NUREG/CR-1397
SAND80-0020

Risk Methodology for Geologic Disposal of Radioactive Waste: Small Sample Sensitivity Analysis Techniques for Computer Models, With an Application to Risk Assessment

Ronald L. Iman, W. J. Conover, James E. Campbell

Printed March 1980

Sandia National Laboratories

SF 2900-Q(3-80)

Prepared for
U. S. NUCLEAR REGULATORY COMMISSION

Available from

GPO Sales Program
Division of Technical Information and Document Control
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

and

National Technical Information Service
Springfield, Virginia 22161

NUREG/CR-1397
SAND80-0020
GF

RISK METHODOLOGY FOR GEOLOGIC DISPOSAL OF RADIOACTIVE WASTE: SMALL SAMPLE SENSITIVITY ANALYSIS TECHNIQUES FOR COMPUTER MODELS, WITH AN APPLICATION TO RISK ASSESSMENT

Ronald L. Iman
W. J. Conover*
James E. Campbell

Date Published: March 1980

Sandia Laboratories
Albuquerque, New Mexico 87185
operated by
Sandia Corporation
for the
U. S. Department of Energy

Prepared for
Probabilistic Analysis Staff
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
Under Memorandum of Understanding DOE 40-550-75
NRC FIN No. A1192

*College of Business Administration, Texas Tech University, Lubbock, Texas 79409

THIS PAGE LEFT BLANK INTENTIONALLY.

TABLE OF CONTENTS

THIS PAGE LEFT BLANK INTENTIONALLY.

THIS PAGE LEFT BLANK INTENTIONALLY.

ABSTRACT

As modeling efforts expand to a broader spectrum of areas the amount of computer time required to exercise the corresponding computer codes has become quite costly (several hours for a single run is not uncommon). This costly process can be directly tied to the complexity of the modeling and to the large number of input variables (often numbering in the hundreds). Further, the complexity of the modeling (usually involving systems of differential equations) makes the relationships among the input variables not mathematically tractable. In this setting it is desired to perform sensitivity studies of the input-output relationships. Hence, a judicious selection procedure for the choice of values of input variables is required. Latin hypercube sampling has been shown to work well on this type of problem.

However, a variety of situations require that decisions and judgments be made in the face of uncertainty. The source of this uncertainty may be lack of knowledge about probability distributions associated with input variables, or about different hypothesized future conditions, or may be present as a result of different strategies associated with a decision making process. In this paper a generalization of Latin hypercube sampling is given that allows these areas to be investigated without making additional computer runs. In particular it is shown how weights associated with Latin hypercube input vectors may be changed to reflect different probability distribution assumptions on key input variables and yet provide an unbiased estimate of the cumulative distribution function of the output variable. This allows for different distribution assumptions on input variables to be studied without additional computer runs and without fitting a response surface. In addition these same weights can be used in a modified nonparametric Friedman test to compare treatments. Sample size requirements needed to apply the results of the work are also considered. The procedures presented in this paper are illustrated using a model associated with the risk assessment of geologic disposal of radioactive waste.

EXECUTIVE SUMMARY

The nation's energy problems have created a need for modeling various physical phenomena. For example, one type of model seeks to simulate the complicated workings of a nuclear reactor in order to determine the operating conditions that optimize the efficiency of the reactor within acceptable safety standards. Another type of model attempts to recreate the physical environment in the vicinity of a proposed burial site for nuclear waste in order to mimic the behavior of potentially harmful nuclides as they migrate through geologic formations and change chemical form over a long period of time. A third type of model incorporates many economic, social, political, and geographical characteristics of our society in order to examine possible relationships among those variables, in an attempt to measure the environmental impact of various alternative sources of energy.

Computer codes that implement the mathematical models for these and other phenomena are in everyday usage by both government and private industry. These codes have several characteristics in common. They represent serious attempts to include all variables that may be important to the process being modeled and therefore each code usually has many input variables, often numbering in the hundreds. The distribution function of these variables is frequently not well known. In addition, the relationships among the variables are usually complex, modeled only by systems of differential equations which are not mathematically tractable. The combination of many variables and the complex relationships among the variables results in a computer code that often requires several hours of computer time tc make a simulation run for a single input vector. Because of the expense and time involved on the computer, only a limited number of simulation runs is feasible. On the basis of these few runs, numbering sometimes between 50 and 100, a complete analysis of the model is desired. The analysis usually includes, but is not limited to, (1) the estimation of the means, variances, and distribution functions of several output variables, (2) an analysis of the model's sensitivity to the various input variables, and (3) the effect that uncertainty regarding the distribution functions of the input variables has upon inferences pertaining to the output variables.

Extraction of the amount of information indicated in the previous paragraph requires the development of new statistical techniques. Latin hypercube sampling, as introduced by McKay, Conover and Beckman (1979), appears to provide a satisfactory method for selecting input variables so that good estimators of the means, variances, and distribution functions of the output variables may be obtained, providing the answer to part (1) of the desired analysis. The model's sensitivity to the various input variables is then handled by partial rank correlation coefficients as described by McKay, Conover and Whiteman (1976). This procedure satisfies part (2) of the desired analysis. In order to handle part (3) of the desired analysis we have extended the development of Latin hypercube sampling in this paper. The generalization of Latin hypercube sampling is presented in Section 2. Its application to the problem at hand enables the distribution functions of the input variables to be changed from those assumed originally, and, without making any computer runs other than the ones used in the earlier analysis, enables estimates of the means, variances, and distribution functions of the output variables to be made. The details of this procedure are given in Section 3.

Starting with Section 4 this paper is concerned with an example, showing how the methods of Sections 2 and 3 are used in a model which depicts the movement of nuclides through geologic media in the vicinity of an underground depository for nuclear waste. Section 4 illustrates the straightforward application of the procedures outlined in these sections. Comparisons are also made among other sampling procedures such as replicated Latin hypercube sampling and random sampling.

Not all models allow for the straightforward application of this or any other method. One desirable property of this procedure is that it is flexible enough to adapt to unusual situations that may develop. For example, in the model we use in Section 4, the movement of nuclides is influenced by conditions that exist in the vicinity of the burial site. However, these conditions may change unexpectedly at some time in the long range future. Since it is not possible to know precisely what these conditions would become, the best we can do is hypothesize what conditions could reasonably exist (call these conditions "scenarios"), model these scenarios and run the code for these scenarios. The purpose of these calculations would be to order the scenarios with respect to their output random variable. Since the number of scenarios could easily reach several hundred, an efficient technique is required for the ordering. In Section 5, we show how the results of Sections 2 and 3 can be used with changing assumptions of distributions on the input variables to obtain the desired ordering.

This work is part of a project to develop a methodology for the examination of the long-term public risk from radioactive waste repositories in deep geologic formations. This project is being conducted at Sandia Laboratories with funding provided by the Nuclear Regulatory Commission

(NRC) and assists the NRC repository liceusing program. It is anticipated that the methodology developed in this project will be used by the NRC staff in the evaluation of proposed radioactive waste repositories.

1. INTRODUCTION

Evaluation of a waste repository site to verify or deny compliance with regulatory standards will amost certainly involve estimates of the long term risk associated with the waste disposal activity. Thus risk analysis may play an important role in the decision to license waste repositories. Because of the long times which must be considered in waste disposal risk analysis, it is necessary to make extensive use of mathematical models in such an analysis. Some of the physical processes which will be represented by mathematical models include; (1) thermal and mechanical effects induced by interactions between the radioactive waste and the host rock, (2) effects of disruptive features on the groundwater flow system, (3) radionuclide migration in groundwater and (4) radionuclide movement through the surface environment and human uptake. The risk results obtained from the use of such models are subject to considerable uncertainties. These uncertainties arise from two principal sources; (1) uncertainty in the values which serve as input to the models and (2) uncertainty in conditions which may exist in the vicinity of the repository in the long term future. For risk results to be useful in the repository licensing process, these uncertainties must be taken into account.

This report presents statistical techniques to account for uncertainties in three important areas of analysis of waste repository sites. These are; (1) sensitivity analysis of computer models, (2) scenario screening, (3) estimation of risk with uncertainties. Definitions of standard statistical terminology, with which some readers may be unfamiliar, may be found in Conover (1980).

1.1 Sensitivity Analysis of Computer Models

The primary purpose of sensitivity analysis is to determine those model input variables whose uncertainties must be accounted for in risk analysis. Sensitivity analysis can also play an important role in directing research toward those site and radioactive waste properties which contribute most to risk uncertainties.

Computer models used in the analysis of radioactive waste disposal sites are often large and complex. Because these codes represent serious attempts to include all variables that may be important to the process being modeled, each code usually has many input variables, often numbering in the hundreds. The distribution function of these variables is frequently not well known. The combination of many variables and the complex relationships among the variables results in a computer code that may require several hours of computer time to make a simulation run for a single input vector. Because of the expense and time involved on the computer, only a limited number of simulation runs is feasible. On the basis of these few runs, numbering sometimes between 50 and 100, a complete analysis of the model is desired. The analysis usually includes, but is not limited to, (1) the estimation of the means, variances, and distribution functions of several output variables, (2) an analysis of the model's sensitivity to the various input variables, and (3) the effect that uncertainty regarding the distribution functions of the input variables has upon output variable distributions.

Extraction of the amount of information indicated in the previous paragraph requires the development of new statistical techniques. Latin hypercube sampling, as introduced by McKay, Conover and Beckman (1979), appears to provide a satisfactory method for selecting input variables so that good estimators of the means, variances, and distribution functions of the output variables may be obtained, providing the answer to part (1) of the desired analysis. The model's sensitivity to the input variables is then handled by partial rank correlation coefficients as described by McKay, Conover and Whiteman (1976). This procedure satisfies part (2) of the desired analysis. In order to handle part (3) of the desired analysis we have extended the development of Latin hypercube sampling in this paper. The generalization of Latin hypercube sampling is presented in Section 2. Its application to the problem at hand enables the distribution functions of the input variables to be changed from those asumed originally, and, without making any computer runs other than the ones used in the earlier analysis, enables estimates of the means, variances, and distribution functions of the output variables to be made. The details of this procedure are given in Section 3.

Starting with Section 4 this paper is concerned with an example, showing how the methods of Sections 2 and 3 are used in a model which depicts the movement of nuclides through geologic waste. Section 4 illustrates the straightforward application of the procedures outlined in these sections.

1.2 Scenario Screening

The risk from radioactive waste disposal is influenced by conditions which exist in the vicinity of a waste repository. However, these conditions may change at some time in the long range future. As it is not possible to know precisely what these conditions would become, the best we can do is to hypothesize what conditions could reasonably exist (call these conditions "scenarios"), model these scenarios and run the code for these scenarios. As the number of scenarios could easily reach several hundred, an efficient technique is needed for ordering and grouping scenarios in terms of their output variable (some appropriate measure of consequence) so that a smaller number of important scenarios can be examined more extensively. In Section 5 we show how the results of Sections 2 and 3 can be used to obtain the desired ordering.

1.3 Estimation of Risk with Uncertainties

Even though sensitivity analysis may have significantly reduced the original number of input variables, risk analysis will still require sampling from appropriate ranges for a large number of model input variables. Furthermore, despite one's best attempts at scenario screening, several tens of scenarios may have to be included in risk analysis. Thus efficient statistical techniques are required to estimate risk with uncertainties. The methods of this paper may be used in estimating risk with uncertainties, in an efficient manner. There is very little direct discussion of risk assessment in this paper; the emphasis is on statistical methods which are useful in the ultimate goal which is risk assessment. However, Figure 5.5 presents estimated risk assessment curves in an example which uses most of the methods presented in this paper.

This work is part of a project to develop a methodology for the examination of the long-term public risk from radioactive waste repositories in deep geologic formations. This project is being conducted at Sandia Laboratories with funding provided by the Nuclear Regulatory Commission (NRC) and assists the NRC repository licensing program. It is anticipated that the methodology developed in this project will be used by the NRC staff in the evaluation of proposed radioactive waste repositories.

2. A GENERALIZATION OF LATIN HYPERCUBE SAMPLING

The material contained in this section and the next section draws heavily upon results which appeared in Conover (1975) and McKay, Conover and Beckman (1979). In most cases these results represent a generalization of the previous results, so that they may apply more easily to the problem of sensitivity analysis.

2.1 The Rationale

The selection of particular values for the input variables to run in a computer code should be done in such a way as to support the original objectives of the computer code as much as possible. The code is designed to simulate the true physical situation, in order to estimate certain real quantities that cannot be measured directly. A good method of selection of values of input variables should make possible;

(a) probability related statements, such as those regarding the mean, variance, or cumulative distribution function of the output variable,

(b) estimates that are close to the values of the quantities being estimated,

(c) an assessment of the relative importance of each input variable,

(d) some means for measuring the sensitivity of the code output with respect to distribution assumptions on the input variables.

Requirement (a) above is met only if all physically reasonable values of the input (and hence output) variable have some chance, however remote, of occurring. If some region of possible values of input variables is excluded from being selected (as would be true for deterministic selection techniques), then the ability to make probability statements concerning the output may be severely limited.

Requirement (b) states that estimators should be close to the real values of the quantities being estimated. The "closeness" of an estimator is usually measured in terms of its "mean square error." When the estimator is unbiased, the mean square error equals the variance of the estimator. The variance of an estimator is closely related to the method of selecting input variables and the particular code being examined. For codes in which the output variable is a monotonic function of one input variable, stratified sampling of the input variable usually results in a substantial decrease in the variance of estimators of interest over that obtained from random sampling. This is because stratified sampling forces the entire range of the input variable to be represented in the set of input variables. The sampling procedure resembles a numerical integration procedure in which the range of the integration variable is divided into tiny pieces (the strata). The value of the integral is the expected value of the estimator, the item of interest.

The same advantages obtained by stratified sampling of one input variable may be obtained when the model has more than one input variable. When there are several input variables, usually some variables are more influential than others on the output variable. If one input variable dominates, then that variable should be sampled according to a stratified sampling scheme, and the method of choosing values for the other input variables is of little importance. However, it is usually not possible, a priori, to determine the most important input variables. Furthermore, the output variable may be a function of time (t), and one input variable may dominate the output for certain values of t, while another input variable may dominate the output for other values of t. Therefore, it makes sense to use stratified sampling for each of the input variables. Then it doesn't matter which variable or variables are most important; they are all sampled in such a way as to reduce the variance of the estimator if they are important.

Every stratum on one variable must have some possibility of appearing in the code coupled with each stratum on each other variable, or else certain regions of input variables are excluded by design from the code, and probability statements concerning the output may be severely limited. Therefore, a random combination of the different strata of the input variables is required. If there are only two input variables this method of sampling is known in sample surveys as a "Latin square." Because we are using more than two input variables, we call this sampling procedure "Latin hypercube sampling." A more precise description appears in the next subsection.

Requirement (c) states that a good sampling scheme should permit an assessment of the relative importance of each input variable. In the case of linear models, the relative importance of each input variable is usually measured using the partial correlation coefficient. In the codes we are discussing, the relationship is usually not linear, but it is reasonable to assume that the input-output relationship is monotonic in most cases. That is, if all other variables are held constant, the output is usually an increasing (or decreasing) function of each input variable. The output may be an increasing function of some input variables and a decreasing function of others. In such cases, a measure of the monotonicity of the input-output relationship is more meaningful than a measure of its linearity. Rank correlation coefficients provide a good means for measuring monotonicity. As a result, the partial correlation coefficient computed on the ranks of the input and output variables, called the partial rank correlation coefficient, may be used as a measure of the relative importance of each input variable.

Requirement (d), which states that a good method of selecting values of input variables should provide some means for measuring the sensitivity of the code to distribution assumptions made on the input variables, is met very nicely by Latin hypercube sampling. Changes in the assumptions regarding the distribution of the input variables may be assessed without running additional points through the code. The method for doing this is discussed in Section 3.

2.2 A Description of the Latin Hypercube Sampling Procedure

We will represent the vector of input variables as

$$\underset{\sim}{X} = (X_1, \ldots, X_K) \tag{2.1}$$

and let

$$Y = h(\underset{\sim}{X}) \tag{2.2}$$

represent the output variable, where $h(\underset{\sim}{X})$ is a deterministic, but unknown function of the input variables. The sample $\{\underset{\sim}{X}_n\}$, n = 1, ...,N of input variables is selected in the following manner.

The range of each of the K components X_k of $\underset{\sim}{X}$ is partitioned into N intervals $\{I_{k,n}\}$, k = 1,...,K; n = 1,...,N. The probability $p_{k,n}$ of each interval is defined as

$$p_{k,n} = P(X_k \in I_{k,n}). \tag{2.3}$$

If X_k is dependent on $X_1, X_2, \ldots, X_{k-1}$ then the intervals $I_{k,n}$ and the probabilities $p_{k,n}$ for X_k are functions of the intervals and probabilities for $X_1, \ldots, X_{k-1}$. Such a dependency does not affect the results which follow, so we proceed in our discussion as if the input variables were independent.

The set of all Cartesian products of the form

$$I_{1n_1} \times I_{2n_2} \times \ldots \times I_{Kn_K} = S_{\underset{\sim}{n}} \tag{2.4}$$

is a partition of the sample space S of $\underset{\sim}{X}$ into N^K cells of respective probability sizes

$$p_{1n_1} \cdot p_{2n_2} \cdot \ldots \cdot p_{Kn_K} = p_{\underset{\sim}{n}} \tag{2.5}$$

where

$$\underset{\sim}{n} = (n_1, n_2, \ldots, n_K) \tag{2.6}$$

identifies the "location" of each cell.

A Latin hypercube sample of size N is obtained by first selecting N cells and then obtaining one observation from each cell in a manner described as follows. The N cells are identified by the coordinates

$$\begin{aligned} \underset{\sim}{n}_1 &= (n_{11}, n_{12}, \ldots, n_{1K}) \\ \underset{\sim}{n}_2 &= (n_{21}, n_{22}, \ldots, n_{2K}) \\ \ldots &\quad \ldots \\ \underset{\sim}{n}_N &= (n_{N1}, n_{N2}, \ldots, n_{NK}) \end{aligned} \tag{2.7}$$

with the condition that the N subscripts $(n_{1k}, n_{2k}, \ldots, \quad)$ represent a permutation of the integers $(1,2,\ldots,N)$, for each value of k from 1 to K. In this way, we are assured that the entire range (i.e., each interval

$I_{k,n}$) of each input random variable is sampled. Furthermore, we randomize so that every combination of cells, eligible under the above restriction, is equally likely to be obtained. This is accomplished by requiring that each of the K permutations be random (equally likely) permutations, and that they be mutually independent permutations. Once the selection of N cells is made, a random selection procedure is used to obtain an observation within each of the N cells, and these constitute the N inputs $\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_N$ to the code. The "random observation" is one realization of the conditional random variable $\underset{\sim}{X}$, given $\underset{\sim}{X}$ is in the selected cell.

In practice, a Latin hypercube sample may be obtained as follows. The range of each input variable is divided into N intervals, and one observation on the input variable is made in each interval using random sampling within each interval. Thus, there are N observations (by stratified sampling) on each of the K input variables. One of the observations on X_1 is randomly selected (each observation is equally likely to be selected), matched with a randomly selected observation on X_2, and so on through X_K. These collectively constitute $\underset{\sim}{X}_1$. One of the remaining observations on X_1 is then matched at random with one of the remaining observations on X_2, and so on, to get $\underset{\sim}{X}_2$. A similar procedure is followed for $\underset{\sim}{X}_3,\ldots,\underset{\sim}{X}_N$, which exhausts all of the observations and results in a Latin hypercube sample.

2.3 A General Estimator and its Mean

Estimators for quantities such as the mean, other moments, and the distribution function for the output variables may be treated in a unified manner. These estimators are special cases of a general estimator T defined in this section. First T is shown to be an unbiased estimator.

Then the estimators of interest are shown to be special cases of T. After reading Theorem 1, the reader who is interested only in the application of the method may proceed directly to Section 3.

Theorem 1. Let g(Y) be a function of the output variable Y, and consider the statistic

$$T = \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} g(Y_i) \tag{2.8}$$

where, as usual, $Y_i = h(\underset{\sim}{X}_i)$, and $p_{\underset{\sim}{n}_i}$ is the probability associated with the cell from which $\underset{\sim}{X}_i$ was obtained, as indicated by Equation (2.5). Then T is an unbiased estimator of the mean of g(Y). That is

$$E(T) = E[g(Y)]. \tag{2.9}$$

Proof. Denote the density function of $\underset{\sim}{X}$ by $f(\underset{\sim}{x})$. Note that $\underset{\sim}{X}$ doesn't need to be continuous, but for convenience of notation we will assume $\underset{\sim}{X}$ is continuous and has a density. Then the density of the conditional random variable $\underset{\sim}{X}$, given $\underset{\sim}{X}$ is in cell $\underset{\sim}{n}$, is

$$\begin{aligned} f_{\underset{\sim}{n}}(\underset{\sim}{x}) &= p_{\underset{\sim}{n}}^{-1} f(\underset{\sim}{x}) \text{ if } \underset{\sim}{x} \in S_{\underset{\sim}{n}} \\ &= 0 \text{ otherwise} \end{aligned} \tag{2.10}$$

Since the probability of selecting $\underset{\sim}{X}$ from cell $\underset{\sim}{n}$ is $(1/N)^K$, and is the same for all cells, we have

$$E(p_{\underset{\sim}{n}_i} g(Y_i)) = \sum_{\text{all cells } q} E(p_{\underset{\sim}{n}_i} g(Y_i) \mid \underset{\sim}{n}_i \text{ is cell } q) P(\underset{\sim}{n}_i \text{ is cell } q)$$

$$= \sum_{\text{all cells } q} \left\{ \int_{\text{cell } q} p_q g(h(\underset{\sim}{x}))(f(\underset{\sim}{x})/p_q)d\underset{\sim}{x} \right\} (1/N)^K \qquad (2.11)$$

where p_q represents $p_{\underset{\sim}{n}_i}$ given that the coordinates $\underset{\sim}{n}_i$ represent a particular cell, indexed as "q". Continuation gives

$$E(p_{\underset{\sim}{n}_i} g(Y_i)) = \sum_{\text{all cells } q} (1/N)^K \int_{\text{cell } q} g(h(\underset{\sim}{x}))f(\underset{\sim}{x})d\underset{\sim}{x}$$

$$= (1/N)^K \int_S g(h(\underset{\sim}{x}))f(\underset{\sim}{x})d\underset{\sim}{x}$$

$$= (1/N)^K E[g(Y)]. \qquad (2.12)$$

Therefore, we have

$$E(T) = \sum_{i=1}^{N} N^{K-1} E\{p_{\underset{\sim}{n}_i} g(Y_i)\}$$

$$= \sum_{i=1}^{N} N^{K-1}(1/N)^K E[g(Y)] = E[g(Y)]. \qquad (2.13)$$

Theorem 1 states that T is an unbiased estimator for $E[g(Y)]$. If

$$g(Y) = Y, \qquad (2.14)$$

T is an unbiased estimator for the mean $E(Y)$. If

$$g(Y) = Y^r \tag{2.15}$$

then T is an unbiased estimator for the r th moment of Y. If

$$\begin{aligned} g(Y) &= 1 \text{ if } Y \leq c \\ &= 0 \text{ if } Y > c, \end{aligned} \tag{2.16}$$

then T is an unbiased estimator for the distribution function $P(Y \leq c) = G(c)$ of Y at the value c, because of

$$E\{g(Y)\} = 1 \cdot P(Y \leq c) + 0 \cdot P(Y > c) = P(Y \leq c). \tag{2.17}$$

As a result of Equation (2.17), an unbiased estimate of the entire cumulative distribution function of Y is given by the weighted empirical distribution function

$$S(y) = \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} \; u(y - Y_i), \; -\infty < y < \infty, \tag{2.18}$$

where the unitary function is

$$\begin{aligned} u(t) &= 1 \text{ if } t \geq 0 \\ &= 0 \text{ if } t < 0. \end{aligned} \tag{2.19}$$

In other words, proceeding from left to right, at each observed value Y_i increase the function $S(y)$ by an amount $N^{K-1} p_{\underset{\sim}{n}_i}$ where $p_{\underset{\sim}{n}_i}$ is the probability contained in cell number $\underset{\sim}{n}_i$ from which $\underset{\sim}{X}_i$ is obtained. Note that this function always starts at zero $(y \to -\infty)$ but may be greater than or less than 1.0 as y gets large $(y \to -\infty)$.

Another unbiased estimate of G(y) may be obtained by using the reverse approach. That is, consider the fact that

$$G(y) = 1 - P(Y>y) \tag{2.20}$$

and let

$$S^*(y) = \sum_{i=1}^{N} N^{K-1} P_{\underset{\sim}{n}_i} [1 - u(y-Y_i)], \; -\infty < y < \infty, \tag{2.21}$$

be the estimator. Then

$$E(S^*(y)) = P(Y > y) = 1 - G(y) \tag{2.22}$$

(the development parallels the previous one) and

$$E(1 - S^*(y)) = G(y) \tag{2.23}$$

which shows that $1 - S^*(y)$ is an unbiased estimator for G(y), but one which equals 1 for large $y(y \to \infty)$, and may be less than zero for small $y(y \to -\infty)$.

Both of the above estimators for G(y) behave unlike G(y), which is bounded between 0 and 1 inclusive. The only time S(y) or S*(y) is bounded between 0 and 1 is when the sum of the cell probabilities is bounded above by $(1/N)^{K-1}$, such as when all cell probabilities are all equal; i.e.,

$$P_{\underset{\sim}{n}_i} = (1/N)^K \text{ for all } \underset{\sim}{n}_i .$$

For this reason, the user may prefer to use a standardized form of S(y) such as

$$S^{**}(y) = S(y)/S(\infty) \tag{2.24}$$

which is monotonically increasing from 0 to 1, but in general, may not be unbiased for G(y).

2.4 The Variance of the Estimator

The variance of T does not seem to have a simple form, so we will look at the variance of T in several different ways. Each different form for expressing the variance of T is useful in its own way, because each form provides a different view of the advantages and disadvantages of Latin hypercube sampling. Recall that for unbiased estimators, such as T, the variance of T is also the mean square error of T, which should be small if possible. In Latin hypercube sampling applications the variance of T is usually smaller than the variance of estimators arising from other sampling schemes, but this result may be closely related to the monotonicity property of the code, as we shall see later.

The notation becomes cumbersome when looking at the variance, so let us fix some notation as a start. Let $S_{1n_1}, S_{2n_2}, \ldots, S_{Nn_N}$ represent the cells from which $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ are sampled, respectively, and let

$$U = (S_{1\underset{\sim}{n}_1}, S_{2\underset{\sim}{n}_2}, \ldots, S_{N\underset{\sim}{n}_N}) \tag{2.25}$$

represent the ordered N-tuple of these disjoint cells.

How many such ordered N-tuples are there? This number M is obtained easily by counting the number of ways $S_{1\underset{\sim}{n}_1}$ may be obtained, multiplying this by the number of ways in which $S_{2\underset{\sim}{n}_2}$ may be obtained once $S_{1\underset{\sim}{n}_1}$ has been selected, etc. Since cell $S_{1\underset{\sim}{n}_1}$ represents the selection of one interval I_{kn} for each of the K random variables, and in each case there are N intervals to choose from, there are N^K ways of selecting $S_{1\underset{\sim}{n}_1}$. Next, $S_{2\underset{\sim}{n}_2}$ is formed by selecting one of the remaining N-1 intervals for each of the K random variables, so $S_{2\underset{\sim}{n}_2}$ may be selected $(N-1)^K$ ways once $S_{1\underset{\sim}{n}_1}$ has been selected. Continuation of this line of reasoning leads to

$$M = N^K(N-1)^K(N-2)^K \ldots (2)^K(1)^K = (N!)^K \tag{2.26}$$

as the number of ordered N-tuples U. We will index U, and the corresponding cells, with superscripts:

$$U^i = (S^{(i)}_{1\underset{\sim}{n}_1}, S^{(i)}_{2\underset{\sim}{n}_2}, \ldots, S^{(i)}_{N\underset{\sim}{n}_N}), \quad i=1,\ldots,M. \tag{2.27}$$

Each of these N-tuples is equally likely,

$$P(U=U^i) = 1/M. \tag{2.28}$$

Using the well known relationship

$$\mathrm{Var}(X) = E[\mathrm{Var}(X|Y)] + \mathrm{Var}\,[E(X|Y)] \tag{2.29}$$

we have

$$\mathrm{Var}(T) = E[\mathrm{Var}(T|U)] + \mathrm{Var}[E(T|U)]. \tag{2.30}$$

Now, from (2.28) we have

$$E[\mathrm{Var}(T|U)] = \sum_{i=1}^{M} \mathrm{Var}(T|U^i)P(U=U^i)$$

$$= \frac{1}{M}\sum_{i=1}^{M} \mathrm{Var}(T|U^i). \tag{2.31}$$

The conditional random variables $\underset{\sim}{X}_1$ given $S_{1\underset{\sim}{n}_1}$, $\underset{\sim}{X}_2$ given $S_{2\underset{\sim}{n}_2}$, etc., are independent, because the cells are fixed (given), and the only variation is within each cell. Therefore

$$\mathrm{Var}(T|U^i) = \mathrm{Var}[\sum_{j=1}^{N} N^{K-1} p_{\underset{\sim}{n}_j} g(h(\underset{\sim}{X}_j))|\underset{\sim}{X}_j \in S^i_{j\underset{\sim}{n}_j}]$$

$$= \sum_{j=1}^{N} \mathrm{Var}[N^{K-1} p_{\underset{\sim}{n}_j} g(h(\underset{\sim}{X}_j))|\underset{\sim}{X}_j \in S^i_{j\underset{\sim}{n}_j}]. \tag{2.32}$$

The terms in this last summation represent the within cell variance of a function of $\underset{\sim}{X}_j$. Substitution of Equation (2.32) into (2.31) results in summing the within cell variance of a function of $\underset{\sim}{X}$ over all cells in the sample space, where (by the symmetry of the situation) each cell is involved in the same number of terms. There are N terms in (2.32), $(N!)^K$ terms in (2.31), and only N^K different cells, so each cell is included $(N!)^K/N^{K-1}$ times. If we eliminate the duplication of cells Equation (2.31) reduces to

$E[Var(T|U)]$

$$= \sum_{\text{all cells } q} \frac{1}{M} \frac{M}{N^{K-1}} Var[N^{K-1} p_q g(h(\underset{\sim}{X})) | \underset{\sim}{X} \in \text{cell } q]$$

$$= \sum_{\text{all cells}} N^{K-1} p_q^2 \, Var[g(h(\underset{\sim}{X})) | \underset{\sim}{X} \in \text{cell } q] \tag{2.33}$$

where, as before, p_q refers to the probability of $\underset{\sim}{X}$ being in cell q. This completes the development of the first term on the right hand side of (2.30), so the variance of T may be written as

$$Var(T) = \sum_{\text{all cells } q} N^{K-1} p_q^2 \; Var(g(h(\underset{\sim}{X})) | \underset{\sim}{X} \in \text{cell } q) + Var\,(E(T|U)). \tag{2.34}$$

2.5 An Illustrative Example

As an example to illustrate how Equation (2.34) may be used let $g(Y) = Y$, so T is an estimator of the mean of Y. Suppose the true model is

$$Y = h(\underset{\sim}{X}) = \sum_{k=1}^{K} a_k X_k \tag{2.35}$$

where $\{a_k\}_{k=1}^{K}$ are some constants, and where $X_1, \ldots, X_k$ are independent random variables, each uniformly distributed on the unit interval (0,1). ...o, suppose the intervals $I_{k,n}$ are of equal width so $I_{k,n} = (\frac{k-1}{N}, \frac{k}{N})$ and

$$p_q = (\tfrac{1}{N})^K \text{ for each cell } q. \tag{2.36}$$

The next factor in Equation (2.34) needing evaluation is

$\mathrm{Var}[g(h(\underset{\sim}{X}))|\underset{\sim}{X} \in \text{cell } q]$ which equals

$$\mathrm{Var}\left[\sum_{k=1}^{K} a_k X_k | \underset{\sim}{X} \in \text{cell } q\right] = \sum_{k=1}^{K} a_k^2 \mathrm{Var}[X_k | X_k \in I_{k,n}] \tag{2.37}$$

because when the cell is fixed, the individual components of $\underset{\sim}{X}$ are independent of one another. Since the conditional random variable $[X_k \mid X_k \in I_{k,n}]$ is uniform over an interval of length 1/N, it has variance $(12N^2)^{-1}$ and Equation (2.37) becomes

$$\mathrm{Var}[g(h(\underset{\sim}{X}))|\underset{\sim}{X} \in \text{cell } q] = \frac{1}{12N^2} \sum_{i=1}^{K} a_k^2 \tag{2.38}$$

which is the same for each cell.

