x_1, x_2 | Pr[A_1\ x_1 + A_2\ x_2 \leq 1] \geq .90\}$ the boundary of this set was approximated in the following manner. For a fixed value of x_1, the value of x_2 was incremented in units of .02 until such time that |ALPHA-900| $\leq$ 5. The procedure was then repeated for incremental (.2) values of x_1. The resulting values of x_1, x_2, and ALPHA are presented in Table 4.4.

Table 4.4

Approximate Boundary Points of Actual Region S

x_1	x_2	ALPHA
0	1.16	898
.2	1.06	903
.4	.96	895
.6	.80	904
.8	.58	901
1	.32	896

These points were used to approximate the true region with the region S shown in Figure 4.1, which also contains the sets $S_S(.90, .95)$, $S_L(.90, .83)$, $S_T(.90)$ and $S_Q(.90)$.

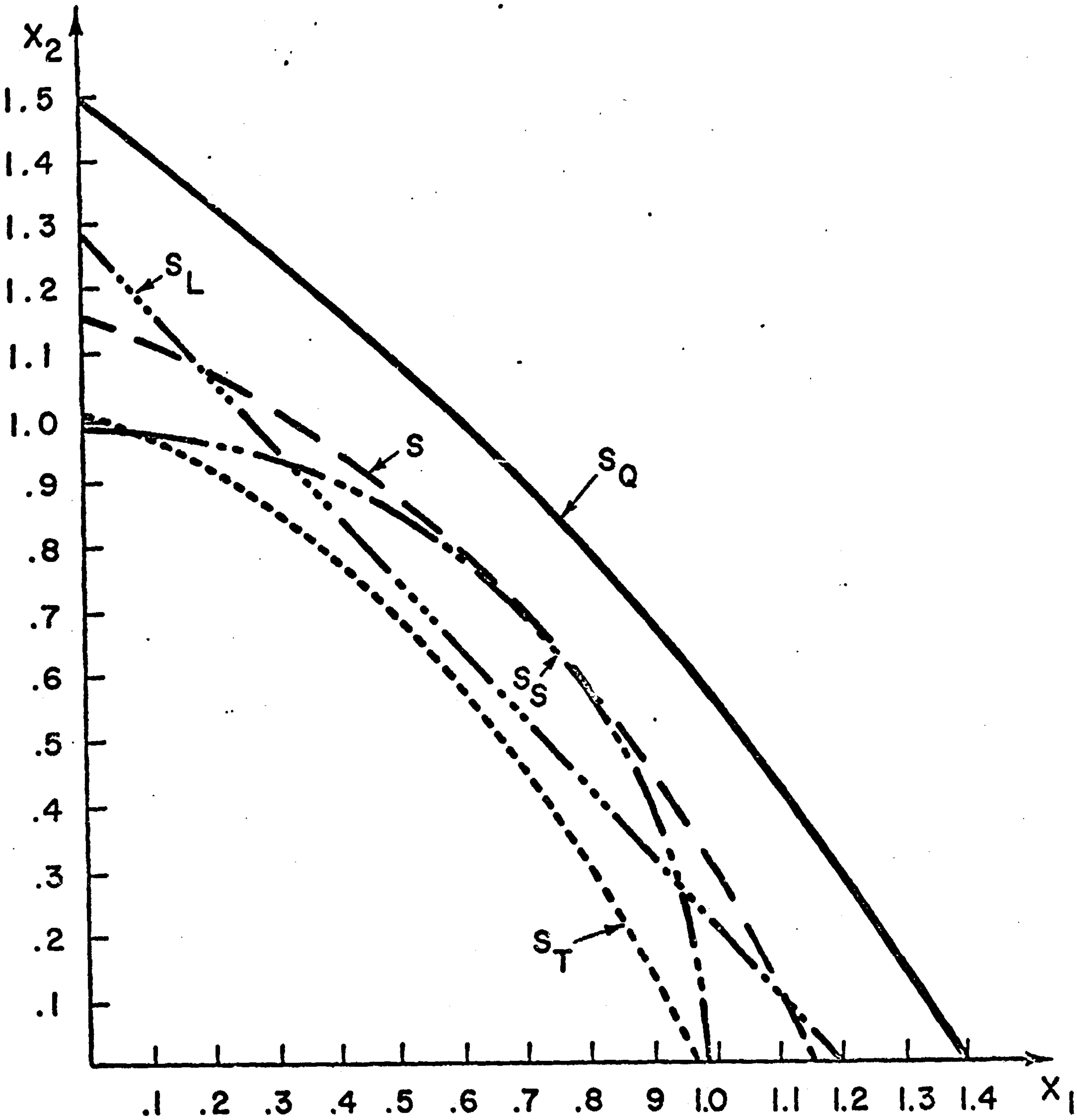


Figure 4.1. Comparison of chance-constrained sets with no knowledge of the underlying distribution.

The following observations are made from this figure.

A. The sets S_S, S_L, S_T are conservative with respect to the degree in which boundary points satisfy the constraint more than 90 percent of the time.

B. The set S_T obtained via the Tchebysheff Method is the most conservative.

C. The set S_Q obtained via the Quantile Method yields a considerably larger region, but points along the boundary will violate the constraint more than 10 percent of the time.

D. The boundaries of the sets S_Q and S_T follow the shape of the true boundary more closely than those of either S_S or S_L.

The first three observations are illustrated numerically by considering various points along the respective boundaries and checking the constraint satisfaction with samples of size N = 1,000. In particular, the points considered are those which maximize the value of

$$z = c_1 x_1 + c_2 x_2$$

