NASA Contractor Report 159318

NASA-CR-159318
1980 0020931

STATISTICAL ASPECTS OF CARBON FIBER RISK
ASSESSMENT MODELING

Donald Gross, Douglas R. Miller,
and Richard M. Soland

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering
Washington, DC 20052

NASA Grant NSG-1556
July 1980

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

TABLE OF CONTENTS

Page

1. INTRODUCTION 1

2. TREATMENT OF ERRORS AND UNCERTAINTY IN CF RISK ANALYSIS 2

3. TREATMENT OF IMPRECISE KNOWLEDGE IN MODEL CONCEPTION 4

3.1 Use of Expected or Deterministic Values in Place of Random Variables 8
3.2 A General Model of Electronic Equipment Failure 13

4. TREATMENT OF PARAMETER ESTIMATION 21

4.1 Estimating the Parameter of the Exponential Failure Model 21
4.2 Testing for an Exponential Failure Distribution 24
4.3 Confidence Regions for IHR Failure Distribution 41
4.4 Sensitivity of the Model to Estimated Parameters and Distributions 47

5. SIMULATION MODEL DESIGN AND TREATMENT OF SAMPLING ERROR 92

5.1 Simulation Model Design 92
5.2 Pointwise Confidence Bounds 94
5.3 Simultaneous Confidence Bounds 100
5.4 Comparison of Procedures 108

6. CONSERVATIVE RISK PROFILES 108

6.1 Stochastic Dominance 110
6.2 Obtaining Conservative Risk Profiles 111
6.3 Summary 117

REFERENCES 118

LIST OF ILLUSTRATIONS

Figure		Page
2-1	Risk profile methodology	5
2-2	Potential errors affecting the risk profile methodology	6
2-3	Approaches to handling errors in risk profile methodology	7
3-1	An example using average cost per failure vs. distribution of cost	11
4-1	The empirical cumulative distribution function	26
4-2	Data Set #1: avionics terminal blocks, 7.5 mm	28
4-3	Data Set #2: Dynaco amplifier, 7 mm	29
4-4	Data Set #15: transponder GY 70, 10 mm	30
4-5	Data Set #16: transponder GY 70, 3 mm	31
4-6	Comparison of exponential and Erlang-2 (both with mean 1) with an empirical CDF generated from the exponential and then normalized to have mean = 1	33
4-7	Power curve for test of exponential vs. Gamma based on cumulative-total-time-on-test statistic	39
4-8	Comparison of exponential and Erlang CDF's	42
4-9	Empirical CDF and 95% K-S upper bound; 7.5 KV insulator pin, 5 mm fibers, Westinghouse data	44
4-10	Empirical cumulative hazard upper bound $\hat{H}_u$, and its greatest convex minorant $\hat{H}_u$, IHR	45
4-11	95% confidence region for IHR failure distribution for data in Figure 3-1	46
4-12	Washington National Airport risk profile	51

Figure		Page

Figure		Page
4-13	Chicago O"Hare risk profile	53
4-14	Denver risk profile	54
4-15	Detroit risk profile	55
4-16	Houston risk profile	56
4-17	Kansas City risk profile	57
4-18	Kennedy risk profile	58
4-19	La Guardia risk profile	59
4-20	New Orleans risk profile	60
4-21	St. Louis risk profile	61
4-22(a)	Philadelphia risk profile	63
4-22(b)	Philadelphia risk profiel	64
4-23(a)	Los Angeles risk profile	65
4-23(b)	Los Angeles risk profile	66
4-24	National conditional risk profile R with changes in the means of the airport damage distribution (26 airports) . .	69
4-25	National conditioanl risk profile R with changes in the standard deviations of the airport damage distributions (26 airports)	70
4-26	Annual risk profile S with changes in the means of the airport damage distributions (26 airports)	71
4-27	Annual risk profile S with changes in standard deviations of the airport damage distributions (26 airports)	72
4-28	Annual risk profile S with changes in the expected number of accidents in a year (26 airports)	73
4-29	National conditional risk profile R with changes in the means of the airport damage distribution (13 airports) . .	86

Figure Page

4-30 National conditional risk profile R with changes in the standard deviations of the airport damage distributions (13 airports) 87

4-31 Annual risk profile S with changes in the means of the airport damage distribution (13 airports) 88

4-32 Annual risk profile S with changes in the standard deviations of the airport damage distributions (13 airports) 89

4-33 Annual risk profile S with changes in the expected number of accidents in a year (13 airports) 90

5-1 K-S bounds 101

5-2 Desired bounds 104

5-3 Simulated risk profile and confidence regions 107

LIST OF TABLES

Table		Page
4-1	Failure Data: 21 Tests of Equipment	25
4-2	Testing for Exponential vs. IHR using Cumulative-Total-Time-On-Test Statistic	37
4-3	Sensitivity of the National Conditional Risk Profile R to Changes in the Means of the Airport Damage Distribution	74
4-4	Sensitivity of the National Conditional Risk Distribution to Changes in the Means of the Airport Damage Distributions	75
4-5	Sensitivity of the National Conditional Risk Profile R to Changes in the Standard Deviations of the Airport Damage Distributions	76
4-6	Sensitivity of the National Conditional Risk Distribution to Changes in the Standard Deviations of the Airport Damage Distributions	77
4-7	Sensitivity of the Annual Risk Profile S to Changes in the Means of the Airport Damage Distributions	79
4-8	Sensitivity of the Annual Risk Distribution to Changes in the Means of the Airport Damage Distributions	80
4-9	Sensitivity of the Annual Risk Profile S to Changes in the Standard Deviations of the Airport Damage Distributions	81
4-10	Sensitivity of the Annual Risk Distribution to Changes in the Standard Deviations of the Airport Damage Distributions	82
4-11	Sensitivity of the Annual Risk Profile S to Changes in the Expected Number of Accidents in a Year	83
4-12	Sensitivity of the Annual Risk Distribution to Changes in the Expected Number of Accidents in a Year	84

Table		Page
5-1	Sample Sizes for a One-Sided Tolerance Limit	99
5-2	Sample Sizes for Upper K-S Confidence Bounds	102
5-3	Sample Sizes for Various Tail Probabilities and Confidences	103
5-4	Required Sample Sizes for 99% Confidence Statements with 10^{-4} Precision	109

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

STATISTICAL ASPECTS OF CARBON FIBER RISK ASSESSMENT MODELING

by

Donald Gross
Douglas R. Miller
Richard M. Soland

1. Introduction

This report deals with considerations of the statistical aspects of carbon fiber risk assessment modeling for fire accidents involving commercial aircraft. There are numerous advantages to using carbon fiber materials on aircraft; strength and weight reduction are two such examples. However, should an aircraft be involved in a fire accident, the possibility exists for a release of free carbon fibers to the atmosphere with a potential effect of some of these fibers infiltrating and shorting out electrical and electronic "machinery".

The ultimate goal of the entire carbon fiber risk assessment program was to determine risk profiles for this phenomenon; that is, curves of potential damage values and their associated probabilities. Very comprehensive reviews of the factors affecting such profiles and of the entire risk assessment program itself are given by Huston (1979, 1980). The present study focuses on the statistical aspects influencing the development of the risk profiles.

The next section of the report presents an overview of the statistical problems encountered in producing risk profiles and identifies the major sources of uncertainty. Sections 3, 4, and 5 treat each of these major uncertainty sources in detail, namely Section 3 deals with imprecise knowledge in establishing the model, Section 4 treats the problems associated with model parameter estimation and Section 5 concentrates on sampling errors in the Monte Carlo simulation analysis and obtaining confidence bounds on the results. Finally, Section 6 provides a general framework for building in and obtaining conservatism in risk profile generation.

2. Treatment of Errors and Uncertainty in CF Risk Analysis

In assessing CF-related damage due to accidents of commercial aircraft, various uncertainties must be dealt with. The cost incurred from a CF incident will vary according to the circumstances surrounding the incident, and the resulting uncertainty in the cost incurred is therefore described by a risk profile. A similar risk profile is used to describe the uncertainty associated with the total cost incurred in a year from CF incidents. The goal of risk analysis is to determine these profiles which reflect the inherent randomness of actual physical phenomena.

In the pursuit of estimating these risk profiles, additional uncertainties (potential sources of error) are encountered. These can be classified into three general categories: (1) imperfections in the mathematical model of the physical phenomena, (2) inexact specification of the numerical and quantitative aspects of the model, (3) statistical error from simulation sampling. A careful modeling effort and a statistically sound methodology can help to control these uncertainties, thereby preventing them from contributing to misleading conclusions about the risk profiles. In the CF analyses performed here, these factors have generally been controlled or else dealt with conservatively, resulting in conservative estimates of the risk profiles.

We will now elaborate on each of the three sources of uncertainty just mentioned and on how they were dealt with in the CF risk analysis.

(1) Imperfect model. By its very nature, the mathematical model will be an approximate description of the physical phenomena. This results from imprecise knowledge of the phenomena, the need for tractability, and factors that may have been overlooked. For example, in the CF model considerable aggregation is necessary, but this is performed cautiously by using conservative numerical quantities for entire classes of equipment, buildings, aircraft, etc. Deterministic cost values were used instead of distributions of costs, but it can be shown that the "law of averages" implies minimal error propagation due to this modeling simplification. Many secondary economic effects (lost production, clean-up, etc.) were included and their costs conservatively estimated; for example, the modeling approach of ORI allowed the entire gross domestic product of an area to be lost. In general, a modeling philosophy of "reasoned conservatism" was followed.

(2) Numerical inputs for the model. The model of CF risk requires many numerical and other quantitative inputs; among these are accident rates, positions of vulnerable equipment, weather frequencies, air traffic projections, building transfer coefficients, failure probability distributions, etc. Most of these quantities must be estimated from data and then projected to the year 1993, which leads to only approximate values of these quantities for use in the model. A conservative approach was generally taken. For example, for factors such as CF release and building penetration, conservatively high values were used. For many other factors, sensitivity analyses showed that imprecise knowledge of their true values had minor effects on the final estimate of the risk profiles. One sensitive parameter, however, is the equipment failure distribution. Equipment failure data and theoretical considerations showed that exponential failure laws were appropriate in many cases and conservative in others. Thus the use of exponential failure laws in the model leads to conservative results. Further, use of vulnerable 1979-vintage equipment results in conservative approximations to 1993 electronic equipment.

(3) Simulation sampling error. The risk model is exercised by simulation. This is equivalent to drawing a statistical sample from a

population, and attributes of the population are then estimated from the sample. When the computer simulation sampling is done properly, it is possible to use modifications of classical statistical techniques to generate accurate bounds on the sampling error, which permits the establishment of confidence bounds on the risk profile.

To summarize: by performing careful analyses it is possible to exert control over the effects of various sources of the added uncertainty (or error) which can influence the estimate of the risk profiles. While the statistical aspects of model building and simulation interpretation are not well developed enough to give precise conclusions on the total error in the final answers, it is possible to analyze the errors individually due to the separate sources and to control them. Figures 2-1, 2-2, and 2-3 illustrate this process. The interaction of the many sources is a problem, but a fairly strong degree of confidence in the conservatism of the final conclusion arises from the conservative and statistically sound approaches taken to control the individual error sources.

3. Treatment of Imprecise Knowledge in Model Conception

Any model is an abstraction of and approximation to the real world. In constructing a model, often a trade-off is necessary between "realism" and tractability. Further, judgements are often required in making certain assumptions concerning how factors behave and relate.

This section deals with two such examples, namely the use of deterministic values in place of random variables in order to gain modeling efficiency, and the choice of an appropriate probability distribution for representing equipment failure.

RISK ASSESSMENT PROCEDURE

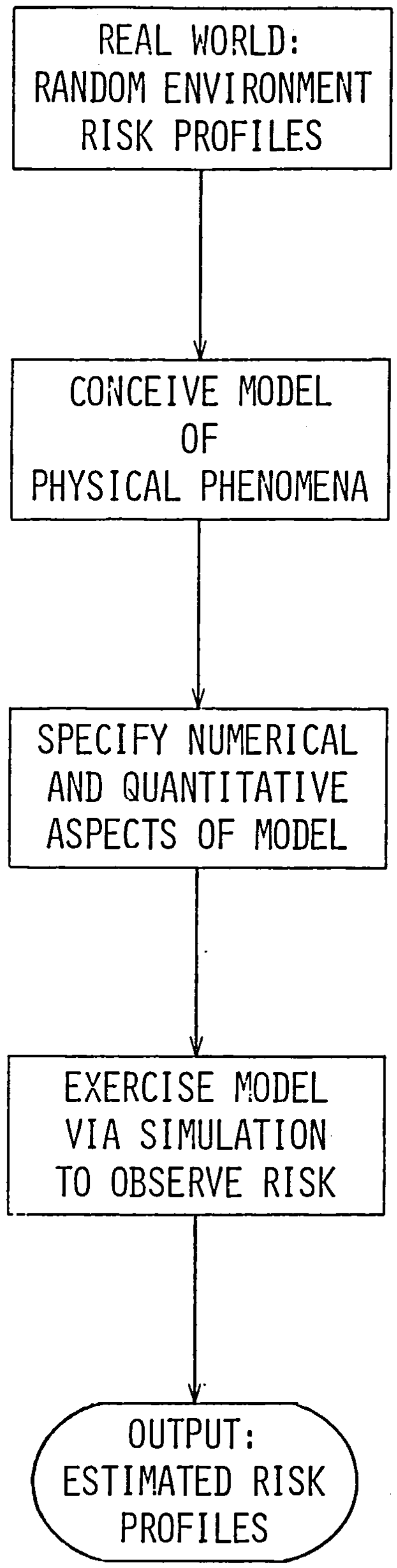


Figure 2-1.--Risk profile methodology.

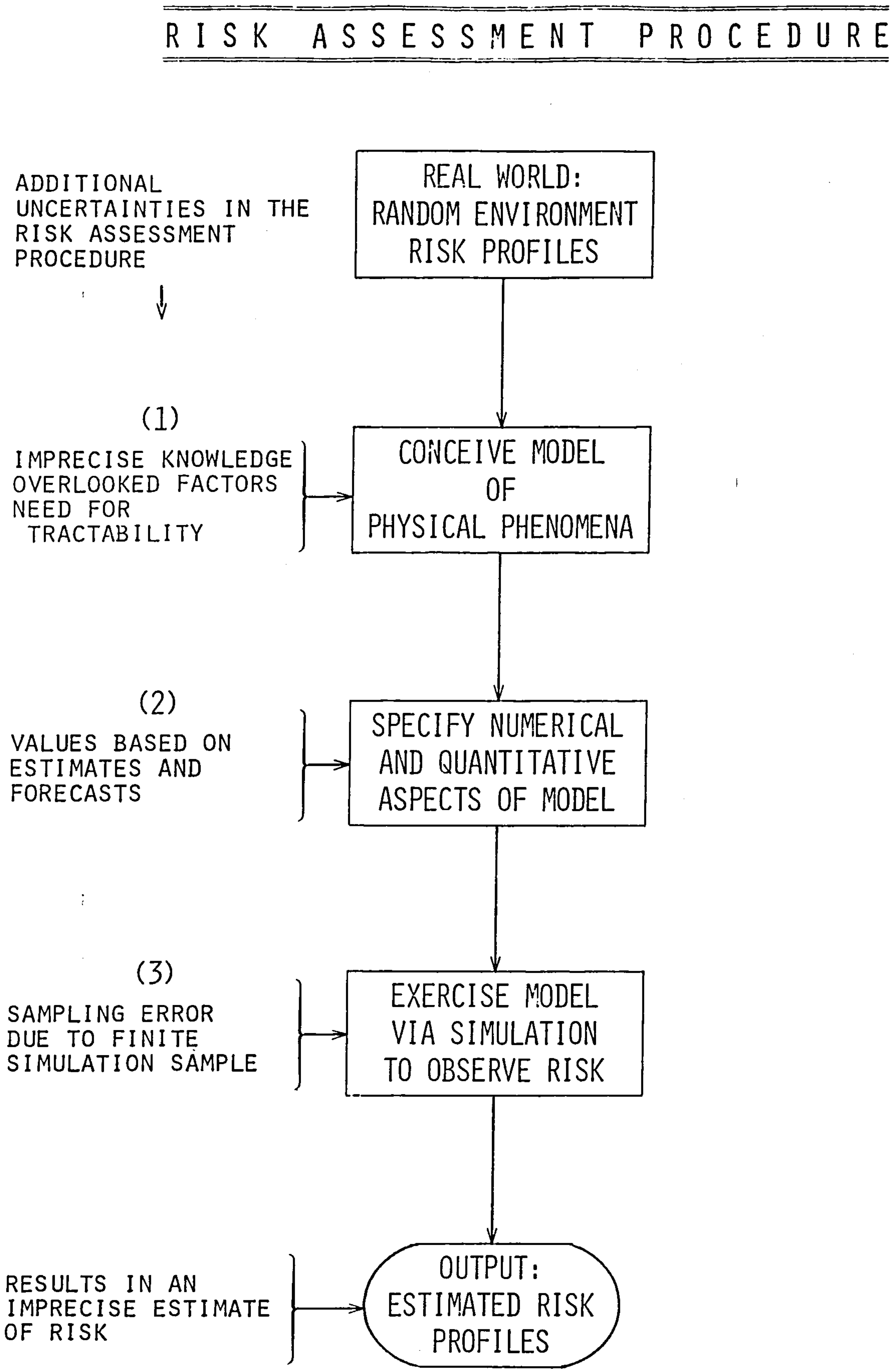


Figure 2-2--Potential errors affecting the risk profile methodology.

RISK ASSESSMENT PROCEDURE

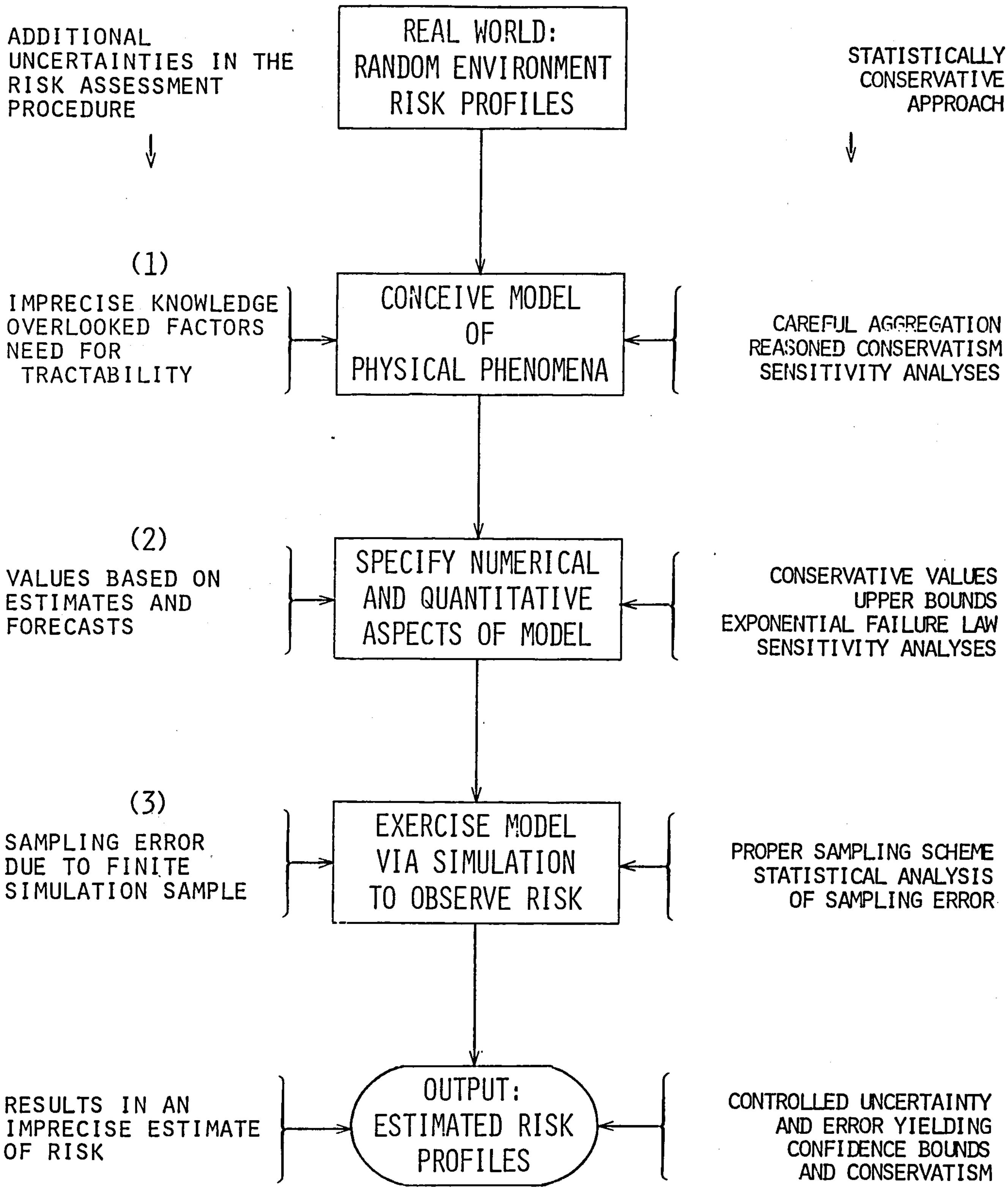


Figure 2-3--Approaches to handling errors in risk profile methodology.

3.1 Use of expected or deterministic values in place of random variables

The simulation analysis of CF damage and its cost is based to a large extent upon various averages, but the goal is to estimate (with confidence) risk probabilities. This section focuses on the effect that variation of input data which are assumed constant can have on the simulation conclusions.

Average values have been used instead of random ones for repair and downtime costs in the current simulation modeling effort. Considering the categories and types of repair and disruption costs, it seems that each would tend to be quite variable. A coefficient of variation (σ/μ) equal to 2 or 3 would not be unreasonable for most of these costs. The question is: What is the effect of using variable costs in inputs instead of expected costs?

The possible magnitude of such a change in the model can be seen for the variance of the national conditional risk profile by considering the following analysis of a much simpler problem. Let X equal the total economic loss or cost, given an accident at some particular airport. Roughly speaking, X equals the sum of a large number of individual costs $L_1, L_2, L_3, \ldots, L_N$, where N is random:

$$X = \sum_{i=1}^{N} L_i \, .$$

Now consider the distribution of X, in particular its mean and variance. Making the simplification for illustrative purposes that the L_i's are independent and have common mean μ_L and standard deviation σ_L, we have

$$\begin{aligned} E[X] &= E[N]\mu_L \\ \mathrm{Var}[X] &= E[X^2] - (E[X])^2 \\ &= E\left[\left(\sum_{i=1}^{N} L_i\right)^2\right] - (E[n]\mu_L)^2 \end{aligned}$$

$$= E\left[\sum_{i=1}^{N} L_i^2 + \sum_{i\neq j}\sum L_i L_j\right] - (E[N])^2 \mu_L{}^2$$

$$= E[N]E[L_i^2] + \sum_{n=0}^{\infty}((n^2-n)\ (E[L])^2 P(N=n) - (E[N])^2\mu_L^2$$

$$= E[N]E[L^2] + (E[N^2]-E[N])\ (E[L])^2 - (E[N])^2\mu_L^2$$

$$= E[N]\left[E[L^2]-(E[L])^2\right] + (E[L])^2\left[E[N^2]-(E[N])^2\right]$$

$$= E[N](\sigma_L^2) + \mu_L^2\sigma_N^2 .$$

Therefore,

$$\mathrm{Var}[X] = \mu_N\sigma_L^2 + \mu_L^2\sigma_N^2 . \tag{3.1}$$

Now consider

$$X' = \sum_{i=1}^{N} E[L] = N\ E[L] = N\mu_L ,$$

which is analogous to computing risk by assuming the costs are fixed at their expected values;

$$E[X'] = E[N]\mu_L ,$$

$$\mathrm{Var}[X'] = \sigma_N^2\mu_L^2 .$$

From the accidental nature of the process generating the cost it is not unreasonable to assume that N has a Poisson distribution. Then

$$\sigma_N^2 = \mu_N .$$

Suppose the coefficient of variation of the cost distribution equals k, that is,

$$\sigma_L/\mu_L = k ,$$

or

$$\sigma_L = k\mu_L .$$

Then

$$\mathrm{Var}[X] = (1+k^2)\mu_N\mu_L^2 ,$$

and hence

$$\mathrm{Var}[X] = (1+k^2)\mathrm{Var}[X'] .$$

Thus if the individual costs have coefficient of variation k, the true variation of risk is $1+k^2$ times the value obtained using expected costs.

The above is not meant to be a precise analysis for the actual system under consideration. However, the two are close enough that the above analysis may roughly apply. Namely, if there is much variation in costs, it can have a pronounced effect on the variation of the total risk. It is not unreasonable to believe the costs could have a coefficient of variation equal to 2 or 3, and in this case the standard deviation of X will be approximately 2.24 or 3.16, respectively, times as great as that of X'.

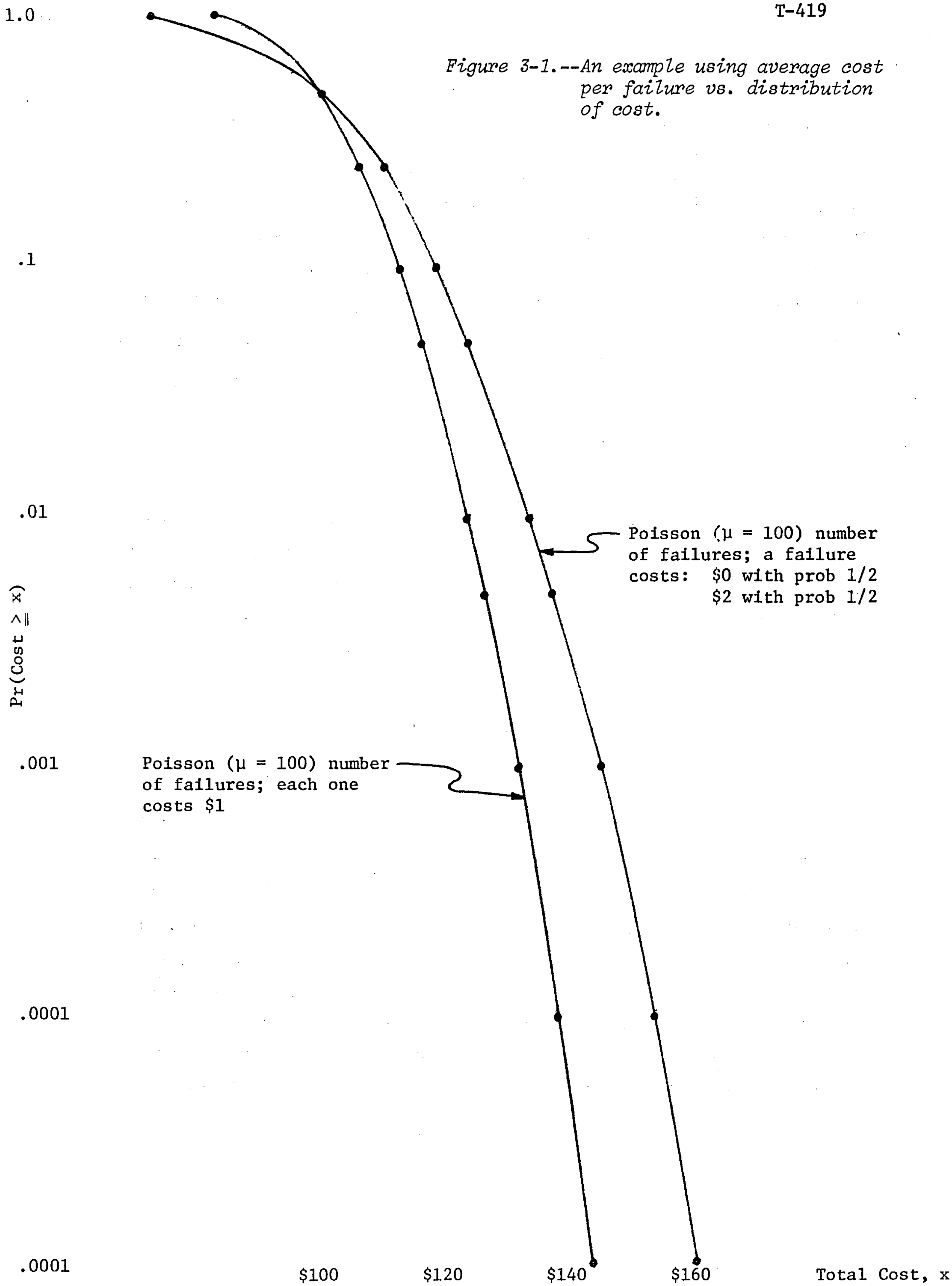
We have seen that in adding up a random number of randomly distributed costs, if the costs are assumed to be fixed and nonrandom, a term $\mu_N\sigma_L^2$ is ignored, and that this can be serious if μ_N is of the same order of magnitude as σ_N^2 and $\sigma_L/\mu_L \geqq 1$. If N is Poisson this is true. To get some idea of what the assumption of fixed costs would mean in that case, risk curves for two cases are plotted in Figure 3-1: one for fixed costs of \$1 per failure and one for a distribution of costs for each failure, \$0 and \$2 being equally likely. A normal approximation to the Poisson is used.

Note that, in a sense, the *severity* of the error by assuming fixed costs depends on how the graphs are used. There are two quantities that can be read off these graphs: tail probabilities and tail percentiles.

Tail probabilities: Suppose one is interested in the probability of the damage exceeding \$140. The correct answer is .0024, the incorrect answer is .000033, almost two orders of magnitude too optimistic--a rather bad error.

Tail percentiles: Suppose one is interested in the 99.99th percentile. The correct answer is \$153. The incorrect answer is \$137--not a very great error.

Figure 3-1.--An example using average cost per failure vs. distribution of cost.



Thus even though the correct variance is double the incorrect variance, for certain purposes it may not be too serious a mistake.

The reason we can obtain a good estimate of percentiles from the incorrect curve is that the slopes are steep. So the question arises: What happens when the slope of the curve based upon fixed costs is not steep? Such will only arise for a risk distribution with a heavy tail which implies a high coefficient of variation σ_N/μ_N ; for illustrative purposes let's suppose $\sigma_N/\mu_N = 3$. However, another factor now must be considered, namely, the magnitudes of σ_N and μ_N . Reconsider Equation (3.1), namely,

$$\mathrm{Var}[X] \;=\; \mu_N\sigma_L^2 + \mu_L^2\sigma_N^2 \,.$$

If we can assume the total cost incurred during a year equals the sum of many small costs (perhaps, on the average, *100* such costs), and further assume that they are roughly independently and identically distributed, then in (3.1) $\mu_N = 100$ and $\sigma_N = 300$. Furthermore, if we assume a worst cast of $\sigma_L/\mu_L = 10$, then the summands of (3.1) satisfy

$$\mu_N\sigma_L^2 = \frac{1}{9}\,\mu_L^2\sigma_N^2 \,.$$

Thus using fixed costs ignores approximately 10% of the variation.

The tentative conclusion is this: If one can assume that the total cost is the sum of many small costs and the number of these small costs incurred is distributed with $\sigma_N/\mu_N > 1$, then one can get a fairly good estimate of the variance by using the average costs rather than the distribution of costs. We believe this could be made precise in a manner analogous to proofs of the central limit theorem for sums of independent but not identically distributed random variables. However, the above possibilities hold only when the true mean costs are used. If these are not known exactly, another source of variation is introduced.

3.2 A general model of electronic equipment failure

We examine a simple, but fairly general, stochastic model of the process by which a piece of electronic equipment (hereinafter called the "item") may fail to function properly because of exposure to carbon fibers (CF). The model assumes the item to be composed of n independent subsystems or "circuits," each of which receives the same amount of exposure E to CF, thereby causing it to receive "shocks" or "hits" according to a Poisson stochastic process. Circuit j is assumed to fail after receiving r_j shocks, $j=1,\ldots,n$, and the item is assumed to fail when s circuits have failed; only the cases $s=1$ and $s=2$ are examined in detail. Also, we only examine closely the cases in which all r_j are the same.

For the various cases considered we compute $F(E)$, the probability that the item fails due to the exposure E . The algebraic form of $F(E)$ indicates the probability distribution of the exposure level at which the item fails; for certain combinations of the parameters n,s and r_j this distribution is either exponential or Erlang-r (for $r > 1$). For other combinations of the parameters the probability distribution is not one of the well-known types.

For all cases considered we examine the asymptotic behavior of $F(E)$ as E decreases to zero. It is found that $F(E)$ is approximately equal to cE^q for small values of E , where the constant c and the positive integer q depend on the particular case as well as the parameters. An exponential failure distribution yields $q=1$ while an Erlang-r failure distribution yields $q=r$.

Our conclusions are: (1) many different failure distributions arise from different choices of the model parameters; (2) examination of $F(E)$ for small values of E is not a good guide for inferring the probability distribution of failure of the item.

3.2.1 The general model

The model is based on the following assumptions. The item consists of n independent circuits, each subjected to an exposure E. Circuit j $(j=1,\ldots,n)$ is characterized by the parameter a_j, where $1/a_j$ is the *average* amount of incremental exposure *between* consecutive shocks of circuit j. More precisely, we assume that for *every* value of E in the interval $(0,\infty)$, the probability distribution of the number of shocks of circuit j during the period of time during which the exposure has accumulated to a value of E is the Poisson distribution given by

$$P_j(x_j) \;=\; e^{-a_jE}(a_jE)^{x_j}/x_j! \;, \quad x_j=0,1,2,\ldots \;. \tag{3.2}$$

In particular, $P_j(0) = e^{-a_jE}$.

It is assumed that circuit j has failed due to the exposure E if it has received r_j or more shocks, where r_j (assumed given) has one of the values $1,2,\ldots$. Let $F_j(E)$ be the probability that circuit j fails due to the exposure E. Then

$$F_j(E) \;=\; \sum_{x_j=r_j}^{\infty} P_j(x_j) \;=\; 1 - \sum_{x_j=0}^{r_j-1} e^{-a_jE}(a_jE)^{x_j}/x_j! \;.$$

It is assumed that the item has failed due to the exposure E if s or more circuits have failed, where s (assumed given) has one of the values $1,2,\ldots,n$. Let $F(E)$ be the probability that the item fails due to the exposure E. We will initially write expressions for $F(E)$ for three fairly general cases and then analyze various cases in detail.