Finally, $E(T|U)$ is a constant as the following development reveals

$$E(T|U) = E\left(\sum_{i=1}^{N} N^{K-1} \left(\frac{1}{N}\right)^K \left(\sum_{k=1}^{K} a_k X_{ki}\right) \Big| U\right)$$

$$= \frac{1}{N} \sum_{k=1}^{K} a_k \sum_{i=1}^{N} E(X_{ki}|U). \tag{2.39}$$

The last summation above includes the sum of the means of each partitioned piece of X_k, so it doesn't matter in which order the sum is taken. That is, the sum is independent of U, the particular cells involved. This implies Equation (2.39) is constant over U, and hence the variance of $E(T|U)$ is zero.

In retrospect, for any additive model

$$Y = h(\underset{\sim}{X}) = \sum_{k=1}^{K} f_k(X_k) \tag{2.40}$$

with equi-probable cells, and with $g(Y) = Y$ we have

$$E(T|U) = E(\sum_{i=1}^{N} N^{K-1} (\frac{1}{N})^K \sum_{i=1}^{K} f_k(X_{ki})|U)$$

$$= \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N} E(f_k(X_{ki})|U) \quad (2.41)$$

which is constant over U, since the same values of $E(f_k(X_{ki}))$ are being added for each choice of U; the different values of U merely change the order of addition, so

$$\text{Var}[E(T \mid U)] = 0 \quad (2.42)$$

holds for all additive models with equi-probable cells.

Substitution of Equations (2.42), (2.38) and (2.36) into (2.34) furnishes

$$\text{Var } (T) = \sum_{\text{all cells } q} N^{K-1} (\frac{1}{N})^{2K} \frac{1}{12N^2} \sum_{k=1}^{K} a_k^2$$

$$= \frac{1}{12N^3} \sum_{k=1}^{K} a_k^2 \quad (2.43)$$

as the variance of the estimator of the mean, for the model (2.35) with the stipulated conditions. Note that this ic a factor of N^{-2} smaller than the variance of the sample mean using simple random sampling, a substantial improvement. Also, note that if all $a_k \geq 0$, the largest possible value of T under Latin hypercube sampling is

$$T_{max} = \frac{1}{N} \sum_{j=1}^{N} \sum_{\kappa=1}^{K} a_k (\frac{j}{N}) = \frac{N+1}{2N} \sum_{k=1}^{K} a_k \quad (2.44)$$

and the smallest possible value is

$$T_{min} = \frac{1}{N} \sum_{j=1}^{K} \sum_{k=1}^{K} a_k \left(\frac{j-1}{N}\right) = \frac{N-1}{2N} \sum_{k=1}^{K} a_k . \tag{2.45}$$

Therefore, the absolute inequality

$$T_{min} < T < T_{max} \tag{2.46}$$

leads to the other absolute inequalities

$$\frac{N-1}{2N} \sum_{k=1}^{K} a_k < T < \frac{N+1}{2N} \sum_{k=1}^{K} a_k \tag{2.47}$$

$$\frac{N}{N+1} T < \frac{1}{2} \sum_{k=1}^{K} a_k < \frac{N}{N-1} T \tag{2.48}$$

and finally the 100 percent confidence interval for the mean of Y

$$\frac{N}{N+1} T < E(Y) < \frac{N}{N-1} T \tag{2.49}$$

because the mean of Y is easily seen from Equation (2.35) to equal (1/2) $\sum a_k$. Note that these results are independent of the constants $\{a_k\}$ involved as long as they are nonnegative. Also note that these results pertain only to the model and assumptions stated in Equation (2.35) and in that vicinity.

As another way of looking at the additive model, consider

$$T = \sum_{i=1}^{N} N^{k-1} p_{\underset{\sim}{n}_i} \, g(h(\underset{\sim}{X}_i)) \tag{2.50}$$

where

$$g(h(\underset{\sim}{X}_i)) = \sum_{k=1}^{K} f_k(X_{ki}) \qquad (2.51)$$

for arbitrary functions f_k. Then if

$$p_{\underset{\sim}{n}_i} = (1/N)^K \qquad (2.52)$$

we have

$$T = \sum_{i=1}^{N} N^{-1} \sum_{k=1}^{K} f_k\ (X_{ki}) = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N} f_k(X_{ki}). \qquad (2.53)$$

Again the function $f_k(X_k)$ is evaluated once in each interval I_{kn} of X_k and summed, so the order in which the random variables are evaluated is irrelevant. That is, T is independent of the particular cells selected in the Latin hypercube sample. Thus

$$\text{Var}\ (T) = \frac{1}{N^2} \sum_{k=1}^{K} \sum_{i=1}^{N} \text{Var}[f_k(X_{ki})]$$

$$= \frac{K}{N}\ [\text{average "within interval" variance of } f_k\ (X) \text{ over all NK intervals}] \qquad (2.54)$$

when the random variables $X_1,\ldots,X_k$ are independent. Equation (2.43) occurs as a special case of (2.54), for $f_k(X) = a_k X$.

The above results may be summarized as follows.

Result 1. As a general result of Latin hypercube sampling when T is given by

$$T = \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} \, g(h(\underset{\sim}{X}_i)) \tag{2.55}$$

we have

$$\mathrm{Var}(T) = \sum_{\text{all cells } q} N^{K-1} p_q^2 \, \mathrm{Var}[g(h(\underset{\sim}{X}) | \underset{\sim}{X} \in \text{cell } q] + \mathrm{Var}[E(T|U)] \tag{2.56}$$

where U represents an ordered random selection of N cells having no cell coordinates in common.

Result 2. Further, if all cells are equi-probable

$$p_q = (1/N)^K, \tag{2.57}$$

if $g(h(\underset{\sim}{X}_i))$ is an additive model

$$Y_i = g(h(\underset{\sim}{X}_i)) = \sum_{k=1}^{K} f_k(X_{ki}) \tag{2.58}$$

for arbitrary functions $f_k(X)$, and if $X_1, \ldots, X_K$ are mutually independent, then

$$\mathrm{Var}(T) = \frac{1}{N^2} \sum_{k=1}^{K} \sum_{i=1}^{N} \mathrm{Var}[f_k(X_{ki})] \tag{2.59}$$

where X_{ki} is the conditional random variable X_k given X_k is in interval I_{ki}, $i = 1,\ldots,N$.

Result 3. Further, if

$$f_k(X) = a_k X \qquad (2.60)$$

and the X_i have a standard uniform distribution, then

$$\text{Var}(T) = \frac{1}{12N^3} \sum_{k=1}^{K} a_k^2 \qquad (2.61)$$

and if all $a_k \geq 0$

$$\frac{N}{N+1} T < E[g(h(\underset{\sim}{X}))] < \frac{N}{N-1} \qquad (2.62)$$

with probability 1.

Next we will prove the following.

Result 4. As a general result of Latin hypercube sampling, when T is given by Equation (2.55),

$$\text{Var}(T) = N^{K-1} \sum_{\text{all cells } q} p_q \int_{\text{cell } q} [g(h(\underset{\sim}{x}))]^2 f(\underset{\sim}{x})d\underset{\sim}{x} - [E\{g(h(\underset{\sim}{X}))\}]^2 +$$

$$+ \left(\frac{N}{N-1}\right)^{K-1} \sum_{(q,r)} \sum_{\varepsilon R} \int_{\text{cell } q} g(h(\underset{\sim}{x}))f(\underset{\sim}{x})d\underset{\sim}{x} \int_{\text{cell } r} g(h(\underset{\sim}{x}))f(\underset{\sim}{x})d\underset{\sim}{x} \tag{2.63}$$

where p_q represents the probability size of cell q, and where R represents the restricted space of pairs of cells (q,r) which have no cell coordinates in common (i.e., all pairs of cells that could occur in the same Latin hypercube sample).

As a special case of Result 4 we have the following result of McKay, Conover, and Beckman (1979).

Result 5. If all cells are equi-probable, (2.63) reduces to

$$\mathrm{Var}(T) = \frac{1}{N} \mathrm{Var}\ [g(h(\underset{\sim}{X}))] + \frac{1}{N^{K+1}(N-1)^{K-1}} \sum_{(q,r)} \sum_{\varepsilon R} (\mu_q - \mu)(\mu_r - \mu) \tag{2.64}$$

where

$$\mu_q = N^K \int_{\text{cell } q} (h(\underset{\sim}{x}))f(\underset{\sim}{x})d\underset{\sim}{x} \tag{2.65}$$

and

$$\mu = \int_S g(h(\underset{\sim}{x}))f(\underset{\sim}{x})d\underset{\sim}{x} = E\ \{f(h(\underset{\sim}{X}))\} \tag{2.66}$$

represent the conditional cell means and the overall mean, respectively.

To derive Result 4 we use a fresh start,

$$V(T) = \mathrm{Var}\{ \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i}\ g(h(\underset{\sim}{X}_i))\}$$

$$= \sum_{i=1}^{N} \mathrm{Var}[N^{K-1} p_{\underset{\sim}{n}_i}\ g(h(\underset{\sim}{X}_i))] +$$

$$+ \sum_{\substack{i=1\\ i\neq j}}^{N} \sum_{j=1}^{N} \text{Cov}[N^{K-1} p_{\underset{\sim}{n}_i} g(h(\underset{\sim}{X}_i)), N^{K-1} p_{\underset{\sim}{n}_j} g(h(\underset{\sim}{X}_j))]$$

$$= N \, \text{Var}[N^{K-1} p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1))]$$

$$+ N(N-1)\text{Cov}[N^{K-1} p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1)), N^{K-1} p_{\underset{\sim}{n}_2} g(h(\underset{\sim}{X}_2))] \qquad (2.67)$$

because of the symmetry of the problem (i.e., the distribution of $\underset{\sim}{X}_1$ is the same as $\underset{\sim}{X}_2$ when the cells are as yet unspecified, and their joint distribution is the same as any other joint distribution).

First consider the term

$$\text{Var}[N^{K-1} p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1))] = E[N^{2K-2} p_{\underset{\sim}{n}_1} g^2(h(\underset{\sim}{X}_1))]$$

$$- [E(N^{K-1} p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1)))]^2. \qquad (2.68)$$

From (2.12) we already have

$$[E(N^{K-1} p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1)))]^2 = N^{-2}[E(g(h(\underset{\sim}{X})))]^2. \qquad (2.69)$$

In the same fashion that Equation (2.12) was derived, we obtain

$$E[N^{2K-2} p^2_{\underset{\sim}{n}_1} g^2(h(\underset{\sim}{X}_1))] = \sum_{\text{all cells } q} E[N^{2K-2} p^2_q g^2(h(\underset{\sim}{X}_1)) \mid \underset{\sim}{X}_1 \text{ is in cell } q] \cdot P(\underset{\sim}{X}_1 \text{ is in cell } q)$$

$$= \sum_{\text{all cells } q} N^{2K-2} p^2_q \int_{\text{cell } q} g^2(h(\underset{\sim}{x}))(f(\underset{\sim}{x})/p_q)d\underset{\sim}{x} \cdot (1/N)^K$$

$$= \sum_{\text{all cells } q} N^{K-2} p_q \int_{\text{cell } q} g^2(h(\underset{\sim}{x}))f(\underset{\sim}{x})dx. \tag{2.70}$$

Substitution of Equations (2.70) and (2.69) into (2.68) gives

$$N \operatorname{Var}[N^{K-1} p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{x}_1))] = N^{K-1} \sum_{\text{all cells } q} \int_{\text{cell } q} g^2(h(\underset{\sim}{x}))f(\underset{\sim}{x})d(\underset{\sim}{x})$$

$$-N^{-1}[E\{g(h(\underset{\sim}{X}))\}]^2. \tag{2.71}$$

Now consider the term

$$\operatorname{Cov}[N^{K-1} p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1)), N^{K-1} p_{\underset{\sim}{n}_2} g(h(\underset{\sim}{X}_2))]$$

$$= N^{2K-2} E[p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1)) p_{\underset{\sim}{n}_2} g(h(\underset{\sim}{X}_2))] - N^{-2} E[g(h(\underset{\sim}{X}))] \tag{2.72}$$

with the aid of Equation (2.12) again. The first part of (2.72) becomes

$$E[p_{\underset{\sim}{n}_1} g(h(\underset{\sim}{X}_1))\, p_{\underset{\sim}{n}_2} g(h(\underset{\sim}{X}_2))]$$

$$= \sum_{\substack{\text{all}\\ \text{cells } q}} \sum_{\substack{\text{all}\\ \text{cells } r}} E[p_q g(h(\underset{\sim}{X}_1)) p_r g(h(\underset{\sim}{X}_2)) \mid \underset{\sim}{X}_1 \text{ and } \underset{\sim}{X}_2 \text{ are from cells } q \text{ and } r \text{ respectively}] \cdot$$

$$\cdot\, P(X_1 \text{ and } X_2 \text{ are from cells } q \text{ and } r \text{ respectively})$$

$$= \sum_{q} \sum_{r}{}_{(q,r)\,\varepsilon R} \int_{\text{cell } q} p_q g(h(\underset{\sim}{x}_1))(f(\underset{\sim}{x}_1)/p_q d\underset{\sim}{x}_1 \int_{\text{cell } r} p_r g(h(\underset{\sim}{x}_2)) \cdot$$

$$\cdot\, (f(\underset{\sim}{x}_2)/p_r) d\underset{\sim}{x}_2 N^{-K}(N-1)^{-K} \tag{2.73}$$

where R represents the space of all pairs of cells having no cell coordinates in common. Substitution of (2.73) into (2.72), and then (2.72) and (2.71) into (2.67) gives Result 4.

To obtain Result 5, use the notation of Equations (2.65) and (2.66) to rewrite (2.63) as

$$\mathrm{Var}(T) = N^{K-1} \sum_{\text{all cells } q} p_q \int_{\text{cell } q} [g(h(\underset{\sim}{x}))]^2 f(\underset{\sim}{x})d\underset{\sim}{x} - \mu^2 + \left(\frac{N}{N-1}\right)^{K-1}\left(\frac{1}{N}\right)^{2K} \sum_{(q,r)\in R}\sum \mu_q\mu_r. \tag{2.74}$$

The identity

$$\sum_{\text{all cells } q} \mu_q = N^K\mu \tag{2.75}$$

is used to obtain

$$\sum_{(q,r)\in R}\sum (\mu_q-\mu)(\mu_r-\mu) = \sum_{(q,r)}\sum_{R}(\mu_q\mu_r - \mu_q\mu - \mu_r\mu + \mu^2)$$

$$= \sum_{(q,r)\in R}\sum \mu_q\mu_r - 2\mu(N-1)^K \sum_{\text{all cells } q} \mu_q + N^K(N-1)^K\mu^2$$

$$= \sum_{(q,r)\in R}\sum \mu_q\mu_r - N^K(N-1)^K\mu^2 \tag{2.76}$$

which is used in (2.74) to obtain

$$\mathrm{Var}(T) = N^{K-1} \sum_{\text{all cells } q} p_q \int_{\text{cell } q} [g(h(\underset{\sim}{x}))]^2 f(\underset{\sim}{x})d\underset{\sim}{x} - N^{-1}\mu^2 + \frac{1}{N^{K-1}(N-1)^{K-1}} \sum_{(q,r)\in R}\sum (\mu_q-\mu)(\mu_r-\mu) \tag{2.77}$$

Equation (2.77) is an alternative form for the variance of T, as valid as

Equation (2.63) in Result 4, and may be preferable to use in some situations. All one needs to do is substitute $p_q = N^{-K}$ into (2.77) to obtain Result 5, in agreement with a similar result presented in McKay, Conover, and Beckman (1979). If one thinks of sampling a pair (q,r) of cells at random from R, the covariance of the cell means thus obtained is

$$\mathrm{Cov}[\mu_q, \mu_r] = \frac{1}{N^K (N-1)^K} \sum_{(q,r) \in R} \sum (\mu_q - \mu)(\mu_r - \mu) \qquad (2.78)$$

which resembles a major term in Equations (2.77) and (2.64). Thus, it is obvious that the variance of T will be less than the variance obtained through random sampling if and only if the covariance of two cell means, randomly selected from R, is negative, under the condition that the cells have the same probability size N^{-K}. A sufficient condition for the negative covariance in (2.78) is if $g(h(\underset{\sim}{X}))$ is a monotonic function of each of the input variables $X_1, \ldots, X_K$, which is proved in McKay, Conover, and Beckman (1979).

2.6 The Sample Variance

We have already introduced an unbiased estimator of the population mean $Eh(\underset{\sim}{X})$ in the form

$$T = \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} g(h(\underset{\sim}{X}_i)) = \bar{Y} \qquad (2.79)$$

which we will call the sample mean. The sample variance

$$S^2 = \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} (h(\underset{\sim}{X}_i) - \bar{Y})^2 \qquad (2.80)$$

may be used to estimate the population variance Var(h(X)). However, S^2 is not in the form of a T estimator as introduced earlier, so special consideration of S^2 is required.

The bias of S^2 depends on the population distribution and the particular cells obtainable under Latin hypercube sampling. That is,

$$E(S^2) = E[\sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} (h^2(\underset{\sim}{X}_i) - 2\bar{Y}h(\underset{\sim}{X}_i) + \bar{Y}^2)]$$

$$= E[\sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} h^2(\underset{\sim}{X}_i)] - 2E[\bar{Y} \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} h(\underset{\sim}{X}_i)]$$

$$+ E[\bar{Y}^2 \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i}]. \tag{2.81}$$

From the unbiased property of T statistics the first expectation is

$$E[\sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} h^2(\underset{\sim}{X}_i)] = E[h^2(\underset{\sim}{X})] = \text{Var}\ [h(\underset{\sim}{X})] + \mu^2 \tag{2.82}$$

The second expectation becomes

$$-2E[\bar{Y}^2] = -2\ \text{Var}\ (\bar{Y}) - 2\mu^2 \tag{2.83}$$

because $E(\bar{Y}) = E(Y) = \mu$. Thus (2.81) becomes

$$E(S^2) = \text{Var}[h(\underset{\sim}{X})] - 2\ \text{Var}(\bar{Y}) + E[\bar{Y}^2 \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i}] - \mu^2 \tag{2.84}$$

The latter terms in (2.84) represent the bias of S^2 as an estimator of Var

$[h(\underset{\sim}{X})]$, the population variance. Note that if all cells have equal size $p_{\underset{\sim}{n}} = (1/N)^K$, Equation (2.84) reduces to

$$E(S^2) = Var[h(\underset{\sim}{X})] - Var[Y] \tag{2.85}$$

in the same form as occurs when random sampling is used. Of course in random sampling Var $(\bar{Y})$ = Var(Y)/N, while here Var$(\bar{Y})$ may be larger or smaller than Var (Y)/N depending on the type of function $h(\underset{\sim}{X})$ involved, as was discussed to some extent in the previous subsection. One purpose of using Latin hypercube sampling is to reduce the variance of estimators such as $\bar{Y}$. If this goal is achieved, the bias of S^2 will be very small. As an illustration, for the linear model with independent, standard uniform input random variables, described in Result 3 of the previous subsection, we obtain

$$E(S^2) = \frac{1}{12} \sum_{k=1}^{K} a_k^2 - \frac{1}{12N^3} \sum_{k=1}^{K} a_k^2$$

$$= (1 - \frac{1}{N^3}) \, Var \sum_{k=1}^{K} a_k X_{ki} \tag{2.86}$$

which has negligible bias for moderate sized N.

3. CHANGING THE DISTRIBUTION FUNCTION OF THE INPUT VARIABLES

One of the features of the Latin hypercube sampling procedure is that it allows one to measure the sensitivity of the code to some of the assumptions of the model behind the code. In particular, after a set of runs is completed under certain distribution assumptions on the input variables the assumed distribution functions may be altered, and an estimate of the corresponding change in the distribution of the output variable may be made without making additional runs on the code. This latter feature is important when the code is no longer available or when it is very costly and time consuming to operate. If the distribution function of the output variable is altered significantly by adjustments in the input distributions, the output may be considered to be sensitive to assumptions regarding the input distributions. The same sort of sensitivity analysis may be performed on the mean of the output, the moments of the output, or any other quantities that can be estimated by statistics such as the ones introduced in Section 2.

3.1 The New Estimator

In Latin hypercube sampling the range of each input variable is divided into N intervals $I_{k,n}$, and the Cartesian product of these intervals results in N^K cells of size

$$p_{\underset{\sim}{n}} = P(\underset{\sim}{X} \in S_{\underset{\sim}{n}}). \tag{3.1}$$

If the density of $\underset{\sim}{X}$ is $f(\underset{\sim}{x})$, the conditional density of $\underset{\sim}{X}$, given $\underset{\sim}{X}$ is in cell $S_{\underset{\sim}{n}}$, is

$$\begin{aligned} f_{\underset{\sim}{n}}(\underset{\sim}{x}) &= p_{\underset{\sim}{n}}^{-1} f(\underset{\sim}{x}) \text{ if } \underset{\sim}{x} \in S_{\underset{\sim}{n}} \\ &= 0 \text{ otherwise} \end{aligned} \tag{3.2}$$

as stated earlier in Equation (2.10). A random selection of N cells is obtained and $\underset{\sim}{X}$ is sampled randomly, according to the distribution (3.2), within each of N cells. An unbiased estimate of some function $g(h(\underset{\sim}{X}))$ of the output $h(\underset{\sim}{X})$ is given by the T estimator, which was given earlier in Equation (2.8) and is repeated here for the convenience of the reader,

$$T = \sum_{i=1}^{N} N^{K-1} p_{\underset{\sim}{n}_i} \, g(h(\underset{\sim}{X}_i)). \qquad (3.3)$$

Now let us assume that the true distribution of $\underset{\sim}{X}$ is not $f(\underset{\sim}{x})$, but some other distribution, say $q(\underset{\sim}{x})$. Then the cells described above don't have probability $p_{\underset{\sim}{n}}$, but rather some other value, say $q_{\underset{\sim}{n}}$, given by

$$q_{\underset{\sim}{n}} = P(\underset{\sim}{X} \in S_{\underset{\sim}{n}} | q(\underset{\sim}{x})). \qquad (3.4)$$

The new estimator of $g(h(\underset{\sim}{X}))$, given by

$$Q = \sum_{i=1}^{N} N^{K-1} q_{\underset{\sim}{n}_i} \, g(h(\underset{\sim}{X}_i)) \qquad (3.5)$$

differs from T only in the factor $q_{\underset{\sim}{n}}$ instead of $p_{\underset{\sim}{n}}$. If $\underset{\sim}{X}$ had been sampled from each cell as if it had the density $g(\underset{\sim}{x})/q_{\underset{\sim}{n}}$, instead of the density $f(\underset{\sim}{x})/p_{\underset{\sim}{n}}$ which was actually used, then Q would be an unbiased estimator of $g(h(\underset{\sim}{X}))$. This is because the situation would be exactly the same as described in the previous section, except for a change in notation.

However, the sampling within each cell was done as if the density were $f(\underset{\sim}{x})/p_{\underset{\sim}{n}}$ instead of $q(\underset{\sim}{x})/q_{\underset{\sim}{n}}$, so the statistic Q is not necessarily an unbiased estimator of $g(h(\underset{\sim}{X}))$. The bias under usual circumstances may be assumed to be small, however, for the following reason. Each cell is likely to be only a small part of the sample space, since there are N^K cells in the sample space. That is, if the cells were originally chosen to be of comparable dimensions with probabilities within an order of magnitude of each other, and if the new probabilities $q_{\underset{\sim}{n}}$ are reasonably close to each other, then each cell represents a small portion of the sample space, in size and in probability content. Furthermore, if the densities $f(\underset{\sim}{x})$ and $q(\underset{\sim}{x})$ are reasonably smooth over the sample space, there will be very little change in $f(\underset{\sim}{x})$ or $q(\underset{\sim}{x})$ within any one cell. That is, the maximum value of $f(\underset{\sim}{x})$ within any particular cell will be approximately the same as the minimum value of $f(\underset{\sim}{x})$ within that same cell, and the same can be said about $q(\underset{\sim}{x})$. So for all practical purposes, the sampling within each cell may be conducted as if the distribution across that cell were uniform, with little effect on the sampling results. The assumption of $q(\underset{\sim}{x})$ rather than $f(\underset{\sim}{x})$ does not affect the sampling within cells very much if $q(\underset{\sim}{x})$ and $f(\underset{\sim}{x})$ have approximately the same range space S and are fairly smooth functions. Thus, the bias in Q as an estimator of $g(h(\underset{\sim}{X}))$ should be small.

The estimator Q is treated in the same manner as T was treated in the previous section. That is if $g(Y) = Y$ as in Equation (2.14), then Q is an estimator of the new mean of Y. If $g(Y) = Y^r$ as in Equation (2.15), Q is an estimator of the r^{th} moment of Y. Finally, an estimator of the new distribution function of the output random variable Y is given by $S'(y)$, analogous to the estimator $S(y)$ of Equation (2.18)

$$S'(y) = \sum_{i=1}^{N} N^{K-1} q_{\underset{\sim}{n}_i} \, u(y - Y_i), \; -\infty < y < \infty \qquad (3.6)$$

where u(t) is the unitary function defined by Equation (2.19). These estimators are essentially unbiased, depending on how well the density $f(\underset{\sim}{x})/p_r$ approximates the density $q(\underset{\sim}{x})/q_r$ as discussed above.

To see what is happening mathematically, consider, as we did in Section 2, Equation (2.11),

$$E(q_{\underset{\sim}{n}_i}\, g(Y_i)) = \sum_{\text{all cells } r} E(q_{\underset{\sim}{n}_i}\, g(Y_i) \mid \underset{\sim}{n}_i \text{ is cell } r) \cdot$$

$$\cdot\, P(\underset{\sim}{n}_i \text{ is cell } r)$$

$$= \sum_{\text{all cells } r} \int_{\text{cell } r} q_r g(h(\underset{\sim}{x}))(f(\underset{\sim}{x})/p_r)d\underset{\sim}{x}(1/N)^K \qquad (3.7)$$

because the sampling within each cell was performed under the earlier assumption that $f(\underset{\sim}{x})$ was the density. If we assume the approximation

$$f(\underset{\sim}{x})/p_r \doteq q(\underset{\sim}{x})/q_r \qquad (3.8)$$

holds for each cell, then Equation (3.7) becomes

$$E(q_{\underset{\sim}{n}_i}\, g(Y_i)) \doteq \sum_{\text{all cells } r} N^{-K} \int_{\text{cell } r} q_r g(h(\underset{\sim}{x}))(q(\underset{\sim}{x})/q_r)d\underset{\sim}{x}$$

$$\doteq \sum_{\text{all cells } r} N^{-K} \int_{\text{cell } r} g(h(\underset{\sim}{x}))q(\underset{\sim}{x})d\underset{\sim}{x}$$

$$\doteq N^{-K} \int_{S} g(h(\underset{\sim}{x}))q(\underset{\sim}{x})d\underset{\sim}{x}$$

$$\doteq N^{-K} E[g(h(\underset{\sim}{X})) \mid q(\underset{\sim}{x})]. \qquad (3.9)$$

Therefore Q is an approximately unbiased estimator,

$$E(Q) \doteq \sum_{i=1}^{N} N^{K-1} E(q_{\underset{\sim}{n}_i}\, g(h(\underset{\sim}{X}_i)))$$

$$\doteq \sum_{i=1}^{N} N^{K-1} N^{-K} E[g(h(\underset{\sim}{X}))\,|\,q(\underset{\sim}{x})]$$

$$\doteq E[g(h(\underset{\sim}{X}))\,|\,q(\underset{\sim}{x})] \tag{3.10}$$

under the assumption (3.8).

The conclusion we can draw from the above is that the same points $g(h(\underset{\sim}{X}_1))$ through $g(h(\underset{\sim}{X}_N))$, obtained from running the code under the assumed distribution $f(\underset{\sim}{x})$, may be used in Q and they will differ little from what one would have obtained if $q(\underset{\sim}{x})$ had been assumed rather than $f(\underset{\sim}{x})$. The estimator Q will not necessarily be the same as the estimator T, because now they are estimating different quantities. The estimator Q may be used to estimate the mean, other moments, and even the entire distribution function of $h(\underset{\sim}{X})$, under the assumption that $\underset{\sim}{X}$ has the density $q(\underset{\sim}{x})$. These quantities being estimated may be much different than the corresponding quantities estimated using T, and under the assumption that $f(\underset{\sim}{x})$ is the density. The variance of Q may be obtained just as before (e.g., Equations (2.56) and (2.63)) with p and $f(\underset{\sim}{x})$ replaced by q and $q(\underset{\sim}{x})$.

