BOEING

(NASA-CR-144245) METHODS FOR COMBINING PAYLOAD PARAMETER VARIATIONS WITH I[illegible]T ENVIRONMENT Draft Interim Report (Boeing Aerospace Co., Seattle, Wash.) 73 p HC $4.50 CSCL 22B G3/18

N76-21267

Unclas
25186

DRAFT INTERIM REPORT

NAS8-31240

31 DECEMB:R, 1975

METHODS FOR COMBINING PAYLOAD PARAMETER VARIATIONS WITH INPUT ENVIRONMENT

D180-19275-1

Prepared by

BO:'NG AEROSPACE COMPANY

Research and Engineering Division

Seattle, Washington 98124

D. H. Merchant, Technical Leader

J. W. Straayer, Program Leader

Prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

George C. Marshall Space Flight Center

Marshall Space Flight Center, Alabama 35812

ABSTRACT

Methods are presented for calculating design limit loads compatible with probabilistic structural design criteria. The approach is based on the concept that the desired "limit load," defined as the largest load occurring in a mission, is a random variable having a specific probability distribution which may be determined from extreme-value theory. The "design limit load," defined as a particular value of this random limit load, is the value conventionally used in structural design. Methods are presented for determining the limit load probability distributions from both time-domain and frequency-domain dynamic load simulations. Numerical demonstrations of the methods are also presented.

KEY WORDS

Monte Carlo method
probabilistic loads
probabilistic structural design criteria
extreme-value theory
structural design loads

ACKNOWLEDGMENTS

The work described in this report was sponsored by the George C. Marshall Space Flight Center under NASA Contract NAS8-31240. The work was performed under the technical direction of Gale R. Ernsberger of the MSFC Systems Dynamics Laboratory. Digital computer programming support was provided by M. W. Ice, Boeing Computer Services, Inc. Analog computer programming support was provided by M. J. Van Dyke, Boeing Computer Services, Inc. The author gratefully acknowledges the contributions made by all these individuals toward the successful completion of this contract.

ILLUSTRATIONS

TABLES

SYMBOLS

α_n, α - extremal intensity function for n observations

$\hat{\alpha}$ - extremal intensity function for n observations of the standardized normal variate.

μ - mean of underlying distribution

m - mean

m_e - median

m_N - mean for N missions

m_o - mode

σ - standard deviation of underlying distribution

s - standard deviation

s_N - standard deviation for N missions

u_n, u - characteristic largest value for n observations

$\hat{u}$ - characteristic largest value for n observations of the standardized normal variate

V - coefficient of variation (s/m)

V_N - coefficient of variation for N missions

$\breve{x}$ - extreme median for n observations of the standardized normal variate

$\tilde{x}$ - extreme mode for n observations of the standardized normal variate

CONTENTS

1.0 INTRODUCTION

The purpose of this report is to describe and numerically demonstrate methods for combining payload parameter variations with the input environment in probabilistic structural design loads analyses. The design loads resulting from these methods are compatible with probabilistic structural design criteria. The approach is based on the concept that the desired "limit load," defined as the largest load occurring in a mission, is a random variable having a specific probability distribution which may be determined from the extreme-value theory of probability. The "design limit load," defined as a particular value of this random limit load, is the value conventionally used in structural design.

The scope of this study was limited in three general areas. First, no attempt was made to include the effects of structural fatigue. The technical theory is concerned only with structural designs corresponding to the single application of an extreme load to an undamaged structure. Second, no attempt was made to define rationale for selecting acceptable probabilities of failure to be used in the structural design criteria. Third, the technical theory is concerned only with the preliminary design/redesign/design verification phases of a project. No attempt was made to address the inverse problem of operational constraints and decisions.

A discussion of a proven general probabilistic structural design approach is presented in Section 2.0 along with some basic results of extreme-value theory which are particularly applicable to structural loads. Section 3.0 presents methods for determining extreme-value limit-load probability distributions from conventional time-domain and frequency-domain dynamic loads analyses. Numerical demonstrations of each of these methods are presented in Section 4.0. Conclusions from the present research and recommended areas for future research are presented in Section 5.0. A comprehensive list of references completes this report.

2.0 THEORETICAL BACKGROUND

The concept of a randomly varying limit load described by a theoretically correct probability distribution and the use of a particular value of this random limit load for structural design purposes are of basic importance in probabilistic structural design criteria. Since the limit load is conventionally defined as the largest load occurring in a mission, the probability theory of extreme values is useful in determining the theoretically correct limit-load probability distribution. Section 2.1 contains some basic results of extreme-value theory which are particularly applicable to structural loads. Since the determination of probabilistic structural loads is meaningful only within the larger context of structural design, Section 2.2 includes details of the application of probabilistic load quantities in a general structural design approach.

2.1 Limit-Load Probability Distributions

The limit load for a structural component is conventionally defined as the largest load occurring during a given mission. The probability that the component load x is the largest value among n independent observations is defined by

$$\Phi_n(x) = [F(x)]^n \tag{1}$$

where F(x) is the underlying cumulative distribution function (CDF) for the load. Thus $\Phi_n(x)$ is, by definition, the cumulative distribution function of the limit load for a mission which has n independent occurrences of applied load. The probability theory of extreme values, as presented by Gumbel (Reference 1), is concerned with describing the limit-load distribution function (Φ_n) for various forms of the underlying distribution (F).

Two parameters frequently used in extreme-value theory are the characteristic largest value and the extremal intensity function. The characteristic largest value (u_n) in a sample of n observations is defined by Gumbel (Reference 1, page 82) in terms of the following equation:

$$F(u_n) = 1 - \frac{1}{n} \tag{2}$$

where $F(u_n)$ is the underlying CDF evaluated at the characteristic largest value. Thus, as indicated by Equation (2), u_n is that value of the random variable which will be equalled or exceeded one time in n observations, on the average. The extremal intensity function (α_n) in a sample of n observations is defined by Gumbel (Reference 1, page 84) as follows:

$$\alpha_n = \frac{f(u_n)}{1-F(u_n)} \tag{3}$$

where $f(u_n)$ is the underlying probability density function (PDF) and $F(u_n)$ is the underlying CDF, both evaluated at the characteristic largest value. The inverse of the extremal intensity function, called Mill's ratio, is tabulated by K. Pearson for the normal distribution (Reference 2, page 11).

The underlying distribution F(x) is said to be of the exponential type if f(x) approaches zero for large $|x|$ at least as fast as the exponential distribution, $f(x) = \lambda e^{-\lambda x}$. For any distribution of the exponential type, Gumbel (Reference 1, page 168) shows that the CDF for large x is approximately equal to

$$F(x) = 1 - \frac{e}{n}^{-\alpha_n(x-u_n)} \tag{4}$$

An asymptotic distribution of extreme largest values can be obtained by substituting Equation (4) into Equation (1) and taking the limit as n becomes infinite

$$\phi_n^{(1)}(x) = \lim_{n\to\infty}\left[1 - \frac{e}{n}^{-\alpha_n(x-u_n)}\right]^n \tag{5}$$

Evaluating this limit by means of the logarithmic series results in the first asymptotic distribution of extreme largest values, subsequently called the extremal type I distribution:

$$\phi^{(1)}(x) = \exp\left(-e^{-\alpha_n(x-u_n)}\right) \tag{6}$$

The corresponding PDF, which is positively skewed, is given by

$$\phi^{(1)}(x) = \alpha_n \exp\left[-\alpha_n(x-u_n) - e^{-\alpha_n(x-u_n)}\right] \tag{7}$$

The most probable value or mode (m_o) of this distribution is equal to the characteristic largest value:

$$m_o = u_n \tag{8}$$

The fifty-percentile value or median (m_e) is given by

$$m_e = u_n - \frac{\ln(-\ln 0.5)}{\alpha} = u_n + \frac{0.36651292}{\alpha} \tag{9}$$

The mean (m) is

$$m = u_n + \frac{C_E}{\alpha} \tag{10}$$

where $C_E = 0.57721566$ is Euler's constant.

The standard deviation (s) is given by

$$s = \frac{\pi}{\sqrt{6}\,\alpha_n} \tag{11}$$

and the coefficient of variation ($V = s/m$) is

$$V = \frac{\pi}{\sqrt{6}(\alpha_n u_n + C_E)} \tag{12}$$

Equations (10) through (12) define parametric values for the extremal type I distribution corresponding to a single mission. Parametric values for the largest load occurring in N missions are as follows:

$$m_N = m + \frac{\sqrt{6}}{\pi} \ln N \cdot s \tag{13}$$

$$s_N = s \tag{14}$$

$$V_N = \frac{V}{(1 + \frac{\sqrt{6}}{\pi} \ln N \cdot V)} \tag{15}$$

These relations are derived in Reference 3 (page 67). Note that the standard deviation (s) and the extremal intensity function (α_n) for the extremal type I distribution are theoretically independent of sample size.

According to Gumbel (Reference 1, page 182) the extremal type I distribution is often satisfactorily represented by the lognormal distribution. The lognormal distribution with coefficient of variation equal to 0.364 is essentially identical to the extremal type I distribution. For coefficients of variation between 0.31 and 0.42, the extremal and lognormal distributions are graphically indistinguishable. An example of the validity of the log-normal approximation to the extremal type I distribution is given in Reference 4. For this analysis, 28 sets of internal load quantities were calculated as the maximum values experienced in each of 100 simulated lunar landings. A Chi-square test of the hypothesis that the loads were lognormally distributed resulted in cumulative probabilities ranging from 5 to 90 percent. The lognormal approximation was therefore considered acceptable since the Chi-square probabilities were less than 90 percent for all 28 internal load quantities. The Chi-square hypothesis is usually accepted for cumulative probabilities as high as 99 percent. The coefficients of variation for these load quantities varied between 0.2 and 0.4.

The extremal type I distribution, defined by Equations (6) through (15), is the theoretically proper distribution for limit loads due to any condition having an exponential-type underlying probability distribution and a sufficiently large number of independent load occurrences. For the exponential distribution, convergence to the asymptotic extremal type I distribution is essentially complete for 100 observations (Reference 1, page 116). For the normal distribution, however, convergence to the asymptotic type I distribution is extremely slow. According to Fisher and Tippett (Reference 5, page 189), close convergence is attained only for sample sizes on the order of 10^{55}. Such large samples correspond to characteristic largest values of the standardized normal variate on the order of 16.

Accurately describing extreme values from an underlying normal distribution is necessary due to the central role of the normal distribution in engineering applications. The theoretical distribution of extreme largest values from variously sized samples of standardized normal variates was tabulated by K. Pearson in Reference 2 (page 162). Plots of these tabulated values on lognormal probability paper indicate that, for certain sample sizes, the theoretical distribution of normal extremes can be adequately approximated by the lognormal probability distribution. In fact, the theoretical distribution plots essentially as a straight line on lognormal probability paper for standardized characteristic largest values ($\hat{u}_n$) of approximately 2.16. This value of $\hat{u}_n$ corresponds to a sample size (n) of approximately 65.

The lognormal PDF may be written for the normal extreme variate (x) as follows:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\delta x} \exp - \frac{1}{2} \left(\frac{\ell nx - \gamma}{\delta} \right)^2 \tag{16}$$

where γ is the mean of ℓnx and

δ is the standard deviation of ℓnx.

The parameters γ and δ used for the lognormal approximation are obtained in terms of the standardized extreme median $(\breve{x})$ and standardized extreme mode $(\tilde{x})$ by means of the following identities:

(1) The mean of $\ell n x$ is the logarithm of the median of x;

(2) The variance of $\ell n x$ is the logarithm of the ratio of median of x to the mode of x (Reference 1, page 18).

Let the underlying normal distribution of interest have mean μ and standard deviation σ. The required lognormal parameters are then given by

$$\gamma = \ell n(\mu + \breve{x}\sigma) \tag{17}$$

$$\delta = \left[\ell n\left(\frac{\mu + \breve{x}\sigma}{\mu + \tilde{x}\sigma}\right)\right]^{1/2} \tag{18}$$

The median $(\breve{x})$ of the standardized normal extreme for n samples is defined by the following equation:

$$[F(\breve{x})]^n = 0.5 \tag{19}$$

Combining Equations (2) and (19) to eliminate n gives the following desired equation for the standardized extreme median $(\breve{x})$ in terms of the standardized characteristic largest value $(\hat{u}_n)$:

$$F(\breve{x}) = \exp[(\ell n\ 0.5)(1 - F(\hat{u}_n))] \tag{20}$$

where F is the normal CDF.

The mode $(\tilde{x})$ of the standardized normal extreme for n samples is defined by Gumbel (Reference 1, page 133) in terms of the following equation.

$$\frac{\tilde{x} \cdot F(\tilde{x})}{f(\tilde{x})} = n-1 \tag{21}$$

where F is the normal CDF, and

f is the normal PDF.

Combining Equations (2) and (21) to eliminate n gives the following desired equation for the standardized extreme mode $(\tilde{x})$ in terms of the standardized characteristic largest value $(\hat{u}_n)$:

$$\frac{\tilde{x}\cdot F(\tilde{x})}{f(\tilde{x})} = \frac{F(\hat{u}_n)}{1-F(\hat{u}_n)} \tag{22}$$

The lognormal approximation to the distribution of normal extremes is defined by Equations (16) through (22). This representation may be considered adequate for values of the standardized characteristic largest value ($\hat{u}_n$) less than 3.

A second approximation to the distribution of normal extremes was proposed by Fisher and Tippett (Reference 4). The proposed CDF is of the form

$$\phi^{(3)}(y) = \exp\left[-\left(\frac{y}{\hat{u}_n}\right)^{-k}\right] \tag{23}$$

$$\text{where} \quad k = \frac{(\hat{u}_n^{\,2} + 1)^2}{(\hat{u}_n^{\,2} - 1)} \tag{24}$$

This general form is denoted by Gumbel (Reference 1, page 298) as the third asymptotic distribution of extreme values or the extremal type III distribution. By inverting Equation (23), approximate percentage points for extremes of the standardized normal variate are obtained as follows in terms of the cumulative probability, p:

$$y = \frac{x - \mu}{\sigma} = \exp\left[\ln \hat{u}_n - \frac{\ln(-\ln p)}{k}\right] \tag{25}$$

A special characteristic of this extremal type III distribution is that it converges for increasing values of the parameter k toward the extremal type I distribution. Thus, in practice, the extremal type III distribution may be used to represent normal extremes for all values of $\hat{u}_n$ greater than 3.

For very large values of $\hat{u}_n$, the extremal type I distribution, which is the theoretical asymptotic distribution of normal extremes, may be used. For the normal standardized variate (y), the type I distribution function is

$$\phi^{(1)}(y) = \exp\left(-e^{-\hat{\alpha}_n(y - \hat{u}_n)}\right) \tag{26}$$

where

$$\hat{\alpha}_n = \hat{u}_n + \hat{u}_n^{-1} - 2\hat{u}_n^{-3} + 10\hat{u}_n^{-5} \tag{27}$$

This expression for the standardized normal extremal intensity function is derived by Gumbel (Reference 1, page 137). The type I extreme value distribution may be used, if desired, to describe normal extremes for standardized characteristic largest values exceeding 8.

2.2 Probabilistic Structural Design Criteria

The extended reliability structural design approach proposed by Ang and Amin (Reference 6) recognizes both the probabilistic nature of limit loads and strength and the analytical uncertainties associated with their evaluation. The uncertainties associated with determining limit loads and strengths can be quantified by a factor of uncertainty (ν) equal to or greater than unity. Thus the event ($S/L \geq \nu$) constitutes a state of structural safety, where S and L represent the strength and limit load associated with a structural component. If S and L are both random variables, then the probability $P[S/L \geq \nu]$ is a proper measure of structural safety. The extended reliability structural design approach is then expressed by the following probabilistic equation for structural safety:

$$P\left[\frac{S}{L} \geq \nu\right] = 1 - P_F \tag{28}$$

where S is the random variable describing the component strength,
L is the random variable describing the component limit load,
ν is the Ang-Amin coefficient of uncertainty for the component, and
P_F is the component probability of failure or acceptable risk.

When the limit-load and strength probability density functions are known, the structural design approach may be expressed in two equivalent forms:

$$1 - P_F = \int_{-\infty}^{\infty} \int_{-\infty}^{y/\nu} f_L(x) \cdot f_S(y) \, dx \, dy \tag{29}$$

$$P_F = \int_{-\infty}^{\infty} \int_{-\infty}^{\nu x} f_S(y) \cdot f_L(x) \, dy \, dx \tag{30}$$

where $f_L(x)$ is the limit-load probability density function (PDF), and
$f_S(y)$ is the strength PDF.