For $s=1$ we have

$$1 - F(E) \;=\; \prod_{j=1}^{n}[1-F_j(E)] \;=\; \prod_{j=1}^{n}\sum_{x_j=0}^{r_j-1} e^{-a_jE}(a_jE)^{x_j}/x_j! \;. \tag{3.3}$$

For $s=2$ we have

$$1-F(E) = \Pr\{0 \text{ circuits failed}\} + \Pr\{1 \text{ circuit failed}\}$$

$$= \prod_{j=1}^{n} [1-F_j(E)] + \sum_{k=1}^{n} F_k(E) \prod_{j\neq k} [1-F_j(E)] \tag{3.4}$$

$$= \sum_{k=1}^{n} \prod_{j\neq k} \sum_{x_j=0}^{r_j-1} e^{-a_jE}(a_jE)^{x_j}/x_j! - (n-1) \prod_{j=1}^{n} \sum_{x_j=0}^{r_j-1} e^{-a_jE}(a_jE)^{x_j}/x_j! \;.$$

For general s we consider only the case in which $r_j=r$ and $a_j=a$ for all $j=1,\ldots,n$. Then

$$F_j(E) = F_*(E) = 1 - \sum_{x=0}^{r-1} e^{-aE}(aE)^x/x! \;, \tag{3.5}$$

and

$$1 - F(E) = \sum_{k=0}^{s-1} \binom{n}{k} [F_*(E)]^k[1-F_*(E)]^{n-k} \;,$$

and

$$F(E) = \sum_{k=s}^{n} \binom{n}{k} [F_*(E)]^k[1-F_*(E)]^{n-k} \;. \tag{3.6}$$

3.2.2 Case 1: $r_j=1$ for all $j=1,\ldots,n$

We first treat the situation in which $s=1$. For $r_j=1$ we have $F_j(E) = 1 - e^{-a_jE}$. Then, from (3.3), we find

$$F(E) = 1 - e^{-naE} \;, \tag{3.7}$$

where

$$a \equiv \frac{1}{n}\sum_{j=1}^{n} a_j \tag{3.8}$$

is the average shock rate. Equation (3.7) shows that the probability distribution of failure is *exponential*.

We now examine the asymptotic behavior of $F(E)$ as E decreases to zero. In general, for these asymptotic analyses, we will write $F(E) \sim cE^q$ to indicate that $F(E)/E^q$ approaches the constant c as E decreases to zero. The method used will almost always be to replace terms of the form e^{-kE} by the power series $\sum_{i=0}^{\infty}(-kE)^i/i!$ and then to collect coefficients of the various powers of E. Details will be omitted. The present case of the exponential distribution (s-1, all $r_j=1$) is the easiest, and yields, from (3.7),

$$F(E) \sim naE \; ; \tag{3.9}$$

i.e., $F(E)$ is asymptotically linear in E.

Now we treat the situation in which $s=2$. Direct substitution into (3.4) yields

$$F(E) = 1 - e^{-naE}\left[\sum_{j=1}^{n} e^{a_j E} - (n-1)\right] . \tag{3.10}$$

This expression does not correspond to the cumulative distribution function (CDF) of any well-known probability distribution. Asymptotic analysis of (3.10) yields

$$F(E) \sim cE^2 \; , \tag{3.11}$$

where $c = (n^2a^2 - \sum_j a_j^2)/2$, so $F(E)$ is asymptotically quadratic in E.

Specializing expressions (3.10) and (3.11) to the case $a_j=a$ for all j leads to simpler expressions for $F(E)$ and c, but the former still does not correspond to the CDF of any well-known distribution.

3.2.3 Case 2: $r_j=2$ for all $j=1,\ldots,n$

Now $F_j(E) = 1-e^{-a_jE}(1+a_jE)$. For the case $s=1$ expression (3.3) leads to

$$F(E) = 1 - e^{-naE} \prod_{j=1}^{n} (1+a_jE) . \qquad (3.12)$$

This is not the CDF of any well-known distribution, even if it is specialized to the case $a_j=a$ for all j . Asymptotic analysis of (3.12) yields

$$F(3) \sim cE^2 , \qquad (3.13)$$

where $c = (\sum_j a_j^2)/2$, so here again $F(E)$ is asymptotically quadratic in E .

For $s=2$ expression (3.4) leads to

$$F(E) = 1 - e^{-naE}\left[\sum_{k=1}^{n} e^{a_kE} \prod_{j\neq k} (1+a_jE) - (n-1) \prod_{j=1}^{n} (1+a_jE)\right], \qquad (3.14)$$

and asymptotic analysis of (3.14) yields

$$F(E) \sim cE^4 , \qquad (3.15)$$

where

$$c = (\sum_{j<k} a_j^2a_k^2)/4 .$$

For the case $a_j=a$ for all j , the constant c simplifies to $n(n-1)a^4/8$.

3.2.4 Case 3: $n=1$

In this case the item is treated as being composed of only one circuit. We now replace r_j by r and a_j by a , and must have $s=1$.

Then expression (3.3) yields

$$F(E) = 1 - \sum_{x=0}^{r-1} e^{-aE}(aE)^x/x! = F_*(E) . \tag{3.16}$$

Differentiation of (3.16) with respect to E yields the density function

$$f(E) = dF(E)/dE = a(aE)^{r-1} e^{-aE}/(r-1)! , \tag{3.17}$$

which shows that the failure distribution is *Erlang-r*. Asymptotic analysis of (3.16) is a bit more complicated in this case, but eventually yields

$$F(E) \sim cE^r , \tag{3.18}$$

where $c = a^r/r!$. Thus $F(E)$ is asymptotically linear in E for $r=1$ (the exponential case), quadratic in E for $r=2$, etc.

3.2.5 Case 4: general s

Expression (3.6) gives $F(E)$ for the case $r_j=r$ and $a_j=a$ for all j , and this cannot be significantly simplified, even for $r=1$. For the asymptotic analysis we note that s of the n circuits must fail in order to cause failure of the item, and each of the s failing circuits has $F_j(E) \sim a^r E^r/r!$. Thus

$$F(E) \sim cE^{rs} , \tag{3.19}$$

where

$$c = \binom{n}{s}\left(a^r/r!\right)^s \tag{3.20}$$

It is not difficult to see that expression (3.19) also holds when the a_j are different; the constant c is then no longer given by (3.20), but is instead more complex to express. What is important, of course, is that $F(E) \sim cE^q$, where $q = rs$.

3.2.6 Case 5: asymptotic results for the general case

For the asymptotic analysis the argument is still valid that s of the n circuits must fail in order to cause failure of the item, and now we have $F_j(E) \sim a_j^{r_j} E^{r_j} / r_j!$. Asymptotically, therefore, it is not difficult to see that

$$F(E) \sim cE^q , \tag{3.21}$$

where c is a constant whose exact value is algebraically messy to write down and

$$q = \sum_{p=1}^{s} r_{j_p} , \tag{3.22}$$

where the r_j are ordered so that

$$r_{j_1} \leqq r_{j_2} \leqq \cdots \leqq r_{j_n} .$$

Expression (3.22) reduces to $q=rs$ if the s smallest r_j values are all equal to r .

3.2.7 Case 6: limiting case for $s=1$ and large n

We consider only the case where $r_j=r$ and $a_j=a$ for all j . As discussed by Mann, *et al.* (1974), pp. 102-108, as n approaches infinity the limiting failure distribution is Weibull with shape parameter r . This has CDF

$$F(E) = 1 - e^{-cE^r} , \tag{3.23}$$

where c is a constant. For $r=1$ the Weibull distribution is the same as the exponential. The asymptotic form of (3.23) is clearly

$$F(E) \sim cE^r . \tag{3.24}$$

3.2.8 Conclusions

The failure model examined leads to different failure distributions when different choices are made for the parameters n , s , r_j , and a_j . For $s=1$ and all $r_j=1$ we obtain the exponential distribution (Case 1). For $n=1$ (so that $s=1$ and $r_1=r$) we obtain the Erlang-r distribution (Case 3). For n large, $s=1$, and $r_j=r$ and $a_j=a$ for all j , the failure distribution is approximately Weibull with shape parameter r (Case 6). For the other combinations examined, the failure distribution is not one of the standard ones. For a particular piece of electronic equipment, examination of the circuitry and physical arrangement of the components may help to establish which combinations of the model parameters are most appropriate (if indeed any of them are).

Asymptotic analysis indicates that the failure probability $F(E)$ behaves like cE^q for very small values of E , where c is a constant and q is a positive integer. The exponential failure distribution is the only one for which $q=1$; for all other cases $q>1$. Both the Erlang-r (Case 3) and Weibull-r (Case 6) distributions yield $q=r$, but these are not the only failure distributions with $q=r$; for example, $q=2$ also arises (a) with $s=2$ and all $r_j=1$ (Case 1), (b) with $s=1$ and all $r_j=2$ (Case 2), and (c) in the general Case 5 with

$$\sum_{p=1}^{s} r_{j_p} = 2 .$$

Since the same asymptotic form of $F(E)$ arises from a number of different failure distributions, it is clearly inappropriate to infer the type of failure distribution from an examination of $F(E)$ [or an estimate $\hat{F}(E)$ of $F(E)$] for small values of E .

It is important to note that, except for the exponential ($s=1$ and all $r_j=1$), all the cases examined here lead to failure distributions with increasing hazard rate (IHR). The analyses and remarks of Section 4.2 on testing for the exponential failure distribution are therefore appropriate.

4. Treatment of Parameter Estimation

A second major source of error in probability modeling is due to having to estimate parameters of probability distributions and in some case the probability distributions themselves. This section treats some estimation problems and the sensitivity of the risk analysis models to some of the estimated values.

4.1 Estimating the parameter of the exponential failure model

For the exponential failure model used in the GFRAP risk analysis it is necessary to know the value of the parameter $\bar{E}$. Since the actual value of the parameter can never really be known, the usual procedure is to use an appropriate estimate obtained from test data. The estimate most commonly used is the maximum likelihood estimate; we develop that here. We will use a Bayesian approach to treat the case in which no failures have been observed.

Suppose $m+n$ identical pieces of equipment have been exposed to graphite fibers (or perhaps the same piece of equipment $m+n$ times) and m of them have failed at exposures $E_1,\ldots,E_m$ while the other n have survived exposures of $E_{m+1},\ldots,E_{m+n}$. Either m or n could be zero. Based on the exponential failure model $P = \Pr(\text{Item survives exposure } E) = 1-e^{-(E/\bar{E})}$, it is straightforward to show that the likelihood of this sample result A is given by

$$\mathcal{L}(A) = \left[\prod_{i=1}^{m} \bar{E}^{-1} e^{-(E_i/\bar{E})}\right]\left[\prod_{i=m+1}^{m+n} e^{-(E_i/\bar{E})}\right]$$

$$= \bar{E}^{-m} e^{-(E^*/\bar{E})} ,$$

where $E^* = \sum_{i=1}^{m+n} E_i$ is the total exposure. The maximum likelihood estimate of $\bar{E}$, call it $\hat{\bar{E}}$, is the value of $\bar{E}$ which maximizes $\mathcal{L}(A)$; it is easily found to be

$$\hat{\bar{E}} = E^*/m ,$$

provided $m > 0$. If $m=0$, i.e., no failures have been observed, there is no maximum likelihood estimate for $\bar{E}$; we turn therefore to the Bayesian approach. The Bayesian approach uses a prior distribution on $\bar{E}$ and combines it with the likelihood $\mathcal{L}(A)$ via Bayes' theorem to obtain a posterior distribution on $\bar{E}$. The mean of the posterior distribution is often taken as a point estimate of $\bar{E}$.

A convenient prior distribution to use is one that is a natural conjugate of the likelihood $\mathcal{L}(A)$, and in this case that is an inverted gamma-1 distribution [see Raiffa and Schlaifer (1961), Chapter 10] of the form

$$f(\bar{E}|m', E') = K'\bar{E}^{-m'-1} e^{-(E'/\bar{E})} ,$$

where $m' > 0$ and $E' > 1$ are parameters and K' is a normalizing constant. The mean of this prior distribution exists if $m' > 1$ and is $E'/(m'-1)$. When this prior distribution is combined with the sample likelihood $\mathcal{L}(A)$ via Bayes' theorem, the posterior distribution is easily found to be inverted gamma-1 also, with parameters $m'' = m'+m$ and $E'' = E'+E^*$. The posterior mean is thus $(E^*+E')/(m+m'-1)$.

How should the prior parameters m' and E' be chosen? The following rationale is admittedly *ad hoc*, but it *is* rational and it leads to a reasonable and useful estimate of $\bar{E}$. For $m > 0$, a Bayesian analysis is not needed since we will use the maximum likelihood estimate $\hat{\bar{E}} = E^*/m$, so we are only concerned with the case $m=0$.

For the case $m=0$ the posterior mean is $(E^*+E')/(m'-1)$. *If one failure had* been observed along with the total exposure of E^* , we would use the maximum likelihood estimate $\hat{\bar{E}} = E^*/1 = E^*$. We will therefore choose the prior parameters m' and E' so that with $m=1$ and the same total exposure E^* , the posterior mean would have this same value E^* . This yields $(E^*+E')/m' = E^*$, or $E' = E^*(m'-1)$.

We will choose m' so that the standard deviation of the prior distribution is k times the prior mean. This requires $m' > 2$, and yields the equation

$$[\mathrm{Var}\,\bar{E}|m',\, E')]^{.5} \equiv \frac{E'}{(m'-1)(m'-2)^{.5}} = \frac{kE'}{(m'-1)} .$$

When this is combined with the previous equation $E' = E^*(m'-1)$, the posterior mean is found to be

$$E^*(2k^2+1)/(k^2+1) .$$

To indicate a good deal of prior uncertainty about the true value of $\bar{E}$ it is appropriate to chose k rather large. Even for $k=2$ the above expression yields a posterior mean of $1.8E^*$, and as k tends toward infinity m' tends toward 2, E' tends toward E^* , and the posterior mean tends toward $2.0E^*$. It is suggested that this value be used as an appropriate estimate of $\bar{E}$ for the case of no failures; i.e., use $\bar{E} = 2E^*$ if $m=0$.

There is an alternate rationale for choosing m' and E' , again *ad hoc,* that leads to the same prior parameters ($m' = 2$ and $E' = E^*$) and hence the same value of the posterior mean, $2E^*$. One may imagine the value E^* as having been chosen as the maximum total exposure for the experimental testing because it was felt that this much total exposure would yield at least one failure. From the relations $m''-1 = (m-1)+m$ and $E'' = E'+E^*$ one can interpret the prior information implied by the choice of m' and E' as being equivalent to having observed $(m'-1)$ failures caused by a total exposure of E' .

The choice of E^* as hypothesized just above then suggests the choices $(m'-1) = 1$ and $E' = E^*$.

Finally, let us note that if no failures have been observed and a conservative estimate of $\bar{E}$ is desired, then $\bar{E} = E^*$ is appropriate since this is equivalent to conservatively assuming that one failure has occurred.

4.2 Testing for an exponential failure distribution

We were given failure data for equipment exposed to carbon fibers. Twenty-one separate experiments were conducted using different electrical components and/or different lengths of types of fibers. Each experiment was replicated several times (the number varied between 3 and 11). The equipment, fiber, and performer of these experiments are summarized in Table 4-1.

Using the data, we wish to test the hypothesis that an exponential distribution is a reasonable failure model for electrical equipment exposed to graphite fibers, i.e., that the probability p of failure is related to the exposure x by the functional relationship $p = F(x) = 1 - \exp(-\lambda x)$, for some λ , where $1/\lambda$ = mean exposure to failure.

The exponential distribution is used widely in reliability studies, and consequently there is a considerable body of literature concerned with estimation and testing related to the exponential distribution. A good discussion is found in the book by Mann, Schafer, and Singpurwalla (1974), Ch. 7.

4.2.1 The Kolmogorov-Smirnov-Lilliefors test

Many test are based on the empirical cumulative distribution function. The most widely used of these appears to be the modification of the Kolmogorov-Smirnov test due to Lilliefors (K-S-L). We now describe it.

TABLE 4-1

FAILURE DATA: 21 TESTS OF EQUIPMENT

1.	Avionics terminal blocks	7.5 mm	AS	BRL
2.	Dynaco amplifier	7 mm		Mike Vogel
3.	Dynaco	7 mm		Mike Vogel
4.	Dynaco	3.5 mm		Vogel
5.	Dynaco	15 mm		Vogel
6.	Dynaco	15 mm		Harvey
7.	Dynaco	3.5 mm		Mike Harvey
8.	Dynaco	7 mm		Mike Harvey
9.	Dynaco	1 mm	GY70	Phillips
10.	LSI-11 computer	4.5 mm		BRL
11.	LSI-11 computer	7 mm		BRL
12.	19" TV	8 mm	HMS	BRL
13.	19" TV	8 mm	T300 sized	BRL
14.	19" TV	8 mm	T300 unsized	BRL
15.	Transponder	10 mm	GY70	Bionetics
16.	Transponder	3 mm	GY70	Bionetics
17.	Sunbeam toaster	7 mm	T300	Bionetics
18.	Sunbeam toaster	3 mm	GY70	Bionetics
19.	Heritage House	12 mm	T300	Bionetics
20.	Heritage House	7 mm	T300	Bionetics
21.	Sunbeam toaster	12 mm	T300	Bionetics

Suppose we test n items. For each item, note the exposure level x_i at which it fails. (Note: all items must be exposed until they fail; the results described here are *not* valid if items are withdrawn before they fail.) Thus we get n failure values: $x_1, x_2, \ldots, x_n$. The empirical distribution function is

$$\hat{F}_n(x) = \frac{1}{n} \times (\text{number of failure values} \leq x)$$

(see Figure 4-1). This function is an unbiased estimate of the underlying CDF, F ; i.e., $E(\hat{F}_n(x)) = F(x)$, for all x . (It has mathematical properties which enable one to use it in test statistics and to make statements about the possible inference errors of such procedures, e.g., probability of false rejection of a true hypothesis.)

The usual procedure is to look at the maximum deviation between $\hat{F}_n$ and the hypothesized CDF F_0 . However, we are not hypothesizing a single distribution, but an entire family, the exponential family

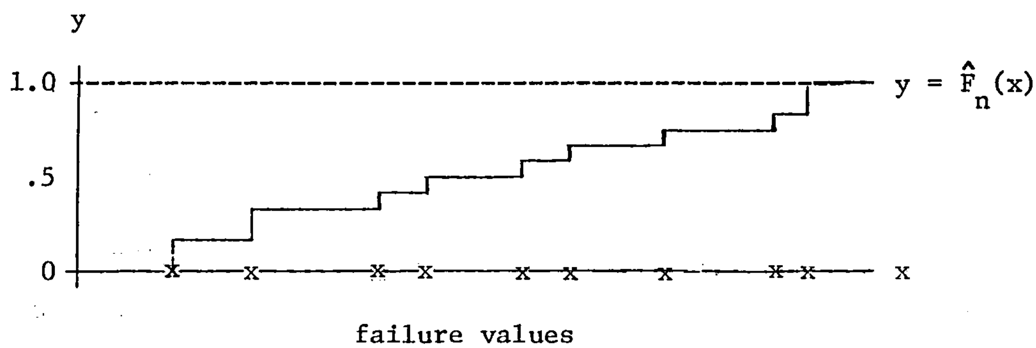


Figure 4-1.--The empirical cumulative distribution function.

$\{F_\lambda : F_\lambda(x) = 1 - e^{-\lambda x}$, any $\lambda\}$. In this case we estimate the parameter λ from the data (for the exponential $\hat{\lambda} = n / \sum_{i=1}^{n} x_i$) , and then look at the maximum deviation between the empirical CDF $\hat{F}_n$ and the exponential CDF $F_{\hat{\lambda}}$, with estimated parameter $\hat{\lambda}$, i.e., we look at

$$d = \sup_{0 \le x < \infty} |\hat{F}_n(x) - F_{\hat{\lambda}}(x)| ,$$

If d is large, we reject the hypothesis of exponentiality. If d is small, we accept it. The critical values depend on n and on α, the desired probability of false rejection. They are given by Lilliefors (1969). For example, if $n=6$ and we desire only a 5% chance of falsely rejecting the hypothesis of exponentiality we should reject if

$$\sup_x |\hat{F}_n(x) - F_{\hat{\lambda}}(x)| > .406 .$$

We have applied the K-S-L test for exponentiality to four of the data sets obtained from NASA Langley. They are graphed in Figures 4-2 through 4-5. On all four, the hypothesized (best-fitting) exponential CDF is plotted, and then upper and lower bounds are given, based on that exponential CDF ± the critical value (.406, in the case of 6 observations). If the empirical CDF falls in this region, we accept the hypothesis of exponentiality. Otherwise we reject it. Note: for set #1 (Figure 4-2), we just barely reject exponentiality at the $\alpha = 5\%$ level (i.e., we would expect such a deviation to occur due to chance alone to be an event with probability <.05); for set #2, we accept exponentiality, even though the best fitting Erlang has shape parameter 2 or 3; for set #15 (Figure 4-4), the data is highly significant: reject exponentiality; for set #16 we accept, but just barely, even though Erlang-4 is the best fitting. In selecting the data sets to analyze, we picked the ones which seemed to deviate the most from being exponential. The results do not constitute an overwhelming rejection of exponentiality. Among 21 sets of data, all drawn from exponential populations, it would *not* be surprising to find 2, say, which are significant at the 5% level.

To give some idea of how difficult it would be to distinguish between an exponential distribution and an Erlang-2, we generated a sample of size 10 on an HP-25 calculator and then normalized the values so the sample mean was 1. The plot of the exponential CDF, the Erlang-2 CDF,

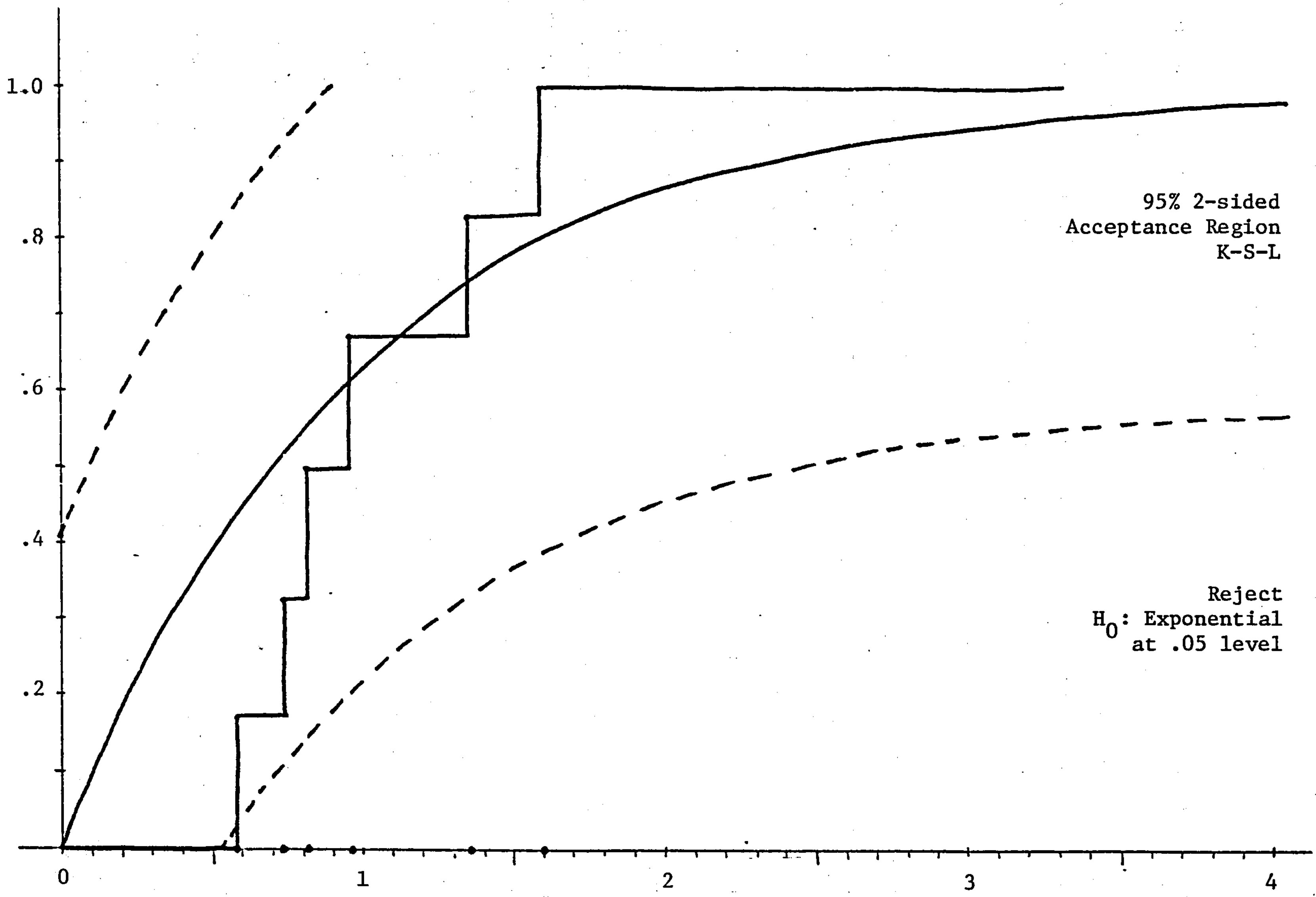


Figure 4-2.--Data Set #1: avionics terminal blocks, 7.5 mm.

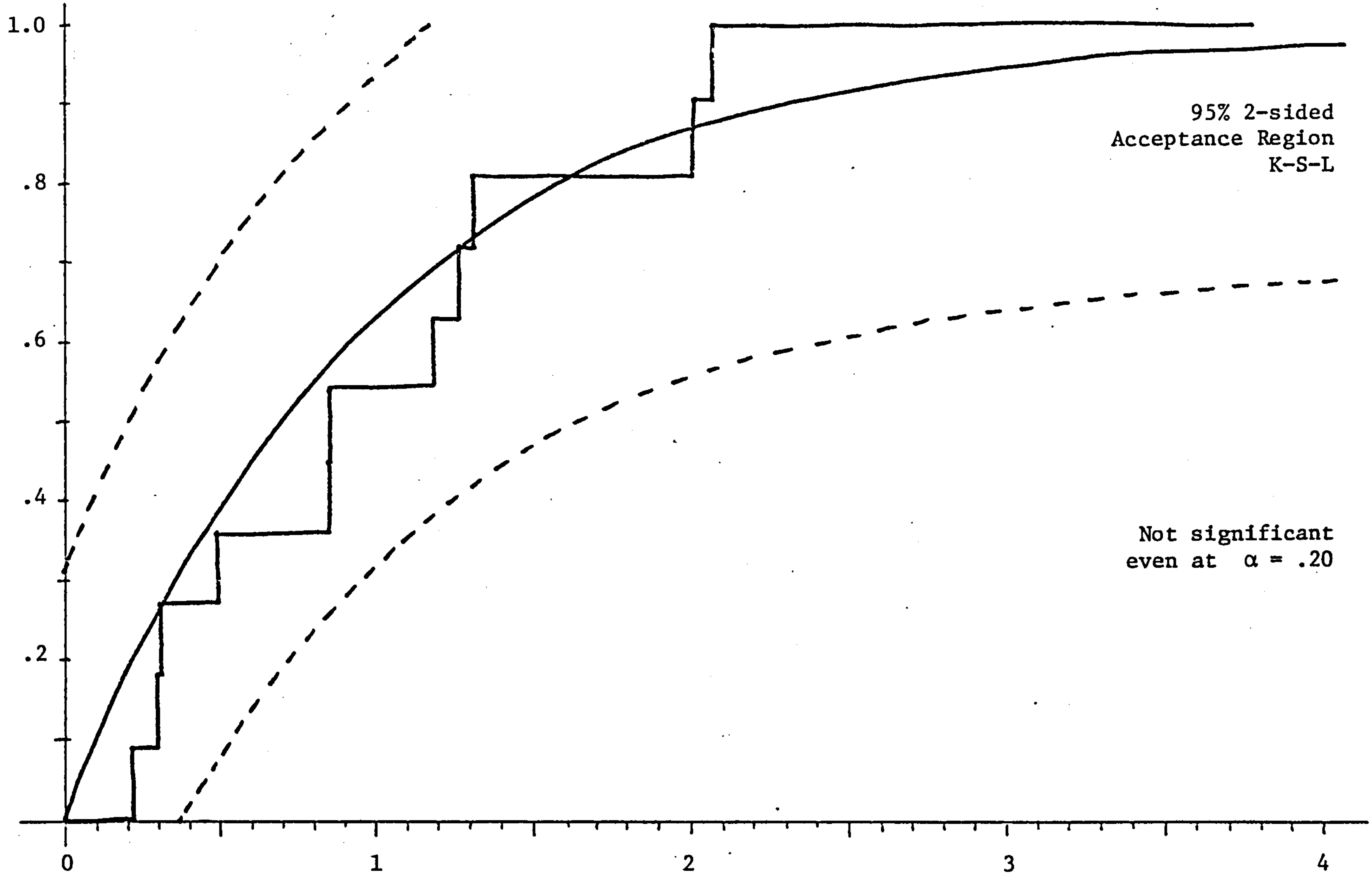


Figure 4-3.--Data Set #2: Dynaco amplifier, 7 mm.

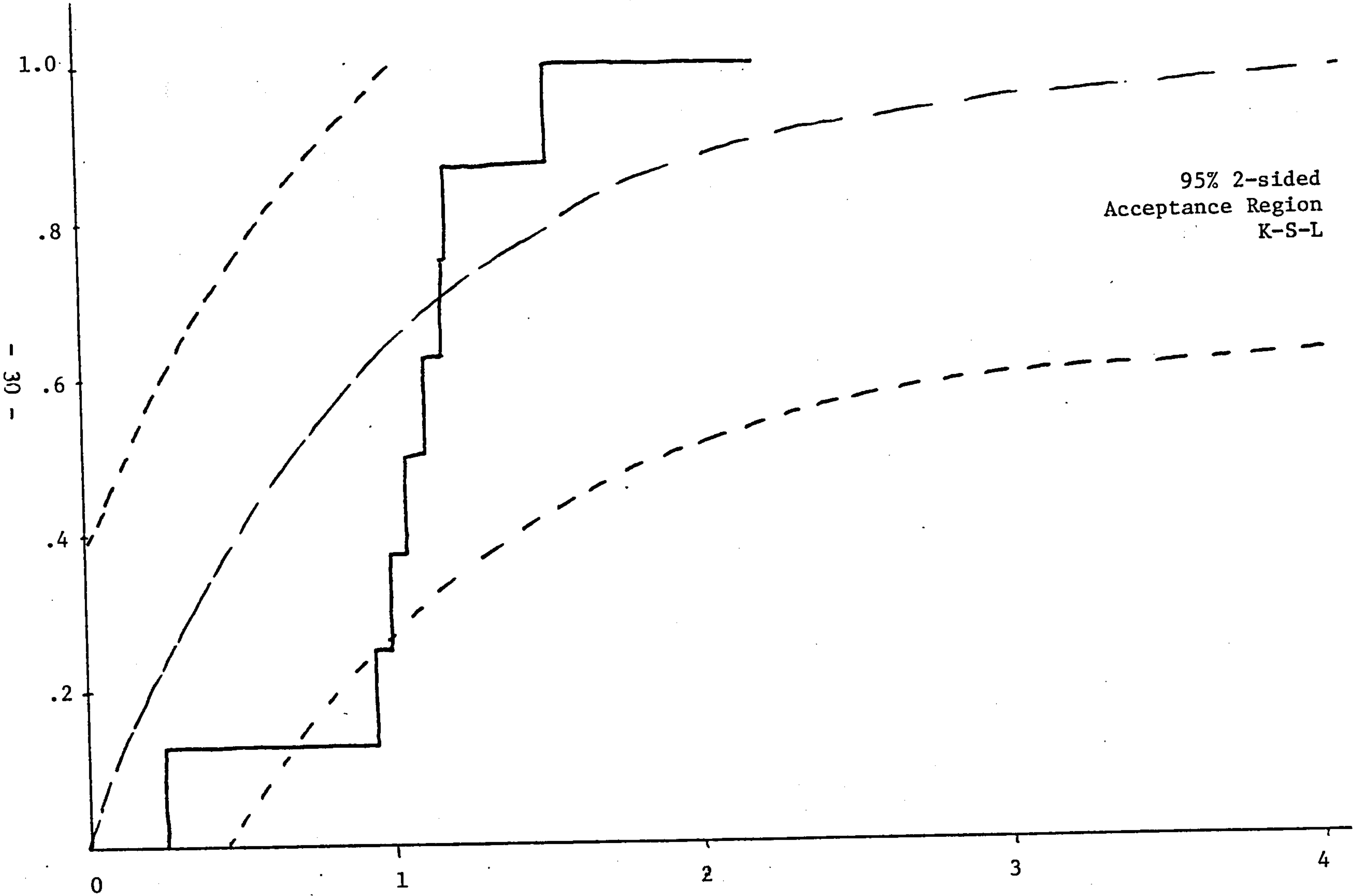


Figure 4-4.--Data Set #15: transponder GY 70, 10mm.