3.2 The Linear Model as an Example

To illustrate the procedure involved when the assumption regarding the distribution of the input variables are changed, consider again the linear model given in Equation (2.60). The assumptions there are

$$h(\underset{\sim}{X}) = \sum_{k=1}^{K} a_k X_{ki}$$

$$I_{k,n} = \left(\frac{n-1}{N}, \frac{n}{N}\right) \text{ for all } k \tag{3.11}$$

$$f(\underset{\sim}{x}) = 1 \text{ if } \underset{\sim}{x} \text{ is in the } [0,1] \text{ hypercube}$$

$$= 0 \text{ elsewhere} \qquad (3.12)$$

As a consequence of these assumptions, the cells are of equal size, geometrically speaking and in terms of their probabilities $p = (1/N)^K$. A Latin hypercube sample is generated by obtaining N cells and sampling once within each cell. A uniform distribution is used for sampling within each cell, and the output observations $h(\underset{\sim}{X}_1),\ldots,h(\underset{\sim}{X}_n)$ are obtained.

Now instead of using the assumption that each component is independently uniformly distributed on (0,1), we wish to investigate the model under the assumption that each component has the triangular density

$$q(x) = 2x,\ 0 < x < 1$$

$$= 0 \text{ elsewhere.} \qquad (3.13)$$

There is a substantial difference between the uniform density $f(x) = 1$, $0 < x < 1$, and the triangular density. If we use Q to estimate the mean $E(h(\underset{\sim}{X}))$, how much bias is involved?

The true mean of $h(\underset{\sim}{X})$ is

$$E[h(\underset{\sim}{X})|q(\underset{\sim}{x})] = E[\sum_{k=1}^{K} a_k X_k] = \sum_{k=1}^{K} a_k E[X_k]$$

$$= \frac{2}{3} \sum_{k=1}^{K} a_k \qquad (3.14)$$

because the mean of a random variable with density q(x) is 2/3. Of course, in most cases the model is far too complex to ever know the true mean. We are using this simple model as a test case to see what the bias of Q might be.

To see the bias of Q consider

$$E(Q) = E[\sum_{i=1}^{N} N^{K-1} q_{\underset{\sim}{n}_i} \sum_{k=1}^{K} a_k X_{ki}]$$

$$= \sum_{i=1}^{N} N^{K-1} \sum_{k=1}^{K} a_k E[q_{\underset{\sim}{n}_i} X_{ki}]$$

$$= N^K \sum_{k=1}^{K} a_k E[q_{\underset{\sim}{n}_1} X_{k1}] \qquad (3.15)$$

where this last step is possible because the unconditional distribution of each $q_{\underset{\sim}{n}_i} X_{ki}$ is the same, i = 1,..., N, due to the symmetry of the situation. To find $E[q_{\underset{\sim}{n}_1} X_{k1}]$ consider

$$E[q_{\underset{\sim}{n}_1} X_{k1}] = \sum_{\text{all cells } r} N^{-K} E[q_r X_{k1} | X_{k1} \text{ is cell } r]. \qquad (3.16)$$

Because X_{k1} being in cell r is dependent only on X_{k1} being in the interval $I_{k,j}$ that contributed one dimension to the forming of cell r, and is independent of the other dimensions of cell r, and because q_r is not random when cell r is specified, the latter expected value in Equation (3.16) may be written as

$$E[q_r X_{kl} | X_{kl} \text{ is in cell } r] = \prod_{i=1}^{K} q_{ij_i} E(X_{kl}) \qquad (3.17)$$

where

$$q_{ij} = P(X_i \in (\frac{j-1}{N}, \frac{j}{N}) | q(x)) = \frac{2j-1}{N^2}. \qquad (3.18)$$

Also note that

$$E(X_{kl} | X_{kl} \in (\frac{j-1}{N}, \frac{j}{N}), f(x)) = \frac{2j-1}{2N} \qquad (3.19)$$

which is the conditional mean of X_{kl}, given the interval within which X_{kl} was randomly obtained in the original sampling plan with its uniform distribution. Putting Equations (3.17), (3.18) and (3.19) into (3.16) gives

$$E[q_{\underset{\sim}{n}_1} X_{kl}] = \sum_{\text{all cells } r} N^{-K} \prod_{i=1}^{K} q_{ij_i} E(X_{kl})$$

$$= \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} \cdots \sum_{j_K=1}^{N} N^{-K} \frac{2j_1-1}{2N} \prod_{i=1}^{K} \frac{2j_i-1}{N^2}$$

$$= N^{-K} \sum_{j_1=1}^{N} (\frac{2j_1-1}{2N})(\frac{2j_1-1}{N^2}) \sum_{j_2=1}^{N} (\frac{2j_2-1}{N^2}) \cdots$$

$$\sum_{j_K=1}^{N} (\frac{2j_K-1}{N^2}). \qquad (3.20)$$

Because of the identities

$$\sum_{i=1}^{N} \left(\frac{2i-1}{N^2}\right) = 1 \tag{3.21}$$

and

$$\sum_{i=1}^{N} \left(\frac{2i-1}{2N}\right)\left(\frac{2i-1}{N^2}\right) = \frac{2}{3} - \frac{1}{6N^2} \tag{3.22}$$

(3.20) becomes

$$E[q_{\underset{\sim}{n}_1} X_{k1}] = N^{-K}\left(\frac{2}{3} - \frac{1}{6N^2}\right) \tag{3.23}$$

and Equation (3.15) becomes

$$E(Q) = \left(\frac{2}{3} - \frac{1}{6N^2}\right) \sum_{k=1}^{K} a_k. \tag{3.24}$$

The comparison of E(Q) given above with the true mean of $h(\underset{\sim}{X})$ given by Equation (3.14) shows that for moderate N, the bias of Q is negligible.

The insignificant bias of Q as an estimator does not tell the whole story. Perhaps more important is the variance of Q. Empirical evidence indicates that when the cells have almost equal probability the variances of Q and T tend to be smaller than when the probabilities vary considerably from one cell to another. This has not been verified analytically, but it is easy to imagine what happens when all of the probability is contained in a few of the cells. When those cells are chosen for the sample, the estimates are large because of large weights conveyed by the q's. When those cells are not in the sample the weights are close to zero because the q's are close to zero, and Q will also be close to zero. We are reminded of the admonition given to Hiawatha by his fellow tribesmen in the poem Hiawatha Designs an Experiment by Maurice G. Kendall. Although Hiawatha's lack of bias in shooting his arrows at a target may be nice to brag about,

> "What resulted in the long run;
> Either he must hit the target
> Much more often than at present
> or himself would have to pay for
> all the arrows that he wasted."

Even though Q may be almost unbiased, and the original observations $h(\underset{\sim}{X}_i)$ were obtained on cells that were thought to have about the same probability size, Q may be of little value as an estimator if the new assumed probability sizes of the cells are extremely unbalanced, because of the large variance of Q. This is not to detract in any way from Latin hypercube sampling, for other sampling methods may be no better. A comparison of Latin hypercube sampling with random sampling and with a sequential procedure in Section 4.6 shows that although the variance of the estimator cannot be estimated with Latin hypercube sampling, this procedure still offers the best estimators in the situations examined.

4. AN APPLICATION

In this section an application is presented which demonstrates the methods described in Sections 2 and 3. For this demonstration we consider the potential escape of radionuclides from a depository for radioactive waste and their migration from the subsurface to the surface environment. The dependent variables of interest are the total discharges and peak discharges to the surface environment of several radionuclides. Radionuclide migration calculations are performed at Sandia Laboratories with the Sandia Waste Isolation Flow and Transport (SWIFT) computer program (Dillon, Lantz, and Pahwa, 1978). In its present form SWIFT requires considerable computer time to simulate radionuclide migration in large systems over long periods of time. This extensive computer time limits the total number of runs that can be made to study the environmental impact of radioactive waste disposal in geologic media. Because of the need to obtain a maximum of information (see Section 2.1) from a few simulated observations, the methods of Sections 2 and 3 are appropriate.

For the demonstration of these methods it is necessary to have a computer program which requires little computer time to run, so that the estimates of the mean, variance, and distribution function (see Sections 2 and 3) may be compared with the actual quantities being estimated. That is, the estimates are obtained from a limited number (200 in this case) of runs, while the quantities being estimated are obtained from a large number of additional runs. For this purpose the Network Flow and Radionuclide Transport (NWFT) model was developed at Sandia (Campbell et. al, 1980). This model and the groundwater flow system it is used to simulate are discussed below.

4.1 The Groundwater Flow System and NWFT Model

The reference radioactive waste repository site and its associated groundwater flow system are described by Campbell et. al. (1978). The reference site is entirely hypothetical yet its setting and geologic properties are analogous to several regions in the continental United States. The site is located in a symmetrical upland valley which is drained by a

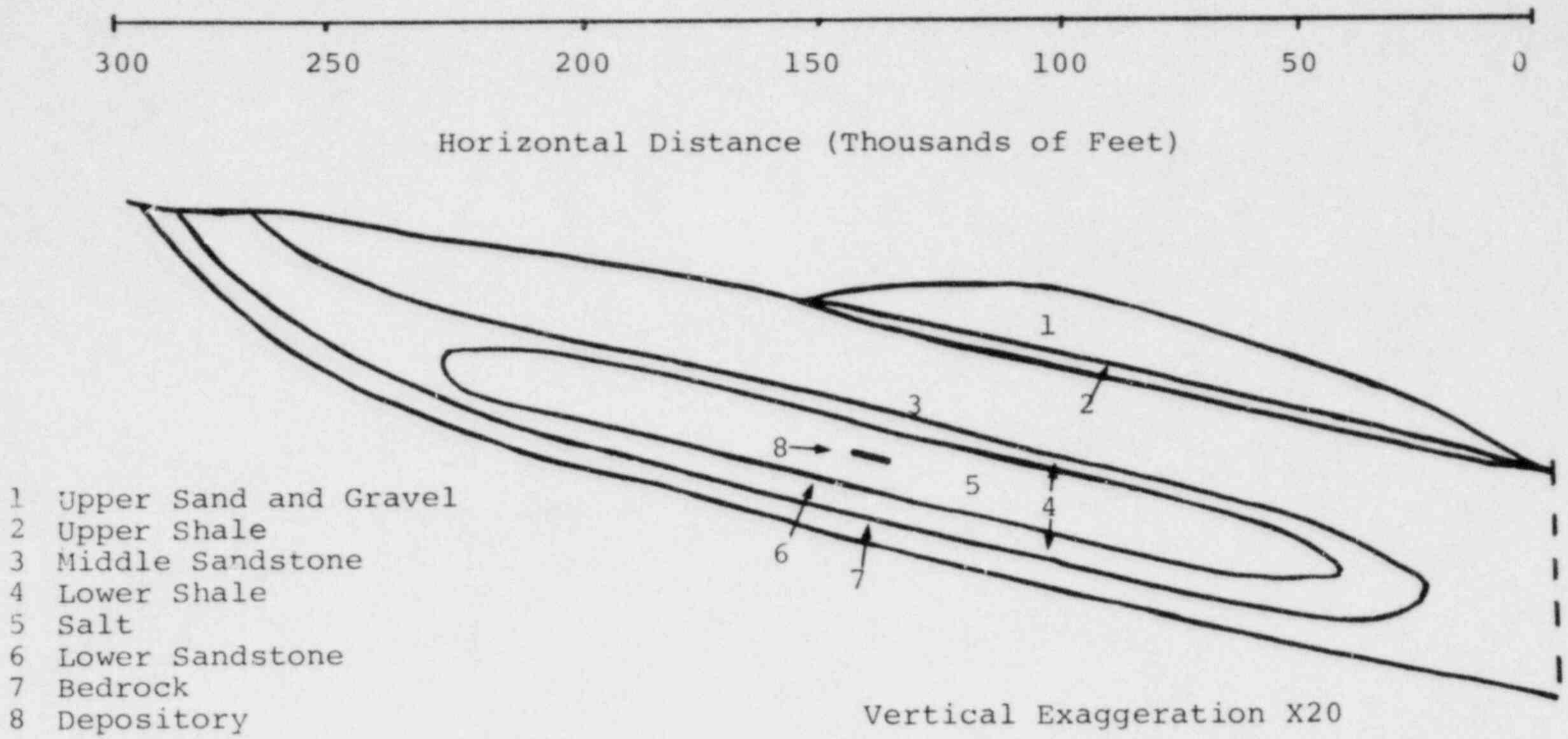


FIGURE 4.1 The Reference Site Geology

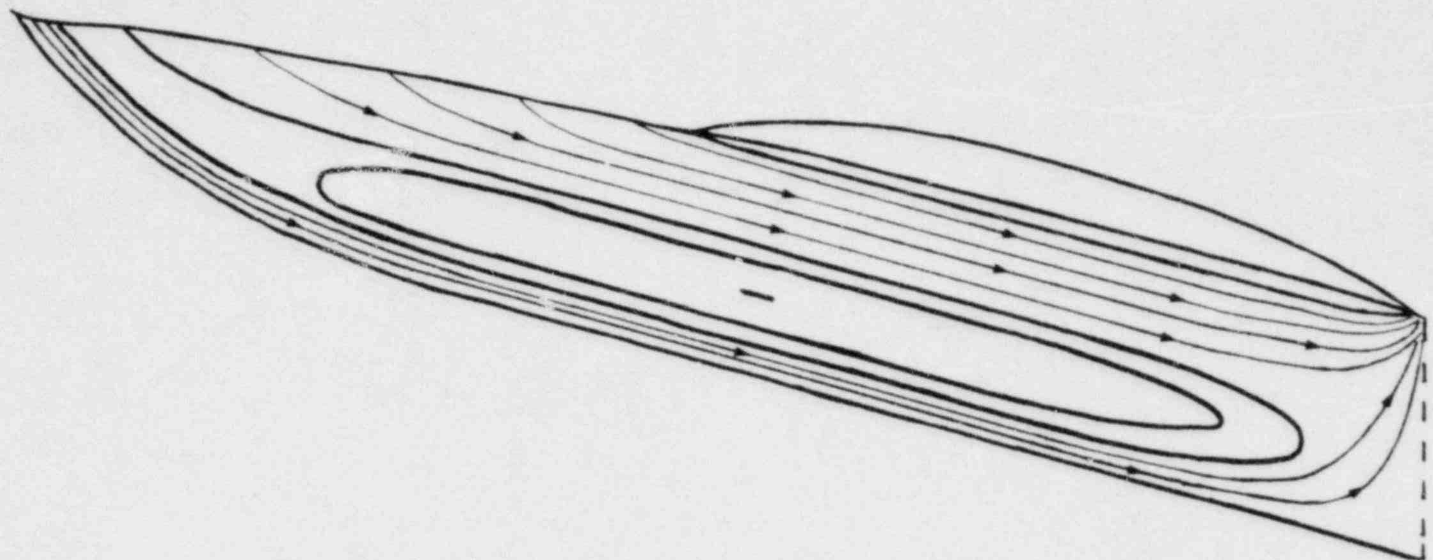


FIGURE 4.2 Flow Lines in Reference System

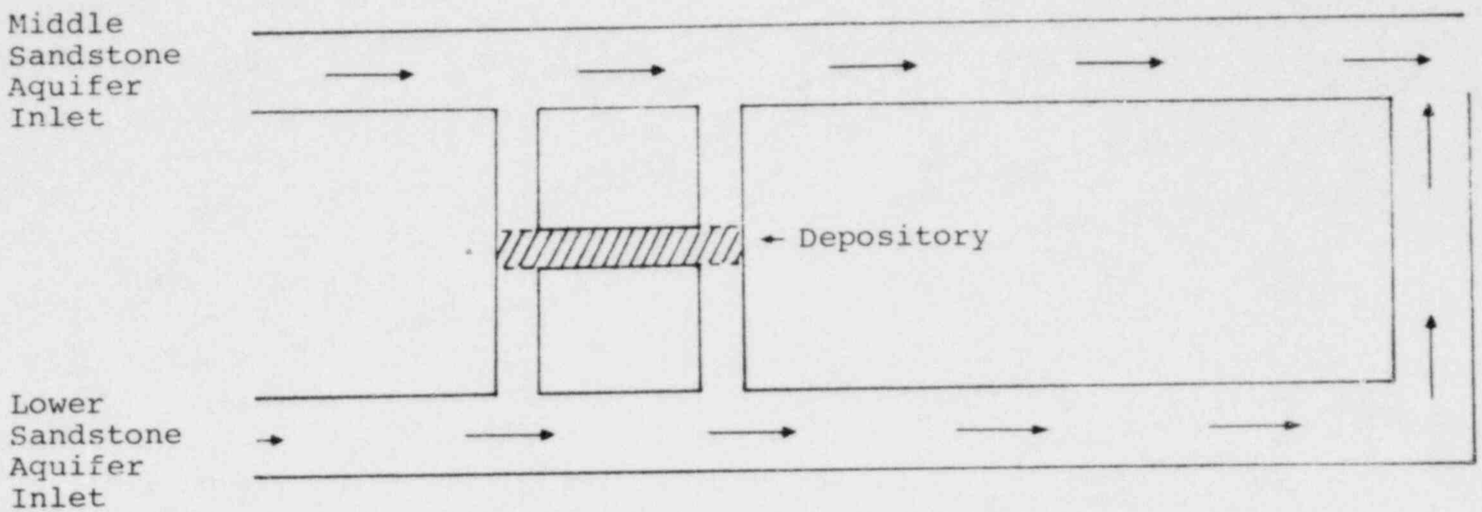


FIGURE 4.3 Network Representation of Reference Site Flow System

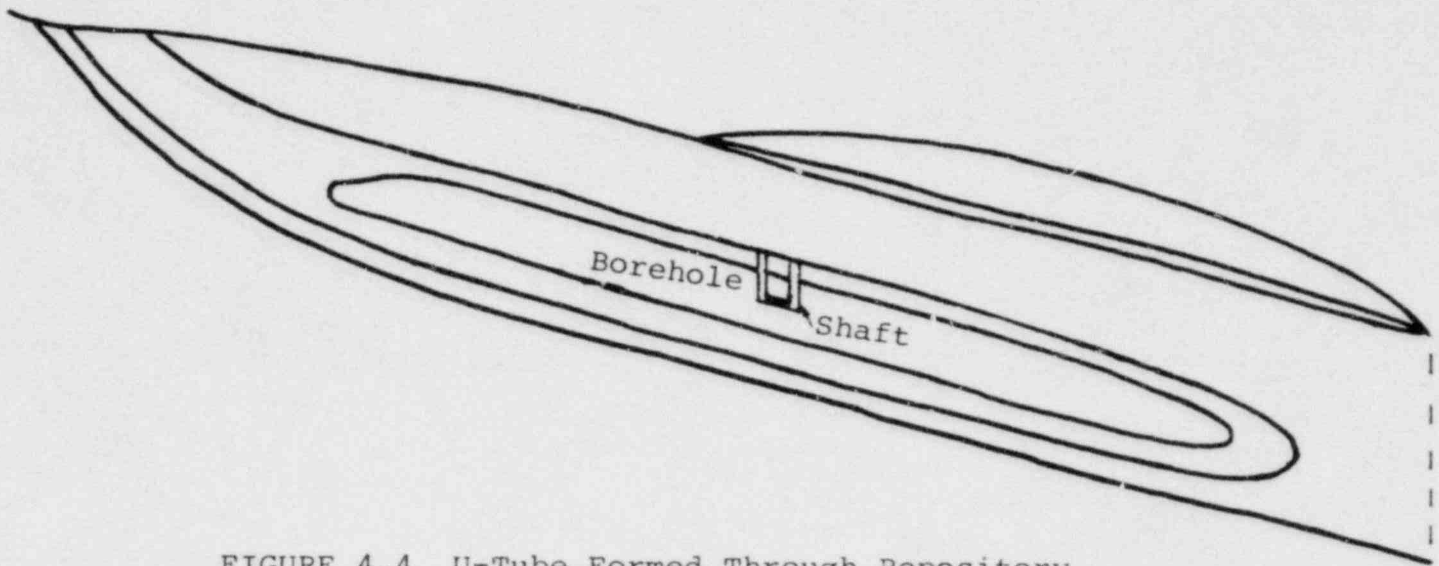


FIGURE 4.4 U-Tube Formed Through Repository

major river. The geology of the area near the site is shown in cross-section in Figure 4.1. The valley is underlain by bedrock which is assumed to be impermeable. Furthermore, barring any disruptive events, the shale and salt layers (layers 4 and 5 in Figure 4.1) have extremely low permeability. Thus in the undisturbed system, groundwater flow is largely confined to the middle and lower sandstone aquifers and is shown schematically in Figure 4.2.

The network flow representation used in NWFT is shown in Figure 4.3. The flow segments, or legs, connecting the waste depository to the middle and lower sandstone aquifers are used to represent various potential disruptive events which could allow radioactive waste to escape the depository. The boundary conditions used at the middle and lower sandstone aquifer inlets and at the discharge point are taken from a two-dimensional simulation of the reference site flow system. For this example, NWFT was used to simulate a U-tube which forms a hydraulic connection between the depository and the overlying (middle sandstone) aquifer. The U-tube is assumed to result from degradation of materials used to seal a borehole and an access shaft to the depository (Figure 4.4). The variables used in this example are as follows:

X_1 = porosity of the overlying aquifer

X_2 = hydraulic conductivity of the overlying aquifer (ft./day)

X_3 = dispersivity (ft)

X_4 = distribution coefficient of the isotope under consideration (cm^3/gm)

X_5 = leach time (years)

X_6 = porosity of the shaft/borehole sealing material

X_7 = hydraulic conductivity of the shaft/borehole sealing material (ft/day)

The input variables are treated as random variables. The reasons for this are:

(1) For some input variables the value of the variable is constant, but unknown, for a given depository. The variable is treated as a random variable with a given distribution function to reflect knowledge concerning possible values for the variable. This application is concordant with the concept of subjective probability.

(2) For some input variables there will be actual unit to unit variation within a given depository. Frequently this variation will be due to location differences within a depository. This type of variation is analogous to sampling variation in the usual statistical context.

(3) For some input variables a small number of measurements will be available. However, they are subject to measurement errors. The distribution function will be developed to reflect uncertainty associated with measurement error.

Frequently the distribution function selected for a given input variable will be developed to accommodate some combination of (1), (2), and (3) above. The purpose of the present analysis is to evaluate the uncertainties in model output (in this case radioactive discharge to the environment) which results from the uncertainty in model input variables and to determine which input variables contribute most to output uncertainties.

As stated earlier, the waste disposal site which is used in the present study to demonstrate sensitivity and risk analysis techniques is entirely hypothetical and, therefore, measurements of input values to the groundwater transport model do not exist. Discussions were held with earth scientists to determine variable ranges and distributions which might be appropriate for sedimentary basins such as the reference site evaluated here. Results of these discussions are shown in the table below. The ranges chosen are global in nature and, therefore, are somewhat broader than one might reasonably expect for a particular site. The exception is variable X_4, the range of which was truncated for reasons given below.

Variable	Range	Probability Distribution
X_1	(.05, .30)	Normal, μ = .175, σ = .04
X_2	(1, 50)	Lognormal, μ = 1.956, σ = .633
X_3	(45, 500)	Uniform
X_4	(10^{-1}, 10^2)	Loguniform
X_5	(10^3, 10^7)	Loguniform
X_6	(.005, .2)	Lognormal, μ = -3.454, σ = .597
X_7	(.01, 50)	Lognormal, μ = -.347, σ = 1.378

The normal probability distributions were truncated at the indicated range values, which were arbitrarily selected to be the .001 and .999 quantiles. Lognormal distributions were obtained from appropriately truncated normal distributions with the indicated parameters. A loguniform random variable over the range (a,b) means a random variable whose logarithm (base 10) is uniformly distributed over the range ($\log_{10}$ a, $\log_{10}$ b).

4.2 Obtaining the Latin Hypercube Sample

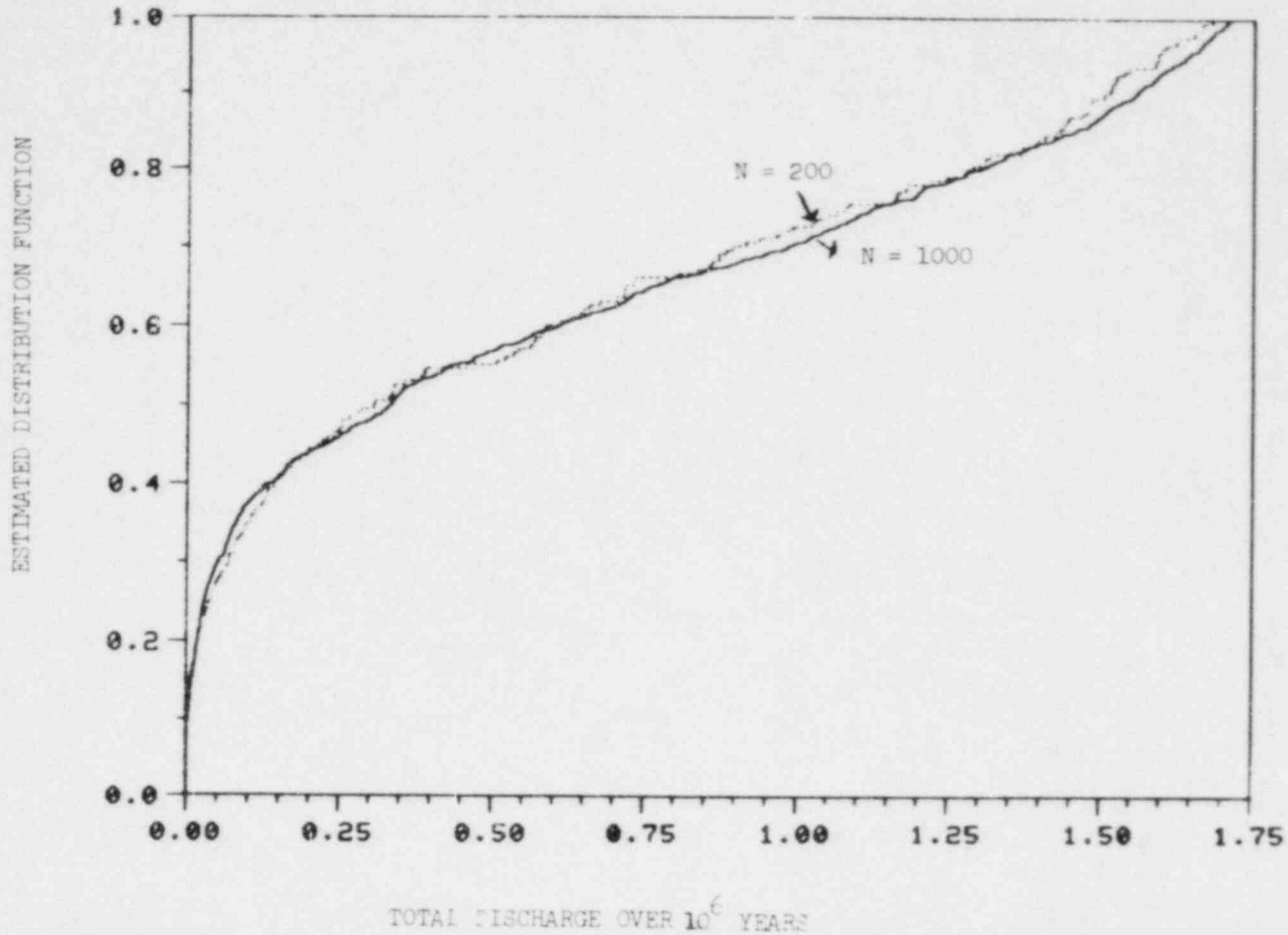
The range of each of the above input variables was divided into 200 intervals of equal probability .005, and one value from each interval was sampled at random. The 200 values thus obtained for X_1 were paired in a random manner (equally likely combinations) with the 200 values of X_2. These 200 pairs were combined in a random manner with the 200 values of X_3 to form 200 triplets, and so on, until two hundred 7-tuples were formed. This is the Latin hypercube sample, which was used as inputs into the model.

The computer program's output variable Y for this example is the total discharge of an isotope in the 10^6 years following burial of the radioactive waste. Actually the output of this isotope, for the ranges of the X_i's supplied by the geologists, included a high percentage of total discharge values equal to zero. Since these values are of little interest for illustrating the methods of Sections 2 and 3, we narrowed the range of the input variable X_4 from the interval $(10^{-2}, 10^5)$ supplied by the geologists to the interval $(10^{-1}, 10^2)$ indicated above. The effect of this restriction on the range of X_4 is the desired result that nearly all of the observed output variable values are nonzero. The 200 observations on the output Y are summarized in Figure 4.5, which presents the empirical distribution function S(y) of Y,

$$S(y) = \frac{1}{200} \sum_{i=1}^{200} u(y-Y_i) \qquad (4.1)$$

obtained from Equation (2.18) with $p_{\underset{\sim}{n}_i} = (1/N)^K = (1/200)^7$. Note that S(y) provides an estimate of the distribution function G(y) of Y. An unbiased estimate of the mean of Y is given by

FIGURE 4.5. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) BASED ON LATIN HYPERCUBE SAMPLES OF SIZE N = 200 AND N = 1000.



$$\bar{Y} = \frac{1}{200} \sum_{i=1}^{200} Y_i = .567 \qquad (4.2)$$

which is obtained from Equation (2.8) by letting $g(y) = Y$. Also appearing in Figure 4.5 is an estimate of $G(y)$ obtained by running the model on a Latin hypercube sample with N=1000. This latter estimate results in a reasonably smooth curve and is useful for comparing the variability associated with a sample of size 200. The estimate of the mean of Y with N=1000 is .580.

4.3 Identification of Influential Input Variables

One of the primary objectives of this data analysis is to assess the relative importance of each input variable. Since plots of the output of this model showed it to be a monotonic nonlinear function of the input variables, we used the techniques of stepwise regression on ranks as described in Iman and Conover (1979) to identify the important variables. The variables used in the stepwise regression were functions of X_1 through X_7. The important variables turned out to be functions of X_4 and X_5. Therefore, X_4 and X_5 were selected for closer examination.

