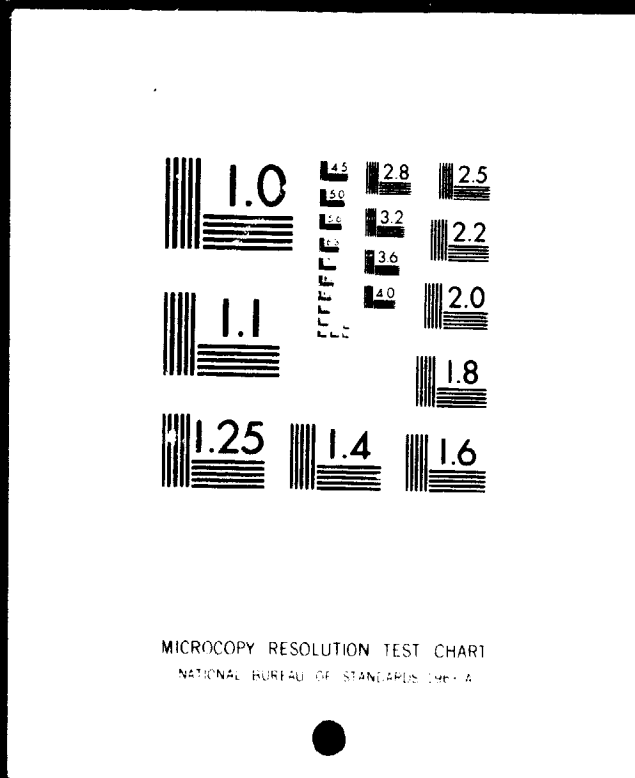


# 1 OF 3

# N82-14531

# UNCLAS



**(NASA-CR-165488) THE APPLICATION OF  
PROBABILISTIC DESIGN THEORY TO HIGH  
TEMPERATURE LOW CYCLE FATIGUE (Arizona  
Univ., Tucson.) 224 p HC A10/MF A01**

**N82-14531**

**Unclas  
CSCL 20K G3/39 08595**

**NASA CR-165488**

**THE APPLICATION OF  
PROBABILISTIC DESIGN THEORY  
TO HIGH TEMPERATURE  
LOW CYCLE FATIGUE**

**by**

**Paul H. Wirsching  
Department of Aerospace  
and Mechanical Engineering  
The University of Arizona  
Tucson, AZ 85721**

**Submitted to:  
NASA/Lewis Research Center  
2100 Brookpark Rd.  
Cleveland, OH 44135**



1. Report No. NASA CR-165488	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The Application of Probabilistic Design Theory to High Temperature, Low-Cycle Fatigue	5. Report Date November 1981	6. Performing Organization Code
	7. Author(s) P. H. Wirsching	8. Performing Organization Report No.
9. Performing Organization Name and Address University of Arizona Tucson, Arizona 85721	10. Work Unit No.	11. Contract or Grant No. NAG3-41
	12. Sponsoring Agency Name and Address National Aeronautics & Space Administration Washington DC 20546	13. Type of Report and Period Covered Topical Report
15. Supplementary Notes Grant Monitor, Robert L. Thompson Structures & Mechanical Technologies Division NASA Lewis Research Center Cleveland, Ohio 44135		

3. Abstract

Metal fatigue under stress and thermal cycling is a principal mode of failure in gas turbine engine hot section components such as turbine blades and disks and combustor liners. Designing for fatigue is subject to considerable uncertainty, e.g. scatter in cycles to failure, available fatigue test data and operating environment data, uncertainties in the models used to predict stresses, etc.

Application of probabilistic design theory to the high-temperature low-cycle fatigue problem is addressed. The focus is on the strain range partitioning model; however, the techniques presented are general and can be applied to other models as well.

Methods of analyzing fatigue test data for probabilistic design purposes are summarized. The general strain life as well as homo- and hetero-scadedastic models are considered.

A review of modern probabilistic design theory is provided. Examples are presented which illustrate application to reliability analysis of gas turbine engine components.

7. Key Words (Suggested by Author(s)) Reliability Probabilistic design Fatigue Statistics Miner's rule	18. Distribution Statement  Unclassified-unlimited		
9. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages	22. Price*

## ACKNOWLEDGEMENT

This study was supported under contract NAG-3-41 issued through the NASA/Lewis Research Center in Cleveland Ohio. The support and encouragement of the NASA staff is gratefully acknowledged. In particular the efforts of Robert L. Thompson as well as the suggestions of Marvin Hirschberg, Gary Halford, C.R. Ensign, and Jim Saltsman have been most helpful.

The work was conducted at the University of Arizona in the Department of Aerospace and Mechanical Engineering of which Professor C.F. Chen is Head. Graduate students who made substantial contributions to this study include Richard Chatfield, Colleen Kelly, Scott Martindale and Kelly Szeto.

Finally the efforts of Velma Chen, Margaret Wheeler and Mary Zornes in typing this manuscript are gratefully acknowledged.

## TABLE OF CONTENTS

1. INTRODUCTION	1-1
1.1 Preliminary Remarks	1-1
1.2 General Comments	1-2
1.3 An Illustration of the Conservatism in Conventional Design Procedures	1-5
1.4 Goals of this Study	1-10
1.5 What is Contained in this Report	1-11
1.6 List of General References for Reliability and Probabilistic Design	1-12
2. METHODS OF DATA ANALYSIS OF A SINGLE VARIABLE	
2.1 Preliminary Remarks	2-1
2.2 Mathematical Tools for Probability Estimates	2-1
2.3 Basic Statistics	2-5
2.4 Statistical Models Used in Fatigue	2-8
2.5 Probability Plots; Rectification of the Distribution Function	2-15
2.6 Comparison Test Using the Least Squares Method	2-19
2.7 Least Squares Analysis of the Three Parameter Weibull	2-25
2.8 Another Comparison Test Based on Fit to the Empirical Distribution Function	2-30
2.9 The Lognormal Assumption	2-32
2.10 The Use of the Tolerance Interval for Establishing a Design Value	2-33
2.11 The Scatter Factor Based on the Tolerance Interval	2-40
2.12 Use of the Prediction Interval in Establishing a Design Value	2-40
2.13 How to Establish a Design Value: Summary Comments	2-44
2.14 References for Chapter 2	2-48
3. ANALYSIS OF STRAIN-LIFE DATA: THE LEAST SQUARES LINE	
3.1 Preliminary Considerations	3-1
3.2 The Least Squares Line (Median Curve through Strain-Life Data)	3-1
3.3 Statistical Distribution of the Least Squares Estimators, $\hat{a}$ and $\hat{b}$	3-8
3.4 Confidence Intervals for the Parameters & Least Squares Line	3-9
3.5 References for Chapter 3	3-14

4.	STRAIN-LIFE DATA ANALYSIS: GENERAL CONSIDERATIONS	
4.1	General Considerations for Design	4-1
4.2	Definition of a Failure Trajectory	4-3
4.3	Linear Life Relationships	4-5
4.4	Uncertainties in the Linear Model	4-7
4.5	Lower $2\sigma$ or $3\sigma$ Limits	4-7
4.6	Implied Risk in Fatigue Design Curves of the ASME Boiler and Pressure Vessel Code	4-13
4.7	The Tolerance Interval Used to Establish a Design S-N Curve	4-16
4.8	A Consistent Method for Characterizing Linear Fatigue Data The Equivalent Prediction Interval (EPI)	4-22
4.9	References for Chapter 4	4-29
5.	STRAIN-LIFE DATA ANALYSIS: SOME ADVANCED TOPICS	
5.1	Preliminary Remarks	5-1
5.2	Linear SN Curve with Variable Scatter	5-1
5.3	Statistical Analysis of the General Strain-Life Relationships	5-8
5.4	A Model for the Distribution of Fatigue Strengths at High Cycle Lives	5-18
5.5	References for Chapter 5	5-26
6.	LIFE PREDICTION EQUATIONS USED FOR FATIGUE DESIGN	
6.1	Preliminary Remarks	6-1
6.2	The Concept of Damage	6-1
6.3	Some Expressions for Damage Used by Designers for Variable Amplitude Loading	6-3
6.4	Damage Using a Fracture Mechanics Approach to Fatigue	6-7
6.5	Damage Expressions for High Temperature Low Cycle Fatigue	6-8
6.6	Damage at Failure--Some General Remarks	6-9
6.7	Fatigue Design Relationships	6-11
6.8	Statistical Considerations of the Palmgren-Miner Fatigue Index	6-16
6.9	Statistical Considerations of Damage at Failure for Strainrange Partitioning	6-26
6.10	Statistical Distribution of $\Delta$	6-31
6.11	Summary Comments on Chapter 6	6-31
6.12	References for Chapter 6	6-34

7. A REVIEW OF MODERN PROBABILISTIC THEORY	
7.1 Introduction	7-1
7.2 Basic Definitions	7-3
7.3 The Full Distributional Approach	7-4
7.4 The Full Distributional Approach: Monte Carlo Methods	7-12
7.5 First Order Second Moment Methods	7-15
7.6 Extended Form of the Generalized Safety Index	7-27
7.7 Summary Comments on Target Reliability Levels and Technique	7-29
7.8 References for Chapter 7	7-30
8. DEMONSTRATION OF RELIABILITY ANALYSIS TO FATIGUE PROBLEMS	
8.1 Preliminary Remarks	8-1
8.2 Example 1; Strainrange Partitioning in a Reliability Format	8-1
8.3 Example 2; A Local Strain Analysis Problem in a Reliability Format	8-6
8.4 Example 3; A Demonstration of the Efficiency of the Rackwitz Algorithm	8-18
8.5 Summary Comments	8-20
8.6 References for Chapter 8	8-20
APPENDIX	
The Lognormal Distribution and Properties of Lognormal Variables	A-1

## CHAPTER 1: INTRODUCTION

### 1.1 Preliminary Remarks

Metal fatigue under stress and thermal cycling is expected to be a principal mode of failure in such engine components as turbine blades and disks, and combustor liners. But fatigue design factors are subject to considerable uncertainty. For example, enormous scatter is observed in cycles to failure fatigue test data with coefficients of variation (standard deviation/mean) ranging typically from 20 to 70%. Furthermore, scatter exists in operating environment data, and uncertainties are present in the models used to predict stresses. Therefore life predictions, which rely on the fatigue models and the data for such models, are also subject to uncertainty. It is suggested that the appropriate mathematical model to describe fatigue design factors is a probabilistic one rather than a deterministic one. Uncertainty in environments and in fatigue resistance imply uncertainty in fatigue life predictions. This uncertainty can be analyzed rationally only using probability theory.

A reliability approach to high temperature fatigue using probabilistic design theory has at least the promise if not the guarantee of producing better engineered design, i.e., components which are more safe, reliable, and cost effective, relative to a deterministic approach. Typically conventional design procedures tend to be conservative and produce inconsistent levels of risk in components of a



system. The payoff for an improved design criteria, e.g., for engine components, would be a savings in weight.

## 1.2 General Comments on Mechanical Reliability

The various uncertainties which exist in a mechanical design problem can be divided into statistical and nonstatistical (or professional) uncertainties as follows

1. Statistical Uncertainty (Data generally available or easy to obtain)
  - a. In basic material behavior, e.g., the scatter observed in basic S-N fatigue data.
  - b. In the estimates of the design parameters from data, e.g., the estimate of the fatigue strength coefficient is a random variable having a significantly large variance when only a small sample of fatigue data is available.
  - c. In the mechanical environment (load, high temperature, corrosion, etc. which affect fatigue life)
2. Nonstatistical Uncertainty (associated with assumptions made in the analysis)
  - a. In the theoretical model used to describe fatigue strength, e.g., linear damage accumulation rules.
  - b. Introduced by the procedure used to compute nominal stresses and temperatures in a component, i.e., assumptions made in the computer analysis routines.
  - c. Introduced by the models used to calculate fatigue strains at the critical points.

d. In strength due to size effects, processing and machining operations, assembly operations, etc.

The above illustrations suggest that we are dealing with a highly unpredictable physical phenomena. Some questions immediately come to mind.

1. Do present analysis procedures provide results which are too conservative? Are components overdesigned?

2. Are analyses telling us to provide too little material, or to recommend improper details to insure adequate protection for fatigue?

3. In general, are the procedures which are currently being used really producing high quality designs?

4. Can we quantify changes in risk (or probability of failure, or percent items which fail) due to modifications in the design?

There is a need to take a hard look at the high temperature, low cycle fatigue design process from a probabilistic viewpoint to attempt to answer these questions and thereby formulate design strategies which could reduce engine weight and provide real cost benefits.

Because of significant uncertainties in fatigue design factors, a probabilistic-statistical approach seems particularly relevant. All design factors are treated as random variables. Recent developments in probabilistic design theory can be utilized to predict distributions of fatigue life, or to establish design rules.

Commonly stated reasons for using a probabilistic approach to design include the following:

1. It is argued that reliability (or its complement, risk or probability of failure) is the most meaningful index of structural performance.

2. The effect on risk of making a design modification can be quantified.

3. A mechanism is provided for explicitly accounting for available statistical data on design factors, e.g., in cycles to failure fatigue data.

4. Factors which have nonstatistical uncertainty (due to assumptions made in analysis) can be treated as random variables. Their effect on design can be quantified and their relative importance assessed. This information can provide guidance for decision makers regarding which elements of the problem require further scrutiny. A rational basis for decisions regarding research programs is available.

5. All components can be designed to a balanced level of risk, thereby producing a more efficient system.

6. Probability based information on mechanical performance can be used to develop rational policies towards pricing, warranties, spare parts requirements, etc.

In summary, probabilistic design theory provides an improved engineering representation of reality.

### 1.3 An Illustration of the Conservatism in Conventional Design Procedures

Design procedures and codes traditionally rely on a factor of safety (applied to material strength) to insure acceptable reliability. Such procedures have performed well, but it is generally thought that requirements are typically overly conservative and that they produce designs having an inconsistent level of risk among the components. Attempts to estimate materials savings which could be realized by probability based design criteria have produced figures of about 10%. Thus, application of reliability methods to vehicle power plant design promises a payoff in weight savings at the same reliability levels.

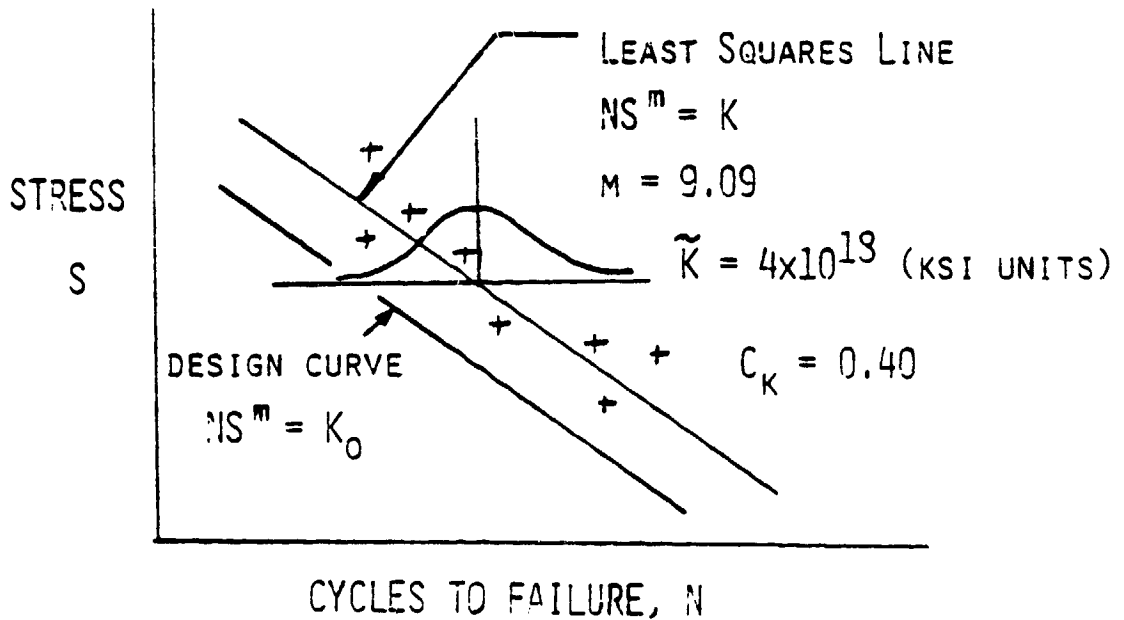
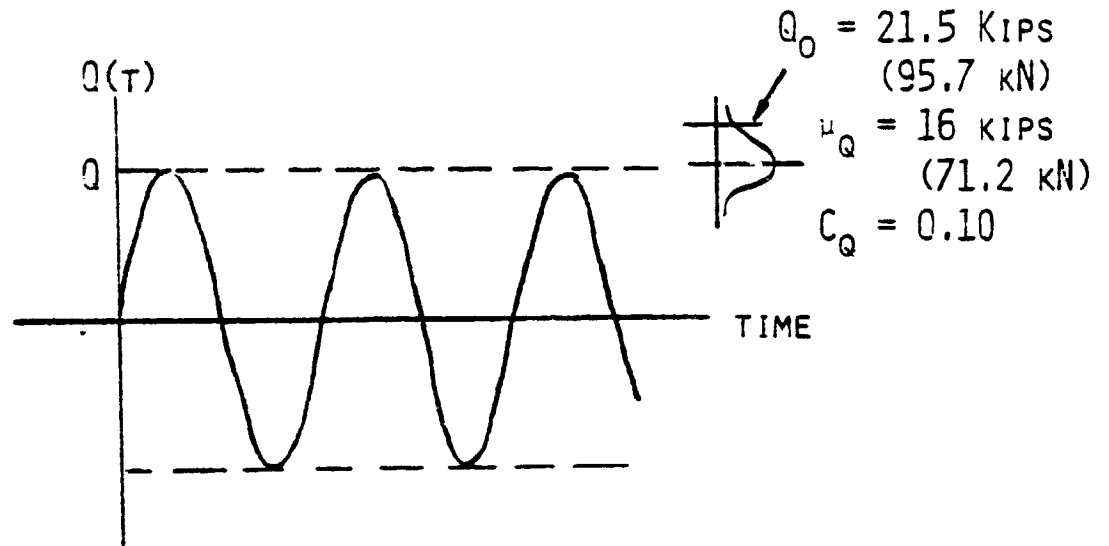
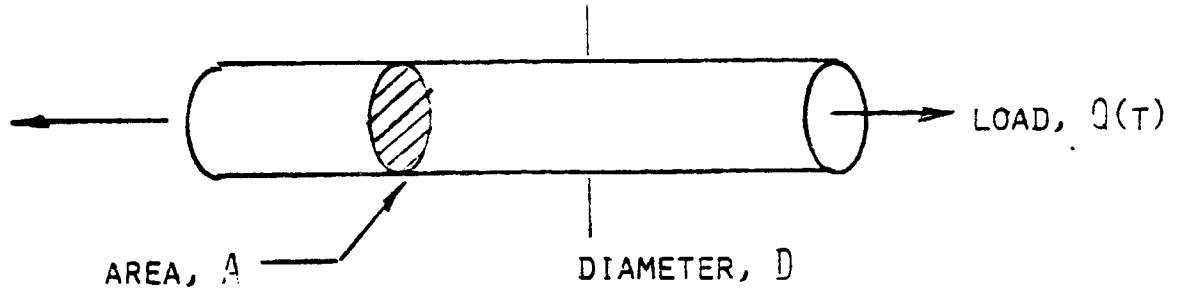
As an example, consider the simple design problem illustrated in Fig. 1.1. The bar is subjected to a harmonic axial load  $Q(t)$ . Fatigue is considered the principal mode of failure. The service life is given as  $N_S = 500,000$  cycles.

Because of analysis procedures, the amplitude  $Q$  is considered to be random. Statistics on  $Q$  are available. The mean value is  $\mu_Q = 16$  kips (71.2 kN) and the coefficient of variation  $C_Q = 0.10$ . The distribution of  $Q$  is lognormal. As a design value, a point on the safe (upper) side of the distribution  $Q_0$  is chosen as the median plus three standard deviations on a log basis. The calculations are summarized in Table 1.1. See also Appendix I.

Data is available on the fatigue behavior of the material as shown in Fig. 1.1. A basic linear model is assumed (See Chapt. 3 and 4) on a log basis, and the least squares curve established. The model for

Fig. 1.1

A DEMONSTRATION OF HOW CONVENTIONAL DESIGN  
CAN PRODUCE CONSERVATIVE RESULTS



cycles to failure is  $N = KS^m$  where the median and coefficient of variation of  $K$ , denoted as  $\tilde{K}$  and  $C_K$  are given in the figure.

But the designer chooses a curve on the safe (lower) side of the distribution as the least squares line minus three standard deviations on a log basis. This curve is defined by  $NS^m = K_0$ .

The conventional approach commonly will require a design life  $N_0$  higher than  $N_S$ . A factor of 2 is used here. The design requirement using conventional methods is given in Table 1.1.

Because statistics are available, the risk inherent in this design can be estimated as shown in Table 1.1. Definitions and details are provided in Chapter 7 and Appendix I. The estimated risk of  $1.92 \times 10^{-6}$  is generally considered to be overly conservative for mechanical and structural components where the consequences of failure are not disastrous. A target value of the probability of failure  $p_f = 10^{-3}$  is accepted as being more reasonable.

Using a value of  $p_f = 10^{-3}$  as the basic requirement, a probability based design produces a smaller component. The estimated weight savings of 15% by using a reliability approach compares with an often quoted estimate of 10% for the level of excess material required by present design codes.

In summary, it was intended that this example illustrate the fact that excessive levels of reliability can be produced by a "pile-up" of safety factors applied to each component of the design algorithm. It was not the intention to imply that all deterministic design codes and procedures are overly conservative.

TABLE 1.1

Calculations: Conventional vs. Probabilistic Design

CONVENTIONAL (DETERMINISTIC) DESIGN

- Establish Design Values (see also Appendix I for mathematics of lognormal variates)

$$\ln Q_o = \ln \tilde{Q} + 3\sigma_{\ln Q}$$

$$\tilde{Q} = \mu_Q / \sqrt{1 + C_Q^2}$$

$$\sigma_{\ln Q} = \sqrt{\ln(1 + C_Q^2)}$$

$$Q_o = 21.5 \text{ kip} \\ = 95.7 \text{ kN}$$

$$C_Q = 0.17$$

$$\ln K_o = \ln \tilde{K} - 3\sigma_{\ln K}$$

$$\sigma_{\ln K} = \sqrt{\ln(1 + C_K^2)}$$

$$K_o = 1.25 \times 10^{18} \text{ (ksi units)} \\ = 5.20 \times 10^{25} \text{ (MPa units)}$$

$$C_K = 0.40$$

- Design Life

$$N_o = 2N_s = 100,000 \text{ cycles}$$

- Design Stress

$$S_o = Q_o/A = 21.5/A \text{ ksi}$$

- Design Equation

$$\text{Cycles to failure, } N = K_o S_o^{-m} > N_o$$

- Solution

$$\text{Area, } A \geq 0.78 \text{ in}^2 \text{ (5.03 cm}^2\text{)}$$

$$\text{Diameter } D \geq 1.00 \text{ in (2.54 cm)}$$

RISK IMPLIED BY THIS DESIGN

- Safety Index

$$\beta = \frac{\ln(\tilde{N}/N_s)}{\sigma_{\ln N}}$$

$$\tilde{N} = \tilde{K}/(\tilde{Q}_r)^m$$

$$\sigma_{\ln N} = \sqrt{\ln[(1 + C_K^2)(1 + C_Q^2)]^m}$$

$$\beta = 4.62$$

- Probability of Failure

$$p_f = \Phi(-\beta)$$

$$p_f = 1.92 \times 10^{-6}$$

... TABLE 1.1 Continued

REDESIGN USING RELIABILITY METHOD

● Target  $p_f = 10^{-3}$  (reasonable level)

Target  $\delta = 3.09$

● Solution (using above equations)

Area,  $A = 0.66 \text{ in}^2$

Diameter  $D = 0.92 \text{ in}$

PERCENT WEIGHT SAVINGS BY USING RELIABILITY METHOD = 15%



#### 1.4 Goals of This Study

The ultimate goal of this study is to recommend methods of reliability analysis for the development of probability based design criteria for fatigue design in general and high temperature fatigue in particular. It is intended that this study provide practical information for engine designers.