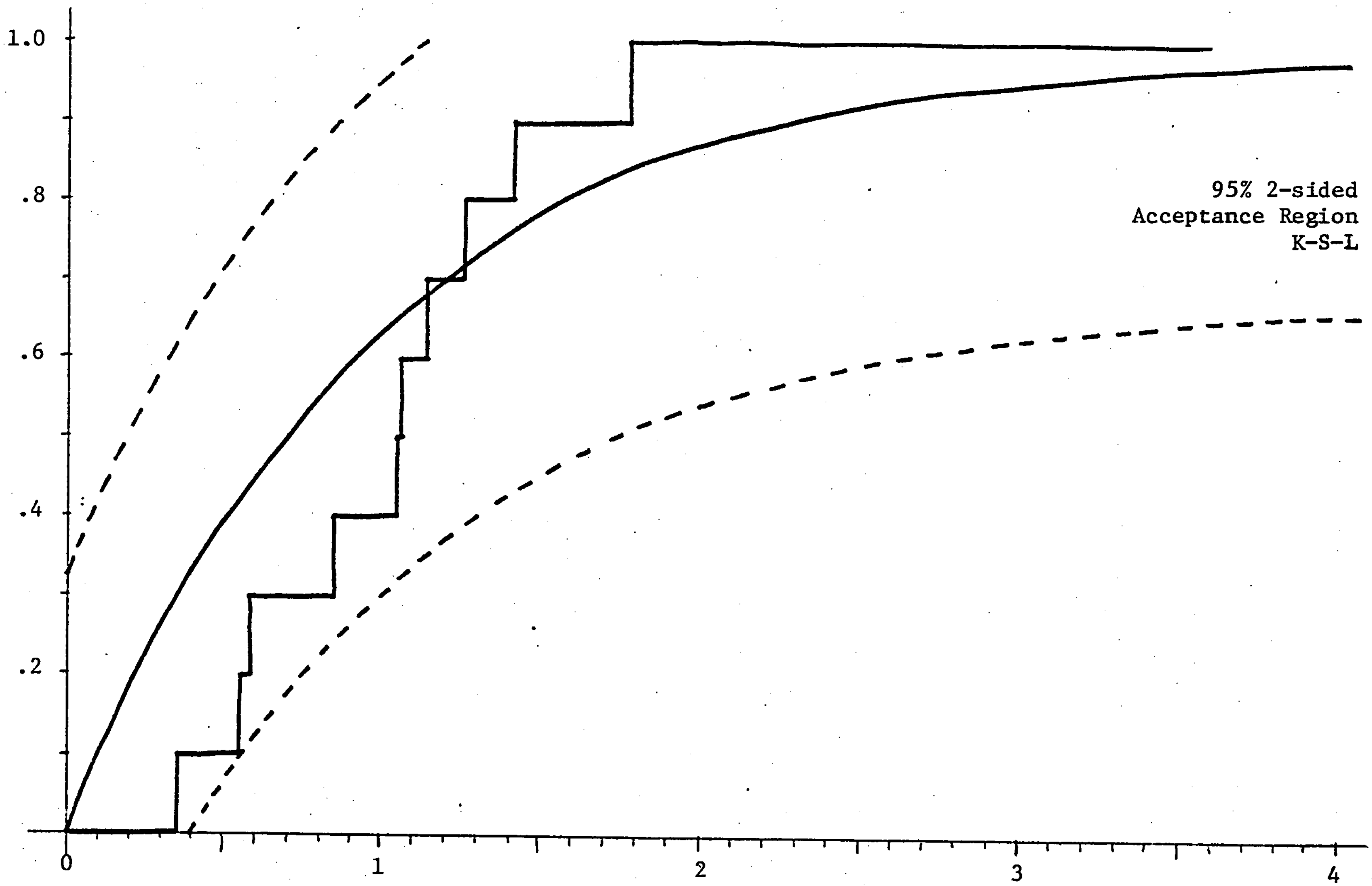


Figure 4-5.--Data Set #16: transponder GY 70, 3 mm.

and the empirical CDF drawn from the exponential are given in Figure 4-6. Note that for such a small sample, neither distribution is an obvious choice.

One property of the K-S tests is that they have very low power. "Power" is defined as the probability of rejecting the hypothesis when it is false. If the true distribution is Erlang-2, we estimate that the chance of rejecting the hypothesis of exponentiality, using a sample of size 10 and $\alpha = .05$, would be .20 or less.

4.2.2 The cumulative total-time-on-test statistic

The low power of the K-S-L test motivates the search for a more powerful test, i.e., one that is more likely to reject the hypothesis of exponentiality when it is false. One way to increase power is to construct a specialized test which is especially good at rejecting exponentiality when a certain class of alternatives is true and to demonstrate that this restricted class of alternatives includes all possibilities: roughly speaking, it is equivalent to saying that you have a better chance of making the right decision if there are fewer alternatives from which to choose.

For the failure distributions under consideration it is reasonable to assume nondecreasing hazard rate (IHR). Let F be the CDF of a random variable and f its density, then $r(x) = f(x)/(1-F(x))$ is the *hazard rate*. It follows that $P\{x \leq X \leq x+dx \mid x \leq X\} \approx r(x)dx$. The distribution is IHR if r is nondecreasing. If r is constant, then we have an exponential distribution, which is considered a boundary case of IHR. IHR is a reasonable assumption for the distribution of failure probabilities as a function of exposure; it is equivalent to saying: if two components (#1 and #2) have survived exposure levels e_1 and e_2, respectively, $e_1 < e_2$, then the next increment of exposure, Δe, is more likely to cause failure to #2 than #1' i.e., $r(e_1)\Delta e < r(e_2)\Delta e$ for $e_1 < e_2$.

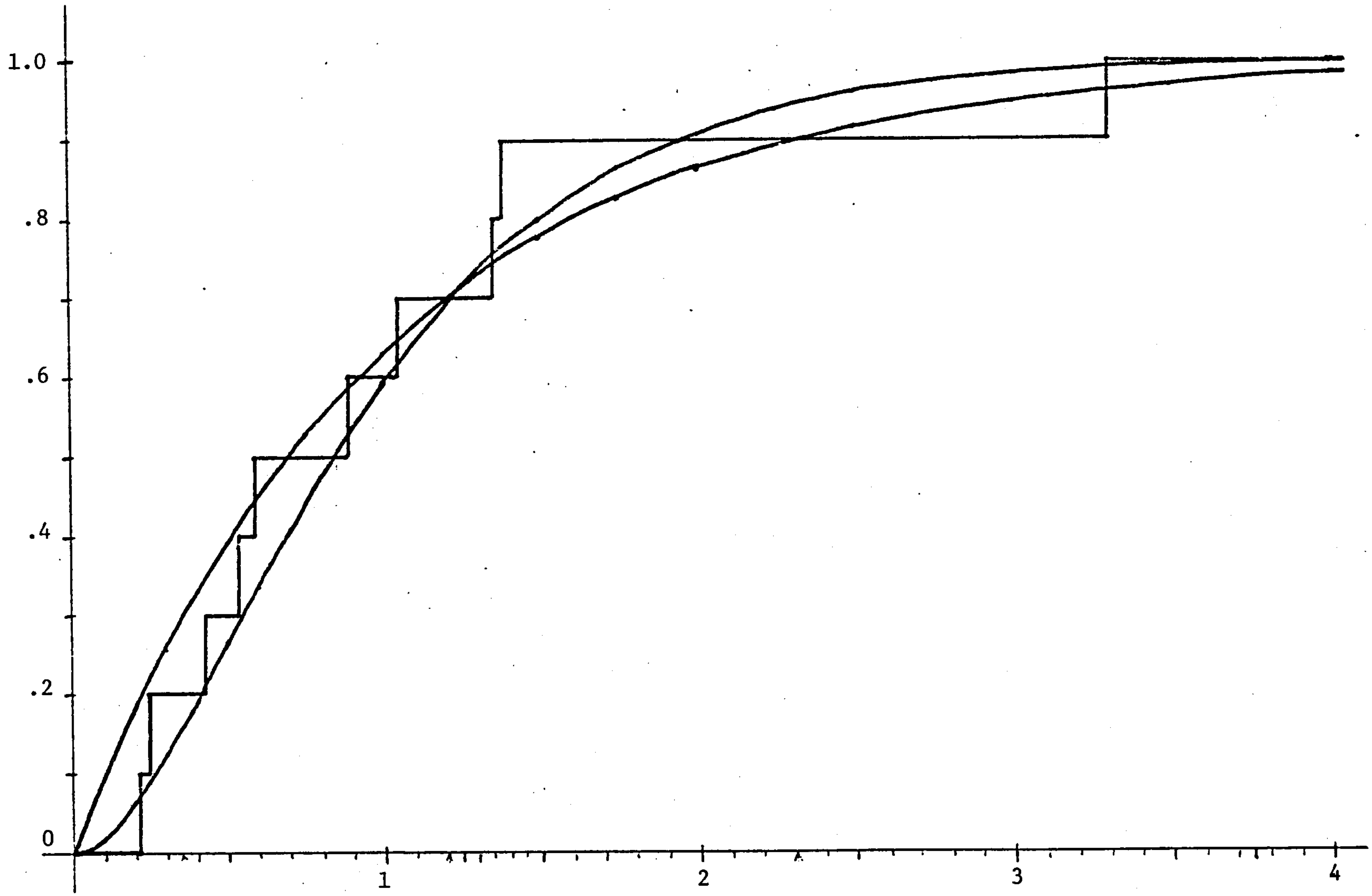


Figure 4-6.--Comparison of exponential and Erlang-2 (both with mean 1) with an empirical CDF generated from the exponential and then normalized to have mean = 1.

Thus we are in a situation of testing the hypothesis of exponentiality versus the alternative of IHR. The most powerful test (we know of) for this situation is discussed by Barlow (1968). It is actually a test for exponential vs. increasing hazard rate average (IHRA). A test which is virtually as powerful is a test based on the time-on-test statistic. This test has other desirable features, such as applicability to censored samples and a nice statistical distribution theory which makes it appear more attractive. It is described in detal by Barlow, Bartholomew, Bremner and Brunk (1972), Secion 6.2. We apply this test to data set #16 (plotted in Figure 4-5), the transponder exposed to 3 mm fibers to GY 70 (Bionetics data). The 10 normalized failure times (exposures) are

.36, .56, .58, .85, 1.05, 1.06, 1.09, 1.26, 1.42, 1.78.

A test statistic is computed as follows: let $X_{i:n}$ be the *ith* ordered failure value, let $D_{i:n} = (n-i+1)(X_{i:n} - X_{i-1:n})$ be the time (exposure) on test accumulated between the (i-1)*st* and the *ith* failure values. This test uses the fact that

$$P\{D_{i:n} > D_{j:n}\} = \frac{1}{2}, \qquad i \neq j ,$$

under the assumption of exponentiality, while under the assumption of IHR,

$$D_{1:n} \underset{st}{\geq} D_{2:n} \underset{st}{\geq} D_{3:n} \underset{st}{\geq} \cdots \underset{st}{\geq} D_{n:n} ;$$

Thus under exponentiality the $D_{i:n}$'s will tend to have 0 slope as a function of i ; under IHR the slope is negative. Thus this becomes a regression problem of testing for zero slope vs. negative slope. The appropriate statistic is

$$V_n \equiv n^{-1} \sum_{i=1}^{n} (i-1) D_{n-i+1:n} \Big/ n^{-1} \sum_{i=1}^{n} D_{i:n}$$

$$= n^{-1} \sum_{i=1}^{n-1} \left[\sum_{j=1}^{i} D_{j:n} \right] \Big/ n^{-1} \sum_{i=1}^{n} D_{i:n}$$

$\equiv$ cumulative-total-time-on-test statistic.

We reject exponentiality for large values of V_n. Critical values of V_n are given by Barlow *et al.* (1972), page 269; for $n=10$ they are as follows:

α	.10	.05	.025	.010	.005
critical value	5.619	5.927	6.189	6.487	6.683 .

For data set #16 we get the following:

i	$X_{i:n}$	$D_{i:n}$	$T_n(X_{i:n}) = \sum_{j=1}^{i} D_{j:n}$
1	.36	3.60	3.60
2	.56	1.80	5.40
3	.58	.16	5.56
4	.85	1.89	7.45
5	1.05	1.20	8.65
6	1.06	.05	8.70
7	1.09	.12	8.82
8	1.26	.51	9.33
9	1.42	.36	9.69
10	1.78	.36	10.05
			67.20

$$V_n = \sum_{i=1}^{n-1} T_n(X_{i:n}) \;/\; T_n(X_{n:n}) = 67.20/10.05 = 6.72 \ .$$

This corresponds to significance at approximately $\alpha = .005$. Thus we reject exponentiality in favor of IHR. (We accepted it using K-S-L.) Thus using this criterion (the cumulative-total-time-on-test statistic) we have observed a deviation from exponentiality which would occur by chance with probability .005. Note that in the K-S-L, the maximum deviation hit the lower edge of a 95% two-sided region; this lower boundary would be an approximate boundary for an $\alpha = 10\%$ level one-sided test. Thus the K-S-L criterion says such a deviation will occur approximately 10% of the time purely by chance. The extra sensitivity of the cumulative-total-time-on-test statistic is due to the fact that it is tailor-made for testing the situation in which we are interested, while the K-S is applicable to a broader range of situations and thus should not be expected to compete favorably with special tests in special situations.

We applied the cumulative-total-time-on-test statistic to all 21 sets of data. The results are tabulated in Table 4-2. Considering each experiment separately, there are four cases where exponentiality can be rejected and three questionable cases. Exponentiality is a plausible hypothesis for the remaining cases.

These conclusions are based on considering each test separately. Exponentiality was rejected when the test statistic took an improbable value (in the 1% extreme tail). However, in 21 trials the chance of an even of probability .01 occurring at least once is $1 - (.99)^{21} = .19$. Thus all 21 populations could have exponential and yet there would be a .19 chance of rejecting exponentiality for at least one of the 21. Thus to be rigorous and be sure that the chance of making such an inference error is .05, say, or less, we must reject at the .00244 level of significance (since $1 - (.00244)^{21} = .05$). In this case exponentiality is rejected for only two cases.

It is possible to construct a joint test of exponentiality using the time-on-test statistic. Note that $V_n \overset{\mathcal{L}}{=} u_1 + u_2 + \dots + u_{n-1}$, where u_i's are independent uniform [0,1] random variables, under the assumption of exponentiality. This fact allows us to construct a test for exponentiality of *all* 21 sets:

H_0 : all are exponential;

H_1 : all are IHR, with at least one being strictly IHR (i.e., not exponential).

Suppose the *ith* data set consists of n_i observations; let $V_{n_i,i}$ be

TABLE 4-2

TESTING FOR EXPONENTIAL VS. IHR USING CUMULATIVE-TOTAL-TIME-ON-TEST STATISTIC

Test #	Number Observ.	Value of Statistic	Expected Value Given H_0	Level of Significance	Conclusion
1	6	3.86	2.5	.015	?
2	11	7.11	5.0	.01	Reject ←
3	5	3.13	2.0	.025	?
4	6	1.90	2.5	.80	Accept
5	5	1.44	2.0	.81	Accept
6	4	1.58	1.5	.44	Accept
7	4	1.02	1.5	.79	Accept
8	4	1.95	1.5	.22	Accept
9	4	1.88	1.5	.25	Accept
10	5	2.88	2.0	.07	?
11	10	5.36	4.5	.17	Accept
12	5	1.09	2.0	.92	Accept
13	4	1.71	1.5	.36	Accept
14	3	.90	1.0	.58	Accept
15	8	5.70	3.5	.002	Reject ←
16	10	6.72	4.5	.005	Reject ←
17	10	8.38	4.5	~ .000008	Reject ←
18	10	3.94	4.5	.73	Accept
19	10	4.39	4.5	.55	Accept
20	10	4.32	4.5	.58	Accept
21	10	2.88	4.5	.98	Accept

the cumulative-total-time-on-test statistic for the *ith* data set. Then

$$V = \sum_{i=1}^{21} V_{n_i,i}$$

is an aggregate statistic, which if H_0 is true is the sum of $\sum_{i=1}^{21}(n_i-1)$ independent uniform [0,1] random variables and consequently approximately Normal with easily calculated mean and variance, from which critical values can be computed.

In light of data set #17 (see Table 4-2), it is obvious that H_0 : all exponential must be rejected. Considering #17 an outlier or an anomaly, we threw it out and tested the hypothesis H_0 : all exponential for the remaining 20 data sets. The test statistic

$$V = \sum_{\substack{i=1 \\ i\neq 17}}^{21} V_{n_i,i} = 63.75 .$$

If H_0 is true, this is an observation from a Normal $(\frac{110}{2}, \frac{110}{12})$ distribution; its level of significance thus equals .002 and H_0 is firmly rejected.

Barlow (1968) has computed the power curve of the cumulative-total-time-on-test statistic when the true distribution is Gamma with shape parameters between 1 and 5 for an $\alpha = .05$-level test based on $n = 10$ observations; see Figure 4-7. Thus, for example, if the true distribution is Erlang-5, the test will reject exponentiality with probability $\simeq 0.69$.

The cumulative-total-time-on-test statistic can be applied to data with censoring. If items are exposed to graphite fibers, but do not fail, they are eventually withdrawn from test. The withdrawal exposure level is noted and the data thus take the form of failure exposure levels and withdrawal exposure levels. The K-S-L approach has no provision for

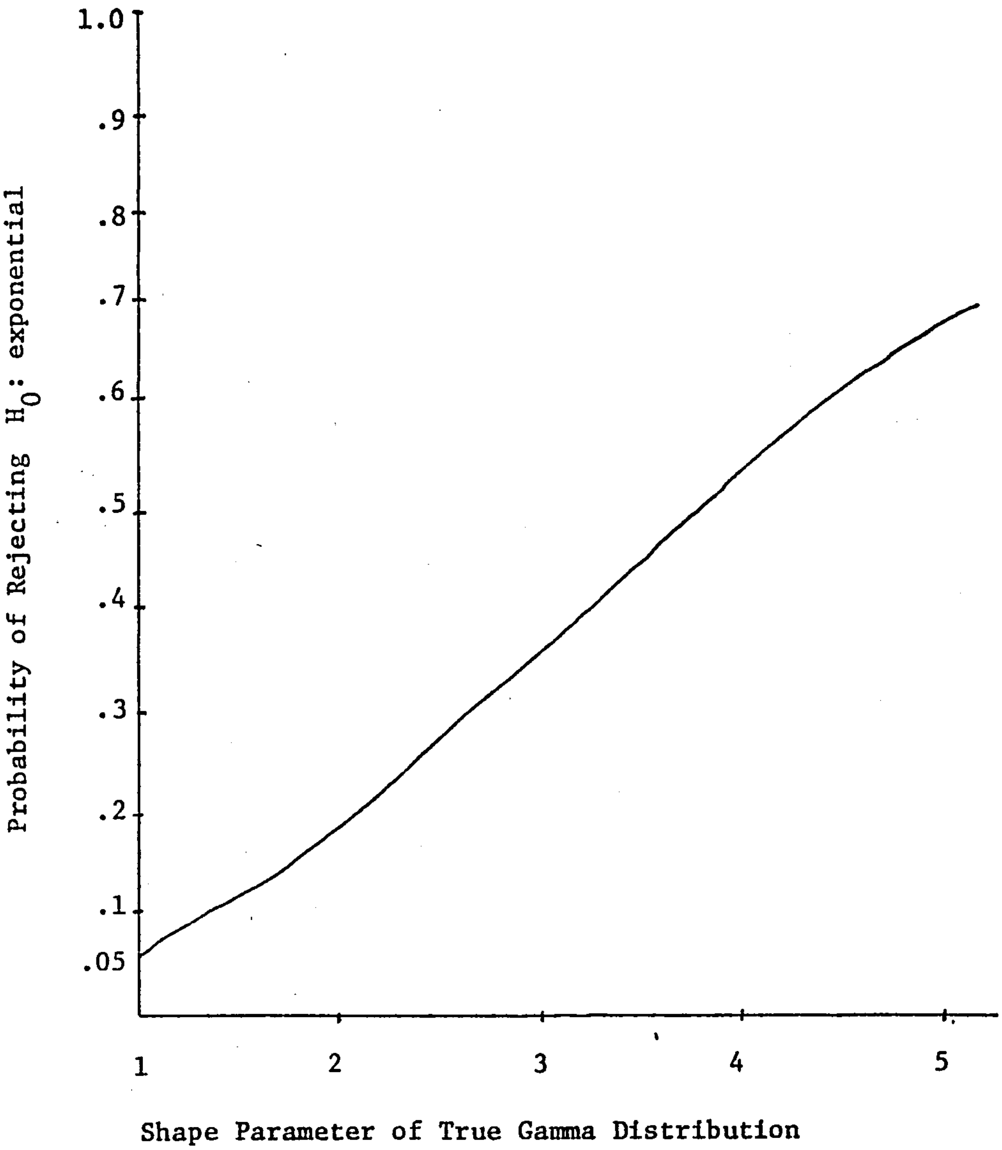


Figure 4-7.--Power curve for test of exponential vs. Gamma based on cumulative-total-time-on-test statistic.

handling these types of data. Fortunately, the cumulative-total-time-on-test approach handles these types of data easily; it is just a matter of redefining $D_{i:n}$, the total-time-on-test between the (i-1)*st* and *ith* failure levels to accommodate withdrawals in this interval. And instead of computing V_n, we compute V_k, where k is the number of observed failures. In view of this it is very important to make the distinction between failure times and times of withdrawal from testing of nonfailed items.

4.2.3 Conclusions

There are good tests for exponential vs. IHR, but it is hard to distinguish between them with small samples. This results in the hypothesis of exponentiality appearing plausible in many instances. Since to accept exponentiality when the distribution is IHR is a conservative error, exponentiality should be used as the failure distribution except in cases where knowledge of the vulnerabilities of the box allow a more detailed model which results in another failure distribution such as an Erlang.

One unexpected feature of the data is that in data sets #15 and #16, the case with length 10 mm deviated more from exponentiality than the case with length 3 mm. However, this can be explained by the randomness inherent in the small samples.

For simulation inputs, what is really needed is some upper confidence bounds on the failure distribution CDF or some method of handling an estimated distribution.

Incorrectly assuming an exponential failure distribution when the distribution is actually IHR (e.g., Erlang-n with $n > 1$) is conservative in the left-hand tail but optimistic in the right-hand tail. That is, the exponential distribution with equal mean exposure to failure gives a higher probability of failure for low exposure levels and a lower probability of failure for high exposure levels. This is illustrated in Figure 4-8, which compares the failure probabilities for exponential, Erlang-2, and Erlang-4 distributions with the same mean value (equal to 1 here). Note that the exponential distribution gives conservative results for all exposures less than about 1.2 times the mean exposure to failure, at which point the failure probability is already very high--

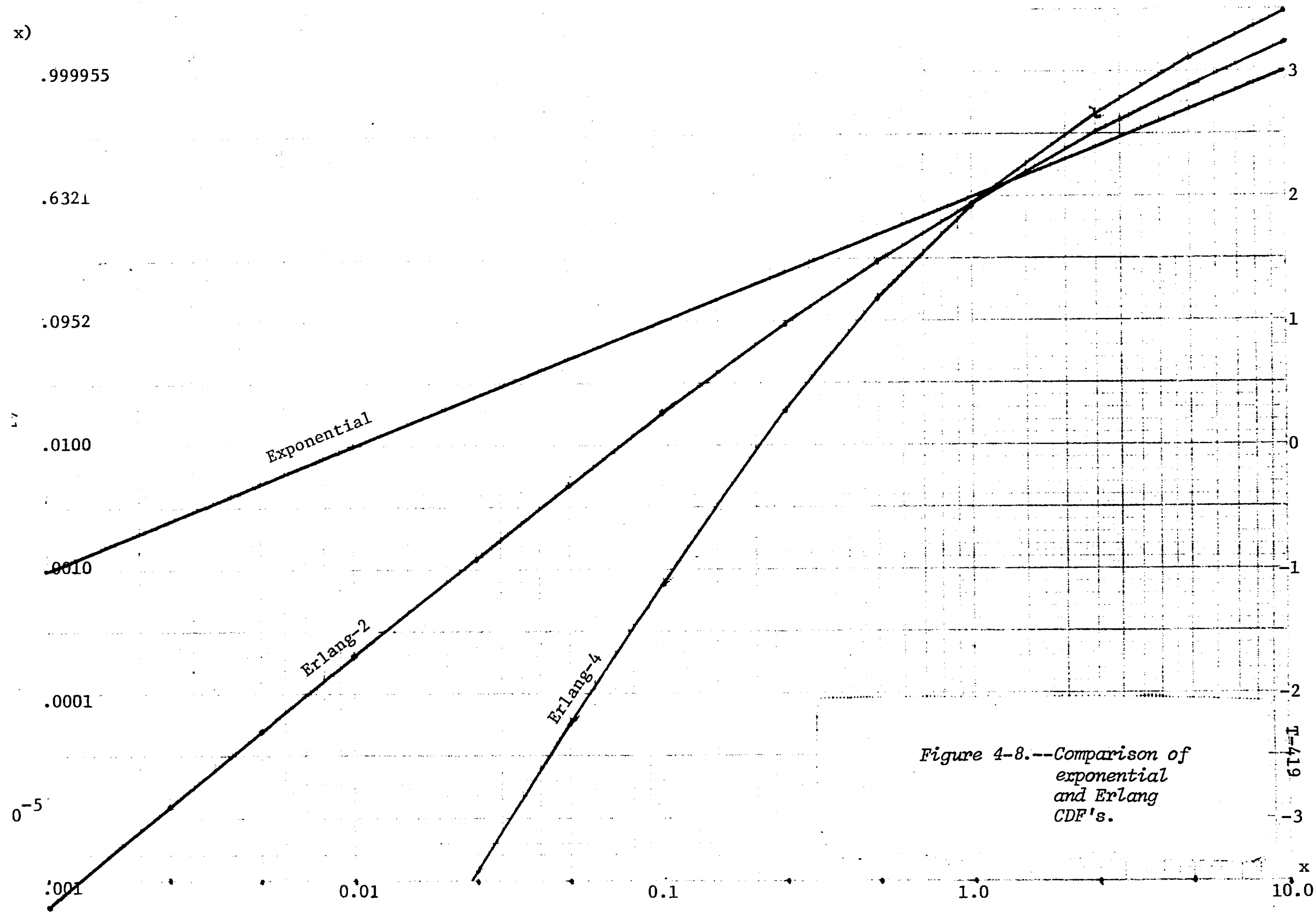


Figure 4-8.--Comparison of exponential and Erlang CDF's.

equal to about 0.7. Thus, for the small exposure levels to which actual electronic equipment will normally be subjected after a CF indident, the use of an exponential failure model produces significantly conservative results.

4.3 Confidence regions for IHR failure distribution

In this risk study it is crucial to know the relationship between exposure level and failure probability for an electrical component. Since this failure probability distribution is estimated from data, there is uncertainty involved and confidence regions should be used. In this case an upper confidence bound on the failure probabilities is needed. Typically, there is a small chance of failure at the exposure levels we expect to encounter, thus we are primarily concerned with estimating (with confidence) the left tail of the failure distribution.

To illustrate this problem, we consider some failure data collected by Westinghouse for a 7.5 KV insulator pin exposed to 5 millimeter fibers. Fifteen tests were performed. The failure data are presented in Figure 4-9 in the form of an empirical distribution function of failure probability versus exposure. In accidental releases we expect to see exposure levels up to 10^5 fiber sec/m^3; consequently, we are concerned with the extreme left tail in Figure 4-9.

We would like to estimate failure probabilities of this component for exposures in the neighborhood of 10^5. There are two classical ways to do this (See Sections 5.2 and 5.3): we can get point estimates based on Binomial probabilities. The fact that 0 of 15 components failed at an exposure level of 10^5 leads to a 95% confidence statement: "Probability of component failure at exposure level of 10^5 fiber sec/m^3 is less than .181." The other method is to compute a Kolmogorov-Smirnov simultaneous 95% confidence region: "$F(e) < \hat{F}_u(e) = \hat{F}_{15}(e) + .304$," where $F(e)$ equals the probability of failure at exposure level e and $\hat{F}_{15}(e)$ equals the estimate of this probability from 15 observed failures. (The boundary $\hat{F}_u$ of this upper confidence region is plotted in Figure 4-9.) The Kolmogorov-Smirnov gives a 95% confidence statement: "Probability of component failure at exposure level 10^5 fiber sec/m^3 is less than .304."

Clearly, we would like to get more precise estimates. This is possible. It can be done by exploiting a physical property of the failure process, namely, that it has an increasing hazard rate (IHR). This fact was discussed in the context of testing for exponential failure laws in Section 4.2.

Let F be the failure distribution and f its density. Define $h(e) = f(e)/(1-F(e))$ and $H(e) = \int_0^e h(t)dt = -\log(1-F(e))$. The function h is the hazard rate and H is the cumulative hazard or log survival function. If the distribution is IHR, then h is monotone nondecreasing and H is convex. Let $\hat{F}_u$ be the K-S upper bound. If we assume an IHR failure distribution, then a 95% confidence region for F will consist of all IHR distributions bounded by $\hat{F}_u$. It turns out that there exists a maximal IHR distribution $\hat{F}_{u,IHR}$ among all distributions bounded by $\hat{F}_u$: Let $\hat{H}_u = -\log(1-\hat{F}_u)$ and $\hat{H}_{u,IHR}$ be the greatest convex minorant of $\hat{H}_u$ (plotted Figure 4-10, then $\hat{F}_{u,IHR} = 1 - \exp(-\hat{H}_{u,IHR})$. If F' is IHR and $F' \leqq \hat{F}_u$, then $F' \leqq \hat{F}_{u,IHR}$.
A 95% confidence region $F \leqq \hat{F}_{u,IHR}$ is indicated in Figure 4-11 for the data of Figure 4-9.

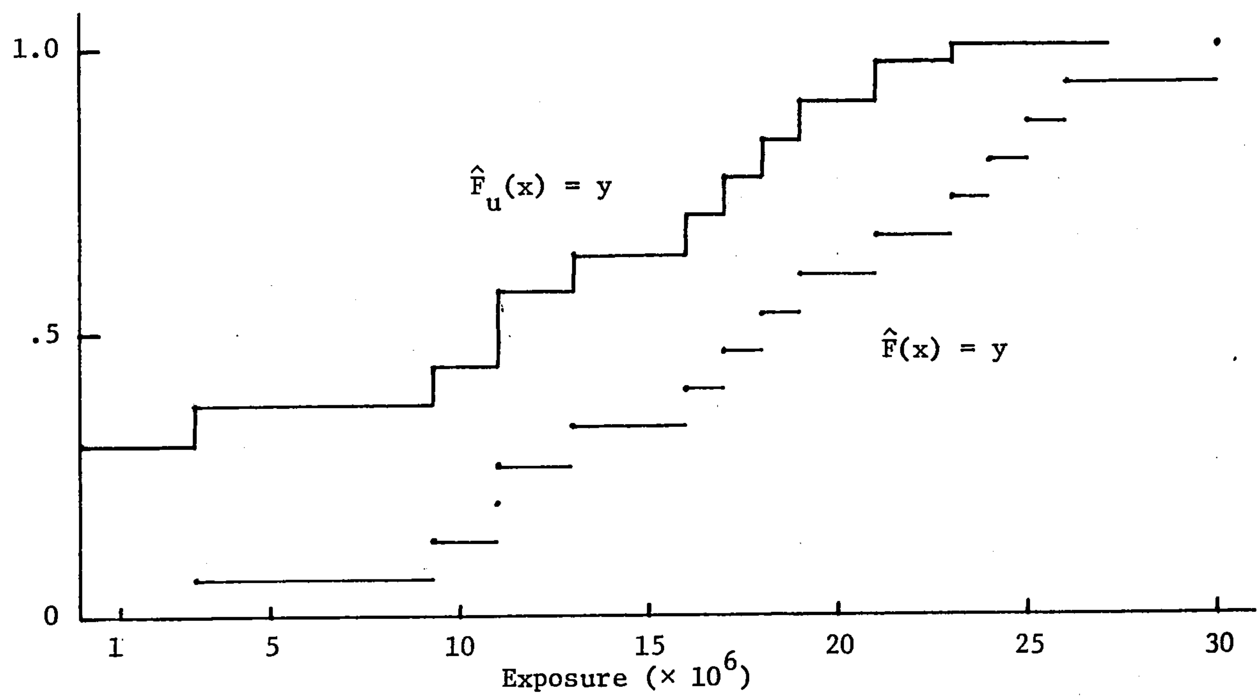


Figure 4-9.--Empirical CDF and 95% K-S upper bound; 7.5 KV insulator pin, 5 mm fibers, Westinghouse data.

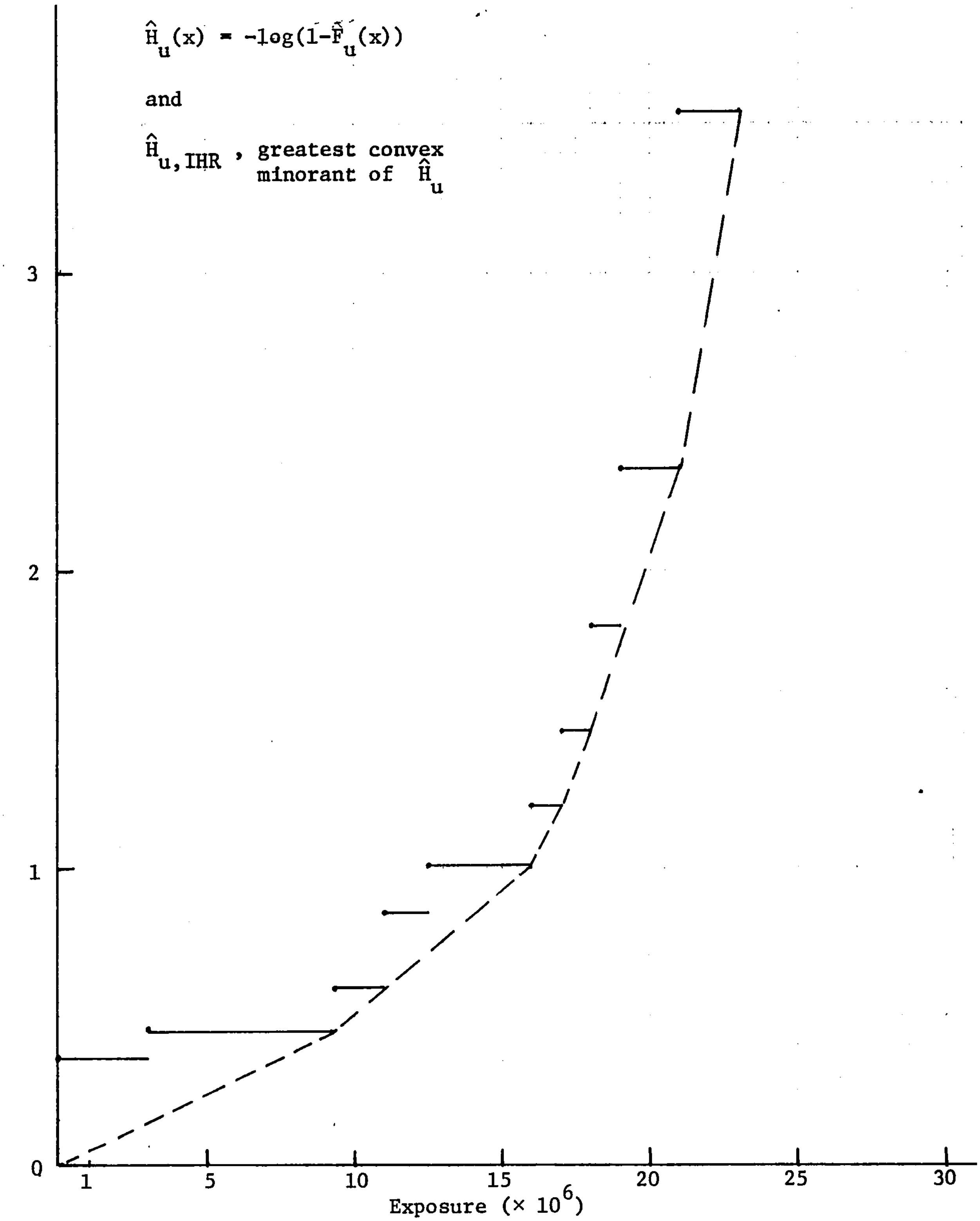


Figure 4-10.--Empirical cumulative hazard upper bound, $\hat{H}_u$ *, and its greatest convex minorant* $\hat{H}_{u,IHR}$.