In the next section we consider the influence on the output distribution function estimate of different distributional assumptions for X_4 and X_5. In order to aid the reader in pairing individual observations with the new weights given in the next section we provide in Table 4.1 a complete listing of the Latin hypercube sample for both X_4 and X_5 as well as the rank of the specific observation and the interval from which the observation was selected.

TABLE 4.1

Actual Latin Hypercube Sample Used With X_4 and X_5

Obs. No.	Rank of X_4	Interval Used		X_4	Rank of X_5	Interval Used		X_5
1	15	.162,	.168	.168	93	69183,	72443	69643
2	145	14.454,	14.962	14.558	151	999999,	1047128	1012668
3	131	8.913,	9.226	9.207	194	7244359,	7585775	7586225
4	176	42.170,	43.652	42.511	61	15848,	16595	16018
5	149	16.596,	17.179	16.723	198	8709635,	9120108	8990016
6	116	5.309,	5.495	5.421	92	66069,	69183	67523
7	58	.716,	.741	.729	103	109647,	114815	112830
8	48	.507,	.525	.512	37	5248,	5495	5385
9	3	.107,	.111	.108	179	3630780,	3801893	3671610
10	82	1.641,	1.698	1.691	152	1047128,	1096478	1087302
11	68	1.012,	1.047	1.023	171	2511886,	2630267	2563516
12	187	61.660,	63.826	62.138	45	7585,	7943	7874
13	191	70.795,	73.282	71.784	51	10000,	10471	10166
14	169	31.989,	33.113	32.807	111	158489,	165958	165566
15	154	19.724,	20.417	19.872	190	6025595,	6309573	6229810
16	188	63.826,	66.069	65.314	159	1445439,	1513561	1472539
17	122	6.531,	6.761	6.552	75	30199,	31622	31539
18	26	.237,	.245	.239	126	316227,	331131	326906
19	184	55.590,	57.544	57.314	156	1258925,	1318256	1292434
20	94	2.483,	2.570	2.498	155	1202264,	1258925	1212429
21	23	.214,	.221	.220	62	16595,	17378	16781
22	129	8.318,	8.610	8.432	154	1148153,	1202264	1200522
23	200	96.605,	100.000	99.031	172	2630267,	2754228	2713817
24	104	3.508,	3.631	3.527	110	151356,	158489	152345
25	198	90.157,	93.325	92.876	31	3981,	4168	4149
26	86	1.884,	1.950	1.918	91	63095,	66069	64574
27	75	1.288,	1.334	1.314	162	1659586,	1737800	1689294
28	143	13.490,	13.964	13.857	144	724435,	758577	726070
29	29	.263,	.272	.267	56	12589,	13182	12791
30	97	2.754,	2.851	2.831	10	1513,	1584	1555
31	43	.427,	.442	.430	95	75857,	79432	78591
32	1	.100,	.104	.101	128	346736,	363078	353477
33	103	3.388,	3.508	3.503	23	2754,	2884	2782
34	140	12.162,	12.589	12.437	196	7943282,	8317637	8301523
35	156	21.135,	21.878	21.551	36	5011,	5248	5024
36	119	5.888,	6.095	5.915	11	1584,	1659	1588

TABLE 4.1 (Continued)

Obs. No.	Rank of X_4	Interval Used		X_4	Rank of X_5	Interval Used		X_5
37	54	.624,	.646	.641	96	79432,	83176	81915
38	9	.132,	.136	.135	160	1513561,	1584893	1527180
39	158	22.646,	23.442	22.733	7	1318,	1380	1351
40	85	1.820,	1.884	1.839	83	43651,	45708	44338
41	18	.180	.185	.185	87	52480,	54954	54556
42	141	12.589,	13.032	12.965	113	173780,	181970	180196
43	11	.141,	.146	.145	46	7943,	8317	7983
44	197	87.096,	90.157	87.977	184	4570881,	4786300	4581385
45	73	1.202,	1.245	1.222	38	5495,	5754	5572
46	21	.200,	.207	.204	197	8317637,	8709635	8655068
47	193	75.858,	78.524	76.472	3	1096,	1148	1103
48	117	5.495,	5.689	5.582	163	1737800,	1819700	1756449
49	88	2.018,	2.089	2.075	125	301995,	316227	311869
50	61	.794,	.822	.806	63	17378,	18197	17901
51	70	1.084,	1.122	1.099	4	1148,	1202	1163
52	173	38.019,	39.355	38.255	78	34673,	36307	35991
53	96	2.661,	2.754	2.676	191	6309573,	6606934	6529586
54	170	34.277,	35.481	35.286	130	380189,	398107	383281
55	64	.881,	.912	.886	153	1096478,	1148153	1113823
56	67	.977,	1.012	.983	137	524807,	549540	544436
57	155	20.417,	21.135	20.485	124	288403,	301995	297175
58	14	.157,	.162	.160	98	87096,	91201	91005
59	4	.111,	.115	.112	42	6606,	6918	6793
60	33	.302,	.313	.308	167	2089296,	2187761	2129153
61	105	3.631,	3.758	3.695	13	1737,	1819	1777
62	6	.119,	.123	.123	85	47863,	50118	48963
63	162	26.002,	26.915	26.061	44	7244,	7585	7297
64	12	.146,	.151	.147	107	131825,	138038	136571
65	81	1.585,	1.641	1.638	180	3801893,	3981071	3955717
66	115	5.129,	5.309	5.277	100	95499,	99999	95521
67	159	23.442,	24.266	23.898	19	2290,	2398	2369
68	125	7.244,	7.499	7.462	33	4365,	4570	4476
69	109	4.169,	4.315	4.206	117	208939,	218776	216454
70	147	15.488,	16.032	15.898	22	2630,	2754	2637
71	41	.398,	.412	.408	68	21877,	22908	22581
72	174	39.355,	40.738	40.087	131	398107,	416869	410922
73	127	7.762,	8.035	7.931	84	45708,	47863	46124
74	17	.174,	.180	.175	88	54954,	57543	57470
75	110	4.315,	4.467	4.375	135	478630,	501187	486553

TABLE 4.1 (Continued)

Obs No.	Rank of X_4	Interval Used		X_4	Rank of X_5	Interval Used		X_5
76	144	13.964,	14.454	14.289	58	13803,	14454	13844
77	163	26.915,	27.861	27.089	148	870963,	912010	896289
78	160	24.266,	25.119	24.964	185	4786300,	5011872	4886404
79	39	.372,	.385	.382	69	22908,	23988	23288
80	36	.335,	.347	.340	116	199526,	208929	201086
81	167	30.903,	31.989	31.511	195	7585775,	7943282	7623814
82	127	7.499,	7.762	7.685	81	39810,	41686	41666
83	199	93.325,	96.605	94.371	143	691830,	724435	714270
84	171	35.481,	36.728	35.615	66	19942,	20892	20478
85	31	.282,	.292	.288	25	3019,	3162	3137
86	34	.313,	.324	.319	29	3630,	3801	3638
87	83	1.698,	1.758	1.758	67	20892,	21877	21067
88	91	2.239,	2.317	2.286	132	416869,	436515	426240
89	148	16.032,	16.596	16.432	32	4168,	4365	4345
90	152	18.408,	19.055	19.044	70	23988,	25118	24415
91	76	1.334,	1.380	1.359	177	3311311,	3467368	3435478
92	90	2.163,	2.239	2.182	24	2884,	3019	2984
93	136	10.593,	10.965	10.841	150	954992,	999999	998005
94	60	.767,	.794	.781	82	41686,	43651	42087
95	146	14.962,	15.488	15.167	182	4168693,	4365158	4342459
96	165	28.840,	29.854	29.225	76	31622,	33113	32961
97	99	2.951,	3.055	2.953	65	19054,	19952	19933
98	142	13.032,	13.490	13.215	199	9120108,	9549925	9243614
99	24	.221,	.229	.222	71	25118,	26302	26183
100	69	1.047,	1.084	1.084	53	10964,	11481	11378
101	50	.543,	.562	.562	149	912010,	954992	939797
102	169	33.113,	34.277	33.356	26	3162,	3311	3255
103	55	.646,	.668	.659	157	1318256,	1380384	1326384
104	123	6.761,	6.998	6.894	6	1258,	1318	1267
105	130	8.610,	8.913	8.754	41	6309,	6606	6549
106	153	19.055,	19.724	19.375	79	36307,	38018	36811
107	30	.272,	.282	.279	188	5495408,	5754399	5567481
108	22	.207,	.214	.212	54	11481,	12022	11560
109	132	9.226,	9.550	9.351	183	4365158,	4570881	4430361
110	51	.562,	.582	.563	146	794328,	831763	808509

TABLE 4.1 (Continued)

Obs. No.	Rank of X_4	Interval Used		X_4	Rank of X_5	Interval Used		X_5
111	183	53.703,	55.590	54.730	161	1584893,	1659586	1646253
112	101	3.162,	3.273	3.247	123	275422,	288403	281998
113	45	.457,	.473	.471	181	3981071,	4168693	4081659
114	7	.123,	.127	.127	112	165958,	173780	173381
115	195	81.283,	84.140	83.000	9	1445,	1513	1513
116	49	.525,	.543	.538	120	239883,	251188	240827
117	190	68.391,	70.795	68.741	129	363078,	380189	372611
118	63	.851,	.881	.879	47	8317,	8709	8638
119	98	2.851,	2.941	2.932	17	2089,	2187	2141
120	128	8.035,	8.318	8.175	192	6606934,	6918309	6714065
121	196	84.140,	87.096	85.434	2	1047,	1096	1056
122	172	36.728,	38.019	36.843	106	125892,	131825	129323
123	47	.490,	.507	.499	169	2290847,	2398832	2357018
124	46	.478,	.490	.476	173	2754228,	2884031	2850385
125	100	3.055,	3.162	3.148	114	181970,	190546	182978
126	28	.254,	.263	.253	136	501187,	524807	512205
127	25	.229,	.237	.236	86	50118,	52480	50160
128	74	1.245,	1.288	1.281	189	5754399,	6025595	5910095
129	164	27.861,	28.840	28.059	104	120226,	125892	125447
130	5	.115,	.119	.115	186	5011872,	5248074	5069598
131	27	.245,	.254	.250	77	33113,	34673	33778
132	8	.127,	.132	.132	141	630957,	660693	638510
133	185	57.544,	59.566	57.815	174	2884031,	3019951	2947853
134	92	2.317,	2.399	2.376	139	575439,	692559	595725
135	38	.359,	.372	.362	50	9549,	9999	9565
136	19	.186,	.193	.190	175	3019951,	3162277	3038191
137	192	73.282,	75.858	74.116	119	229086,	239883	232779
138	108	4.027,	4.169	4.112	97	83176,	87096	84980
139	139	11.749,	12.162	11.843	72	26302,	27542	26817
140	166	29.854,	30.903	30.602	73	27542,	28840	28422
141	95	2.570,	2.661	2.645	99	91201,	95499	94989
142	124	6.998,	7.244	7.095	1	999,	1047	1005
143	177	43.652,	45.186	45.030	101	100000,	104712	104513
144	16	.168,	.174	.173	121	251188,	263026	253175
145	71	1.122,	1.161	1.122	43	6918,	7244	7188

TABLE 4.1 (Continued)

Obs No.	Rank of X_4	Interval Used		X_4	Rank of X_5	Interval Used		X_5
146	80	1.531,	1.585	1.584	30	3801,	3981	3909
147	161	25.119,	26.002	25.345	170	2398832,	2511886	2510679
148	78	1.429,	1.479	1.469	104	114815,	120226	118327
149	53	.603,	.624	.609	168	2187761,	2290867	2237451
150	72	1.161,	1.202	1.173	20	2398,	2411	2403
151	150	17.179,	17.783	17.460	127	331131,	346736	342584
152	35	.324,	.335	.326	8	1380,	1445	1422
153	66	.944,	.977	.975	5	1202,	1258	1238
154	106	3.758,	3.890	3.778	80	38018,	39810	39350
155	179	46.774,	48.417	47.114	74	28840,	30199	29243
156	13	.141,	.157	.153	21	2411,	2630	2513
157	102	3.273,	3.388	3.275	35	4786,	5011	5002
158	114	4.955,	5.129	5.000	193	6918309,	7244359	7206330
159	87	1.950,	2.018	1.984	118	218776,	229085	223413
160	10	.136,	.141	.137	69	14454,	15135	15063
161	57	.692,	.716	.696	12	1659,	1737	1704
162	151	17.783,	18.408	18.391	200	9549925,	9999999	9872012
163	77	1.380,	1.429	1.386	142	660693,	691830	689137
164	79	1.479,	1.531	1.527	16	1995,	2089	2033
165	134	9.886,	10.233	9.923	48	8709,	9120	8890
166	37	.347,	.359	.358	176	3162277,	3311311	3222083
167	178	45.186,	46.774	45.455	64	18197,	19054	18542
168	56	.668,	.692	.684	49	9120,	9549	9514
169	42	.412,	.427	.425	166	1995262,	2089296	2042173
170	182	51.880,	53.703	52.864	138	549540,	575439	553825
171	44	.442,	.457	.442	165	1905460,	1995262	1911211
172	84	1.758,	1.820	1.759	60	15135,	15848	15390
173	121	6.310,	6.531	6.418	27	3311,	3467	3410
174	118	5.689,	5.888	5.775	57	13182,	13803	13347
175	107	3.890,	4.027	3.976	133	436515,	457088	449309
176	180	48.417,	50.119	49.775	164	1819700,	1905460	1878104
177	113	4.786,	4.955	4.919	102	104712,	109647	106948
178	62	.822,	.851	.828	134	457088,	478630	461400
179	133	9.550,	9.886	9.678	158	1380384,	1445439	1385105
180	2	.104,	.107	.105	89	57543,	60255	59322

TABLE 4.1 (Continued)

Obs. No.	Rank of X_4	Interval Used		X_4	Rank of X_5	Interval Used		X_5
181	65	.912,	.944	.922	14	1819,	1905	1847
182	40	.385,	.398	.393	145	758577,	794328	764214
183	89	2.089,	2.163	2.127	40	6025,	6309	6257
184	186	59.566,	61.660	60.109	115	190546,	199526	195136
185	135	10.233,	10.593	10.391	34	4570,	4786	4755
186	111	4.467,	4.624	4.562	90	60255,	63095	61100
187	120	6.095,	6.310	6.153	140	602559,	630957	611131
188	175	40.738,	42.170	41.962	108	138038,	144543	138559
189	20	.193,	.200	.197	122	263026,	275422	270133
190	32	.292,	.302	.296	147	831763,	870963	853889
191	189	66.069,	68.391	66.259	187	5248074,	5495408	5308941
192	52	.582,	.603	.591	18	2187,	2290	2189
193	112	4.624,	4.786	4.641	15	1905,	1995	1911
194	59	.741,	.767	.749	109	144543,	151356	148721
195	137	10.965,	11.350	11.194	52	10471,	10964	10802
196	157	21.878,	22.646	22.046	94	72443,	75857	74513
197	93	2.399,	2.483	2.476	39	5754,	6025	6007
198	181	50.119,	51.880	51.444	55	12022,	12589	12250
199	194	78.524,	81.283	78.915	28	3467,	3630	3583
200	138	11.350,	11.749	11.438	178	3467368,	3630780	3601253

4.4 Determination of the Sensitivity of the Output to Distributional Assumptions on Influential Input Variables

Because of the influence on Y of X_4 and X_5 the assumptions regarding the probability distributions of X_4 and X_5 become of particular interest. That is, how sensitive is the probability distribution of Y to the particular distributional assumptions made on X_4 and X_5? In particular, if the distributions of X_4 and X_5 were actually lognormal instead of loguniform, with the range space remaining the same, how much would the distribution of Y be affected? Changing the entire form of the parent distribution would probably be considered as quite extreme. It seems more likely that the parameters of the distribution rather than the form of the distribution would be subject to question. The methods of Section 3 enable us to investigate these areas (within reason) without making new computer runs under the changed input distribution(s).

We now consider three cases for purposes of illustration; (1) the distribution of X_4 is assumed to be lognormal while X_5 remains loguniform, (2) the distribution of X_5 is assumed to be lognormal while X_4 remains loguniform, and (3) the distributions of both X_4 and X_5 are assumed to be lognormal. Each of the 200 input vectors used in this study carries an initial weight of $(1/200)^7 \quad 200^{7-1} = .005$ since the original Latin hypercube sample was based on equal weights. The weights associated with these input vectors will no longer necessarily be .005 as the intervals in Table 4.1 will be associated with new probabilities according to the lognormal distribution. These new weights are given in Table 4.2 for each of the above 3 cases. Examination of Table 4.2 shows these new weights (or step heights for the new estimate of the output c.d.f.) to range from .000109 to .012350 for both cases (1) and (2), and from .000004 to .030042 for case

TABLE 4.2

New Weights Assigned to Input Vectors

For the 3 Cases Under Consideration

Obs. No.	Case 1	Case 2	Case 3	Obs. No.	Case 1	Case 2	Case 3
1	.000377	.012024	.000906	36	.010490	.000270	.000566
2	.004799	.003656	.003509	37	.004400	.012233	.010765
3	.007923	.000190	.000301	38	.000227	.002279	.000103
4	.000813	.005865	.000953	39	.002548	.000190	.000097
5	.004018	.000132	.000106	40	.011013	.010672	.023506
6	.011013	.011933	.026284	41	.000479	.011322	.001086
7	.005215	.012315	.012844	42	.005645	.011464	.012943
8	.003314	.001802	.001194	43	.000270	.002992	.000161
9	.000132	.000652	.000017	44	.000145	.000443	.000013
10	.010490	.003482	.007[illegible]06	45	.008609	.001914	.003295
11	.007460	.001152	.001718	46	.000605	.000145	.000018
12	.000347	.002839	.000197	47	.000208	.000132	.000005
13	.000248	.003835	.000190	48	.010846	.001914	.004151
14	.001403	.011718	.003289	49	.011464	.009274	.021264
15	.003150	.000270	.000170	50	.005865	.006312	.007404
16	.000319	.002411	.000154	51	.007923	.000145	.000230
17	.009906	.009056	.017940	52	.001005	.0[illegible]700	.001949
18	.000873	.009056	.001581	53	.012233	.000248	.000606
19	.000443	.002839	.000251	54	.001231	.008153	.002008
20	.012105	.002992	.007243	55	.006539	.003314	.004334
21	.000702	.006087	.000855	56	.007229	.006539	.009454
22	.008382	.003150	.005281	57	.002992	.009489	.005678
23	.000109	.001076	.000024	58	.000347	.012315	.000855
24	.012280	.011831	.029055	59	.000145	.002411	.000070
25	.000132	.001231	.000033	60	.001403	.001496	.000420
26	.011172	.011831	.026435	61	.012233	.000319	.000782
27	.009056	.002030	.003677	62	.000174	.011013	.000383
28	.005215	.005005	.005220	63	.002030	.002691	.001093
29	.001076	.004799	.001033	64	.000294	.012105	.000711
30	.012280	.000248	.000608	65	.010301	.000605	.001245
31	.002548	.012174	.006205	66	.011172	.012350	.027595
32	.000109	.008609	.000188	67	.002411	.000518	.000250
33	.012315	.000702	.001730	68	.009274	.001403	.002603
34	.005865	.000159	.000186	69	.001933	.010846	.025885
35	.002839	.001695	.000962	70	.004400	.000652	.000574

TABLE 4.2 (Continued)

Obs. No.	Case 1	Case 2	Case 3	Obs. No.	Case 1	Case 2	Case 3
71	.002279	.007460	.003400	106	.003314	.009906	.006565
72	.000937	.007923	.001485	107	.001152	.000319	.000074
73	.008833	.010846	.019162	108	.000652	.004400	.000574
74	.000443	.011464	.001015	109	.007691	.000479	.000737
75	.011831	.006998	.016558	110	.003835	.004597	.003526
76	.005005	.005215	.005220	111	.000479	.002152	.000206
77	.001914	.004207	.001610	112	.012350	.009700	.023959
78	.002279	.000409	.000186	113	.002839	.000560	.000318
79	.002030	.007691	.003123	114	.000190	.011596	.000441
80	.001695	.011013	.003734	115	.000174	.000227	.000008
81	.001496	.000174	.000052	116	.003482	.010301	.007174
82	.009056	.010301	.018657	117	.000270	.008382	.000452
83	.000120	.005215	.000125	118	.006312	.003150	.003977
84	.001152	.006998	.001612	119	.012315	.000443	.001091
85	.001231	.000813	.000200	120	.008609	.000227	.000391
86	.001496	.001076	.000322	121	.000159	.000120	.000004
87	.010672	.007229	.015428	122	.001076	.012174	.002620
88	.011831	.007691	.018199	123	.003150	.001315	.000828
89	.004207	.001315	.001106	124	.002992	.001005	.000601
90	.003482	.007923	.005518	125	.012350	.011322	.027966
91	.009274	.000756	.001402	126	.001005	.006768	.001360
92	.011718	.000756	.001771	127	.000813	.011172	.001816
93	.006768	.003835	.005191	128	.008833	.000294	.000519
94	.005645	.010490	.011843	129	.001802	.012233	.004409
95	.004597	.000518	.000477	130	.000159	.000377	.000012
96	.001695	.009274	.003144	131	.000937	.009489	.001778
97	.012338	.006768	.016701	132	.000208	.005645	.000235
98	.005428	.000120	.000131	133	.000409	.000937	.000077
99	.000756	.008153	.001232	134	.011933	.006087	.014528
100	.007691	.004207	.006471	135	.001914	.003656	.001399
101	.003656	.004018	.002938	136	.000518	.000873	.000091
102	.001315	.000873	.000230	137	.000227	.010490	.000476
103	.004597	.002691	.002474	138	.012024	.012280	.029531
104	.009700	.000174	.000337	139	.006087	.008382	.010204
105	.008153	.002279	.003716	140	.001593	.008609	.002743

TABLE 4.2 (Continued)

Obs. No.	Case 1	Case 2	Case 3	Obs. No.	Case 1	Case 2	Case 3
141	.012174	.012338	.030042	176	.000605	.001802	.000218
142	.0094[illegible]	.000109	.000208	177	.011464	.012338	.028289
143	.000[illegible]	.012350	.0018[illegible]7	178	.006087	.007229	.008801
144	.000409	.010106	.00082[illegible]	179	.007460	.002548	.003802
145	.008153	.002548	.004155	180	.000120	.011596	.000279
146	.010106	.001152	.002328	181	.006768	.000347	.000470
147	.002152	.001231	.000530	182	.002152	.004799	.002066
148	.009700	.012280	.023822	183	.011596	.002152	.004991
149	.004207	.001403	.001181	184	.000377	.011172	.000842
150	.008382	.000560	.000939	185	.006998	.001496	.002094
151	.003835	.008833	.006775	186	.011718	.011718	.027[illegible]54
152	.001593	.000208	.000066	187	.010301	.005865	.012083
153	.006998	.000159	.000222	188	.000873	.012024	.002100
154	.012174	.010106	.024608	189	.000560	.009906	.001110
155	.000652	.008833	.001152	190	.001315	.004400	.001157
156	.000319	.000605	.000039	191	.000294	.000347	.000020
157	.012338	.001593	.003932	192	.004018	.000479	.000385
158	.011322	.010672	.024166	194	.005428	.011933	.012955
159	.011322	.010672	.024166	194	.005428	.011933	.012955
160	.000248	.005428	.000269	195	.006539	.004018	.005255
161	.005005	.000294	.000294	196	.002691	.012105	.006515
162	.003656	.000109	.000080	197	.012024	.002030	.004883
163	.009489	.005428	.010302	198	.000560	.004597	.000515
164	.009906	.000409	.000810	199	.000190	.001005	.000038
165	.007229	.003314	.004791	200	.006312	.000702	.000887
166	.001802	.000313	.000293		1.000000	1.000000	.986947
167	.000702	.006539	.000918				
168	.004799	.003482	.003343				
169	.002411	.001593	.000768				
170	.000518	.006312	.000654				
171	.002691	.001695	.000912				
172	.010846	.005645	.012246				
173	.010106	.000937	.001894				
174	.010672	.005005	.010683				
175	.012105	.007460	.018061				

(3). This indicates that the proposed changes in distributional assumptions have not dramatically changed the probabilities associated with the intervals in Table 4.1, for if they had we might observe a large number of near zero weights which are dominated by a few large values.

For ease of comparison the results contained in Table 4.2 are presented in graphical form in Figures 4.6 through 4.8. The weights in the column labeled Case 1 in Table 4.2 are used in the estimator S'(y) defined by Equation (3.6). The graph of S'(y) appears in Figure 4.6, along with S(y) which appeared earlier in Figure 4.5. Recall that S(y) estimates the distribution function of the output under base case conditions, while S'(y) estimates the distribution function of the output after changing the assumed distribution on X_4, but without making any additional runs on the computer. In Figure 4.5 it is seen that the estimate S(y) agrees well with the "true" distribution function. Now the question naturally arises, "How well does S'(y) estimate the new distribution function of the output?" Because we are using the simplified version of the transport model, the question is relatively easy to answer. The procedure outlined in Section 4.2 was repeated to obtain 1000 points, with the new distributional assumption on X_4. These 1000 points were run on the simplified model, and the empirical distribution function thus obtained was used as the "true" distribution function for purposes of evaluating S'(y). The resulting curve appears in Figure 4.6 with S(y) and S'(y). The estimate S'(y) appears to be a reasonable estimate of the true distribution function, but it is not possible to tell from this one example whether the size of the differences between S(y) and the "true" c.d.f. is what one might expect due to sampling fluctation. For this reason two other cases are examined in this subsection, and the effect of using different sample sizes is examined

FIGURE 4.6. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 200) AND CORRESPONDING CASE 1 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 1 "TRUE" CURVE BASED ON N = 1000.

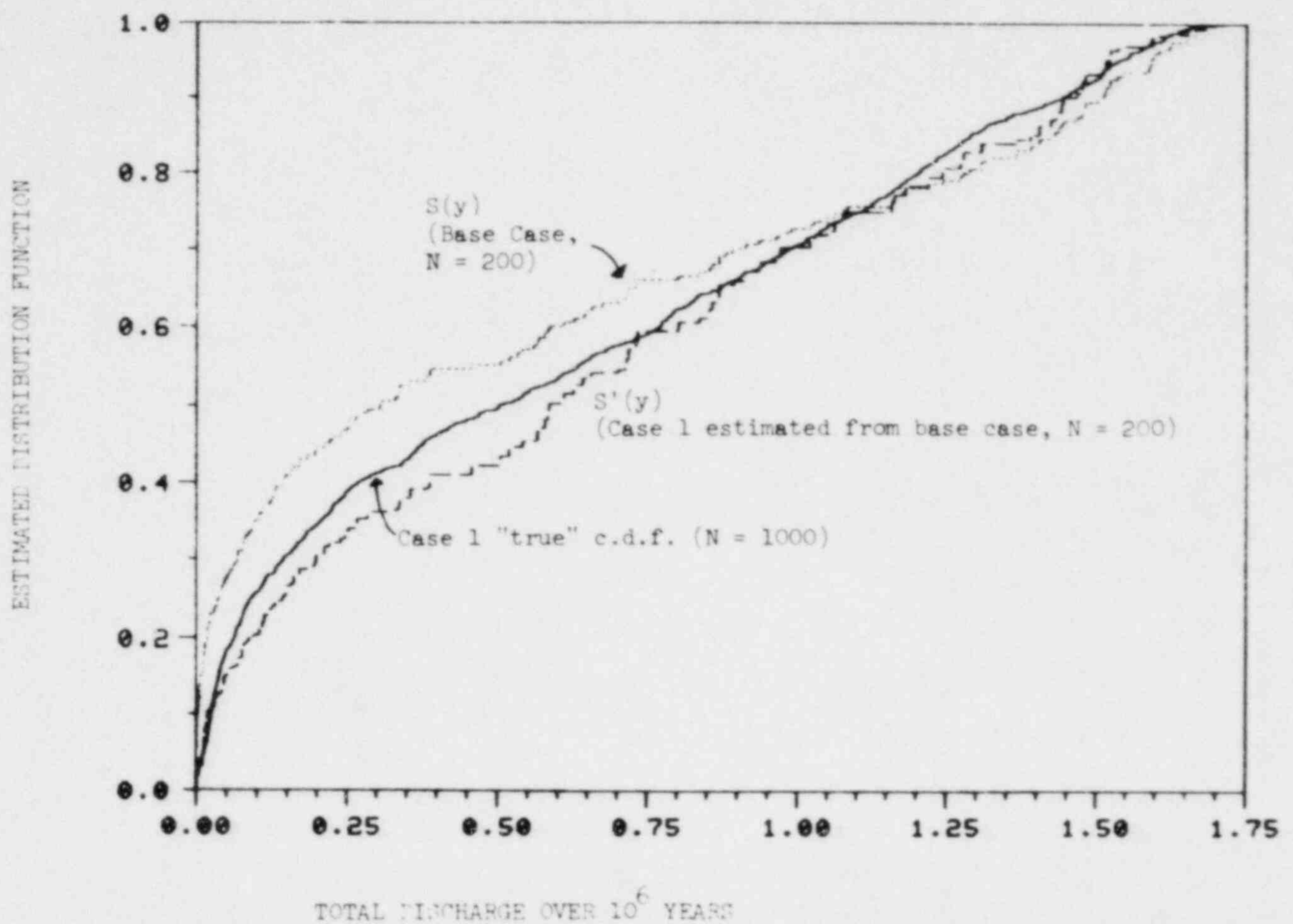


FIGURE 4.7. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 200) AND CORRESPONDING CASE 2 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 2 "TRUE" CURVE BASED ON N = 1000.