Specific goals include the following.

1. To develop methods of providing statistical summaries of data for design purposes, i.e.,
  - a. Design values and/or design curves
  - b. Appropriate statistics which are required for a reliability analysis.
2. To recommend methods of probabilistic design for a complete reliability analysis of a component subjected to high temperature, i.e., the problem where several random design factors appear in an equation and where it is required to evaluate the risk.
3. To present the results of this study in a format for easy implementation for engine designers .

## 1.5 What is Contained in This Report

This report summarizes the results of the first year of a comprehensive study on fatigue/creep reliability supported by the NASA/Lewis Research Center. It is intended that this report provide practical information useful for designers.

Chapter 2 summarizes methods available for statistical analysis of data on a single variable. Various schemes for establishing a design value are presented.

Chapters 3, 4 and 5 summarize procedures for providing statistical summaries of S-N fatigue data. Linear model analysis is presented and applied to both homo- and heteroscedastic data. Preliminary considerations of data analysis using the general strain-life model are included.

Attempts are made in Chapter 6 to present a coordinated overview of the three basic approaches to formulating a fatigue design equation. Performance of Miner's rule and strainrange partitioning are described by statistical summaries.

Chapter 7 provides a summary of available reliability methods for fatigue design. Two examples are presented in Chapter 8. One is a strainrange partitioning example. The other is of a local strain analysis model.

This report is considered preliminary. Work is continuing in all areas, and future reports will provide more complete information.

## 1.6 List of General References for Reliability and Probabilistic Design

The following references include some of the more important sources of information in elementary applied probability theory and statistics, probabilistic design and fatigue reliability. The list is by no means exhaustive and does not include many excellent texts in specialized areas of reliability engineering, quality control, random process theory, etc.

..... for basic probability and statistics

- Ang, A. H.-S., and Tang, W., Probability Concepts in Engineering Planning and Design, Wiley, 1975. Excellent text for engineering problems and statistics, but oriented to civil engineering.
- Benjamin, J. R., and Cornell, C. A., Probability, Statistics and Decision for Civil Engineers, McGraw-Hill, 1970. An excellent reference for engineers of all types, uses modern techniques with many examples, probably one of the best general references available for engineers to date.
- Bowker, A. H., and Lieberman, G. J., Engineering Statistics, Prentice-Hall, 1972. An excellent reference, strong on testing hypotheses and quality control.
- Hines, W. H., and Montgomery, D. C., Probability and Statistics in Engineering and Management Science, 2nd Ed., Wiley, 1980
- Lipson, C., and Sheth, N. J., Statistical Design and Analysis of Engineering Experiments, McGraw-Hill, 1973. Treats problems from mechanical engineering design theory.
- Meyer, P. L., Introductory Probability and Statistical Applications, Addison-Wesley, 1970. An excellent introduction to modern probability theory and mathematical statistics.
- Mood, A. M., and Graybill, F. A., Introduction to the Theory of Statistics, McGraw-Hill, 1963. Intermediate statistics, and a little difficult for a beginner.
- Natrella, M. G., Experimental Statistics, Dept. of Commerce, 1963. An excellent statistics "cookbook" full of examples on how-to-do-it in testing hypotheses and many other tests.

- ..... for probabilistic design
- Ang, A. H.-S., and Cornell, C.A., "Reliability Basis for Structural Safety," Journal of the Structural Division, ASCE, Sept. 1974.
- Cornell, C.A., "A Probability Based Structural Code," ACI Journal, Dec. 1969.
- Dialog, Second International Workshop on Code Formats, Mexico City, Jan. 1976, Danmarks Ingeniorakademi, Building 373, 2800 Lyngby, Denmark.
- Ellingwood, B., et al., "Development of a Probability Based Load Criterion for American National Standard A58," NBS Special Publication 577, June 1980.
- First Order Reliability Concepts for Design Codes, Bulletin D'Information 112, Comite Europeen du Beton. Munich, July 1976.
- Freudenthal, A.M., Garrelts, J.M., and Shinozuka, M., "The Analysis of Structural Safety", Journal of the Structural Division, ASCE, Vol.92, No ST1, Feb. 1966.
- Galambos, T.V., Ravindra, M.K., (and others),... a series of eight papers on Load and Resistance Factor Design (LRF) in the Journal of the Structural Division, Vol. 104, No. ST9, Sept. 1978.
- Hasofer, A.M., and Lind, N.C., "An Exact and Invariant First Order Reliability Format," Journal of the Engineering Mechanics Division, ASCE, February 1974.
- Haugen, E.B., Probabilistic Mechanical Design, Wiley, 1980.
- Haugen, E.B., and Wirsching, P.H., "Probabilistic Design," a five part series in Machine Design, starting April 17, 1975.
- Kececioglu, D.B., and Cormier, D., "Designing a Specified Reliability into a Component," Proceedings of the Third Reliability and Maintainability Conference, Washington, D.C., 1964.
- Probabilistic Mechanics and Structural Reliability, (ed. A.H.-S. Ang and M. Shinozuka), ASCE Tucson, January 1979.
- Probabilistic Methods in Structural Engineering (ed. M. Shinozuka and J.T.P. Yao), ASCE, 1981.
- Rationalisation of Safety and Serviceability Factors in Structural Codes, Report 63, CIRIA, Construction Industry Research and Information Association, 6 Storey's Gate, London SW1P 3AU.

..... for fatigue reliability (many other references cited in text)

Ang, A. H.-S., and Munse, W.H., "Practical Reliability Basis for Structural Fatigue," ASCE National Structural Engineering Conference, Preprint 2494, April 1975.

Ellingwood, B., "Probabilistic Assessment of Low Cycle Fatigue Behavior of Structural Welds," Journal of Pressure Vessel Technology, ASME, Paper 75-PVP-29, February 1976.

Fatigue Reliability: A State of the Art Review, a four part series

1. "Fatigue Reliability: Introduction"
  2. "Fatigue Reliability: Quality Assurance and Maintainability"
  3. "Fatigue Reliability: Variable Amplitude Loading"
  4. "Fatigue Reliability: Development of Criteria for Design"
- Journal of the Structural Division, ASCE, Vol. 108, No. ST1, Jan. 1982.

Fong, J.T., "Uncertainties in Fatigue Life Prediction and a Rational Definition of Safety Factors", Nuclear Engineering and Design, Vol. 51, 1978.

Freudenthal, A.M., and Gumbel, E.J., "On the Statistical Interpretation of Fatigue Tests," Proceedings of the Royal Society of London, Series A, Vol. 216, 1953, pp. 309-322.

Freudenthal, A.M. and Gumbel, E.J., "Physical and Statistical Aspects of Fatigue," Advances in Applied Mechanics, Vol. 4, 1956.

Kececioglu, D.B., and Chester, L.B., "Combined Axial Stress Fatigue Reliability for AISI 4130 and 4340 Steels," ASME Paper 75-WA/DE-117, 1975.

Little, R.E., and Jebe, E.H., Statistical Design of Fatigue Experiments, Applied Science, 1975.

Whittaker, I.C., Besuner, "A Reliability Analysis Approach to Fatigue Life Variability of Aircraft Structures," Wright-Patterson Air Force Base Technical Report AFML-TR-69-65, April 1969.

Wirsching, P.H., "Fatigue Reliability of Welded Joints in Offshore Structures," Proceedings of the Offshore Technology Conference OTC 3380, 1979, also published in the International Journal of Fatigue, April 1980.

Yang, J.N., "Statistical Approach to Fatigue and Fracture Including Maintenance Procedures" Fracture Mechanics, U. of Virginia Press, 1980.

Yao, J.T.P., "Fatigue Reliability and Design", Journal of the Structural Division, ASCE, Vol. 100, No. ST9, September 1974.

## Chapter 2 METHODS OF DATA ANALYSIS OF A SINGLE VARIABLE

### 2.1 Preliminary Remarks

Given a random set of observations on a design factor, it is necessary to interpret the data in order to make a design decision. For example the random sample of Table 2.1 is cycles to failure fatigue data for tests at a single stress level. In order to insure an acceptably low level of risk (of probability of failure) it may be necessary to specify either a "design value" on the safe (lower) side of the distribution or the statistics of the variable, depending upon the design strategy used.

This chapter summarizes various statistical tools which can be employed to provide designers with information that they need to make decisions on a single variable. Later chapters deal with analysis of SN data. The focus of this report is on fatigue, but techniques described herein have wide appreciation.

### 2.2 Mathematical Tools for Probability Estimates

Because observed cycles to failure has significant scatter, it is suggested that cycles to failure, denoted as  $N$ , be treated as a random variable. Therefore a probability density function (pdf)  $f_N(n)$  is defined such that

$$P(n_A \leq N \leq n_B) = \int_{n_A}^{n_B} f_N(n) dn \quad (2.1)$$

where  $P(\cdot)$  denotes the probability of the event in parentheses.

The cumulative distribution function (cdf)  $F_N(n)$  is defined as

$$F_N(n) = P(N \leq n) = \int_{-\infty}^n f_N(x) dx \quad (2.2)$$

where  $n$  denotes a specific value of the random variable  $N$ .

Table 2.1

Cycles to Failure of Specimens Tested at the Same Stress Level [Ref: Evans (13)]

$N_i$   
Cycles to Failure ( $10^3$  cycles)

15.4

22.2

17.3

23.6

14.4

12.3

16.5

25.7

17.5

27.0

20.5

21.3

14.0

27.0

23.0

23.6

16.0

14.7

10.3

16.0

13.0

31.0

---

Sample Size, $n$	=	22
Sample Mean, $\bar{N}$	=	19.2 ( $10^3$ cycles)
Sample Standard Deviation, $S_N$	=	5.59 ( $10^3$ cycles)
Sample Coefficient of Variation	=	29.1%

An example of a pdf and corresponding cdf is given in Figure 2.1. Note that all of the information about probabilities of  $N$  is contained in the pdf and cdf.

The mean  $\mu_N$  and standard deviation  $\sigma_N$  are defined as,

$$\mu_N = \int_{-\infty}^{\infty} n f_N(n) dn \quad (2.3)$$

$$\sigma_N = \sqrt{\int_{-\infty}^{\infty} (n - \mu_N)^2 f_N(n) dn} \quad (2.4)$$

The mean is a measure of the central tendency and the standard deviation is an index of the degree of variability. Another measure of central tendency is the median,  $\tilde{N}$ , defined as the 50% point of the distribution

$$F_N(\tilde{N}) = P(N \leq \tilde{N}) = 0.50 \quad (2.5)$$

The coefficient of variation  $C_N$  is defined as

$$C_N = \sigma_N / \mu_N \quad (2.6)$$

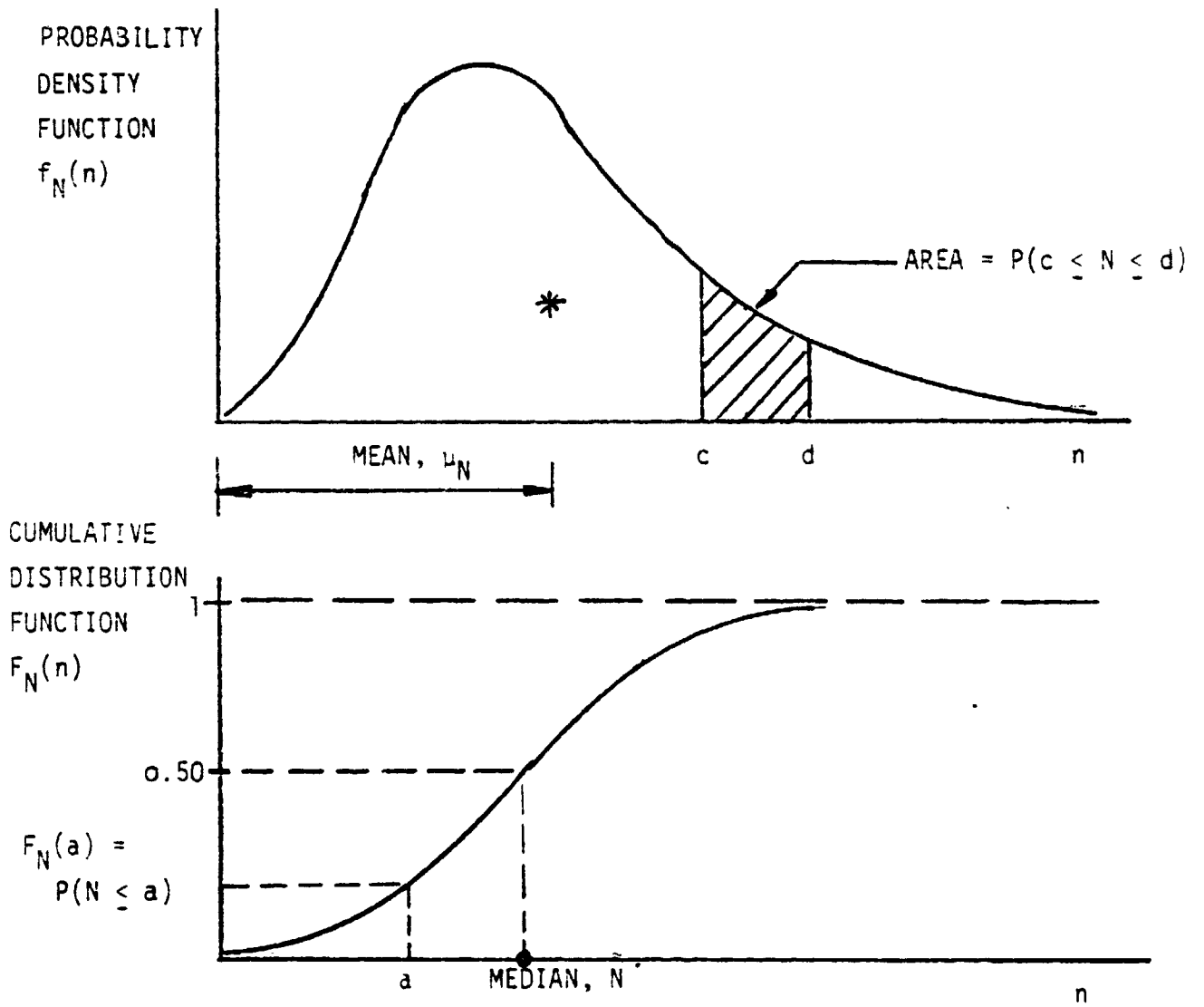
This term, commonly used in probabilistic design literature, is a non-dimensional measure of variability.

Ultimately we wish to make statements regarding risk (or probability of failure) which we can do if we have  $F_N(n)$  and/or  $f_N(n)$ . But at this stage all we have is the random sample  $N_i$ ,  $i = 1, k$  of Table 2.1. The remainder of this chapter deals with the problem of statistics, i.e. how we take the random sample and make probability statements for design purposes. The example used herein is of cycles to failure data, but the analysis presented applies to any random variable.



Figure 2.1

An illustration of the Probability Density Function (pdf) and Corresponding Cumulative Distribution Function (cdf)



### 2.3 Basic Statistics

The sample mean  $\bar{N}$  is a measure of the central tendency

$$\bar{N} = \frac{1}{k} \sum_{i=1}^k N_i \quad (2.6)$$

where  $k$  = sample size.  $\bar{N}$  is an estimate of  $\mu_N$ .

The sample standard deviation  $s_N$  is a measure of the dispersion or scatter in the data

$$s_N = \left[ \frac{1}{k-1} \sum_{i=1}^k (N_i - \bar{N})^2 \right]^{1/2} \quad (2.7)$$

$s_N$  is an estimate of  $\sigma_N$ .

For the data of Table 2.1,  $\bar{N} = 19.2$  and  $s_N = 5.59$ . The sample coefficient of variation is  $C_N = s_N/\bar{N} = 0.291$ .

This value of  $C_N$  is typical of fatigue data at relatively low lives as illustrated in Table 2.2. The values listed give evidence of the relatively large scatter which exists in fatigue strength data. For example, yield and tensile strengths for a wide variety of materials are typically less than 10% and usually about 5%.

The data of Table 2.2 also suggest that scatter is greater at lower stress levels and longer lives. This seems to be a general rule for smooth specimen data.

The empirical cumulative distribution function, an estimate of  $F_N(n)$ , can be established as demonstrated in Table 2.3 (the data from Table 2.1). A smooth curve through these points is an estimate of  $F_N(n)$ . Examples of how the empirical cdf is used are given in Section 2.5.

Table 2.2

Typical Values of the Coefficient of Variation  $C_N$  in Cycles to Failure Data Observed by Various Investigators at Different Stress Levels

	C(%)
Ang and Munse analysis from the 27 structural details in the AISC code (1); mostly welded joints *	52
Whittaker's suggested values based on extensive study of fatigue data of metals used in aircraft application (2,3); for lives less than $4 \times 10^5$ cycles. Steel, UTS < 240 ksi	36
Steel, UTS > 240 ksi	48
Aluminum alloys	22
Titanium alloys	36
Sinclair and Dolan rotating beam tests on specimens of 7075-T6 aluminum alloy (4); at different stress levels Stress amplitude, $S=50$ ksi Avg. life $\approx 5 \times 10^4$ cycles	28
Stress amplitude, $S=30$ ksi Avg. life $\approx 10^7$ cycles	128
Saltsman and Halford on high temperature low cycle fatigue tests of AISI 316(5)	32
Wirsching analysis of the amalgamated data which provides the basis for the AWS-X design curve on tubular welded joints (6); elastic range only (lives mostly above $10^4$ cycles)	136

\* Numbers in parentheses refer to references at the end of each chapter.

Table 2.3

The Empirical Cumulative Distribution Function for the Data of Table 2.1  
(n = 22)

Order Number, $i$	Ordered Data, $N_{(i)}$	Empirical cdf*, $F_i = \frac{i - 1/2}{n}$
1	10.3	.023
2	12.3	.068
3	13.0	.114
4	14.0	.159
5	14.4	.205
6	14.7	.250
7	15.4	.295
8	16.0	.341
9	16.0	.386
10	16.5	.432
11	17.3	.477
12	17.5	.523
13	20.5	.568
14	21.3	.614
15	22.2	.659
16	23.0	.705
17	23.6	.750
18	23.6	.795
19	25.7	.841
20	27.0	.886
21	27.0	.932
22	31.0	.977

\*Many other forms which give similar results have been suggested, but this one (sometimes called the "Hazen formula") seems to generally perform well and is widely used by engineers.

## 2.4 Statistical Models Used in Fatigue

In order to make design decisions based on a set of observations of a design factor, it is necessary to describe the distribution of that factor. In that regard, statistical models are usually employed. The random variable  $N$  denoting cycles to failure is often described with a two parameter Weibull or lognormal model. A summary of these models as well as the normal and the three parameter Weibull is provided in Tables 2.4 through 2.7. The normal is included in this discussion primarily for reference. The three parameter Weibull has been used as a model for fatigue data; practical and analytical difficulties associated with the use of this distribution are summarized later.

The lognormal and two parameter Weibull are most commonly used to describe  $N$ . Use of the lognormal distribution has been based primarily on arguments of mathematical expediency. However physical arguments favor the Weibull for most material strength variables because it is an asymptotic distribution of minima of a sample (7). If failure of a structural element is precipitated by failure of the first of a large number of sub-elements, then the Weibull is likely a "good" model.

Moreover, it has been pointed out by Gumbel (8) and demonstrated in Figure 2.2a that the hazard function decreases for large  $N$  which violates our physical understanding of progressive deterioration resulting from the fatigue process.

Also note from Figure 2.2b that for the same statistics, the Weibull gives larger probabilities in the left tail. Use of the Weibull should produce conservative designs relative to the lognormal.

Nevertheless, the lognormal is often much easier to use. Methods of linear model analysis commonly used on SN data rely on a lognormal assumption for  $N$ . Probabilistic design procedures in a lognormal format have been developed. Furthermore, this author has found, more often than not, the lognormal provides a better fit than the Weibull to real fatigue data.

TABLE 2.4

## A SUMMARY OF THE NORMAL DISTRIBUTION

Probability Density Function

$$f_N(n) = \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{n - \mu_N}{\sigma_N} \right)^2 \right]$$

Statistical parameters

 $\mu_N$  = mean value of N $\sigma_N$  = standard deviation of N

Distribution Function; Probability Calculations

$$F_N(n) = P(N \leq n_0) = \Phi \left( \frac{n_0 - \mu_N}{\sigma_N} \right)$$

where  $\Phi(\cdot)$  = standard normal distribution function  
(definition, e.g. Reference 14)

 $n_0$  = any specific value of N

How the Parameters are Estimated from Data

Data:  $\underline{N} = (N_1, N_2 \dots N_k)$ Compute the sample mean  $\bar{N}$  as an estimate of  $\mu_N$ Compute the sample standard deviation  $s_N$  as the estimate of  $\sigma_N$

Table 2.5

A SUMMARY OF THE LOGNORMAL DISTRIBUTION\*

N is said to have a lognormal distribution if  $X = \ln N$  (or  $X = \log_{10} N$ ) has a normal distribution.

•Probability Density Function (base e)

$$f_N(n) = \frac{1}{\sqrt{2\pi}\sigma_x n} \exp \left[ -\frac{(\ln n - \mu_x)^2}{2\sigma_x^2} \right]$$

Statistical parameters

$\mu_x$  = mean value of  $X$

$\sigma_x$  = standard deviation of  $X$

•Distribution Function Probability Calculations

$$F_N(n_0) = P(N \leq n_0) = \Phi \left( \frac{\ln n_0 - \mu_x}{\sigma_x} \right)$$

or  $\Phi \left( \frac{\log_{10} n_0 - \mu_x}{\sigma_x} \right)$

where  $\Phi(\cdot)$  = standard normal distribution function

\*See Appendix 1 for a comprehensive summary of the lognormal distribution and its application in probabilistic design.

Table 2.5 (continued)

Relationships between Parameters and Moments (i.e., mean and standard deviation)

Base e	Base 10
$\mu_X = \ln \mu_N - \frac{1}{2} \ln(1 + C_N^2)$ $= \ln \tilde{N} \quad (\text{See note below})$	$\mu_X = \log_{10} \mu_N - \frac{1}{2} \log_{10}(1 + C_N^2)$
$\sigma_X^2 = \ln(1 + C_N^2)$	$\sigma_X^2 = 0.434 \log_{10}(1 + C_N^2)$
$\mu_N = \exp\left[\mu_X + \frac{1}{2}\sigma_X^2\right]$	$\mu_N = 10^{\left\{\mu_X + 1/2(\sigma_X^2/.434)\right\}}$
$\sigma_N^2 = \mu_N^2(e^{\sigma_X^2} - 1)$	$C_N = \sqrt{10^{\left(\frac{\sigma_X^2}{.434}\right)} - 1}$
$C_N = \sqrt{\exp(\sigma_X^2) - 1}$	

How the Parameters are Estimated from Data

Data  $N = (N_1, N_2, \dots, N_k)$

Let  $X_i = \ln N_i$  (or  $\log_{10} N_i$ )

Let sample mean  $\bar{X}$  be the estimate of  $\mu_X$

Let sample standard deviation  $s_X$  be the estimate of  $\sigma_X$

Note: A useful relationship between the median  $\tilde{N}$  and the mean  $\mu_N$  can be derived as

$$\tilde{N} = \mu_N / \sqrt{1 + C_N^2}$$



Table 2.6

A SUMMARY OF THE TWO PARAMETER WEIBULL DISTRIBUTION

Distribution Function:  $F_N(n) = P(N \leq n)$

$$F_N(n) = 1 - \exp \left[ - \left( \frac{n}{\beta} \right)^\alpha \right] \quad n \geq 0$$

Statistical Parameters

$\alpha$  = shape parameter

$\beta$  = scale parameter

Probability Density Function

$$f_N(n) = (\alpha/\beta)(n/\beta)^{\alpha-1} \exp \left[ - \left( \frac{n}{\beta} \right)^\alpha \right]$$

Relationship between Parameters and Moments

$$\mu_N = \beta \Gamma(1/\alpha + 1)$$

$$C_N = \left[ \frac{\Gamma(2/\alpha + 1)}{\Gamma^2(1/\alpha + 1)} - 1 \right]^{1/2}$$

where  $\Gamma(\cdot)$  = the gamma function

A useful approximation  $\alpha \approx C_N^{-1.08}$  for  $0.02 \leq C_N \leq 2.00$

How to Estimate the Parameters (Method of Moment Estimators)

Data  $\underline{N} = (N_1, N_2, \dots, N_k)$

Compute sample mean  $\bar{N}$  and sample standard deviation of  $s_N$ .