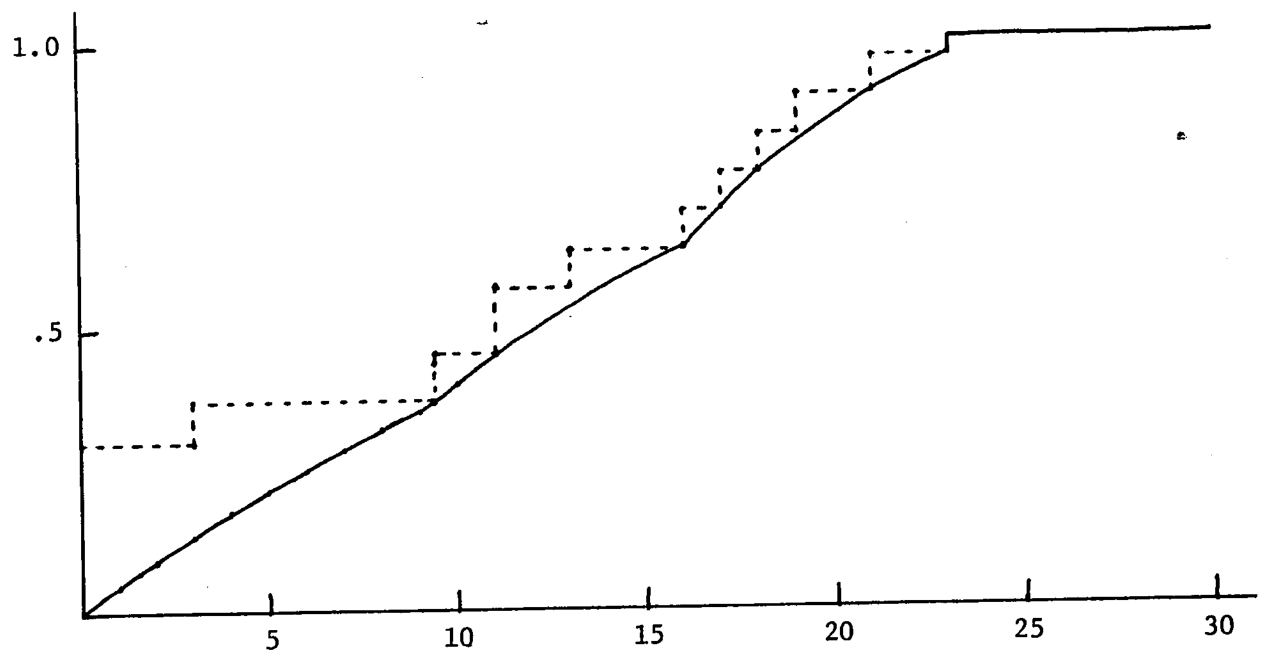


Figure 4-11--95% confidence region for IHR failure distribution for data in Figure 3-1.

Note that Figure 4-11 is a much more accurate region, especially in the left tail. This accuracy was gained by assuming a more restricted model, namely, an IHR failure distribution. For the previously considered exposure level of 10^5 fiber sec/m^3, we are 95% confident that the probability of failure is less than .0048. This is great improvement over the other estimates.

4.4 Sensitivity of the model to estimated parameters and distributions

We have written a special simulation program designed to provide fast and economical sensitivity analyses of the risk profiles obtained as output of the GFRAP risk analysis. Moreover, this program follows a modified simulation approach which yields statistically valid confidence bounds on the risk profiles obtained. The program requires as input the probability distributions of damage per accident at each of the major airports being considered, and we used analytic approximations to the empirical damage distributions reported by Arthur D. Little, Inc. (ADL) in Kalelkar, *et al.* (1979). We have performed an extensive series of sensitivity runs of the program, and of an additional computer program written to provide partial analytic results when the analytic approximations just mentioned take the form of Lognormal distributions. This section describes the simulation program, details our efforts to analytically approximate the different airport damage distributions, and finally presents the results of the sensitivity runs.

4.4.1. The simulation program

For ease in handling simulation error (to be discussed in Section 5.1) we recommend the following procedure for the simulation:

Step 1: Generate by Monte Carlo methods the number of accidents in a year.

Step 2: Determine by Monte Carlo methods at which airport each of the accidents generated in Step 1 occurs.

Step 3: For each accident, determine the cost.

Step 4: Total the cost of all the accidents for the year.

Step 5: Repeat the above four steps n times (yielding a sample size of n years).

Step 6A: Compute the empirical national annual risk profile directly for the sample of n years from the Step 4 values.

Step 6B: Compute the empirical national conditional (given one accident) risk profile directly for the sample of n years' worth of accidents from the Step 3 values.

We have followed this procedure in our specialized simulation program, and will now give the essential details of that program, referring appropriately to the steps above.

In Step 1 we assume that the number of accidents in a year is a random variable following a Poisson distribution; thus only the mean μ of that distribution is required as an input quantity. While the assumption of a Poisson distribution seems to be generally accepted, and there are some theoretical bases for it, other discrete distributions or a more complex relationship could be accommodated by the program without significant alteration. For example, the mean μ of the Poisson distribution could be treated as a random variable following a specified probability distribution.

In Step 2 it is assumed that each accident occurs at or in the vicinity of one of a given number, say N_a, of airports. Each of these is characterized by a probability, say P_i for airport i, that an accidents occurring at one of the N_a airports in fact occurs at airport i. That is, P_i = Prob(accident occurs at airport i | it occurs at one of the N_a airports). These probabilities pertain to *every* simulated accident, regardless of the airports at which other simulated accidents occur; i.e., the various accidents are treated independently with respect to where they occur. The numerical values of the P_i $(i=1,\ldots,N_a)$ are based on historical data (and possibly on future projections as well) pertaining to the weather conditions and numbers of operations at the various airports.

Step 3 is the critical one and must be dependent on the simulation model used to generate the random costs of accidents that take place at each of the N_a airports. Such a model must be rather complex in order to account for various types and locations of accidents, CF dispersion under local weather conditions, local types and quantities of housing and industry, etc. Indeed, the development of such a model is one of the most significant aspects of the GFRAP, and both ADL and ORI have expended considerable effort to this end. Our specialized simulation program therefore requires the inclusion of subroutines that generate random values of accident cost that are solely dependent on the airport concerned. For the sake of future reference, let F_{di} denote the CDF of the damage (cost) per accident at airport i , $i=1,\ldots,N_a$. To test our simulation program and to obtain sensitivity results we had to make some specific choices for the F_{di} . What we did is described below in Section 4.4.2.

Steps 4 and 5 are self-explanatory and require no comment. Steps 6A and 6B are carried out by constructing, in each case, an aggregated relative frequency distribution and then converting this to the aggregated complementary cumulative relative frequency distribution that constitutes the empirical risk profile. The number and size of the cells used to construct the aggregated distributions are specified as input to the program. It would not be practical, for large sample sizes, to save all the observed values because of their number and the computational cost of ordering then (there are n observed values that enter into Step 6A and approximately $n\mu$ observed values that enter into Step 6B). The first four moments of the empirical distributions are computed from the actual observed values.

4.4.2 Approximating the airport damage distributions

The only information that has been available about airport damage distributions is that contained in Table 10-1 of Kalelkar, *et al.* (1979). Recent results of both ADL and ORI indicate that more accurate and current damage distributions have associated with them significantly smaller damage values. As such results were not available when our program was developed and when our sensitivity runs were made, and are still not readily available, we based our work on the empirical damage distributions presented in ADL's Table 10-1. The distributions given there for 26 different airports are based on 300 values each generated by the ADL damage simulation model.

For each of the 26 distributions, ADL's Table 10-1 provides the mean, the standard deviation, the minimum and maximum observed values, and the following percentiles: 5th, 10th, 25th, 50th, 75th, 90th, and 95th. All the minima are zero, as are many of the 25th percentiles. Using the given percentiles, we plotted all 26 distributions on log-log paper in the form of risk profiles (complementary CDF's). One of these, that for Washington, D.C.'s National Airport, is shown in Figure 4-12. The right-hand portion of the curve is shown as a dotted line to emphasize the fact that this portion is heavily dependent upon the maximum value observed and said maximum might vary significantly among samples of size 300. (We associated with the maximum value a risk probability of 1/300.)

We considered using piecewise-linear approximations to the observed distributions in Step 3 of our simulation program, but decided against it for two reasons. First, and most important, this would not provide us with a convenient mechanism for performing sensitivity analyses through controlled changes in the damage distributions used in Step 3. Second, the smooth form of the curve in Figure 4-12 (and of the ones for other airports) suggests that a piecewise log-log or log-linear approximation might be more appropriate. We therefore concentrated on analytic approximations to the observed damage distributions.

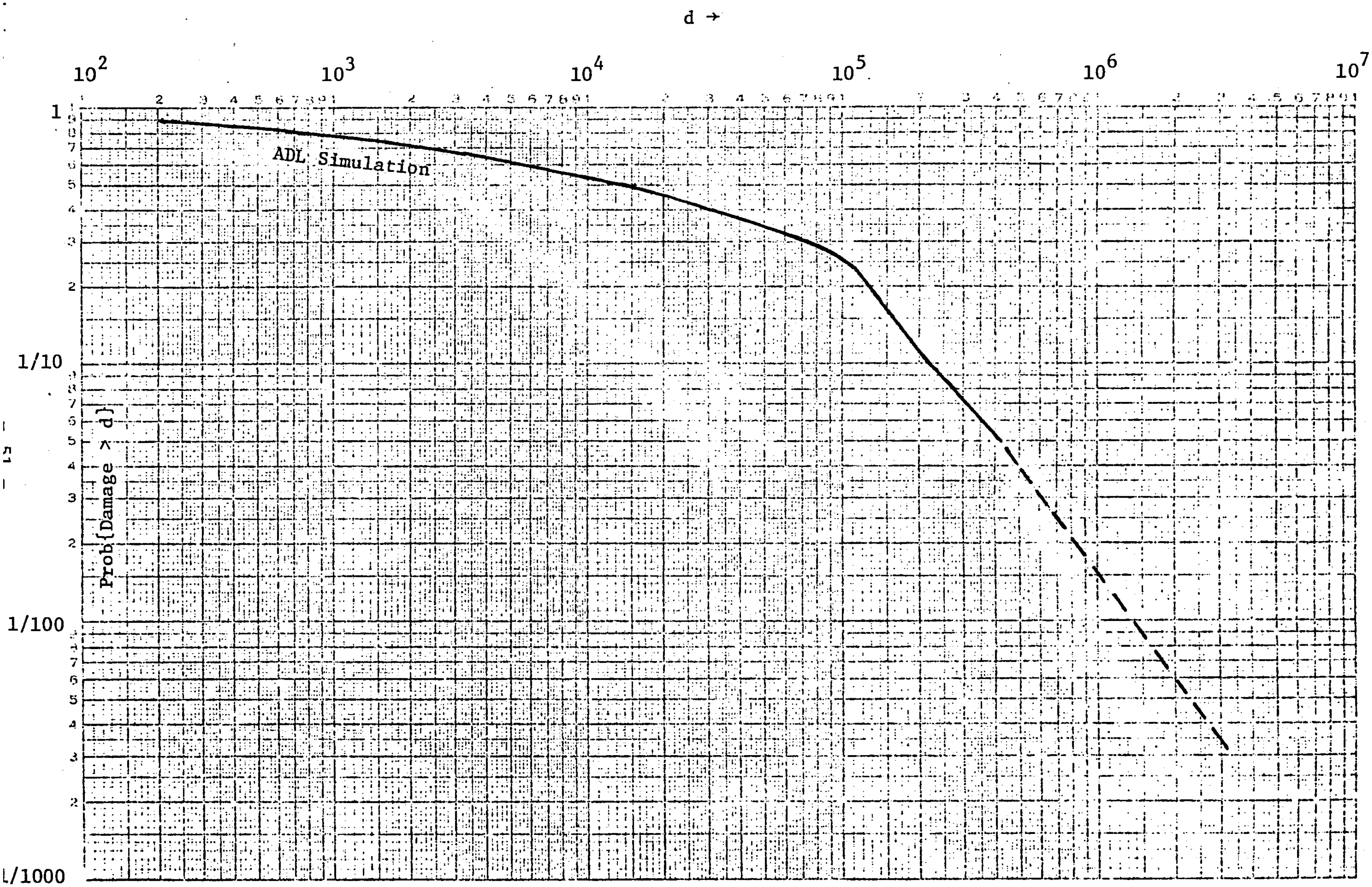


Figure 4-12.--Washington National Airport risk profile.

Examination of the 26 empirical distributions in Table 10-1 of Kalelkar *et al.* (1979) shows them to be considerably skewed to the right; the coefficients of variation range from 1.33 to 3.40 and the ratios of mean to median range from 1.32 to 133. We therefore considered several families of distributions which possess the characteristics of skewness to right and nonnegativity (of the associated random variable). Such families include the Weibull, Gamma, Pareto, and Lognormal families [see Mann *et al.* (1974)]. Our choice of families to examine was partly influenced by the manner in which we decided to fit the empirical distributions and verify the adequacy of the fit. We wanted to determine the two parameters of the theoretical distribution from the mean and standard deviation of the empirical one, and then to compare the two risk profiles. This is difficult to do with the Weibull and Gamma distributions for the following reasons: In the case of the Weibull distribution one cannot find directly the two parameters as closed-form functions of the empirical mean and standard deviation; a nonlinear equation involving the Gamma function must be solved. In the case of the Gamma distribution one cannot obtain the CDF in closed form so calculation of the theoretical risk profile is extremely difficult.

Both the Lognormal and Pareto distributions fit our computational requirements, and examination of the 26 corresponding Lognormal risk profiles showed many of them to fit the empirical risk profiles reasonably well. The Pareto distribution was tried for several cases but did not fit quite as well as the Lognormal.

For the purposes of testing the simulation program we therefore decided to use Lognormal distributions with means and variances equal to those of the empirical distributions. Figures 4-13 through 4-21 show a representative selection of the corresponding risk profiles--the ADL empirical ones and the Lognormal ones used to fit them.

We did not close the door on attempts to fit the empirical distributions; as a next step we examined the possibility of fitting the empirical distributions by mixtures of Lognormal distributions. That is,

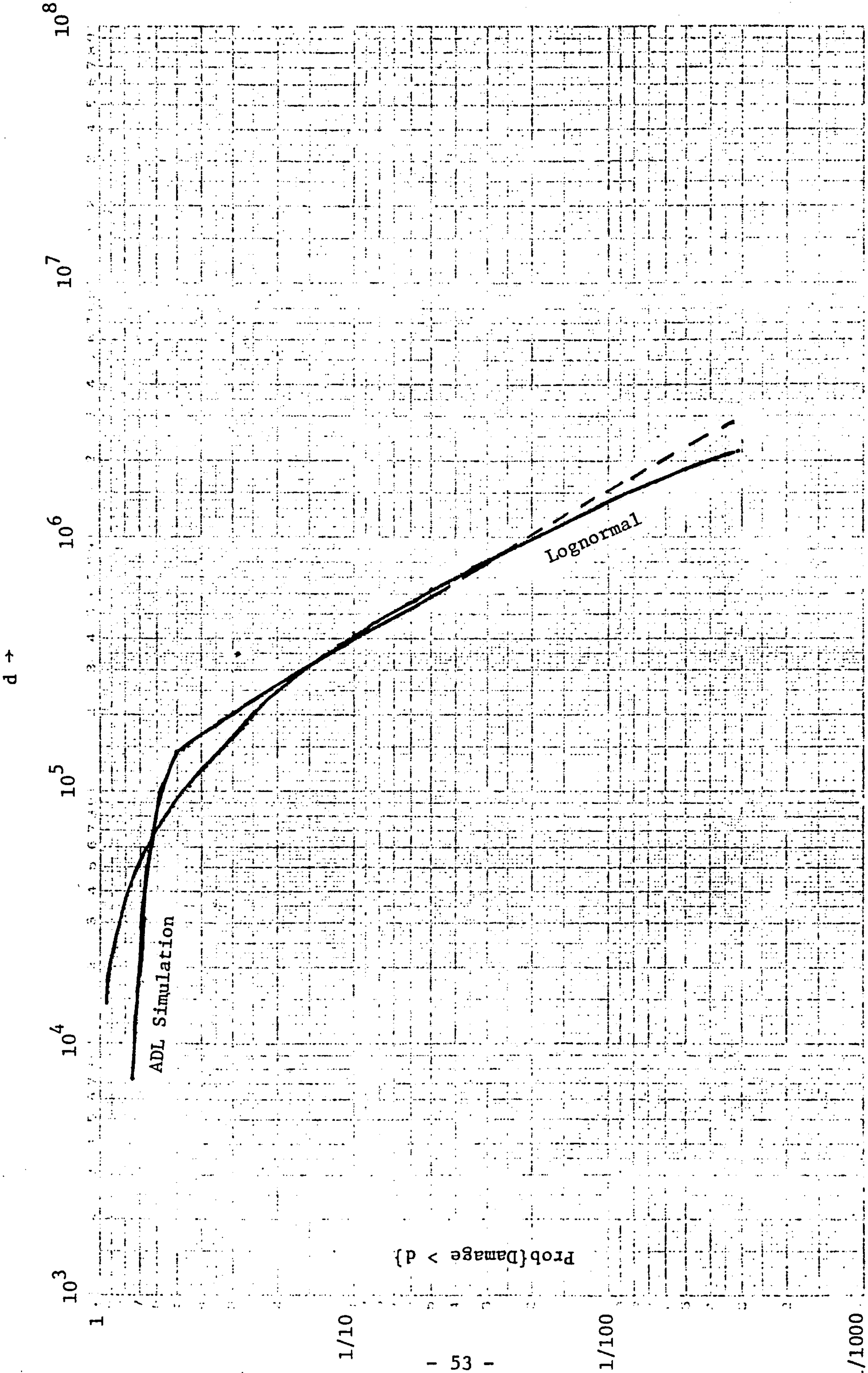


Figure 4-13.--Chicago O'Hare risk profile.

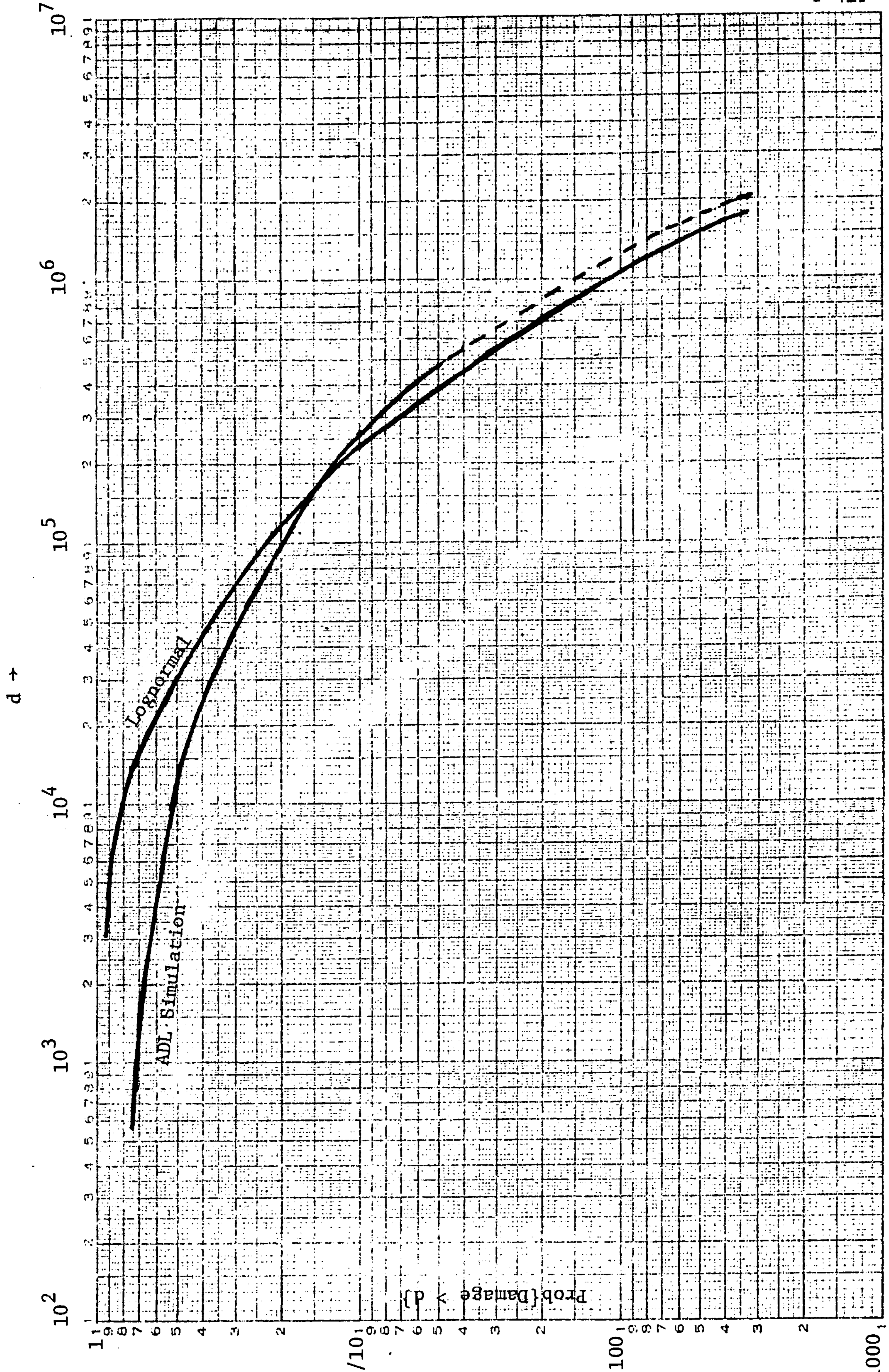


Figure 4-14.--Denver risk profile.

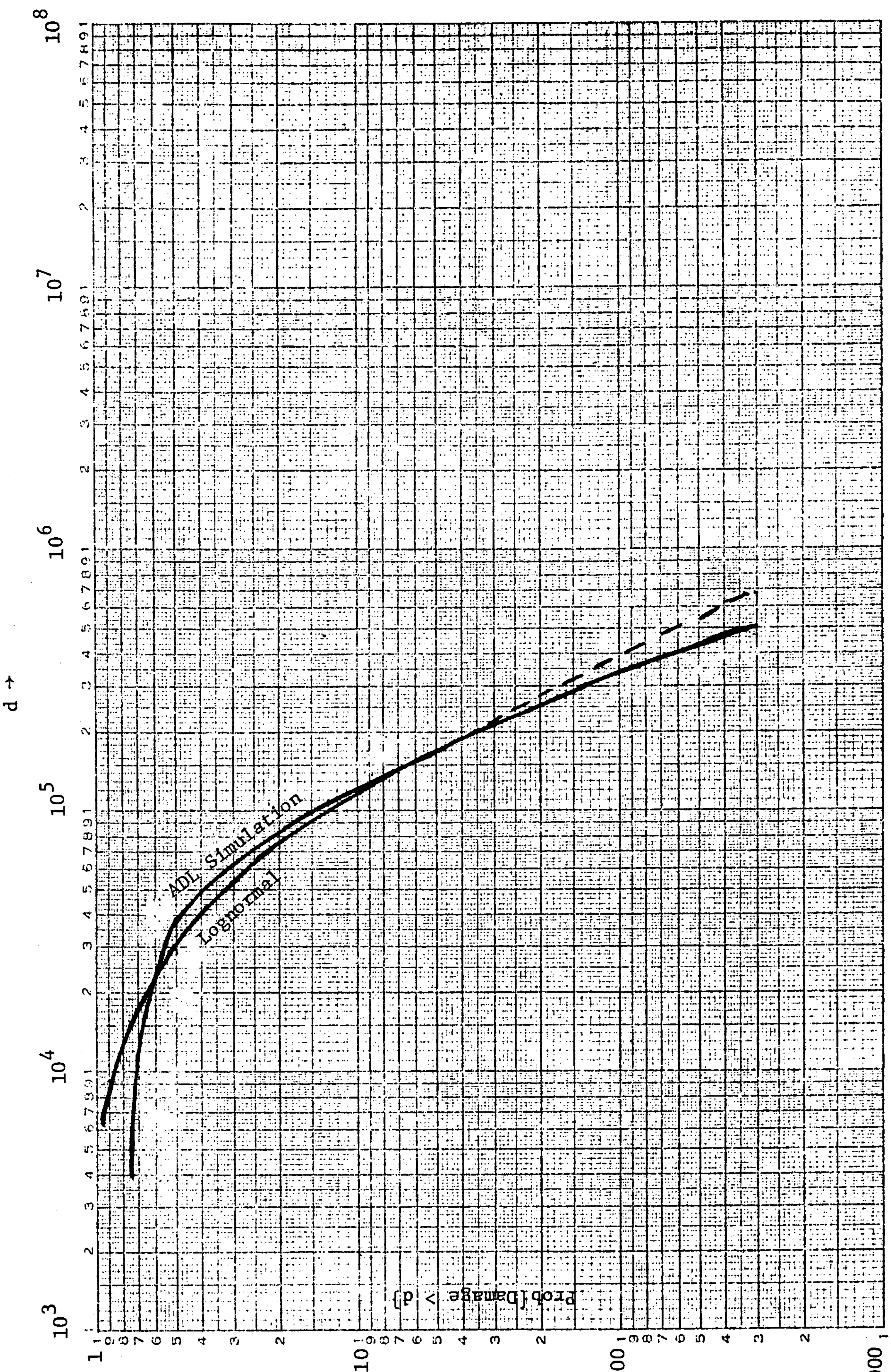


Figure 4-15--Detroit risk profile.

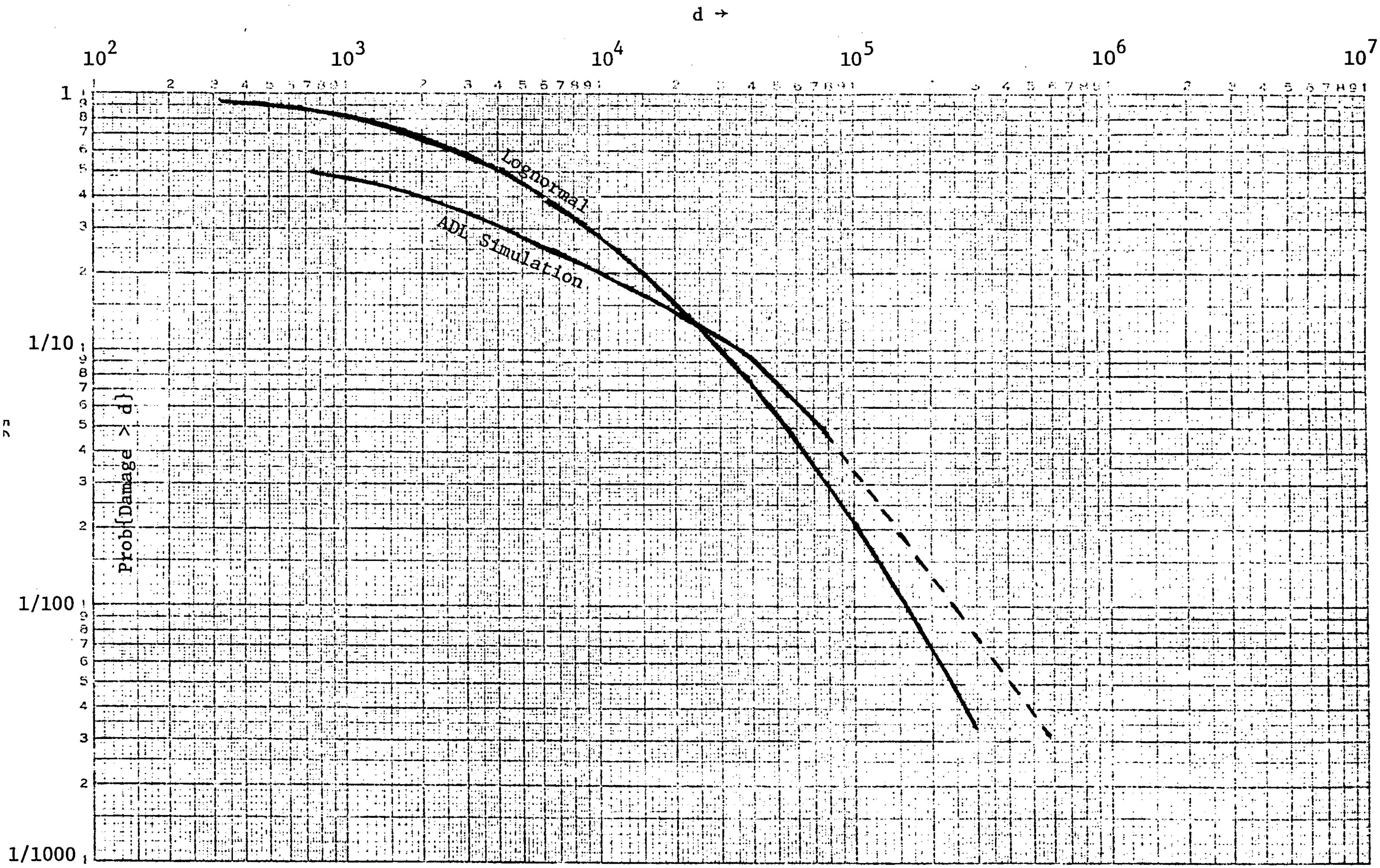


Figure 4-16.--Houston risk profile.

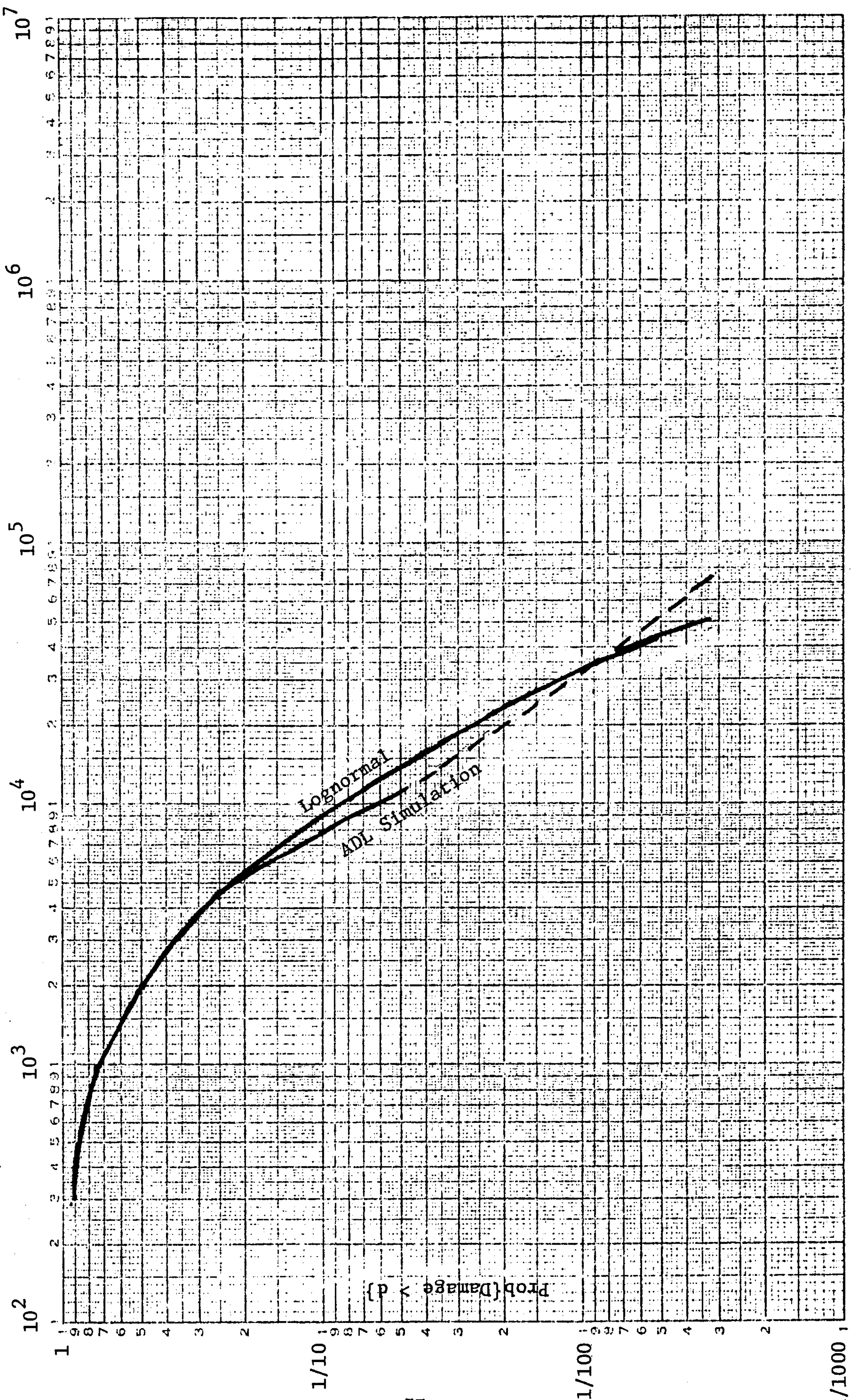


Figure 4-17.—Kansas City risk profile.