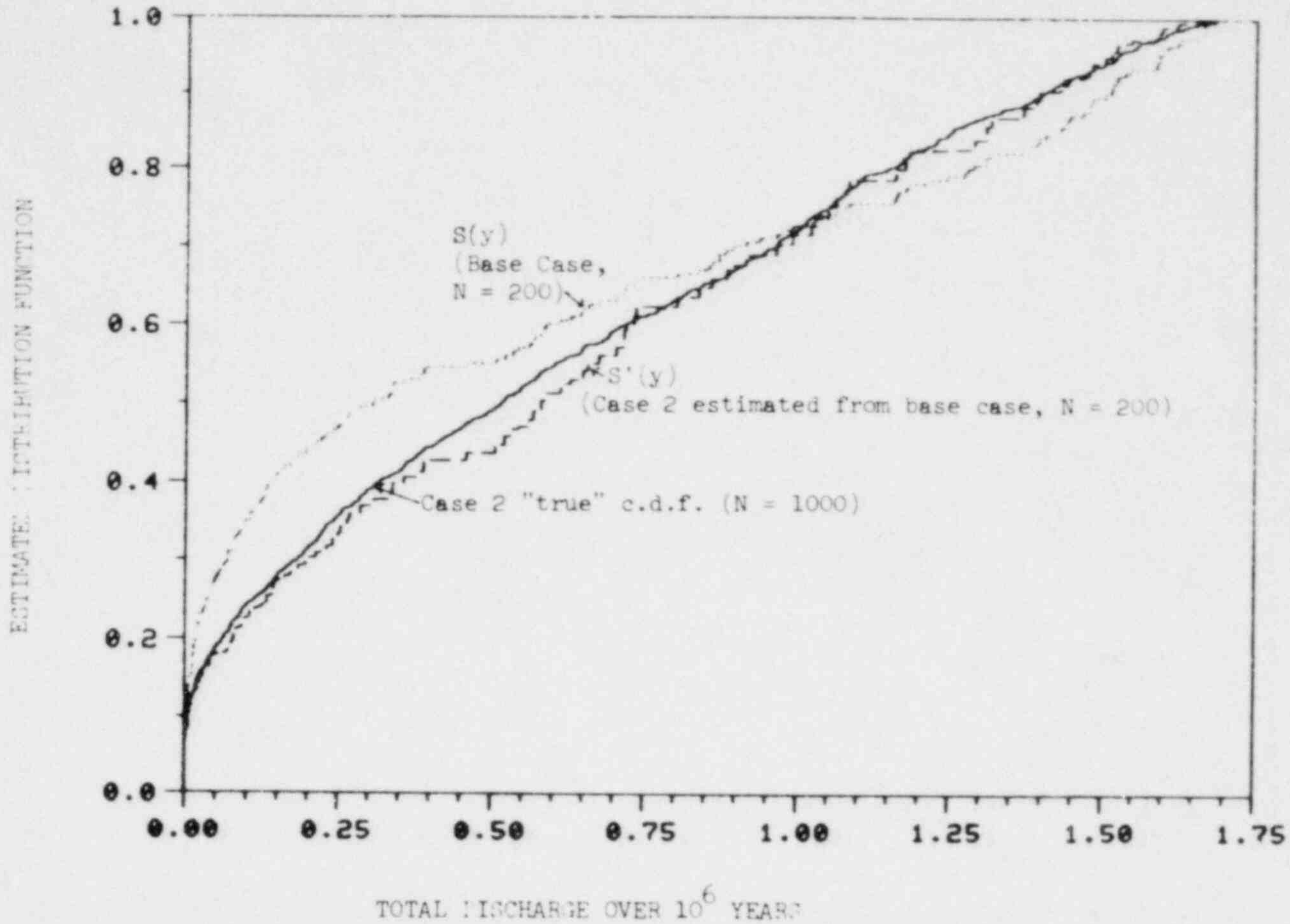
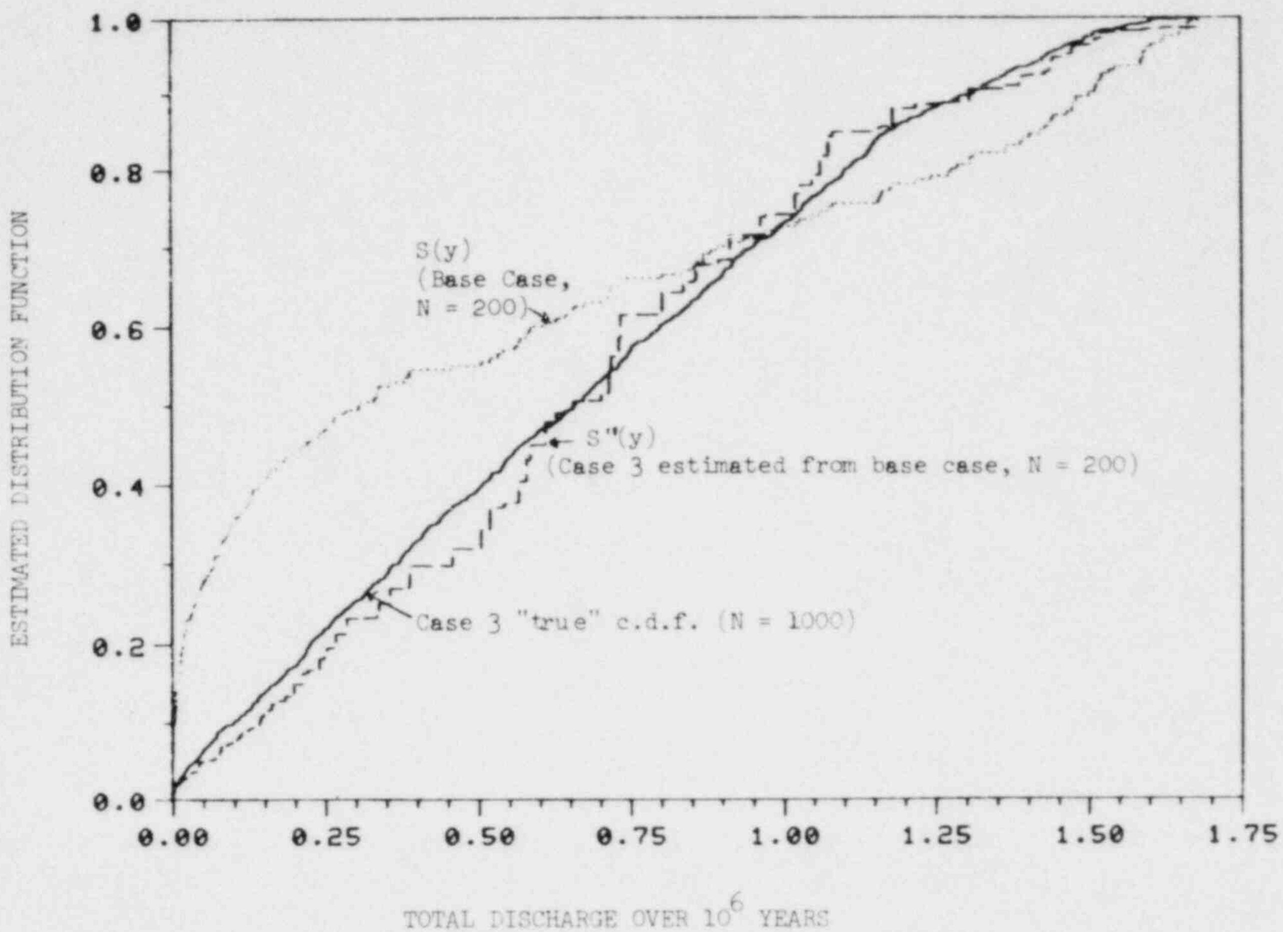


FIGURE 4.8. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 200) AND CORRESPONDING CASE 3 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 3 "TRUE" CURVE BASED ON N = 1000.



in the next subsection.

A similar analysis of the effect of changing X_5's distribution function is summarized in Figure 4.7. In that figure appear the same function S(y) as before, the estimate S'(y) of the new distribution function using the weights listed under Case 2 in Table 4.2, and a graph of the "true" distribution function obtained by running an additional 1000 points through the simplified model, using Latin hypercube Sampling. There is less difference between S'(y) and the distribution function here than appeared in Figure 4.6. In fact the agreement appears to be pretty good.

For the third and final comparison both the X_4 and X_5 input distributions were changed. The original estimator S(y) from Figure 4.5 appears again in Figure 4.8 as a point of reference. The new estimator in this case was actually

$$S''(y) = \frac{S'(y)}{S'(\infty)} = \frac{S'(y)}{.986947} \tag{4.3}$$

to obtain an estimator that increased from 0 to 1, as suggested in Section 2. The new estimator S"(y) is seen in Figure 4.8 to agree well with the "true" distribution function, obtained using an additional Latin hypercube sample in the simplified model.

4.5 The Effect of Sample Size on Estimated c.d.f.'s

The results of the previous section were all based on a Latin hypercube sample with N = 200. Given that the predictions turned out to be in reasonably good agreement with actual results one might be led to wonder what results smaller sample sizes would produce. In this subsection we investigate the effect of sample size for the simplified transport model. We would like to emphasize that the results of this subsection apply to the simplified transport model and these results may not apply directly to just any model since sample size requirements are a function of the model complexity and number of variables used. However it would be reasonable to expect improvement with increased sample sizes.

In this subsection we consider sample sizes 50 and 100. Figures 4.9 and 4.10 contain estimates of the base c.d.f. for Latin hypercube sample sizes 50 and 100 respectively. These two figures can be compared with Figure 4.5 where N = 200. The estimates seem to improve between N = 50 and 100.

The results reported in the previous subsection (Figure 4.6) for Case 1, N = 200, may be compared directly with the results depicted in Figure 4.11, N = 50, and Figure 4.12, N = 100. The actual relationship between the "true" base case c.d.f. and "true" case 1 c.d.f. is more easily discernible with N = 100 and 200 than with N = 50. In like manner, Figure 4.7 of the previous subsection may be compared directly with Figure 4.13, N = 50, and Figure 4.14, N = 100, to see the effect of various sample sizes in Case 2. Once again there is a slight but definite improvement in the quality of the estimators as the sample size increases.

In Figure 4.8 for Case 3 of the previous subsection, where there is a more dramatic change in the c.d.f.'s, the estimates based on N = 200 are much closer to the "true" c.d.f. for Case 3 than are the corresponding estimates for N = 50 (Figure 4.15) and N = 100 (Figure 4.16). Note that here as in the previous subsection the estimator

$$S''(y) = \frac{S'(y)}{S'(\infty)} \tag{4.4}$$

FIGURE 4.9. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) BASED ON LATIN HYPERCUBE SAMPLES OF SIZE N = 50 AND N = 1000.

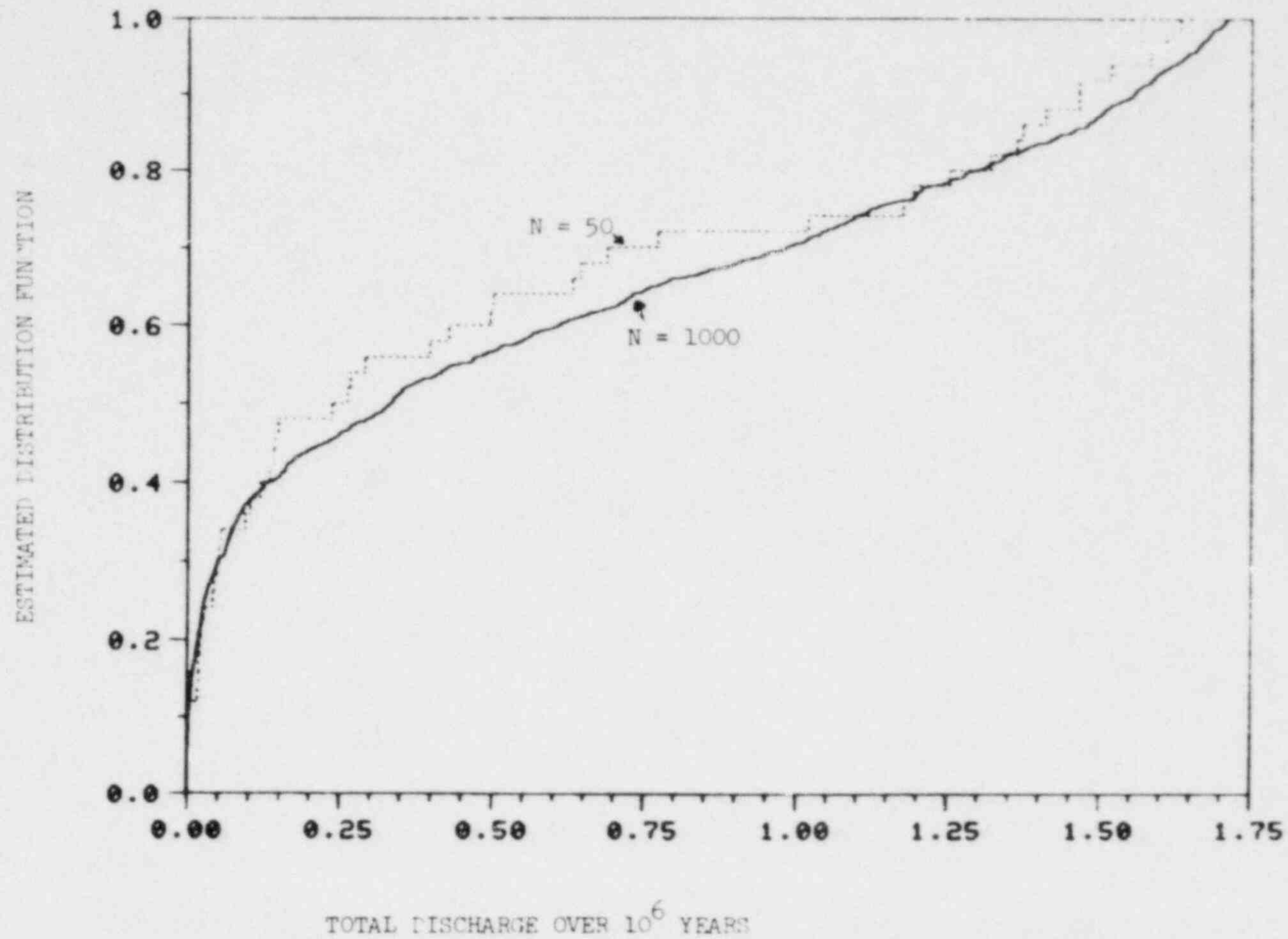


FIGURE 4.10. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) BASED ON LATIN HYPERCUBE SAMPLES OF SIZE N = 100 AND N = 1000.

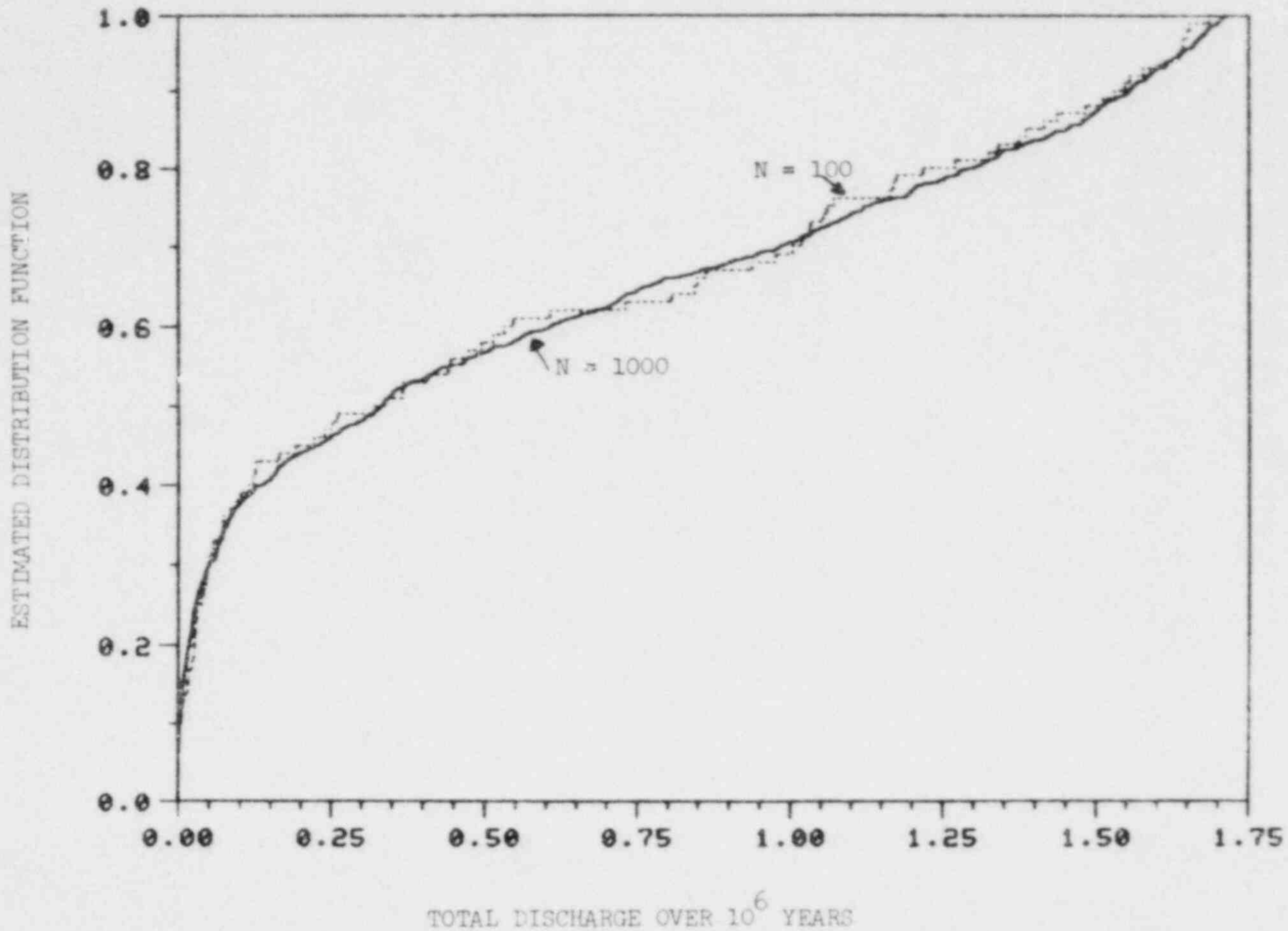


FIGURE 4.11. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 50) AND CORRESPONDING CASE 1 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 1 "TRUE" CURVE BASED ON N = 1000.

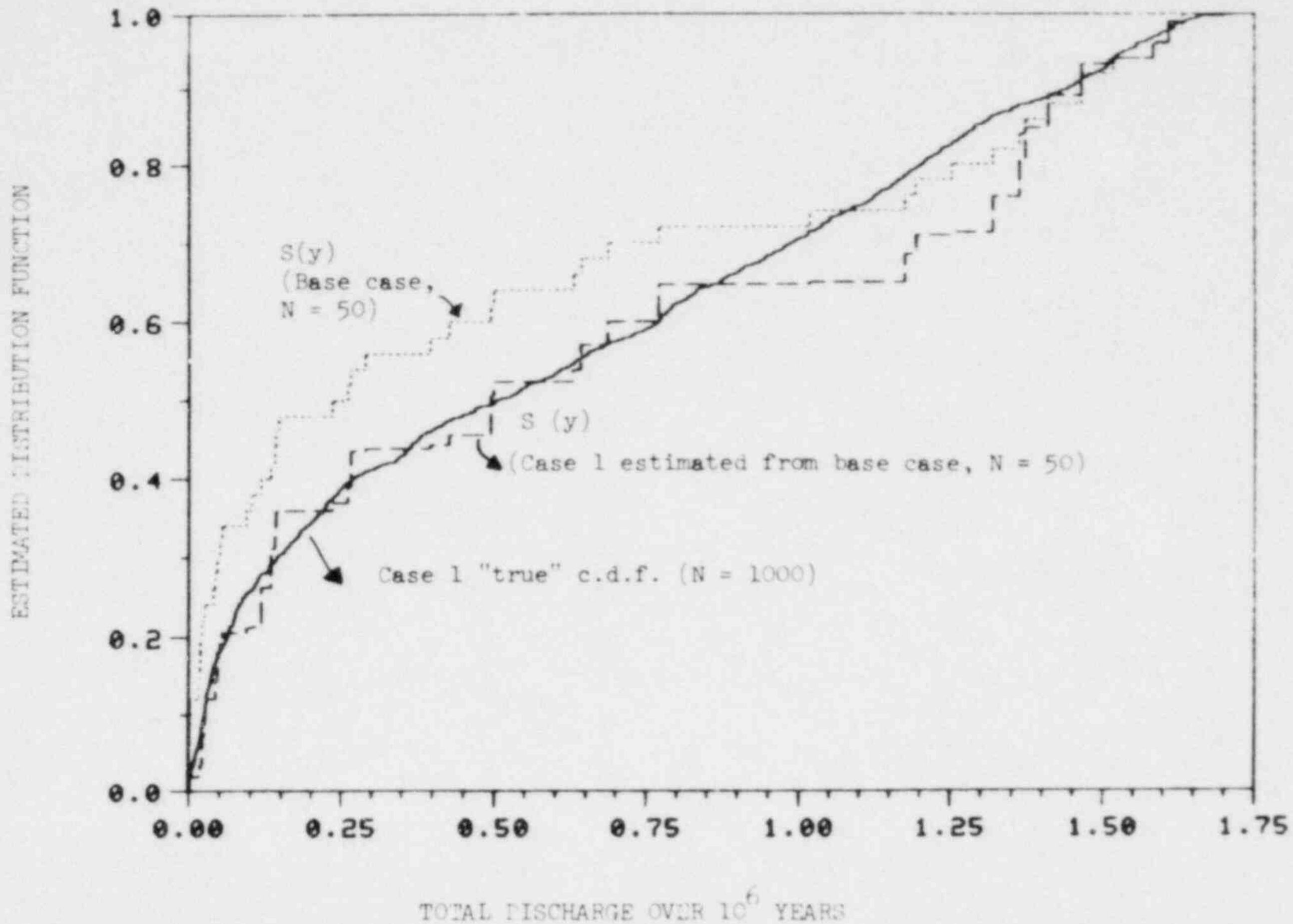


FIGURE 4.12. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 100) AND CORRESPONDING CASE 1 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 1 "TRUE" CURVE BASED ON N = 1000.

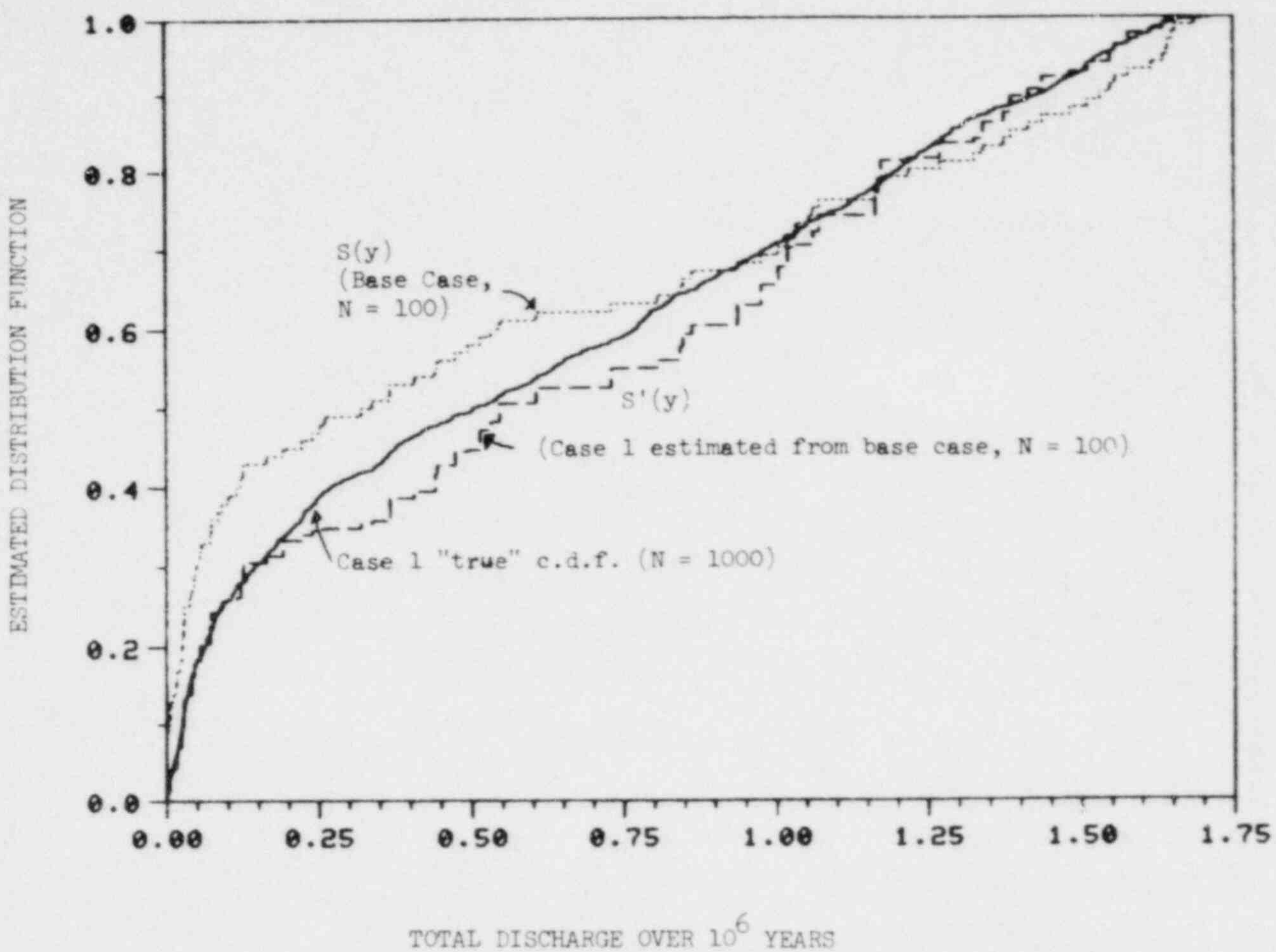


FIGURE 4.13. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 50) AND CORRESPONDING CASE 2 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 2 "TRUE" CURVE BASED ON N = 1000.

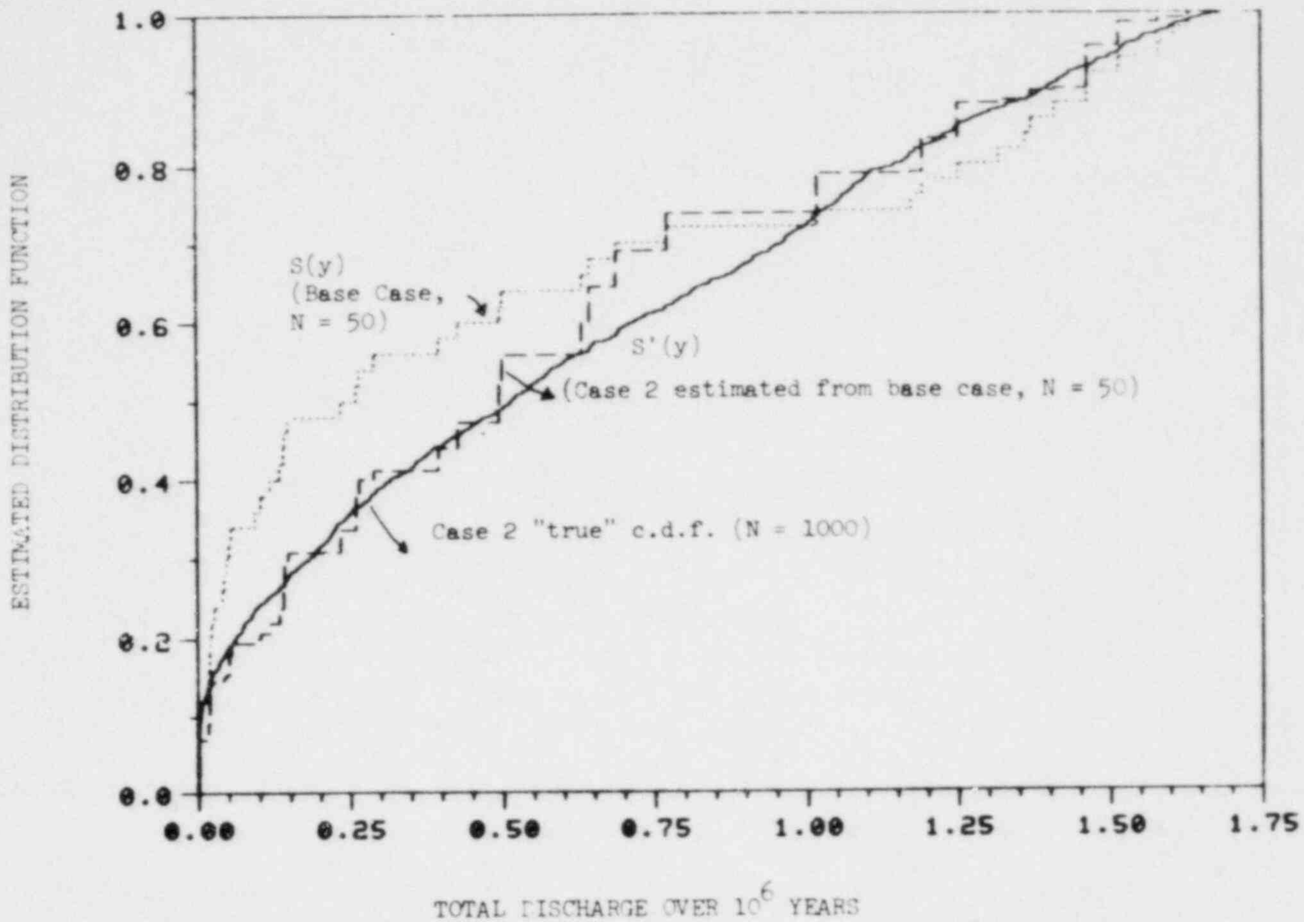


FIGURE 4.14. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 100) AND CORRESPONDING CASE 2 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 2 "TRUE" CURVE BASED ON N = 1000.