Let  $\tilde{\alpha} = (s_N/\bar{N})^{-1.08}$  be an estimate of  $\alpha$  and  $\tilde{\beta} = \bar{N}/\Gamma(1/\tilde{\alpha} + 1)$  be an estimate of  $\beta$

(Method of Moments)

Table 2.7

A SUMMARY OF THE THREE PARAMETER WEIBULL

Distribution Function:  $F_N(n) = P(N \leq n)$

$$F_N(n) = 1 - \exp \left[ - \left( \frac{n - \gamma}{\beta} \right)^\alpha \right] \quad n \geq \gamma$$

Probability Density Function

$$f_N(n) = \frac{\alpha}{\beta} \left( \frac{n - \gamma}{\beta} \right)^{\alpha - 1} \exp \left[ - \left( \frac{n - \gamma}{\beta} \right)^\alpha \right]$$

Relationship Between Parameters and Moments

$$\mu_N = \gamma + \beta \Gamma(1 + 1/\alpha)$$

$$\sigma_N^2 = \beta^2 [\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)]$$

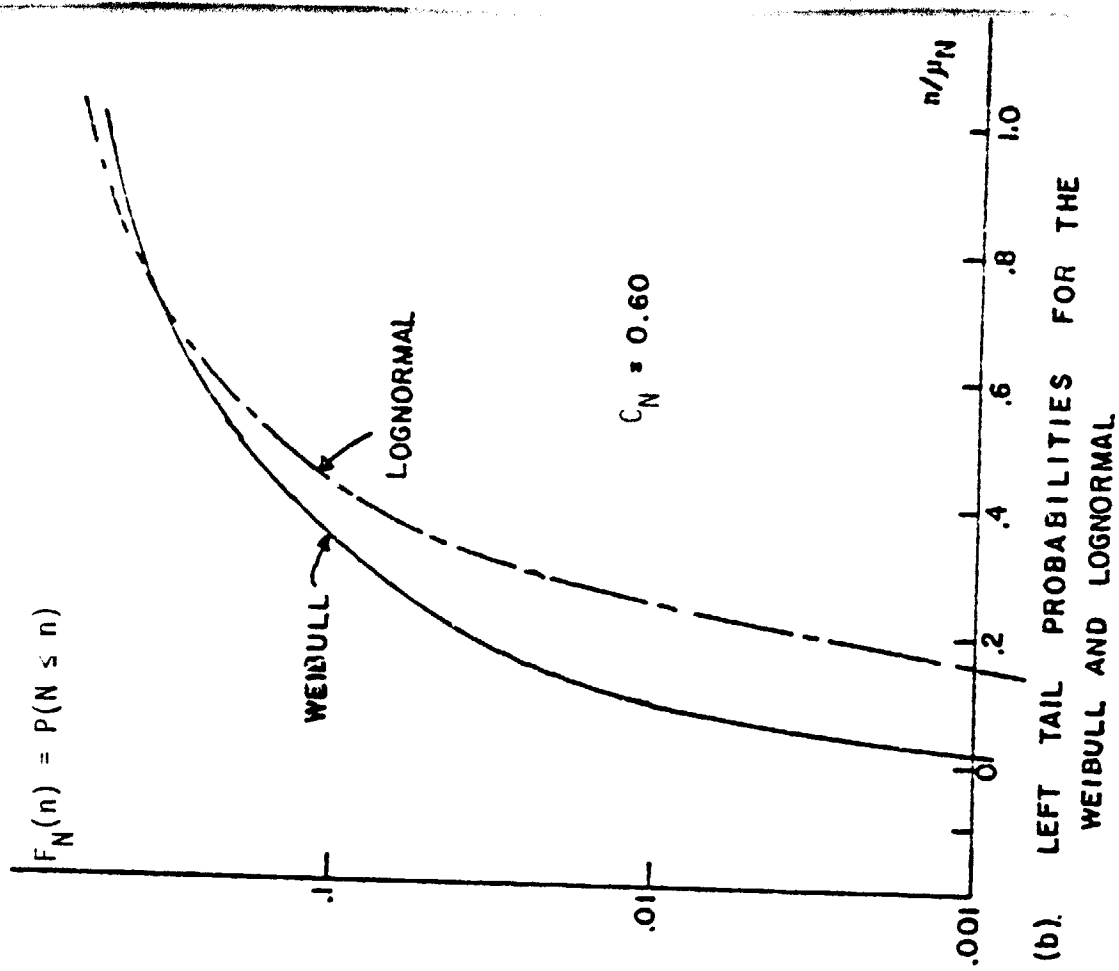
How to Estimate the Parameters

Maximum Likelihood Estimators

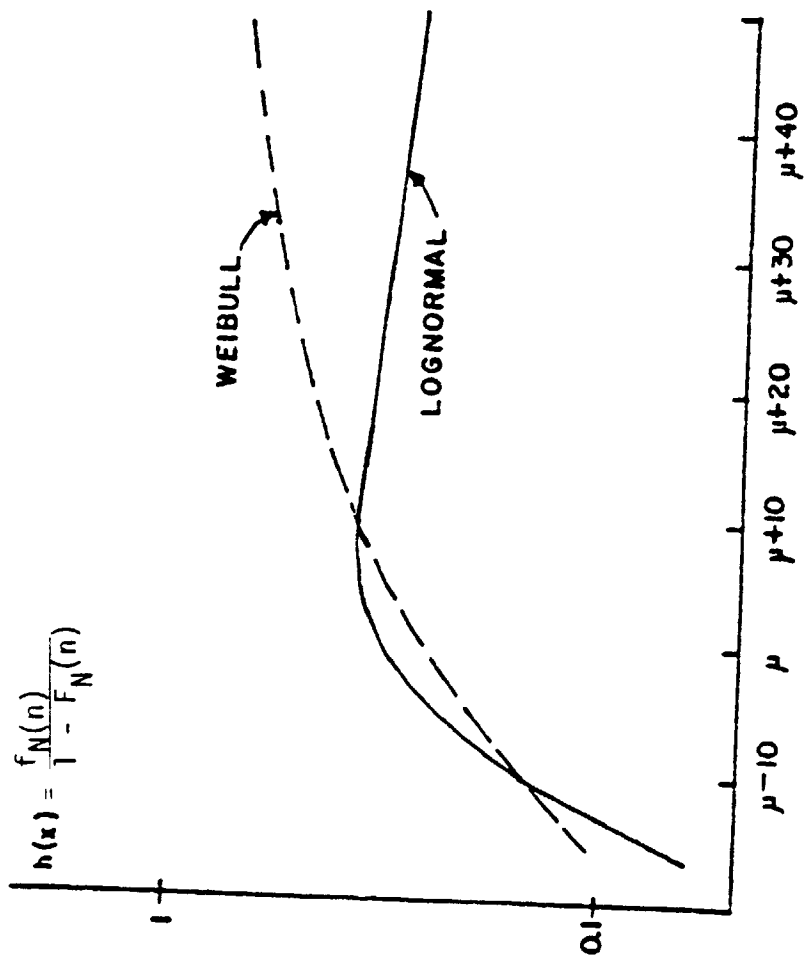
Ref: Mann et al (9)

Least Squares Estimators

See Sections 2.5 and 2.6



(b). LEFT TAIL PROBABILITIES FOR THE WEIBULL AND LOGNORMAL



(c) HAZARD FUNCTIONS FOR  $C_N = 0.60$  AND  $\mu = 10$

FIGURE 2.2 COMPARISON OF THE LOGNORMAL AND WEIBULL DISTRIBUTIONS

## 2.5 Probability Plots; Rectification of the Distribution Function

A probability plot can be used as a basis for making a subjective judgement on how well a set of data fit a particular distribution. Following are the transformations on the normal, lognormal and Weibull (2 and 3 parameters) which permit their distribution functions to be plotted as straight lines.

(a) Normal distribution; the distribution function is,

$$F_N(n) = \Phi \left( \frac{n - \mu}{\sigma} \right) \quad (2.8)$$

where  $\Phi$  is the standard normal cdf, and  $\mu$  and  $\sigma$  are the mean and standard deviations of the random variable  $N$ .

Inverting, this expression

$$\frac{n - \mu}{\sigma} = \Phi^{-1}(F) \quad (2.9)$$

and

$$\Phi^{-1}(F) = \left( \frac{1}{\sigma} \right) n - \frac{\mu}{\sigma}. \quad (2.10)$$

A linear relationship  $Y = AX + B$  has been established where

$$Y \equiv \Phi^{-1}(F), \quad X = n, \quad A = 1/\sigma, \quad B = -\mu/\sigma \quad (2.11)$$

(b) Lognormal distribution; the distribution function is.

$$F_N(n) = \Phi \left( \frac{\ln n - \mu_X}{\sigma_X} \right) \quad (2.12)$$

where  $X = \ln N$ . Here the statistical parameters are  $\mu_X$  and  $\sigma_X$  which are the

mean and standard deviations of  $X = \ln N$ . Inverting,

$$\frac{\ln n - \mu_X}{\sigma_X} = \phi^{-1}(F) \quad (2.13)$$

$$\phi^{-1}(F) = \left[ \frac{1}{\sigma_X} \right] \ln N - \left[ \frac{\mu_X}{\sigma_X} \right] \quad (2.14)$$

A linear relationship has been developed for  $\phi^{-1}(F)$  and  $\ln(n)$ . Here

$$Y = \phi^{-1}(F), \quad X = \ln N, \quad A = 1/\sigma_X, \quad B = -\mu_X/\sigma_X. \quad (2.15)$$

The mean and coefficient of variation of  $N$  is,

$$\mu_N = \exp \left[ \mu_X + \frac{1}{2}\sigma_X^2 \right] \quad (2.16)$$

$$C_N = \sqrt{\exp(\sigma_X^2) - 1}$$

(c) Weibull (2 parameter) distribution; the distribution function is,

$$F_N(n) = 1 - \exp[-(n/\beta)^\alpha] \quad (2.17)$$

where  $\alpha$  and  $\beta$  are the statistical parameters. Inverting,

$$-(n/\beta)^\alpha = \ln(1 - F). \quad (2.18)$$

Taking the log of both sides,

$$\alpha(\ln n - \ln \beta) = \ln[-\ln(1 - F)] \quad (2.19)$$

Let,

$$Y = \ln [-\ln (1 - F)] \quad (2.20)$$

$$X = \ln N.$$

Then, Equation 2.17 becomes linear,

$$Y = AX + B \quad (2.21)$$

where

$$A = \alpha, \quad B = -\alpha \ln \beta \quad (2.22)$$

The mean and coefficient of variation of  $N$  in terms of  $\alpha$  and  $\beta$  are

$$\mu_N = \beta \Gamma\left(\frac{1}{\alpha} + 1\right) \quad (2.23)$$

$$C_N = \alpha^{-.925}$$

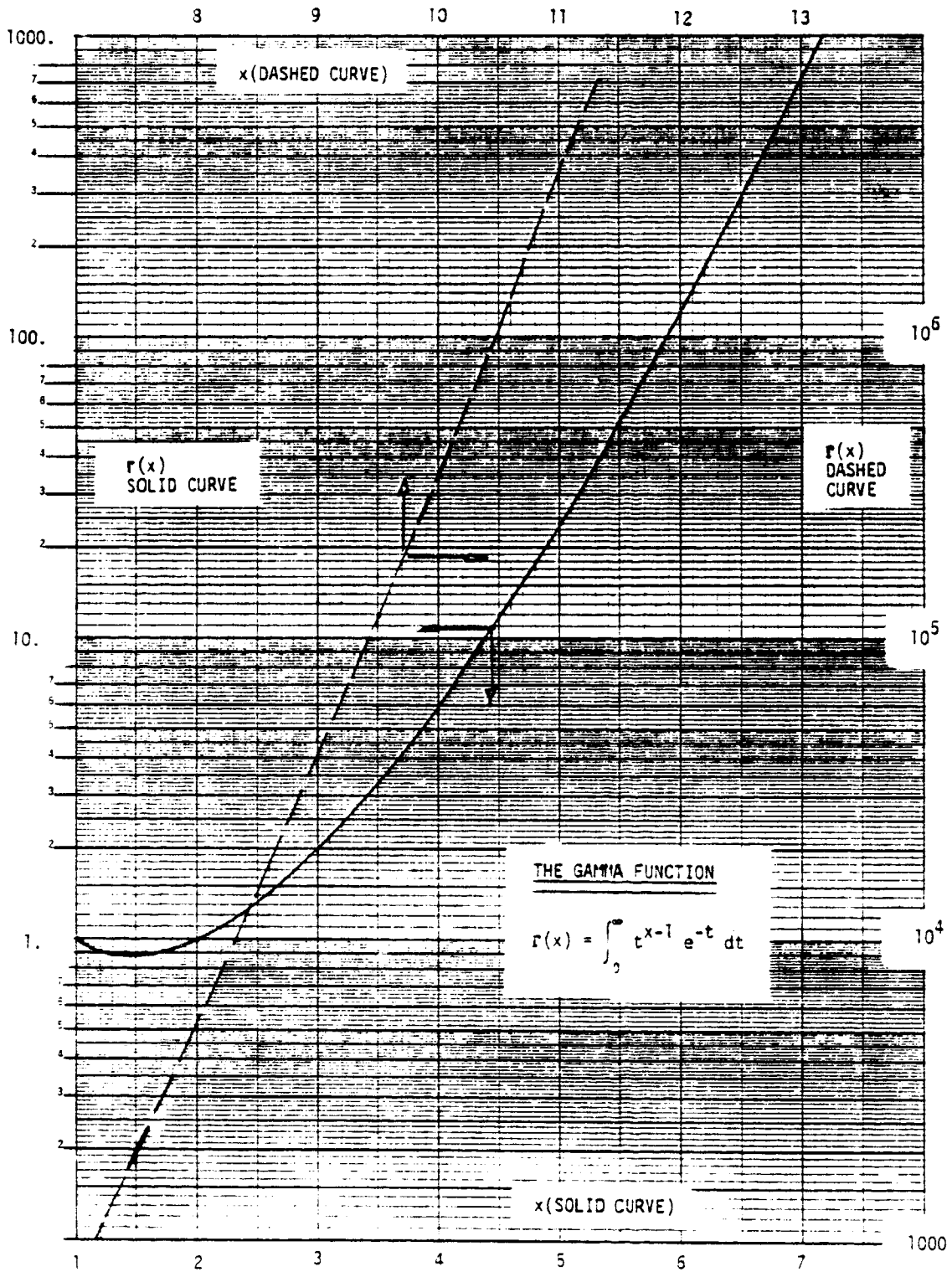
where  $\Gamma(\cdot)$  is the gamma function. A chart which enables convenient evaluation of the gamma function is provided (Figure 2.3). The expression for  $C_N$  is approximate, but the error is small for  $0.02 < C_N < 2.00$ .

(d) Weibull (3 parameter) distribution; The transformation described above for the two parameter Weibull applies directly to the three parameter model when the following transformation is made

$$Y = \ln(N - \gamma) \quad (2.24)$$

The parameter  $\gamma$  must be known in advance (see discussion below on the least squares method).

The Gamma Function



Using the above information for the transformation of each statistical model, one can plot  $F$  vs.  $N$  as a straight line on rectangular paper with the appropriate variable representation on both  $X$  and  $Y$  axes. Incidentally, this operation is performed routinely using appropriate probability paper but automatic plotting was used herein.

The empirical distribution function  $F$  for the example defined in Table 2.2 was plotted in Figure 2.4 on "lognormal paper". Note that  $F_i$  is an estimate of the actual cdf,  $F_N(n)$ . Thus, because the data tends to plot as a straight line in Figure 2.4, the lognormal may be a "reasonable" model for  $N$ . This test is of course, entirely subjective.

Addressed now will be the question of how to determine quantitatively which model, e.g. lognormal or Weibull, best fits the data. This issue is discussed in Sections 2.6 and 2.7.

## 2.6 Comparison Test Using the Least Squares Method

Using probability plots, comparison of the fit of each model can be made analytically by (a) computing the least squares estimators thereby defining the least squares line, (b) measuring the amount of deviation of the data from the least squares line and (c) choosing that model which has the least deviation.

Using the least square method, a straight line can be fitted through these data points (see Sec. 3.2.1). The least squares line is an estimate of

$$E(Y|X) = A + BX$$

$$\hat{Y} = \hat{A} + \hat{B}X \quad (2.25)$$

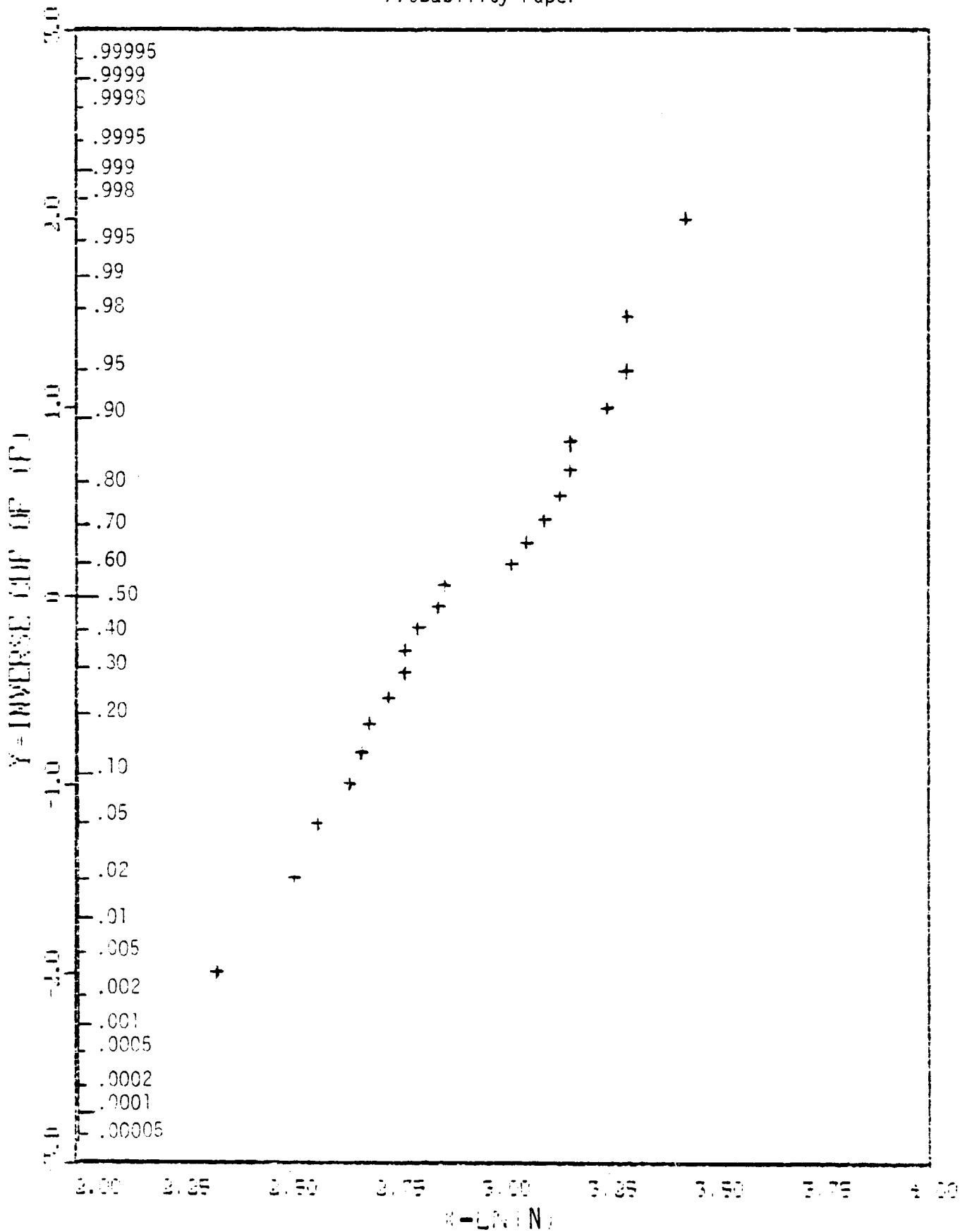
where the least squares estimates  $\hat{A}$  and  $\hat{B}$  of  $A$  and  $B$  are,

$$\hat{B} = \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n(\bar{X})^2} \quad (2.26)$$



Figure 2.4

Fatigue Data of Table 2.1 Plotted on Lognormal  
Probability Paper



$$\hat{A} = \bar{Y} - \hat{B}\bar{X}$$

and where  $\bar{X}$  and  $\bar{Y}$  are the sample means of  $X$  and  $Y$ , respectively.

The least squares line is shown for normal, lognormal, and Weibull plots of the data of Table 2.2 in Figures 2.5, 2.6 and 2.7. The purpose of such an analysis is to provide a test (albeit subjective) for establishing which model provides the best fit of the F-N data. The model for which the data seems to best plot as a straight line is the one which is the "best fit".

Such an exercise is not always successful in identifying the best model as Figures 2.5, 2.6 and 2.7 illustrate. All three models seem to provide a reasonable representation. Therefore, an objective test is necessary.

A computer program (LESQUE-I) has been developed to analyze F-N data. The output includes, (a) determination of the least squares estimators  $\hat{A}$  and  $\hat{B}$ , (b) the parameters of the model from the least squares line and (c) the sample correlation coefficient  $\hat{\rho}$ . The coefficient of correlation is used as a measure of goodness fit. The sample correlation coefficient, denoted by  $\rho$ , is expressed as,

$$\rho = \frac{1}{n} \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{s_X} \right) \left( \frac{Y_i - \bar{Y}}{s_Y} \right) \quad (2.27)$$

with  $-1 \leq \rho \leq 1$  and where  $\bar{X}$  and  $\bar{Y}$  are the sample means and

$$s_X = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2, \quad (2.28)$$

$s_Y$  has a similar expression.

$\rho$  is computed for each model. If all the data fell on a straight line,  $\rho = 1$  (or  $-1$ ). If the data indicates no linear relationship,  $\rho = 0$ . Therefore the model having a larger  $\rho$  suggests a better fit.