The conventional factor of safety is defined as the ratio of the allowable stress (S_A) to the design limit load (L_D):

$$FS = \frac{S_A}{L_D} \tag{31}$$

where S_A is the value of the random strength corresponding to a specified exceedance orobability (P_A), and

L_D is the value of the random limit load corresponding to a specificed non-exceedance probability (P_D).

The purpose of the factor of safety in the structural design procedure is to locate the strength PDF elative to the given limit-load PDF so that Equation (29) or (30) results in the required component probability of failure. This concept is illustrated in Figure 2.2-1. For most probability distributions, the integral of Equation (29) or (30) must be evaluated numerically and the required factor of safety determined by trial-and-error procedures. However, for certain specific distributions, closed-form evaluations leading to convenient design formulas are possible.

A particularly convenient design factor-of-safety equation occurs when both limit loads and strengths are assumed to follow the lognormal probability law. As discussed in Section 2.1, the lognormal distribution often accurately represents the theoretically proper distribution for limit loads. Moreover, for much existing strength data, the lognormal distribution also is a satisfactory representation, due perhaps to the deletion of low-strength values by quality-control procedures.

The component factor-of-safety expression for lognormal limit loads and strengths is derived in Reference 3 in the following form:

$$FS = \frac{S_A}{L_D} = \nu \cdot \exp - \left[F^{-1}(P_F) \sqrt{\ln[(1 + V_L{}^2)(1 + V_S{}^2)]} + F^{-1}(P_D) \sqrt{\ln(1 + V_L{}^2)} + F^{-1}(P_A) \sqrt{\ln(1 + V_S{}^2)} \right] \tag{32}$$

where ν is the Ang-Amin coefficient of uncertainty,

P_F is the probability of failure or acceptable risk,

P_D is the non-exceedance probability for design limit load (L_D),

P_A is the exceedance probability for allowable stress (S_A),

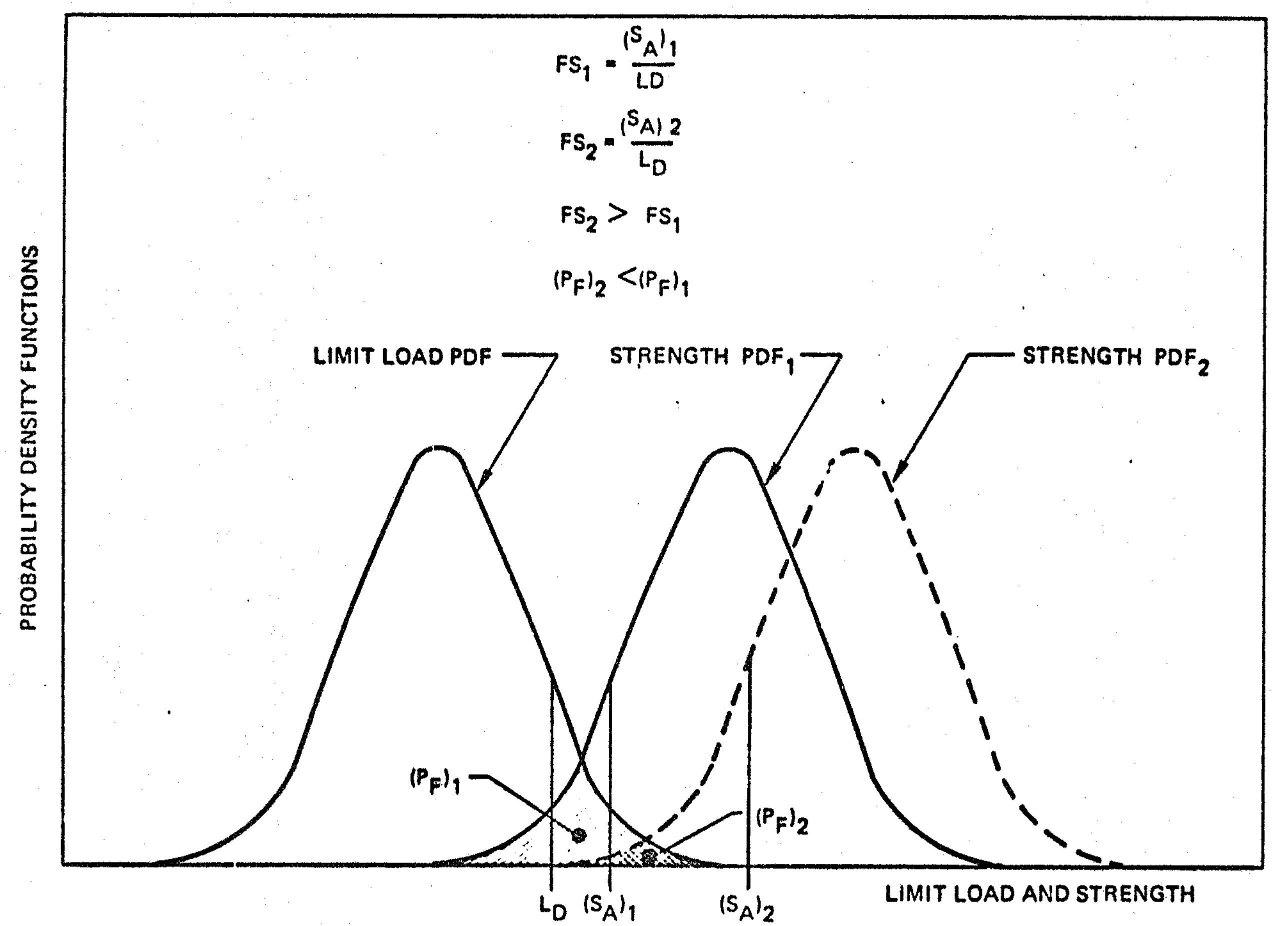

Figure 2.2-1. Graphical Representation of Factor of Safety

V_L and V_S are limit-load and strength coefficients of variation,

$F^{-1}(P)$ is the inverse of the standardized normal cumulative distribution function given by

$$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{F^{-1}(P)} e^{-\frac{1}{2}t^2} dt \qquad (33)$$

The numerical behavior of the lognormal/lognormal factor of safety is shown graphically in Figure 2.2-2. For this plot, the defining probabilities for design limit load and allowable strength are both taken as 99 percent, and the coefficient of uncertainty is taken as unity. The factor of safety is seen to increase monotonically with decreasing probability of failure for given load and strength coefficients of variation.

From Equation (32), the component factor of safety corresponding to a specified probability of failure may be computed. The allowable strength is then determined, from Equation (31), as the product of the factor of safety times the design limit load. Additional details regarding the application of this probabilistic design approach are presented in Reference 3. Procedures for determining the basic limit-load probability distributions from which the specific design limit load is selected are discussed in the following section.

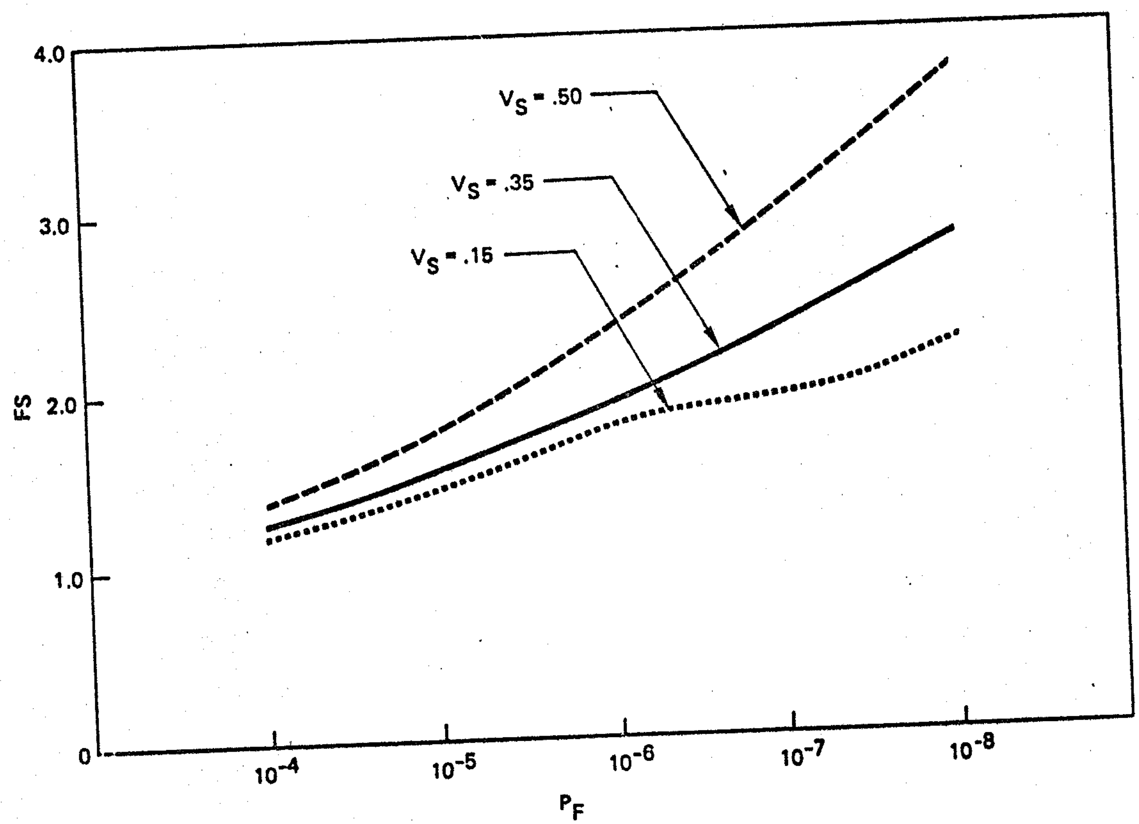

Figure 2.2-2. Factor of Safety (FS) vs Probability of Failure (P_F) *Limit-Load Coefficient of Variation V_L = .3*

3.0 METHODOLOGY DEVELOPMENT

Structural dynamic analyses resulting in design limit loads may be performed either in the time domain or in the frequency domain. The Taylor's series method and the Monte Carlo method are two widely used techniques for determining limit loads from time-domain analyses. The Taylor's series method, described in Section 3.1, is an extension of the parameter variation study often performed to evaluate sensitivity to parametric data uncertainties. The Monte Carlo method, described in Section 3.2, is a simulation of the loading condition using a random combination of vehicle parameters and environments. For each load quantity of interest, the maximum value occurring in each simulated mission is identified and recorded. The maximum load data from a number of simulated missions approximates the desired extreme-value limit-load distribution.

In Section 3.3, a new method is presented for determining the extreme-value limit-load distribution from a frequency-domain analysis. This method determines the probability distribution of the extreme largest load value, for a stationary Gaussian random process, occurring within a given mission length.

3.1 Limit Loads from Taylor's Series Analyses

A detailed discussion of the use of the Taylor's series method to estimate limit-load probability distributions for aerospace launch vehicles is presented by Lovingood (Reference 7). This application involves first analytically simulating the structural loads and responses encountered by a nominal launch vehicle flying through a moderately severe synthetic wind profile. The resulting loads are considered to be the nominal or mean values for the limit load probability distribution. The peak or design limit load values, which are defined as the "3-σ" values having non-exceedance probabilities of 0.9987, are next obtained by computing the variations in load due to 3-σ variations in the significant vehicle parameters, taking the root-sum-square variations of each load quantity, and adding these to the corresponding mean values.

This method is useful for efficiently predicting preliminary and interim structural design loads. However it has the disadvantage of requiring a synthetic wind profile defined such that the mean values of all the limit loads of interest are produced by the analytical simulations. Besides the difficulty of defining this proper synthetic environment, the Taylor's series method is based on three fundamental assumptions which may not be valid for particular applications. These assumptions will be discussed in the brief derivation which follows. A similar derivation in Reference 7 is somewhat more detailed.

The distribution of a nonlinear function of several random variables may be obtained by approximating the desired function as a linear function in the region of interest. The mean and standard deviation of a linear function of several independent random variables are known from elementary probability theory (Reference 8, page 48). If $X_1, X_2, \ldots, X_n$ are independent random variables having means $m_1, m_2, \ldots, m_n$ and variances $s_1^2, s_2^2, \ldots, s_n^2$, respectively, and if $a_1, a_2, \ldots, a_n$ are constants, then a linear random function may be defined as follows:

$$f(X_1, X_2, \cdots, X_n) = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n \quad (34)$$

The mean of f is

$$m_f = a_1 m_1 + a_2 m_2 + \cdots + a_n m_n \quad (35)$$

Thus the mean of a linear combination of random variables is equal to the linear combination of the means. This result is valid even if the X's are dependent.

The variance of f is

$$s_f^2 = a_1^2 s_1^2 + a_2^2 s_2^2 + \cdots + a_n^2 s_n^2 \quad (36)$$

Thus the variance of a linear combination of independent random variables is equal to the sum of the products of variances and squared constants. In addition, if the X's are normally distributed, then f is also normally distributed with mean m_f and variance s_f^2.

A nonlinear function may be expanded in a Taylor's series about any given point as follows (Reference 8, page 62):

$$\begin{aligned} f(X_1, X_2, \cdots, X_n) &= f(m_1, m_2, \cdots, m_n) \\ &+ (X_1 - m_1) \left.\frac{\partial f}{\partial X_1}\right|_{m_1, m_2, \cdots, m_n} + \cdots \\ &+ (X_n - m_n) \left.\frac{\partial f}{\partial X_n}\right|_{m_1, \cdots, m_n} + \text{higher order terms} \quad (37) \end{aligned}$$

If the higher order terms are negligible, the mean of f is, according to Equation (35), approximately equal to:

$$m_f \approx f(m_1, m_2, \cdots, m_n) \tag{38}$$

If, in addition, the X's are independent, the variance of f is, according to Equation (36), approximately equal to

$$s_f^2 \approx s_1^2 \left[\frac{\partial f}{\partial X_1} \bigg|_{m_1, m_2, \cdots, m_n} \right]^2 + \cdots$$

$$+ s_n^2 \left[\frac{\partial f}{\partial X_n} \bigg|_{m_1, m_2, \cdots, m_n} \right]^2 \tag{39}$$

Futhermore, if the X's are normally distributed, f is approximately normally distributed. If the X's are normally distributed and if the function is linear so that Equation (37) contains no higher order terms, then the mean and variance are exactly as given by Equations (38) and (39) and the theoretical distribution of the function is the normal distribution (Reference 9, page 90).

The three assumptions in the use of the Taylor's series method are as follows:

(1) that the higher-order terms in the Taylor expansion are negligible compared with the first-order terms,
(2) that the X's are independent, and
(3) that the X's are normally distributed.

The accuracy of design limit loads determined by the Taylor's series method depends in part upon how well the particular physical simulation is represented by these three assumptions. In Section 4.1, a discussion of the effects of these assumptions is presented along with numerical demonstrations of the method. So long as the potential disadvantages of this method are recognized, it remains an efficient and useful tool for estimating preliminary and interim design limit loads.

3.2 Limit Loads from Monte Carlo Simulations

The Monte Carlo method is a powerful and general tool for predicting structural design loads. The method has been gaining wider acceptance for dynamic load studies of aerospace vehicles (References 4, 10, 11. 12). For this application, the method consists essentially of simulating a random loading phenomenon by combining deterministic and probabilistic variables. The limit-load probability distribution for each load quantity is then the distribution of the largest loads occurring in each simulated mission. For the launch vehicle load simulations described in Reference 10, the deterministic variables included such vehicle parameters as mass and geometry, structural dynamic characteristics, propellant slosh parameters, and control-system parameters. The probabilistic variables for this study were restricted to descriptions of the wind environment. The wind was represented both by detailed measured wind profiles including turbulence and by filtered measured wind profiles with the turbulence considered separately using power spectral density (PSD) methods.

In general, probabilistic variables may include any factors not deterministically known, including initial conditions, propulsion characteristics, alignment tolerances, and mass properties. For time-domain simulations, sample values of individual random variables may be generated using digital random number generators such as those described in References 13 and 14. Sample time histories of random processes such as wind turbulence can be generated from PSD data using the technique described in Reference 15. Of course, actual sample values or sample time histories from test data may be used directly as the random inputs to a Monte Carlo time-domain simulation.