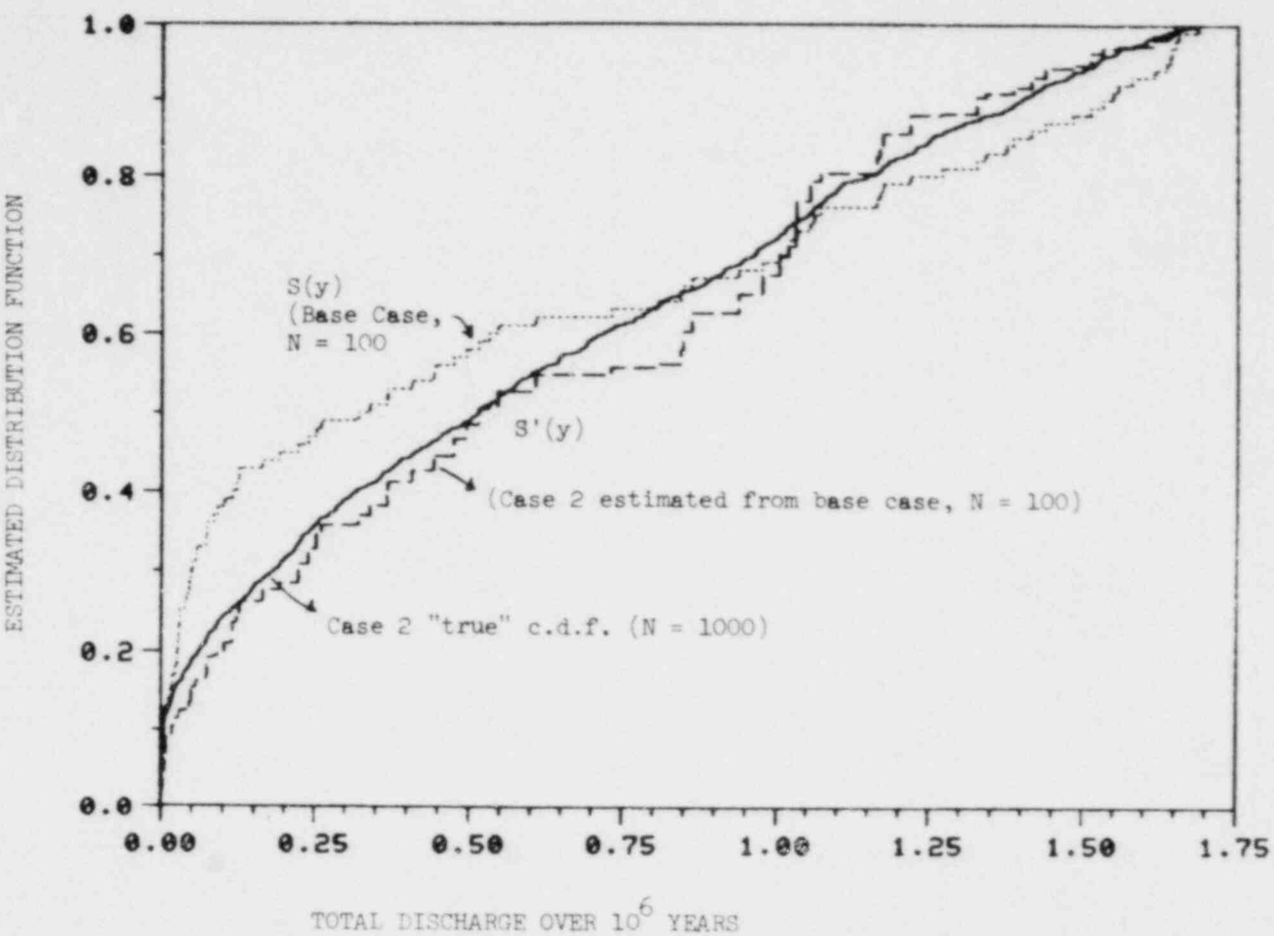


FIGURE 4.15. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 50) AND CORRESPONDING CASE 3 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 3 "TRUE" CURVE BASED ON N = 1000.

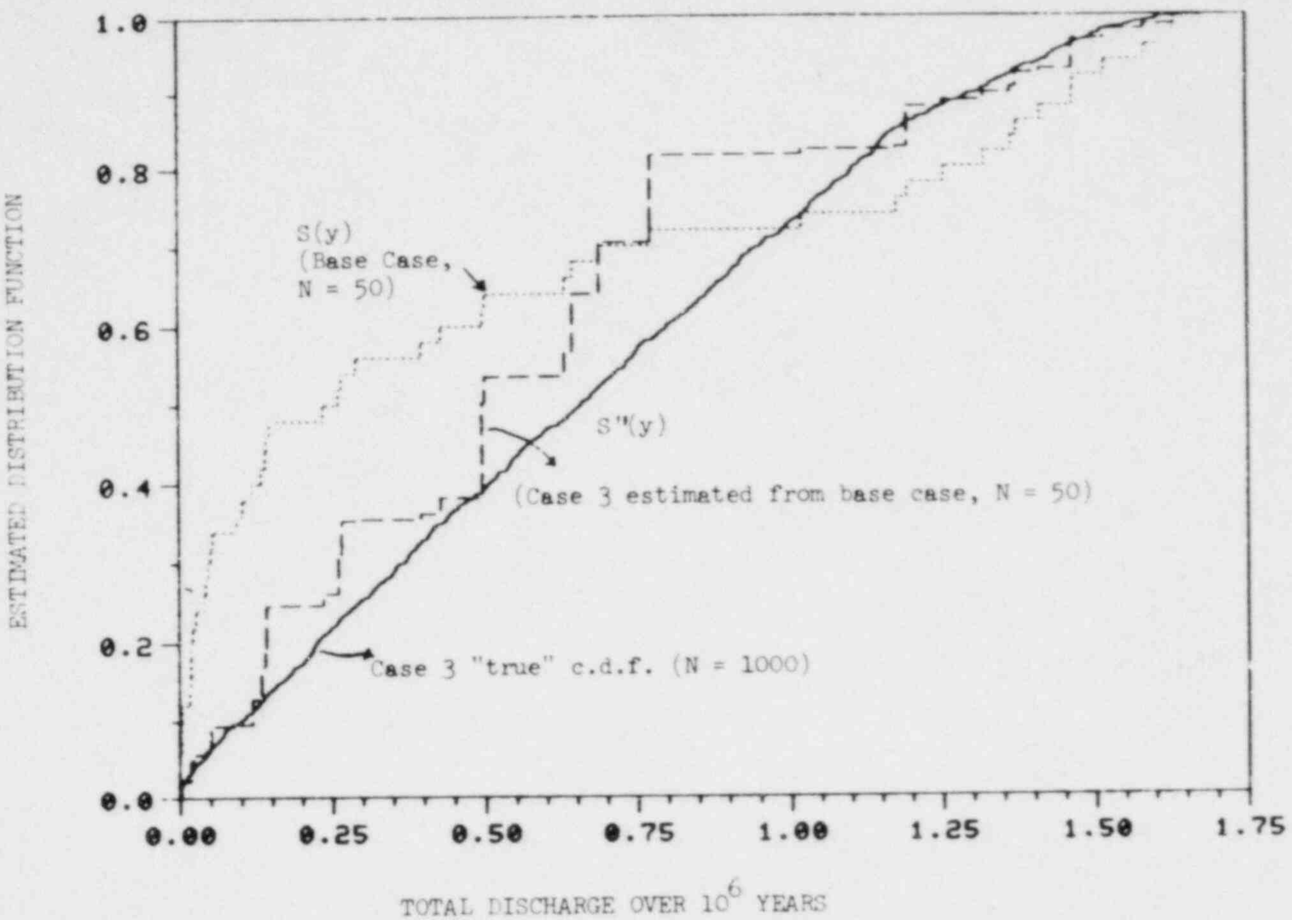
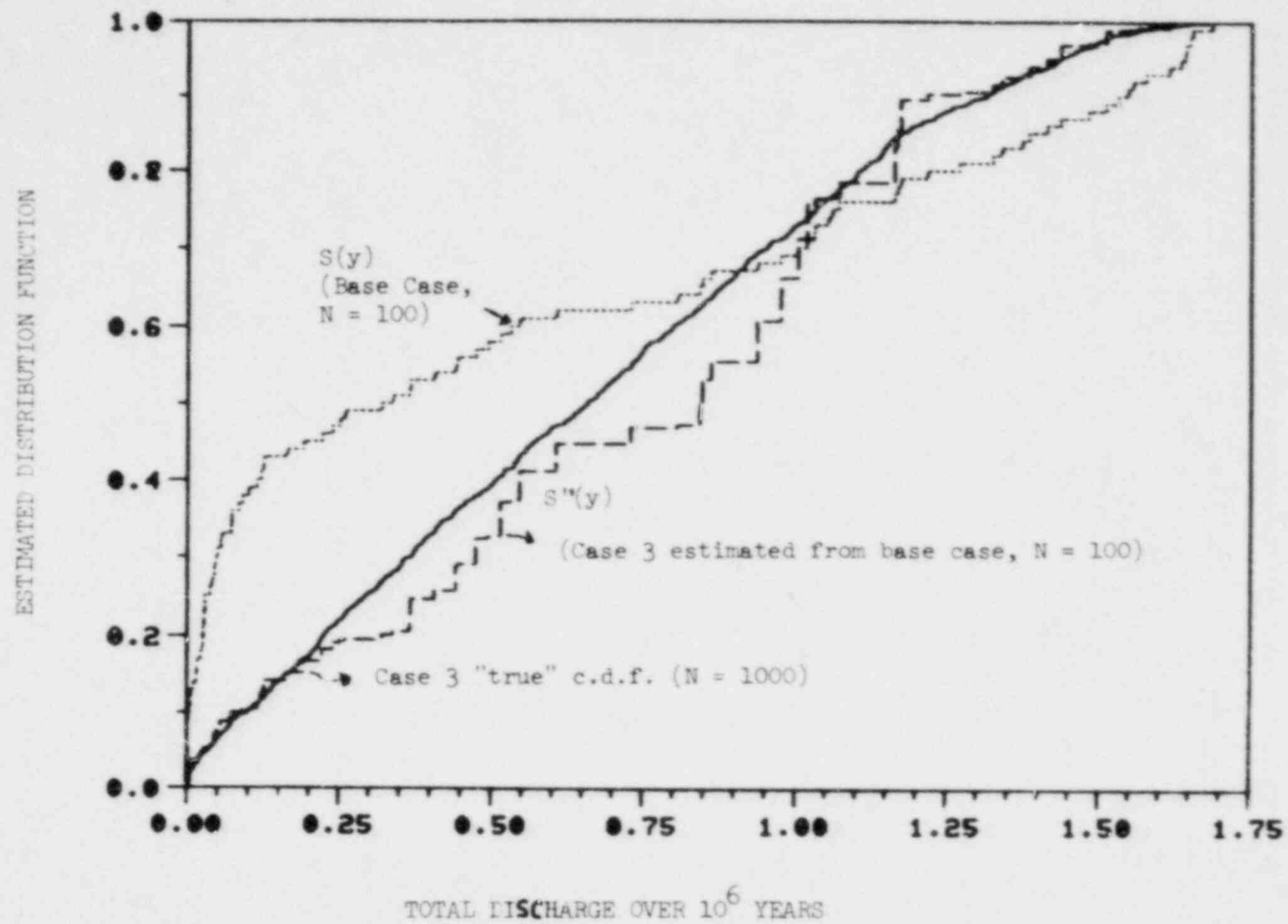


FIGURE 4.16. ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION G(y) FOR THE BASE CASE (N = 100) AND CORRESPONDINC CASE 3 ESTIMATE. ALSO INCLUDED FOR COMPARISON IS THE CASE 3 "TRUE" CURVE BASED ON N = 1000.



is used in Figures 4.15 and 4.16. The sum of the cell probabilities, $S'(\infty)$, equals 0.9914 for N = 50 and 1.1221 for N = 100. Division by $S'(\infty)$ has negligible effect for N = 50, as it did for N = 200 where $S'(\infty)$ was .9869. There is no reason to make a similar adjustment in Cases 1 or 2 because $S'(\infty)$ always equals 1 when only one input distribution function is altered.

An interesting question may be discussed at this time. Suppose someone is interested in estimating the output distribution function under four different sets of assumptions on the input distributions, and can run only 200 points. Would it be more informative to run 200 points under one set of conditions and estimate the other distribution functions using the methods of Section 3, or run 50 points under each of the four sets of conditions? In this example it appears that the former procedure results in better estimates, as seen by comparing Figures 4.6, 4.7, and 4.8 with Figure 4.9. These results depend heavily upon the fact that the weights were not substantially different in the four cases examined. More substantial changes in the weights would probably favor the second procedure. However, if more than 4 cases were to be considered the former procedure may be preferred because of its greater flexibility.

4.6 A Comparison of Latin Hypercube Sampling with Random Sampling and Replicated Latin Hypercube Sampling

In the previous subsection the sample sizes were varied to see what kind of results may be obtained with smaller sample sizes. In each case a sample size was fixed, without any prior information, and a single sample of that size was obtained. This approach may be unsatisfactory for some applications. The user may prefer to use a much smaller sample size, and repeat the procedure several times as a replicated Latin hypercube sampling (RLHS) procedure. For example, rather than obtaining a single LHS with 100 runs, the user may wish to proceed in a sequential manner, with samples of size 10, stopping when satisfactory results are obtained or when funds are exhausted. Results of a comparison of the two procedures are given in this subsection. Both LHS and RLHS are compared with simple random sampling, to provide a point of reference more familiar to the reader. The bases of comparison include point estimates of the mean, variance, and standard error of the estimator of the mean, in addition to comparisons of the resulting estimates of the entire output distribution function. A different isotope is used in this section so the c.d.f. of the output random variable will not necessarily resemble the c.d.f.'s of Section 4.1. Comparisons are first made under the base-case input distribution assumptions of Section 4.1, and then under the changed distribution assumptions considered in Section 4.4 for case 3.

Estimators $\hat{\mu}$ and $\hat{\sigma}^2$ for the mean and variance of Y have already been introduced for Latin hypercube sampling in Equations (2.79) and (2.80) respectively. The estimator of the c.d.f. is given in Equation (2.24). These estimators are easily adjusted for changed input distributions by replacing the former cell probabilities with new cell probabilities, as explained in Section 3. To obtain estimates of the standard errors of these three estimators, the sampling procedure is repeated 50 times. The means and standard deviations of the 50 values of $\hat{\mu}$, $\hat{\sigma}^2$ and S(y) are given in Table 4.3 and Figures 4.17 through 4.20. Note that N = 100.

The replicated Latin hypercube procedure involved 10 subsamples of size 10 each. A subsample of size 10 is a Latin hypercube sample, for which $\hat{\mu}$,

$\hat{\sigma}^2$ and S(y) are computed. The procedure is repeated 10 times and the arithmetic averages of the various estimates are used as estimators. This procedure has one distinct advantage over a single Latin hypercube sample of size 100; the standard error of each estimator may be estimated by computing standard deviation of the 10 observations on $\hat{\mu}$, $\hat{\sigma}^2$ and S(y) respectively, and dividing by $\sqrt{10}$. The sample means of the 50 estimated squared standard errors of $\hat{\mu}$ are also reported in Table 4.3.

Estimates of the mean, variance and distribution function for random samples of size 100 follow the classical lines. In addition, the standard error of $\hat{\mu}$ is estimated in the classical manner, by $(\hat{\sigma}^2/100)^{1/2}$. The actual observed means and standard errors of these estimators, over 50 repetitions, are presented in Table 4.3. When the joint input density function is changed from $f(\underset{\sim}{x})$ to $q(\underset{\sim}{x})$, each observed output $Y = h(\underset{\sim}{X})$ is associated with a weight $W_i = q(\underset{\sim}{X})/f(\underset{\sim}{X})$. Then the new estimators are

$$\hat{\mu}_0 = \frac{1}{N} \sum Y_i W_i \tag{4.5}$$

$$\hat{\sigma}^2_0 = \frac{1}{n-1} \sum (Y_i - \hat{\mu}_0)^2 W_i \tag{4.6}$$

and

$$S^2_0 = \frac{1}{n} \sum u(y - Y_i)\, W_i . \tag{4.7}$$

In this way estimates of the mean, variance, and distribution function can be obtained from the original random samples without rerunning the code. The means and standard errors of these estimators are reported in Table 4.3 and Figures 4.17 through 4.20.

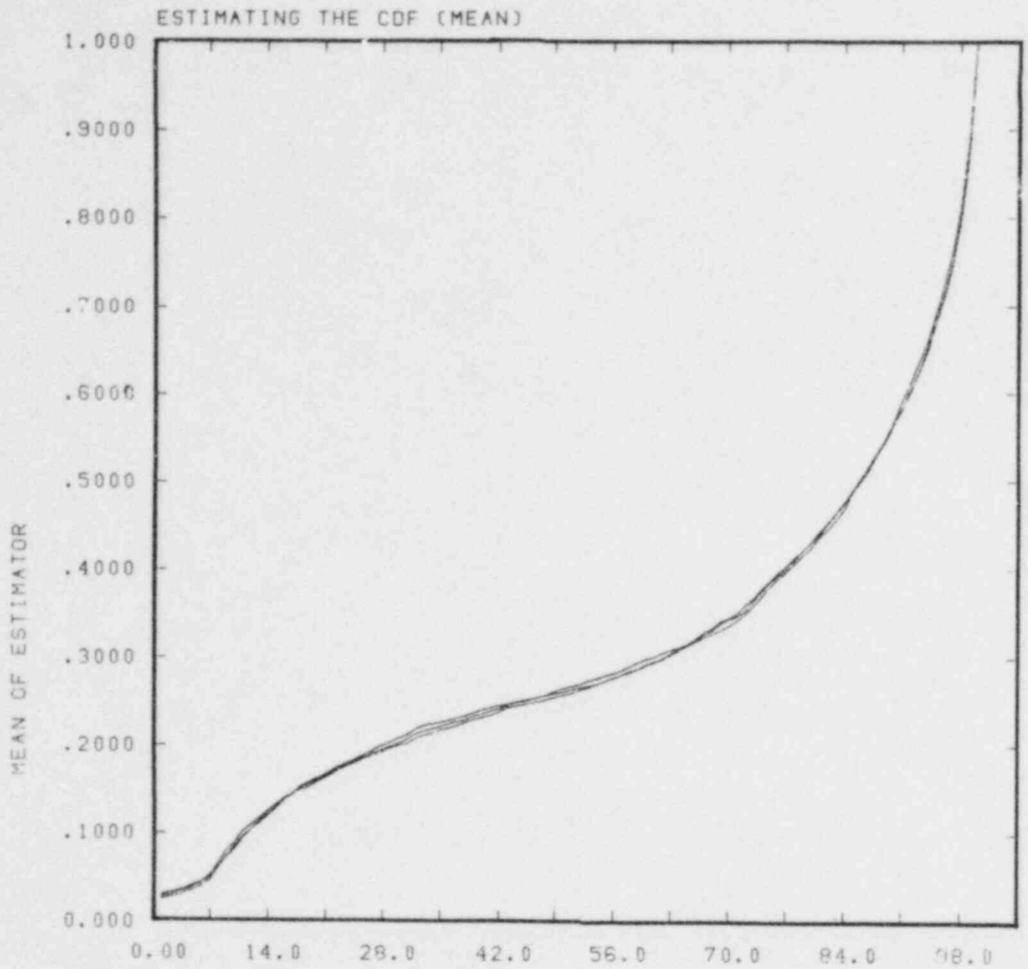
An examination of Table 4.3 reveals many interesting comparisons of the three sampling procedures. Latin hypercube sampling provides, as expected, the best estimates of the mean and variance of the output Y. The standard error of $\hat{\mu}$ increases about 64% for replicated LHS, while random sampling has about five times the standard error of LHS, in the original distributions case. For changed input distributions the increases in standard error are not as pronounced, but still present.

TABLE 4.3

Means and Standard Errors of 50 Observations on Estimators of E(Y), Var(Y), and Var($\hat{\mu}$), using Latin Hypercube Sampling, Replicated Latin Hypercube Sampling, and Random Sampling.

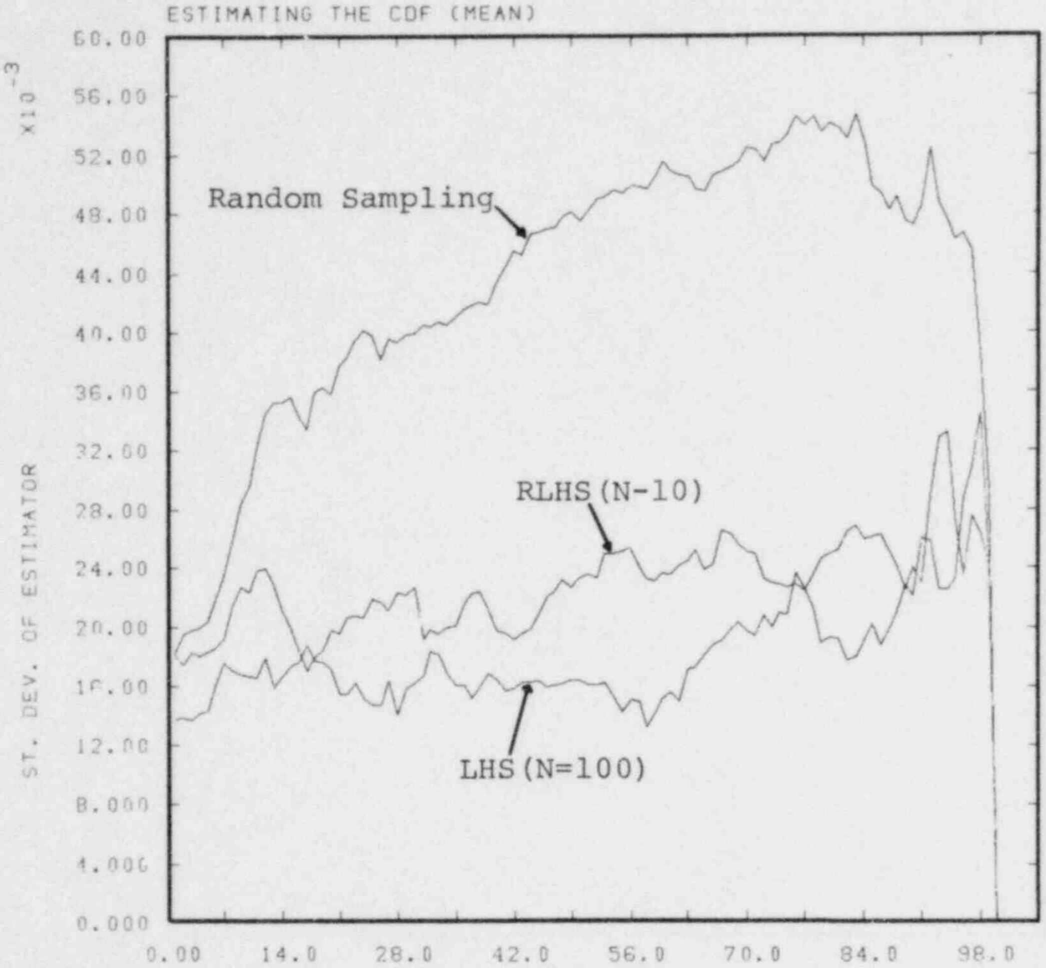
	LHS (N=100)	RLHS (N=10) 10 replications	Random Sample (N=100)
$\hat{\mu}$			
mean	0.6999	0.6972	0.7003
standard error	0.00748	0.01225	0.03556
$\hat{\sigma}^2$			
mean	0.1090	0.1107	0.1103
standard error	0.00588	0.00800	0.01369
$\hat{Var}(\hat{\mu})$			
mean	unobservable	.00011	.00103
(adjusted for changed input distributions)			
$\hat{\mu}$			
mean	0.8735	0.8746	0.8886
standard error	0.06554	0.09000	0.14165
$\hat{\sigma}^2$			
mean	0.0274	0.0712	0.0435
standard error	0.00427	0.02765	0.02646
$\hat{Var}(\hat{\mu})$			
mean	unobservable	0.00561	0.00435

Table 4.3 also shows that the standard error for estimates of the



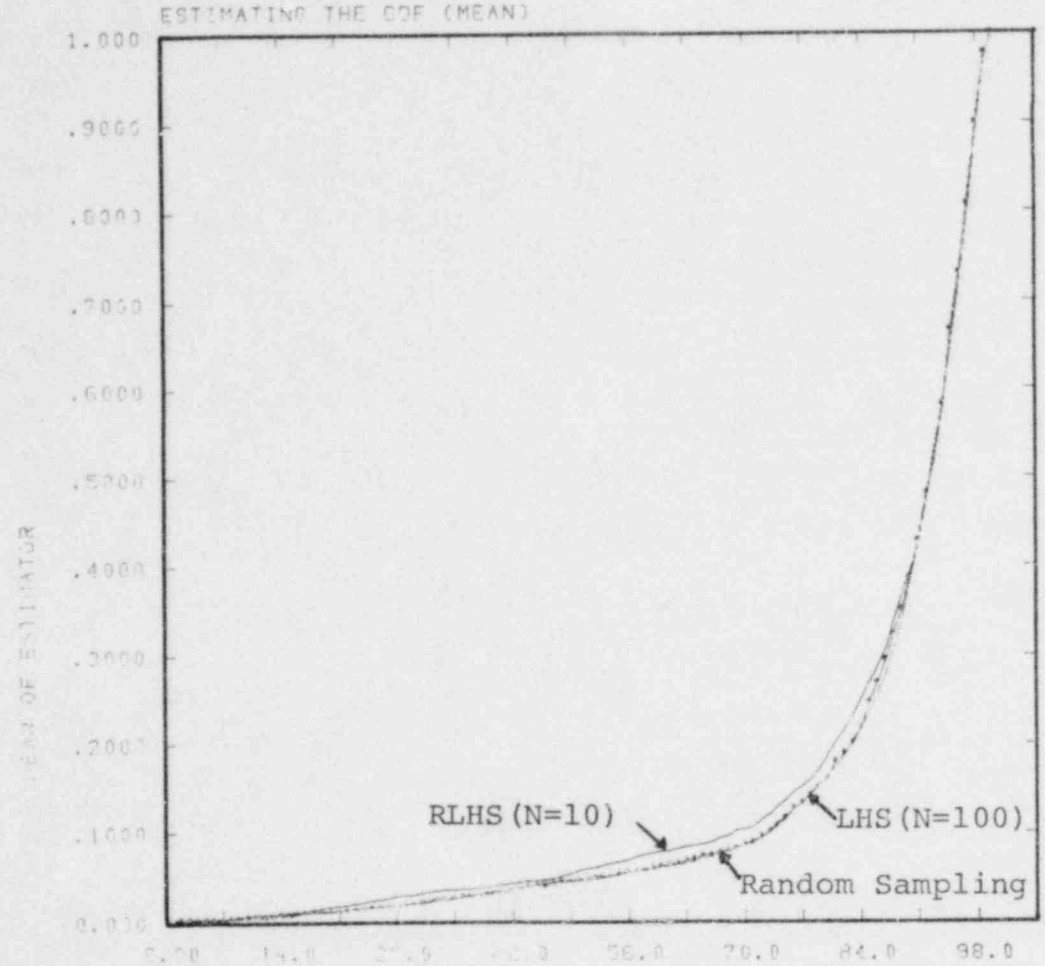
TOTAL DISCHARGE RELATIVE TO SOURCE (PERCENT)

FIGURE 4.17. THE MEAN OF 50 ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION FOR THE BASE CASE USING LHS (N=100), RLHS (N=10, 10 REPS), AND RANDOM SAMPLING (N=100)



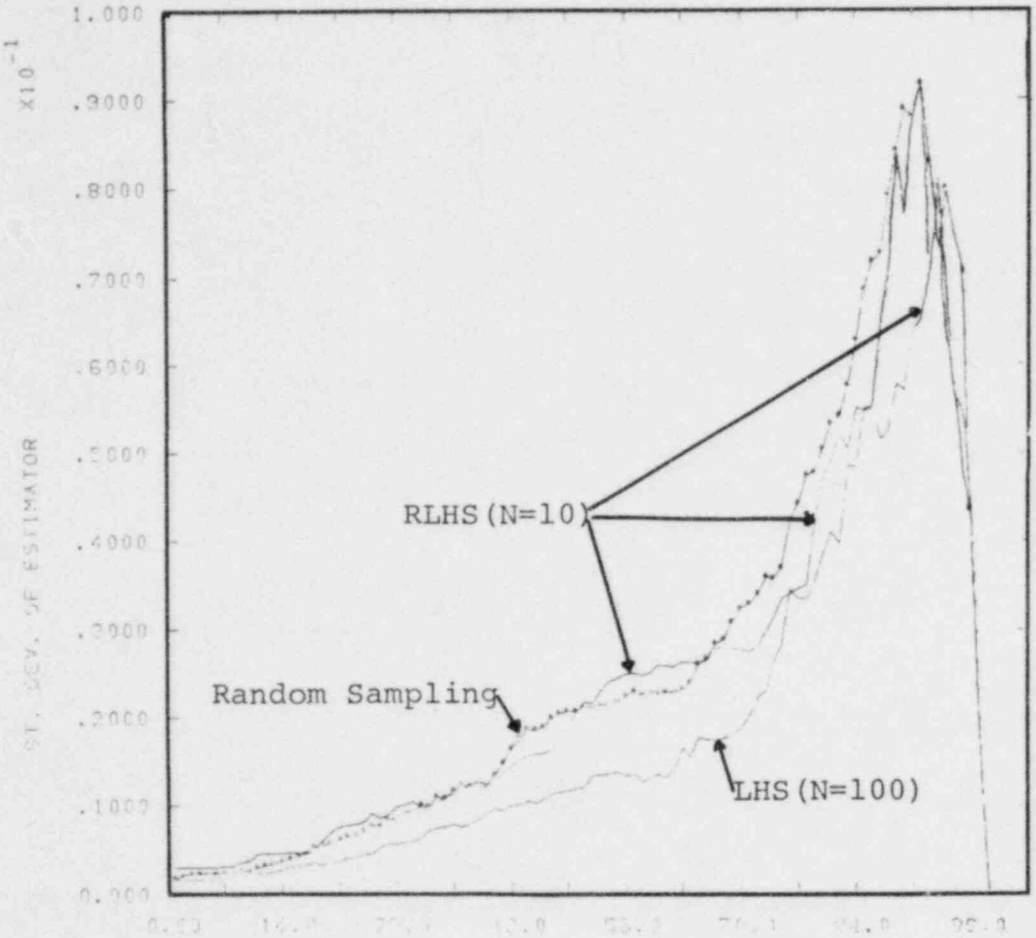
TOTAL DISCHARGE RELATIVE TO SOURCE (PERCENT)

FIGURE 4.18. THE STANDARD DEVIATION OF 50 ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION FOR THE BASE CASE USING LHS (N=100), RLHS (N=10, 10 REPS) AND RANDOM SAMPLING (N=100)



TOTAL DISCHARGE RELATIVE TO SOURCE (PERCENT)

FIGURE 4.19. THE MEAN OF 50 ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION FOR CASE 3 USING LHS (N=100), RLHS (N=10, 10 REPS) AND RANDOM SAMPLING (N=100)



TOTAL DISCHARGE RELATIVE TO SOURCE (PERCENT)

FIGURE 4.20. THE STANDARD DEVIATION OF 50 ESTIMATES OF THE OUTPUT DISTRIBUTION FUNCTION FOR CASE 3 USING LHS (N=100), RLHS (N=10, 10 REPS) AND RANDOM SAMPLING (N=100)

variance are smallest with Latin hypercube sampling and largest with random sampling under the original input distributions. When the input distributions are changed, the variance estimates become unstable, exhibited by the apparent bias and very large variability of the estimator, for all but the LHS.