for values of $\underline{c}$ = (1,1), (2,1) and (4,1). The resulting values x_1, x_2 and corresponding values of ALPHA for each set are presented in Table 4.5.

To determine if the observations A, B and C could be made for higher dimensions, a similar analysis is first performed for the case of n = 10. The boundary points considered

Table 4.5

Empirical Constraint Satisfaction

c	Set	x_1	x_2	ALPHA
(1,1)	S_S	.69	.69	905
	S_L	.00	1.24	892
	S_T	.54	.66	962
	S_Q	.70	.92	781
(2,1)	S_S	.88	.42	917
	S_L	1.12	.00	926
	S_T	.92	.12	976
	S_Q	1.40	.00	826
(4,1)	S_S	.96	.24	922
	S_L	1.12	.00	926
	S_T	.97	.00	973
	S_Q	1.40	.00	826

are those which maximize the value of $z = \sum_{j=1}^{10} x_j$. The solutions are obtained using the Sequential Unconstrained Minimization Technique (SUMT) developed by Fiacco and McCormick [29]. The resulting values of z are presented in Table 4.6 along with the corresponding values of ALPHA. These results relate to A, B and C in the following manner.

A'. The sets S_S and S_T are still conservative, but it is now possible to generate a set S_L which can contain points violating the constraint more

Table 4.6

Values of ALPHA and z with n = 10

Run	ALPHA S_S	S_L	S_T	S_Q
1	982	874	984	764
2	988	919	994	1,000
3	984	920	994	913
4	991	774	999	912
5	977	900	989	841
6	979	943	999	918
7	988	883	987	886
8	995	824	999	958
9	995	908	991	968
10	989	882	995	954

Run	z S_S	S_L	S_T	S_Q
1	1.526	1.309	1.494	1.793
2	1.473	1.132	1.391	1.119
3	1.513	1.178	1.411	1.652
4	1.327	1.541	1.312	1.655
5	1.550	1.251	1.458	1.710
6	1.539	1.346	1.348	1.664
7	1.479	1.275	1.463	1.675
8	1.418	1.425	1.324	1.569
9	1.419	1.204	1.441	1.518
10	1.468	1.292	1.412	1.567

than the preassigned level of $1-\alpha = .10$. [This is to be expected, since the level of confidence is only equal to .001 (see Table 3.1).]