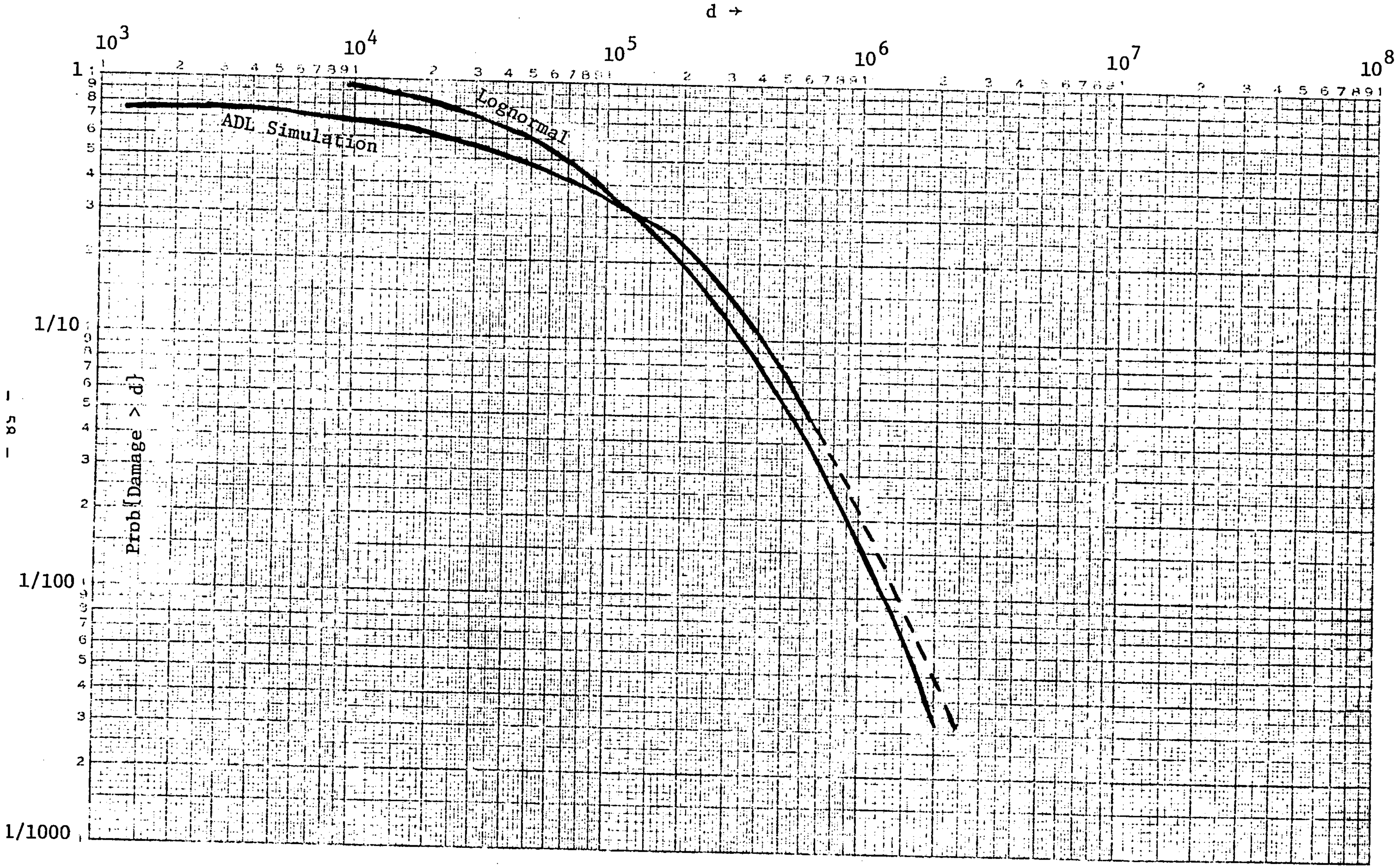


Figure 4-18.--Kennedy risk profile.

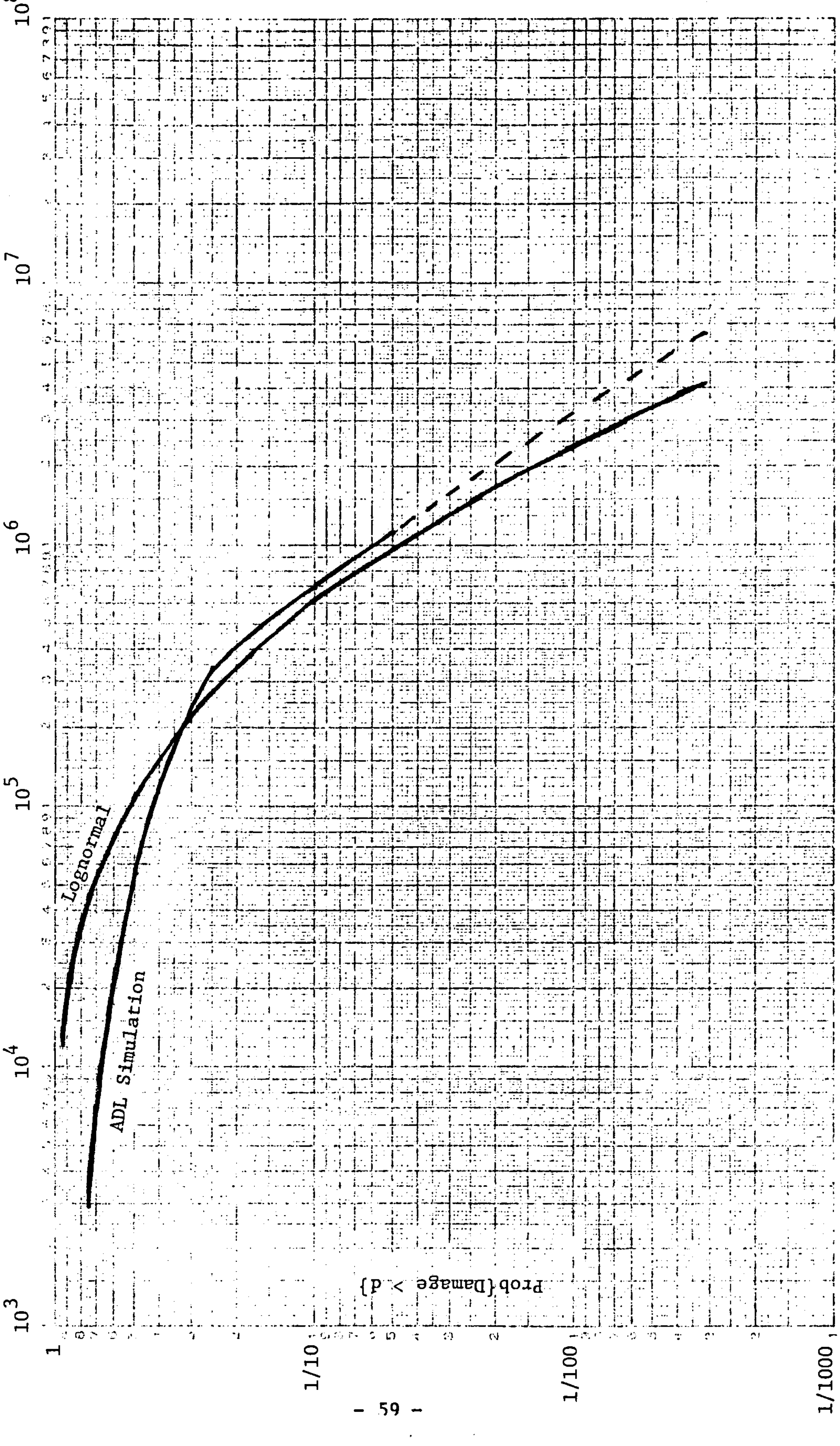


Figure 4-19.--LaGuardia risk profile.

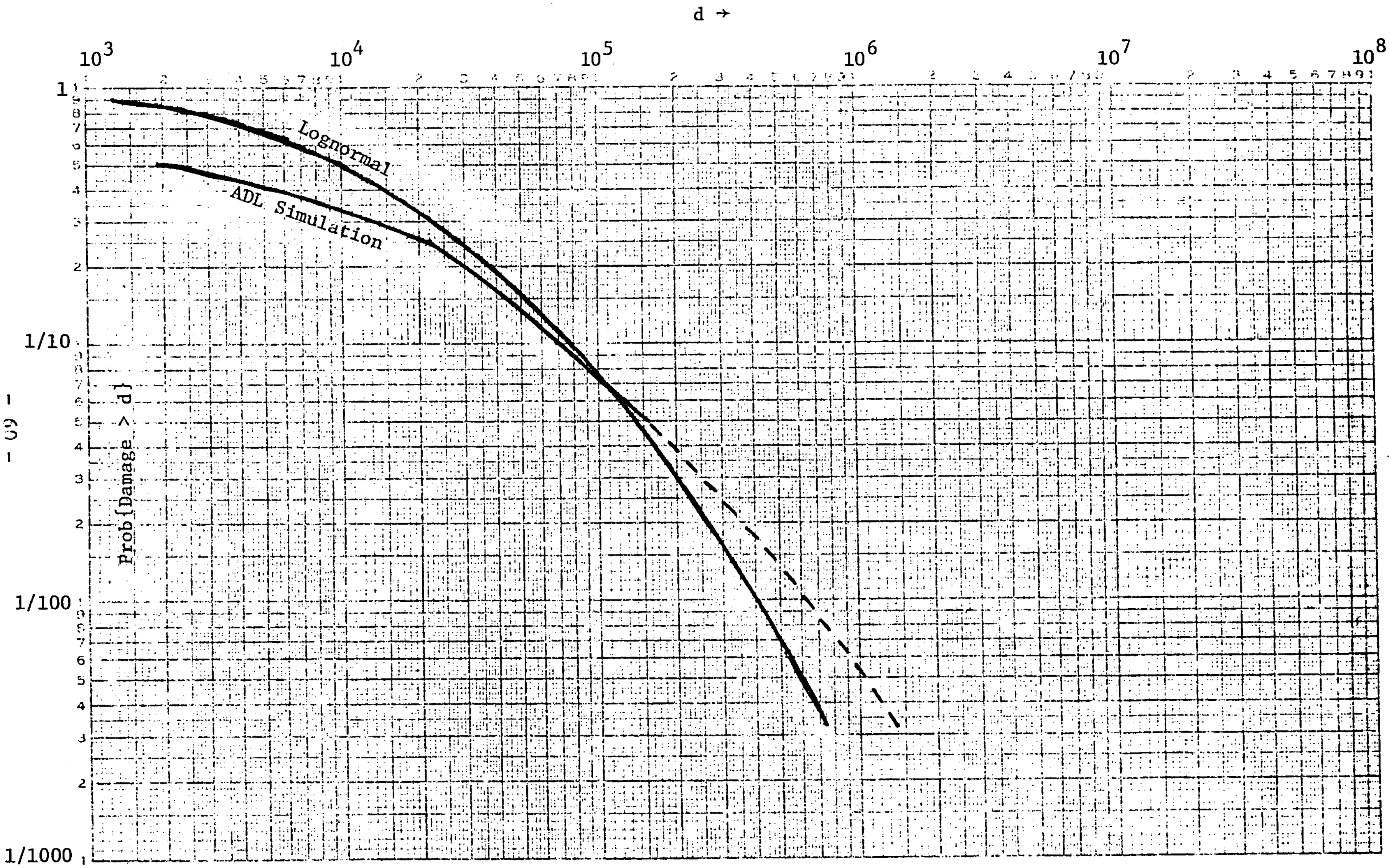


Figure 4-20.--New Orleans risk profile.

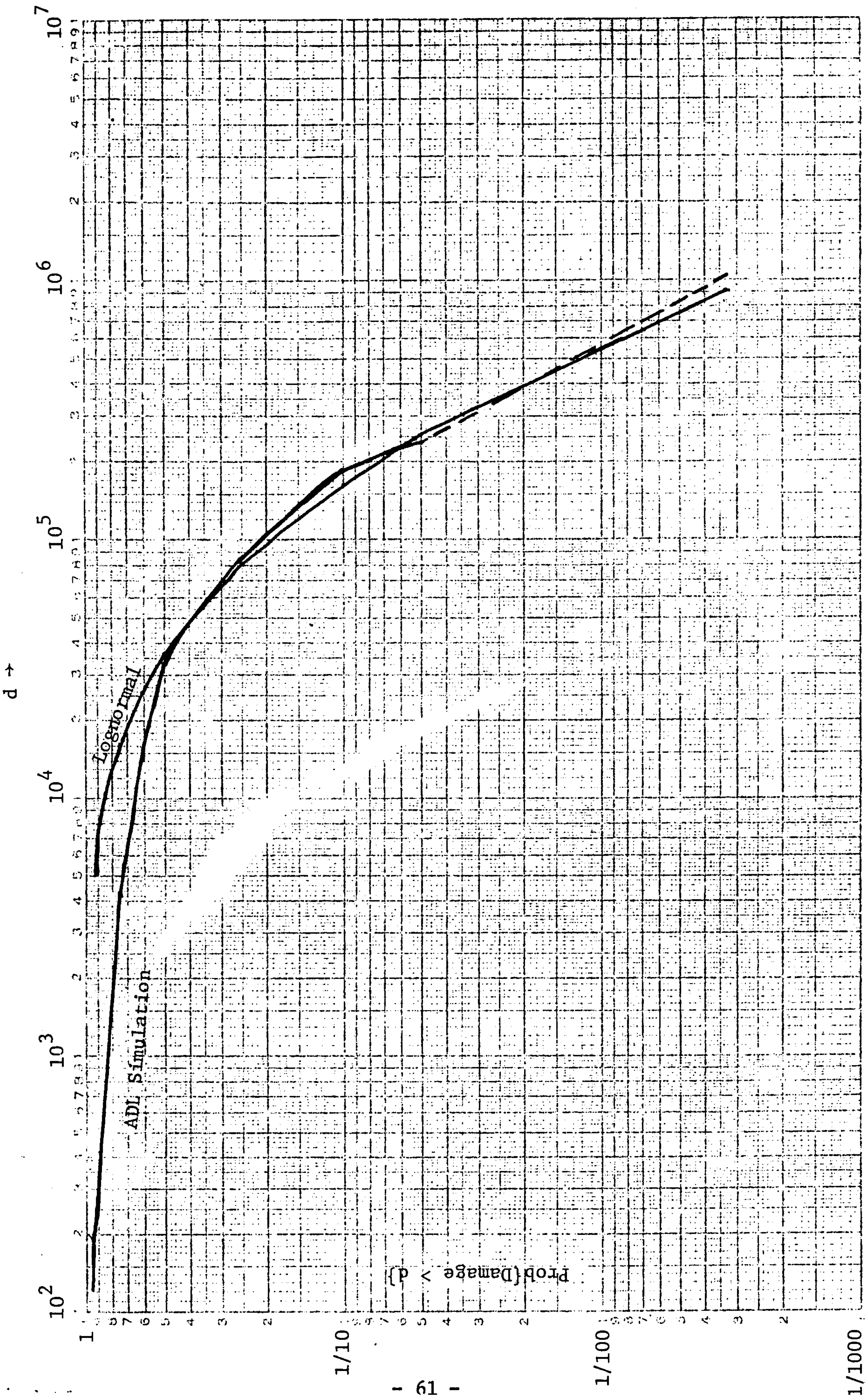


Figure 4-21.--St. Louis risk profile.

instead of approximating the empirical CDF $\hat{F}_{di}$ for airport i by a Lognormal CDF $F_{\ell i}$, we approximated it by the convex combination of Lognormal CDF's

$$\sum_{j=1}^{I_i} q_{ij} F_{\ell ij} ,$$

where each $F_{\ell ij}$ is a Lognormal CDF, each I_i is a small positive integer such as 2 or 3 (or even 1), the q_{ij} are positive, and $\sum_j q_{ij} = 1$. A difficulty with this approach is the lack of a clear criterion to use in choosing I_i, the q_{ij}, and the $F_{\ell ij}$. Even with a clear criterion, the computational effort would be considerable. There is no doubt, however, that decidedly better fits could be obtained this way. Figures 4-22(a) and 4-22(b) show fits to the Philadelphia airport risk profile by a single Lognormal and by a mixture of two Lognormals, respectively; the parameters in the case of the mixture were found heuristically. The improved fit is readily apparent. Figures 4-23(a) and 4-23(b) show similar fits to the Los Angeles airport risk profile.

An interesting result is the following: Let F be the national conditional (given one accident) CDF whose corresponding risk profile is estimated in Step 6B of the specialized simulation program. If each CDF F_{di} is assumed to be Lognormal, say $F_{\ell i}$, then

$$F = \sum_{i=1}^{N_a} p_i F_{\ell i} ,$$

so that F is a mixture of Lognormals. If instead each F_{di} is assumed to be the mixture of Lognormals $\sum_j q_{ij} F_{\ell ij}$, then

$$F = \sum_{i=1}^{N_a} \sum_{j=1}^{I_i} p_i q_{ij} F_{\ell ij} ,$$

so that F is again a mixture of Lognormals.

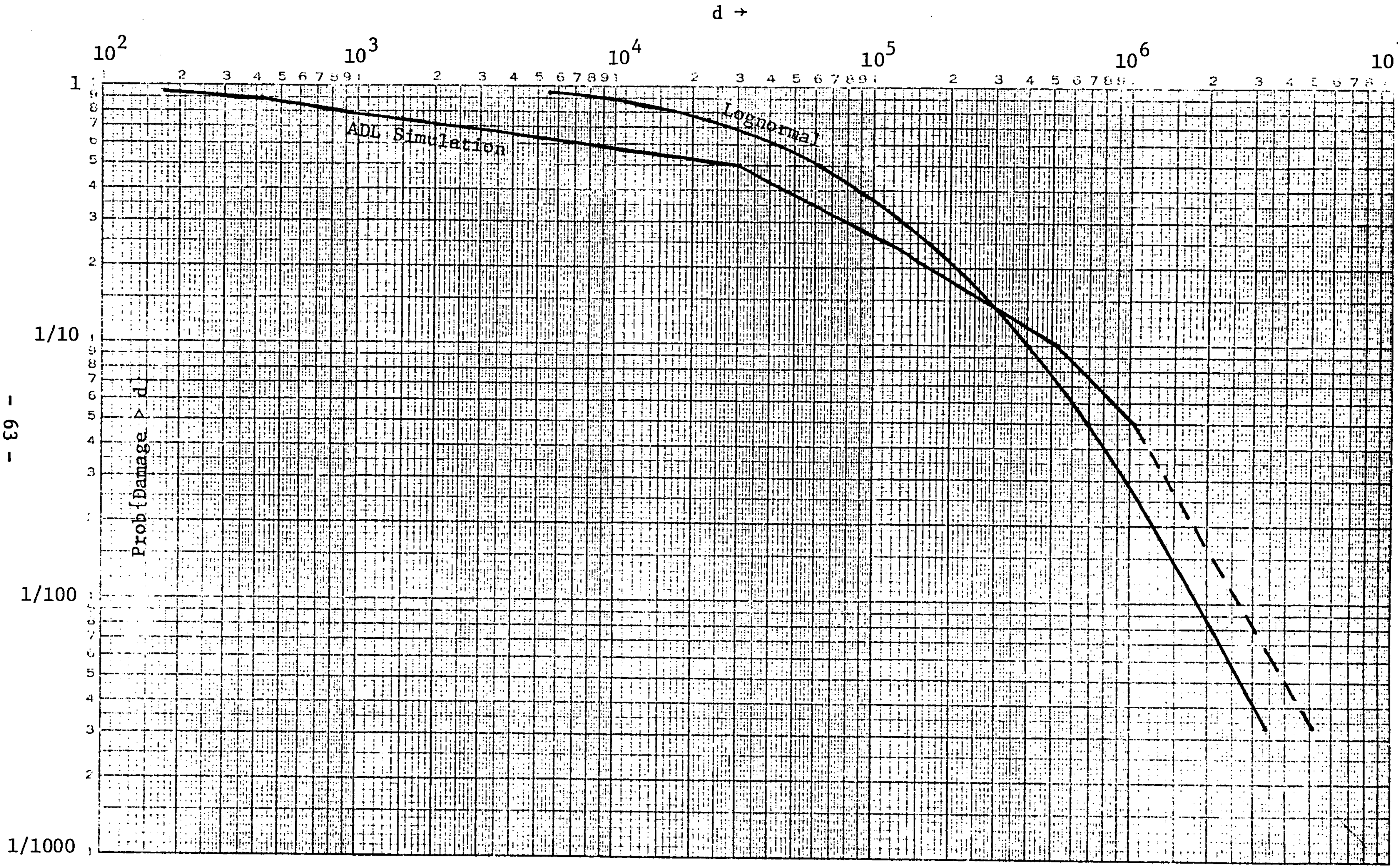


Figure 4-22(a).--Philadelphia risk profile.

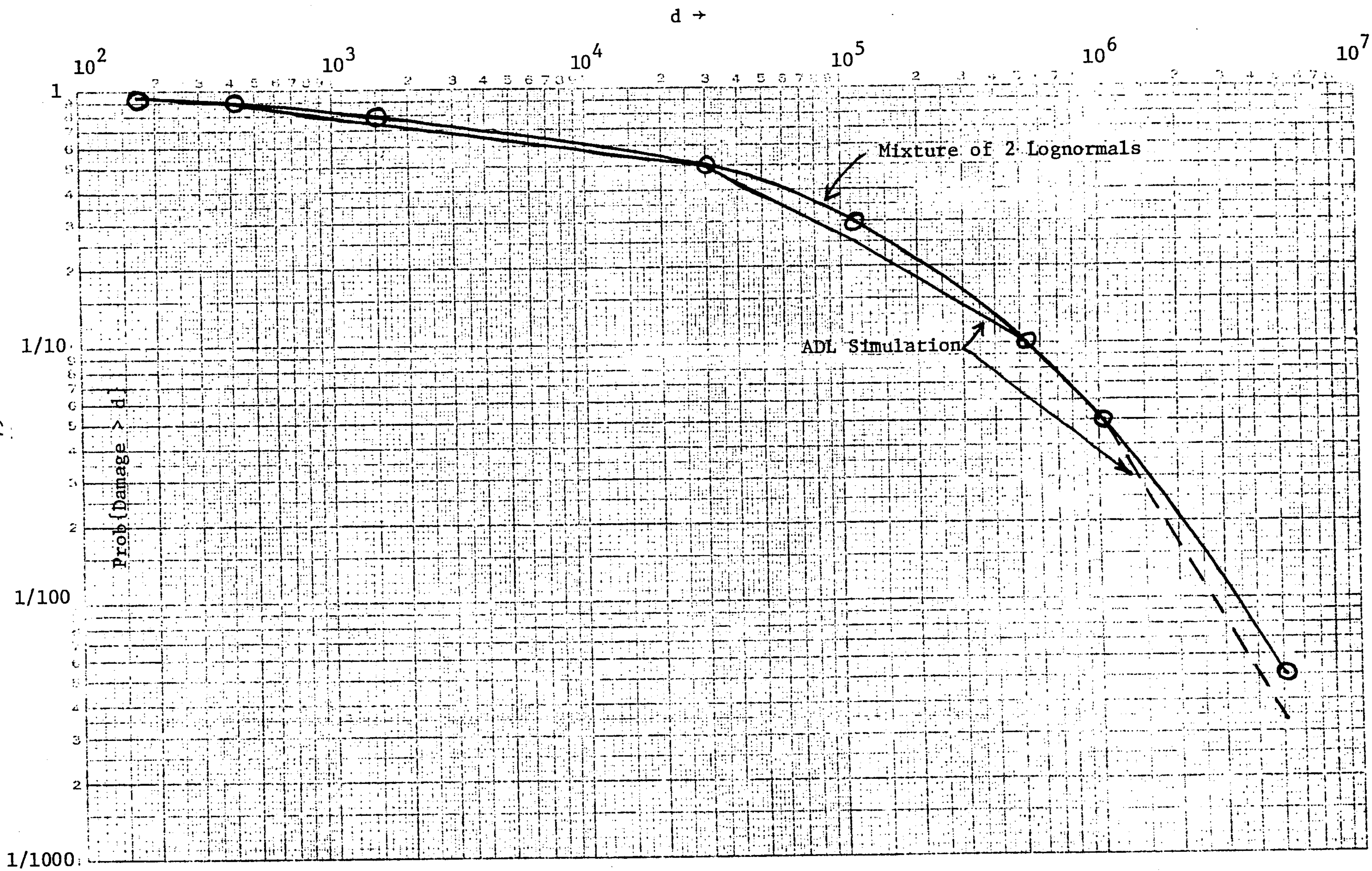


Figure 4- 22(b).--Philadelphia risk profile.

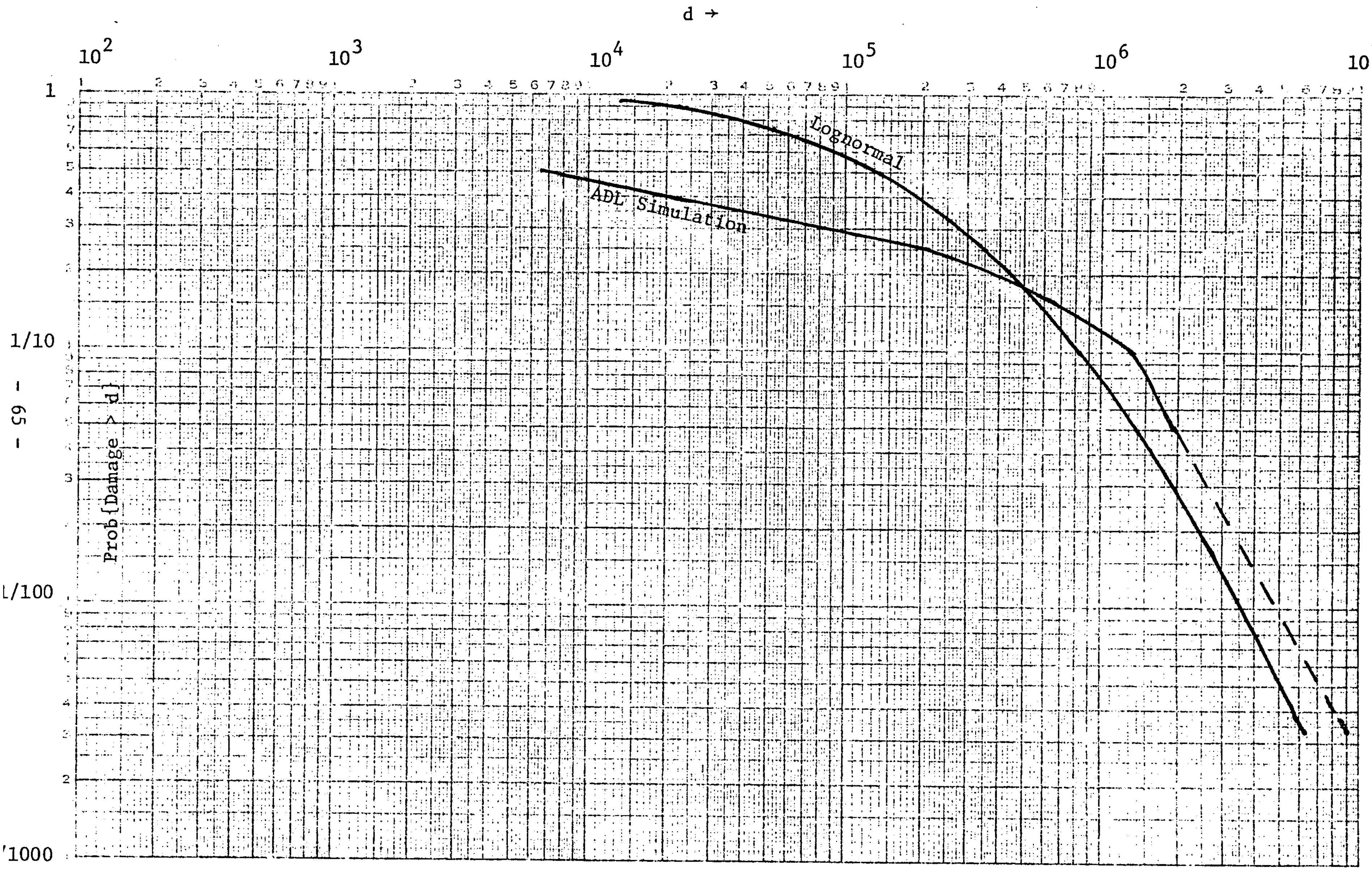


Figure 4-23(a).--Los Angeles risk profile.

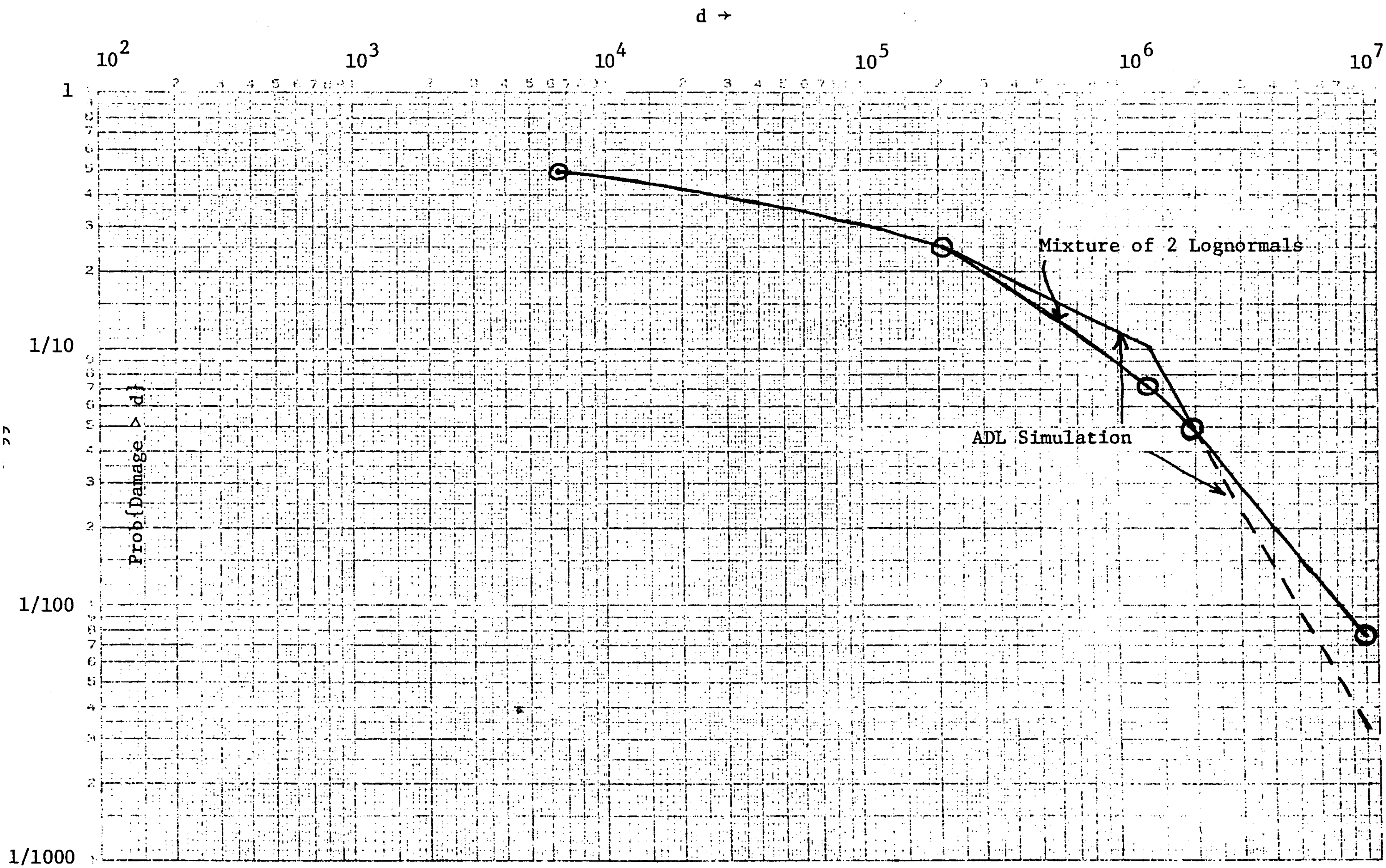


Figure 4-23(b).—Los Angeles risk profile.

Because the empirical distributions given in Table 10-1 of Kalelkar, *et al.* (1979) were known to be outdated, we decided against any further attempt to fit them with analytical distributions. We decided to use 26 Lognormal distributions with the same means and variances as those of the empirical distributions as the basis of our sensitivity analyses. We did, however, investigate the effect of replacing the Lognormal distributions by mixtures of Lognormal distributions. Exactly what was done will be detailed below, but the results may be summarized by the statement that replacing the Lognormal distributions by mixtures of Lognormal distributions had a negligible effect on the national conditional and annual risk profiles obtained, even though the mixture of Lognormals yields a better fit of the individual airport damage distributions.

4.4.3 The sensitivity analyses

Following the ADL analyses we used $N_a = 26$ and used for the p_i the values given in the last column of Table C-4 of Kalelkar *et al.* (1979). For our base case we chose the 26 Lognormal distributions to have means and standard deviations equal to those of the empirical distributions generated by ADL, and we set $\mu = 2.6$. This parameter μ is the expected number of accidents in a year, and the value of 2.6 corresponds to the agreed-upon projection for 1993.