Figure 2.5

Fatigue Data of Table 2.1 Plotted on Normal  
Probability Paper

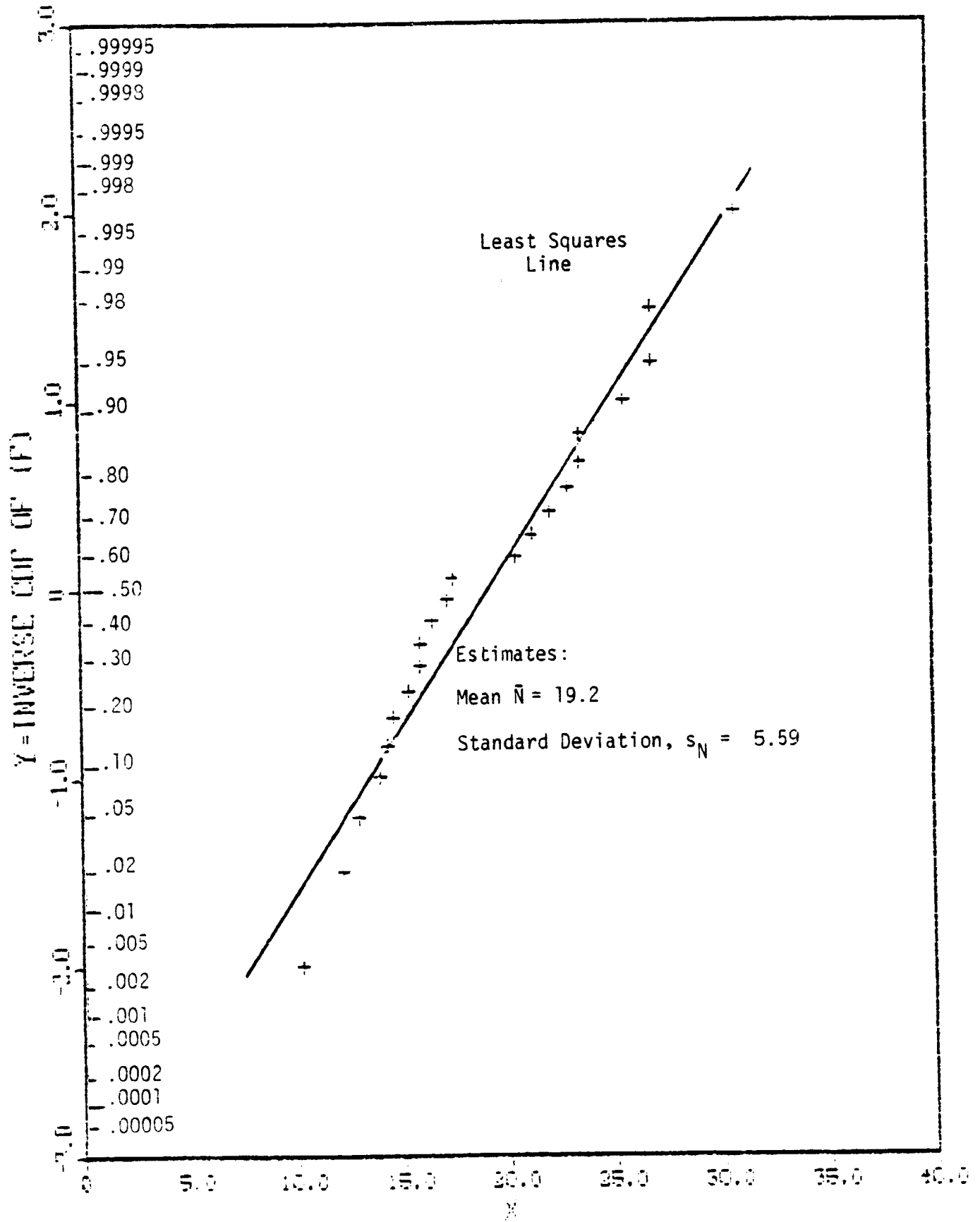


Figure 2.6  
 Fatigue Data of Table 2.1 Plotted on Lognormal  
 Probability Paper

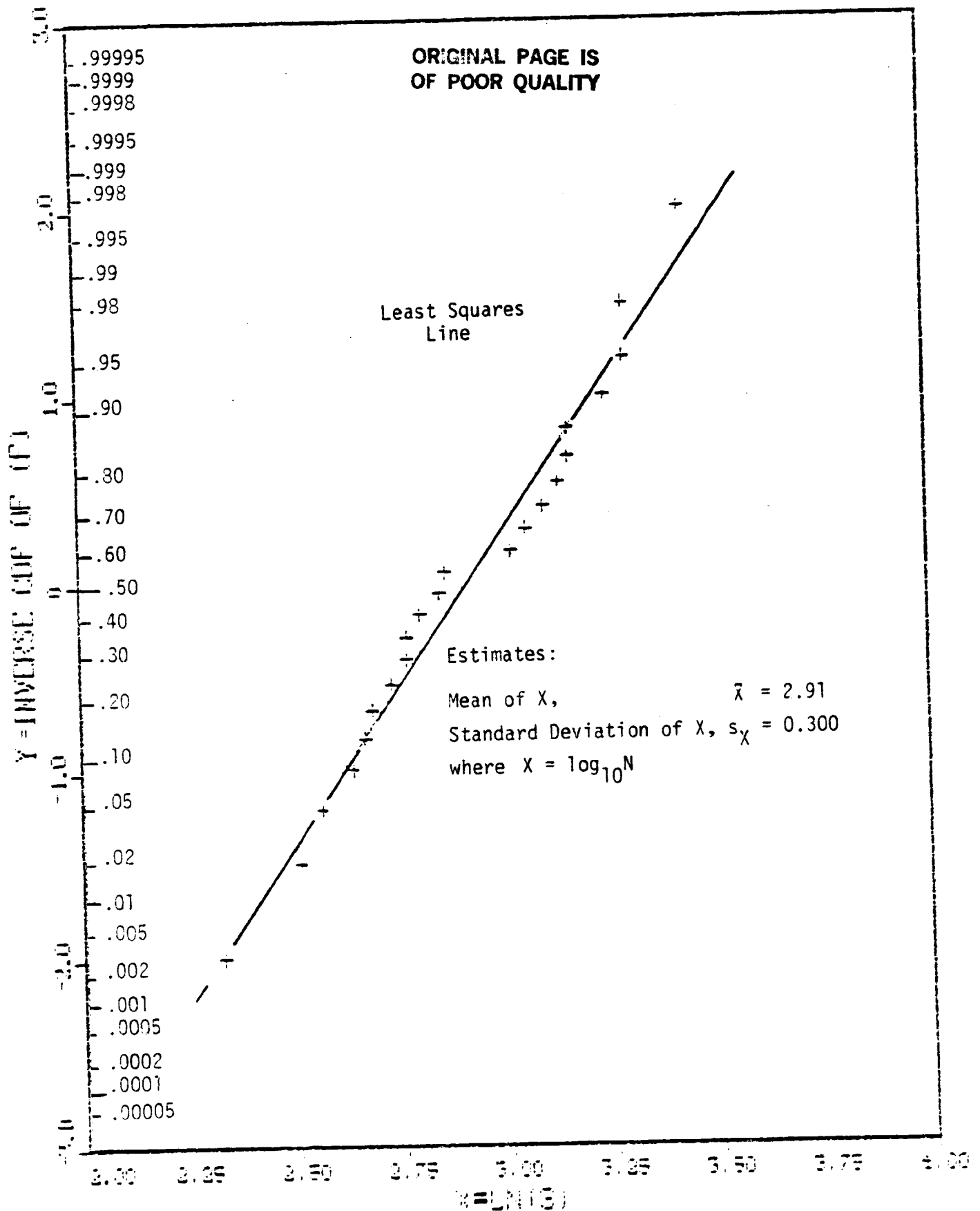
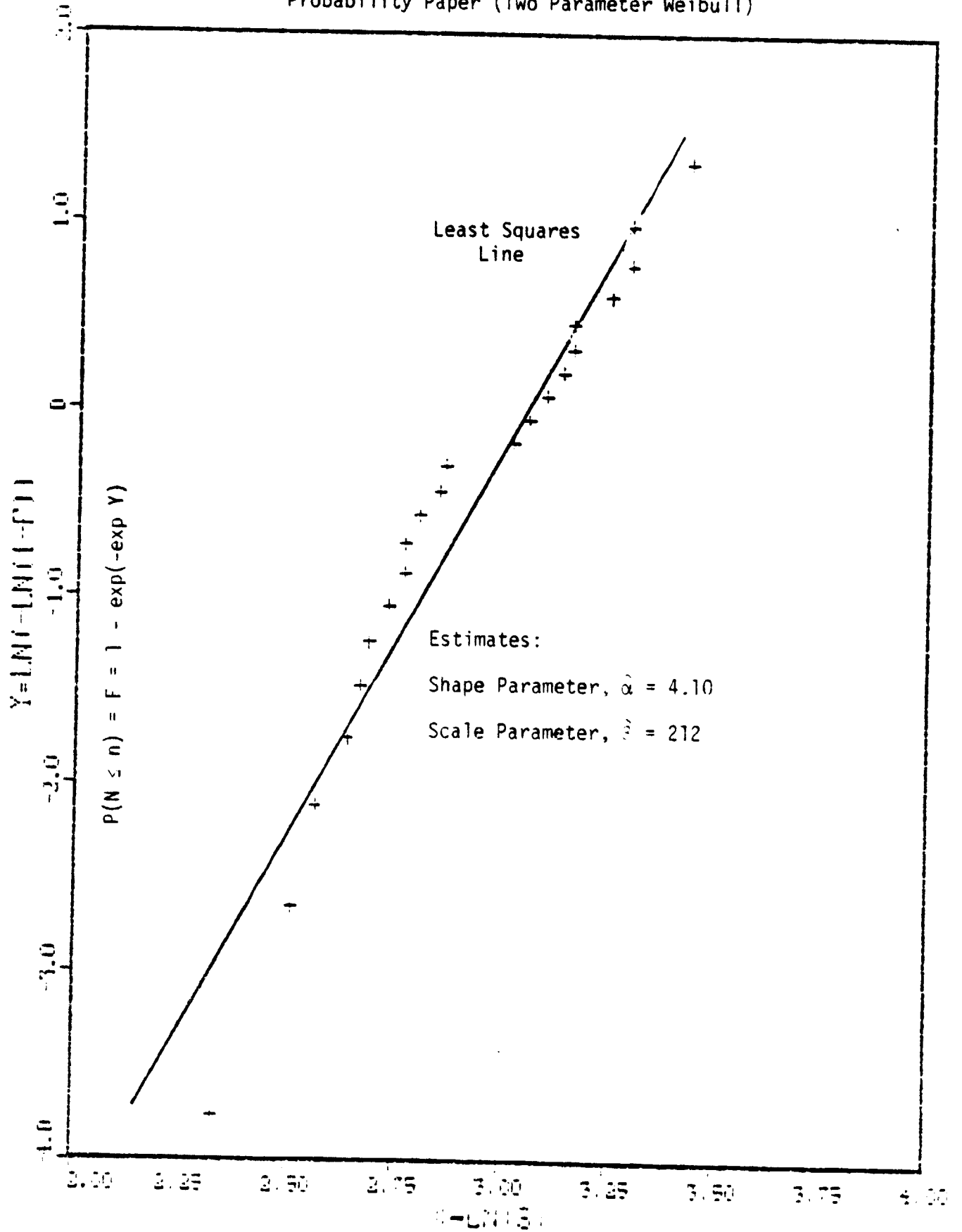


Figure 2.7

Fatigue Data of Table 2.1 Plotted on Weibull  
Probability Paper (Two Parameter Weibull)



The program was used to analyze the data of Table 2.2 for the two parameter models, normal, lognormal and Weibull. Upon examination of the results in Table 2.8 it is seen that the values of  $\rho$  are relatively close, and the lognormal is the winner. A more discriminating index for comparison of these data is a correlation parameter defined as

$$\epsilon = \sqrt{1 - \rho^2} \quad (2.29)$$

A much larger separation exists between these values as shown in the table.  $\epsilon$  is inversely proportional to  $\rho$ . Thus smaller  $\epsilon$  indicates a better fit, and the lognormal is the best fit of these two parameter models.

## 2.7 Least Squares Analysis of the Three Parameter Weibull -- Some Editorial Comments

A least squares analysis routine (LESQUE-II) for the three parameter Weibull (TPW) was developed independently. Analysis is similar to the two parameter case, except that the parameter  $\gamma$  must be specified before analysis can proceed. An iterative scheme is used for obtaining the least squares estimator of  $\gamma$ , denoted as  $\hat{\gamma}$ . A trial value of  $\gamma$  is chosen and  $\rho$  is computed. The least squares estimator  $\hat{\gamma}$  is obtained by repeating the process to find the value of  $\gamma$  which makes  $\rho$  a maximum.

An analysis of the data of Table 2.1 was performed. The probability plot is given in Figure 2.8. Values of  $\rho$  and  $\epsilon$  are given in Table 2.8.

The three parameter Weibull (TPW) is frequently used as a model for cycles to failure (e.g. Fong (16)) and the ASTM has recommended the use of the TPW(17). The TPW is attractive because the location parameter  $\gamma$  defines a non-zero lower bound on the sample space. In theory, such a model seems more realistic than the two parameter models which permit values (albeit with small probability) down to zero.

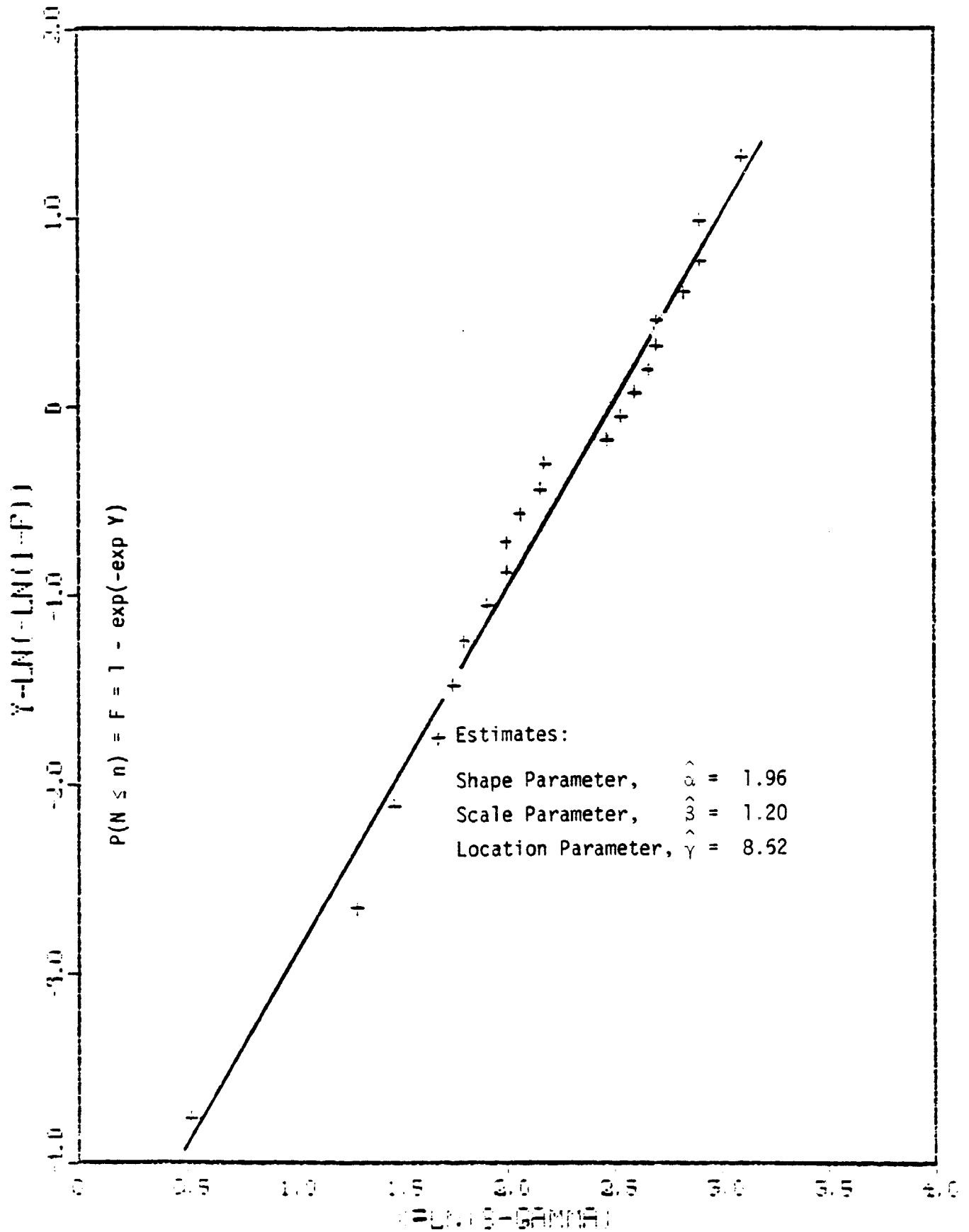
Table 2.8

Summary of Results of the Least Squares Comparison Test  
on Data of Table 2.2

	Correlation Coefficient $\rho$	Correlation Parameter $\epsilon = \sqrt{1 - \rho^2}$
Two Parameter Models		
Normal	.9797	0.200
Lognormal **	.9876	0.157
Weibull	.9683	0.250
Three Parameter Model		
Weibull	.9908	0.135

\*\* Best fit of the two parameter models

Figure 2.8  
 Fatigue Data of Table 2.1 Plotted on Weibull  
 Probability Paper (Three Parameter Weibull)

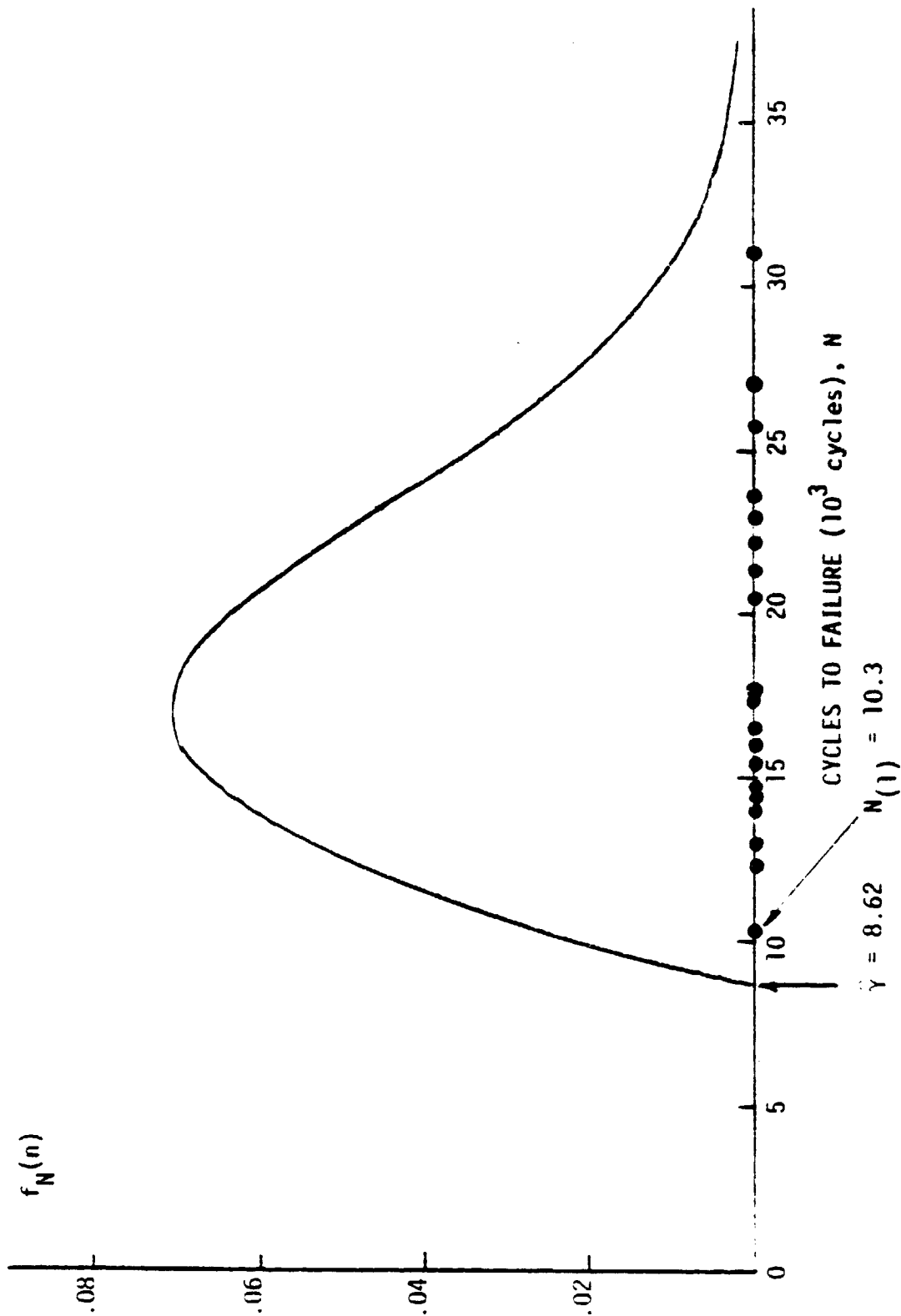




There are however some undesirable features of the TPW which may make its use impractical in certain cases:

1. The TPW is a very difficult distribution to use because
  - a) A complicated iterative program is required to compute estimates of the parameters, e.g. the maximum likelihood estimates (9).
  - b) Distributions of the estimators for routine statistical operations such as confidence intervals, testing hypotheses, etc. are difficult if not impossible to obtain.
  - c) Complicated numerical analysis is required (integration or Monte Carlo) for any reliability analysis involving any other random design factors.
2. The results given in Table 2.8 show clearly that the TPW provides a better fit to the Table 2.1 data than the three two-parameter models studied, but this comes as no surprise. More parameters produce a better fit. Why not use a four-parameter model, ... or five?
3. Figure 2.9 show the density function of the fitted TPW relative to the data. Values of the location parameter  $\hat{\gamma}$ , and the smallest sample point  $N_{(1)}$  are as shown. Note that  $\hat{\gamma}$  is only slightly below  $N_{(1)}$ . Moreover, recall that  $\hat{\gamma}$  is a point below which we are absolutely certain that no future values will fall. Examination of these values of  $\hat{\gamma}$  in light of the scatter in the small sample suggests that to define  $\hat{\gamma}$  as a lower bound is risky to say the least (e.g., the normal model predicts that 3% of the data will fall below  $\hat{\gamma}$ ).
4. Because of the truncation of the distribution at a level which may be too high, use of the TPW may result in risk estimates which are too low. Thus it is likely that the model may produce unconservative designs.

Figure 2.9  
 The Probability Density Function of the Three Parameter Weibull (TPW)  
 Shown Relative to the Data



## 2.8 Another Comparison Test Based on Fit to the Empirical Distribution Function

Wirsching and Carlson (10) developed an objective method for determining which of several competing statistical models best describes the data. Their test is based on deviations between the empirical and hypothesized distribution functions.

Let  $F_i$  denote the empirical cdf, which is determined from the sample and is an estimate of the cdf of  $X$ . The cdf of the  $j^{\text{th}}$  model considered is  $G_j(x; \hat{\theta})$  where  $\hat{\theta}$  is the parameter as estimated by the data.

The rationale for the test is that if  $G_j(x; \hat{\theta})$  is the sampling distribution of  $X$ , or a good approximation thereof, and  $F_i$  will tend to be "close to"  $G_j(x; \hat{\theta})$ . The  $W$ -statistic, a measure of this closeness, is based upon a form similar to the Cramer-von Mises statistic used for goodness-of-fit tests

$$W_j^2 = \frac{1}{n} \sum_{i=1}^n D_{ij}^2 \quad (2.30)$$

where

$$D_{ij} = G_j(x_i; \hat{\theta}) - F_i \quad (2.31)$$

Deviations associated with the  $W$ -statistic are shown in Figure 2.10.

The value of  $W_j$  is computed for each of the competing statistical models. The model having the smallest  $W_j$  is considered to be the best fit. The  $W$ -statistic is computed for the normal, lognormal and Weibull using the data of Table 2.1.