B'. The set S_T is still more conservative in the majority of the trials, but not substantially so when compared with the set S_S.

C'. Depending upon the particular sample values drawn, the corresponding set S_Q may or may not contain points on the boundary which violate the constraint more than 10 percent of the time.

The above observations are further supported for the case of $n = 25$, as shown by the results presented for this case in Table 4.7.

In observation D, the boundaries of the sets S_Q and S_T were much more representative of the shape of the true boundary.

The distribution-free boundaries were not nearly as representative, since even before the samples were drawn it was known that the resulting sets S_S and S_L would be circular and linear, respectively. Although this will always be the case for the latter set, it need not be for the former set because the shape of this set can be controlled by the choice of the vector $\underline{d}$. There is an infinite number of choices for the values of the elements in this vector, and there is no way of telling prior to sampling which choice yields a more representative shape of the true

Table 4.7

Values of ALPHA and z with n = 25

Run	ALPHA S_S	S_L	S_T	S_Q
1	995	502	1,000	859
2	972	797	990	930
3	992	440	987	859
4	1,000	795	989	865
5	996	904	999	885
6	998	696	989	832
7	961	777	967	900
8	996	662	982	911
9	993	756	988	931
10	991	883	992	956

Run	z S_S	S_L	S_T	S_Q
1	1.582	2.103	1.411	1.819
2	1.672	1.469	1.577	1.726
3	1.614	2.272	1.594	1.829
4	1.430	1.471	1.616	1.828
5	1.569	1.234	1.477	1.753
6	1.457	1.760	1.605	1.831
7	1.702	1.516	1.652	1.778
8	1.573	1.745	1.535	1.781
9	1.598	1.604	1.583	1.748
10	1.625	1.299	1.480	1.726

boundary. It should be noted, however, that the use of the true mean values has worked exceptionally well. That is to say that the set S'_S generated by the cutting function $\phi = |(A_1, A_2) - (.5, .5)|$ follows the shape of the true boundary of S'. This is illustrated in Figure 4.2. This figure also contains the sets S_S, S and S'_T, where S_S and S are as in Figure 4.2 and S'_T is a set obtained from the Tchebysheff method using actual means and variances. From Figure 4.2 it is seen that while the set S'_S gives a better approximation of the shape of the actual region, it is also more conservative than the set S_S. The set S'_T still remains the most conservative.