The estimate of the standard error of $\hat{\mu}$, which is useful for forming confidence intervals, is obtained for each of the 50 repetitions of RLHS and random sampling. These estimates appear to be unbiased, as shown by the mean of $\text{Var}(\hat{\mu})$, which is approximately equal to the square of the actual observed standard error of the 50 values of $\hat{\mu}$.

Replicated Latin hypercube sampling appears to be a viable alternative to Latin hypercube sampling when some estimate of the standard error of the estimator is desired. RLHS appears to provide smaller standard errors than random sampling in most cases, but not as small as LHS. Some increase in standard error, over LHS, accompanies the privilege of obtaining an estimate of that standard error. RLHS is a more general sampling plan than the other two procedures examined. That is, when RLHS has only one replication it reduces to LHS. When RLHS consists of N replicated samples of size one each, it is random sampling.

The estimates of the output distribution function follow the same pattern as the point estimates just discussed. The average of the 50 empirical c.d.f.'s is shown in Figure 4.17 for all three sampling schemes. Obviously they are all unbaised and estimate the same function. The standard errors of those estimators are graphed in Figure 4.18. The c.d.f. estimate furnished by LHS is the best, as measured by its standard error. The RLHS procedure shows a slight increase in standard error, while random sampling produces 2 or 3 times the standard error of the LHS procedure.

The variance of the empirical c.d.f. is estimated using S(y)(1-S(y))/100 using random sampling. The sample variance of the ten estimators obtained in RLHS, divided by 10, provides an estimate of the variance of the S(y) obtained from RLHS. Both of these quantities were obtained for each of the 50 replications. Their mean values agreed with the actual observed standard errors, and so are not reported here. The conclusions are the standard error of the estimate S(y) can be estimated using RLHS and random sampling; the standard error is about twice as large in this example using random sampling; but the smallest standard error belongs to LHS, although no estimate of that standard error is obtainable from a single sample.

The estimates of the c.d.f., when the input distributions are changed, were obtained for each of the three sampling procedures, for the 50 runs. The averages for the three procedures are in remarkable agreement (Figure 4.19). The differences in their standard errors, evident in Figure 4.18 for the original distribution functions, have narrowed considerably in this case. It is not clear which procedure has the smallest variance, although LHS seems to be consistently better than random sampling. The net effect of having unequal cell probabilities seems to be one of narrowing the differences between Latin hypercube sampling and random sampling.

5. A COMPARISON OF SCENARIOS

The methods of sections 2 and 3 can be applied to a variety of situations where decisions must be made in the face of uncertainty. That is, several different strategies may be compared simultaneously using these methods, or several models with substantial differences among them may be compared. In this section we demonstrate these methods with an application involving a single basic model which has a multitude of substantial variations called scenarios.

The physical quantities involved in the assessment of any particular site which is selected as a candidate for geologic disposal of radioactive waste will have uncertainties associated with them. The last section discussed how these uncertainties arise from lack of knowledge about probability distributions associated with input variables. Another important source of uncertainty is introduced by an inability to predict exactly what conditions will exist at the site in the long range future. For example, if future generations lose administrative control over the site then exploratory drilling for minerals and water could take place. A drill hole through or near the depository could establish hydraulic communication between the depository and either the underlying or overlying aquifer.

The above example is by no means the only way by which a release of radionuclides from the depository could take place. However, rather than discussing future conditions we will introduce first the concept of a scenario. A scenario is a set of conditions which could exist at or near the depository. The conditions may or may not eventually lead to a discharge of radionuclides into the environment (biosphere). For a risk assessment of a particular site to be credible a large number of scenarios needs to be examined (perhaps hundreds). Computer time considerations will not allow for an extensive investigation of every possible scenario. Yet, the scenarios need to be evaluated and compared on the basis of their output random variables. In this section we explain how one might accomplish such a comparison with a limited number of computer runs. We also show how the methods of the previous sections can be used to assess the effects of different input distribution assumptions on scenario comparisons.

5.1 Scenarios and Latin Hypercube Sampling

We assume that in the analysis of an actual site the set of scenarios has been carefully defined through the efforts of experts in various fields such as geology, hydrology, physics, and engineering. Our immediate concern is to obtain enough information on each scenario to enable the "high consequence" scenarios to be identified, equivalent scenarios to be combined into groups, and "low consequence" scenarios to be eliminated from further study. In this way the number of scenarios can be reduced to a manageable number which may be studied more extensively.

Initially a decision must be made on how many computer runs are needed on each scenario, taking into consideration such items as reasonable computer time and the power associated with a test for differences in the scenario output random variable. The number of runs (i.e., the sample size) should be large enough to provide good separation (or grouping) of scenarios, and yet should be within the inherent time and cost constraints. The mechanics of obtaining output observations for a particular scenario require the selection of a set of input vectors. We feel that Latin hypercube sampling provides a viable method of selecting these input vectors such that the desired information can be efficiently obtained. Furthermore, the same set of input vectors is used in each scenario to enable a direct comparison among scenarios. This approach assures that each scenario will be run under exactly the same input conditions, and that any differences observed will be due to scenario differences and not sampling variation.

5.2 The Scenarios Used

For purposes of illustration, we again use the analytical transport model which was introduced in Section 4. Whereas the model was used in Section 4 under fixed conditions (i.e., one scenario), consideration is now given to 9 scenarios. The selection of the 9 scenarios is based more on computational simplicity than any attempt to represent all possible modes by which radionuclides could potentially escape a waste depository. Furthermore, even though no attempt is made here to assign probabilities to these scenarios, we believe that the probabilities would be quite low. The 9 scenarios are described briefly below.

Scenario 1. This scenario represents a U-tube which forms a hydraulic connection between the depository and the overlying aquifer (Figure 5.1). These connecting legs are assumed to have relatively low transmissivity. Such a scenario could result from degradation of materials used to seal access and ventilation shafts to the depository or from exploratory drill holes which penetrate to the depository.

A complete listing of variables with ranges and distributions assumed for this analysis is given in Tables 5.1 and 5.2. Those variables used to calculate radionuclide discharge in the individual scenarios are also identified in Table 5.2.

Scenario 2. This scenario is identical to scenario 1 except that a small portion (1%) of the flow in the overlying aquifer is allowed to discharge at a point A, located 10,000 feet downstream from the repository (Figure 5.1). Such discharge could result from water wells placed into the overlying aquifer (e.g., for irrigation or a municipal water supply).

Scenario 3. Scenario 3 is identical to Scenario 1 except that the con-

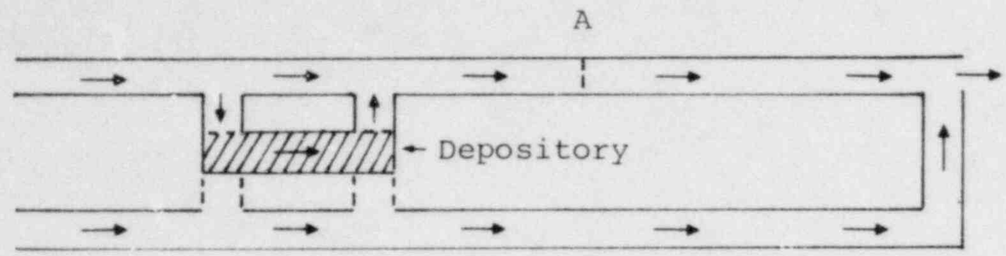


FIGURE 5.1 U-Tube to Overlying Aquifer, Used in Scenarios 1-4

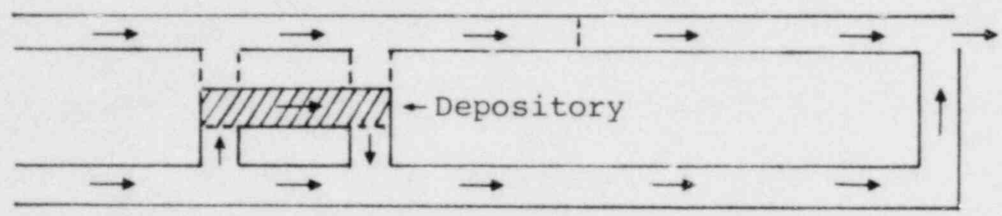


FIGURE 5.2 U-Tube to Underlying Aquifer, Used in Scenario 5

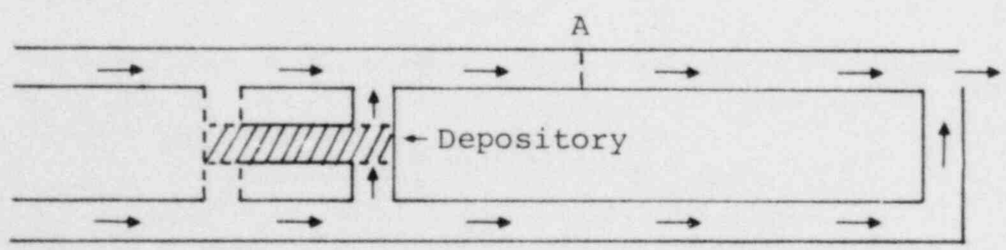


FIGURE 5.3 Connection Through Depository From Underlying to Overlying Aquifer, Used in Scenarios 6-9

TABLE 5.1.

Input Variables Used in the Scenarios

X_1 = hydraulic conductivity of the overlying sandstone aquifer (ft/day)

X_2 = hydraulic conductivity of the underlying sandstone aquifer (ft/day)

X_3 = hydraulic conductivity of the low transmissivity connection (ft/day)

X_4 = hydraulic conductivity of the high transmissivity connection (ft/day)

X_5 = area of the low transmissivity connection (ft^2)

X_6 = area of the high transmissivity connection (ft^2)

X_7 = porosity of the overlying sandstone aquifer

X_8 = porosity of the underlying sandstone aquifer

X_9 = porosity of the low transmissivity connection

X_{10} = porosity of the high transmissivity connection

X_{11} = dispersivity (ft)

X_{12} = distribution coefficient of the sandstone (cm^3/gm)

X_{13} = distribution coefficient of the connection (cm^3/gm)

X_{14} = radionuclide solubility limit (gm/gm)

necting legs are assumed to have high transmissivity. These high transmissivity connections could represent fractures created by mechanical or thermal stresses induced by the presence of the depository and the radioactive waste.

Scenario 4. Scenario 4 is identical to scenario 3 except that 1% of the flow in the overlying aquifer is allowed to discharge at a point A,

TABLE 5.2.

Properties of the Input Variables Used in the Scenarios

Variables	Range	Probability Distribution	Scenarios Using This Variable
X_1	(1,50)	Loguniform	1-4, 6-9
X_2	(1,50)	Loguniform	5
X_3	(1,50)	Loguniform	1, 2, 8, 9
X_4	(1,1000)	Loguniform	3-7
X_5	(1,1000)	Loguniform	1, 2, 8, 9
X_6	$(10^4,10^7)$	Loguniform	3-7
X_7	(.05,.30)	Normal, $\mu=.175, \sigma=.04$	1-4, 6-9
X_8	(.05,.30)	Normal, $\mu=.175, \sigma=.04$	5-9
X_9	(.005,.20)	Loguniform	1, 2, 8, 9
X_{10}	(.0001,.01)	Loguniform	3-7
X_{11}	(20,500)	Uniform	1-9
X_{12}	$(10,10^4)$	Loguniform	1-9
X_{13}	$(10^{-2},10^2)$	Loguniform	1-9
X_{14}	$(10^{-9},10^{-6})$	Loguniform	1-9

located 10,000 feet downstream from the depository.

Scenario 5. This scenario represents the formation of a U-tube to the underlying aquifer (Figure 5.2). The legs of the U-tube are assumed to have high transmissivity. As in scenario 3, these high transmissivity connections could result from fracture formation.

Scenario 6. This scenario represents a hydraulic connection through the depository from the underlying to the overlying aquifer (Figure 5.3). The connecting legs are assumed to have high transmissivity. Such a scenario could result from faulting or fractures through the depository.

Scenario 7. This scenario is identical to scenario 6 except that 1% of the flow in the overlying aquifer is allowed to discharge at a point A, located 10,000 feet downstream from the depository (Figure 5.3).

Scenario 8. This scenario is identical to Scenario 6 except that the connecting legs are assumed to have low transmissivity. Such a scenario could result from exploratory drill holes which penetrate through the depository to the underlying aquifer.

Scenario 9. Scenario 9 is identical to scenario 8 except that 1% of the flow in the overlying aquifer is allowed to discharge at a point A, located 10,000 feet downstream from the depository.

5.3 Ordering of Scenarios by Use of the Friedman Test

Since the discharge rates associated with each of the above scenarios are calculated by using identical input vectors in the computer model as explained in subsection 5.1, a blocking effect is created across scenarios. Therefore, the discharge rates of the scenarios are set up according to a randomized complete block design. The non-normality of the output random variable suggests the use of the nonparametric Friedman test to test whether the scenarios have identical consequences. This test requires that the discharge rates be ranked from 1 to 9 within each block (assigning average ranks in the case of ties). The ranks assigned to each scenario are summed over all blocks as part of the computation for the desired test statistic. This test is sensitive to differences in relative orderings within blocks.

The Friedman test provides a convenient method of distinguishing between scenarios on the basis of their location parameters (i.e., means or medians) on the basis of only a few observations. Other procedures may be more appropriate if other quantities are of interest, such as the 95th percentile or only observations above some critical value. However inferences regarding these other quantities are likely to require much large sample sizes.

Time and cost constraints inherent in large models like SWIFT ordinarily limit the analysis to a small number of blocks. In order to estimate the number of blocks required for a reliable ordering of scenarios, a larger number of blocks can be run on a simplified replacement model such as NWFT. This provides a "proper" ordering of scenarios against which the results from a smaller number of runs may be compared.

To illustrate this point we chose to start with 100 blocks using the NWFT model. Latin hypercube sampling based on equal probability intervals of 1/100 is used for each of the 14 variables listed in subsection 5.2. The corresponding output random variables for each of these 100 input vectors are assigned ranks as explained above with the results shown in Table 5.3.

The results show that within block number 1, scenario number 5 had the largest (rank 9) discharge rate while scenario number 1 had the smallest (rank 1). In block number 2 scenarios 1, 5, and 8 all tied for the smallest discharge, hence they all receive the average rank of (1+2+3)/3 = 2. Likewise, in block 3 scenarios 1, 3, 6, and 8 all receive rank

TABLE 5.3.

Ranks Assigned Within Blocks

Scenario No.	Run (Block) Number 1	2	3	...	100	Sum
1	1	2	2.5	...	1	193.5
2	2	6	6	...	2	462.0
3	5	5	2.5	...	6	461.0
4	6	9	8	...	7	824.0
5	9	2	5	...	5	425.0
6	7	4	2.5	...	8	446.0
7	8	8	9	...	9	814.0
8	3	2	2.5	...	3	284.5
9	4	7	7	...	4	590.0

(1+2+3+4)/4 = 2.5. The test statistic is then based on the sum of the ranks across the row for each scenario.

These rank totals indicate that scenarios 4 and 7 tend to give the highest discharges and scenarios 1 and 8 the smallest. The F statistic computed on the ranks, as recommended by Iman and Davenport (1980) as an alternative form for the Friedman test, is found to be F = 179.8. When this value is compared with tables of the F distribution for 9-1=8 and (9-1)(100-1) = 792 degrees of freedom the significance level associated with this outcome is $<<.00001$. This means that significant differences do exist among the discharge rates for the 9 scenarios. Fisher's least significant difference procedure computed on the ranks, as outlined in Conover (1980), is used to make multiple comparisons. This procedure gives the least significant difference at the 10% level for separation of these rank sums as 37.0. Any group of rank sums must differ by at least 37.0 to be declared significantly different. To demonstrate, the scenarios are ordered according to their rank sums and equivalent groupings are noted.

Scenario	Rank Sum
1	193.5
8	284.5
5	425.0 |
6	446.0 |
3	461.0 |
2	462.0 |
9	590.0
7	814.0 |
4	824.0 |

Since the rank sums of scenarios 1 and 8 differ by 91 (which is more than 37.0) these scenarios are declared to have significantly different discharge rates. However, scenarios 5, 6, 3, 2 are spanned by the measuring stick of 37.00, and hence considered to be in the same group. Likewise, scenarios 7 and 4 can be grouped together.

Since Latin hypercube sampling was used to obtain the input vectors, the empirical distribution function of the output random variable provides an unbiased estimate of the cumulative distribution function for each

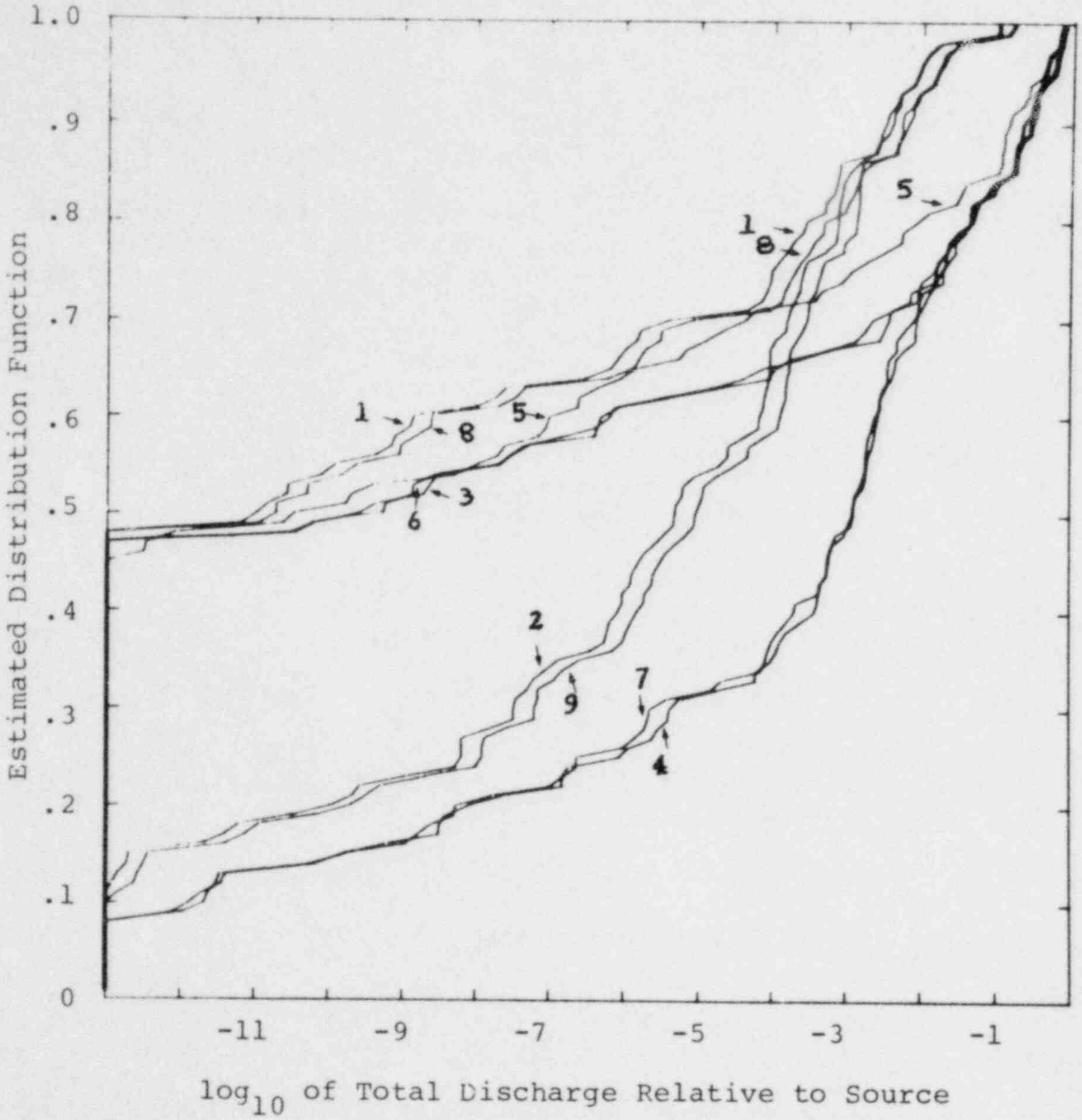


FIGURE 5.4 Estimated c.d.f.'s for Each of the Scenarios

scenario. Graphs of these empirical distribution functions are given in Figure 5.4. These graphs provide additional information pertinent to the interpretation of the above analysis. For example, the output of scenario 1 has been declared significantly lower than the output of scenario 8. From a statistical standpoint this is true as can be seen from Figure 5.4 which shows that the c.d.f. for scenario 1 is always to the left of the c.d.f. for scenario 8. However, from a practical point of view one would be hard pressed to claim that there is any real difference between the output of scenarios 1 and 8. Therefore, one should realize that these procedures have a lot of statistical power to detect scenario differences, but one always has to be aware of the practical interpretation of the results.

We feel that the sample size of 100 is large enough to obtain a reliable ordering of the scenarios. Use of this large sample size was made possible by use of the simplified analytic transport model (NWFT). In the next sub-section the results of ordering the scenarios based on smaller sample sizes are given and compared with the results of this sub-section.

5.4 Scenario Ordering with Smaller Sample Sizes

The results of the ordering of scenarios given in the above sub-section involve large enough sample sizes to give reliable groupings. It would be desirable to reproduce those results using smaller sample sizes and hence work within the realm of available (and feasible) computer time. In order to determine how small a sample might be used, samples of size 2, 3, 4, 5, and 10 were studied. For each sample size the value of the Friedman F statistic and associated significance level is examined to determine if differences exist. If significant differences exist (at the 10% level or less) the scenarios are ordered according to their rank sums and similar groupings are noted with horizontal lines.

Sample size = 2 F = 2.67, Sign. level ≐ .10

Scenario:	1	8	2 = 5	6	3	9	7 = 4	
Rank Sum:	4	6	8	8	9	10	11	17	17

Least significant difference at the 10% level is 7.2

Sample size = 3 F = 2.14, Sign. level ≐ .10

Scenario:	8	6	1 = 3		5	2 = 9		7	4
Rank Sum:	6.5	11.5	12.5	12.5	13	16	16	22	25

Least significant difference at the 10% level is 9.4.

Sample size = 4 F = 6.43, Sign. level ≐ .0001

Scenario:	1	8	5	3 = 6		2	9	4 = 7	
Rank Sum:	9	11	16	17	17	19	25	33	33

Least significant difference at the 10 % level is 8.2.

Sample size = 5 F = 10.9, Sign. level <<.0001

Scenario:	1	8	5	3 = 6		2	9	7	4
Rank Sum:	10	12	19	21	21	27	32	41	42

Least significant difference at the 10% level is 8.4.

Sample size = 10 F = 12.88, Sign. level <<.0001

Scenario:	1	8	5	6	2	3	9	7	4
Rank Sum:	24	30	41	42	47	50	55	79	82

Least significant difference at the 10% level is 1[illegible].[illegible].

Examination of these results shows that these smaller sample sizes do not always give as clear and sharp grouping of the scenarios as is experienced with the larger sample size of the previous subsection. For example with 2 blocks the significance level is large, and although scenarios 7 and 4 show a tendency to produce output values larger than those from the other scenarios, there is still overlap with scenarios 3 and 9. For 5 blocks the significance level is small and scenarios 7 and 4 have separated from the remaining scenarios. Also 1 and 8 (as well as 9) show a tendency to separate from 5, 3, 6, and 2. These results are in excellent agreement with those for 100 blocks. The results for 10 blocks do not show any apparent improvement over 5 blocks. Based upon these results we felt somewhat satisfied using 5 blocks and decided to investigate this sample size further to see what sampling variation might be expected.

Ten additional samples using 5 blocks each produced the following results. The summary below gives only the value of the F statistic, associated significance level and scenario grouping.

Run No. 1 F = 8.93 S.L.<<.0001
Scenario: 1 8 5 2 3 = 6 9 4 7

Run No. 2 F = 8.73 S.L.<<.0001
Scenario: 1 = 8 6 3 = 5 2 9 7 4

Run No. 3 F = 16.77 S.L.<<.0001
Scenario: 1 8 5 6 3 2 9 7 4

Run No. 4 F = 8.75 S.L.<<.0001
Scenario: 1 = 8 3 = 5 6 2 9 7 4

Run No. 5 F = 8.09 S.L.<<.0001
Scenario: 1 8 3 = 6 5 2 9 4 7

Run No. 6 F = 5.68 S.L. ≐ .0005
Scenario: 1 8 5 2 6 3 = 9 7 4

Run No. 7 F = 13.22 S.L.<<.0001
Scenario: 1 8 6 3 2 5 9 7 4

Run No. 8 F = 5.48 S.L. ≐ .0005
Scenario: 1 8 5 6 2 = 3 9 7 4

Run No. 9 F = 5.30 S.L. $\doteq$ .0005

Scenario: 1 8 5 6 2 3 9 7 4

Run No. 10 F = 8.78 S.L. $\ll$ .0001

Scenario: 1 8 5 6 3 2 9 4 7

The absolute ordering of these scenarios is very consistent with 1 and 8 low and 4 and 7 on the high end just above 9. The other four seemed to be mixed somewhat as we might expect since the results of the sample of size 100 indicated no significant difference among these scenarios. These results reinforce our conclusion that a sample of size 5 is adequate to provide an ordering of the scenarios in this application.

5.5 Effect of Input Distribution Assumptions on Scenario Ordering

The Friedman test of subsection 5.3 can be modified by assigning weights to each of the blocks. In this sub-section we use the methods of sections 2 and 3 to generate these weights. In particular these weights reflect the different probability assumptions on key input variables and allow us to estimate a new ordering of scenarios without rerunning the computer model.

The sensitivity of the computer model's output to the distribution assumptions on key input variables may be assessed by obtaining a new set of input vectors, drawn from the new probability distributions, and running them through the various scenarios in a repeat performance of the original study. When the code is time consuming and expensive to run, this approach is sometimes not feasible. If the new input distribution functions do not differ radically from the previous assumptions, the approach outlined in Section 3 may be used to estimate the output distributions and the (possibly) new ordering of the scenarios. The new probability weights associated with the input vectors are used in a modified form of the Friedman test (as appears in Quade (1979) and Conover (1980)) to order the scenarios.

An outline of the procedure will now be given for the convenience of the reader. Let X_{ij} represent the discharge rate for the $i\underline{th}$ run in scenario j where $i = 1, \ldots, b$ and $j = 1, \ldots, k$. Hence, the experiment again consists of b blocks and k treatments. As before, the X_{ij} are assigned ranks $R(X_{ij})$ from 1 to k within each block. The modification of the Friedman test starts by assigning ranks to each block based on the sample range which is obtained from the original data.

$$\text{Range in block } i = \max_j (X_{ij}) - \min_j (X_{ij}).$$

Let $Q_1, Q_2, \ldots, Q_b$ represent the ranks (weights) assigned to each block $1, 2, \ldots, b$ respectively. Next, replace each $R(X_{ij})$ with the product S_{ij}, where

$$S_{ij} = Q_i \left(R(X_{ij}) - \frac{k+1}{2}\right). \tag{5.1}$$

This amounts to weighting each ranked observation by the relative size of the sample range of the block containing the observation.

Let us comment briefly on the rationale behind this weighting. Under the null hypothesis of no treatment differences, each assignment of ranks within a block is equally likely. However, if differences do exist among treatments, they are more easily identified in those blocks with the largest range (corresponding to the largest separation of treatments). This procedure makes use of this information by assigning large weights (i.e., the largest Q_i's) to these blocks. For those readers familiar with the nonparametric Wilcoxon signed rank test for paired data (blocks of size 2), this modification of the Friedman test represents a generalization of the Wilcoxon signed ranks test to k treatments.

Analogous to the rank sum used in the computation of the Friedman F statistic, a weighted sum is now found for each treatment as

$$S_j = \sum_{i=1}^{b} S_{ij} \text{ for } j = 1, \ldots, k. \tag{5.2}$$

The modified Friedman F statistic is computed as

$$F = \frac{(b - 1)B}{A - B} \tag{5.3}$$

where

$$B = \frac{1}{b} \sum_{j=1}^{k} S_j^2 \text{ and } A = \sum_{i=1}^{b} \sum_{j=1}^{k} S_{ij}^2. \tag{5.4}$$

The F statistic is approximately F-distributed, so it is compared with a tabled F-distribution having $k - 1$ and $(b - 1)(k - 1)$ degrees of freedom to find the corresponding significance level. For purposes of multiple comparisons, two treatments i and j are declared to be significantly different if the F statistic is first found to be significant at a preselected significance level and

$$|S_i - S_j| > t_{1-\alpha/2} \left(\frac{2b(A - B)}{(b - 1)(k - 1)}\right)^{1/2} \tag{5.5}$$

where the t statistic has (b - 1) (k - 1) degrees of freedom. The right-hand side of the above inequality is called the least significant difference (LSD).

We have indicated that the block weights, Q_i, are obtained by assigning ranks 1 to b to the block ranges. Actually, the procedure is more general than this and the values Q_i may be almost any weight that the user desires. In particular the weights associated with the Latin hypercube input vectors may be used. For the initial set of input vectors these weights are all the same (since the Latin hypercube sample was based on equal probability) and the test reduces to the unweighted Friedman test of Section 5.3. However, if one wishes to investigate the effect of different probability distribution assumptions on the input vectors these weights will be changed to reflect the new probability distribution assumptions on the same range space and intervals as determined by the orginal Latin hypercube sample.

For pu   illustration, the sample of size 10 from the previous subsection is used. As in Section 4, stepwise regression on ranks as given in Iman and Conover (1979) was used to determine the important variables associated with each of the 9 scenarios. This analysis showed variable X_1 to be relatively important in all scenarios except number 5. The original distribution assumed on X_1 given in subsection 5.2 was loguniform on (1, 50). We decided to investigate the effect (if any) on the scenario ordering of changing this assumption to a uniform distribution on the same range space for X_1. The new weights, Q_i, based on this assumption, are given in Table 5.4 for the original Latin hypercube intervals.

The same data used in the example given in the previous subsection for a sample size of 10 are now reanalyzed using the modified Friedman test

TABLE 5.4.

Weights Obtained From Changing the Distribution Assumption on X_1.