A major consideration in the general application of the Monte Carlo method is to reduce the required cost of simulation as much as possible. In Reference 16 (page 146), H. Kahn describes several such techniques. Two of these (Russian Roulette and Use of Expected Values) have been used successfully in structural load analyses. Russian Roulette involves concentrating the computational effort on cases of special interest. For a landing dynamics analysis, the cases of interest may be those having the largest initial kinetic energy which therefore result in the largest structural loads. For a flight loads analysis, the cases of interest may be those having the wind profiles resulting in largest loads; the critical profiles are identified using very greatly simplified flight simulations. These cases identified as being of special interest are then analyzed using the more detailed simulation methods. The Use of Expected Values is merely a separation of computational tasks into what can be efficiently calculated analytically and what must be simulated by Monte Carlo methods. An example of this technique is the separation of the wind profile into small-scale turbulence (efficiently treated by PSD methods) and large-scale variations as described in Reference 10.

Another technique which may be used successfully for determining probabilistic design limit loads is the statistical estimation method. As an extension of the normal confidence limit concept, this method is based on the generally valid representation of random limit loads by the lognormal probability law. The expression of the one-sided normal confidence limit as derived in Reference 4 is valid for Monte Carlo samples of 50 or more observations. This expression can be simply modified as follows to be valid for samples as small as 20 observations.

Let y_1, y_2, ..., y_n be n independent observations of a normal random variable with mean m_y and standard deviation s_y. The unbiased estimates of the sample mean and variance, which are stochastically independent, are given by

$$m_y^* = \frac{1}{n} \sum_{i=1}^{n} y_i \qquad (40)$$

$$s_y^{*2} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - m_y^*)^2 \qquad (41)$$

According to Wilks (Reference 17, page 208), the sample mean (m_y^*) is normally distributed with mean (m_y) and standard deviation (s_y/n) and the sample variance is distributed as follows

$$(n-1)\frac{s_y^{*2}}{s_y^2} \sim \chi^2_{n-1} \tag{42}$$

The Chi-square distribution with k degrees of freedom (χ_k^2) is approximately normal for large k (Reference 17, page 189). However, a much more rapidly converging approximation is given by Bowker and Lieberman (Reference 8, page 556):

$$\sqrt{2\chi_k^2} \sim N\left(\sqrt{2k-1},\, 1\right) \tag{43}$$

The close convergence of this approximation for 20 degrees of freedom is shown in Figure 3.2-1. Combining Equations (42) and (43) results in the following approximate distribution for the sample standard deviation for n as small as 20:

$$s_y^* \sim N\left(s_y\sqrt{\frac{2n-3}{2n-2}},\, \frac{s_y}{\sqrt{2(n-1)}}\right). \tag{44}$$

Define the true α x 100 percentile load by

$$F_\alpha = m_y + K_\alpha \cdot s_y \tag{45}$$

where $$\alpha = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{K_\alpha} e^{-\frac{1}{2}t^2}\,dt$$

The statistical estimate of F_α is

$$\hat{F} = m_y^* + K\cdot s_y^* \tag{46}$$

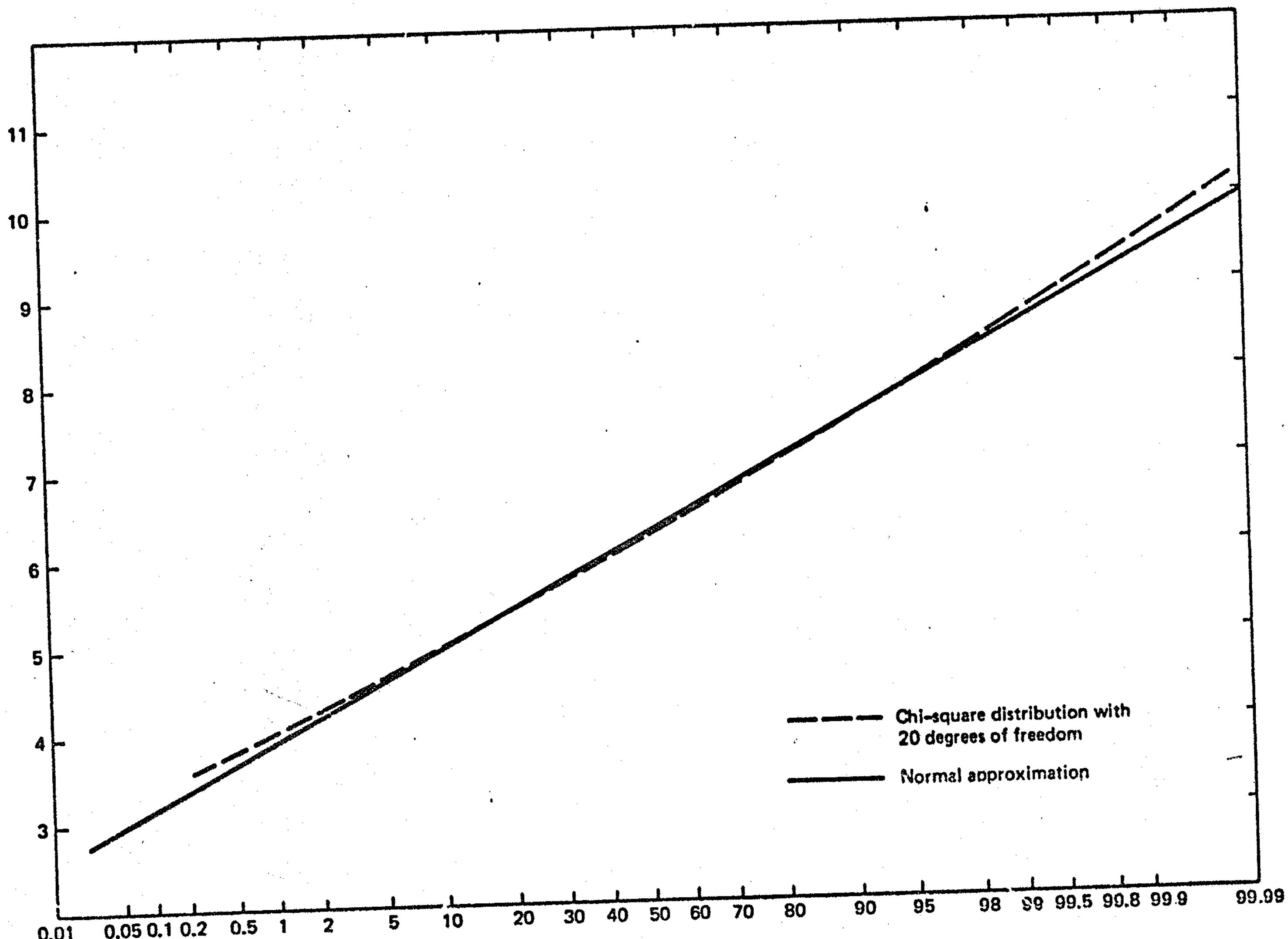

Figure 3.2-1. Normal Approximation to Chi-square Distribution

From Equations (35), (36), and (44), the mean and variance of $\hat{F}$ are

$$E[\hat{F}] = m_Y + K \cdot s_Y \sqrt{\frac{2n-3}{2n-2}} \tag{47}$$

$$Var[\hat{F}] = \frac{s_Y^2}{n} + \frac{K^2 \cdot s_Y^2}{2(n-1)} \tag{48}$$

The one-sided confidence limit equation is

$$P[F_\alpha < \hat{F}] = \beta \tag{49}$$

Equation (49) implies that

$$-K_\beta = \frac{F_\alpha - E[\hat{F}]}{\sqrt{Var[\hat{F}]}} \tag{50}$$

where

$$\beta = \int_{-K_\beta}^{\infty} e^{-\frac{1}{2}t^2} dt$$

Substituting Equations (45), (47), and (48) into Equation (50) and solving for the appropriate root of K yields

$$K = \frac{K_\alpha \cdot \sqrt{C} + K_\beta \cdot \sqrt{B + A/n}}{A} \tag{51}$$

where

$$A = C - K_\beta^2/2(n-1)$$
$$B = K_\alpha^2/2(n-1)$$
$$C = (2n-3)/(2n-2)$$

Equation (51) may be used with Equation (46) to determine the one-sided confidence limit for any probability level (α) and confidence level (β) so long as the sample size (n) is at least 20.

The statistical estimation method for use in estimating the limit load probability distribution from at least 20 unbiased Monte Carlo observations is described as follows for each load quantity of interest:

(1) Calculate the sample mean and standard deviation of the natural logarithms of the observed loads using Equations (40) and (41).

(2) Calculate the one-sided confidence limits for several different probability levels (α) for a given confidence level (β) using Equations (46) and (51)

$$\hat{F}(\alpha 1, \beta) = m_Y^* + K(\alpha 1, \beta) \cdot s_Y^*$$
$$\hat{F}(\alpha 2, \beta) = m_Y^* + K(\alpha 2, \beta) \cdot s_Y^*$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

(3) Solve for the mean and standard deviation of the logarithms which provide the least-squares fit to the following equations:

$$\hat{F}(\alpha 1, \beta) = m_Y(\beta) + K_{\alpha 1} \cdot s_Y(\beta)$$
$$\hat{F}(\alpha 2, \beta) = m_Y(\beta) + K_{\alpha 2} \cdot s_Y(\beta)$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

(4) Convert $m_y(\beta)$ and $s_y(\beta)$ to lognormal mean $m_x(\beta)$ and coefficient of variation $V_x(\beta)$ using the following standard expressions

$$V_x(\beta) = \left[\exp\left(s_Y(\beta)^2\right) - 1\right]^{\frac{1}{2}} \qquad (52)$$

$$m_x(\beta) = \left[1 + V_x(\beta)^2\right]^{\frac{1}{2}} \cdot \exp\left(m_Y(\beta)\right) \qquad (53)$$

Equations (52) and (53) are consistent with the following notation:

$$Y = \ln X$$

where
$$X \sim \text{Lognormal}\left(m_X, V_X = s_X/m_X\right)$$
$$Y \sim \text{Normal}\left(m_Y, s_Y\right)$$

This statistical estimation method provides conservative estimates of the lognormal parameters of limit loads determined from at least 20 Monte Carlo simulations. The degree of conservatism in the estimated parameters is, of course, dependent on the confidence level (β) chosen. The estimated parametric values are also somewhat dependent on the particular probability levels (αi) chosen for the least-squares fit. A numerical demonstration of this method is presented in Section 4.2.

3.3 Limit Loads from Frequency - Domain Simulation

In the practical solution of random vibration problems, the dynamic characteristics of a structural system are usually assumed to be linear and deterministic, and the excitation is assumed to be random. Furthermore, the random excitation is usually assumed to be stationary and Gaussian with zero mean value, since the random process for the response can then be completely characterized by its power spectral density function (Reference 9, page 89). Solutions to two random vibration problems for this special case of stationary Gaussian response are available in the literature (Reference 18, page 293). The threshold-crossing problem is concerned with the expected rate at which a random process $X(t)$ exceeds a certain value. The peak-distribution problem is concerned both with the probability distribution of peak magnitudes in $X(t)$ and with the expected rate of occurrence of the peaks. However, neither of these available solutions provides the extreme-value probability distribution required for probabilistic ultimate strength design. The objective of the present study is to determine the probability distribution of the extreme largest value, for a stationary Gaussian random process $X(t)$, occurring within a given mission length. This required limit-load probability distribution will be expressed in terms of the power spectral density function (PSD) of the calculated load.

The real autocorrelation function associated with a real-valued stationary random process $X(t)$ may be defined by

$$R_{xx}(\tau) \equiv \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)\cdot X(t+\tau)\ dt \tag{54}$$

Equations relating the autocorrelation function and the power spectral density function (PSD) are known as the Wiener-Khintchine relations (Reference 19, page 579). For a real-valued random process, such as the random load in a structural member, the defining equations may be written as follows:

$$R(\tau) = \int_0^\infty G(\omega)\cos\omega\tau \, d\omega \qquad (55)$$

$$G(\omega) = \frac{2}{\pi}\int_0^\infty R(\tau)\cos\omega\tau \, d\tau \qquad (56)$$

where $G(\omega)$ is the load PSD with frequency (ω) in radians/second.

The load PSD may alternatively be written with frequency in Hz as follows

$$R(\tau) = \int_0^\infty \Gamma(f)\cos 2\pi f\tau \, df \qquad (57)$$

$$\Gamma(f) = 4\int_0^\infty R(\tau)\cos 2\pi f\tau \, d\tau \qquad (58)$$

where $f = \omega/2\pi$ in Hz.

For some applications, the load PSD may be more conveniently defined in terms of spatial frequency (radians per unit distance) and spatial distance instead of circular frequency (radians per second) and time. Equations (55) and (56) with appropriate notation changes may be used as the defining Wiener-Khintchine relations for such applications.

With no loss of generality, a stationary random process may be assigned a zero mean value. The variance of such a real-valued random process is obtained from Equations (55) and (57) by evaluating the autocorrelation function for zero time lag,

$$R(0) = \sigma^2 = \int_0^\infty G(\omega)d\omega = \int_0^\infty \Gamma(f)\, df \qquad (59)$$

Equations (54) through (59) form a consistent set of definitions for use in harmonic analysis of stationary random processes. Since many authors use alternate forms of the Wiener-Khintchine relations (Reference 19, page 580), special care is required when applying formulas for random vibration analysis.

Standard methods are available for computing the PSD of loads in a linear structure due to stationary Gaussian excitation (References 18, 20, 21). The output response PSD for the r^{th} calculated load quantity is given by the following general equation:

$$G_r(\omega) = \lfloor L_{ri}^*(j\omega) \rfloor \; [G_f(j\omega)]\{L_{ir}(j\omega)\} \tag{60}$$

where $\{L_{ir}(j\omega)\}$ is the column matrix of complex frequency responses for the r^{th} load quantity and for i excitation points,

$\lfloor L_{ri}^*(j\omega) \rfloor$ is a row matrix of the complex conjugates of $L_{ir}(j\omega)$, and

$[G_f(j\omega)]$ is the PSD matrix of input power spectral density functions for each of the i excitation points and cross-power spectral densities between the excitation points.

The following development converts the Gaussian load PSD typically defined by Equation (60) into an extreme-value limit-load probability distribution required for probabilistic structural design.

The critical parameter in the three distributions used for describing extreme normal variates is the characteristic largest value (u). Its magnitude increases with sample size until, as n becomes very large, it converges to the most probable value (mode) of the asymptotic extremal type I distribution (Reference 1, page 172). However, as described in Section 2.1, the convergence of the normal extremes to the type I distribution is so slow that the lognormal and extremal type III distributions must be used for small and moderately sized samples. The following development is based on expressing the characteristic largest value in terms of Rice's theorem for the expected number of threshold crossings per unit time.

According to Rice (Reference 22, page 192), the expected rate of zero crossings from below for a stationary Gaussian process with zero mean is given by

$$E[N_+(0) = \left[\frac{\int_0^\infty f^2\,\Gamma(f)\,df}{\int_0^\infty \Gamma(f)\,df}\right]^{1/2} \tag{61}$$

With the PSD defined in radians per second (or radians per unit distance), Equation (61) becomes

$$E[N_+(0)] = \frac{1}{2\pi}\left[\frac{\int_0^\infty \omega^2 \, G(\omega) \, d\omega}{\int_0^\infty G(\omega) \, d\omega}\right]^{1/2} \tag{62}$$

With the random structural load PSD defined by Equation (60) the integrals of Equations (61) and (62) will converge whenever the input PSD has a finite variance.

The equation for the expected number of times per unit time or distance that the Gaussian load passes through the threshold value (ξ) with positive slope is given by Rice (Reference 22, page 192) as follows:

$$E[N_+(\xi)] = E[N_+(0)] \exp\left(\frac{-\xi^2}{2\sigma^2}\right) \tag{63}$$

where $E[N_+(0)]$ is defined by Equation (61) or (62) and σ^2 is defined by Equation (59).

Equation (63) may also be found in Reference 18 (page 297), Reference 21 (page 42), and Reference 23 (page 5.121) among many other sources. It is restricted to stationary Gaussian random processes having zero mean values. Since the Gaussian model is commonly used to represent inflight atmospheric turbulence (Reference 23, page 5.116) and transonic buffeting (Reference 20), this restriction is not significant to most current engineering applications.