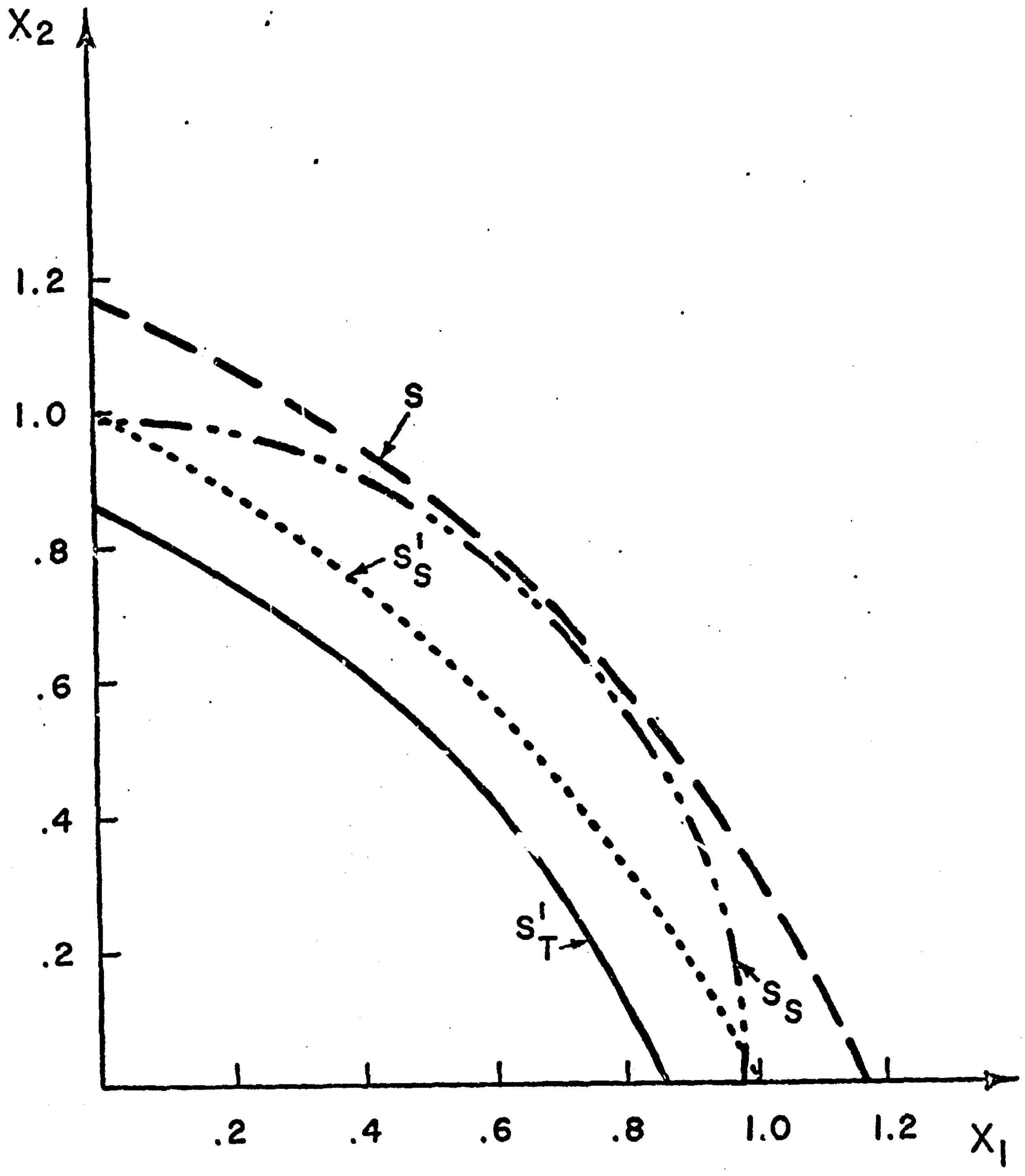


Figure 4.2. Comparison of chance-constrained sets with knowledge of means and variances.

CHAPTER V

CONCLUSIONS AND EXTENSIONS

In this research methods were developed to deal with the chance-constrained set, $S = \{\underline{x} | Pr[\underline{A}\ \underline{x} \leq B] \geq \alpha\}$, when any information concerning the random variables $A_1,\ldots,A_n$ and B must be derived from actual samples. When existing techniques are employed, it is not possible to relate the accuracy of sample information to actual constraint satisfaction. The distribution-free methods which were developed as a result of this research alleviate the problem by providing a lower bound on the confidence β, that one can associate with a value of $\underline{x}$ satisfying the chance-constraint at the preassigned probability level, α. The sample size, N, required to meet the desired confidence is readily available in tabular or graphical form.

Two methods of approximating the set S were developed using the theory of distribution-free tolerance regions. The resulting sets, $S_L(\alpha,\beta)$ and $S_S(\alpha,\beta)$, have the property that any $\underline{x}$ contained in them satisfies the chance-constraint, $Pr[\underline{A}\ \underline{x} \leq B] \geq \alpha$, with levels of confidence β_L and β_S. The advantage of the set $S_L(\alpha,\beta)$ is that it is a linear constraint with exactly the same number of coefficients of the original constraint. Furthermore, the values for these coefficients can be determined directly by inspection of the random samples. The disadvantage of the set $S_L(\alpha,\beta)$ is that for fixed levels of α and β, the required sample size increases rapidly as n, the dimension of $\underline{A} = (A_1,\ldots,A_n)$, increases. The convex set $S_S(\alpha,\beta)$, on the other hand, does not possess this functional relationship between N and n. Another advantage is the flexibility which is provided for choosing the general shape of the resulting distribution-free set.