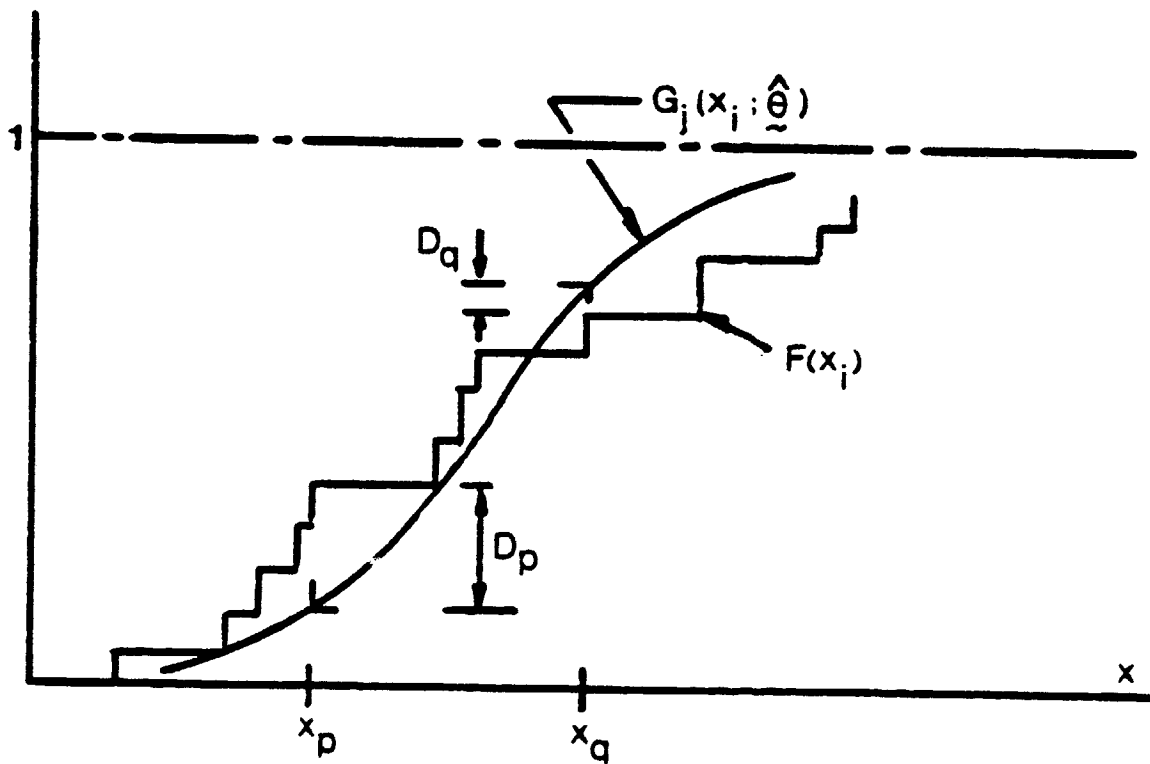
### W-STATISTIC OF DATA OF TABLE

Normal	0.0591
Lognormal	0.0491
Weibull	0.0525

Figure 2.10

Deviations Associated with the W-Statistic

Note that the deviation associated with each  $x_i$  is the vertical distance from the top of the step of empirical cdf,  $F(x_i)$ , to the hypothesized cdf  $G_j(x_i)$



This test agrees with the least squares comparison test by ranking the lognormal as the best. However the two tests disagree on the normal and Weibull. It is the opinion of the author that the W-statistic is the more powerful test.

## 2.9 The Lognormal Assumption

For the most part it will be assumed herein that the lognormal is the governing distribution for not only cycles to failure data, but also for the other fatigue design factors. See Appendix I for details of method. The reasons for the use of the lognormal are:

1. The lognormal generally has been shown to provide a reasonable description for the distribution of a wide variety of design variables. For example, upon analysis of cycles to failure data, using methods described above, the author has found that the lognormal consistently provides a better fit for fatigue data than does the Weibull or normal.
2. Statistical properties of the lognormal are well defined. All of the theory developed for normal, e.g. confidence intervals for the mean, apply equally to the lognormal.
3. The lognormal is easy to use. For example if fatigue life  $T$  is a multiplicative function of several lognormal random variables, then  $T$  is exactly lognormal and it is a simple matter to compute failure probabilities.
4. Reliability formats using the lognormal can easily accommodate design variables having relatively large coefficients of variation. Some formats rely on small variance assumptions for the design factors.
5. The lognormal is already widely used in the design profession. For example, commonly used methods of linear model analysis for characterizing S-N fatigue data implicitly assumes that cycles to failure has a lognormal distribution (e.g., see Chapters 3 and 4).

## 2.10 The Use of the Tolerance Interval of Establishing a Design Value

The "tolerance interval" can be used to establish consistent and rational design values on the safe side of data. (For a general reference, see Natrella (11)). Design values in MIL-HDBK-5, for example, are based on tolerance intervals (12). The following discussion describes the rationale and demonstrates the operation of establishing the lower tolerance limit as a design value. Unfortunately, while this analysis can specify a lower bound to SN data, it fails to provide a mechanism for consideration of other factors in the fatigue equation which are subject to uncertainty.

As an example, consider the cycles to failure data of Table 2.1. It is assumed that  $N$  has a lognormal distribution so that  $Y = \log N$  has a normal distribution. Values of  $Y_i$ , as well as the sample mean  $\bar{Y}$  and sample standard deviation  $s$ , of  $Y$  are given in Table 2.9.

Suppose, for example, a decision has been made to establish a design value  $Y_D$  (or  $N_D = 10^{Y_D}$ ) as the value below which it is anticipated that no more than 1% ( $\alpha = 0.01$ ) of future measurements would be expected to fall. Figure 2.11 illustrates  $Y_D$  relative to the distribution of  $Y$ . The value of  $Y_D$  and the corresponding  $N_D$  are easily calculated, assuming that  $Y$  has a normal distribution with mean  $\bar{Y}$  and standard deviation  $s$ .

### Design Value for Y:

In general

$$Y_D = \bar{Y} - z_{1-\alpha} s \quad (2.32)$$

In this example

$$Y_D = 1.27 - (2.33)(0.127) = 0.9741$$

Table 2.9

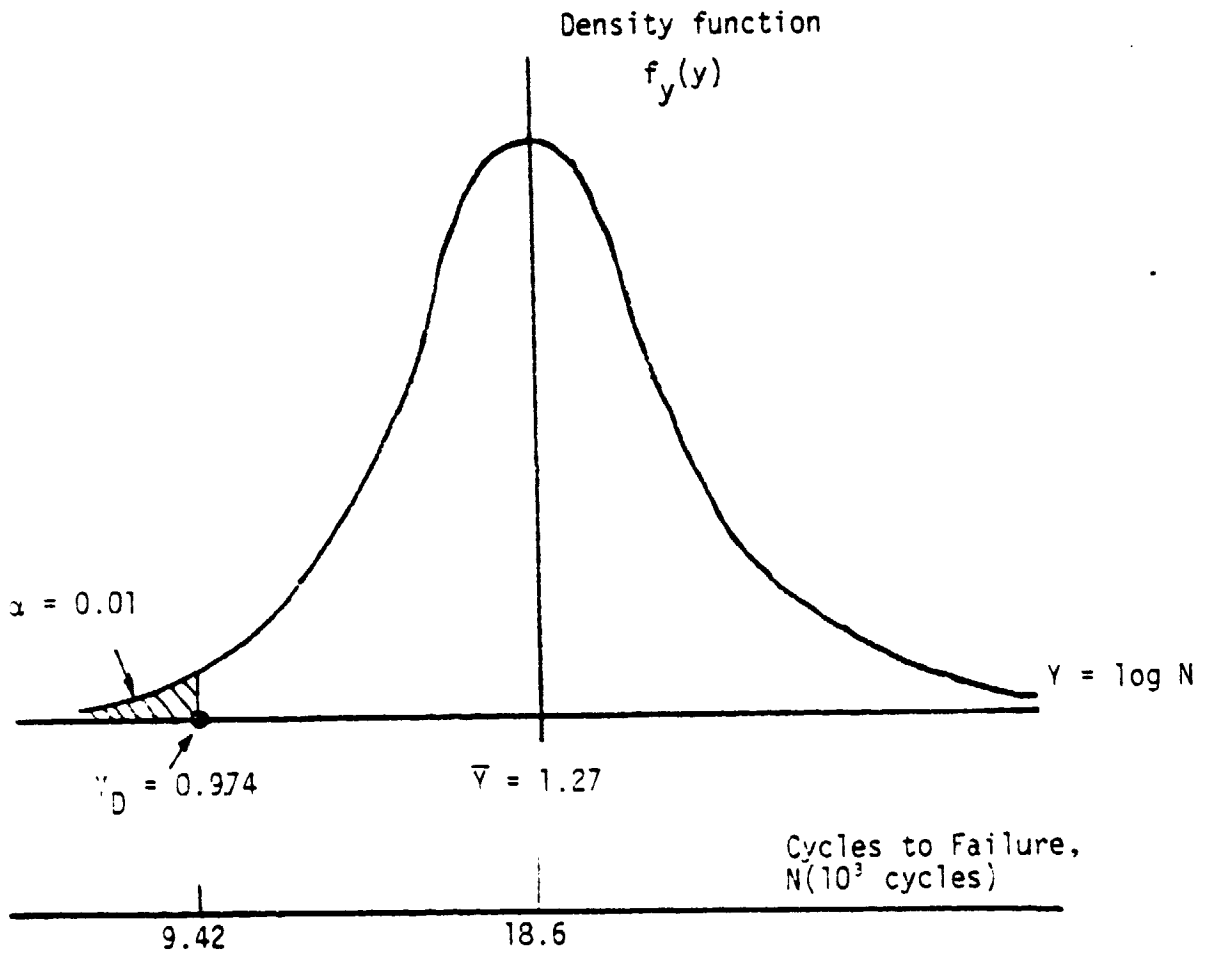
Cycles to Failure of Specimens Tested at the Same Stress Level

$N_i$ Cycles to Failure ( $10^3$ cycles)	$Y_i = \log_{10} N_i$
15.4	1.19
22.2	1.35
17.3	1.24
23.6	1.37
14.4	1.16
12.3	1.09
16.5	1.22
25.7	1.41
17.5	1.24
27.0	1.43
20.5	1.31
21.3	1.33
14.0	1.15
27.0	1.43
23.0	1.36
23.6	1.37
16.0	1.20
14.7	1.17
10.3	1.01
16.0	1.20
13.0	1.11
31.0	1.49
n=22	$\bar{Y}=1.27$ s=0.127

Figure 2.11

Design Value for N Based on Assumption that

$Y$  is Normal  $(\bar{Y}, s)^*$



\*As explained in the text, this approach gives non-conservative design values because it fails to account for the fact that  $\bar{Y}$  and  $s$  are themselves random variables.



Design Value for N:

In general

$$N_D = 10^{Y_D} \quad (2.33)$$

In this example

$$N_D = 10^{0.9741} = 9.42 (10^3 \text{ cycles})$$

It is estimated that there is only a 1% chance that a specimen selected at random would have N cycles to failure less than 9.42 thousand cycles. Stated another way, for a large number of specimens we would expect about 1% of them to have cycles to failure less than 9.42 thousand cycles.

But the analysis of Figure 2.11 fails to account for the fact that  $\bar{Y}$  and  $s$  are themselves random variables. Thus,  $Y_D$ , as seen from Eq 2.32, is a random variable, and  $N_D$  from Eq. 2.33 is also a random variable.

For example, if several labs independently conducted this same test on 22 specimens, each would obtain a different value of  $N_D$  because of inherent randomness in the process. Therefore,  $N_D$  being a random variable has its own distribution function  $F(N_D)$  and it makes sense to say, for example, that there exists a value of  $N_{.01}$  corresponding to  $F(N_D) = .05$ . This value denoted as  $N_\alpha$  and called a "tolerance limit" is the value above which we may predict with 95% confidence that 99% of the population will lie. This scheme is commonly used to establish design values (e.g., Reference 12).

In general,  $N_\alpha$  is the point below which we expect proportion  $\alpha$  to lie with confidence  $\gamma$ .

To find a single value  $Y_\alpha$  above which we may predict with confidence  $\gamma$  that a proportion  $\alpha$  of the population will lie

$$Y_\alpha = \bar{Y} - K_{\alpha,\gamma} s \quad (2.34)$$

where  $K_{\alpha,\gamma}$  is found from Table 2.10. Using the above example,  $n=22$ ,  $\gamma=0.95$ ,  $\alpha=0.01$  and  $K_{.01,.95}=3.233$  from Table 2.10 (noting that  $p=1-\alpha$ )

$$\begin{aligned} Y_\alpha &= 1.27 - (3.233)(0.127) \\ &= .8594 \end{aligned} \quad (2.35)$$

and the corresponding point for  $N$ , denoted as  $N_\alpha$  for  $\gamma=.95$ ,  $\alpha=0.01$  is

$$\begin{aligned} N_\alpha &= 10^{Y_\alpha} = 10^{0.8594} \\ N_\alpha &= 7.23 \text{ thou-cycles} \end{aligned} \quad (2.36)$$

Consider another example. To establish a "safe life" in turbine discs, the following criteria is used for civil engines in the UK (13). The safe life is established at minus three standard deviations from the population mean. Then there is a 95% confidence that the probability of failure at the safe life does not exceed 1 in 750. Note that for the normal distribution, the tail area beyond three standard deviations is 0.00135 or approximately 1 in 750.

For  $n=22$ ,  $\gamma=0.95$  and  $\alpha=.00135$ , the value of  $K_{\alpha,\gamma}$  is found from interpolation of Table 2.10.  $K_{.00135,.95}=4.20$  Using values from the above example,

$$\begin{aligned} Y_\alpha &= 1.27 - (4.20)(0.127) \\ &= 0.7367 \end{aligned} \quad (2.37)$$

The "safe life" is

$$\begin{aligned} N_\alpha &= 10^{Y_\alpha} = 10^{0.7367} \\ &= 5.45 \text{ thou-cycles} \end{aligned} \quad (2.38)$$

**TABLE 2.10 FACTORS FOR ONE-SIDED TOLERANCE LIMITS FOR NORMAL DISTRIBUTIONS \***

Factors  $K$  such that the probability is  $\gamma$  that at least a proportion  $P$  of the distribution will be less than  $\bar{X} + Ks$  (or greater than  $\bar{X} - Ks$ ), where  $\bar{X}$  and  $s$  are estimates of the mean and the standard deviation computed from a sample size of  $n$ .

$\begin{matrix} \nearrow 1-\alpha \\ \searrow P \\ n \end{matrix}$	$\gamma = 0.75$					$\gamma = 0.90$				
	0.75	0.90	0.95	0.99	0.999	0.75	0.90	0.95	0.99	0.999
3	1.464	2.501	3.152	4.396	5.805	2.602	4.258	5.310	7.340	9.651
4	1.256	2.134	2.680	3.726	4.910	1.972	3.187	3.957	5.437	7.128
5	1.152	1.961	2.463	3.421	4.507	1.698	2.742	3.400	4.666	6.112
6	1.087	1.860	2.336	3.243	4.273	1.540	2.494	3.091	4.242	5.556
7	1.043	1.791	2.250	3.126	4.118	1.435	2.333	2.894	3.972	5.201
8	1.010	1.740	2.190	3.042	4.008	1.360	2.219	2.755	3.783	4.955
9	0.984	1.702	2.141	2.977	3.924	1.302	2.133	2.649	3.641	4.772
10	0.964	1.671	2.103	2.927	3.858	1.257	2.065	2.568	3.532	4.629
11	0.947	1.646	2.073	2.885	3.804	1.219	2.012	2.503	3.444	4.515
12	0.933	1.624	2.048	2.851	3.760	1.188	1.966	2.448	3.371	4.420
13	0.919	1.606	2.026	2.822	3.722	1.162	1.928	2.403	3.310	4.341
14	0.909	1.591	2.007	2.796	3.690	1.139	1.895	2.363	3.257	4.274
15	0.899	1.577	1.991	2.776	3.661	1.119	1.866	2.329	3.212	4.215
16	0.891	1.566	1.977	2.756	3.637	1.101	1.842	2.299	3.172	4.164
17	0.883	1.554	1.964	2.739	3.615	1.085	1.820	2.272	3.136	4.118
18	0.876	1.544	1.951	2.723	3.595	1.071	1.800	2.249	3.106	4.078
19	0.870	1.536	1.942	2.710	3.577	1.058	1.781	2.228	3.078	4.041
20	0.865	1.528	1.933	2.697	3.561	1.046	1.765	2.208	3.052	4.009
21	0.859	1.520	1.923	2.686	3.545	1.035	1.750	2.190	3.028	3.979
22	0.854	1.514	1.916	2.675	3.532	1.025	1.736	2.174	3.007	3.952
23	0.849	1.508	1.907	2.665	3.520	1.016	1.724	2.159	2.987	3.927
24	0.845	1.502	1.901	2.656	3.509	1.007	1.712	2.145	2.969	3.904
25	0.842	1.496	1.895	2.647	3.497	0.999	1.702	2.132	2.952	3.882
30	0.825	1.475	1.869	2.613	3.454	0.966	1.657	2.080	2.884	3.794
35	0.812	1.458	1.849	2.588	3.421	0.942	1.623	2.041	2.833	3.730
40	0.803	1.445	1.834	2.568	3.395	0.923	1.598	2.010	2.793	3.679
45	0.795	1.435	1.821	2.552	3.375	0.908	1.577	1.986	2.762	3.638
50	0.788	1.426	1.811	2.538	3.358	0.894	1.560	1.965	2.735	3.604

Adapted by permission from *Industrial Quality Control*, Vol. XIV, No. 10, April 1958, from article entitled "Tables for One-Sided Statistical Tolerance Limits" by G. J. Lieberman.

\*Natrella, M. G., *Experimental Statistics*, NBS Handbook 91, 1963

ORIGINAL PAGE IS  
OF POOR QUALITY

TABLE 2.10(Continued). FACTORS FOR ONE-SIDED TOLERANCE LIMITS FOR  
NORMAL DISTRIBUTIONS\*

\*The two starred values have been corrected to the values given by D. B. Owen in "Factors for One-Sided Tolerance Limits and for Variables Sampling Plans", Sandia Corporation Monograph SCR-607, available from the Clearing House for Federal Scientific and Technical Information, U.S. Department of Commerce, Springfield, Va. 22151. The Owen Tables indicate other errors in the table below, not exceeding 4 in the last digit.

n	$\gamma = 0.95$					$\gamma = 0.99$				
	0.75	0.90	0.95	0.99	0.999	0.75	0.90	0.95	0.99	0.999
3	3.804	6.158	7.655	10.552	13.857	—	—	—	—	—
4	2.619	4.163	5.145	7.042	9.215	—	—	—	—	—
5	2.149	3.407	4.202	5.741	7.501	—	—	—	—	—
6	1.895	3.006	3.707	5.062	6.612	2.849	4.408	5.409	7.334	9.550*
7	1.732	2.755	3.399	4.641	6.061	2.490	3.856	4.730	6.411	8.348
8	1.617	2.582	3.188	4.353	5.686	2.252	3.496	4.287	5.811	7.566
9	1.532	2.454	3.031	4.143	5.414	2.085	3.242	3.971	5.389	7.014
10	1.465	2.355	2.911	3.981	5.203	1.954	3.048	3.739	5.075	6.603
11	1.411	2.275	2.815	3.852	5.036	1.854	2.897	3.557	4.828	6.284
12	1.366	2.210	2.736	3.747	4.900	1.771	2.773	3.410	4.633	6.032
13	1.329	2.155	2.670	3.659	4.787	1.702	2.677	3.290	4.472	5.826
14	1.296	2.108	2.614	3.585	4.690	1.645	2.592	3.189	4.336	5.651
15	1.268	2.068	2.566	3.520	4.607	1.596	2.521	3.102	4.224	5.507
16	1.242	2.032	2.523	3.463	4.534	1.553	2.458	3.028	4.124	5.374
17	1.220	2.001	2.486	3.415	4.471	1.514	2.405	2.962	4.038	5.268
18	1.200	1.974	2.453	3.370	4.415	1.481	2.357	2.906	3.961	5.167
19	1.183	1.949	2.423	3.331	4.364	1.450	2.315	2.855	3.893	5.078
20	1.167	1.926	2.396	3.295	4.319	1.424	2.275	2.807	3.832	5.003
21	1.152	1.905	2.371	3.262	4.276	1.397	2.241	2.768	3.776	4.932
22	1.138	1.887	2.350	3.233	4.238	1.376	2.208	2.729	3.727	4.866
23	1.126	1.869	2.329	3.206	4.204	1.355	2.179	2.693	3.680	4.806
24	1.114	1.853	2.309	3.181	4.171	1.336	2.154	2.663	3.638	4.755
25	1.103	1.838	2.292	3.158	4.143	1.319	2.129	2.632	3.601	4.706
30	1.059	1.778	2.220	3.064	4.022	1.249	2.029	2.516	3.446	4.508
35	1.025	1.732	2.166	2.994	3.934	1.195	1.957	2.431	3.334	4.364
40	0.999	1.697	2.126	2.941	3.866	1.154	1.902	2.365	3.250	4.255
45	0.978	1.669	2.092	2.897	3.811	1.122	1.857	2.313	3.181	4.168
50	0.961	1.646	2.065	2.863	3.766	1.096	1.821	2.269*	3.124	4.096

\*Natrella, M. G., Experimental Statistics, NBS Handbook 91, 1963

## 2.11 The Scatter Factor Based on the Tolerance Interval

The "scatter factor" is a term which has been used to describe the variability in cycles to failure data for purposes of predicting a safe life. It is analogous to a safety factor. For example, if  $\tilde{N}$  is the median cycles to failure at a given stress level, a "design value"  $N_\alpha = \tilde{N}/\zeta$  where  $\zeta$  is a scatter factor associated with an appropriate tail probability  $\alpha$ .

As an example, consider the problem described previously, the data of which is given in Table 2.9. The design value  $N_\alpha$  is defined as the point above which at least  $P=99\%$  ( $\alpha=1\%$ ) of the values are expected to fall with confidence of  $\gamma=95\%$ . That value has been established as  $N_\alpha=7.23$  thou-cycles. The median cycles to failure is

$$\begin{aligned}\tilde{N} &= 10^{\bar{Y}} = 10^{1.27} \\ &= 18.6 \text{ thou-cycles}\end{aligned}\tag{2.39}$$

and thus the scatter factor is

$$\begin{aligned}\zeta &= \tilde{N}/N_\alpha = 18.6/7.23 \\ &= 2.58\end{aligned}\tag{2.40}$$

The general form of the scatter factor can be derived by substituting the expressions of  $\tilde{N}$  and  $N_\alpha$  in terms of the statistics for  $Y$ . Thus

$$\zeta = \tilde{N}/N_\alpha = 10^{\bar{Y}}/[10^{(\bar{Y}-K_{\alpha,\gamma} s)}]\tag{2.41}$$

## 2.12 The Use of the Prediction Interval in Establishing a Design Value

Consider a random variable  $Y$  which is normally distributed with mean and standard deviation  $\mu$  and  $\sigma$  (as the above example in Sec 2.10 and 2.11). A random sample of size  $n$  is taken, and the sample mean  $\bar{Y}$  and sample standard deviation,  $s$ , are computed.

It is required to make probability predictions about future observations. Following is a derivation of the prediction interval and a discussion of how it can be used to establish design values as well as provide statistics for reliability analysis (see References 14 and 15).

Consider the distribution of  $(Y_0 - \bar{Y})$  where  $Y_0$  is a random variable denoting a future observation and  $\bar{Y}$  is the sample mean.  $Y_0$  and  $\bar{Y}$  are normally distributed and have mean and standard deviation of  $(\mu, \sigma)$  and  $(\mu, \sigma/\sqrt{n})$  respectively. Furthermore  $Y_0$  and  $\bar{Y}$  are independent in a probability sense because  $\bar{Y}$  is derived from the first  $n$  observations and  $Y_0$  corresponds to a future observation. Hence  $(Y_0 - \bar{Y})$  is normally distributed with mean zero and standard deviation

$$\sigma' = \sigma \sqrt{1 + \frac{1}{n}} \quad (2.42)$$

Note that  $(n-1)s^2/\sigma^2$  has a  $\chi^2$  distribution with  $(n-1)$  degrees of freedom. Thus it follows that

$$\frac{Y_0 - \bar{Y}}{s \sqrt{1 + \frac{1}{n}}}$$

has a student's  $t$  distribution with  $(n-1)$  degrees of freedom. (14)

A rational method of establishing a design value can be based on a prediction interval. Let  $Y_D$  denote the value below which we expect the next observation to fall with probability  $\alpha$ . Then

$$Y_D = \bar{Y} - G(\alpha, n)s \quad (2.43)$$

where

$$G(\alpha, n) = t_{\alpha; n-1} \sqrt{1 + \frac{1}{n}}$$

where  $t_{\alpha, n-1}$  is the Student's  $t$  variate with  $n - 1$  degrees of freedom at probability level  $\alpha$

\*This is a simplified version of the Equivalent Prediction Interval concept developed by Wirsching and Hsieh (15): See also Section 4.7.