The expected number of threshold crossings in a given time or distance interval (T) is obtained simply by modifying Equation (63) as follows:

$$E[N_+(\xi)]_T = T \cdot E[N_+(0)] \exp\left(\frac{-\xi^2}{2\sigma^2}\right) \tag{64}$$

where T defines the length of a mission.

The desired characteristic largest value in a sample of size n, u, is defined as follows by Gumbel (Reference 1, page 82): "In n observations, the expected number of values equal to or larger than u is unity." Thus, by definition, the characteristic largest value for a mission of length T is determined from Equation (64) by setting the expected number of threshold crossings to unity. The required characteristic largest value for the standardized normal variate is then

$$\hat{u} = \frac{u}{\sigma} = [2\ \ell n(T \cdot E[N_+(0)])]^{1/2} \tag{65}$$

where σ is defined by Equation (59) and

$E[N_+(0)]$ is defined by either of Equations (61) or (62).

This characteristic largest value for a stationary Gaussian random process having zero mean is sufficient to completely define any of the three distributions used for normal extremes.

As discussed in Section 2.1, the convergence of normal extremes to the type I asymptotic extreme-value distribution is extremely slow. The type I distribution is therefore recommended for describing normal extremes only when the characteristic largest value for the standardized variate exceeds 8. The extremal intensity function ($\hat{\alpha}$) corresponding to the standardized characteristic largest value ($\hat{u}$) is given by Gumbel (Reference 1, page 137) as follows for normal extremes:

$$\hat{\alpha} = \hat{u} + \hat{u}^{-1} - 2\hat{u}^{-3} + 10\hat{u}^{-5} \tag{27}$$

The extremal type I cumulative distribution function for the standardized variate (y) is given by

$$\phi^{(1)}(y) = \exp(-e^{-\hat{\alpha}(y-\hat{u})}) \tag{26}$$

where $\hat{u}$ and $\hat{\alpha}$ are defined by Equations (65) and (27), respectively.

The extremal type III distribution, which converges to the type I distribution with increasing $\hat{u}$, is recommended for characteristic largest values between 3 and 8 for the standardized normal variate. The type III distribution function for normal extremes was first suggested by Fisher and Tippett (Reference 5). A suitable form of this distribution for the standardized variate is

$$\phi^{(3)}(y) = \exp\left[-\left(\frac{y}{\hat{u}}\right)^{-k}\right] \tag{23}$$

$$\text{where } k = \frac{(\hat{u}^2+1)^2}{(\hat{u}^2-1)} \tag{24}$$

The percentage points of this distribution as a function of the cumulative probability, p, are given by Gumbel (Reference 1, page 299) as

$$y = \frac{x}{\sigma} = \exp\left[\ln \hat{u} - \frac{\ln(-\ln p)}{k}\right] \tag{25}$$

As discussed in Section 2.1, for characteristic largest values of the standardized normal extremes close to 2, the lognormal distribution is essentially identical to the actual distribution of normal extremes calculated by Tippett and plotted in Reference 1 (page 129). For values less than 3, the lognormal approximation is generally more accurate than the extremal type III approximation and is therefore recommended for this range. The lognormal probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\ \delta x} \exp - \frac{1}{2} \left(\frac{\ln x - \gamma}{\delta} \right)^2 \tag{16}$$

$$\text{where } \gamma = \ln \breve{x} + \ln \sigma \tag{66}$$

$$\delta = \left[\ln\left(\frac{\breve{x}}{\tilde{x}} \right)\right]^{1/2} \tag{67}$$

$\breve{x}$ is the median of the standardized normal extreme, and

$\tilde{x}$ is the mode of the standardized normal extreme.

The required median ($\breve{x}$) of the standardized normal extreme is obtained from the following equation:

$$F(\breve{x}) = \exp[0.69315(F(\hat{u})-1)] \tag{20}$$

where $F(\hat{u})$ is the normal cumulative distribution function evaluated at $\hat{u}$. Equation (20) is a modified form of the equation for medians of extreme values given by Gumbel (Reference 1, page 79). The Gaussian probability functions are tabulated, for example, in Reference 8 (page 555) and Reference 24 (page 33).

The required mode ($\tilde{x}$) of the standardized normal extreme is obtained from the following equation:

$$\frac{\tilde{x} \cdot F(\tilde{x})}{f(\tilde{x})} = \frac{F(\hat{u})}{1-F(\hat{u})} \tag{22}$$

Equation (22) is a modified form of the equation for modes of normal extremes given by Gumbel (Reference 1, page 133). Tabulated values of $F(x)/f(x)$ may be found in Reference 2 (page 11).

4.0 NUMERICAL DEMONSTRATION

The methods previously described for determining limit-load probability distributions from time-domain and frequency-domain analyses have certain limitations which may be best illustrated by numerical examples. Section 4.1 presents three numerical examples of the Taylor's series method which demonstrate the effects of the method's fundamental assumptions. Section 4.2 demonstrates the statistical estimation method which may be used to reduce the required number of Monte Carlo simulations. These numerical examples are based on sets of random numbers generated by a digital computer. The method for determining limit loads from a frequency-domain analysis is demonstrated using numerical data obtained from an analog Gaussian noise generator. Section 4.3 presents the results of three examples of this method.

4.1 Examples of Taylor's Series Method

As discussed in Section 3.1, the Taylor's series method for estimating the probability distribution of a nonlinear function of several random variables is based on the following three assumptions:

(1) that the higher-order terms in the Taylor expansion of the function are negligible compared with the first order terms,
(2) that the individual random variables are mutually independent, and
(3) that the individual random variables are each normally distributed.

The following is a brief discussion of the implications of these assumptions with numerical examples.

Consider a function of four random variables

$$f(W, X, Y, Z) = W^2 \cdot X / Y + Z$$

where

$$\begin{array}{ll} W \sim \text{Normal} & (m_W = 20,\ s_W = 2) \\ X \sim \text{———} & (m_X = 2,\ s_X = 0.57735) \\ Y \sim \text{Normal} & (m_Y = 2,\ s_Y = 0.2) \\ Z \sim \text{Normal} & (m_Z = 0,\ s_Z = 50) \end{array}$$

By the Taylor's series method, the estimates of the mean, variance, and standard deviation of the function are as follows:

$$m_f \approx m_W^2 \cdot m_X / m_Y + m_Z = 400$$

$$s_f^2 \approx \left(2\frac{m_W \cdot m_X}{m_Y}\right)^2 \cdot s_W^2 + \left(\frac{m_W^2}{m_Y}\right)^2 \cdot s_X^2 + \left(\frac{m_W^2 \cdot m_X}{m_Y^2}\right)^2 \cdot s_Y^2 + s_Z^2$$

$$= 23{,}833$$

$$\sim 1544$$

The stochastic behavior of this function was studied for three different cases. Case 1 involved dependent variables (ρ_{WY} = -0.5) and a non-normal variable with the variable X being uniformly distributed in the range 1 to 3. Case 2 involved a non-normal variable ($X \sim U\ (1,3)$) but all variables were independent. Case 3 involved all normal and independent variables. The mean and standard deviation and the cumulative distribution function (CDF) were determined from a Monte Carlo simulation using a sample size of 2000 for each of the three cases. The Monte Carlo simulations were performed with the Boeing Generalized Statistics Program (GESP) described in References 13 and 14. The resulting means and standard deviations are presented in Table 4.1-1 for comparison with the Taylor's series estimates. Results of significance tests of the hypothesis that the Monte Carlo parameters are identical to the Taylor's series parameters are also presented in Table 4.1-1 along with the results of a Chi-square test for normality (Reference 8, page 366). The hypothesis test for the mean was performed using Student's t statistic (Reference 8, page 127). The hypothesis test for the standard deviation was performed using the Chi-square statistic (Reference 8, page 138). The acceptance probabilities for such hypothesis tests are usually established at either one percent or five percent levels. Values of the 37-degree-of-freedom Chi-square statistic corresponding to these probability levels are 59 and 52, respectively.

For this particular function, the Monte Carlo means and standard deviations are seen to approach the Taylor's series parameters as the assumptions of independence and normality of the individual variables are better satisfied. The hypothesis tests indicate that the mean determined by the Taylor's series method is sufficiently accurate regardless of normality and independence of

Table 4.1-1 Numerical Evaluation of Taylor's Series Method

Method	m_f	s_f	$P[Y_m > m_f]$	$P[Y_s > s_f]$	χ^2_{37}
Taylor's series	400.0	154.4	-	-	-
Monte Carlo Case 1	409.4	172.3	0.008	0	200.1
Monte Carlo Case 2	405.9	159.0	0.049	0.028	133.3
Monte Carlo Case 3	402.8	156.8	0.210	0.164	127.6

the individual random variables; the standard deviation determined by the Taylor's series method is sufficiently accurate only when the individual random variables are independent. However, for none of the three cases was the hypothesis of normality verified by the Chi-square test. In Figure 4.1-1, the cumulative distribution function determined from the Monte Carlo simulation for Case 3 is plotted versus the Taylor's series normal distribution to illustrate the results of the Chi-square test.

These numerical examples are consistent with the theory discussed in Section 3.1. An accurate estimate of the mean requires only that the higher-order terms in the Taylor's series expansion are negligible, whereas an accurate estimate of the variance requires the additional assumption of independence among the individual random variables. All three assumptions must be satisfied in order that the function be approximately normally distributed. For the function studied, the second and higher partial derivatives are negligible or zero except with respect to the Y variable. The numerical influence of the neglected non-zero terms on the Taylor's series estimate of the mean and standard deviation appears to be small.

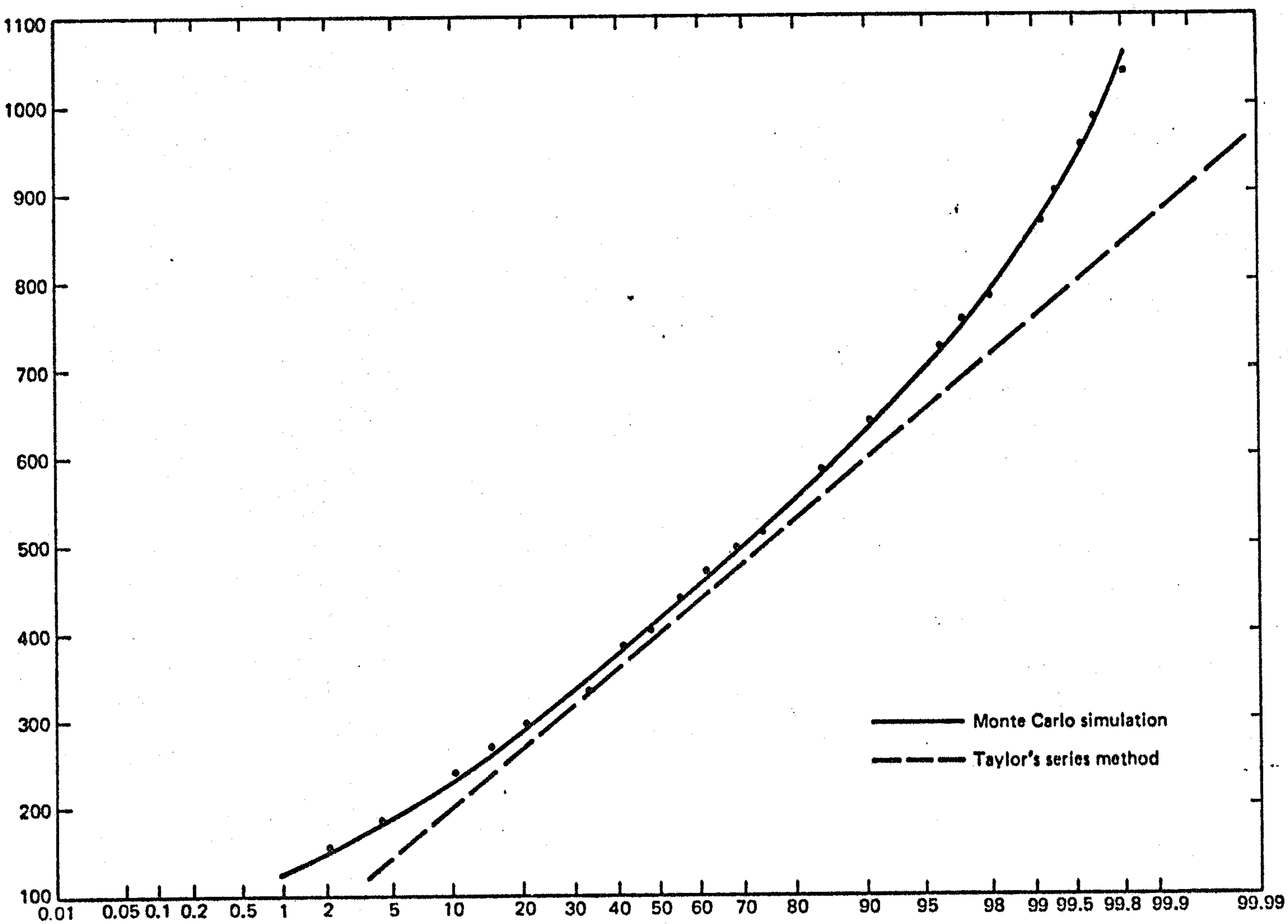

Figure 4.1-1. Taylor's Series and Monte Carlo Distributions for Case 3

4.2 Examples of Monte Carlo Method

As described in Section 3.2, the statistical estimation method is a technique for conservatively estimating the lognormal limit-load parameters from a small sample of observed loads from a Monte Carlo simulation. The numerical demonstration of this method is based on a simulated analysis in which the limit load is defined as the largest load occurring in 100 independent observations of a standardized normal variate. By means of the GESP random number generator (References 13 and 14), 2000 simulated limit loads were generated. The limit-load distribution was approximately lognormal ($P[\chi^2_{37} > 49] = 0.095$) with mean equal to 2.509 and coefficient of variation equal to 0.1715. Ten data sets of 20 values each were statistically analyzed to determine the sample mean of the logarithms (m_y^*) and the standard deviation of the logarithms (s_y^*). The best-fit mean m_y ($\hat{a}$) and standard deviation s_y ($\hat{\beta}$) of the logarithms were then conservatively estimated using the 90% one-sided confidence limit (Equation 51) for two sets of probability levels. The probability levels designated confidence fit "a" were biased to positive values: $K_\alpha = 1, 2, 3, 4, 5$. The probability levels designated confidence fit "b" were unbiased: $K_\alpha = -4, -3, -2, -1, 0, 1, 2, 3, 4, 5$. The sample data and the conservative estimates for confidence fits "a" and "b" are presented in Tables 4.2-1 and 4.2-2, respectively. For comparison purposes, the "true" sample mean and standard deviation of the logarithms based on 2000 values are $m_y = 0.905$ and $s_y = 0.1703$.

The data presented in Tables 4.2-1 and 4.2-2 are plotted on normal probability paper in Figures 4.2-1 through 4.2-10. Each plot shows, for each data set, the conservatively estimated distributions based on 20 values along with the "true" distribution based on 2000 values. Both conservative distributions result in values larger that the "true" values for the probability range of interest. Values from the biased confidence fit "a" suggest that most of the conservatism is in the estimate of the standard deviation. Values from the unbiased confidence fit "b" show a more balanced approximation to the "true" distribution.