The superiority of the set $S_S(\alpha,\beta)$ over the sets $S_Q(\alpha)$ and $S_T(\alpha)$ obtained via the Quantile and Tchebysheff Methods was demonstrated using simulated gamma variates. The Quantile Method, with normal variates, is superior since the set $S_Q(\alpha)$ is equivalent to the desired set, S, whereas the sets $S_S(\alpha,\beta)$ and S_T are only small subsets of S. However, when the normality assumption does not hold, it is possible for the set $S_Q(\alpha)$ to contain points which do not satisfy the constraint at the desired level, α, as demonstrated in Chapter IV. Thus, before employing the Quantile Method, it is essential that the normality assumption be carefully checked. If it is found that the underlying distribution is definitely non-normal, then a distribution-free approach should be considered over the Tchebysheff Method for the following reasons.

1. It provides a way of measuring effect of the sample size, N, upon the confidence, β, associated with attainment of the desired probability level, α. With the Tchebysheff Method, it is difficult to decide on an appropriate sample size to estimate the required parameters.

2. The results of Chapter IV indicate that the set $S_S(\alpha,\beta)$ is not as conservative as $S_T(\alpha)$, even for the relatively high level of confidence level of $\beta = 95$. This means that if points in $S_T(\alpha)$ are expected to satisfy the constraint with a probability of at least α, they actually satisfy them at least $100\alpha_T$ percent of the time, where $\alpha_T >> \alpha$. The corresponding value for the set $S_S(\alpha,\beta)$ is closer to the desired level, α. This can be seen in Chapter IV by comparing the value of ALPHA obtained using the above methods.

Although the empirical results of Chapter IV were based upon independently distributed random variates, the procedure for constructing the sets $S_S(\alpha,\beta)$ and $S_L(\alpha,\beta)$ for dependent variates is the same. This is not the case for the set $S_T(\alpha)$, which requires estimates of the covariances. It could be argued that the Tchebysheff Method is superior to a distribution-free method on the grounds that the former is able to take advantage of more information regarding the interdependence of the random variables in question. In real-world situations, however, estimation of covariances is much more difficult than that of means and variances, and the problem of assessing the effect of bad estimates upon the set $S_T(\dot{\alpha})$ is made considerably more difficult.

The simulated random variables in Chapter IV were continuously distributed. Had discrete variates been used, the only deviation from the method constructing distribution-free sets would have arisen in the case of ties; that is, two or more sample points would yield the same maximum value of the particular cutting function employed. In such a case, the ties could be broken using lexicographical ordering rules as discussed in Chapter II. It should be noted that the values of α and β do not depend upon the continuity of the variables in question.

The problem of increasing the size of a distribution-free set was investigated. With the Quantile or Tchebysheff Methods, the size of the chance-constrained set can be expanded by decreasing the level of probability level, α. For a distribution-free method, the same goal can be attained by repeating the construction procedure at lower levels of α and/or β, with reduced sample sizes. If resampling is not possible, then one must work with randomly chosen subsets of the original sample. While this does not guarantee an expanded set, the only alternative is to take additional cuts on the original tolerance region. This is not recommended since the resulting confidence level is dependent upon a fixed but unknown quantity.

There are several possible extensions to the work presented in this paper. Certainly there is a need for more experimentation with distribution-free chance-constrained sets using simulated data from distributions other than gamma. Perhaps an even better insight into the usefulness of these sets could be derived by applying them to real-world linear programming problems with random coefficients.

Further research is needed in determining appropriate values for the elements of the shaping vector $\underline{d}$ for the set $S_S(\alpha,\beta)$. This problem was investigated briefly in Chapter IV, where it was shown that the choice of sample means (as opposed to $\underline{d} = \underline{0}$) resulted in a set $S'(\alpha,\beta)$ whose shape was very close to that of the true chance-constrained set.