Original Interval Assuming X_1 as Loguniform	Original Weight	New Weight, Q_i, Assuming X_1 as Uniform
1 to 1.479	.1	.010
1.479 to 2.187	.1	.014
2.187 to 3.234	.1	.021
3.234 to 4.782	.1	.032
4.782 to 7.071	.1	.047
7.071 to 10.456	.1	.069
10.456 to 15.462	.1	.102
15.462 to 22.865	.1	.151
22.865 to 33.812	.1	.223
33.812 to 50	.1	.330
	1.0	1.000

with weights Q_i instead of the equal probability weights inherent in the unmodified Friedman test. Table 5.5 presents the values of S_{ij} for scenario 1 only, and the resulting value of S_1.

Similar calculations yield the other S_j's from which the following computations are made.

$$B = \frac{1}{10}(45.471) = 4.547$$

$$A = 12.012$$

$$F = \frac{9(4.547)}{12.012 - 4.547}\ 5.482. \text{ Sign. Level} \doteq .0001 \tag{5.6}$$

TABLE 5.5.

Calculation of S_1 Reflecting the Change in Distribution Assumption on X_1

Block No. (i)	$R(X_{i1})$ Ranks Assigned to X_{i1}	Q_i (Arranged by the Rank of Latin Hypercube Interval Assigned to Block i)	$S_{i1} = Q_i(R(X_{i1}) - 5)$
1	3	.021	- .042
2	2	.102	- .306
3	1	.069	- .276
4	5	.010	.000
5	2.5	.032	- .080
6	2.5	.047	- .118
7	3	.151	- .302
8	2	.330	- .990
9	1	.014	- .056
10	2	.223	- .669
			S_1 = -2.839

The least significant difference for multiple comparisons is

$$LSD_{.10} = 1.665 \left(\frac{2 \cdot 10(12.012 - 4.547)}{(10 - 1)(9 - 1)}\right)^{1/2} = 2.398 \qquad (5.7)$$

which leads to the following grouping of scenarios:

Scenario No.:

	5	1	8	2	6	9	3	7	4
S_j:	-3.5009	-2.839	-1.439	-.933	-.214	.546	.992	3.046	3.849

This represents our estimate of the correct order of the scenarios when X_1 has a uniform distribution, obtained without rerunning any input vectors throughout the model. When this ordering is compared with the original ordering for the sample of size 10 in the previous subsection, we note the largest change has been in scenario 5, which has shifted down to be about tied with scenario 1. Other minor changes include scenarios 3 and 9 interchanging positions but still in equivalent groups as is true also for scenarios 6 and 2.

To see how valid the results might be we reran the Latin hypercube sample of size 100 with the distribution of X_1 changed to uniform. Our results are as follows:

Scenario No.:

1	5	8	2	6	3	9	7	4

Rank Sum:

194	290	296	456	482.5	512.5	586	825	858

A comparison of these results with the previous ordering of scenarios with X_1 loguniform, shows that the most dramatic change has been in scenario 5 which has a decrease in its rank sum from 425 to 290. Also, with the loss of scenario 5 from the formerly observed equivalent grouping of scenarios 5, 6, 3, and 2, the ordering of 6, 3, and 2 has changed to 2, 6, and 3. In addition, scenarios 3 and 9 have moved closer together with 3 showing an increase in rank sum from 461 to 512.5, while 9 shows a slight decrease from 590 to 586. The results are in good agreement with the predictions made above from a sample of size 10.

The value of the F statistic associated with this modification of the Friedman test is 5.482 while the results on the original sample showed F = 12.88. This leads to the question, "Is there a loss of power associated with this weighted procedure in this application?" The answer is probably "yes," as we found that the original sample of size 5 was unable to provide us with a different ordering and also showed a marked decrease in the size of the F statistic. (This analysis is not shown in this paper.) The reason for this apparent loss of power likely stems from the fact that small samples may provide too large of a gradation on the range space of the input variables with respect to the number of Latin hypercube intervals to allow for much flexibility in changing distribution assumptions. For example, if a normal distribution is assumed for X_1 in the above example with 10 runs, the corresponding 4 largest Q_i's are .483, .267, .147, and .074, which total .971. This means that any decision about scenario differences would be based essentially on the 4 input vectors providing these weights. With a sample of size 5 the problem is further compounded.

5.6 Effect of Input Distribution Assumptions on Risk Assessment

Thus far we have indicated how a scenario ordering might be accomplished and have mentioned that the benefit of such an ordering is that the scenarios have effectively been screened by the grouping of scenarios that produce similar consequences. This reduced set of scenarios can now be investigated more extensively. That is, we would want to make a larger number of computer runs on each of the scenarios resulting from the screening process and then use the weights (probabilities) associated with each scenario to form a risk assessment curve which is based on the combined outputs of the scenarios. In turn this curve can be used to compare against quantiles (standards) which represent "acceptable" levels of risk as defined by various governing agencies. The method of obtaining such a curve is explained in this subsection. Once again the methods of the previous sections are used to determine the effect of different input distribution assumptions with respect to risk assessment.

Consider Y_{ij} as the output of the i^{th} run of scenario j. Let p_j represent an expert judgment of the probability associated with the occurance of scenario j. Compute the mean output for the i^{th} run as

$$\bar{Y}_i = \sum_{j=1}^{k} p_j Y_{ij}, \quad i=1, \ldots, b. \tag{5.8}$$

The $\bar{Y}_i$ can be plotted in the form of an empirical distribution function to provide an estimated risk assessment curve.

The 9 scenarios defined in subsection 5.3 are used for purposes of illustration. For simplicity assume that only 9 scenarios exist (i.e., $\sum_{j=1}^{9} p_j = 1$), that all scenarios are of interest, and furthermore that these 9 scenarios all occur with equal probability (i.e., $p_1 = p_2 = \ldots = 1/9$). In reality there may or may not be exactly 9 scenarios contained in the subset of interest, and they almost certainly would not occur with equal probability. In fact, a scenario which has no discharge (one which we haven't considered in this paper) would most likely have a much large probability associated with it than all other scenarios combined. However, these simplifying assumptions will not affect the general application of the procedure.

The first risk assessment curve is given in Figure 5.5 and is labeled as $X_1 \sim$ Loguniform, n = 100. This curve reflects the pooling together of the output results used to generate the estimated distribution functions in Figure 5.4. A second risk assessment curve given in Figure 5.5 uses the results of subsection 5.5 for n = 100 and $X_1 \sim$ Uniform. A comparison of these two solid curves show them to differ by about a half of an order of magnitude at the medians, but by well over 3 orders of magnitude at the .20 quantiles. The remaining two curves in Figure 5.5 appear as dashed lines and are associated with the sample of size 10 for which weights were given in Table 5.4. That is to say, the curve labeled $X_1 \sim$ Loguniform, n = 10 is the estimated risk curve obtained by pooling together the scenario results in the manner described above for the runs that were initially given in subsection 5.4 for ordering scenarios with a sample of size 10. The agreement with the curve based on a sample of size 100 is good (within sampling variation). The second curve in Figure 5.5 with dashed lines uses the weights from Table 5.4 for $X_1 \sim$ Uniform to provide an estimated risk curve (from the sample of size 10) for the case where $X_1 \sim$ Uniform. Although we have demonstrated this technique on the "subset" of scenarios for n = 10, in practice one would probably follow the above recommendation and use a larger sample size if feasible.

In closing this section we would like to emphasize that the determination of the quantiles of a risk assessment curve can be greatly

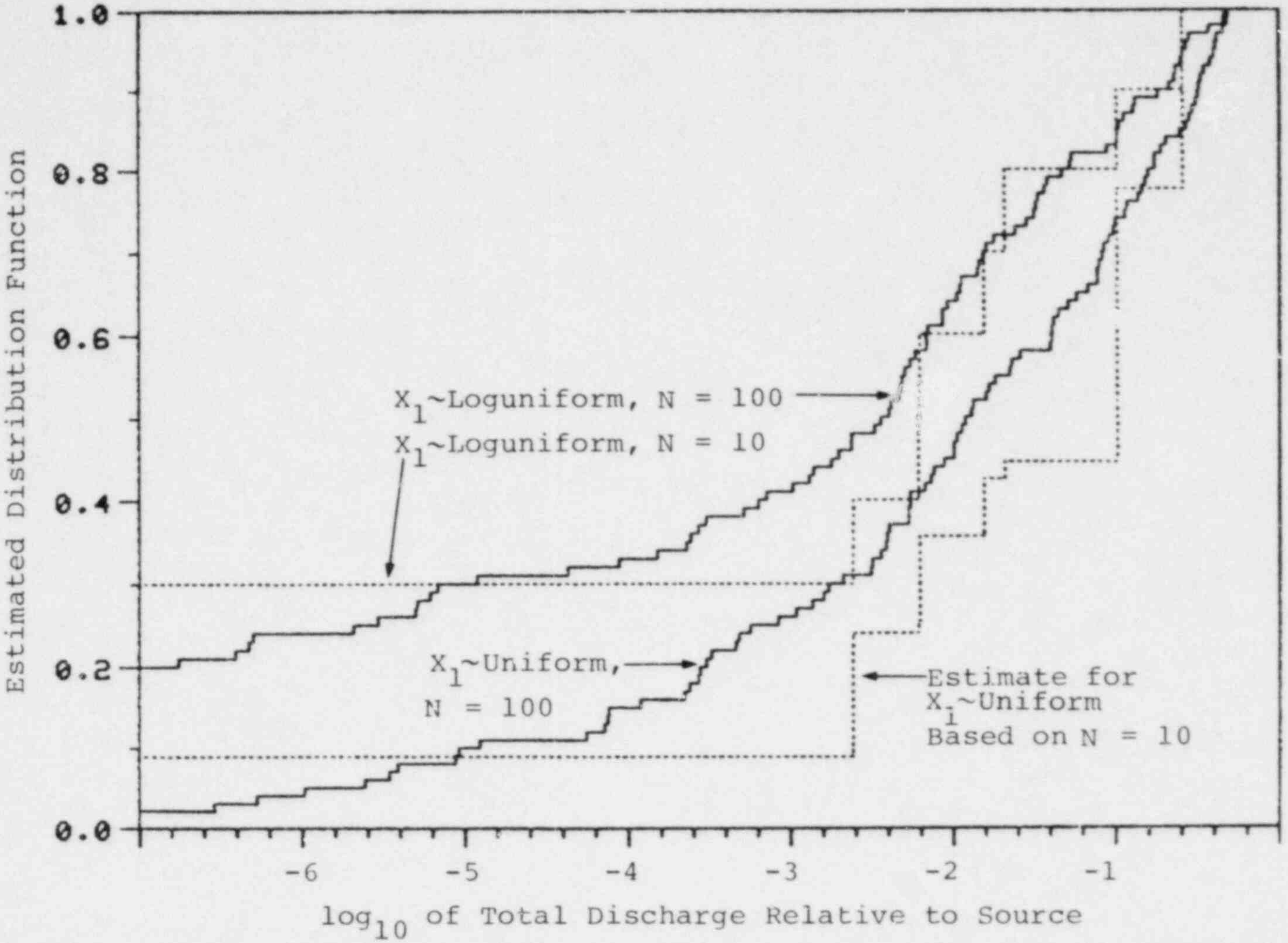


FIGURE 5.5 Estimated Risk Assessment Curves Based on Scenarios 1-9

influenced by distribution assumptions on key input variables as well as probability assumptions associated with scenarios. It is reasonable to expect that in an assessment of a real site disagreement will exist about distribution assumptions (or parameter values associated with particular distributions) and about the assignment of scenario probabilities. We feel that the techniques presented in this paper provide for a great deal of flexibility in handling these questions and do so in an efficient and accurate manner, thus lending credibility to the risk assessment.

6. SUMMARY AND CONCLUSIONS

State of the art modeling efforts have created several difficult and interesting problem areas for individuals concerned with sensitivity studies of the input-output relationships for computer codes which implement these models. Two of these areas which are of particular interest are the following:

(1) The level of complexity of the modeling frequently is based on a series of differential equations which cause the corresponding computer codes to be quite time consuming, perhaps taking several hours for a single run. Hence, a judicious selection procedure for the choice of the values of the input variables is mandated.

(2) A variety of situations require that decisions and judgments be made in the face of uncertainty. The source of this uncertainty may be lack of knowledge about probability distributions associated with input variables or perhaps lack of knowledge about future conditions. At the same time the constraints indicated in (1) make a large number of computer runs under various conditions

prohibitive.

In particular a good selection of values of input variables should make possible the following:

(a) probability related statements, such as those regarding the mean, variance, or cumulative distribution function of the output variable,

(b) estimates that are close to the real values of the quantities being estimated,

(c) an assessment of the relative importance of each input variable,

(d) some means for measuring the sensitivity of the code output with respect to distribution assumptions on the input variable.

If the input-output relationship is monotonic then the genera ation of Latin hypercube sampling provided in this paper provides an inexpensive and reliable way of answering (a) through (d). Latin hypercube sampling has this increased flexibility over other sampling schemes since each input value can be associated with a particular interval defined on the range space of the corresponding input variable. Initially the weights associated with these intervals are all equal as Latin hypercube sampling is usually based on equal probability intervals. However, the weights associated with these intervals can be changed to study the effect of different distributional assumptions on key input variables. Further, these weight changes allow accurate estimates of the output cumulative distribution function to be obtained without making additional (costly) computer runs. In addition these same weights can be used in a modified nonparametric Friedman test in order to examine the effect of different input distributional assumptions on treatment orderings where the treatments may be different hypothesized future conditions, or for that matter different strategies that could be used in a decision making process.

Recommendations concerning sample size requirements are specific to the problem being investigated. For the particular application investigated in this paper we found that a sample size of between 100 and 200 is sufficient to provide good estimates of the empirical distribution functions, under the original assumed input distributions as well as under changed distribution assumptions, provided the changes are not extreme. A much smaller sample size, about 4 or 5, appears to be sufficient to provide a comparison of scenarios. Potential users of Latin hypercube sampling may want to use these sample sizes as guides to their experimental designs, or they may wish to follow our example and use a simplified version of the computer code for a thorough analysis of the sample sizes required for a meaningful analysis.

ACKNOWLEDGEMENTS

Several individuals have made valuable comments and contributions which have led to the final form of this paper. Thanks is due to INTERA Environmental Consultants of Houston for their help in the scenario modeling and subsequent discussions. We also deeply appreciate the review given this document in its development stages at Sandia Laboratories by John M. Wiesen, Richard R. Prairie and especially Robert G. Easterling, whose comments led to the inclusion Section 4.6.

BIBLIOGRAPHY

Campbell, J.E., Dillon, R.T., Tierney, M.S., Davis, H.T., McGrath, P.E., Pearson, F.J., Shaw, H.R., Helton, J.C., and Donath, F.A. (1978). Risk Methodology for Geologic Disposal of Radioactive Waste: Interim Report. SAND78-0029, Sandia Laboratories, Albuquerque, NM.

Campbell, J.E., Kaestner, P.C., Langkopf, B.S., and Lantz, R.B. (1980). Risk Methodology for Geologic Disposal of Radioactive Waste: The Network Flow and Radionuclide Transport (NWFT) Model. SAND79-1920, Sandia Laboratories, Albuquerque, N.M.

Conover, W.J. (1975). On a Better Method for Selecting Input Variables. Unpublished manuscript.

Conover, W.J. (1980). Practical Nonparametric Statistics, 2nd ed., New York: John Wiley and Sons, Inc.

Dillon, R.T., Lantz, R.B., and Pahwa, S.B. (1978). Risk Methodology for Geologic Disposal of Radioactive Waste: The Sandia Waste Isolation Flow and Transport (SWIFT) Model. SAND78-1267, Sandia Laboratories, Albuquerque, NM.

Iman, R.L., and Conover, W.J. (1979). The Use of the Rank Transformation in Regression. Technometrics, 21, 499-509.

Iman, R.L., and Davenport, J.M. (1980). Approximations of the Critical Region of the Friedman Statistic. Communications in Statistics, A9.

Iman, R.L., Helton, J.C., and Campbell, J.E. (1978). Risk Methodology for Geologic Disposal of Radioactive Waste: Sensitivity Analysis Techniques. SAND78-0912, Sandia Laboratories, Albuquerque, NM.

McKay, M.D., Conover, W.J., and Beckman, R.J. (1979). A Comparison of

Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics, 21, 239-245.

McKay, M.D., Conover, W.J., and Whiteman, D.E. (1976). Report on the Application of Statistical Techniques to the Analysis of Computer Codes. Informal Report, LA-NUREG-6526-MS, Los Alamos Scientific Laboratory, Los Alamos, NM.

Quade, D. (1979). Using Weighted Rankings in the Analysis of Complete Blocks with Additive Block Effects. Journal of the American Statistical Association, 74, 680-683.

LIST OF SYMBOLS

Page	Symbol	Meaning
18	(a,b)	the interval from a to b
18	a_k	arbitrary nonnegative constants
103	A or B	a sum of squares used in the weighted Friedman test
102	b	the number of blocks in the weighted Friedman test
97	c.d.f.	cumulative distribution function
26	$Cov(X,Y)$	the covariance of X and Y; i.e., $E(XY) - E(X)E(Y)$
11	$E(\cdot)$	the expected value (mean) of the quantity within the parenthesis
11	$E(\cdot \mid C)$	the conditional mean of the quantity in parenthesis before the $\mid$, given C
11	$f(\underset{\sim}{x})$	the multivariate density function of $\underset{\sim}{X}$
11	$f_n(\underset{\sim}{x})$	the conditional density function of $\underset{\sim}{X}$, given $\underset{\sim}{X}$ is in $S_{\underset{\sim}{n}}$
95	F	the statistic used in analysis of variance
11	$g(Y)$	an arbitrary function of Y, used in defining a class of estimators
50	$G(y)$	the distribution function of Y; i.e. $P(Y \leq y)$
8	$h(\underset{\sim}{X})$	the deterministic function defined within the computer code
8	$I_{k,n}$	interval from which the nth observation on variable X_k is sampled
102	k	the number of scenarios, when used in the weighted Friedman test
8	K	the number of input variables (components of X)
78	LHS	Latin hypercube sampling
104	LSD	least significant difference
107	LSD_α	the LSD comparison made at a level of significance = α

Page	Symbol	Meaning
16	M	the total number of ordered N-tuples U
9	$\underset{\sim}{n}$	a vector of indices, locates the cell $S_{\underset{\sim}{n}}$
8	N	the number of observations
42	NWFT	Network Flow and Radionuclide Transport model
109	p_j	the probability associated with scenario j, when used in the weighted Friedman test
9	$p_{\underset{\sim}{n}}$	probability associated with the cell $S_{\underset{\sim}{n}}$
8	$p_{k,n}$	probability associated with interval $I_{k,n}$
11	$P(\cdot)$	the probability of the event stated within the parentheses
18	q	an index to denote a particular cell S_q
33	$q_{\underset{\sim}{n}}$	the probability associated with $S_{\underset{\sim}{n}}$ when $q(x)$ is the density of $\underset{\sim}{X}$
33	$q(\underset{\sim}{x})$	a density function of $\underset{\sim}{X}$, different than $f(\underset{\sim}{x})$
33	Q	a function of Y, used as a general estimator when $q(\underset{\sim}{x})$ is the density of $\underset{\sim}{X}$
102	Q_i	the rank of the range of block i, when used in the weighted Friedman test
25	r	an index to denote a particular cell S_r
25	R	the restricted space of all pairs of cells which have no cell coordinates in common
102	$R(X)$	the rank of the random variable X
78	RLHS	replicated Latin hypercube sampling
29	s^2	the weighted sample variance of Y
9	$S_{\underset{\sim}{n}}$	a hypercube in the sample space of X
102	s_{ij}	a score, or quantity, assigned to a random variable, when used in the weighted Friedman test
103	S_j	the sum of scores in scenario j, when used in the weighted Friedman test
13	$S(y)$	an empirical distribution function

Page	Symbol	Meaning
34	$S'(y)$	an empirical distribution function, when $q(\underset{\sim}{X})$ is the density of $\underset{\sim}{X}$
67	$S''(y)$	the standardized form of $S'(y)$
14	$S^*(y)$	an empirical exceedance probability function
15	$S^{**}(y)$	a standardized empirical distribution function
100	S.L.	significance level
42	SWIFT	Sandia Waste Isolation Flow and Transport computer program
103	$t_{1-\alpha/2}$	the upper $\alpha/2$ critical value of student's t distribution
10	T	a function of Y, used as a general estimator
13	$u(t)$	an indicator function, = 1 when $t \geq 0$, = 0 when $t < 0$
15	U	an ordered N-tuple of cells $S_{\underset{\sim}{n}}$
16	U^i	values of U with an index i
16	$Var(\cdot)$	the variance of the quantity within the parentheses; i.e., $Var(X)=E(X-E(X))^2$
17	$Var(\cdot \mid C)$	the conditional variance of the quantity within the parentheses, given the condition C
80	$Var(\cdot)$	an estimate of the variance of the quantity in parentheses
8	X_k	individual input variable, k=1,...,K
8	$\underset{\sim}{X}$	vector of input variables
8	$\{\underset{\sim}{X}_n\}$	sample of input vectors, n=1,...,N
8	Y	the output random variable, equals $h(\underset{\sim}{X})$
29	$\bar{Y}$	the weighted sample mean of $Y_1,\ldots,Y_N$
11	ϵ	"is an element of", in set notation
25	μ	the mean of $g(Y)$
25	μ_q	the conditional mean of $Y=h(\underset{\sim}{X})$, given $\underset{\sim}{X}$ is in S_q

Page	Symbol	Meaning
78	$\hat{\mu}$	an estimator of μ
78	$\hat{\sigma}^2$	an estimator of the variance of Y
11	Σ	the summation symbol
12	$\int$	the integral symbol
110	$\sim$	"is distributed as", distribution notation
95	$<<$	"much less than"
35	$\doteq$	"approximately equal to"

Distribution:

U. S. Nuclear Regulatory Commission
NRC Standard Distribution GF (310 copies)
Division of Document Control
Distribution Services Branch
7920 Norfolk Avenue
Bethesda, MD 20014

Probabilistic Analysis Staff (36)
Office of Nuclear Regulatory Research
U. S. Nuclear Regulatory Commission
Washington, DC 20555
Attn: M. Cullingford (35)
F. Rowsome

Division of Safeguards, Fuel Cycle and Environmental Research (2)
Office of Nuclear Regulatory Research
U. S. Nuclear Regulatory Commission
Mail Stop 1130SS
Washington, DC 20555
Attn: F. Arsenault
C. Jupiter

High Level and Transuranic Waste Branch (3)
Division of Fuel Cycle and Material Safety
Office of Nuclear Material Safety and Standards
U. S. Nuclear Regulatory Commission
Washington, DC 20555
Attn: J. Malaro
R. Boyle
S. Scheurs

U. S. Geologic Survey (2)
U. S. Department of Interior
Denver Federal Center
Denver, CO 80225
Attn: D. B. Grove
L. F. Konikow

The Analytical Sciences Corporation
6 Jacob Way
Reading, MA 01867
Attn: J. W. Bartlett
C. Koplik

INTERA Environmental Consultants, Inc. (5)
11511 Katy Freeway, Suite 630
Houston, TX 77079
Attn: R. B. Lantz (4)
D. Ward

John Buckner
E. I. Dupont
Savannah River Laboratory
Aiken, SC 29801

Hans Haggblom
Studsvik Energi Teknik AB
S-611 82 Nykoping
Sweden

Donald E. Wood
Rockwell Hanford Operations
202-S Bldg.
200 West Area
P. O. Box 800
Richland, WA 99352

Larry Rickartsen
Science Applications, Inc.
P. O. Box 843
Oak Ridge, TN 37830

Dan H. Holland, President
New Millennium Associates
1129 State Street, Suite 32
Santa Barbara, CA 93101

V. K. Barwell
Environmental Research Branch
Atomic Energy of Canada Limited
Research Company
Chalk River, Ontario
Canada KOJ1JO

R. Budnitz
Office of Nuclear Regulatory Research
U. S. Nuclear Regulatory Commission
Washington, DC 20555

D. Egan
Office of Radiation Programs (ANR-460)
U. S. Environmental Protection Agency
Washington, DC 20464

R. H. Moore
Applied Statistics Branch
U. S. Nuclear Regulatory Commission
Washington, DC 20555

R. K. Waddell
USGS, MS416
Denver Federal Center
Denver, CO 80225

George Brethauer
USGS
1536 Cole Blvd.
Golden, CO 80225

R. De Batist
Materials Science Department
S.C.K./C.E.N.
Boeretang 200
B-2400 MOL (Belgium)

Lawrence Livermore Laboratory (2)
P. O. Box 808
Livermore, CA 94550
Attn: A. Kaufman, L-156
Dana Isherwood, L-224

David Hodgkinson
Theoretical Physics Division
Bldg. 8.9
AERE Harwell
Oxfordshire OX110RA
England

A. J. Soinski
California Energy Commission
Nuclear Assessments Office
1111 Howe Avenue, MS #35
Sacramento, CA 95820

Science Applications, Inc. (1)
1200 Prospect Street
P. O. Box 2351
La Jolla, CA 92037
Attn: E. Straker

Pierre Pages
Boite Postale No. 48
92260 Fontenay-Aux Roses
France

D. R. Proske
Duetsche Gesellschaft fur Wiederaufarbeitung
von Kernbrennstoffen mbH
Bunteweg 2, 3000 Hannover 71
West Germany

Cathy Fore
Ecological Sciences Information Center
Oak Ridge National Laboratory
P. O. Box X
Oak Ridge, TN 37830

Stephan Ormonde
Quantum Systems, Inc.
P. O. Box 8575
Albuquerque, NM 87198

Lynn Gelhar
Dept. of Geoscience
New Mexico Tech
Socorro, NM 87801

Los Alamos Scientific Lab (5)
Group S1, MS606
Los Alamos, NM 87545
Attn: R. A. Waller
M. D. McKay
R. J. Beckman
M. E. Johnson
G. L. Tietjen

M. Mazumdar
Westinghouse Electric Corp.
Research Laboratories
Beulah Road, Churchill Borough
Pittsburgh, PA 15235

G. Apostolakis
Chemical, Nuclear, and Thermal Eng. Dept.
University of California
5532 Boelter Hall
Los Angeles, CA 90024

W. J. Conover (15)
College of Business Administration
P. O. Box 4320
Lubbock, TX 79409

G. Schwarz
ORNL
P. O. Box X
Oak Ridge, TN 37830

Richard Mensing
L-300
Lawrence Livermore Laboratory
Livermore, CA 94550

Larry George
L-116
Lawrence Livermore, CA 94550

Mark Yang
Department of Statistics
University of Florida
Gainesville, FL 32611

D. J. Barron
Test Planning and Analysis
G. M. Proving Grounds
Milford, MI 48042

D. L. Shaeffer
117 Dana Drive
Oak Ridge, TN 37830

Dave Sylvester
Statistics Program
113 Votey Building
University of Vermont
Burlington, VT 05405

M. M. Arjunan
Energy and Safety Systems
1820 Dolley Madison Blvd.
McLean, VA 22102

S. P. Lan
311 Pittsboro St.
UNC - 256H
University of North Carolina
Chapel Hill, NC 27514

T. C. Wu
3379 Rt. 46, #12N
Parsippany, NJ 07054

David Andress
7813 Garland Avenue
Takoma Park, MD 20012

J. M. Davenport
Dept. of Mathematics
P. O. Box 4319
Lubbock, TX 79409

Dept. of Statistics (2)
Kansas State University
Manhattan, KS 66506
Attn: G. A. Milliken
D. E. Johnson

D. A. Gardiner (2)
Computer Sciences Div., UCND
P. O. Box Y
Oak Ridge, TN 37830

Paul Baybutt
Battelle Columbus Laboratories
505 King Avenue
Columbus, OH 43201

Al Bartolucci
Dept. of Biostatistics
U. of Alabama in Birmingham
Birmingham, AL 35294

L. K. Chan
Dept. of Mathematics
University of Western Ontario
London, Ontario, Canada

D. E. Bennett
Biomedical Division
L-452
Lawrence Livermore Laboratory
Livermore, CA 94550

M. J. Egger
Biostatistics BC107
University of Utah Medical Center
Salt Lake City, UT 84112

Kenneth Gerald
Rockwell International
P. O. Box 464, Bldg. 881
Golden, CO 80401

K. S. Crump
Dept. of Mathematics
Louisiana Tech University
Ruston, LA 71270

L. A. Thibodeau
Dept. of Biostatistics
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115

Lon Hixon
Pritsker and Associates
P. O. Box 2413
West Lafayette, IN 47906

J. A. Hasson
4601 Iris Place
Rockville, MD 20853

Jay Beder
8715 First Avenue
Apt. 23[illegible]C
Silver Spring, MD 20910

Lee Abramson
Nuclear Regulatory Commission
Washington, DC 20555

W. G. Hunter
Dept. of Statistics
University of Wisconsin
Madison, WI 53706

400 C. Winter
1223 R. R. Prairie
1223 R. G. Easterling
1223 R. L. Iman (25)
1223 D. D. Sheldon
1223 I. J. Hall
1418 C. A. Trauth, Jr.
1425 E. E. Ard
1425 J. R. Ellefson
1425 F. W. Muller
1425 F. Spencer
1758 C. E. Olson
4000 A. Narath
4400 A. W. Snyder
4410 D. J. McCloskey
Attn: J. W. Hickman
G. V. Varnado
L. D. Chapman

4413 N. R. Ortiz
4413 J. E. Campbell (20)
4413 R. M. Cranwell
4413 F. Donath
4413 N. C. Finley
4413 J. C. Helton
4413 D. Longsine
4413 R. E. Pepping
4413 W. B. Murfin
4413 S. J. Niemczyk
4413 L. T. Ritchie
4413 A. W. Frazier
4443 D. A. Dahlgren
4510 W. D. Weart
4511 L. R. Hill
4511 G. E. Barr
4514 M. L. Merritt
4514 J. P. Brannen
4530 R. W. Lynch
4536 D. M. Talbert
4537 B. S. Langkopf
4538 R. C. Lincoln
4538 S. Sinnock
4538 H. P. Stephens
4538 L. D. Tyler
4737 P. C. Kaestner
8320 T. S. Gold (2)
8346 C. J. Decarli
3141 T. L. Werner (5)
3151 W. L. Garner (3)
For DOE/TIC (Unlimited Release)
3154-3 R. P. Campbell (25)
For NRC distribution to NTIS
8266 E. A. Aas

☆U.S. Government Printing Office: 1980- 677— 116/526

Org.	Bldg.	Name	Rec'd by *	Org.	Bldg.	Name	Rec'd by *

* Recipient must initial on classified documents.