Table 4.2-1 Parameters for Statistical Estimation Demonstration Using 90% Confidence Fit "a"

Data Set	m_y^*	s_y^*	$m_y(\beta)$	$s_y(\beta)$
1	0.886	0.1830	0.915	0.2301
2	0.804	0.1669	0.830	0.2099
3	0.893	0.2320	0.929	0.2918
4	0.872	0.1888	0.901	0.2374
5	0.910	0.1819	0.938	0.2287
6	0.901	0.1459	0.924	0.1835
7	0.877	0.1467	0.900	0.1845
8	0.930	0.1543	0.954	0.1940
9	0.815	0.1684	0.841	0.2118
10	0.859	0.1750	0.887	0.2200

Table 4.2-2 Parameters for Statistical Estimation Demonstration Using 90% Confidence Fit "b"

Data Set	m_y^*	s_y^*	$m_y(\beta)$	$s_y(\beta)$
1	0.886	0.1830	1.004	0.1995
2	0.804	0.1669	0.912	0.1820
3	0.893	0.2320	1.043	0.2530
4	0.872	0.1888	0.994	0.2058
5	0.910	0.1819	1.027	0.1983
6	0.901	0.1459	0.995	0.1591
7	0.877	0.1467	0.972	0.1600
8	0.930	0.1543	1.030	0.1682
9	0.815	0.1684	0.924	0.1836
10	0.859	0.1750	0.973	0.1907