The notion of a distribution-free tolerance region might prove to be beneficial in other areas of stochastic linear programming. For example, in distribution problems, the distribution of the optimal objective function value is derived explicitly or by numerical approximation, then decision rules are based on features of the distribution. The alternative distribution-free approach would be to base decision rules on distribution-free tolerance limits.

APPENDIX

PROOF OF THEOREM 3.3

Theorem 3.3

Let U^m be the coverage of the region $T(\alpha,\beta)$ constructed from a sample of size N by removing m blocks. Let $U^{m+m'}$ be the coverage of the region $T'(\alpha',\beta')$ formed by removing m' additional blocks from $T(\alpha,\beta)$. Then

$$\Pr[U^{m+m'} \geq \alpha] = 1 - I_{\alpha'/U^m}(N-m-m'+1,\ m') = \beta'$$

Proof

Let U_1 be the coverage of θ_1[as defined by (2.3) in Chapter II]. Assign zero probability to θ_1 and normalize to unity the portion of the original population contained in $\overline{\theta}_1$. Let U_2' be the conditional coverage of θ_2 given U_1. Continuing in this manner, a sequence of conditional coverages U_1, $U_2',\ldots,U_N'$ is obtained for which the probability element (p.e.) can be shown [12] to be

$$N!(1 - U_1)(1 - U_2')^{N-2} \ldots (1 - U_N')^{N-N}\, dU_1\, dU_2' \ldots dU_N'$$

In particular, the p.e. of the distribution of the conditional coverages $U_{m+1}',\ldots,U_n'$ is

$$(N - m)!(1 - U_{m+1}')^{N-m+1} \ldots (1 - U_N')\, dU_{m+1}' \ldots dU_N'$$

Now the coverages $U_{m+1},\ldots,U_N$ of the blocks $\theta_{m+1},\ldots,\theta_N$ are related to the above conditional coverages in the following manner.

$$U'_{m+1} = U_{m+1}/U^m$$

$$U'_{m+2} = U_{m+2}/(U^m - U_{m+1})$$

$$\vdots \qquad \vdots$$

$$U'_N = U_N/(U^m - U_{m+1} - U_{m+2} - \cdots - U_{N-1})$$

The Jacobian of this transformation is

$$(U^m)^{N-m}\,(1 - U'_{m+1})^{N-m+1} \cdots (1 - U'_{N-1})$$

and the corresponding p.d.f. of the coverages

$$= \begin{cases} \dfrac{(N-m)!}{(U^m)^{N-m}}\, dU_{m+1} \cdots dU_N\,, & \sum\limits_{i=1}^{N-m} U_{m+i} \le U^m \\ 0 & \text{otherwise} \end{cases}$$

Making the transformation $U_{m+i} = U^m V_i$, $i = 1,\ldots,N - m$, the p.d.f. of the random variables, $V_1,\ldots,V_{N-m}$ is

$$= \begin{cases} (N - m)!\, dV_1 \cdots dV_{N-m}\,, & \sum\limits_{i=1}^{N-m} V_i \le 1 \\ 0 & \text{otherwise} \end{cases}$$

which is the (N-m)-variate Dirichlet.

The coverage $U^{m+m'}$ can now be expressed as $U^{m+m'} = U^m \cdot U^{m'}$ where $U^{m'}$ is the sum the m' additional coverages removed from $T(\alpha,\beta)$. It follows that

$$
\begin{aligned}
\Pr[U^{m+m'} \geq \alpha'] &= \Pr[U^{m'} \leq U^{m} - \alpha'] \\
&= \Pr[V^{m'} \leq 1 - (\alpha'/U^{m})] \\
&= I_{1-\alpha'/U^{m}}\,(m',\, N-m-m'+1)
\end{aligned}
$$

LIST OF REFERENCES

1. Kataoka, S., "A Stochastic Programming Model," Econometrica, 31 [1-2], 181-196 (1963).

2. Charnes, A., and W. W. Cooper, "Deterministic Equivalents for Optimizing and Satisficing Under Chance Constraints," Operations Research, 11 [1], 18-39 (1963).