Because of the assumption of Lognormal distributions of damage per accident at the various airports, the national conditional risk profile (given one accident) is the complementary cumulative distribution function (CDF) of a mixture of Lognormal CDF's. That is,

$$F = \sum_{i=1}^{N_a} p_i F_{\ell i} ,$$

where $F_{\ell i}$ is the Lognormal CDF of damage per accident at airport i and $R = (1-F)$ is the national conditional risk profile (given one accident). It is thus possible to compute the values of the risk profile R exactly with the aid of a table or computer program that provides CDF values of the standard normal distribution. (This is not true for the annual risk profile, however.) A separate computer program was therefore written to compute R this way, and it was used in adddition to check the corresponding risk profiles generated by the Monte Carlo simulation program (which generates both risk profiles).

The annual risk profiles for various sensitivity runs were generated by the simulation program; in each case the year 1993 was simulated 4000 times.

We performed three separate types of sensitivity analyses. In the first, we varied the means of the Lognormal distributions of damage per accident at the various airports and kept the standard deviations the same. In the second, we varied the standard deviations of the Lognormal distributions of damage per accident at the various airports and kept the means the same. In both cases all 26 means or standard deviations were changed by the same percentage. In the third type of sensitivity analysis we varied μ, the mean number of accidents in a year. This third type of sensitivity analysis affects the annual risk profile but not the national conditional risk profile (given one accident), whereas the first two types affect both risk profiles.

The results are shown in Figures 4-24 through 4-28. Figures 4-24 and 4-25 show the national conditional risk profiles R for various changes in the means (Figure 4-24) and standard deviations (Figure 4-25) of the distributions of damage per accident at each airport. Each risk profile is identified by the ratio of the value of the parameter (mean or standard deviation) to its value in the base case. Tables 4-3 and 4-5 show numerically some of the results shown graphically in Figures 4-24 and 4-25; no exact values are given since they are not likely to be truly meaningful, but instead ratios are given to indicate the effects of the change in means or standard deviations of the distributions of damage per accident at each airport. Tables 4-4 and 4-6 show the ratios of the means and standard deviations of the national conditional risk

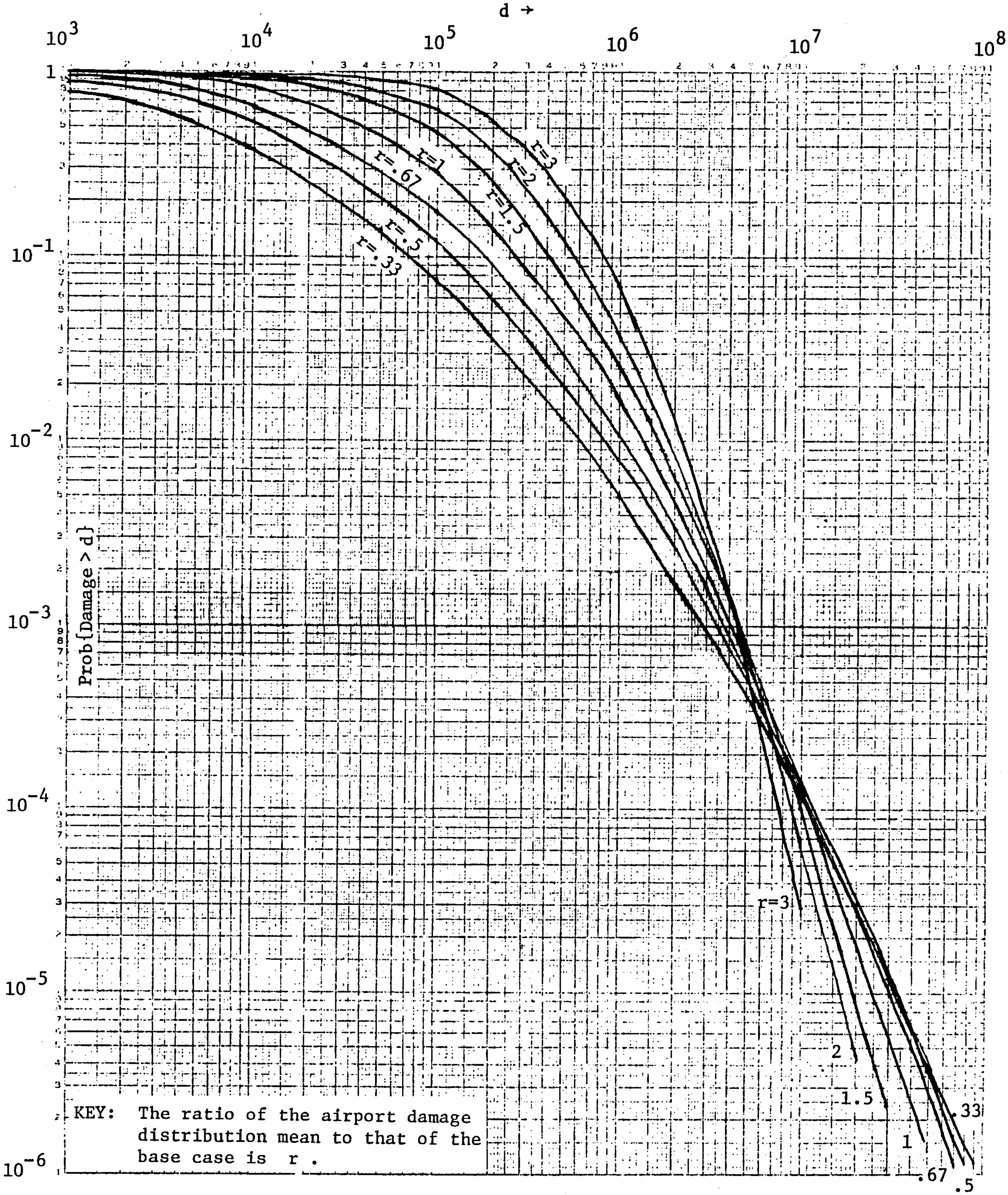


Figure 4-24--National conditional risk profile R with changes in the means of the airport damage distributions (26 airports).

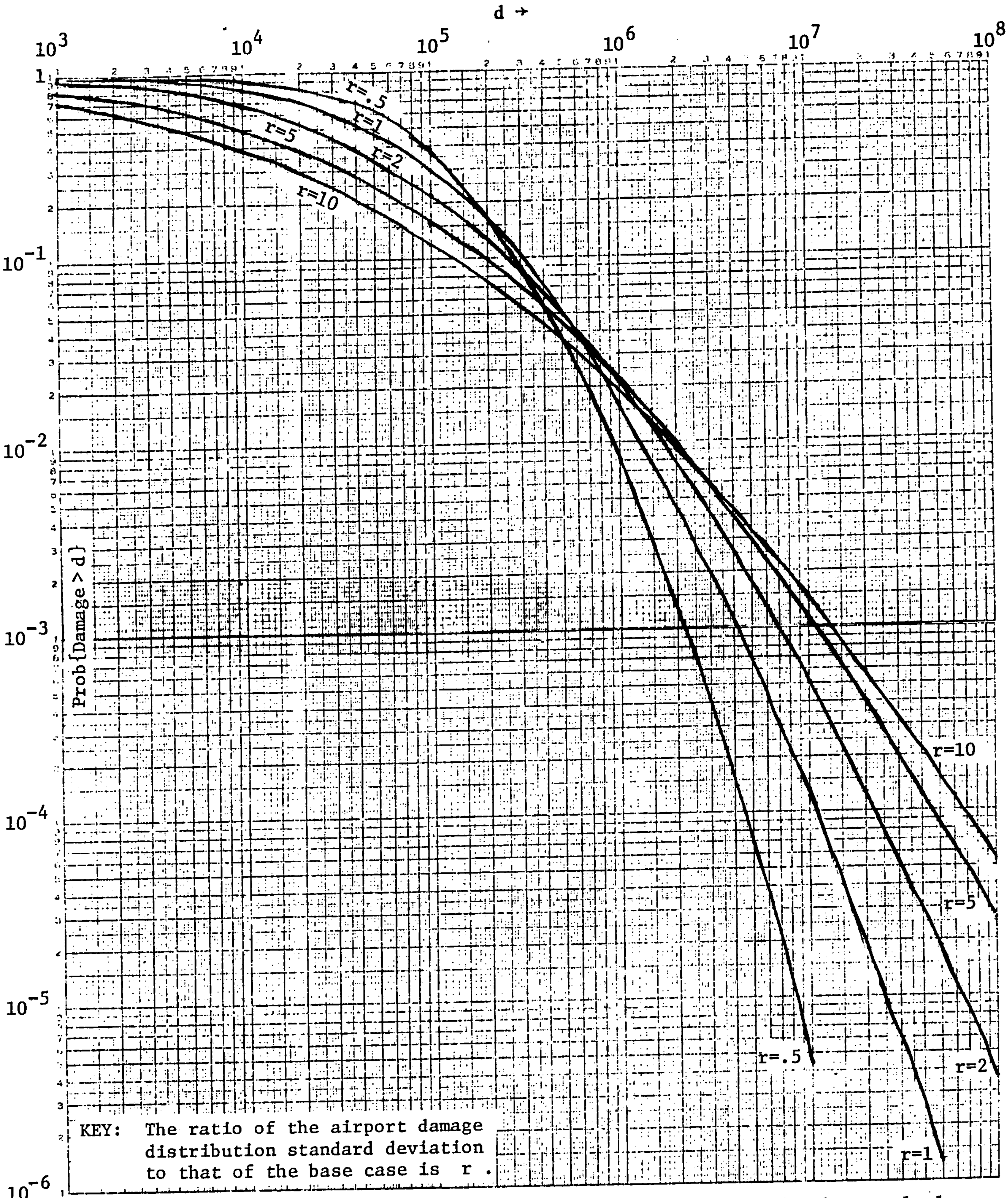


Figure 4-25.--National conditional risk profile R with changes in the standard deviations of the airport damage distributions (26 airports).

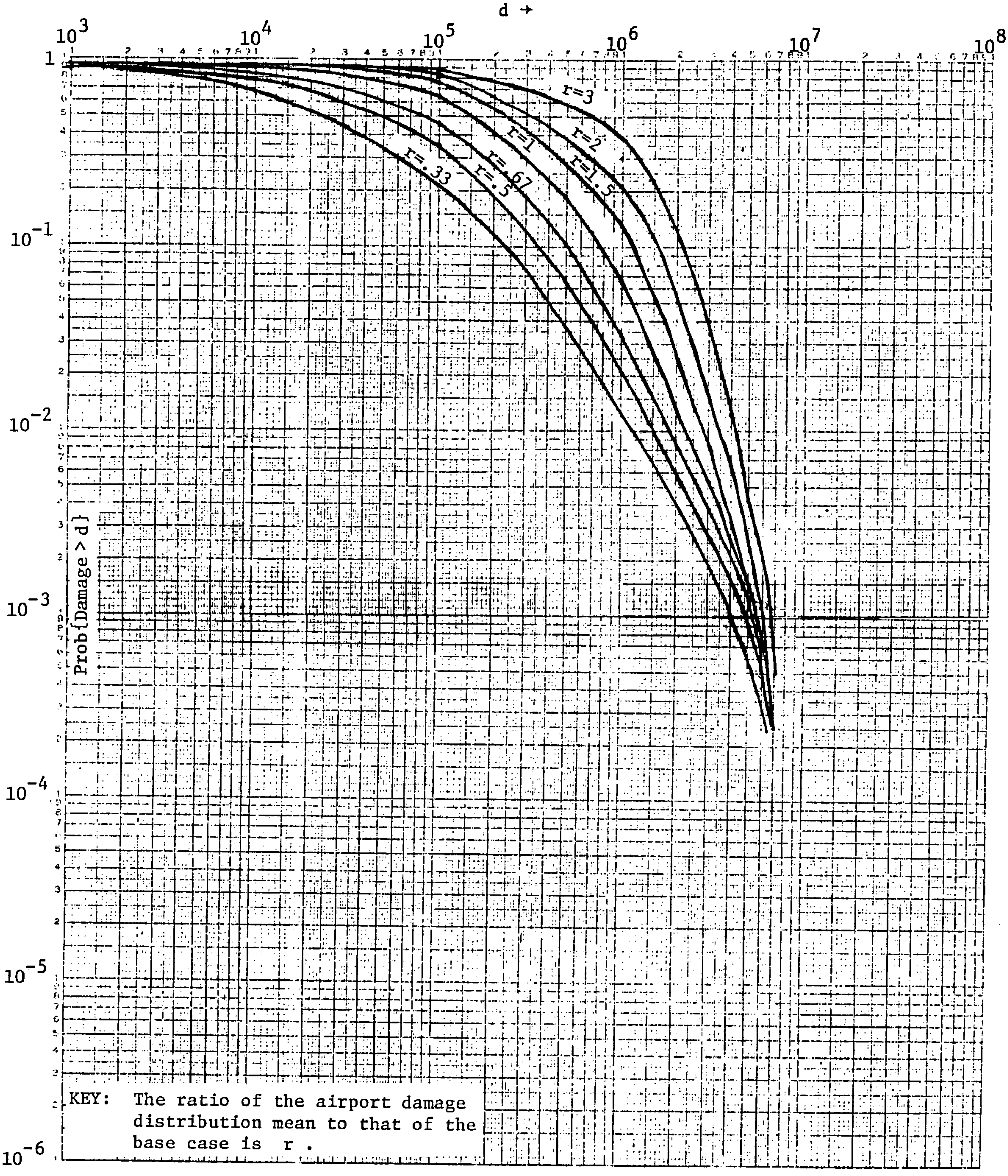


Figure 4-26.--Annual risk profile S with changes in the means of the airport damage distributions (26 airports).

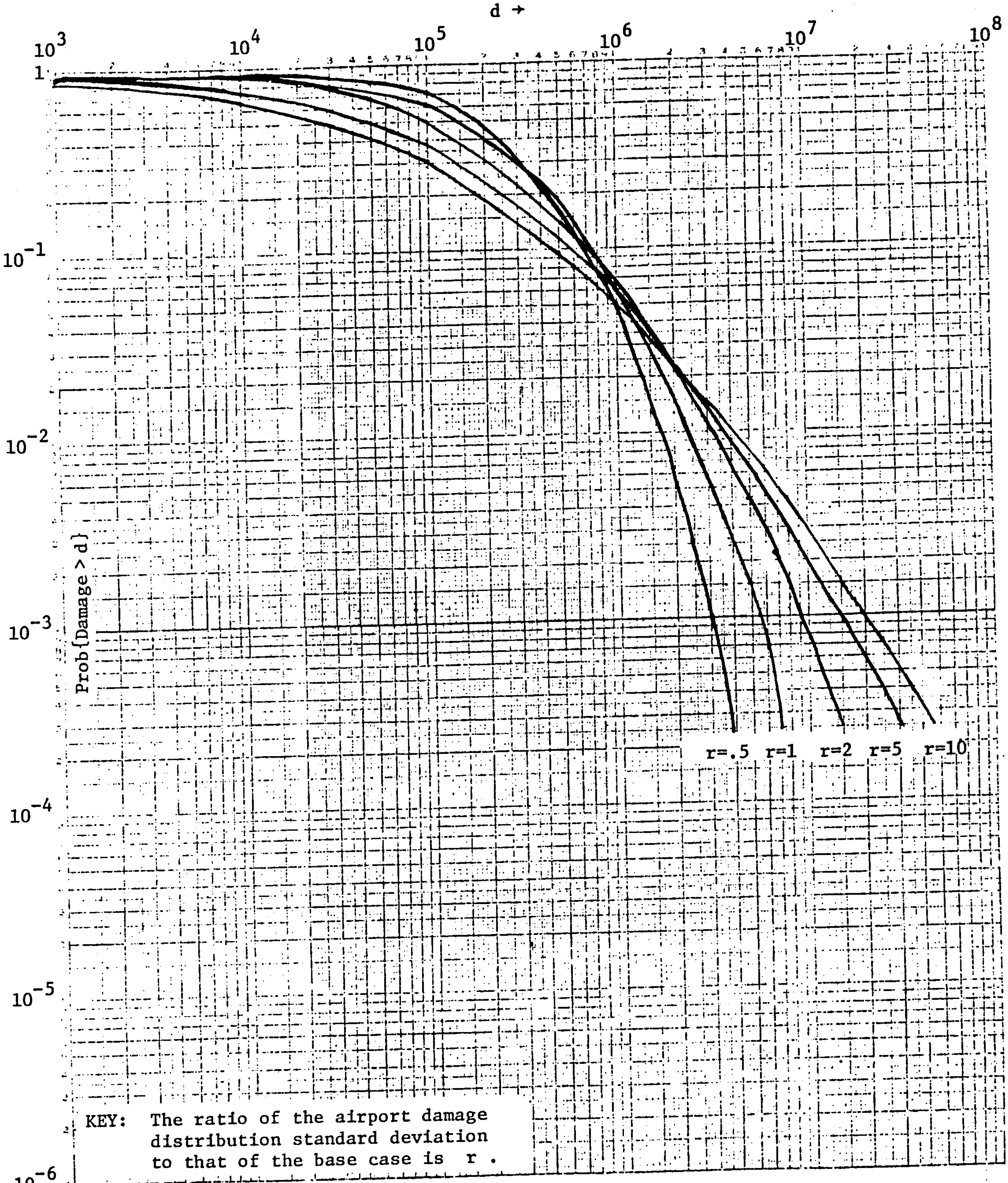


Figure 4-27. -Annual risk profile S with changes in the standard deviations of the airport damage distributions (26 airports).

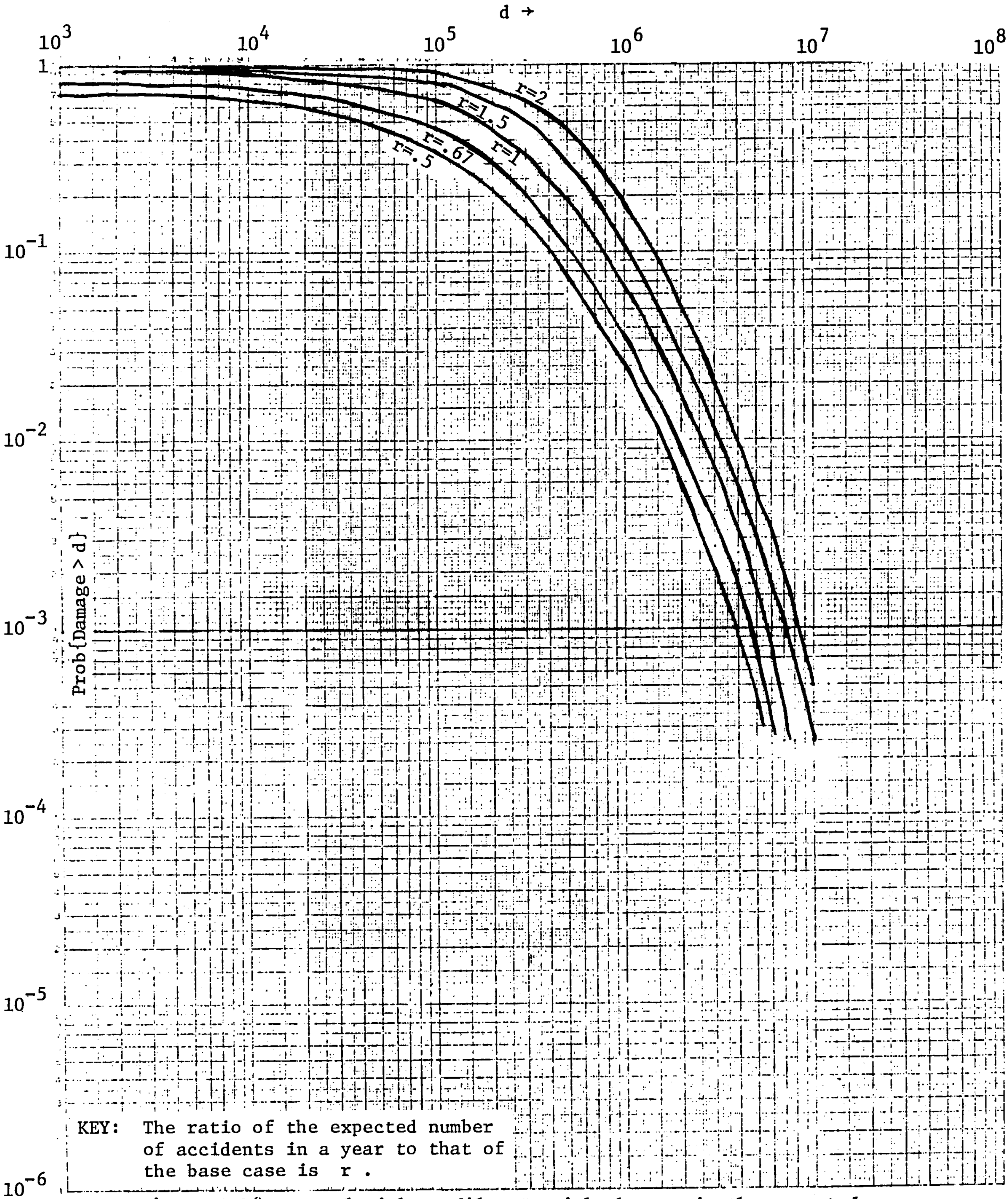


Figure 4-28r--Annual risk profile S with changes in the expected number of accidents in a year (26 airports).

TABLE 4-3

SENSITIVITY OF THE NATIONAL CONDITIONAL RISK PROFILE R TO CHANGES IN THE MEANS OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the means are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution mean to that of the base case is r and d is a numerical damage value.

P{Damage > d}	*(a) Damage Value d* r = 0.5	r = 1.5	r = 2.0	r = 3.0
3×10^{-1}	.31	1.87	2.69	4.06
1×10^{-1}	.45	1.53	2.07	2.98
3×10^{-2}	.52	1.30	1.60	2.06
1×10^{-2}	.62	1.21	1.39	1.63
3×10^{-3}	.75	1.13	1.20	1.31
1×10^{-3}	.84	1.07	1.10	1.12
3×10^{-4}	1.00	1.06	1.06	.88
1×10^{-4}	1.04	.91	.81	.69
3×10^{-5}	1.20	.83	.72	.60

Damage Value d	*(b) P{Damage > d}* r = 0.5	r = 1.5	r = 2.0	r = 3.0
$.03 \times 10^{6}$	.50	1.36	1.53	1.65
$.1 \times 10^{6}$	.42	1.65	2.21	2.81
$.3 \times 10^{6}$	.42	1.76	2.76	4.33
1×10^{6}	.46	1.53	2.21	4.00
3×10^{6}	.65	1.30	1.50	1.95
10×10^{6}	1.07	.79	.48	.23

TABLE 4-4

SENSITIVITY OF THE NATIONAL CONDITIONAL RISK DISTRIBUTION TO CHANGES IN THE MEANS OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the means are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution mean to that of the base case is r .

Ratio r	Mean of the National Conditional Risk Distribution	Standard Deviation of the National Conditional Risk Distribution
.33	.33	.97
.50	.50	.98
.67	.67	.98
1.5	1.50	1.04
2.0	2.00	1.09
3.0	3.00	1.23

TABLE 4-5

SENSITIVITY OF THE NATIONAL CONDITIONAL RISK PROFILE R TO CHANGES IN THE STANDARD DEVIATIONS OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the standard deviations are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution standard deviation to that of the base case is r and d is a numerical damage value.

	(a) Damage Value d			
P{Damage > d}	r = 0.5	r = 2.0	r = 5.0	r = 10.0
3×10^{-1}	1.23	.60	.35	.19
1×10^{-1}	.92	.84	.65	.44
3×10^{-2}	.79	1.06	1.04	.89
1×10^{-2}	.72	1.33	1.56	1.44
3×10^{-3}	.60	1.49	2.02	2.23
1×10^{-3}	.52	1.64	2.59	3.12
3×10^{-4}	.47	1.87	3.41	4.53
1×10^{-4}	.41	2.03	4.14	6.04
3×10^{-5}	.38	2.38	5.48	8.33
	(b) P{Damage > d}			
Damage Value d	r = 0.5	r = 2.0	r = 5.0	r = 10.0
$.03 \times 10^{6}$	1.27	.76	.55	.41
$.1 \times 10^{6}$	1.30	.74	.54	.40
$.3 \times 10^{6}$	.86	.85	.68	.54
1×10^{6}	.51	1.34	1.34	1.19
3×10^{6}	.21	2.13	2.99	2.99
10×10^{6}	.03	3.86	8.48	10.98

TABLE 4-6

SENSITIVITY OF THE NATIONAL CONDITIONAL RISK DISTRIBUTION
TO CHANGES IN THE STANDARD DEVIATIONS
OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the standard deviations are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution standard deviation to that of the base case is r .

Ratio r	Mean of the National Conditional Risk Distribution	Standard Deviation of the National Conditional Risk Distribution
0.5	1.0	.55
2.0	1.0	1.95
5.0	1.0	4.84
10.0	1.0	9.68

profile R to those of the base case when the indicated changes are made in the means (Table 4-4) or standard deviations (Table 4-6) of the individual airport damage distributions.

Figures 4-26 and 4-27 are comparable to Figures 4-24 and 4-25, respectively, but are for the annual risk profile, call it S , instead. Tables 4-7 through 4-10 show the results numerically, and are comparable to Tables 4-3 through 4-6, respectively.

Figure 4-28 shows the changes in the annual risk profile S as μ changes; note that increasing values of μ produce increasingly conservative annual risk profiles as defined later. Table 4-11 shows numerically some of the results shown graphically in Figure 4-2 ; again, only ratios are given. Table 4-12 shows the ratios of the means and standard deviations of the annual risk profile S to those of the base case when the indicated changes are made in μ .

4.4.4 Investigation of the effect of approximating the airport damage distributions by lognormal distributions

As indicated above, the empirical individual airport damage distributions are better fit by mixtures of Lognormal distributions than by a single Lognormal distribution. We wished to see whether using such better approximations to the airport damage distributions would significantly alter the results of the sensitivity analyses. It was not practical to carry out the process of fitting all 26 empirical distributions by mixtures of Lognormal distributions, so we decided to proceed differently. We assumed, in effect, that there are only 13 airports, all of whose damage distributions can be fit by mixtures of two Lognormal distributions. This yields 26 Lognormal distributions in all, which we took to be the ones utilized for the base case described above in Section 4.4.3. That base case becomes, in terms of the 13 fictitious airports, the most exact representation available, in that each of the

TABLE 4-7

SENSITIVITY OF THE ANNUAL RISK PROFILE S TO CHANGES IN THE MEANS OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the means are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution mean to that of the base case is r and d is a numerical damage value.

	(a) Damage Value d			
P{Damage > d}	r = 0.5	r = 1.5	r = 2.0	r = 3.0
3×10^{-1}	.38	1.75	2.52	3.87
1×10^{-1}	.46	1.54	2.03	2.89
3×10^{-2}	.59	1.35	1.76	2.45
1×10^{-2}	.69	1.22	1.53	1.89
3×10^{-3}	.79	1.08	1.26	1.48
1×10^{-3}	.99	1.06	1.15	1.27

	(b) P{Damage > d}			
Damage Value d	r = 0.5	r = 1.5	r = 2.0	r = 3.0
$.03 \times 10^{6}$	.76	1.10	1.13	1.13
$.1 \times 10^{6}$	.52	1.17	1.26	1.38
$.3 \times 10^{6}$	.41	1.53	1.97	2.56
1×10^{6}	.34	2.11	3.32	5.92
3×10^{6}	.54	1.46	3.34	7.80

TABLE 4-8

SENSITIVITY OF THE ANNUAL RISK DISTRIBUTION TO CHANGES IN THE MEANS OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the means are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution mean to that of the base case is r .

Ratio r	Mean of the Annual Risk Distribution	Standard Deviation of the Annual Risk Distribution
.33	.33	.92
.50	.50	.93
.67	.67	.95
1.5	1.50	1.10
2.0	2.00	1.22
3.0	3.00	1.53

TABLE 4-9

SENSITIVITY OF THE ANNUAL RISK PROFILE S TO CHANGES IN THE STANDARD DEVIATIONS OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the standard deviations are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution standard deviation to that of the base case is r and d is a numerical damage value.

P{Damage > d}	(a) *Damage Value d* r = 0.5	r = 2.0	r = 5.0	r = 10.0
3×10^{-1}	1.02	.71	.47	.34
1×10^{-1}	.89	.95	.78	.62
3×10^{-2}	.81	1.17	1.17	1.10
1×10^{-2}	.72	1.31	1.52	1.71
3×10^{-3}	.62	1.58	2.05	2.58
1×10^{-3}	.54	2.00	4.06	6.00
Damage Value d	**(b) *P{Damage > d}* r = 0.5**	**r = 2.0**	**r = 5.0**	**r = 10.0**
$.03 \times 10^{6}$	1.10	.95	.74	.61
$.1 \times 10^{6}$	1.16	.81	.61	.50
$.3 \times 10^{6}$	1.03	.76	.59	.47
1×10^{6}	.85	1.12	1.00	.85
3×10^{6}	.18	2.00	2.47	2.62

TABLE 4-10

SENSITIVITY OF THE ANNUAL RISK DISTRIBUTION TO CHANGES IN THE STANDARD DEVIATIONS OF THE AIRPORT DAMAGE DISTRIBUTIONS

KEY: All entries in the table are ratios of (1) the values obtained when the standard deviations are changed to (2) the corresponding values for the base case. The ratio of airport damage distribution standard deviation to that of the base case is r .

Ratio r	Mean of the Annual Risk Distribution	Standard Deviation of the Annual Risk Distribution
0.5	1.0	.61
2.0	1.0	1.87
5.0	1.0	4.58
10.0	1.0	9.14

TABLE 4-11

SENSITIVITY OF THE ANNUAL RISK PROFILE S TO CHANGES IN THE EXPECTED NUMBER OF ACCIDENTS IN A YEAR

KEY: All entries in the table are ratios of (1) the values obtained when the expected number of accidents in a year is changed to (2) the corresponding values for the base case. The ratio of the expected number of accidents to that of the base case is r and d is a numerical damage value.

P{Damage > d}	*(a) Damage Value d* r = 0.50	r = 0.67	r = 1.5	r = 2.0
3×10^{-1}	.39	.61	1.51	2.12
1×10^{-1}	.53	.67	1.36	1.83
3×10^{-2}	.59	.73	1.29	1.67
1×10^{-2}	.62	.74	1.23	1.52
3×10^{-3}	.61	.76	1.19	1.48
1×10^{-3}	.68	.81	1.25	1.47
Damage Value d	***(b) P{Damage > d}* r = 0.50**	**r = 0.67**	**r = 1.5**	**r = 2.0**
$.03 \times 10^{6}$	.69	.82	1.11	1.20
$.1 \times 10^{6}$	.54	.70	1.23	1.38
$.3 \times 10^{6}$	.45	.61	1.52	1.97
1×10^{6}	.40	.55	1.69	2.85
3×10^{6}	.28	.49	1.62	2.75

TABLE 4-12

SENSITIVITY OF THE ANNUAL RISK DISTRIBUTION TO CHANGES IN THE EXPECTED NUMBER OF ACCIDENTS IN A YEAR

KEY: All entries in the table are ratios of (1) the values obtained when the expected number of accidents in a year is changed to (2) the corresponding values for the base case. The ratio of the expected number of accidents to that of the base case is r .

Ratio r	Mean of the Annual Risk Distribution	Standard Deviation of the Annual Risk Distribution
0.50	.50	.71
0.67	.67	.82
1.5	1.50	1.22
2.0	2.00	1.41

13 damage distributions is represented by a mixture of two Lognormal distributions.

We then approximated each of the 13 damage distributions by a single Lognormal distribution. To do this we first randomly combined the 26 Lognormal distributions to obtain 13 pairs. For each such pair we found the first two moments of the mixed distribution (with mixing probabilities proportional to the appropriate p_i), and then approximated the mixed distribution by a single Lognormal distribution with the given first two moments. This yielded Lognormal approximations for the damage distributions at the 13 fictitious airports.

Using the 13 Lognormal distributions we exercised the two computer programs discussed above for the base case and also for all of the sensitivity analysis runs previously described for the case of 26 airports.

The results are shown in Figures 4-29 through 4-33, which correspond to Figures 4-24 through 4-28, respectively. Not only do they correspond, but in each of the five cases the two sets of curves are almost carbon copies. That is, with respect to the scenario of 13 airports, approximations of the 13 airport damage distributions by (a) Lognormal distributions and by (b) mixtures of two Lognormal distributions produce almost exactly the same results. This certainly lends credibility to the validity of the sensitivity analysis results obtained for the scenario of 26 airports. Those results will now be discussed.

4.4.5 Discussion of the results of the sensitivity analyses

This discussion is divided into three parts, corresponding to changes in (a) the means of the airport damage distributions, (b) the standard deviations of the airport damage distributions, and (c) the expected number of accidents in a year. It is based on Tables 4-3 through 4-12.