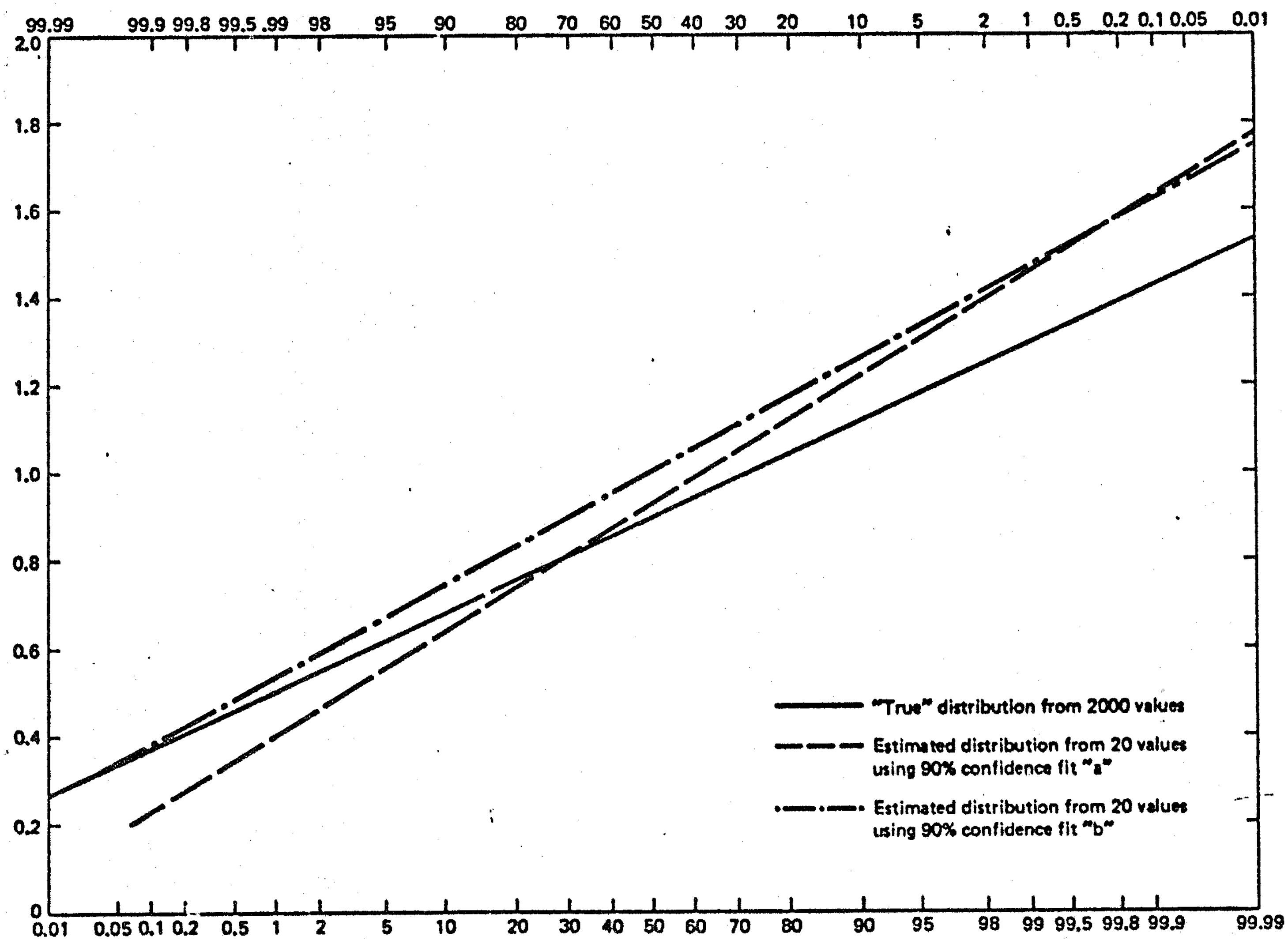

Figure 4.2-1. Example of Statistical Estimation Method – Data Set 1

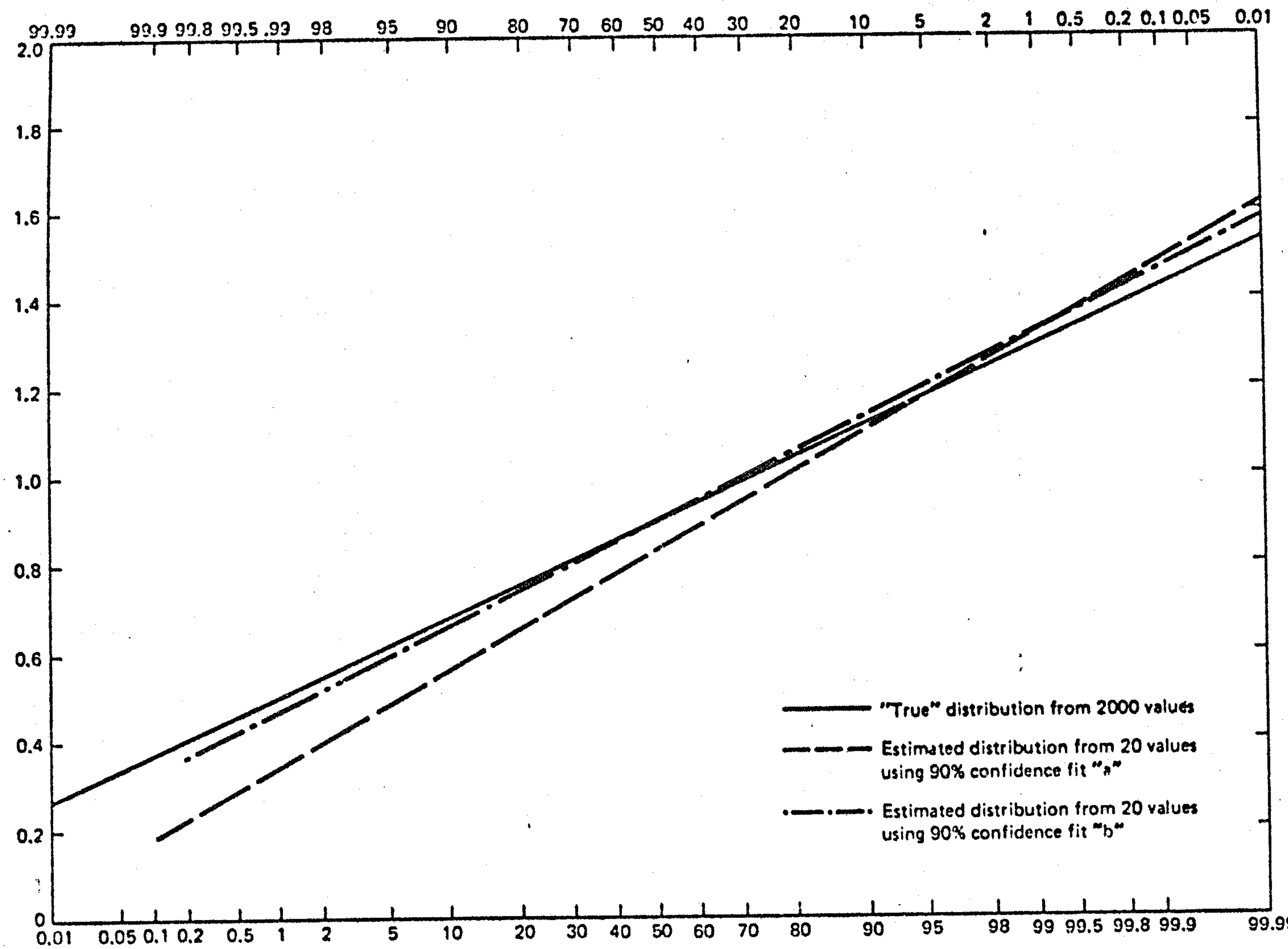

Figure 4.2-2. Example of Statistical Estimation Method — Data Set 2

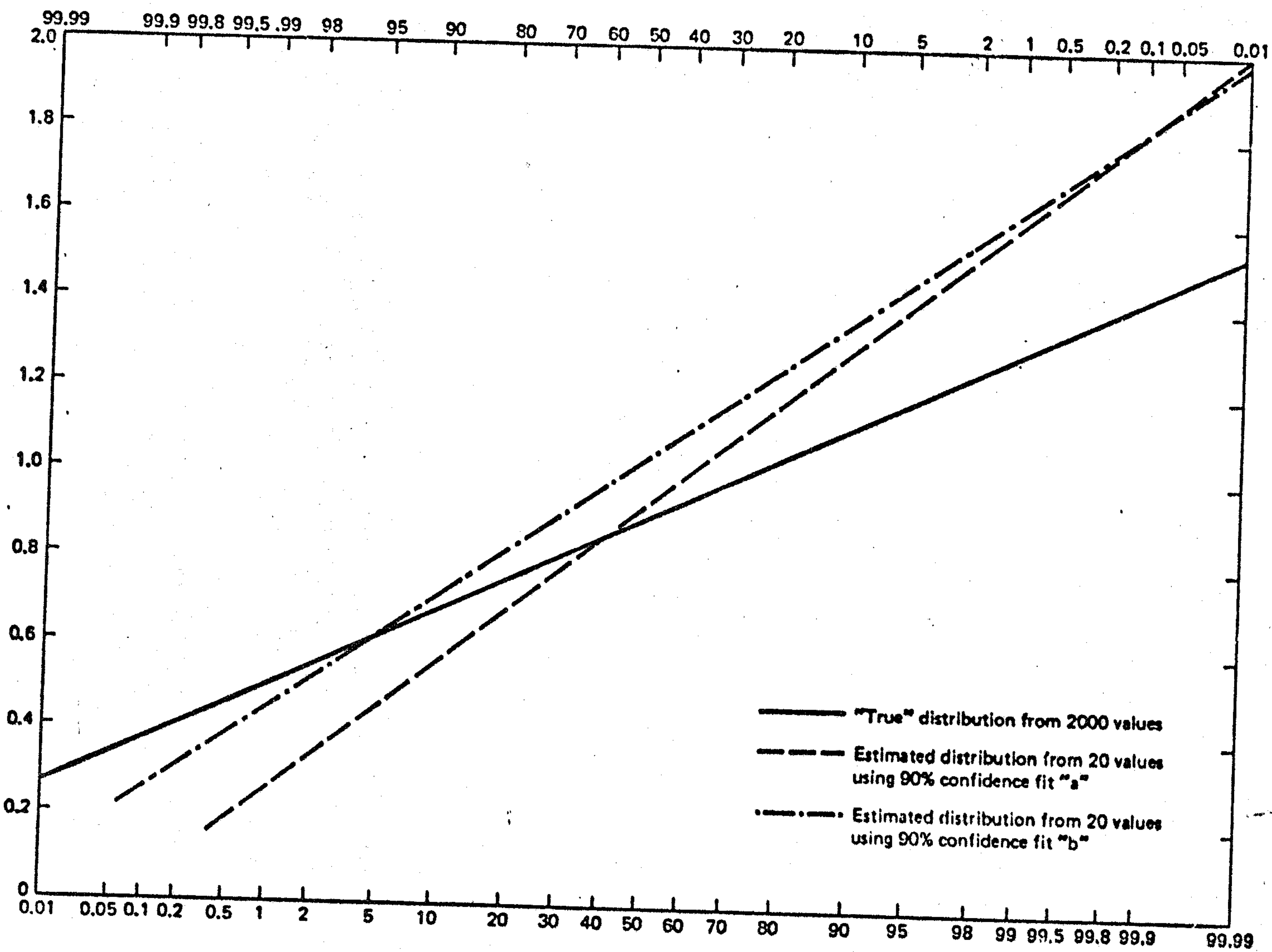

Figure 4.2-3. Example of Statistical Estimation Method — Data Set 3

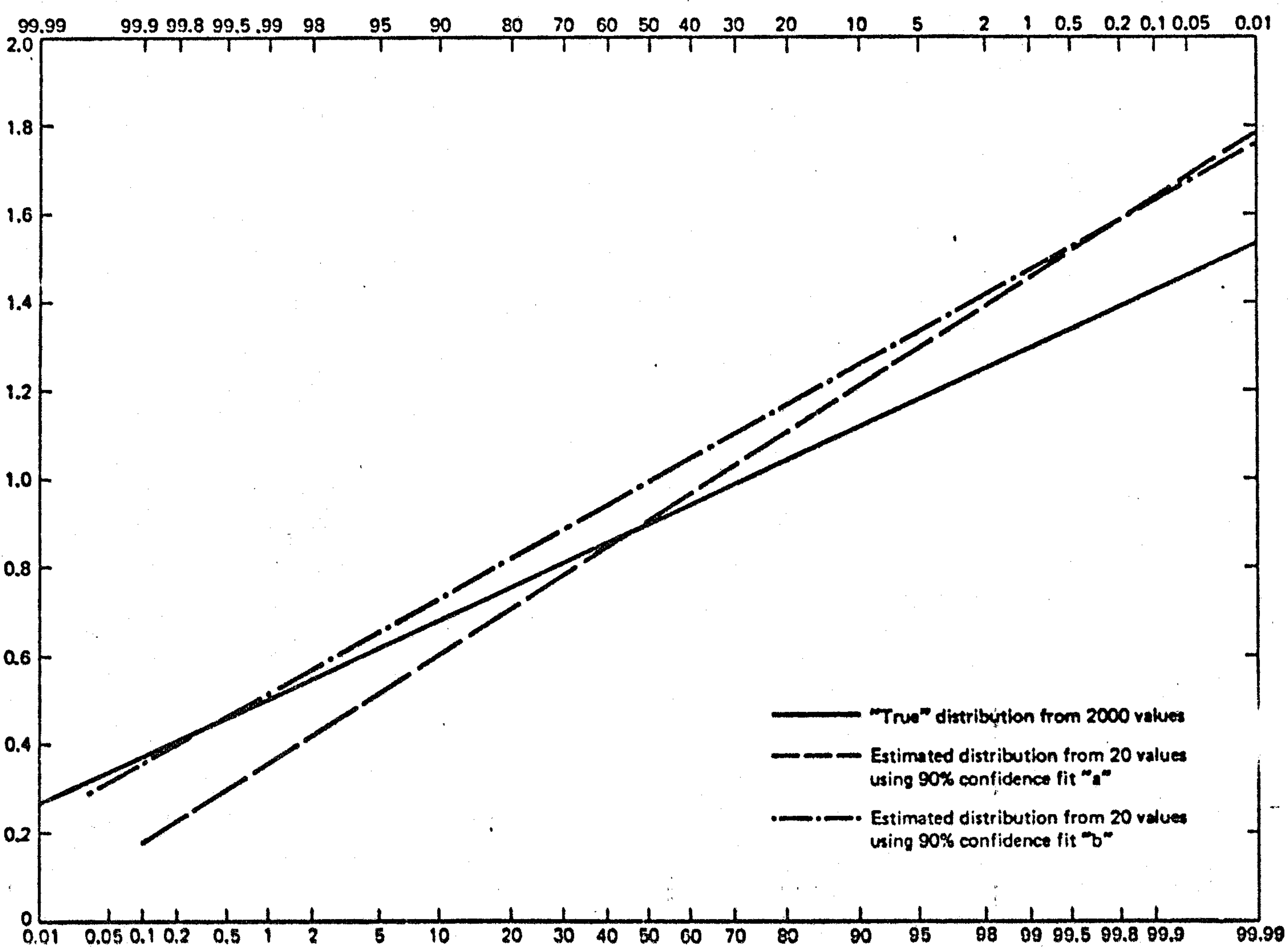

Figure 4.2-4. Example of Statistical Estimation Method — Data Set 4

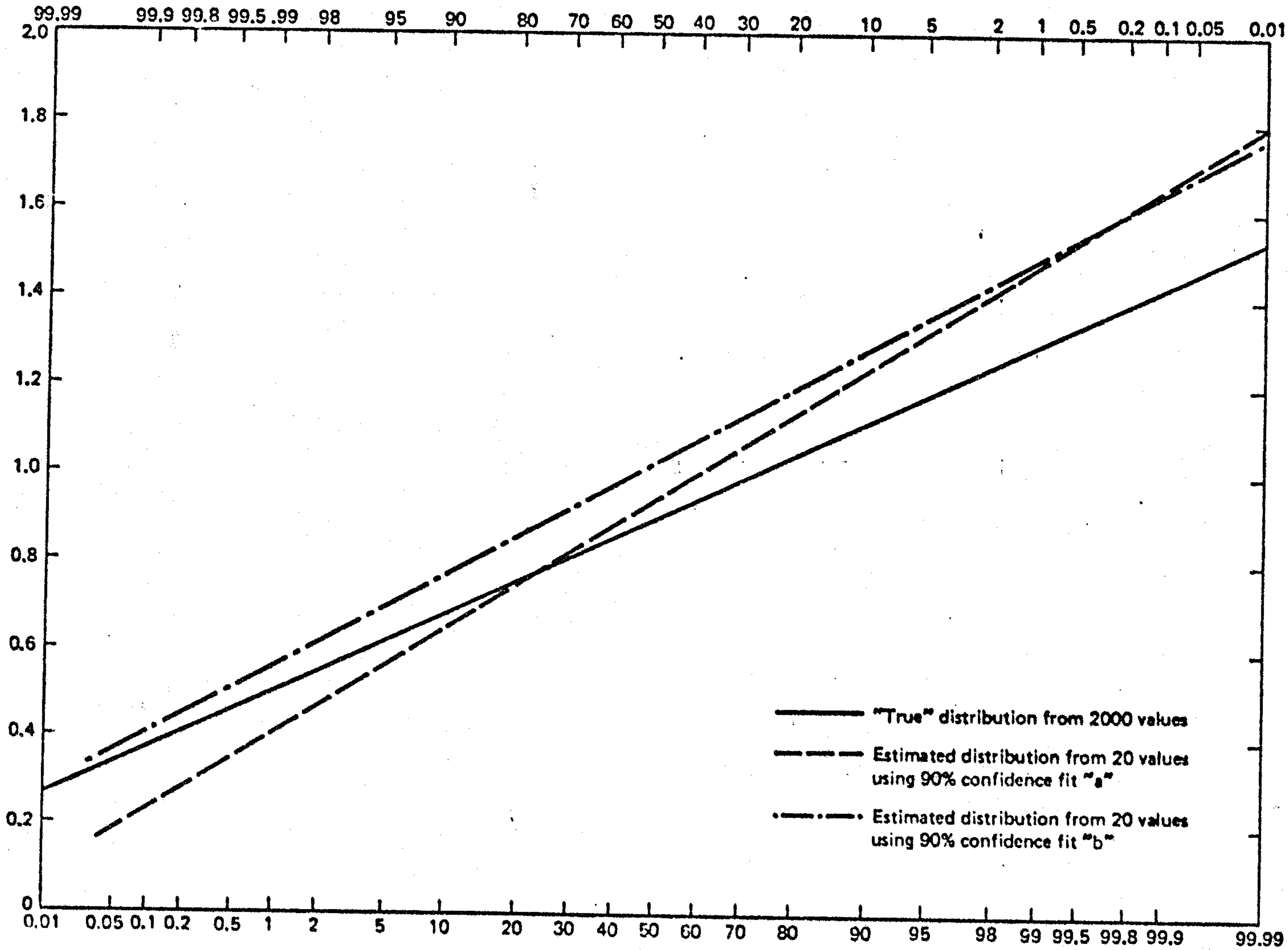

Figure 4.2-5. Example of Statistical Estimation Method — Data Set 5

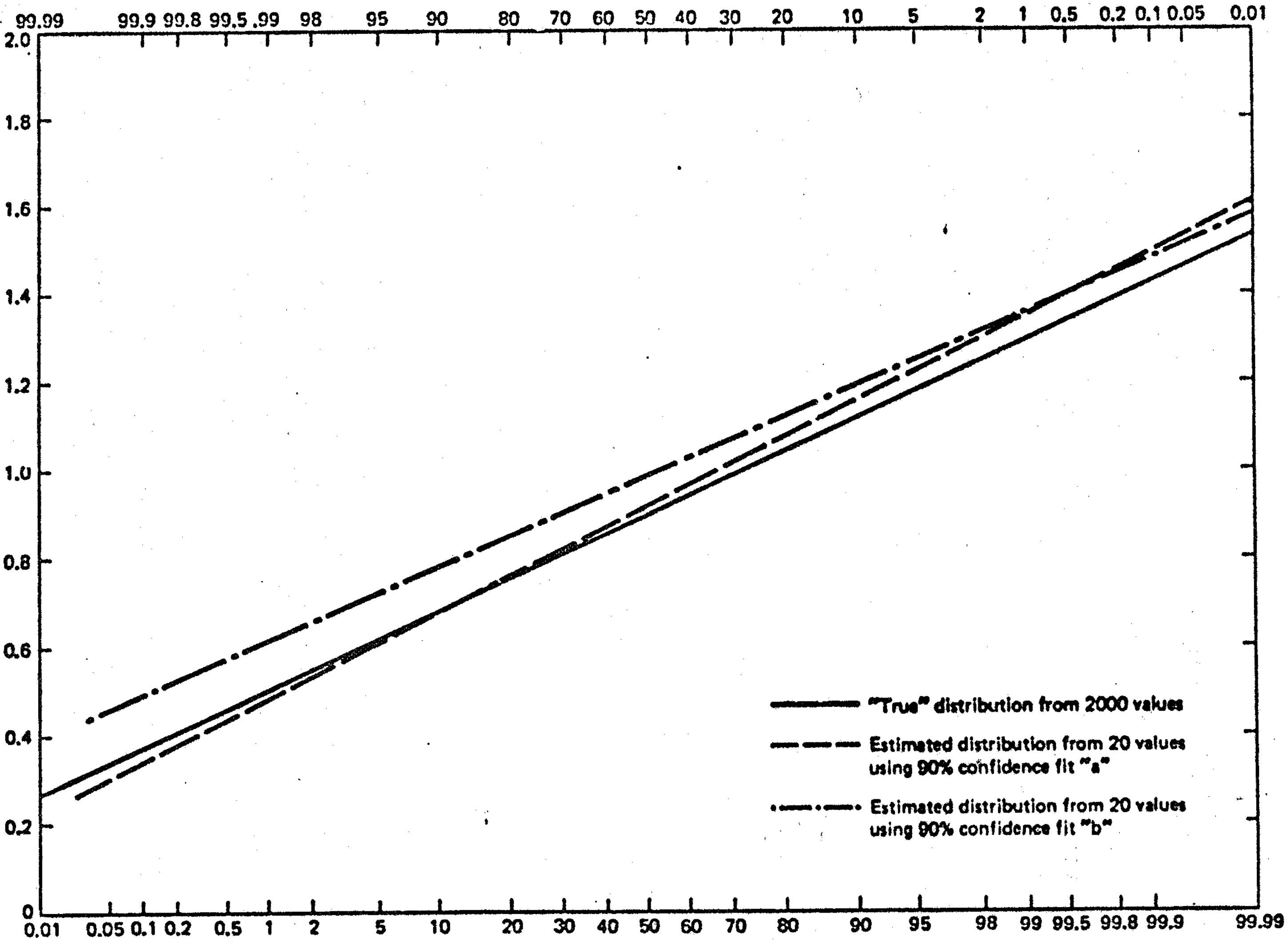

Figure 4.2-6. Example of Statistical Estimation Method – Data Set 6

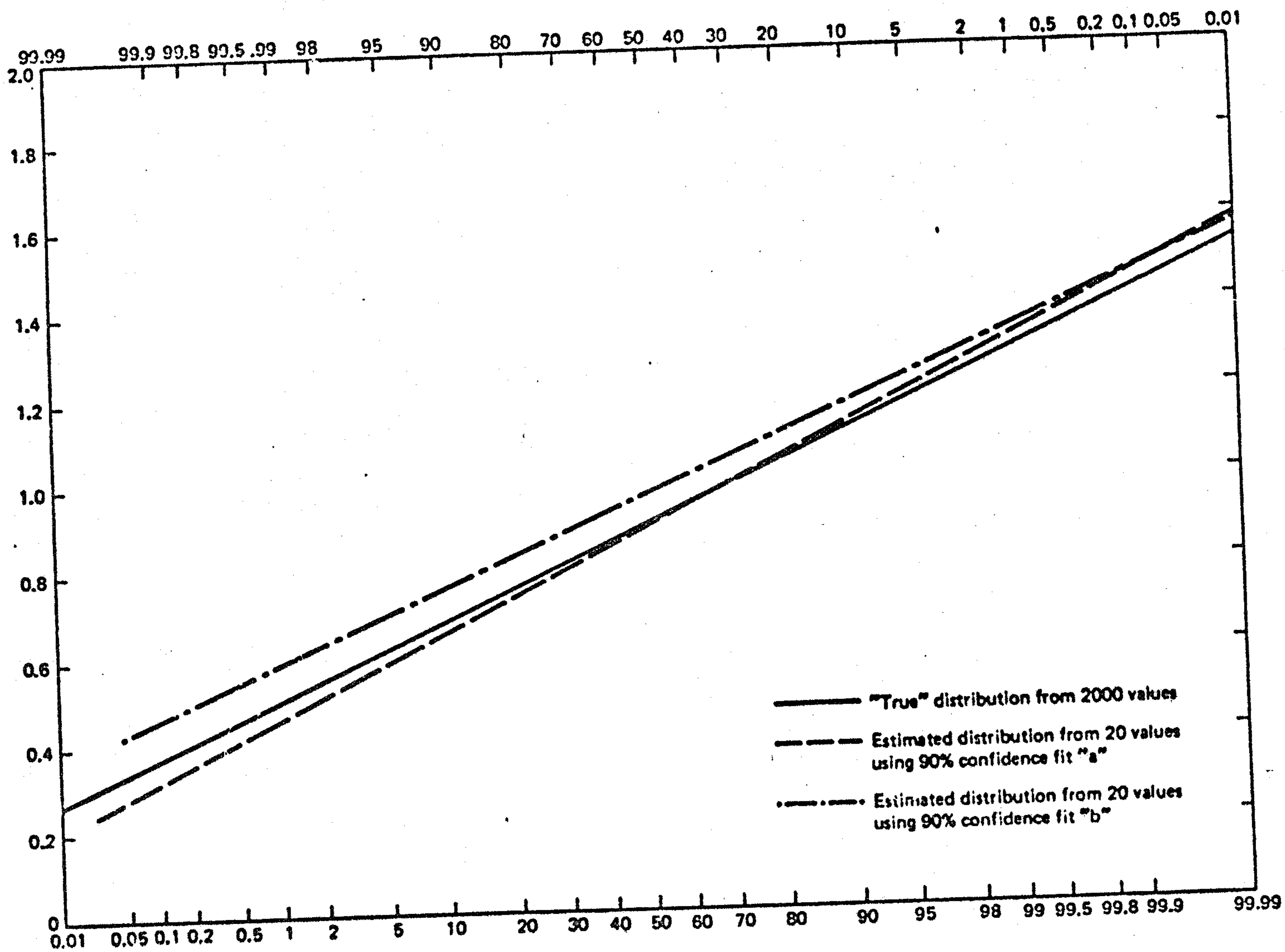

Figure 4.2-7. Example of Statistical Estimation Method — Data Set 7

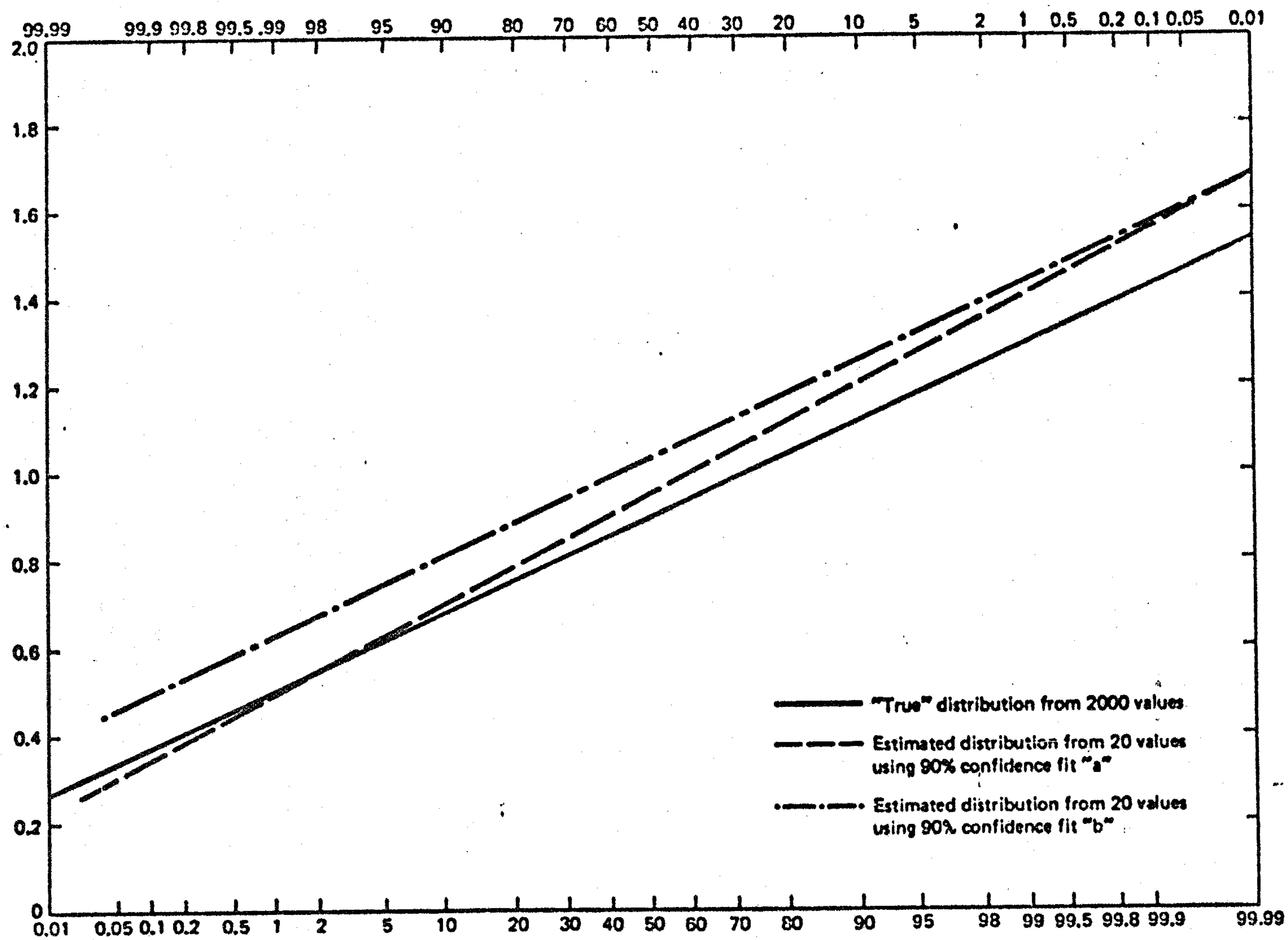

Figure 4.2-8. Example of Statistical Estimation Method – Data Set 8

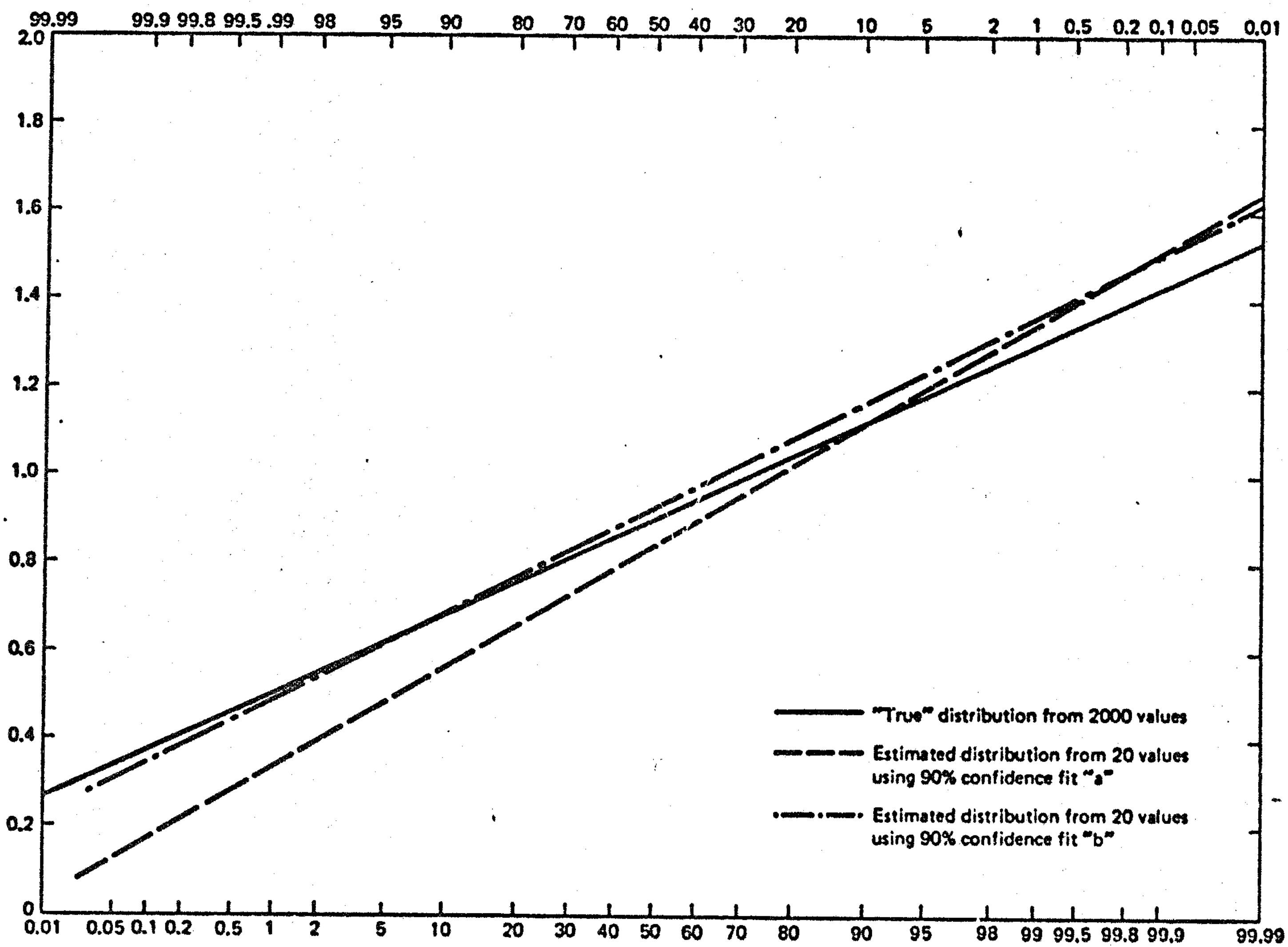

Figure 4.2-9. Example of Statistical Estimation Method – Data Set 9

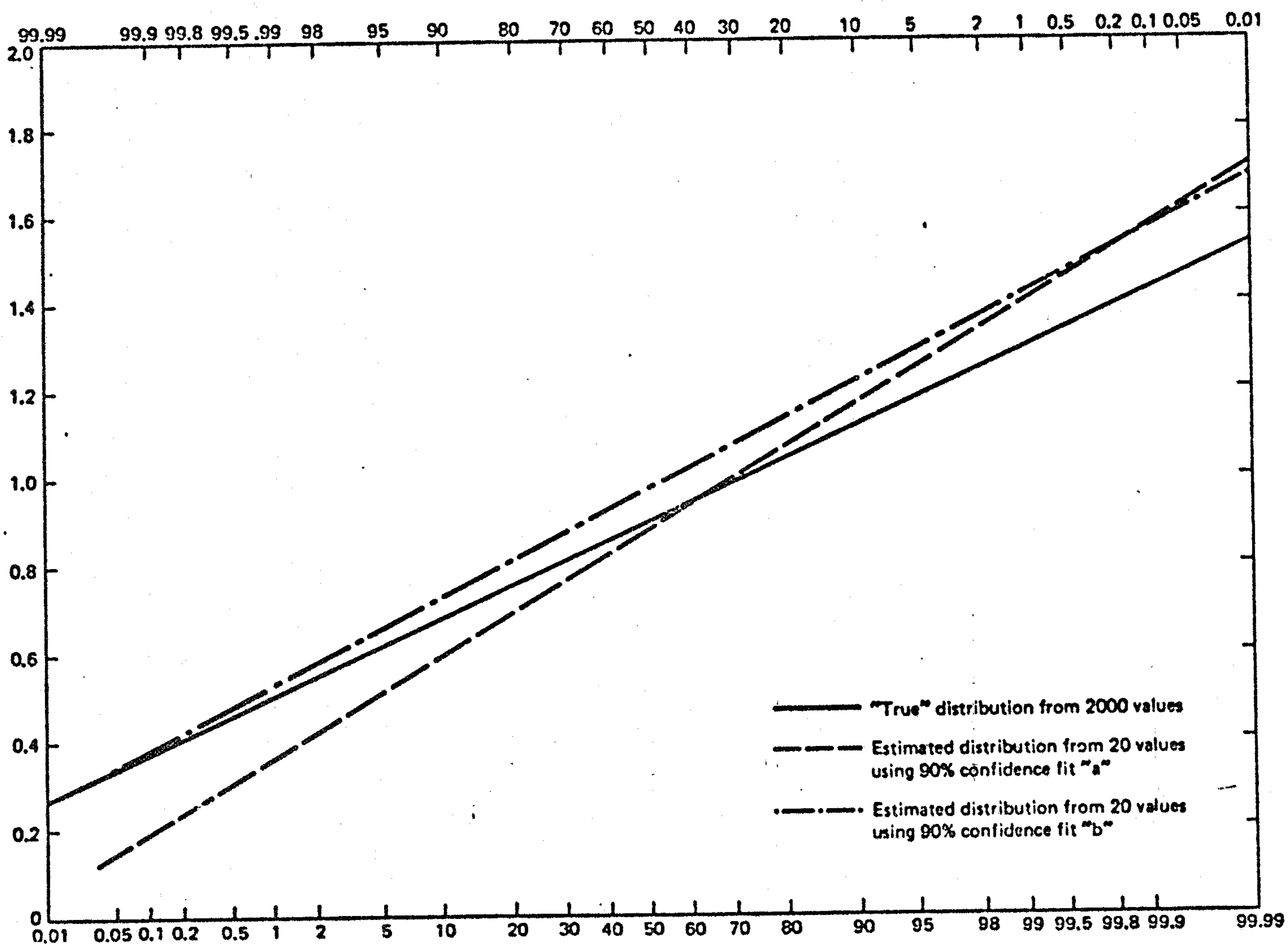

Figure 4.2-10. Example of Statistical Estimation Method — Data Set 10

4.3 Examples of Frequency - Domain Method

Numerically demonstrating the method for determining the limit-load probability distribution from the power spectral density function (PSD) of a Gaussian random process has two general aspects. The first is demonstrating that the extreme values from a continuous Gaussian time series of specified duration behave mathematically as extremes from a population of discrete normal variates. The second is demonstrating that the lognormal and extremal type III distributions provide valid representations of the actual distribution of extremes when based on the following expression for the standardized characteristic largest value derived in Section 3.3:

$$\hat{u} = \frac{u}{\sigma} = [2\ \ln(T \cdot E[N_+(0)])]^{1/2} \qquad (65)$$

Both of these aspects will be demonstrated using numerical data obtained from an analog Gaussian noise generator.

The Elgenco Model 311A Gaussian Noise Generator was used to obtain the required random time histories. This electronic device provides a stable and reliable source of Gaussian random noise having the following characteristics:

(1) The output PSD is uniform to ± 0.1 dB from 0 to 35 Hz; the output falls off rapidly above 40 Hz.

(2) The amplitude probability density function is Gaussian (normal) to less than ± 1 percent.

The output of the Gaussian noise generator was passed through three first-order filters, all having cutoff frequencies of 25 Hz. The purpose of this filtering was to specify accurately the high-frequency roll-off so that the actual PSD could be precisely defined. The PSD used for the numerical demonstration is defined as

$$\Gamma(f) = \frac{a^2 f_c^{\,6}}{(f^2+f_c^{\,2})^3} \qquad \text{for } 0 \le f \le 40 \tag{68}$$

$$= 0 \qquad \text{for } f > 40$$

where f_c = 25 Hz, and

a^2 = 2.785 is the magnitude factor determined empirically from the generated output.

Figure 4.3-1 presents time histories of the unfiltered random noise produced directly by the Elgenco Noise Generator and of the random noise after it was passed through three 25 Hz filters.

To determine the actual mean and standard deviation of the generated time series, a statistical analysis of the filtered output time history was performed, based on the assumptions of ergodicity and stationarity. A twenty-second duration of the output from the noise generator was sampled at 0.02-second intervals to provide 1000 data points. The mean and standard deviation of this large sample were then computed with the following results:

$$\mu = -0.279$$
$$\sigma = 6.418$$

These statistical estimates were assumed to be the true parameters of the generated time series for all subsequent studies.

A Chi-square goodness-of-fit test (Reference 8, p. 365) was also performed with the sample of 1000 data points to verify that the generated output was Gaussian. The Chi-square statistic, based on a division of the data into 19 cells, was 19. This value corresponds to a Chi-square cumulative probability of less than 75 percent. Therefore, the random time histories obtained from the noise generator may be justifiably considered Gaussian with parametric values as estimated.

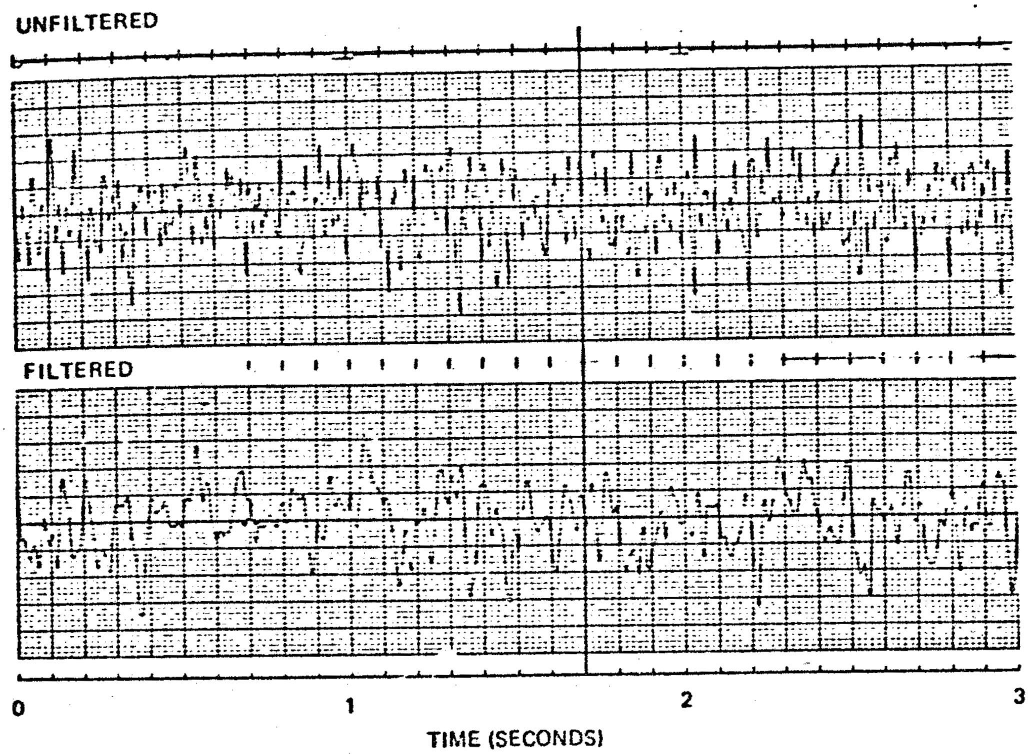

Figure 4.3-1. Time Histories of Unfiltered and Filtered Gaussian Random Noise