3. Charnes, A., and M. J. L. Kirby, "Some Special P-Models in Chance Constrained Programming," Management Science, 14 [3], 183-195 (November, 1967).

4. Charnes, A., W. W. Cooper and G. H. Symonds, "Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil," Management Science, 4 [3], 235-263 (1958).

5. McQuillan, C. E., "Stochastic Programming: A Survey," Technical Report Title II-C-1228, University of Wisconsin Water Resource Center, March, 1969.

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

7. Vajda, S., "Stochastic Programming," Integer and Nonlinear Programming, American Elsevier Publishing Co., New York, 1970 (ed. by J. Abadie).

8. Sengupta, J. K., "Safety-First Rules Under Chance-Constrained Linear Programming," Operations Research, 17 [1], 112-132 (1969).

9. Sinha, S. M., "Stochastic Programming," Ph.D. Thesis, University of California, Berkeley, 1963.

10. Fréchet, M., Recherches Théoriques Modernes sur la Théorie des Probabilités, Paris, 1937.

11. Braswell, R. N., and F. M. Allen, "Non-Parametric Sensitivity Analysis in Linear Programming," AIIE Transactions, March, 1969, pp. 17-23.

12. Wilks, S. S., Mathematical Statistics, John Wiley and Sons, New York, 1963.

13. Shewhart, W. A., Economic Control of Quality of Manufactured Products, D. Van Nostrand Co., New York City, 1931.

14. Fraser, D., "Non-Parametric Tolerance Regions," Ann. Math. Stat., 24, 44-55 (1953).

15. Somerville, P. N., "Tables for Obtaining Non-Parametric Tolerance Limits," Ann. Math. Stat., 29, 599-601 (1958).

16. Noether, G. E., Elements of Non-Parametric Statistics, John Wiley and Sons, New York, 1967.

17. Bradley, J. V., Distribution-Free Statistical Tests, Prentice-Hall, Englewood Cliffs, New Jersey, 1968.

18. Wilks, S. S., "Determination of Sample Sizes for Setting Tolerance Limits," Ann. Math. Stat., 12, 91-96 (1941).

19. Robbins, H., "On Distribution-Free Tolerance Limits in Random Sampling," Ann. Math. Stat., 15, 214-219 (1944).

20. Pearson, K., Tables of the Incomplete Beta Function, Cambridge University Press, Cambridge, England, 1934.

21. Murphy, R. B., "Non-Parametric Tolerance Limits," Ann. Math. Stat., 19, 581-589 (1948).

22. Scheffe, H., and J. Tukey, "Non-Parametric Estimation I. Validation of Order Statistics," Ann. Math. Stat., 16, 187-192 (1945).

23. Wald, A., "An Extension of Wilk's Method for Setting Tolerance Limits," Ann. Math. Stat., 14, 45-55 (1943).

24. Tukey, J., "Non-Parametric Estimation II. Statistically Equivalent Blocks and Tolerance Regions - The Continuous Case," Ann. Math. Stat., 18, 529-539 (1947).

25. Fraser, D., "Sequentially Determined Statistically Equivalent Blocks," Ann. Math. Stat., 22, 372-381 (1951).

26. Kemperman, J., "Generalized Tolerance Limits," Ann. Math. Stat., 27, 180-186 (1957).

27. Tukey, J. W., "Non-Parametric Estimation III. Statistically Equivalent Blocks and Multivariate Tolerance Regions - The Discontinuous Case," Ann. Math. Stat., 19, 30-39 (1948).

28. Naylor, T. H., J. L. Balintfy, D. S. Burdick and K. Chu, Computer Simulation Techniques, John Wiley and Sons, New York, 1966.

29. Fiacco, A. V., and G. P. McCormick, "Computational Algorithm for the Sequential Unconstrained Minimization Technique for Non-linear Programming," Management Science, 10 [4], 601-617 (1964).