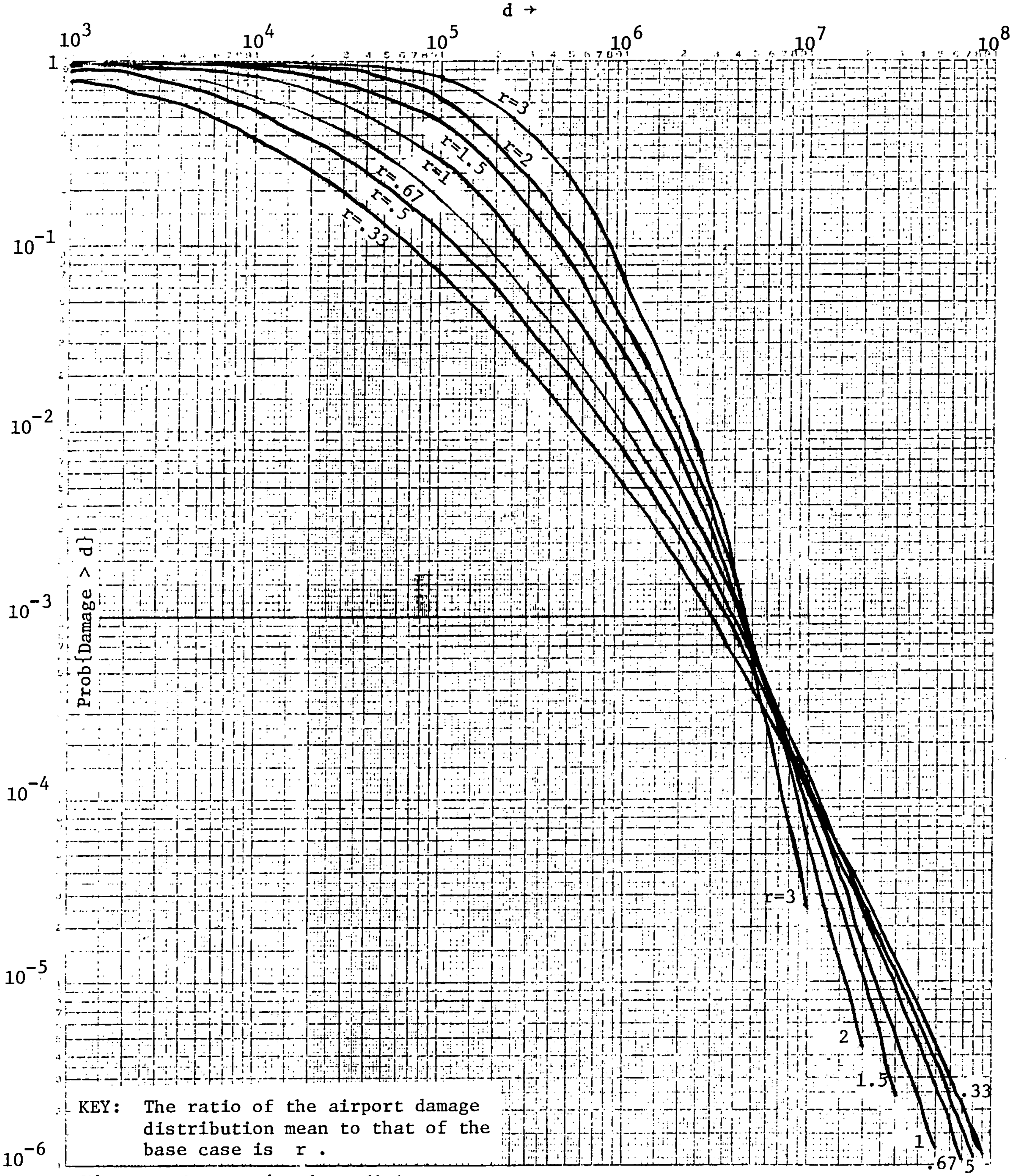


Figure 4-29.--National conditional risk profile R with changes in the means of the airport damage distributions (13 airports).

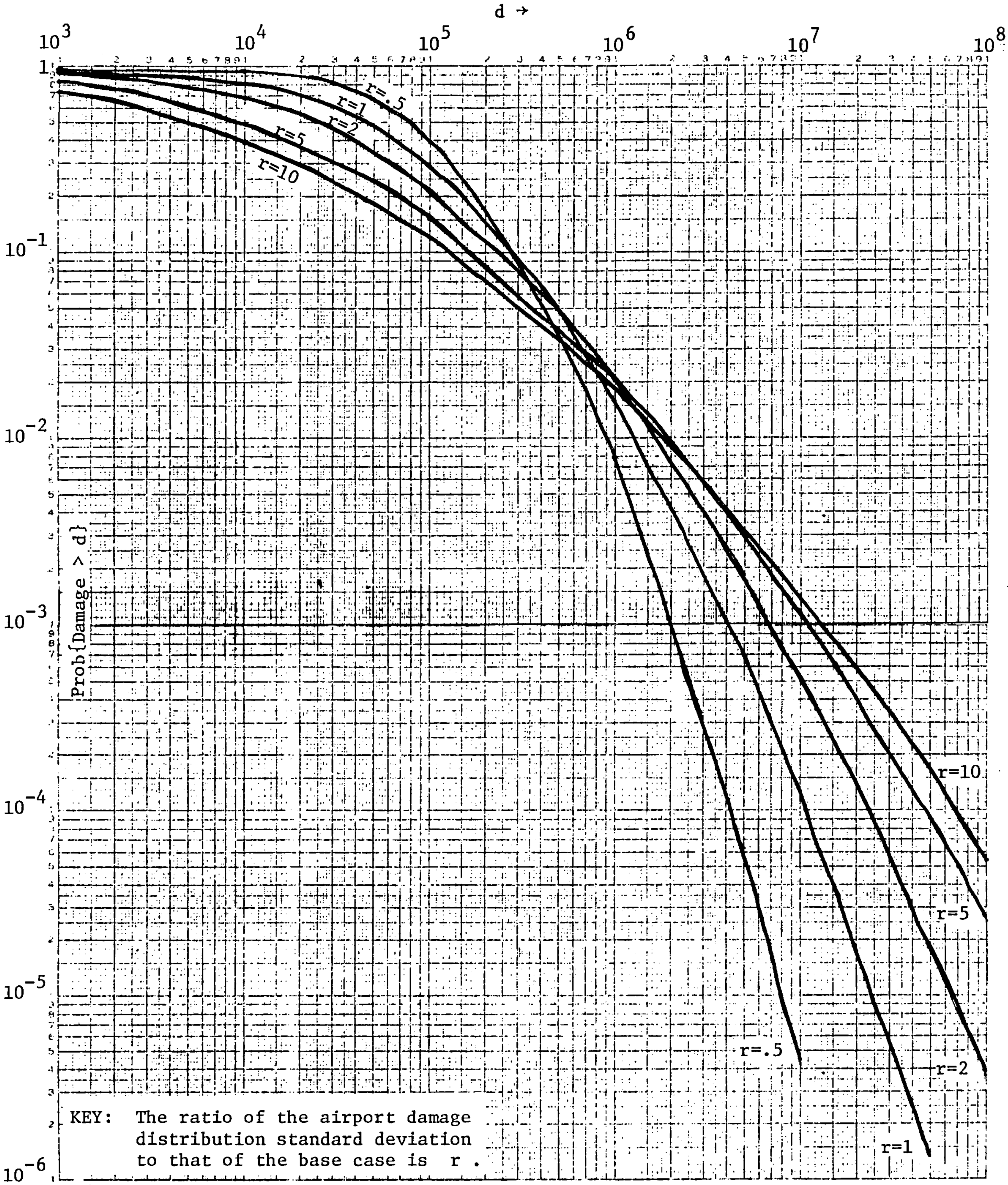


Figure 4-30.--National conditional risk profile R with changes in the standard deviations of the airport damage distributions (13 airports).

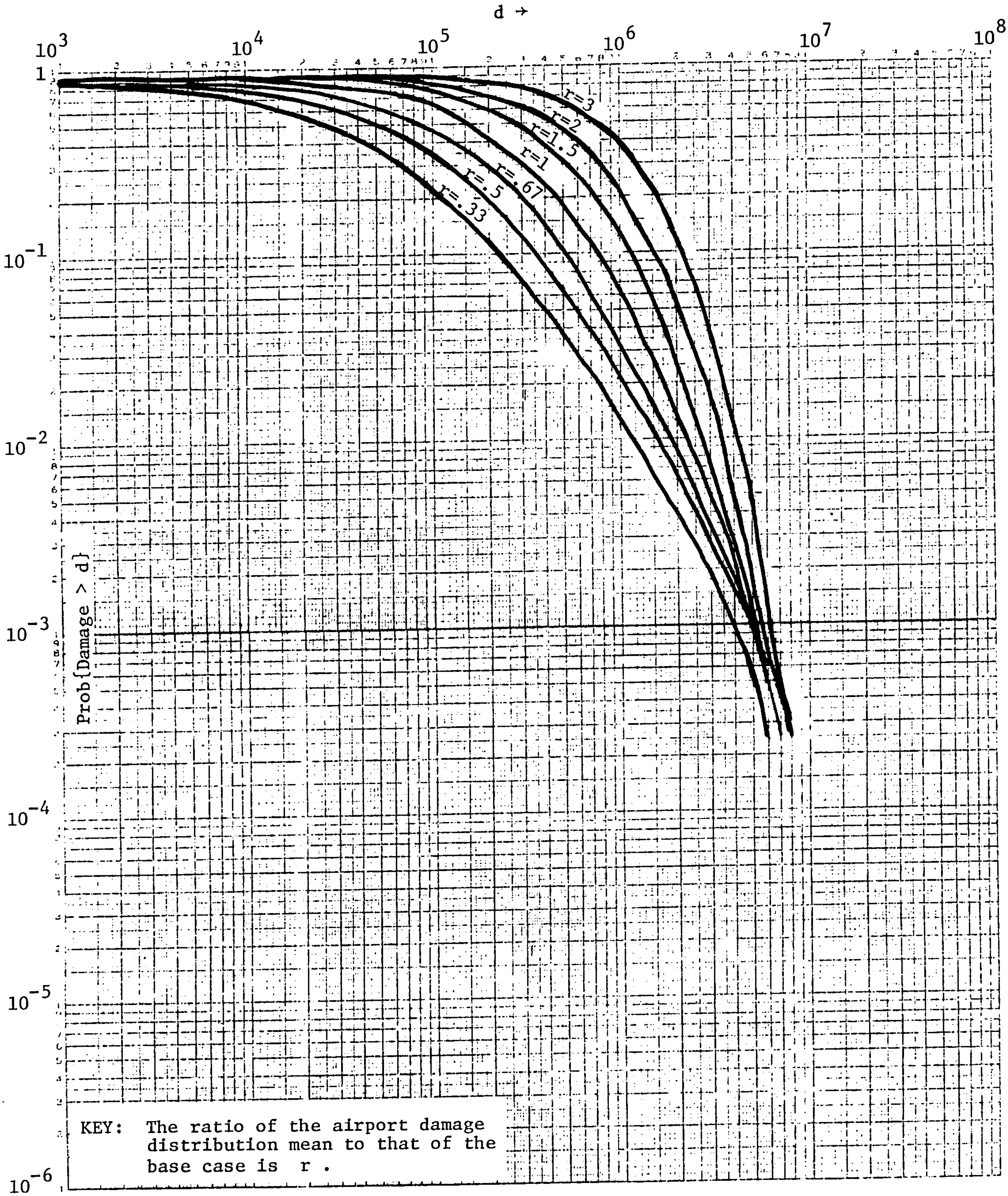


Figure 4-31.--Annual risk profile S with changes in the means of the airport damage distributions (13 airports).

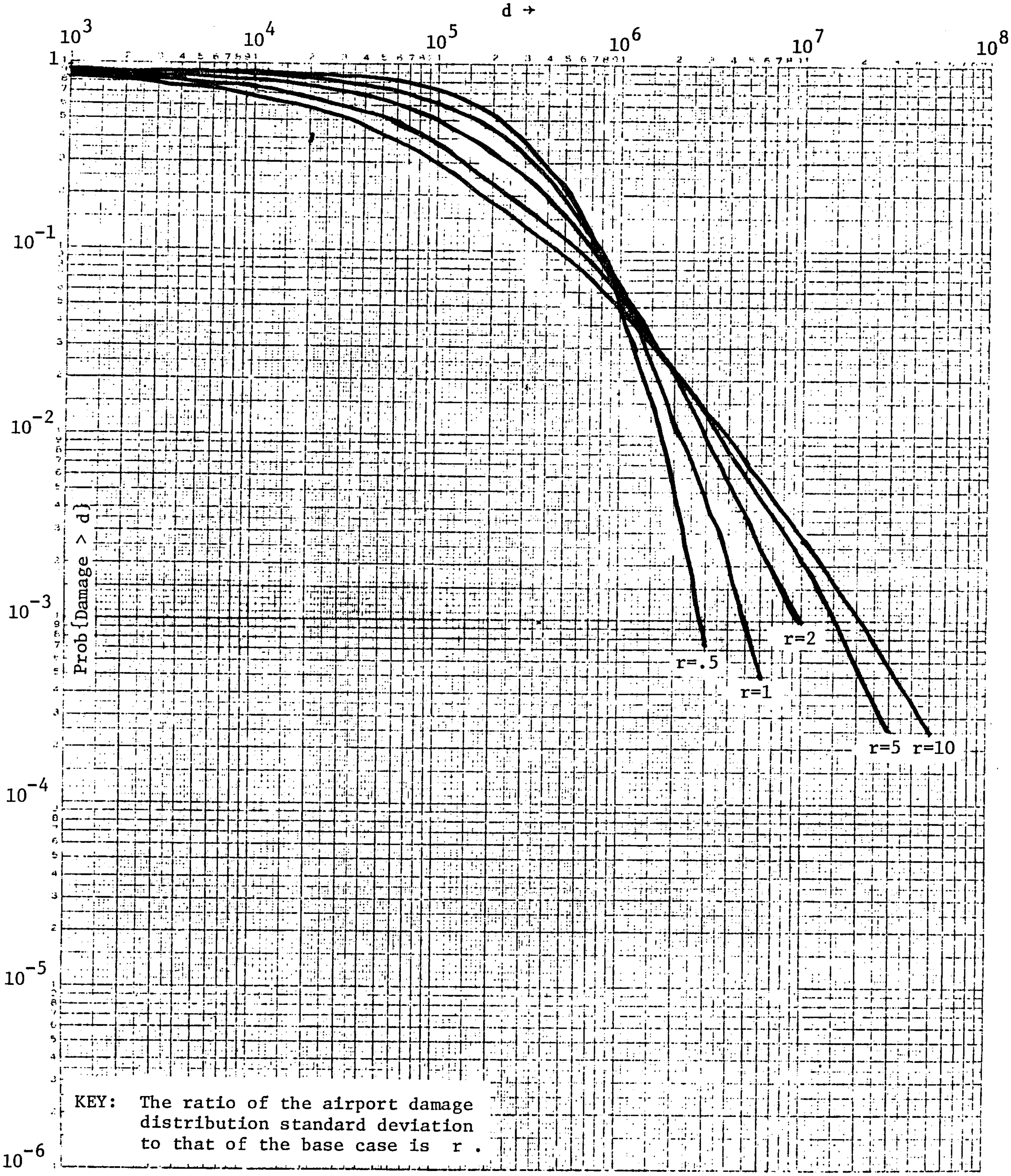


Figure 4-32.--Annual risk profile S with changes in the standard deviations of the airport damage distributions (13 airports).

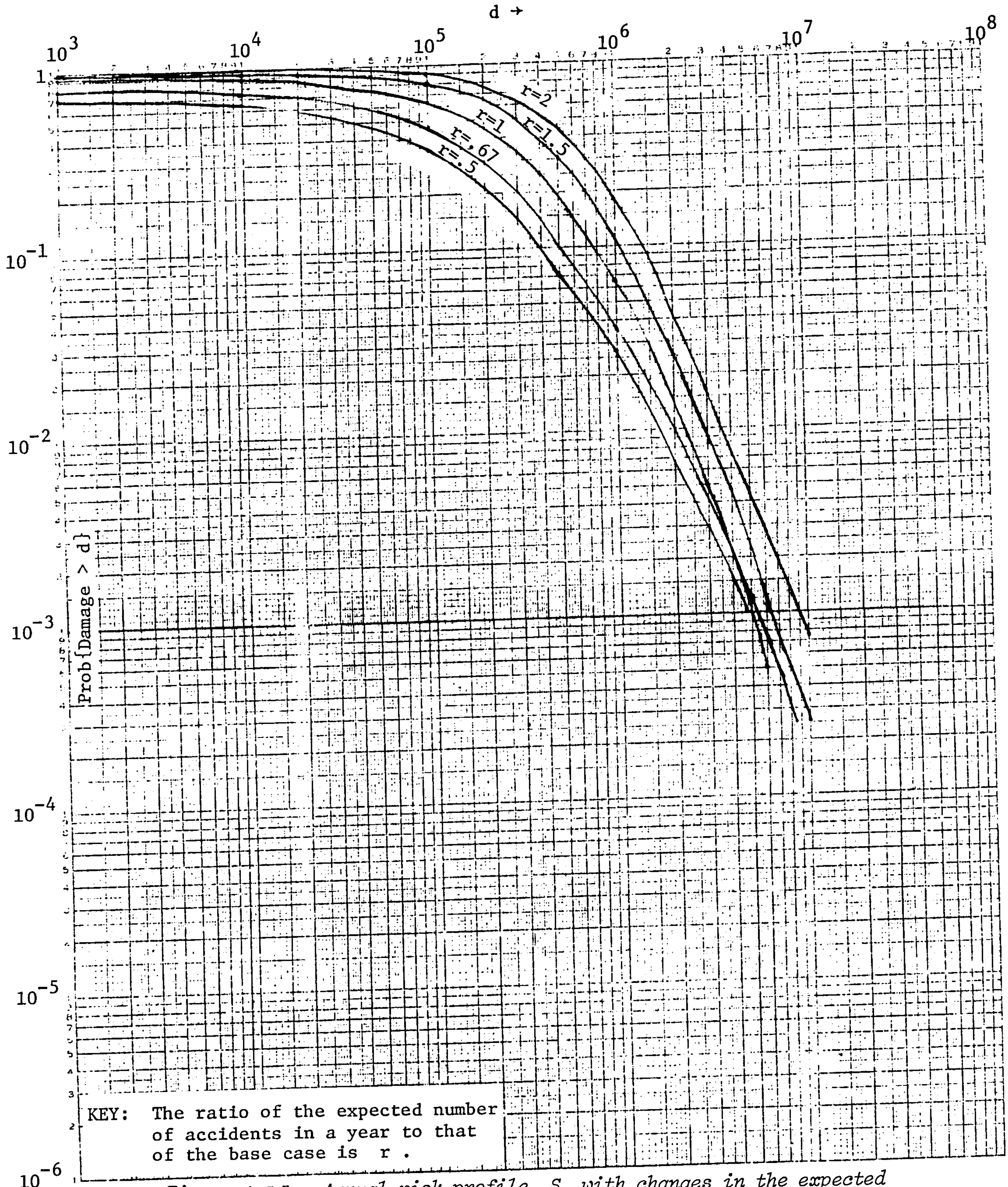


Figure 4-33.--Annual risk profile S with changes in the expected number of accidents in a year (13 airports).

Tables 4-4 and 4-8 show that when the means of all the airport damage distributions are changed in the same proportion, then the means of the national conditional and annual damage distributions are changed in that same proportion but the standard deviations of these two distributions are changed to a much smaller degree. The standard deviation of the annual damage distribution is changed somewhat more than that of the national conditional damage distribution. With respect to the two risk profiles, Tables 4-3 and 4-7 can be summarized by stating that the damage value (for a given exceedence probability) and the exceedence probability (for a given damage value) are both changed in roughly the same proportion as the change in the means of the airport damage dis-istributions.

Tables 4-6 and 4-10 show that when the standard deviations of all the airport damage distributions are changed in the same proportion, then the means of the two damage distributions are unchanged, whereas both their standard deviations are changed in almost (but slightly less than) that same proportion. With regard to the two risk profiles, Tables 4-5 and 4-9 indicate that, in general, the damage value (for a given exceedence probability) and the exceedence probability (for a given damage value) are both changed in somewhat less than the same proportion as the change in the standard deviations of the airport damage distributions.

A change in μ, the expected number of accidents in a year, has no effect, of course, on the national conditonal damage distribution, but Table 4-12 shows that it changes the mean of the annual damage distribution in the same proportion and the standard deviation of this distribution in a smaller proportion (the square root of the previous one). Table 4-11 shows that the effect on the annual risk profile is roughly proportional to the change in μ.

In overall summary, then, our sensitivity analyses show that changes in the means or standard deviations of the airport damage distributions or in the expected number of accidents in a year produce roughly proportional changes in the national conditional and annual risk profiles. Since changes in these risk profiles of less than a factor of 5 are probably not deemed very significant, it seems fair to say that the risk profiles are not overly sensitive to the changes that were investigated.

5. Simulation Model Design and Treatment of Sampling Error

The third major source of error in the risk analysis modeling stems from sampling error in exercising the simulation. Providing the simulation is done properly, statistical techniques are available to treat this type of error.

5.1 Simulation model design

One current simulation model generates, by Monte Carlo simulation, a conditional (given an accident) risk profile for each airport. Denoting the CDF generated for airport i by $F_i^{(1)}(x)$ and letting p_i represent the conditional probability that an accident occurs at airport i, given that it happens at some airport in the U.S., then the conditional (given an accident somewhere in the United States) risk profile, denoted by $1-F^{(1)}(x)$, is obtained from

$$F^{(1)}(x) = p_1 F_1^{(1)}(x) + p_2 F_2^{(1)}(x) + \ldots + p_{26} F_{26}^{(1)}(x) .$$

Even assuming that exact confidence bounds could be found for the $F_i^{(1)}(x)$, and that the p_i's were exact and not estimated, it would still remain a most difficult (if at all possible) task to obtain exact confidence bounds on $F^{(1)}(x)$.

Further, to get the unconditional national annual risk profile, the $F^{(1)}(x)$ are convolved as follows:

$$F(x) = P(0) + P(1)F^{(1)}(x) + P(2)F^{(2)}(x) + \dots ,$$

where $F^{(n)}(x)$ is the n-fold convolution of $F^{(1)}(x)$ with itself and $P(i)$ is the probability of i accidents in a year. Even assuming the $P(i)$ are exact and that confidence bounds could somehow de determined for $F^{(1)}(x)$, the convolution procedure makes it very difficult to obtain confidence bounds on $F(x)$ (especially if exact bounds are desired).

In order to take advantage of the available statistical theory for calculating *valid* confidence bounds, it is necessary to modify the Monte Carlo simulation procedure sketched above. As indicated, two major problems which prevent the use of valid statistical procedures in the above modeling design are (1) probabilistic mixing of individual airport conditional (on one accident) risk profiles to get a national conditional risk profile, and (2) convolving the national conditional risk profile to obtain the unconditional national risk profile. The mixing and convolving of CDFs (or complementary CDFs such as the risk profiles) invalidate the available statistical theory.

To get around these problems, we suggest the modified simulation procedure of Section 4.4.1, which we repeat here:

Step 1: Generate by Monte Carlo methods the number of accidents in a year.

Step 2: Determine by Monte Carlo methods at which airport each of the accidents generated in Step 1 occurs.

Step 3: For each accident, determine the cost.

Step 4: Total the cost of all the accidents for the year.

Step 5: Repeat the above four steps n times (yielding a sample size of n years).

Step 6A: Compute the empirical national annual risk profile directly for the sample of n years from the Step 4 values.

Step 6B: Compute the empirical national conditional (given one accident) risk profile directly for the sample of n years' worth of accidents from the Step 3 values.

In this way, the statistical theory discussed in the following sections can be applied to both the national conditional risk profile and the unconditional national annual risk profile. This may result in larger sample sizes than presently being used, but the advantage is that the sample sizes required for the amount of confidence and precision desired can be computed in advance. The following sections illustrate these computational procedures.

5.2 Pointwise confidence bounds

The following methodology allows for confidence bound statements at a single point only.

5.2.1 Binomial bounds on the risk for a single value

If it is desired to obtain bounds on the risk for a particular value x_0 , then the binomial distribution can be used to obtain well-accepted approximate bounds. Assuming the number of independent simulation runs, n , is large enough for the normal approximation to the binomial ($[nF(x_0)]$ and $n[1-F(x_0)]$ should be $\geqq 5$), approximate $100(1-\alpha)\%$ confidence bound are

$$1-\hat{F}_n(x_0) \pm \frac{z_{\alpha/2}}{\sqrt{n}} \sqrt{\hat{F}_n(x_0)(1-\hat{F}_n(x_0))}.$$

Note that the band width of the bound gets smaller for large x_0 , which is the property we desire; but this confidence statement is good at a single point x_0 only, and not for all x_0 simultaneously.

If the normal approximation to the binomial is not adequate (which is probably the case for the sample sizes anticipated and the small tail probability of interest) then either the Poisson approximation to the binomial or the exact binomial itself must be used. For example, suppose r values greater than x_0 are observed in the sample of size n. Then, we desire to find the largest value of p for a binomial distribution such that $P(X \leqq r \mid n,p) \geq \alpha$, where X is a binomial random variable with parameters n and p, and $1-\alpha$ is the confidence level desired. Denoting this value by $\hat{p}$, a one sided $100(1-\alpha)\%$ confidence interval estimate of $R(x_0)$ is $(0,\hat{p})$. As an illustration, let us assume that in n simulation runs no values greater than x_0 are observed. Then $r=0$ and

$$P(X \leqq 0 \mid n,p) = P(X=0 \mid n,p) = (1-\hat{p})^n = \alpha ;$$

hence,

$$\hat{p} = 1-\alpha^{1/n} .$$

For values of $r>0$ numerical search procedures would be necessary to find $\hat{p}$.

5.2.2 Nonparametric tolerance limits

The prediction approach using tolerance limits as described here is also distribution-free. Let $X_n^{(1)}, X_n^{(2)}, \ldots, X_n^{(n)}$ be the order statistics from a sample of n observations from the distribution with CDF $F(x)$. The problem is to predict the $(n+1)st$ observation, X_{n+1}, which occurs in the future. Intervals of the form

$$X_n^{(r)} < X_{n+1} < X_n^{(s)}$$

are used [see Aitchison and Dunsmore (1975)]. There are two measures of precision: mean coverage and guaranteed coverage (tolerance limit) intervals.

5.2.2.1 Mean coverage: Considering that

$$P\left\{X_n^{(r)} < X_{n+1} < X_n^{(s)}\right\} = \frac{s-r}{n+1} ,$$

is follows that "on the average" (hence the name mean coverage) the interval $(X_n^{(r)}, X_n^{(s)})$ will cover the next observation with a proportion $(s-r)/(n+1)$ of the instances when the procedure is repeated. Note that $P(X_{n+1} > X_n^{(n)}) = 1/(n+1)$. Care must be taken in applying this procedure; for example, one possible misinterpretation would be to look at the data, note that $X_{300}^{(300)} = \$9M$ and then conclude that $P(X_{n+1} > \$9M) = 1/301$. The problem is that

$$P\left(X_{n+1}<m \mid X_n^{(n)}=m\right) \neq P\left(X_{n+1} < X_n^{(n)}\right) .$$

The left-hand side equals $P(x_{n+1}<m) = F(m)$ and is not distribution-free.

5.2.2.2 Guaranteed coverage (tolerance limits): In this case two values, α and γ , are specified, where γ is the probability of coverage and $1-\alpha$ is the guarantee or confidence. The desired interval satisfies

$$P\left\{F(X_n^{(s)}) - F(X_n^{(r)}) = \gamma\right\} = 1-\alpha .$$

Thus we are $100(1-\alpha)\%$ confident that $100\gamma\%$ of the population will fall in $(X^{(r)}, X^{(s)})$. The probability

$$P\left\{F(X_n^{(s)}) - F(X_n^{(r)}) = \gamma\right\}$$

is distribution-free and can be expressed in terms of an incomplete beta function. See Aitchison and Dunsmore (1975), David (1970), and Walpole and Myers (1978).

5.2.3 An upper binomial bound for $R(x_0)$, for a fixed x_0

Using a normal approximation to the binomial, the half-width of such a sonfidence region will be

$$z_\alpha \sqrt{\frac{\hat{R}(x_0)[1-\hat{R}(x_0)]}{n}}$$

to achieve $100(1-\alpha)\%$ confidence. In the tails this will be approximately $z_\alpha \sqrt{\hat{R}(x_0)/n}$. If we are looking at x_0 which corresponds to a tail probability of 10^{-d} and wish to be $100(1-\alpha)\%$ confident that our estimation error is also less than 10^{-d} , then we must approximately satisfy

$$z_\alpha \sqrt{10^{-d}/n} = 10^{-d} .$$

For example, if we want to be 99% confident when the tail probability is of the order of 10^{-4} , then n satisfies $z_{.01}\sqrt{10^{-4}/n} = 10^{-4}$ or $2.326/\sqrt{n} = 10^{-2}$, which implies $\sqrt{n} = 232.6$; and hence $n = 54{,}103 \cong 5.4\times10^4$. In general, to make $100(1-\alpha)\%$ confidence statements about $R(x_0)$ when it is of the order of 10^{-d} , one needs $n = z_\alpha^2 10^d$ observations, for fixed x_0 . If it turns out that $nR(x_0)$ is likely to be less than five--so that the normal approximation might not be valid--then the exact binomial distribution would have to be used, necessitating a numerical search procedure to find n .

5.2.4 Mean coverage prediction intervals

If a sample of size n is drawn and

$$X_n^{(1)}, X_n^{(2)}, \ldots, X_n^{(n)}$$

are the order statistics, then the prediction interval $(X_n^{(r)}, X_n^{(s)})$ has mean coverage of the $(n+1)st$ observation equal to $(s-r)/(n+1)$. An interval of the form $(0, X_n^{(n)})$ has mean coverage equal to $n/(n+1)$. Thus, in order to get mean coverage of $1 - 10^{-d}$, approximately 10^d observations are required. Hence for a probability of 10^{-4} for the next observation exceeding the largest of the n samples, n must be 10^4.

The interpretation of mean coverage is that if this procedure is used many times the resulting intervals (one for each repetition of the procedure) will cover the $(n+1)st$ observation a proportion of times equal to the mean coverage or, equivalently, will not cover it a proportion equal to one minus the mean coverage.

5.2.5 Guaranteed coverage prediction (tolerance) intervals

We are interested in distribution-free prediction intervals of the form $(0, X_n^{(s)})$ which satisfy

$$P\left\{F(X_n^{(s)}) \geqq \gamma\right\} = 1-\alpha$$

for a specified coverage value γ and guarantee probability $1-\alpha$ (this is a one-sided version of the type discussed in Section 5.2.2.2. Here $F(X_n^{(s)})$ has a beta distribution with parameters $s-1$ and $n-s$ with density

$$f_{X_n^{(s)}}(u) = \frac{n!}{(s-1)!(n-s)!} u^{s-1} (1-u)^{n-s} .$$

It is asymptotically normal with mean $s/(n+1)$ and variance $s(n-s+1)/[(n+1)^2(n+2)]$. If $s/n = p$, i.e., $X_n^{(s)}$ is the p*th* sample fractile, then $F(x_n^{(s)})$ is asymptotically normal $(p, [p(1-p)]/n)$.

Consider the case of using $(0, X_n^{(n)})$ as a guaranteed coverage prediction interval; then

$$P\left\{F(X_n^{(n)}) \geq \gamma\right\} = 1 - (1-\gamma)^n = 1-\alpha$$

and

$$n = \frac{\log(\alpha)}{\log(\gamma)} .$$

The sample sizes required to achieve preassigned values of $1-\alpha$ and γ are given in Table 5-1.

TABLE 5-1

SAMPLE SIZES FOR A ONE-SIDED TOLERANCE LIMIT

$1-\alpha$	γ .99	.995	.999	.9995	.9999
.95	299	598	2994	5990	29956
.99	458	919	4603	9208	46049
.995	527	1057	5296	10594	52980
.999	687	1378	6904	13812	69074

Thus, using this procedure, if we wish to deal with tail probabilities of 10^{-4} (or coverage of .9999) with 99% guarantee, we need 46,000 observations. Tables similar to Table 5-1 may be found in Walpole and Myers (1978), page 550.

A misuse of the above procedure that is tempting but not correct is the following. Suppose a particular break-off point such as \$10M annual cost is of interest. Take a sample of size n and note the smallest order statistic that is greater than 10M, calling it $X_n^{(s)}$. Using this, compute $P\{F(X_n^{(s)}) \geqq \gamma_0\}$ for some desired γ_0, and obtain a guarantee probability $1-\alpha_0$. Then, make the statement that (0,\$10M) has a $100(1-\alpha)\%$ guarantee of coverage of $100\gamma_0\%$. The above is not justified; it is called *data snooping*.

5.3 Simultaneous confidence bounds

The following methodology allows for simultaneous confidence bound statements over the entire risk profile.

5.3.1 Kolmogorov-Smirnov (K-S) type confidence bounds

The K-S statistic gives the maximum deviation between an empirical and a true CDF and is denoted by D_n, so that

$$D_n = \sup_x |\hat{F}_n(x) - F(x)| .$$

Durbin (1973), Hoel, Port, and Stone (1971); Dixon and Massey (1969); and Breiman (1973) include discussions of this statistic, both for testing H_0: $F = F_0$ and for constructing simulataneous confidence bounds on the unknown F. Dixon and Massey (1969) refer to the K-S statistic as a *d-statistic*. Lilliefors (1969) treats the case of testing H_0: $F =$ exponential using a modified K-S statistic where the mean of the hypothsized exponential is estimated from the sample.

Confidence bounds using D_n can easily be obtained by noting that

$$\Pr\left\{|\hat{F}_n(x)-F(x)| < d_{\alpha/2}(n)\right\} = 1-\alpha \quad , \tag{5.1}$$

where $d_{\alpha/2}(n)$ is tabulated and depends on the level of confidence desired $(1-\alpha)$ and the sample size (n). Rewriting 5.1 we have

$$\Pr\left\{-d_{\alpha/2}(n) < \hat{F}_n(x)-F(x) < d_{\alpha/2}(n)\right\} = 1-\alpha \quad ,$$

or

$$\Pr\left\{1-\hat{F}_n(x)-d_{\alpha/2}(n) < 1-F(x) < 1-\hat{F}_n(x)+d_{\alpha/2}(n)\right\} = 1-\alpha \quad ,$$

yielding as the lower and upper confidence bound curves $1-\hat{F}_n(x) \pm d_{\alpha/2}(n)$. Denoting the risk profile $1-\hat{F}_n(x)$ by $\hat{R}_n(x)$, we have confidence bounds of $\hat{R}_n(x) \pm d_{\alpha/2}(n)$.

One problem is that the upper confidence bound approaches a constant value for small values of $\hat{R}_n(x)$ since $d_{\alpha/2}(n)$ is a fixed constant added to the empirical value $\hat{R}_n(x)$. A plot of the K-S bound would look like that shown in Figure 5-1.