The magnitude factor of 2.785 used in the PSD expression defined by Equation (68) is consistent with the empirically determined standard deviation. With this PSD expression, the expected rate of zero crossings per second with positive slope was calculated to be 12.77. This compares well with the observed average of 13.0 taken from 26 seconds of the generated random time history.

Because the theory developed in Section 3.3 applies only to Gaussian random processes having zero mean values, the DC bias of -0.229 was subtracted from the generated time histories for all subsequent numerical comparisons. Each desired extreme-value observation was defined as the largest positive value occurring in a specified time duration, T. The theoretical symmetry of the Gaussian distribution was used to assure uniformity of the extreme-value data. Thus, observations of extremes were taken equally from the positive and negative peaks, and the absolute values of the two data sets were combined into one total sample. The positive and negative data values were selected from different sections of the random time history, except for the data corresponding to T = 100 seconds. According to Gumbel (Reference 1, p. 110), the extreme largest and extreme smallest values are asymptotically independent for large samples. Thus, the combining of positive and negative extreme values from the same time-history section for T = 100 seconds is believed to be justified theoretically.

Verifying that observed extreme values from a generated time series behave mathematically as normal extremes was accomplished by comparing observed and theoretical cumulative probability distributions. The theoretical probability distribution, representing the largest individual in samples of size n taken from a standardized normal population, was tabulated by K. Pearson in Reference 2, (page 162). The basic relation between sample size (n) and characteristic largest value (u) for a specified distribution is given by Gumbel (Reference 1, page 82) by

$$F(u) = 1 - \frac{1}{n} \qquad (2)$$

The theoretical distribution of standardized normal extremes for 200 discrete samples, which corresponds to a standardized characteristic largest value of 2.5758, was selected for comparison with an observed extreme-value distribution obtained from the generated time series. From Equation (65), the value of T corresponding to this characteristic largest value is 2.161 seconds for the specified PSD. The observed probability distribution was obtained from 200 samples of largest values occurring in time intervals of 2.161 seconds from the generated random time history. The 200 values were modified to eliminate the DC bias and then ranked in increasing order. The cumulative probability assigned to the ith observed value was i/201.

The comparison between the observed and theoretical distributions is shown in Figure 4.3-2. The excellent agreement is corroborated by a Chi-square test of the hypothesis that the observed distribution is identical to the theoretical distribution of normal extremes. The Chi-square statistic, based on a division of the data into 20 cells, was 22.8. This value corresponds to a Chi-square cumulative probability of approximately 80 percent. Therefore, the observed extreme values obtained from the generated random time series may be considered as normal extremes from a discrete sample of size n, where n is determined from Equation (2) given the characteristic largest value.

Verifying the accuracy of the lognormal and extremal type III distributions to represent normal extremes in terms of the standardized characteristic largest value from Equation (65) was accomplished by comparing the approximating distributions with the distribution of observed extremes for two different time intervals, T. For T = 1.0 second, the standardized characteristic largest value corresponding to the specified PSD is 2.257. This is in the range where the lognormal distribution provides a nearly perfect representation of the actual distribution of normal extremes. Therefore, the required lognormal parameters may be obtained from Equations (20), (22), (66), and (67) of Section 3.3. The straight line in Figure 4.3-3 corresponds to $\gamma = 2.733$ and $\delta = 0.193$. The lognormal approximation is seen to provide a very good representation of the observed distribution plotted from 500 data points.