The function  $G$  is shown in Fig. 2.12 for  $\alpha = .01$  and  $.05$ . As  $n \rightarrow \infty$ , the  $t$  variate approaches the standard normal  $z$  variate (also shown on Fig. 2.12); the radial term approaches one. Thus  $G$  approaches  $z$  as shown.

As an example, consider the data of Table 2.9;  $\bar{Y} = 1.27$  and  $s = 0.127$ . Find the lower 1% prediction interval. For  $n = 22$ ,  $G(.01, 22) = 2.57$ .

Thus

$$\begin{aligned} Y_D &= 1.27 - 2.57 (.127) \\ &= 0.943 \end{aligned}$$

The design value of cycles to failure  $N$  is given as,

$$N_D = 10^{0.943} = 8.77 (10^3 \text{ cycles})$$

The prediction interval also provides a mechanism for specifying the mean and standard deviation for a reliability approach. Use of  $\bar{Y}$  and  $s$  directly in a reliability format may lead to significant errors for small sample sizes because  $\bar{Y}$  and  $s$  are only estimates of the population  $\mu$  and  $\sigma$ . The uncertainty in the estimates can be accounted for by expanding the sample standard deviation  $s$ .

An alternative expression for the design value is

$$Y_D = \bar{Y} - Z_\alpha \sigma_0 \quad (2.44)$$

where  $\sigma_0$  is defined as the "equivalent standard deviation". Comparing Eq 2.43 and 2.44, it follows that

$$\sigma_0 = g(n, \alpha) s \quad (2.45)$$

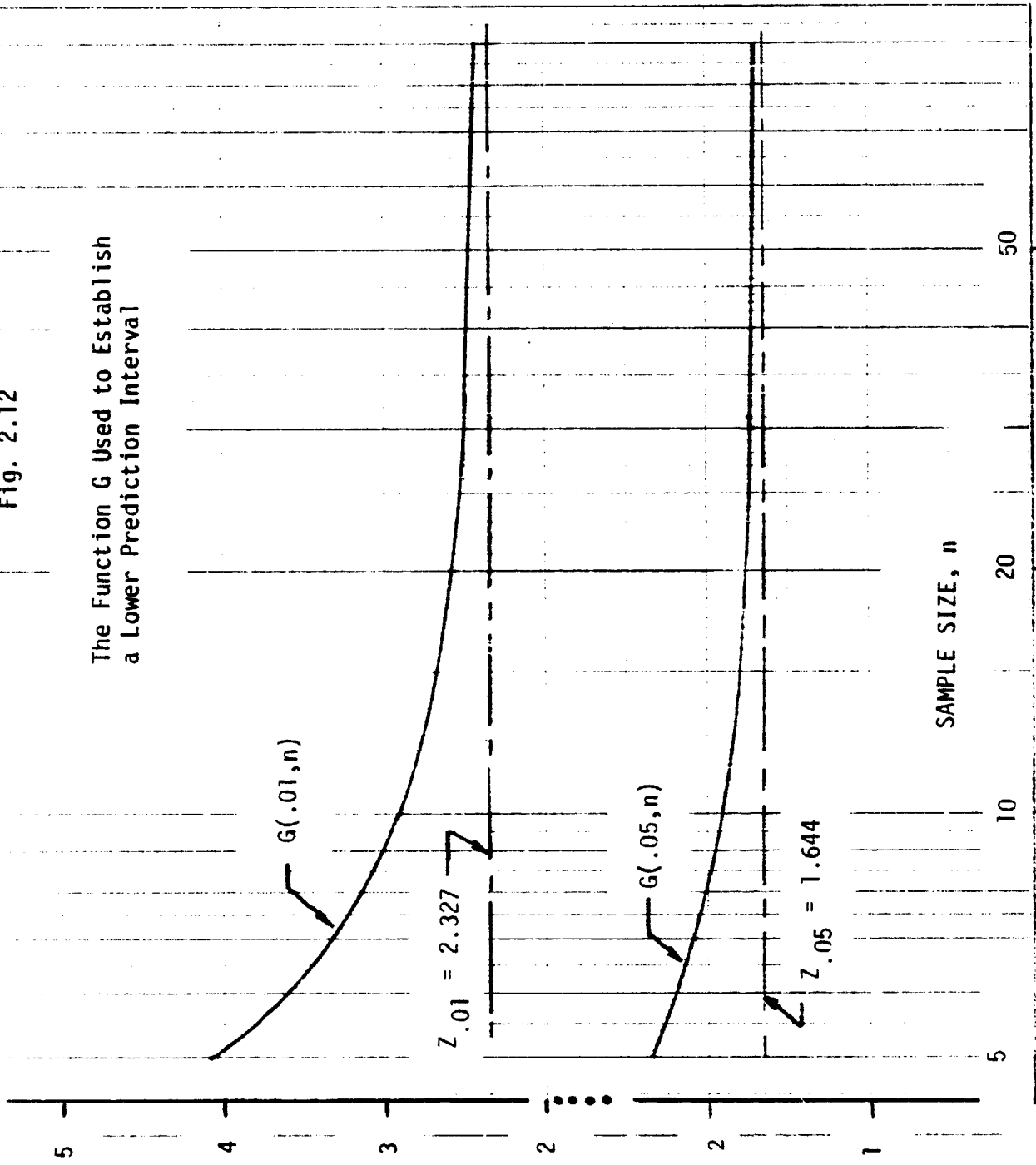
where

$$\begin{aligned} g(n, \alpha) &= t_{\alpha; n-1} \sqrt{1 + \frac{1}{n}} / z_\alpha \\ &= G(n, \alpha) / z_\alpha \end{aligned}$$

Values of  $G(n, \alpha)$  can be established from Fig. 2.12., which is then divided by  $z_\alpha$  to obtain  $g(n, \alpha)$

Fig. 2.12

The Function G Used to Establish  
a Lower Prediction Interval





The value  $\alpha_0$  can be interpreted as the expanded standard deviation to account for the fact that the estimates  $\bar{Y}$  and  $s$  are random variables. For use in reliability analysis one would state that  $Y$  has normal distribution with a mean and standard deviation  $(\bar{Y}, \sigma_0)$

For the data of Table 2.9 and Figure 2.12

$$\begin{aligned} g(.01, 22) &= G(.01, 22) / z_{.01} \\ &= 2.57 / 2.33 = 1.10 \end{aligned}$$

Thus,  $\sigma_0 = 1.10 (0.127) = 0.140$

Thus we could state in a reliability format that  $Y \sim N(1.27, 0.140)$ .

It is necessary in this approach to specify the "reference value" of  $\alpha$ . As a general recommendation, the value of  $\alpha = 0.01$  should provide reasonable results.

### 2.13 How to Establish a Design Value: Summary Comments

Given a random sample of data, it is often required to specify a design value on the safe side of that data. If the term is a "strength" variable (i.e. design becomes dangerous if variable gets too small), then the design value should be on the left side. On the other hand, if the term is a "stress" variable (i.e. design becomes dangerous if the variable gets too large) then the design value should be on the right side. Thus  $Y_D = \bar{Y} + Ks$ .

There are a number of schemes currently used in industry to specify design values for a "strength" variable. These methods are illustrated in Figure 2.13 using fatigue data presented in Table 2.11. A summary of the methods and computed values are presented in Table 2.12.

It is interesting to compare the design values with the actual population from which the data is sampled. This particular sample of size  $n = 10$  had values larger than average. Thus the lower 1% value (No. 5), based on the assumption that  $\mu_Y = \bar{Y}$  and  $\sigma_Y = s$  and using normal probabilities, was too high.

Figure 2.13

A SUMMARY OF METHODS TO LOCATE DESIGN VALUE ON SAFE SIDE (LOWER) OF DISTRIBUTION

Population:  $\mu = 1.25$       Sample:  $\bar{Y} = 1.28$   
 $\sigma = 0.12$                        $s = 0.104$

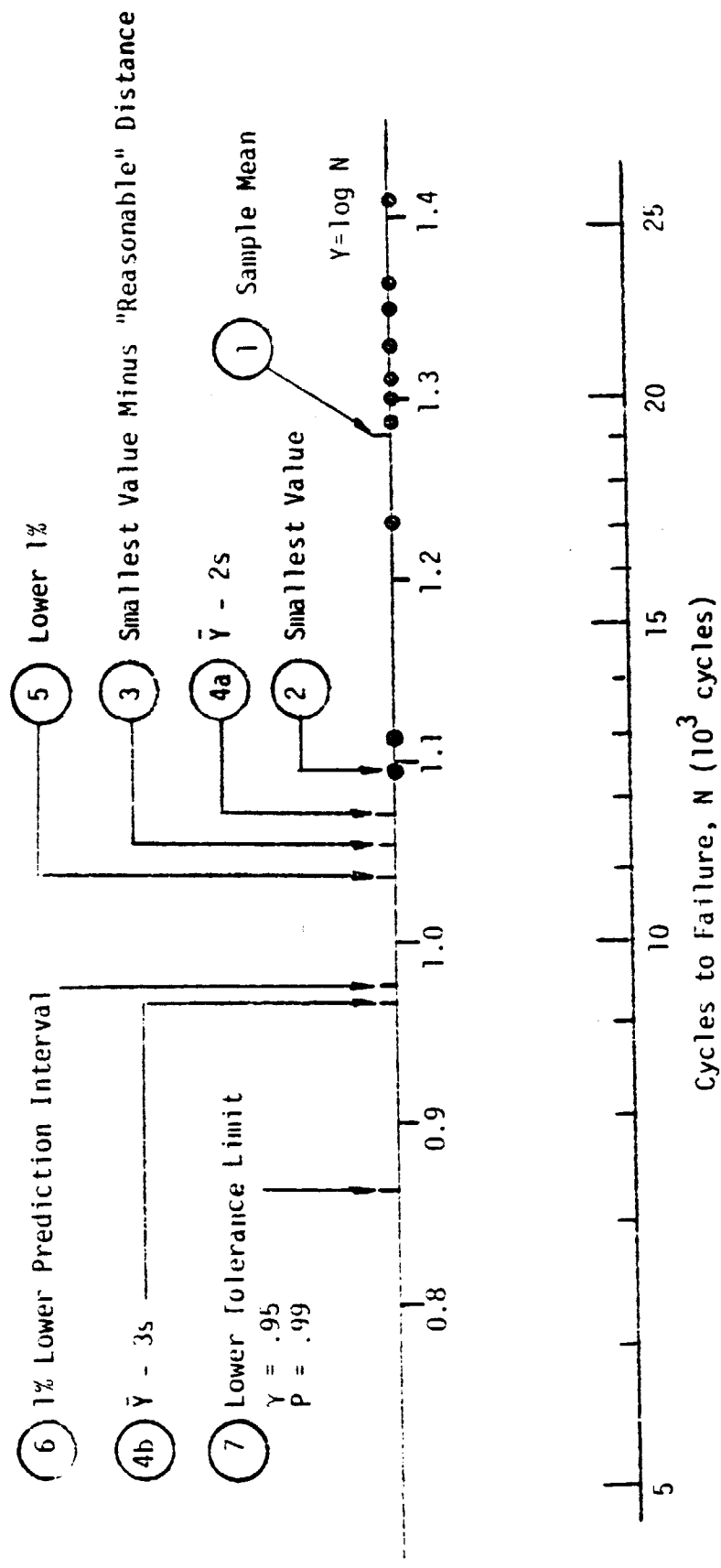


Table 2.11

Data for Example to Illustrate How Design Values  
are Established

This data on  $Y$  was actually sampled from a normal distribution having mean  $\mu_Y = 1.25$  and standard deviation  $\sigma_Y = 0.12$ .

$Y_i = \log_{10} N_i$	Cycles to Failure, $N_i$ ( $10^3$ Cycles)
1.298	19.86
1.351	22.46
1.328	21.30
1.312	20.54
1.095	12.46
1.415	25.98
1.364	23.14
1.232	17.07
1.112	12.95
1.292	19.58

Sample Mean,  $\bar{Y} = 1.280$

Sample Std. Dev.,  $s = .1047$

Table 2.12

An Example of the Various Methods to Establish Design Values  
 (Basic Assumption: Y is Normally Distributed)

$$Y_D = \bar{Y} - Ks$$

Method	Value of K	Design Value, $Y_D$
1. Mean	0	1.28
2. Smallest Value in Sample		1.095
3. Smallest Value Minus Reasonable Distance (choice arbitrary)		1.050
4. Mean - $Ks$	$K = 2$ $K = 3$	1.07 0.966
5. Lower 1%; assumes that $\mu_Y = \bar{Y}$ and $\sigma_Y = s$	2.33	1.036
6. Lower 1% Prediction Interval	2.90	0.976
7. Lower Tolerance Limit = .95 = .1%	3.98	0.863

The message here is that the statistical distribution of  $\bar{Y}$  and  $s$  should be accounted for in establishing the design value. The prediction interval (No. 6) and tolerance interval (No. 7) are the only methods which do this.

In summary it is recommended that the prediction interval or tolerance interval be used to define design values. The prediction interval can easily be used to provide the data needed for reliability analysis. Tolerance intervals are commonly used, but they tend to produce conservative values as suggested in Fig. 2.13.

#### 2.14 References for Chapter 2

1. Ang, A.H.-S. and Munse, W. H., "Practical Reliability Basis for Structural Fatigue", presented at ASCE Structural Engineering Conference, New Orleans, LA, April 1975, Preprint 2494.
2. Whittaker, I. C. and Besuner, P. M., "A Reliability Analysis Approach to Fatigue Life Variability of Aircraft Structures," Wright-Patterson Air Force Base Technical Report AFML-TR-69-65, April 1969.
3. Whittaker, I. C., "Development of Titanium and Steel Fatigue Variability Model for Application of Reliability Approach to Aircraft Structures," AFML-TR-72-236, Wright-Patterson AFB, Ohio, October 1972.
4. Sinclair, G. M. and T. J. Dolan, "Effect of Stress Amplitude on the Variability in Fatigue Life of 75S-T6 Aluminum Alloy," Transactions of the ASME, July, 1953.
5. Saltsman, J. F. and G. R. Halford, "Application of Strainrange Partitioning to the Prediction of Creep-Fatigue Lives of AISI Types 304 and 316 Stainless Steel," Paper No. 76-PVP-26, 1976.
6. Wirsching, P. H., "Probability Based Fatigue Design for Ocean Structures," Final Report, American Petroleum Institute, PRAC Project #15," Aug. 1979.
7. Gumbel, E. J., Statistics of Extremes, Columbia U., Press 1958.
8. Gumbel, E. J., "Parameters in the Distribution of Fatigue Life", Journal of the Engineering Mechanics Division, ASME, Vol. 89, No. EM5, Oct. 1963.
9. Mann, N. R. et al, Methods for Statistical Analysis of Reliability and Life Data Wiley, 1974.
10. Wirsching, P. H. and J. R. Carlson, "Model Identification for Engineering Variables," Journal of the Engineering Mechanics Division ASCE, Vol. 103, No. EM 1, Feb. 1977.

11. Natrella, M. G., Experimental Statistics NBS Handbook 91, 1963.
12. MIL-HDBK-5B, Metallic Materials and Elements for Aerospace Vehicle Structures, U.S. Department of Defense.
13. W. J. Evans, "Fatigue Limitations in the Design of Gas Turbine Discs", Memorandum 79028, National Gas Turbine Establishment, Pyestock Farnsbrough Harts, England, May 1979.
14. Bowker, A. H., and Lieberman, Engineering Statistics, Prentice Hall, 1972.
15. Wirsching, P. H. and S. Hsieh, "Linear Model in Probabilistic Fatigue Design," The Journal of the Engineering Mechanics Division, ASCE, Vol. 106, No. EM6, Dec. 1980.
16. Fong, J.T., "Uncertainties in Fatigue Life Prediction and a Rational Definition of Safety Factors", Nuclear Engineering and Design, Vol. 51, pp. 45-54, 1978.
17. A Guide for Fatigue Testing and the Statistical Analysis of Fatigue Data, ASTM STP 91-A, 1963.

## Chapter 3 ANALYSIS OF STRAIN-LIFE DATA: THE LEAST SQUARES LINE

### 3.1 Preliminary Considerations

Statistical analyses of data from a single random variable for the purpose of establishing a design value was described in Chapter 2. Examples of analyses of cycles to failure data at single stress levels were presented. But generally, it is required to establish  $N$  as a function of  $S$ . Therefore, tests are performed at various stress levels.

Typical fatigue data might consist of inelastic strain range versus  $N$  cycles to failure as illustrated in Figure 3.1. It is necessary to characterize this data for design purposes. Two basic methods are employed:

- a. The common approach is to define a design curve on the safe (lower) side of the data.
- b. A statistical summary can be presented for a reliability analysis or probabilistic design approach.

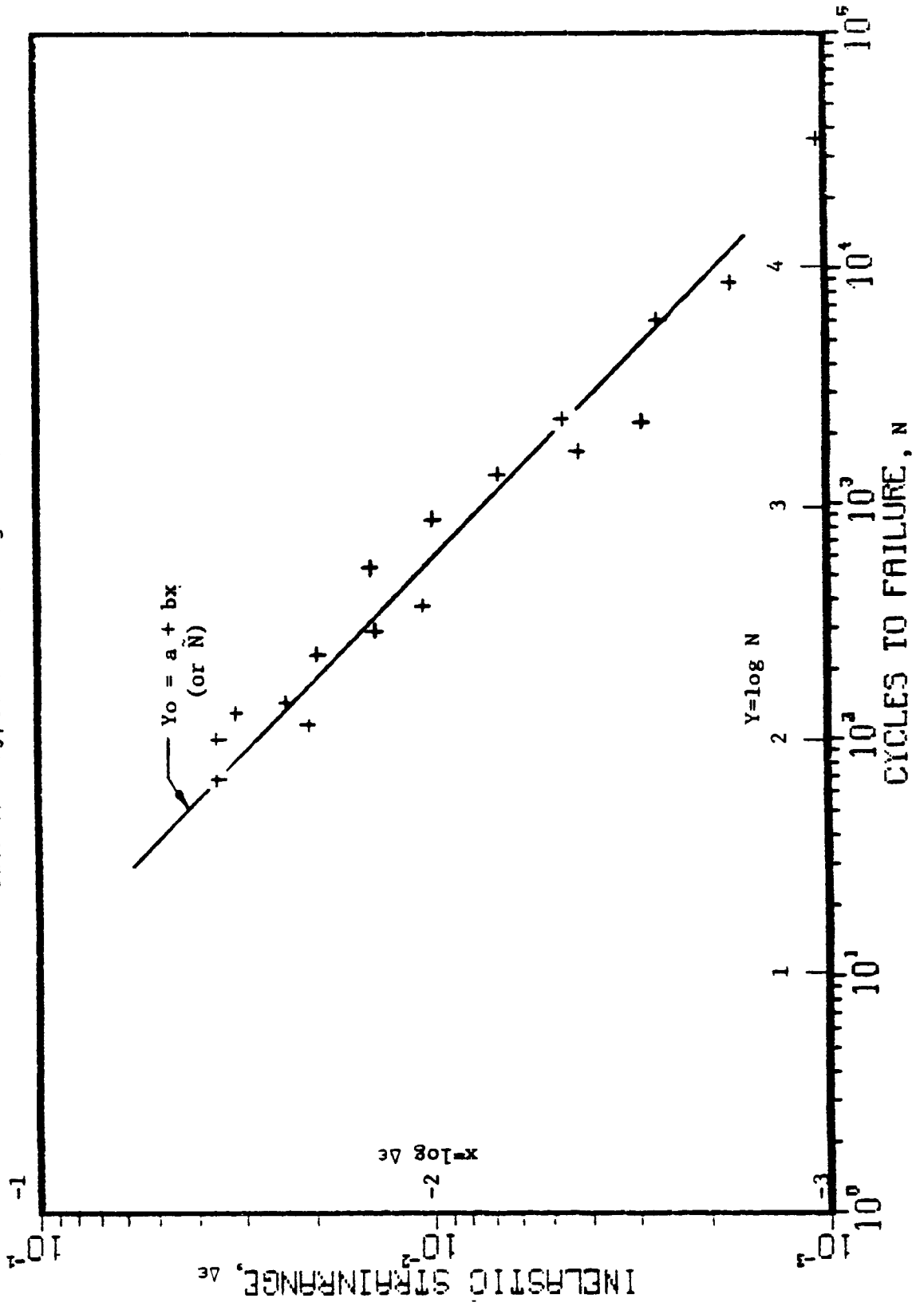
A key to strain-life data analysis is the determination of the least squares line, an estimate of the median value of  $N$  for a given strain (or stress) level. This chapter discusses the least squares line and some of its characteristics. Chapters 5 and 6 then show how to establish design values from strain-life data.

### 3.2 The Least Squares Line (Median Curve Through Strain-Life Data)

#### 3.2.1 Introduction

Methods for analyzing strain-life  $\Delta\epsilon$ - $N$  (or stress-life,  $S$ - $N$ ) data are discussed in this chapter and in the next two. The goal of such data analysis is to provide a characterization of the  $\Delta\epsilon$ - $N$  relationship in a form suitable for design purposes. General references on the least squares method include References 1, 2 and 3.

Figure 3.1  
Data from Hypothetical Fatigue Test





### 3.2.2 Median Life Curve

Consider a constant amplitude strain controlled fatigue test in which pairs of data  $(\Delta\epsilon_i, N_i)$   $i=1, n$  are collected.  $N_i$  are the cycles to failure associated with strain range  $\Delta\epsilon_i$ , and  $n$  is the sample size.  $\Delta\epsilon$  is the independent (or controlled) variable and  $N$  is the dependent variable. Data from a hypothetical test are shown in Fig. 3.1 plotted on log-log paper.

Methods of basic linear model analysis are typically used to analyze fatigue data, and these methods will be used herein to describe life relationships.

Consider first a log transformation of variables. Let

$$Y = \log N \qquad X = \log(\Delta\epsilon). \qquad (3.2)$$

Thus  $X$  is the independent variable,  $Y$  is the dependent variable.

Clearly there is no functional relationship between  $Y$  and  $X$ , but there does seem to be some kind of relation. It will be assumed that the data (e.g. Fig. 3.1) is a random sample from the following model

$$Y(x) = Y_0(x) + \delta \qquad (3.3)$$

in which  $\delta$  is a normally distributed random variable with mean equal to zero and standard deviation equal to  $\sigma$  and where

$$Y_0(x) = a + bx \qquad (3.4)$$

where  $a$  and  $b$  are constants. Thus for a specified  $X$ ,  $Y$  is a random variable normally distributed having mean and standard deviation

$$E(Y|X) = Y_0 = a + bx \qquad (3.5)$$

$$\sigma(Y|X) = \sigma \qquad (3.6)$$

Note the assumption that  $\sigma$  is a constant, not a function of  $X$ . The line  $Y_0 = a + bx$ , being the mean of  $Y$ , will pass through the "center" of the data.