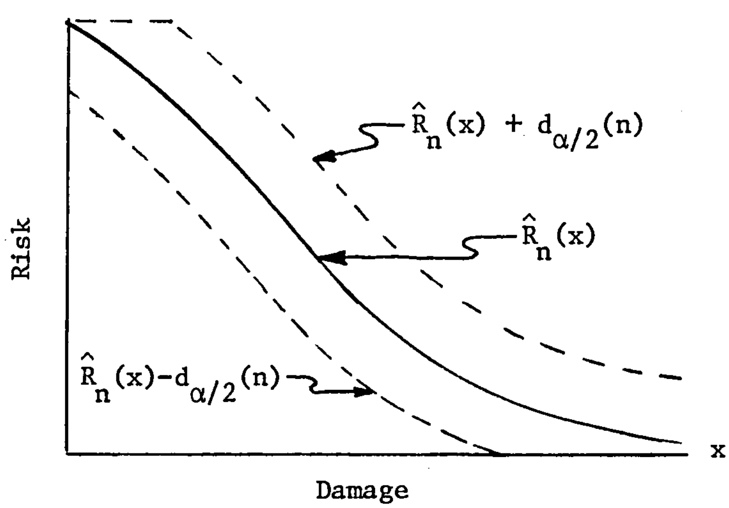


Figure 5-1.--K-S bounds.

It should be pointed out that the type of statement we can make from such bounds is the following:

"With confidence $100(1-\alpha)\%$, the probability of damage exceeding x_0 lies between $[1-\hat{F}_n(x_0)] \pm d_{\alpha/2}(n)$ for *all* values x_0."

This is an extremely powerful statistical statement.

5.3.2 Upper K-S bounds for the risk profile $R(x)$.

We would generally be interested in only one-sided confidence bounds, namely, the upper bound on the risk profile. Since sample sizes are large, the asymptotic theory can be used--in which case the one-sided critical values when using the K-S statistic are as shown in Table 5-2. Thus, for example, $P\{R(x) - \hat{R}_n(x) \leq 1.52/\sqrt{n}\} \cong .99$, and to deal with tail probabilities of the order of 10^{-4} with 99% confidence requires a sample satisfying

$$\frac{1.52}{\sqrt{n}} = 10^{-4} \quad \text{or} \quad n = 2.3 \times 10^8 .$$

TABLE 5-2

SAMPLE SIZES FOR UPPER K-S CONFIDENCE BOUNDS

Confidence Level	Half-Width
.95	$1.22 / \sqrt{n}$
.99	$1.52 / \sqrt{n}$
.995	$1.63 / \sqrt{n}$
.999	$1.86 / \sqrt{n}$
.9995	$1.95 / \sqrt{n}$
.9999	$2.15 / \sqrt{n}$

Some other sample sizes required to obtain various confidence levels in probabilities of certain orders are given in Table 5-3.

TABLE 5-3

SAMPLE SIZES FOR VARIOUS TAIL PROBABILITIES AND CONFIDENCES

Confidence	Tail Probability of Interest 10^{-2}	10^{-3}	10^{-4}	10^{-5}
.95	1.5×10^{4}	1.5×10^{6}	1.5×10^{8}	1.5×10^{10}
.99	2.3×10^{4}	2.3×10^{6}	2.3×10^{8}	2.3×10^{10}
.995	2.7×10^{4}	2.7×10^{6}	2.7×10^{8}	2.7×10^{10}
.999	3.5×10^{4}	3.5×10^{6}	3.5×10^{8}	3.5×10^{10}
.9995	3.8×10^{4}	3.8×10^{6}	3.8×10^{8}	3.8×10^{10}
.9999	4.6×10^{4}	4.6×10^{6}	4.6×10^{8}	4.6×10^{10}

Roughly speaking, to be able to make confidence statements about probabilities of the order of 10^{-d}, the width of the band must also be of that order, 10^{-d}. To achieve a one-sided K-S bound of half-width 10^{-d} requires approximately 10^{2d} observations.

5.3.3 Modified Kolmogorov-Smirnov confidence regions

One problem with K-S bounds is that for very small values of risk, the term $d_\alpha(n)$ overwhelms $\hat{R}_n(x)$. It seems desirable to obtain confidence bounds such that $d_\alpha(n)$ decreases with increasing x so that one obtains a picture such as that shown in Figure 5-2. There are some generalizations of the K-S statistic which might help in this egard. The Anderson-Darling (A-D) statistic can have the property of the bounds coming "in" at large (and small) values of x but these bounds are only approximate. Discussion of the A-D statistic may be found in Anderson and Darling (1952), and Durbin (1973). Another generalization of K-S, called the generalized D^+ statistic, has the bounds partially coming in at the ends--a sort of situation between K-S and the approximate A-D [see, for example, Dempster (1959) and Dwass (1959)].

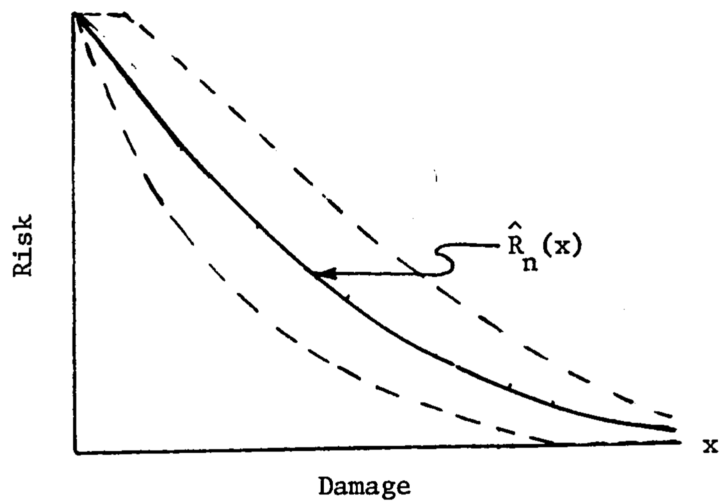


Figure 5-2.--Desired bounds.

In this section we investigate a method of constructing a simultaneous confidence region which is narrower in the right tail. This enables one to be more confident in the right--tail probabilities at the expense of less confidence in the central and left-tail probabilities.

Let $R(\cdot)$ be the cumulative risk function: $R(x) = P\{\text{risk exceeds } x\}$. Let $F(x) = 1 - R(x)$ be the CDF of the risk; R is estimated by the empirical risk function $\hat{R}_n$, based on a simulated sample of n accidents (or n years).

It is desired to estimate R with a confidence region based upon $\hat{R}_n$, or equivalently F with a confidence region based on $\hat{F}_n$. Because we are primarily interested in the right tail, we would like a region which is narrower in the tail. A very desirable region would be

$$R < \hat{R}_n + \beta \hat{R}_n \ , \tag{5.2}$$

where $\beta > 0$. If $\beta = 9$, we are one order of magnitude wide. Unfortunately it is impossible to get a region of the form (5.2). Suppose $X_n^{(n)}$ is the largest of the n simulated values, then $\hat{R}_n(x) = 0$ for $x > X_n^{(n)}$; but $P\{R(X_n^{(n)} = 0\} < 1$ and if R is continuous $P\{R(X_n^{(n)} = 0\} = 0$. Thus $P\{R(x) < \hat{R}_n(x) + \beta \hat{R}_n(x)$ for all $x\} = 0$.

However, it is possible to construct a confidence region of the form

$$R < \hat{R}_n + \beta \hat{R}_n + \alpha \ , \tag{5.3}$$

where typically we might choose $= 10^{-2}, 10^{-3}$, or 10^{-4} and $1+\beta = \sqrt{10}, 10, \sqrt{100}$; i.e., ½, 1, or 1½ orders of magnitude. The equivalent region for $F = 1 - R$ would be

$$\left\{1-F < (1-\hat{F}_n)\ (1+\beta) + \alpha\right\} = \left\{\hat{F}_n < \frac{1}{1+\beta} F + \frac{\alpha+\beta}{1+\beta}\right\} .$$

Thus

$$P\left\{R(x) < \hat{R}_n(x) + \beta R_n(x) + \alpha, \text{ for all } x\right\} = P\left\{\hat{F}_n(x) < \frac{1}{1+\beta} F(x) + \frac{\alpha+\beta}{1+\beta}\right\} . \tag{5.4}$$

By the standard distribution-free argument, (5.4) is independent of F , so for computation we can let $F(x) = x$, the uniform [0,1] distribution. Let $\hat{U}_n(\cdot)$ be the empirical distribution function of uniform order statistics, then

$$\begin{aligned} P\left\{R < \hat{R}_n + \beta \hat{R}_n + \alpha\right\} &= P\left\{\hat{U}_n(x) \quad \frac{1}{1+\beta} x + \frac{\alpha+\beta}{1+\beta}, \text{ for all } x\right\} \\ &= P\left\{\hat{U}_n(x) < bx + a, \text{ for all } x\right\} \\ &= p_n(a,b) \ , \end{aligned} \tag{5.5}$$

where $b = 1/(1+\beta)$ and $a = (\alpha+\beta)/(1+\beta)$. The probability (5.5) can be calculated. Explicit formulae for it are given in Durbin (1973).

It is possible to find the sample size n which achieves

$$1-\gamma = p_n\left(\frac{\alpha+\beta}{1+\beta}, \frac{1}{1+\beta}\right)$$

for various combinations of α, β, γ, e.g., $\gamma=.01, .05$; $1+\beta = \sqrt{10}$, 10, $\sqrt{100}$; $\alpha=10^{-2}, 10^{-3}, 10^{-4}$. This enables one to take a large enough simulated sample to achieve any desired precision in a simultaneous confidence region.

To illustrate the usefulness of the above modified K-S region, we consider an example. Figure 5-3 shows a typical simulated risk profile. If this is based on a sample of 500 observations, then the usual K-S 95% confidence region is

$$R < \hat{R} + \frac{1.22}{\sqrt{500}} = \hat{R} + .05456 .$$

This region is shown in Figure 5-3. The modified region $R < 2\hat{R} + d$ can be computed using a formula for $p_n(a,b)$ from Durbin (1973):

$$p_n(a,b) = \frac{a+b-1}{b^n} \sum_{j=[na+1]}^{n} \binom{n}{j}\left(\frac{j}{n} - a\right)^j \left(a+b-\frac{j}{n}\right)^{n-j-1}$$

Numerical methods can be used to solve the equation

$$p_{500}(a,.5) = .95 .$$

The solution is $a = .50375$. This gives a modified K-S 95% confidence region

$$R < 2\hat{R} + .0075 .$$

This region is also included in Figure 5-3. Note that it gives much greater precision in the right tail. This is achieved at the sacrifice of precision in the right tail.

It should be noted that the tail of an empirical CDF behaves like a Poisson process and consequently the properties of these modified K-S regions can be approximated by properties of Poisson processes; see Dwass (1974) and Pyke (1959).

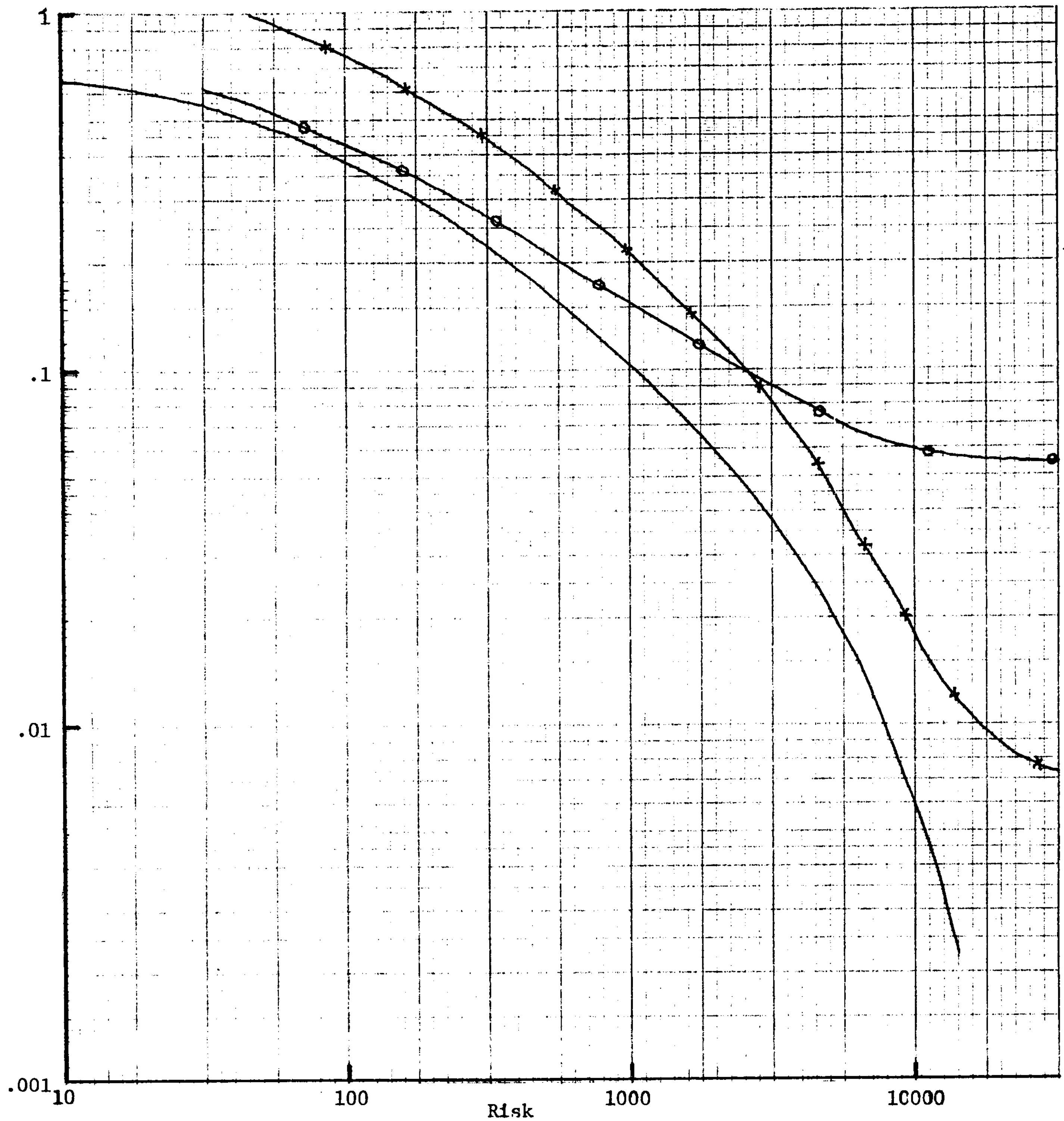


Figure 5-3.--Simulated risk profile and confidence regions: risk profile $\hat{R}_n$ (———) based on 500 simulated observations; Kolmogorov-Smirnov 95% upper bound $\hat{R}_n + .056$ (-o——o——); modified Kolmogorov-Smirnov 95% upper bound $2\hat{R}_n + .0075$ (-x—x——).

5.4 Comparison of procedures

The K-S procedure gives the greatest flexibility because it provides simultaneous bounds, but this is paid for with a huge sample size requirement. The modified K-S procedure reduces this price somewhat.

If a single cut-off value x_0 can be fixed, for which a confidence region on $R(x_0)$ is useful, this can be obtained using the binomial distribution, with considerably fewer observations than the K-S. However, it can be done *only* for a single point.

If a desired mean coverage probability or coverage-guarantee pair can be predetermined, then the prediction interval approach may be used. However, data snooping is not allowed.

A comparison of sample sizes required to achieve a precision of 10^{-4} with 99% confidence is given in Table 5-4. (Note that Mean Coverage statements do not involve confidence.)

6. Conservative Risk Profiles

Engineering design has long used "safety factors" to ensure the adequacy and the safety of physical and electronic systems. This may be looked at as a heuristic, but operationally viable, way to deal with the uncertainty of the environment in which a given system will operate. Not having sufficient knowledge about the environment to design for it precisely, one takes a "conservative" approach in using safety factors to design for extreme conditions that might arise. One can then say that if there is an "error" in the design, then it is surely an "error on the conservative side."

It seems desirable to follow a similar principle of conservatism in generating and reporting results of the GFRAP. Assuming that the GFRAP will ultimately report the risks associated with the use of CF in commerical aircraft to be "small," the results of the risk analysis will be more readily accepted if it can be stated that they overestimate the

TABLE 5-4

REQUIRED SAMPLE SIZES FOR 99% CONFIDENCE STATEMENTS WITH 10^{-4} PRECISION

• Kolmogorov-Smirnov	
$R \leqq \hat{R}_n + 10^{-4}$	, $n = 2.3 \times 10^8$
$R \leqq \hat{R}_n + 10^{-2}$	, $n = 2.3 \times 10^4$
• Modified Kolmogorov-Smirnov (Poisson approximation)	
$R \leqq 2\hat{R}_n + 10^{-4}$	, $n = 4.6 \times 10^4$
• Binomial $(\hat{R}_n(x_0) = 10^{-4})$	
$R(x_0) \leqq \hat{R}_n(x_0) + 10^{-4}$	, $n = 5.4 \times 10^4$
• Mean Coverage	
P{next observation > $X_n^{[L]}$} $= 10^{-4}$	, $n = 10^4$
• Guaranteed Coverage	
P{Damage $\geqq X_n^{[L]}$} $\leqq 10^{-4}$	, $n = 4.6 \times 10^4$

true risks. It should be stated that it is not really possible to present the "true risks" because of modeling limitations, data inadequacies, and limited financial resources.

The results of the risk analysis are summarized in two risk profiles, the national annual risk profile and the national conditional (given one accident) risk profile. We therefore use the following concept of conservatism for risk profiles. If two risk profiles do not cross, it is justifiable to say that the one above and to the right of the other is more conservative (or pessimistic, or represents a greater amount of risk). Figure 4-25 (in Section 4.4.3) illustrates such a case. The risk profile for $r=2$ is above and to the right of that for $r=1$. That the former is more pessimistic folllows from the following two observations: (1) For every damage value d, the profile for $r=2$ shows a higher probability of exceeding d than does the profile for $r=1$. (2) For every probability p, the profile for $r=2$ shows a higher value d such that Prob{damage > d} = p than does the profile for $r=1$.

Thus if we present GFRAP results in the form of risk profiles which are known to be above and to the right of the "true" unknown risk profiles, we present results which are conservative. This section is concerned with methods by which such conservative risk profiles may be obtained. Section 6.1 formalizes the above concept of two risk profiles which do not cross through the concept of stochastic dominance. Section 6.2 shows how this concept can be used operationally in the GFRAP in order to obtain conservative results, and Section 6.3 summarizes and interprets the preceding developments. Several theorems are stated in Section 6.2 without proof in order to conserve space.

6.1 Stochastic dominance

Stochastic dominance is a concept which has become very important in the area of decision making under uncertainty, but it also has useful application here to risk profiles. The following definitions and properties may be found in Whitemore and Findlay (1978) unless noted otherwise. We adopt the notation that $\overline{F} = 1-F$ for any CDF F.

The basic definition of first degree stochastic dominance is the following: Let F and G be CDF's. Now $F \overset{st}{\geqq} G$ (F "stochastically dominates" G) if and only if $\overline{F}(x) \geqq \overline{G}(x)$ for all real values x. This says that $F \overset{st}{\geqq} G$ if and only if the risk profile corresponding to the CDF F lies above (or at least not below) and to the right (or at least not to the left) of that corresponding to G; i.e., the risk profile corresponding to F is conservative relative to that corresponding to G.

Another notation for indicating first degree stochastic dominance is in terms of random variables having the indicated CDF's. Let X and Y be random variables with respective CDF's F and G. One writes $X \geqq_1 Y$ if and only if $F \overset{st}{\geqq} G$.

A useful result is the folllowing: Let U_1 be the set of nondecreasing functions on the real line. Then $F \overset{st}{\geqq} G$ if and only if $\int u dF \geqq \int u dG$ for all functions u in U_1 for which these integrals exist.

Suppose $\overline{G}$ represents one of the two "true" risk profiles that are desired among the final outputs of the GFRAP risk analysis. Not being able to determine $\overline{G}$, as indicated above, we want to report a risk profile $\overline{F}$ such that $F \overset{st}{\geqq} G$. The general idea is to replace all estimated and projected quantities and probability distributions which enter the simulation model by "conservative" ones. The next section provides details.

6.2 Obtaining conservative risk profiles

We deal first with the national annual risk profile and relate it to the national conditional (given one accident) risk profile. Denote these by $\overline{H}$ and $\overline{F}$, respectively, and define the following random variables:

X_i = the damage done by the *ith* accident that occurs during the year, $i=1,2,\ldots$;

S = the total damage done by all accidents that occur during the year;

Z = the number of accidents that occur during the year.

Then S has CDF H , each X_i has CDF F , the X_i are mutually independent, and

$$S = \sum_{i=0}^{Z} X_i ,$$

where X_0 is a random variable which is identically zero.

Let the CDF of the discrete random variable Z be G , and consider what happens if G is replaced by G^* , where G^* is the CDF of another discrete nonnegative random variable, say Z^* , such that $G^* \overset{st}{\geqq} G$; i.e., $Z^* \geqq_1 Z$. Let

$$S^* = \sum_{i=0}^{Z^*} X_i$$

and let H^* be the CDF of S^* . Theorem 6.1 then states that $H^* \overset{st}{\geqq} H$ so that a conservative risk profile is obtained by replacing the CDF of Z by a CDF which is stochastically greater.

Theorem 6.1: If $G^* \overset{st}{\geqq} G$, then $H^* \overset{st}{\geqq} H$, or equivalently, $S^* \geqq_1 S$.

We have assumed that Z has a Poisson distribution with mean μ . The next theorem states that increasing the value of μ yields a CDF of G^* for Z^* such that $G^* \overset{st}{\geqq} G$.

Theorem 6.2: Suppose Z and Z^* both have Poisson distributions with respective means μ and μ^* . Then $Z^* \geqq_1 Z$, i.e., $G^* \overset{st}{\geqq} G$, if and only if $\mu^* \geqq \mu$.

Theorems 6.1 and 6.2 together indicate that the use of a mean number of accidents per year greater than the true mean will yield a conservative risk profile.

Now suppose that the CDF F corresponding to the national conditional risk profile is replaced by a CDF $F^{\#}$ such that $F^{\#} \overset{st}{\geqq} F$. Clearly this will lead to a conservative risk profile, say $\overline{H}^{\#}$. Formally, let $X_i^{\#}$ (i=1,2,...) be mutually independent random variables with CDF $F^{\#}$. Now define

$$S^{\#} = \sum_{i=0}^{Z} X_i^{\#} ,$$

and let $H^{\#}$ be the CDF of $S^{\#}$.

Theorem 6.3: If $F^{\#} \overset{st}{\geqq} F$, then $H^{\#} \overset{st}{\geqq} H$, or equivalently, $S^{\#} \geqq_1 S$.

For the GFRAP, the national conditional risk profile is the weighted average of the single-accident risk profiles at the N_a different airports. Using the notation of Section 4.4.1, we have

$$F(x) = \sum_{i=1}^{N_a} p_i F_{di}(x) ,$$

or equivalently, $\overline{F} = \sum_i p_i \overline{F}_{di}$. If some of the F_{di} are replaced by CDF's which dominate them it is clear that a conservative national conditional risk profile is obtained. Formally, we have Theorem 6.4.

Theorem 6.4: If $F_{id}^{\#} \overset{st}{\geqq} F_{id}$ for $i=1,\ldots,N_a$, and $F^{\#} \equiv \sum_i p_i F_{id}^{\#}$, then $F^{\#} \overset{st}{\geqq} F$.

In Theorem 6.4 the p_i values are held fixed. If they are allowed to vary, some must decrease while others increase since their sum remains equal to one. It is intuitive that if the p_i's which increase correspond to CDF's which dominate the CDF's corresponding to the p_i's which decrease, then a conservative national conditional risk profile is obtained. This is now formalized. Let $N = \{1,2,\ldots,N_a\}$ and let $I^+, \bar{I}$ be disjoint subsets of N . Define a new set of probability values $p_i^{\#}$, $i \in N$, such that

$$p_i^{\#} > p_i \,, \quad \text{if} \quad i \in I^+ \,,$$

$$p_i^{\#} < p_i \,, \quad \text{if} \quad i \in I^- \,,$$

$$p_i^{\#} = p_i \,, \quad \text{otherwise.}$$

Now let $F^{\#} = \sum_i p_i F_{id}^{\#}$; we obtain:

Theorem 6.5: If $F_{id} \overset{st}{\geqq} F_{jd}$ for every pair (i,j) such that $i \in I^+$ and $j \in I^-$, then $F^{\#} \overset{st}{\geqq} F$.

Now we step back and concentrate on the CDF F_{id} of the damge per accident at airport i . We will show how to obtain a conservative estimate of its corresponding risk profile $\bar{F}_{id}$. We first drop the subscript i , for convenience, and to indicate that the analysis applies at each airport.

Consider the random variable X , defined to be the damage done by an arbitrary accident at a particular airport; it has CDF F_d . This X is the sum of the costs due to the failure of various electronic and electrical components and systems. We number all the components and systems which could be affected by an accident from 1 to M (a large finite number), and define the random variables

$$Y_j = 1 \text{ if component or system } j \text{ fails,}$$

$$\quad = 0 \text{ otherwise}$$

$$C_j = \text{the cost incurred if } Y_j = 1 \,.$$

We assume the C_j to be mutually independent and independent of the Y_j. Now we write the damage X as

$$X = \sum_{j=1}^{M} C_j Y_j .$$

From this expression it is clear that replacing the CDF of any cost C_j by a CDF that dominates it will lead to a risk profile that is conservative relative to $\overline{F}_d$. Formally, let $X^* = \sum_j C_j^* Y_j$. Then we have:

Theorem 6.6: $C_j^* \geqq_1 C_j$ for all $j=1,\ldots,M$ implies $X^* \geqq_1 X$.

We clearly have an analogous result if the Y_j are replaced by Y_j^* such that $Y_j^* \geqq_1 Y_j$. Note that this means we replace the failure probability of component j by a value at least as large. Now let $X^* = \sum_j C_j Y_j^*$.

Theorem 6.7: $Y_j^* \geqq_1 Y_j$ for all $j=1,\ldots,M$ implies $X^* \geqq_1 X$.

Now we will concentrate on conditions which imply $Y_j^* \geqq_1 Y_j$ for all j . According to the exponential failure model the probability distribution of Y_j is specified by $\Pr(Y_j = 1) = 1 - \exp(-W_j)$, where the random variable W_j is defined by

$$W_j = T_j E_j / \overline{E}_j ,$$

and $\overline{E}_j$ is the mean exposure to failure of component j , E_j is the outside exposure applicable to component j , and T_j is the overall transfer coefficient (or transfer function) applicable to component j in its individual environment. Here T_j , E_j , and $\overline{E}_j$ are all treated as (independent) random variables; E_j is indeed a random variable, and

T_j and $\overline{E}_j$ are treated as such because of a lack of sufficiently precise information to specify their values exactly. Note that use of the exponential failure model is itself a conservative assumption for the GFRAP (refer to Section 4.2.3).

The next theorem follows from the result given in Section 6.1.

Theorem 6.8: Specify Y_j^* by $\Pr(Y_j^* = 1) = 1 - \exp(-W_j^*)$. Then $W_j^* \geq_1 W_j$ implies $Y_j^* \geq_1 Y_j$.

Now we concentrate on W_j . Let $W_j^* = T_j^* E_j^* / \overline{E}_j^*$. Since T_j , T_j^* , E_j , E_j^* , $\overline{E}_j$, and $\overline{E}_j^*$ are all positive random variables one obtains:

Theorem 6.9: If $T_j^* \geq_1 T_j$, $E_j^* \geq_1 E_j$ and $\overline{E}_j \geq_1 \overline{E}_j^*$, then $W_j^* \geq_1 W_j$ and so $Y_j^* \geq_1 Y_j$.

The last thing we wish to do here is relate the exposure E_j to certain characteristics of the accident that produces this exposure. It is clearly reasonable to assume that, with other factors (such as wind direction and velocity, aircraft type, duration of the burn, etc.) held constant,

$$E_j = KF_r F_i Q_{cf} ;$$

that is, E_j is directly proportional to (K is the proportionality constant) the product of Q_{cf} , the quantity of CF composite on the aircraft, F_i , the fraction of the CF composite carried that is involved in the fire, and F_r , the fraction of the CF burned that is released. Treating Q_{cf} , F_i , and F_r as random variables, and defining E_j^* as the exposure obtained in place of E_j when Q_{cf} , F_i , and F_r are replaced by Q_{cf}^* , F_i^* , and F_r^* , respectively, we obtain the following:

Theorem 6.10: $Q^*_{cf} \geqq_1 Q_{cf}$, $F^*_i \geqq_1 F_i$, and $F^*_r \geqq_1 F_r$ implies $E^*_j \geqq E_j$ for all $j=1,\ldots,M$.

6.3 Summary

We have tried in this section to give some precision to the notion of a "conservative risk analysis" through the concepts of stochastic dominance and conservative risk profiles. In essence, we have shown that replacing the probability distributions used in the GFRAP risk simulation model by probability distributions which stochastically dominate them will lead to a conservative risk analysis. Given that modeling limitations, data inadequacies, etc. prevent us from determining the "true risk profiles," we think it appropriate to report risk profiles that can be stated to be conservative relative to the "true" ones.

In more concrete terms, for example, if the mean number of accidents per year μ is projected to be, say, 2.6, but it is recognized that this projection may be in error, then an appropriately determined higher value, e.g., 3.0, should be used instead in order to produce a conservative national annual risk profile. It would be appropriate to state that it is confidently believed that the true value of μ does not exceed 3.0.

The situation is somewhat different when it is desired to replace a random variable by a specific numerical value. For example, consider the case of F_r , the fraction of the CF burned that is released. It is known that F_r varies from accident to accident and should be treated as a random variable. But suppose that the experimental evidence presently available indicates that values of F_r above 0.01 are extremely improbable. It is then acceptable, and conservative, to use $F_r = 0.01$, stating that it is sufficiently confidently believed that F_r will not exceed 0.01 that such possibility has been deleted from the modeling effort.

REFERENCES

AITCHISON, J. and I. R. DUNSMORE (1975). *Statistical Prediction Analysis.* Cambridge University Press, Cambridge, England.

ANDERSON, T. W. and D. A. DARLING (1952). Asymptotic theory of certain goodness of fit criteria based on stochastic processes. *Ann. Math. Statist. 23*, 193-212.

BARLOW, R. E. (1968). Likelihood ratio tests for restricted families of probability distributions, *Ann. Math. Statist. 39*, 547-560.

BARLOW, R.E., D. J. BARTHOLOMEW, J. M. BREMNER, and H. D. BRUNK (1972). *Statistical Inference Under Order Restrictions.* John Wiley and Sons, Inc., New York.

BREIMAN, L. (1973). *Statistics: With a View Toward Application.* Houghton-Mifflin, Boston, Massachusetts.

DAVID, H. A. (1970). *Order Statistics.* John Wiley and Sons, Inc., New York.

DEMPSTER, A. P. (1959). Generalized D_n^+ statistics. *Ann. Math. Statist. 30*, 593-597.

DIXON, W. J. and F. J. MASSEY (1969). *Introduction to Statistical Analysis* (3rd ed.). McGraw-Hill, Hightstown, New Jersey.

DURBIN, J. (1973). *Distribution Theory for Tests Based on the Sample Distribution Function.* SIAM, Philadelphia.

DWASS, M. (1959). The distribution of a generalized D_n^+ statistic. *Ann. Math. Statist. 30*, 1024-1028.

DWASS, M. (1974). Poisson process and distribution-free statistics, *Adv. Appl. Prob. 6*, 359-375.

HOEL, P. B., S. C. PORT, and C. J. STONE (1971). *Introduction to Statistical Theory.* Houghton-Mifflin, Boston, Massachusetts.

HUSTON, R. J. (ed.) (1979). *Carbon Fiber Risk Analysis.* NASA CP-2074. (Proceedings of an industry-government briefing held at Langley Research Center, Hampton, Virginia, October 31 - November 1, 1978.)

HUSTON, R. J. (ed.) (1980). *Assessment of Carbon Fiber Electrical Effects.* NASA CP-2119. (Proceedings of an industry/government briefing hald at Langley Research Center, Hampton, Virginia, December 4-5, 1979.)

KALELKAR, A. S., J. FIKSEL, P. P. K. RAJ, and D. B. ROSENFIELD (1979). An assessment of the risks presented by the use of carbon fiber composites in commercial aviation, NASA Contractor Report 158989.

LILLIEFORS, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown, *J. Amer. Statist. Assoc. 62,* 399-402.

MANN, N. R., R. E. SCHAFER, and N. D. SINGPURWALLA (1974). *Methods for Statistical Analysis of Reliability and Life Data.* John Wiley and Sons, Inc., New York.

PYKE, R. (1959). The supremum and infimum of the Poisson process, *Ann. Math. Statist. 30,* 568-576.

RAIFFA, H. and R. SCHLAIFER (1961). *Applied Statistical Decision Theory.* Graduate School of Business Administration, Harvard University.

WALPOLE, R. E. and R. H. MYERS (1978). *Probability and Statistics for Engineers and Scientists* (2nd ed.). Macmillan, New York.

WHITEMORE, G. A. and M. C. FINDLAY (eds.) (1978). *Stochastic Dominance.* D. C. Heath and Co., Lexington, Massachusetts.

1. Report No. NASA CR-159318	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle STATISTICAL ASPECTS OF CARBON FIBER RISK ASSESSMENT MODELING		5. Report Date July 1980
		6. Performing Organization Code
7. Author(s) Donald Gross, Douglas R. Miller and Richard M. Soland		8. Performing Organization Report No.
		10. Work Unit No.
9. Performing Organization Name and Address The George Washington University School of Engineering and Applied Science Institute for Management Science and Engineering Washington, DC 20052		11. Contract or Grant No. NSG-1556
		13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546		14. Sponsoring Agency Code

15. Supplementary Notes

Langley Technical Monitor: Wolf Elber
Langley Program Manager: Robert J. Huston

16. Abstract

This study deals with the probabilistic and statistical aspects of the carbon fiber risk assessment modeling of fire accidents involving commercial aircraft. Three major sources of uncertainty in the modeling effort are first identified. These are: (1) imprecise knowledge in establishing the model, (2) parameter estimation, and (3) Monte Carlo sampling error. All three sources of uncertainty are treated in detail, and statistical procedures are utilized and/or developed to control them wherever possible.

17. Key Words (Suggested by Author(s)) Carbon fiber; Risk analysis - Risk profile; Operations research; Statistical analysis; Parameter estimation; Sampling error	18. Distribution Statement Unclassified - Unlimited. Subject Category 24

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	127	A07

N-305

*For sale by the National Technical Information Service, Springfield, Virginia 22161

End of Document