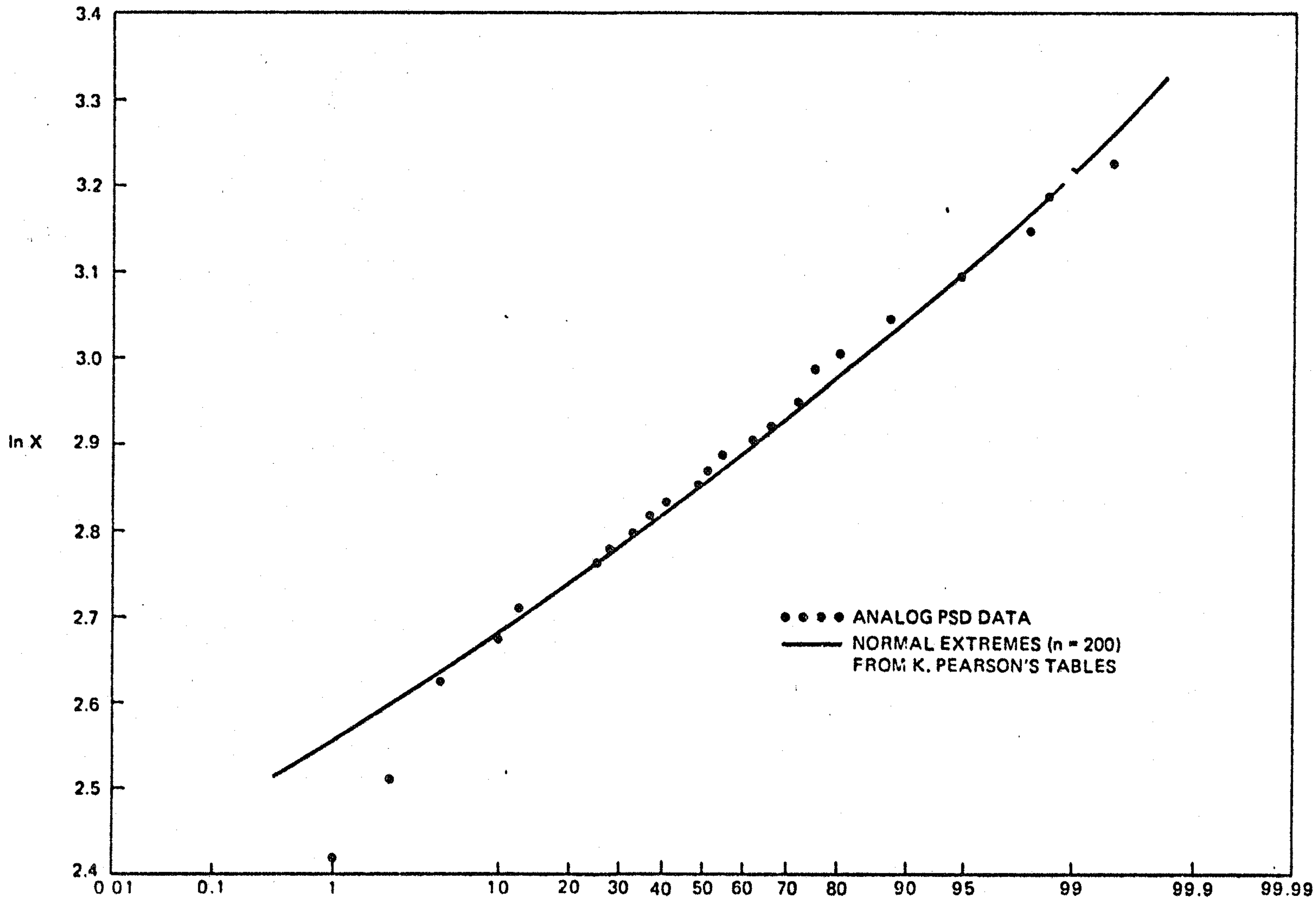

Figure 4.3-2. Distribution of Normal Extremes for T = 2.161 Seconds ($\hat{u}$ = 2.5758)

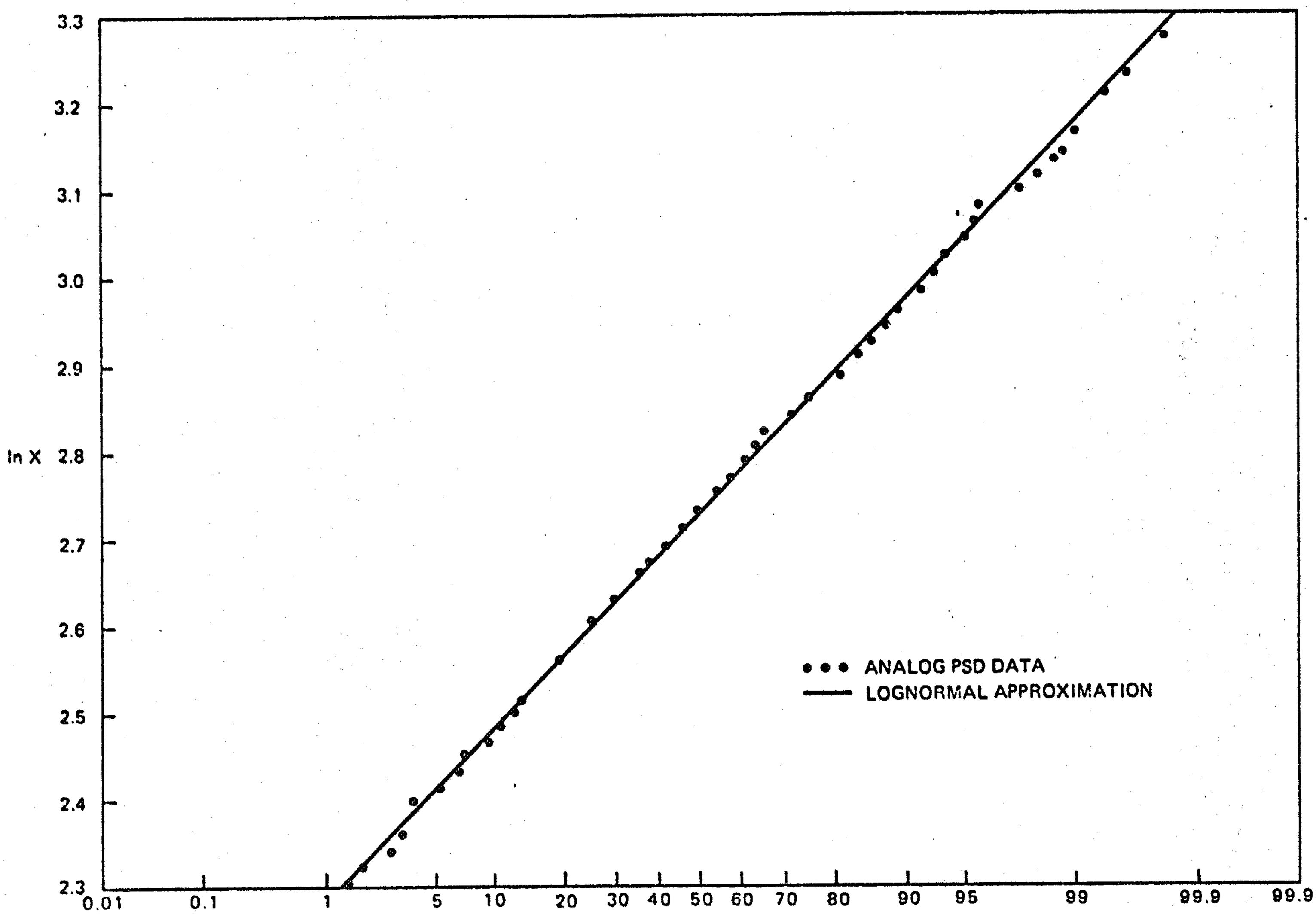

Figure 4.3-3. Distribution of Extremes from PSD for T = 1.0 Second ($\hat{u}$ = 2.257)

For T = 100 seconds, the standardized characteristic largest value corresponding to the specified PSD is 3.782. This is in the range where the extremal type III distribution provides a somewhat better representation of the actual distribution of normal extremes than the lognormal distribution. The required extremal type III parameter for the standardized variate, obtained from Equation (24) of Section 3.3, is $k = 17.604$. The corresponding distribution, obtained from Equation (25) of Section 3.3, is shown in Figure 4.3-4. The lognormal distribution also shown in this figure corresponds to $\gamma = 3.213$ and $\delta = 0.140$. The extremal type III is seen to provide an adequate representation of the observed distribution plotted from 118 data points. The standard deviation for the lognormal approximation, which determines the slope of the straight line, is seen to be too large. Since the median value is accurately determined for both approximating distributions, the characteristic largest value calculated from the Gaussian PSD by Equation (65) is seen to be the proper value for defining the desired extreme-value distribution.

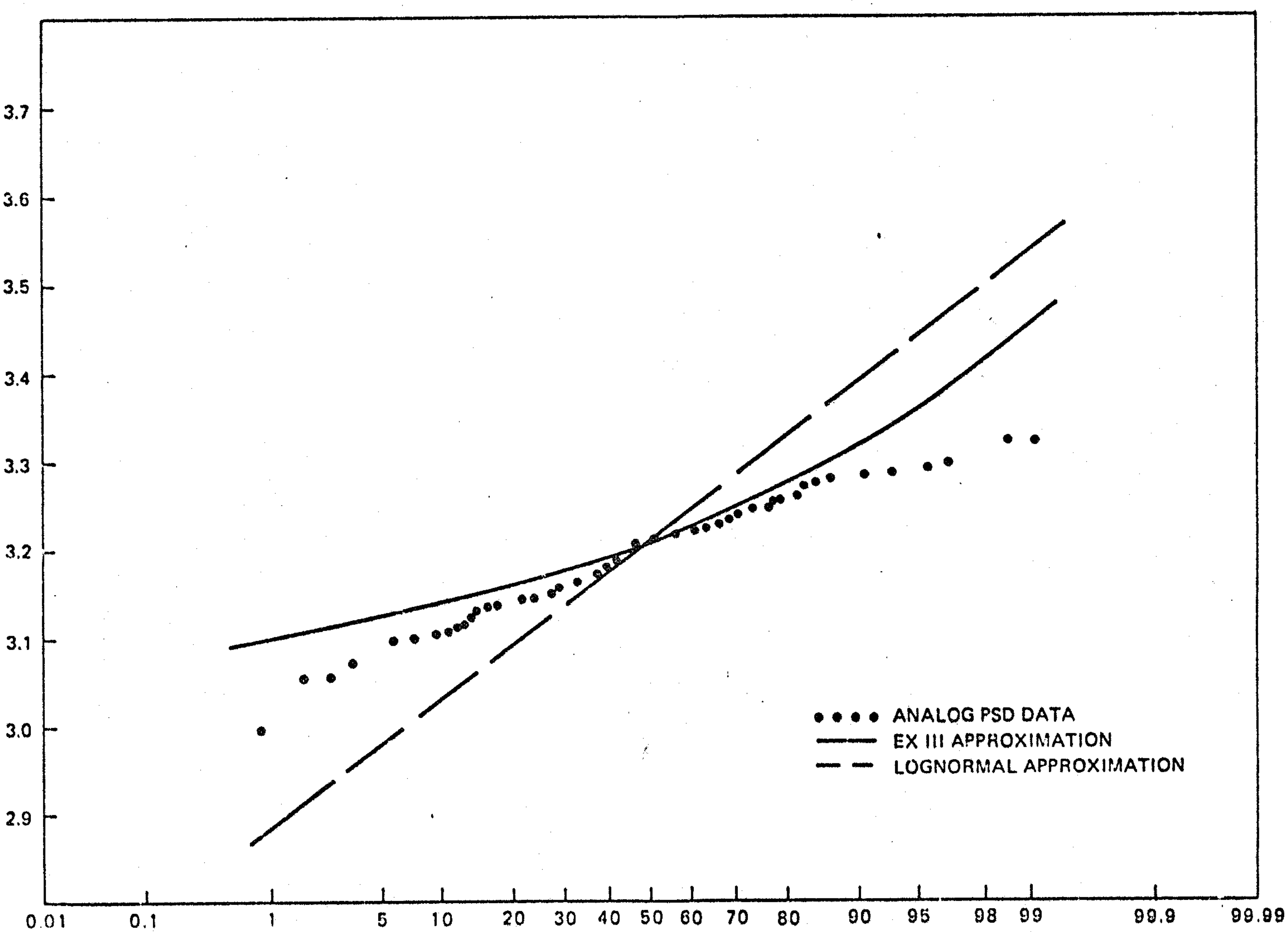

Figure 4.3-4. Distribution of Extremes From PSD for T = 100 Seconds ($\hat{u}$ = 3.782)

5.0 CONCLUSIONS AND RECOMENDATIONS

The methods for determining limit-load probability distributions from time-domain and frequency-domain dynamic loads analyses have been described and numerically demonstrated. The primary contribution is obtaining the extreme-value probability distributions from the Gaussian PSD of a frequency-domain analysis. Another contribution is obtaining conservative estimates of the limit-load probability distributions from a small number of Monte Carlo simulations.

Recommended areas for additional research include the following:

1. Improve the accuracy of the estimate of the parameter δ in Equation (67) for the lognormal approximation to the distribution of normal extremes when $\hat{u} > 3$.

2. Study the possibility of combining the Monte Carlo technique of Russian Roulette and the statistical estimation method to determine more efficient estimates of the limit-load parameters from Monte Carlo analyses.

3. Develop a method for obtaining the extreme-value distributions from a combination of time-domain and frequency-domain dynamic loads analyses.

4. Study methods of accounting for payload mass and stiffness variations in dynamic loads analyses.

REFERENCES

1. E. J. Gumbel, Statistics of Extremes, Columbia University Press, New York, 1958.

2. K. Pearson, Tables For Statisticians and Biometricians, Vol. II, Cambridge University Press, London, 1931.

3. D. H. Merchant, et al, "Study of Ground Handling Equipment Design Factors," Boeing document D180-18503-1, dated 3 February 1975.

4. D. H. Merchant and D. T. Sawdy, "Monte Carlo Dynamic Analysis for Lunar Module Landing Loads," Journal of Spacecraft and Rockets, Vol. 8, No.1, January 1971, pp. 48-55.

5. R. A. Fisher and L. H. C. Tippett, "Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample," Cambridge Philosophical Society Proceedings, 24:180, 1928, pp. 180-190.

6. A. Ang and M. Amin, "Safety Factors and Probability in Structural Design," Journal of the Structural Division, ASCE, Vol. 95, No. ST7, Proc. Paper 6667, July 1969, pp. 1389-1405.

7. J. A. Lovingood, "A Technique for Including the Effects of Vehicle Parameter Variations in Wind Response Studies," NASA TM X-53042, May 1, 1964.

8. A. H. Bowker and G. J. Lieberman, Engineering Statistics, Prentice-Hall, Englewood Cliffs, 1959.

9. E. Parzen, Stochastic Processes, Holden-Day, San Francisco, 1962.

REFERENCES (Cont'd)

10. R. S. Ryan and A. W. King, "The Influential Aspects of Atmospheric Disturbances on Space Vehicle Design Using Statistical Approaches for Analysis," NASA TM X-53565, January 13, 1967.

11. W. Holland, "SSME Side Loads Analysis for Flight Configuration," Revision A, NASA TM X-64841, September 1974.

12. D. N. Counter, "Splash Evaluations of SRB Designs," NASA TM X-64910, October 1974.

13. H. S. Epply, "Analytical Techniques for the Generalized Statistics Program (GESP)," Boeing document D2-118420-1, dated March 13, 1972

14. B. J. Bethune, "Generalized Statistics Program (GESP) Computer System Description," Boeing document D2-118353-2, dated August 24, 1972.

15. M. Shinozuka and C. M. Jan, "Digital Simulation of Random Processes and Its Application," Journal of Sound and Vibration, Vol. 25, No. 1, 1972, pp. 111-128.

16. H. Kahn, "Use of Different Monte Carlo Sampling Techniques," in Symposium on Monte Carlo Methods, H. A. Meyer, ed., Wiley, New York, 1956, pp. 146-190.

17. S. S. Wilks, Mathematical Statistics, Wiley, New York, 1962.

18. Y. K. Lin, Probabilistic Theory of Structural Dynamics, McGraw-Hill, New York, 1967.

REFERENCES (Cont'd)

19. G. A. Korn and T. M. Korn, Mathematical Handbook for Scientists and Engineers, McGraw-Hill, New York, 1961.

20. J. S. Andrews, "Effects of Transonic Buffeting on a Hammerhead-Shaped Payload," Journal of Spacecraft and Rockets, AIAA, Vol. 3, No. 8, August 1966, pp. 1285-1291.

21. B. E. Clingan, R. M. Gates, and J. S. Andrews, "Dynamic Loads During Boosted Flight," U.S. Air Force Systems Command, Aeronautical Systems Division, ASD-TDR-63-302, May 1963.

22. S. O. Rice, "Mathematical Analysis of Random Noise," reprinted in N. Wax (ed.), Selected Papers on Noise and Stochastic Processes, Dover, New York, 1954.

23. G. E. Daniels (ed.), "Terrestrial Environment (Climatic) Criteria Guidelines for Use in Space Vehicle Development, 1971 Revision," NASA TM X-64589, 10 May 1971.

24. A. Hald, Statistical Tables and Formulas, Wiley, New York, 1952.