Because  $Y$  is normally distributed,  $N$  (given  $\Delta\epsilon$ ) will be lognormal. Thus the median of  $N$ , denoted as  $\tilde{N}$ , is given by  $Y_0 = \log \tilde{N}$ . In terms of the original coordinates, the  $Y$  line can be written as

$$\Delta\epsilon = C\tilde{N}^{\zeta} \quad (3.7a)$$

in which it follows from the above definitions that

$$a = -\frac{1}{\zeta} \log C \quad b = 1/\zeta \quad (3.7b)$$

### 3.2.3 The Least Squares Line

The parameters  $a$  and  $b$  (and thus  $C$  and  $\zeta$ ) and  $\sigma$  are not known in advance and must be estimated from the data  $(\Delta\epsilon_i, N_i)$ ,  $i=1, n$ . Equation 3.2 is used to translate the data into  $(X_i, Y_i)$   $i=1, n$ .

Using the method of least squares,  $a$ ,  $b$ , and  $\sigma$  are estimated by  $\hat{a}$ ,  $\hat{b}$ , and  $\hat{s}$  respectively (3),

$$\hat{b} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (3.8)$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} \quad (3.9)$$

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n [Y_i - (a + bX_i)]^2 \quad (3.10)$$

where  $\bar{X}$  and  $\bar{Y}$  are the sample means of  $X$  and  $Y$  respectively. Because each  $Y_i$  is a random variable, the estimates  $\hat{a}$ ,  $\hat{b}$ , and  $s$  are also random variables.

The "best fit" line

$$\hat{Y} = \hat{a} + \hat{b}X \quad (3.11)$$

is called the least squares line.  $\hat{Y}$  is the estimate of  $Y_0$ , i.e. the mean of  $Y$  given  $X$ .

As an example, the PP strainrange data for AISI 316 presented by Saltsman and Halford (4), and shown in Fig. 3.2, will be analyzed. This data is presented in Table 3.1, along with statistics associated with this

\* PP strainrange is tensile plasticity reversed by compressive plasticity; no creep.

Figure 3.2  
 The Least Squares Line for Fatigue Data  
 [AISI 316; after Saltzman and Halford (4)]

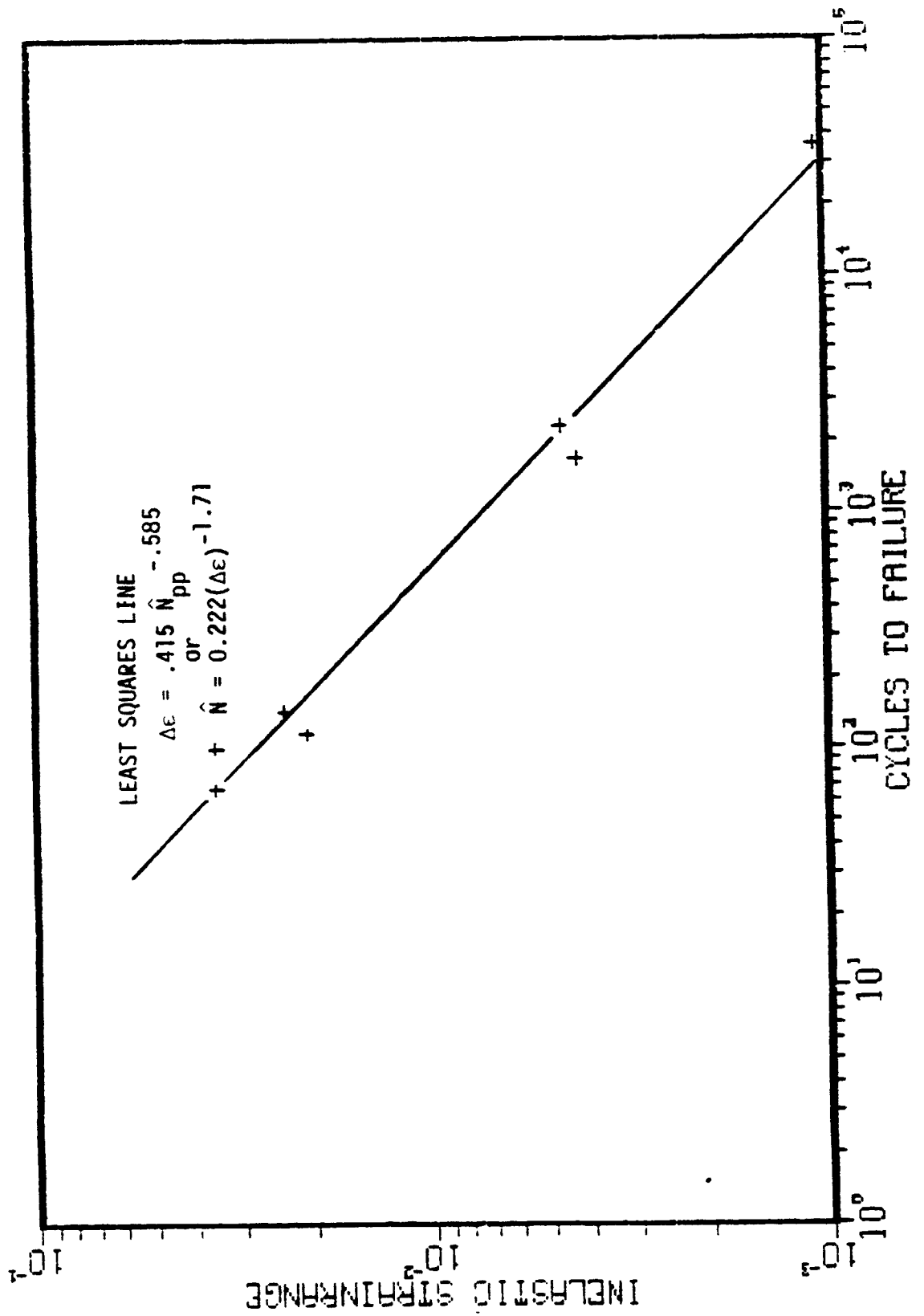


Table 3.1

Statistical Analysis of AISI 316 PP Data

[after Saltsman and Halford (4)]

Sample Size n=7

Strain Range ( $\Delta\epsilon$ ) <sub>i</sub>	Cycles to Failure N <sub>i</sub>	$X_i = \log(\Delta\epsilon)_i$	$Y_i = \log N_i$
.00424	1700	-2.373	3.230
.00105	35600	-2.979	4.551
.03508	120	-1.455	2.079
.03496	68	-1.564	1.832
.00466	2333	-2.332	3.368
.02066	116	-1.685	2.064
.02360	146	-1.627	2.164

Sample Mean of X

$$\bar{X} = -1.986$$

Estimate of a

$$\hat{a} = -.6530$$

Sample Mean of Y

$$\bar{Y} = 2.746$$

Estimate of b

$$\hat{b} = -1.711$$

Other Statistics

$$s = 0.1427$$

$$\sum x_i^2 = 29.663$$

$$s_X^2 = \frac{1}{n} \sum x_i^2 - (\bar{X})^2 = .2934$$

$$s_X = 0.5416$$

Equations 3.8, 3.9 and 3.10 can be used to obtain  $\hat{a}$ ,  $\hat{b}$ , and  $s$ , the estimates of  $a$ ,  $b$ , and  $\sigma$ , respectively.

$$\hat{a} = -0.6530 \quad \hat{b} = -1.7108 \quad s = 0.1427 \quad (3.12)$$

Least squares estimators  $\hat{\zeta}$  and  $\hat{C}$  are obtained by inverting Equation 3.7b and solving for  $\zeta$  and  $C$ .

$$\hat{\zeta} = 1/\hat{b} = 0.5845 \quad \hat{C} = 10^{-\hat{a}/\hat{b}} = .4152 \quad (3.13)$$

Least squares estimators  $\hat{a}$  and  $\hat{b}$  are obtained from Equation 3.7b

$$\begin{aligned} \hat{\zeta} &= 1/\hat{b} = 0.5845 \\ \hat{C} &= 10^{-\hat{a}/\hat{b}} = .4152 \end{aligned} \quad (3.13)$$

The least squares line,  $\Delta\epsilon = \hat{C}\hat{N}^{\hat{\zeta}}$  is plotted on Fig. 3.2.  $\hat{N}$  is defined as the estimate of the median  $\tilde{N}$ .

Note that (a)  $Y$  given  $X$  is normal and (b) the least squares line is the estimate of the mean of  $Y|X^*$ . Therefore, it follows that (a)  $N|\Delta\epsilon$  is lognormal and (b) the least squares line,  $N$ , is the estimate of the median of  $N|\Delta\epsilon$ .

#### 3.2.4 An Alternate Form

The form of the  $\Delta\epsilon$ - $N$  relationship given above is most commonly used in SRP literature. However, for probabilistic design purposes, it is more convenient to express the " $Y_0$  line" (Equation 3.7) as

$$N = K(\Delta\epsilon)^m \quad (3.14)$$

Comparing Equations 3.7 and 3.14 the constants  $K$  and  $m$  in terms of  $C$  and  $\zeta$  are

$$K = C^{-1/\zeta} \quad m = 1/\zeta \quad (3.15)$$

The transformation to linear form involves the substitutions,

$$Y = \log \Delta\epsilon \quad X = \log K \quad a = \log K \quad b = m \quad (3.16)$$

The least squares estimates of  $K$  and  $m$  are

$$\hat{K} = 10^{\hat{a}} \quad \hat{m} = \hat{b} \quad (3.17)$$

\*  $Y|X$  means the random variable  $Y$  given a specific value of the random variable  $X$ .

For the example of Fig. 3.2

$$\hat{k} = .2223 \qquad \hat{m} = -1.7108 \qquad (3.18)$$

The estimate of the median curve is

$$\hat{N} = 0.222(\Delta\varepsilon)^{-1.71} \qquad (3.19)$$

### 3.3 Statistical Distribution of the Least Squares Estimators, $\hat{a}$ and $\hat{b}$

The estimators  $\hat{a}$  and  $\hat{b}$  will have a bivariate normal distribution. The expected values are,

$$E(\hat{a}) = a \qquad E(\hat{b}) = b \qquad (3.20)$$

and the covariance matrix is

$$C(\hat{a}, \hat{b}) = (\sigma^2/n) \begin{bmatrix} 1 + (\bar{x}/s_x)^2 & -(\bar{x}/s_x^2) \\ -(\bar{x}/s_x^2) & 1/s_x^2 \end{bmatrix} \qquad (3.21)$$

where  $\bar{x}$  is the sample mean of  $X$  and

$$\begin{aligned} s_x^2 &= \frac{1}{n} \sum (x_i - \bar{x})^2 \\ &= \left( \frac{1}{n} \sum x_i^2 \right) - \bar{x}^2 \end{aligned} \qquad (3.22)$$

EXAMPLE: Consider the PP strainrange data of Saltsman and Halford (4).

Substituting appropriate values from Table 3.1 into Equation 3.21, including the assumption that  $s^2 = \sigma^2$ , it follows that an estimate of  $C(\hat{a}, \hat{b})$  is

$$C(\hat{a}, \hat{b}) = \begin{bmatrix} 0.042 & 0.020 \\ 0.020 & 0.010 \end{bmatrix} \qquad (3.23)$$

The variance of  $\hat{a}$  and  $\hat{b}$  are respectively 0.042 and 0.010 (standard deviations 0.205 and 0.100). The covariance of  $\hat{a}$  and  $\hat{b}$  of 0.020 indicates that  $\hat{a}$  and  $\hat{b}$  are correlated. The correlation coefficient

$$\rho = \frac{\text{Cov}(\hat{a}, \hat{b})}{\sigma_{\hat{a}} \sigma_{\hat{b}}} \quad (3.24)$$

$$= \frac{0.020}{(.205)(0.10)} = 0.976 \quad (3.25)$$

There exists a high degree of correlation between  $\hat{a}$  and  $\hat{b}$ . This dependency leads to operational difficulties when we attempt to present fatigue data summaries in a format useful for designers. Practical representations of these design factors will be the subject of Chapter 4.

### 3.4 Confidence Intervals for the Parameters and the Least Squares Line

Because  $\hat{a}$  and  $\hat{b}$  are estimates of  $a$  and  $b$ , they are random variables. For example, suppose that a sample of size  $n=7$  of AISI 316 PP data were collected in addition to that of Fig. 3.2. Because it is a random sample, a different value of  $\hat{a}$  and  $\hat{b}$  and a different least squares line will be obtained. Indeed each time a sample of size  $n$  is chosen, a different  $\hat{a}$ ,  $\hat{b}$  and therefore  $\hat{Y}_0$  will result. This point is important because the uncertainty in  $a$  and  $b$  has a significant effect on risk estimates.

The consequences of having  $\hat{a}$  and  $\hat{b}$  as correlated random variables is reflected in the character of the confidence intervals of the least squares line. Summarized in this section is how to compute confidence intervals for  $K$  and  $m$  as well as for  $Y_0(x)$ .

The AISI 316 PP data given in Fig. 3.2 and Table 3.1 is used as an example. Table 3.1 summarizes basic statistical analyses.

The  $1-\alpha$  confidence limits for  $a$  and  $b$  are (3)

$$\text{(For } a) \quad \hat{a} + t_{\alpha/2; n-2} s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{ns_x^2}} \quad (3.26)$$

$$\text{(For } b) \quad \hat{b} + t_{\alpha/2; n-2} s / (\sqrt{ns_x}) \quad (3.27)$$

where  $t_{\alpha/2; n-2}$  is the 100  $\alpha/2$  percentage point of Student's distribution with  $n-2$  degrees of freedom and using the data of Table 3.1, the 99% confidence limits ( $\alpha = .01$ ) are computed as follows

$$t_{.005, 5} = 4.032$$

$$\begin{aligned} \text{(For a)} \quad & -0.6530 \pm (4.032)(.1427) \sqrt{\frac{1}{7} + \frac{(-1.986)^2}{(7)(.2934)}} \\ & = -0.6530 \pm .8265 \\ & \quad (-1.479, +0.1735) \end{aligned} \tag{3.28}$$

$$\begin{aligned} \text{(For b)} \quad & -1.711 \pm (4.032)(.1427) / \sqrt{7(.5416)} \\ & = -1.711 \pm .401 \\ & \quad (-2.111, -1.310) \end{aligned} \tag{3.29}$$

The 99% confidence limits can be described by the probability statement,

$$P(-1.479 < a < 0.1735) = .99 \tag{3.30}$$

$$P(-2.111 < a < -1.310) = .99 \tag{3.31}$$

Because  $K$  and  $m$  are monotonic functions of  $a$  and  $b$ , the confidence limits can be evaluated by direct substitution. Results are given in Table 3.2.

Table 3.2

Confidence Limits for  $K$  and  $m$  (AISI 316 PP Data)

	Least Squares Estimate	99% Confidence Limits	
		Lower	Upper
$m$	-1.711	-2.111	-1.310
$k$	0.223	.0331	1.491

Because much of the work in SRP has used the form of  $\Delta\epsilon = C(\Delta\epsilon)^\zeta$  it is of interest to consider the confidence limits of  $C$  and  $\zeta$  for the AISI 316 PP data. Results are summarized in Table 3.3.

Confidence limits on  $Y_0$ , the mean of  $Y$  given  $X$ , are computed using

$$\hat{Y}_0 \pm t_{\alpha/2; n-2} s \sqrt{\frac{1}{n} + \frac{(x-\bar{X})^2}{ns_x^2}} \tag{3.32}$$



For the AISI 316 PP strain data of Fig. 3.2 and Table 3.1, the 99% confidence intervals on  $Y_0$  are plotted in Fig. 3.3.

Table 3.3  
Confidence Limits for  $\zeta$  and  $c$  (AISI 316 PP Data)

	Least Squares Estimate	99% Confidence Limits	
		Lower	Upper
$\zeta$	-.585	-.437	-.763
$c$	.415	.0743*	1.208*

\* Values obtained from Equation 3.13 by noting that the correlation coefficient between  $\hat{a}$  and  $\hat{b}$   $\rho = -1$ . Upper limit on  $a$  was combined with lower on  $b$  and vice versa.

The purpose of the plot of Fig. 3.3 is to illustrate the fact that when dealing with a phenomena that has as much scatter as fatigue, sample sizes larger than  $n=7$  are required for a reliable prediction of the least squares line. Assuming that identical statistics ( $\hat{a}$ ,  $\hat{b}$ ,  $s$ ,  $\bar{X}$ ,  $\sqrt{s_x}$ ) resulted from experiments having different sample sizes, an illustration of the effect of sample size can be provided. Fig. 3.4 shows 99% confidence intervals for  $Y_0$  for  $n=7$ , 15 and 30.

For design purposes however, it is customary not to use the least squares curve to characterize the data, but rather a curve on the safe (lower) side of the data. This approach will be pursued in Chapter 4.

Figure 3.3

99% Confidence Interval for Least Squares Line  
[AISI 316 Data; after Saltzman and Halford (4)]

### PP TYPE STRAINRANGE

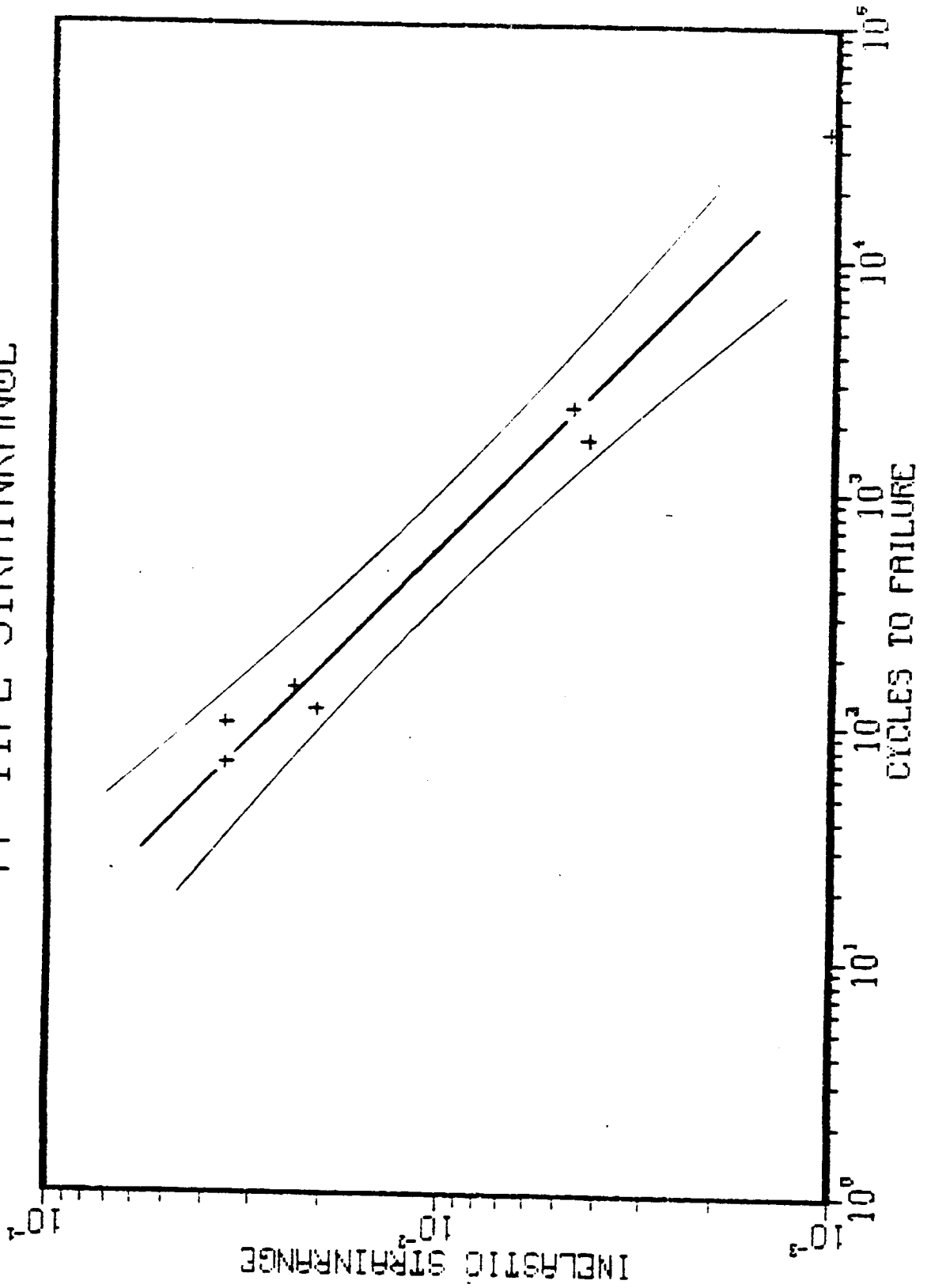
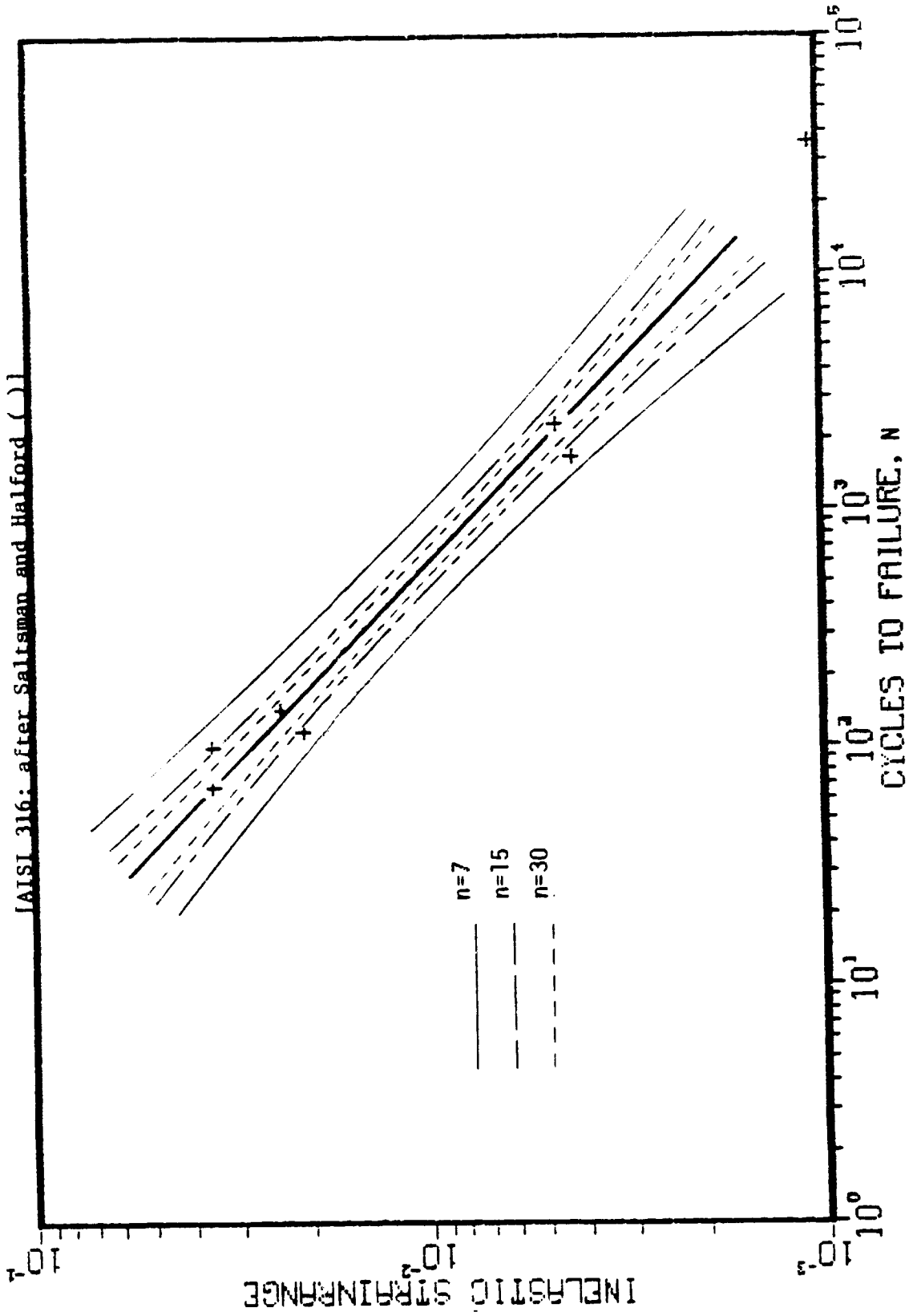


Figure 3.4  
 99% Confidence Intervals for the Least Squares Line  
 for N = 7, 15 and 30 Using Same Statistics

PP TYPE STRAINRANGE

[AISI 316: after Saltsman and Halford (1)]



### 3.5 References for Chapter 3.

1. Ang, A.H.S. and Tang, W.H., Probability Concepts in Engineering Planning and Design, Wiley, 1975.
2. Natrella, M.G., Experimental Statistics, NBS Handbook No. 91, 1963.
3. Bowker, A.H. and Lieberman, G.J., Engineering Statistics, Prentice-Hall, 1972.
4. Saltsman, J.F. and Halford, G.R., "Application of Strainrange Partitioning to the Prediction of Creep Fatigue Lives of AISI Types 304 and 316 Stainless Steel", ASME Paper No. 75-PVP-26, 1976.

#### 4.1 General Considerations for Design

The general problem of analysis of fatigue data is illustrated by Figure 4.1. The data typically will be curved even on a log-log plot and will have a scatter (standard deviation of  $N$  given  $\Delta\epsilon$ , denoted as  $N|\Delta\epsilon$ ) which increases with smaller  $\Delta\epsilon$ .

A median curve can be established through the data, typically using a least squares method. But the median curve only defines the 50% point. A low risk design would require definition of a design curve on the safe side of the data as shown in Figure 4.1.

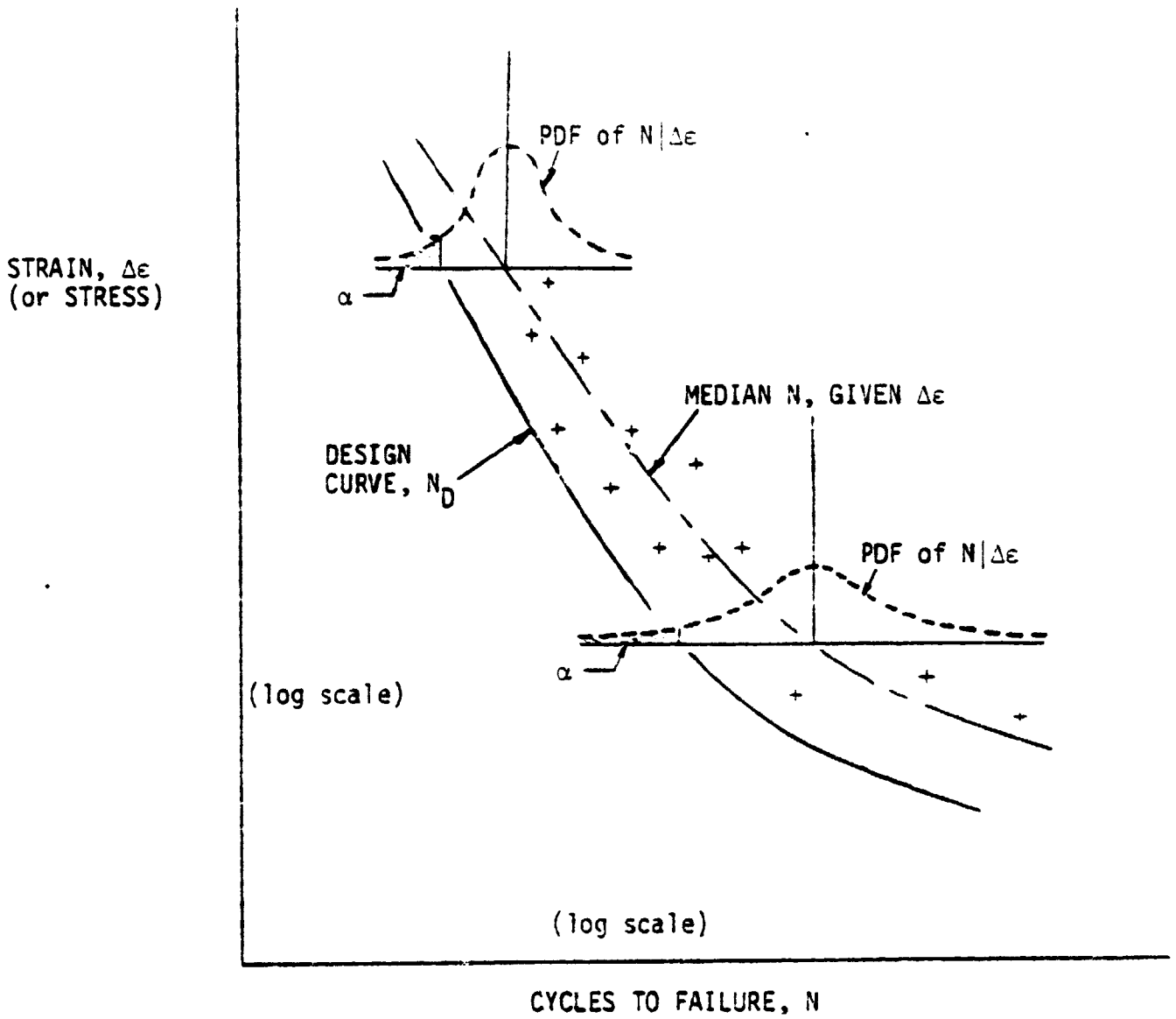
The design curve of Figure 4.1, denoted as  $N_D$ , is defined as a curve below which one would expect the next data point to fall with probability  $\alpha$ . Given that  $N$  is a random variable denoting cycles to failure,  $N_D$  can be defined by the probability statement

$$P[N < N_D | \Delta\epsilon] = \alpha \quad (4.1)$$

for any  $\Delta\epsilon$ . This left tail area,  $\alpha$ , is the estimate of the fraction of points that would be expected to fall below  $N_D$ . It can also be thought of as a "risk level" although it would be equal to the probability of failure only in the case where all other design factors, including strain history, were exactly known and the factor of safety equal one. In practice, the designer must select the value of  $\alpha$ . Typical values which have been used range from  $\alpha = 0.1\%$  to  $5\%$ .

Figure 4.1

An Illustration of a Design Curve Based on Risk Level  $\alpha$



An approach using methods of modern probabilistic design theory does not require a design curve. However statistics of the design parameters must be given. For example, it will be shown later for a probabilistic fatigue design approach that required information from the linear  $\Delta\varepsilon$ -N model,  $N=K(\Delta\varepsilon)^m$ , includes the value of  $m$  and the median and coefficient of variation of  $K$ .

In summary, two approaches to fatigue design can be used. Each requires a different way of characterizing  $\Delta\varepsilon$ -N data but both require statistical analyses.

1. A deterministic approach in which a design curve ( $\Delta\varepsilon$ -N) is specified. All other design factors are treated as deterministic or constant.
2. A probabilistic approach in which statistics of the model parameters are stated. All design factors are treated as random variables.

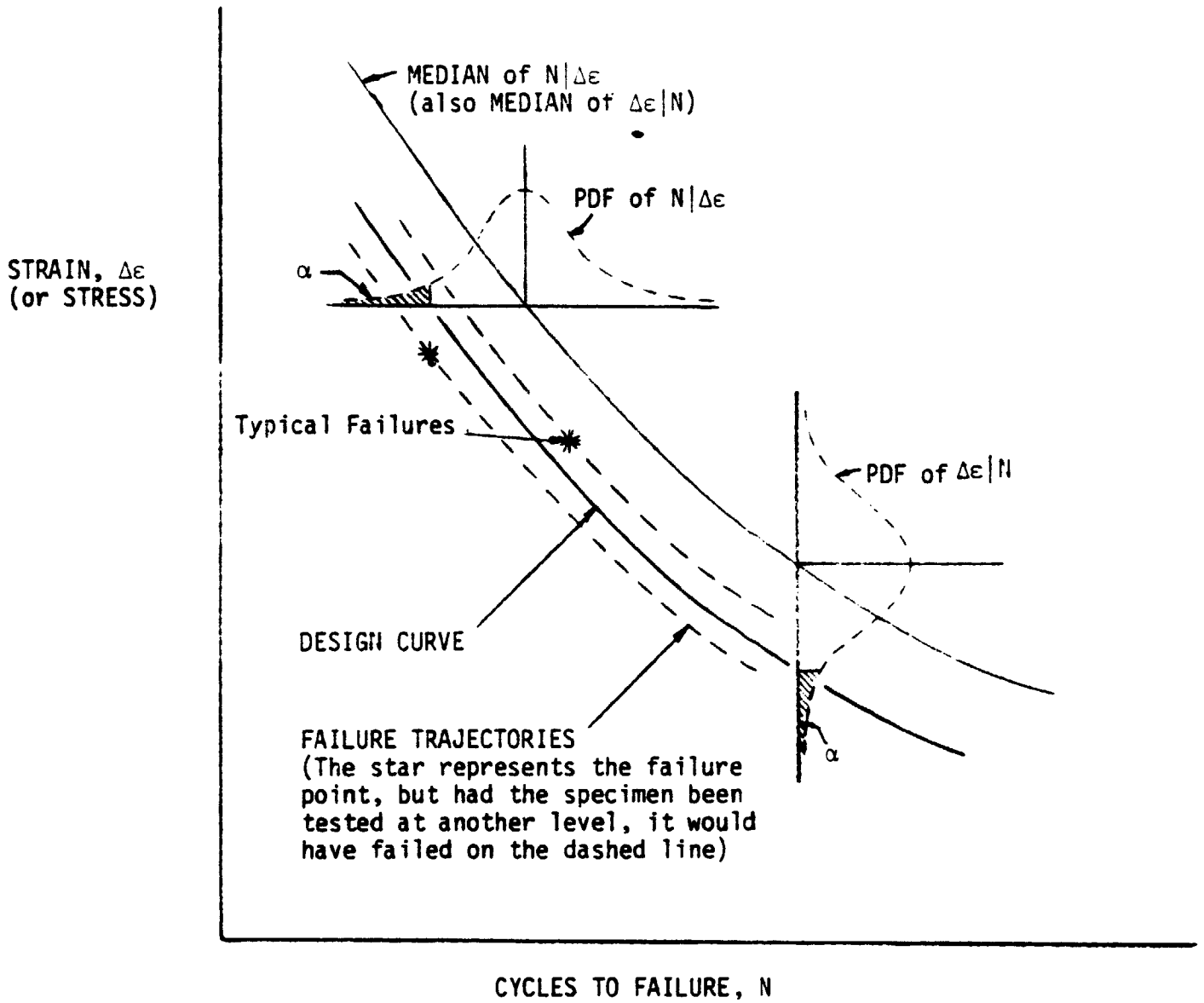
Both methods are described herein. This chapter summarizes the analysis for the simple linear model case. The following chapter explores this problem for the more difficult non-linear problem illustrated in Figure 4.1 in which the scatter is a function of strain level.

#### 4.2 Definition of a Failure Trajectory

The concept of a failure trajectory has been very useful in analyzing fatigue data for design purposes. In Figure 4.2, the star represents a specimen which has failed. The dashed lines (failure trajectories) are drawn through the point according to some predefined rule. Typically in this case the failure trajectories are parallel to the median curve. It is argued that the curve defines the cycles to failure of that specimen if it had been tested at a different stress level. For example, a specimen which is "weak" at a high stress level would also be weak at a low stress level.

Figure 4.2

Assumed Relationship Between the Distributions of  $N|\Delta\epsilon$  and  $\Delta\epsilon/N$





The assumption that failure trajectories accurately describe material behavior leads to many useful statistical descriptions that otherwise would not be possible. For example we can construct the distribution of  $\Delta\varepsilon$  given  $N$  as shown in Figure 4.2. The design curve defines the lower  $\alpha\%$  of failures in both directions. This statement cannot be made without the failure trajectory assumption.

As will be demonstrated in the following sections, not always will this concept be employed, but often it can be quite useful.

#### 4.3 Linear Life Relationships

It is assumed in this chapter that the strain-life relationship will be described by the basic linear model (Equation 3.3)

$$Y = a + bx + \delta \quad (4.2)$$

where  $\delta$  is a random variable having mean value of zero and standard deviation  $\sigma$ . Let

$$Y = \log N \quad x = \log S \quad a = \log K \quad m = b \quad (4.3)$$

Define the random variable  $K_0$  by the relation

$$\delta = \log K_0 \quad (4.4)$$

Combining Equations 4.2, 4.3, and 4.4 yields,

$$N = K_0 K (\Delta\varepsilon)^m \quad (4.5)$$

Because  $\delta \sim N(0, \sigma)$ , it follows that  $K_0$  is a lognormally distributed random variable. The median and coefficient of variation of  $K_0$  are for base 10 logs,

$$\tilde{K}_0 = 1 \quad c_{K_0} = \sqrt{10^{(\sigma^2/0.434)} - 1} \quad (4.6)$$

(See Appendix 1 for a description of the lognormal distribution.)

Note that because  $K_0$  is a random variable,  $N$  is a random variable. The median life,  $\tilde{N}$  is Equation 4.5 with the median value of  $K_0$ ;  $\tilde{K}_0=1$ . Thus

$$\tilde{N} = K(\Delta\epsilon)^m \quad (4.7)$$

In chapter 3 we saw how  $K$  and  $m$  are estimated by the least squares estimators  $\hat{K}$  and  $\hat{m}$  and that the least squares line

$$\hat{N} = \hat{K}(\Delta\epsilon)^{\hat{m}} \quad (4.8)$$

is an estimate of the median curve  $\tilde{N}$ .

For example in the strainrange partitioning (SRP) method, there are four such strain-life curves depending upon the strain type. See Manson, Halford, and Hirschberg (1), or Saltsman and Halford (2) for details on SRP.

$N_{pp} = G(\Delta\epsilon)^Y$	tensile plasticity, compressive plasticity
$N_{cp} = H(\Delta\epsilon)^\eta$	tensile creep, compressive plasticity
$N_{pc} = K(\Delta\epsilon)^\rho$	tensile plasticity, compressive creep
$N_{cc} = M(\Delta\epsilon)^\xi$	tensile creep, compressive creep

Least squares analysis in each of the four cases can be performed as described above. The median curve for each is simply the relationship with the least squares estimators.

#### 4.4 Uncertainties in the Linear Model

Upon employing the basic linear model with the least squares line for design, it is important to note that uncertainties stem from four sources (3)

1. Basic scatter inherent in the material behavior, as described by  $\sigma$
2. The sample standard deviation,  $s$ , and estimate of  $\sigma$ , is a random variable.
3. The estimate of  $a$  is  $\hat{a}$ , a random variable.
4. The estimate of  $b$  is  $\hat{b}$ , a random variable and is correlated to  $\hat{a}$ .

As the sample size  $n$  becomes large,  $\hat{a} \rightarrow a$ ,  $\hat{b} \rightarrow b$  and  $\hat{Y}_0 \rightarrow Y_0$ , and  $s$  approaches  $\sigma$ . For all practical purposes  $\hat{a}$ ,  $\hat{b}$  and  $s$  could be treated as constants for large  $n$  (typically the assumption would be reasonable for  $n > 50$ ). However, because of the expense associated with fatigue testing, sample sizes will generally be small. Therefore in general it is necessary to give full consideration to statistical distributions of these estimators.

Uncertainties in the estimators significantly complicate the development of a model suitable for design purposes. In the case that the data can be described by a linear model, the equivalent prediction interval described in this chapter provides an approximate solution.

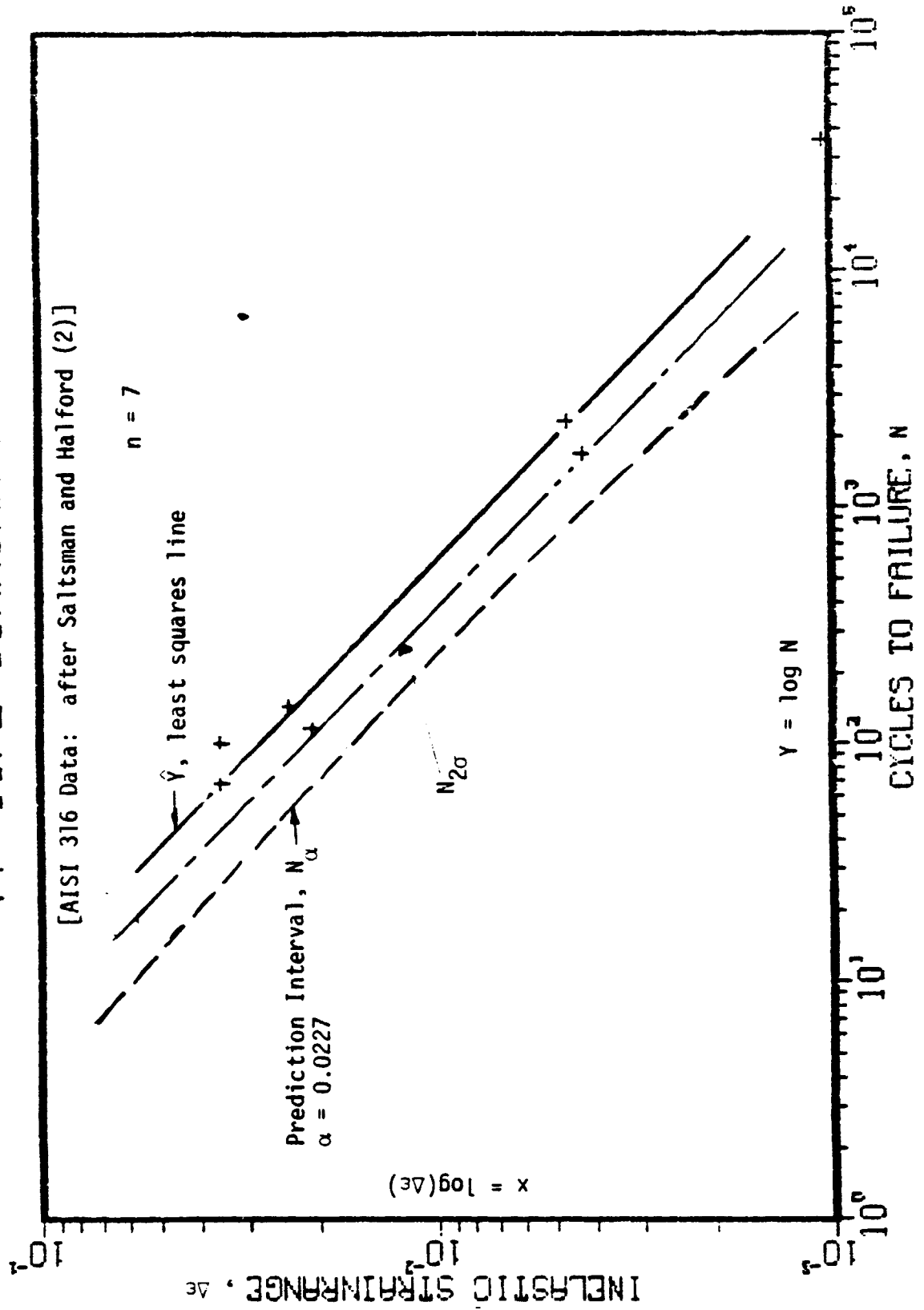
This chapter summarizes statistical methods of analysis commonly used for linear data and provides commentary regarding the performance of each.

#### 4.5 Lower $2\sigma$ or $3\sigma$ Limits

It is common practice to establish a design curve as a line which is parallel to the median  $\Delta\epsilon$ - $N$  curve and which lies a distance typically 2 or  $3\sigma$  to the left of the median curve. The value  $\sigma$  is taken to be equal to the sample standard deviation,  $s$  of Equation 3.10. Figure 4.3 illustrates a  $2\sigma$  curve, denoted as  $N_{2\sigma}$ .

Figure 4.3

# An Illustration of the $2\sigma$ Criterion PP TYPE STRAINRANGE



Use of such a criterion implies a risk level  $\alpha$  as follows:

Distance to the left of the Median*	Implied risk level, $\alpha$
$2\sigma$	0.0227
$3\sigma$	0.00135

\*Note that  $\sigma$ =standard deviation of  $Y=\log N$

This approach to a design curve can produce inconsistent levels of risk. Some of the pitfalls are illustrated in the following discussion and example.

On the basis of the available data, the statistical distribution of the next observation  $Y$  is described using the concept of prediction interval. The distribution of  $Y|X$  is Student's  $t$  with  $n-2$  degrees of freedom\*. To compute probabilities of  $Y$  for a given  $X$ , the 100  $\alpha\%$  lower prediction interval can be used (7,8)

$$\alpha = P[Y(x) \leq (\hat{a} + \hat{b}x) - t_{1-\alpha;n-2} s \sqrt{\frac{1+n}{n} + \frac{(x-\bar{X})^2}{ns_x^2}}] \quad 4.10$$

In Equations 4.10,  $s$  accounts for the scatter inherent in the physical phenomena, the term with the radical sign accounts for uncertainty in  $a$  and  $b$ , and the use of the Student's  $t$  distribution (rather than the normal) accounts for the uncertainty in  $s$ .

As an example, letting  $\alpha = 0.227$ , the same risk as implied by  $2\sigma$  curve results in the  $N_\alpha$  curve as shown in Figure 4.3. This curve given by

$$\log N_\alpha = \hat{Y} = t_{1-\alpha;n-2} s \sqrt{\frac{1+n}{n} + \frac{(x-\bar{X})^2}{ns_x^2}} \quad 4.11$$

\* e.g. See Reference 3 or 8.

is obtained from Equation 4.10. The  $N_\alpha$  curve is interpreted as follows. On the basis of the available data, there is 2.27% chance that the next point (Y,X) selected would fall below this curve. This illustration demonstrates that (a) the distribution of  $Y|X$  is a function of  $X$ , (b) because  $N_\alpha$  lies below  $N_{2\sigma}$  the use of the  $2\sigma$  criterion results in risks which are higher than advertised.

The  $3\sigma$  criterion is illustrated in Figure 4.4. The curve  $N_\alpha$  is for  $\alpha=0.00135$ . Again we see a significant difference between the  $N_{3\sigma}$  curve which implies a risk of  $\alpha$ , and the curve  $N_\alpha$  which actually defines  $\alpha$ .

However, as  $n \rightarrow \infty$ ,  $\hat{a} \rightarrow a$ ,  $\hat{b} \rightarrow b$ ,  $s \rightarrow \sigma$ , the estimators being consistent. Moreover,  $t_{1-\alpha;n-2} \rightarrow z_\alpha$  and the radical term of Equation 4.11  $\rightarrow 1$ . Therefore, for large  $n$ ,  $N_\alpha$  approaches  $N_{2\sigma}$ . In summary, the difference in  $N_\alpha$  and  $N_{2\sigma}$  of Figure 4.3 (and  $N_\alpha$  and  $N_{3\sigma}$  of Figure 4.4) is due to the uncertainty in the estimators.

Another way of looking at the problem is to consider the actual risk level  $\alpha$  as a function of sample size  $n$  for the  $2\sigma$  and  $3\sigma$  curves. The curves shown in Figure 4.5 illustrate that the actual risk level depends upon the sample size and can substantially differ from the implied risk level.

In summary, a  $2\sigma$  or  $3\sigma$  criterion can produce consistent probability levels in  $\Delta\epsilon$ -N data only if sample sizes are sufficiently large. As a rule of thumb, a large sample would typically be  $n \geq 50$ . For smaller samples, due consideration should be given to the fact that  $\hat{a}$ ,  $\hat{b}$  and  $s$  are random variables. Proposed in Sections 4.7 and 4.8 are models for consistent description of  $\Delta\epsilon$ -N data.

Figure 4.4  
 An Illustration of the  $3\sigma$  Criterion  
 PP TYPE STRAINRANGE

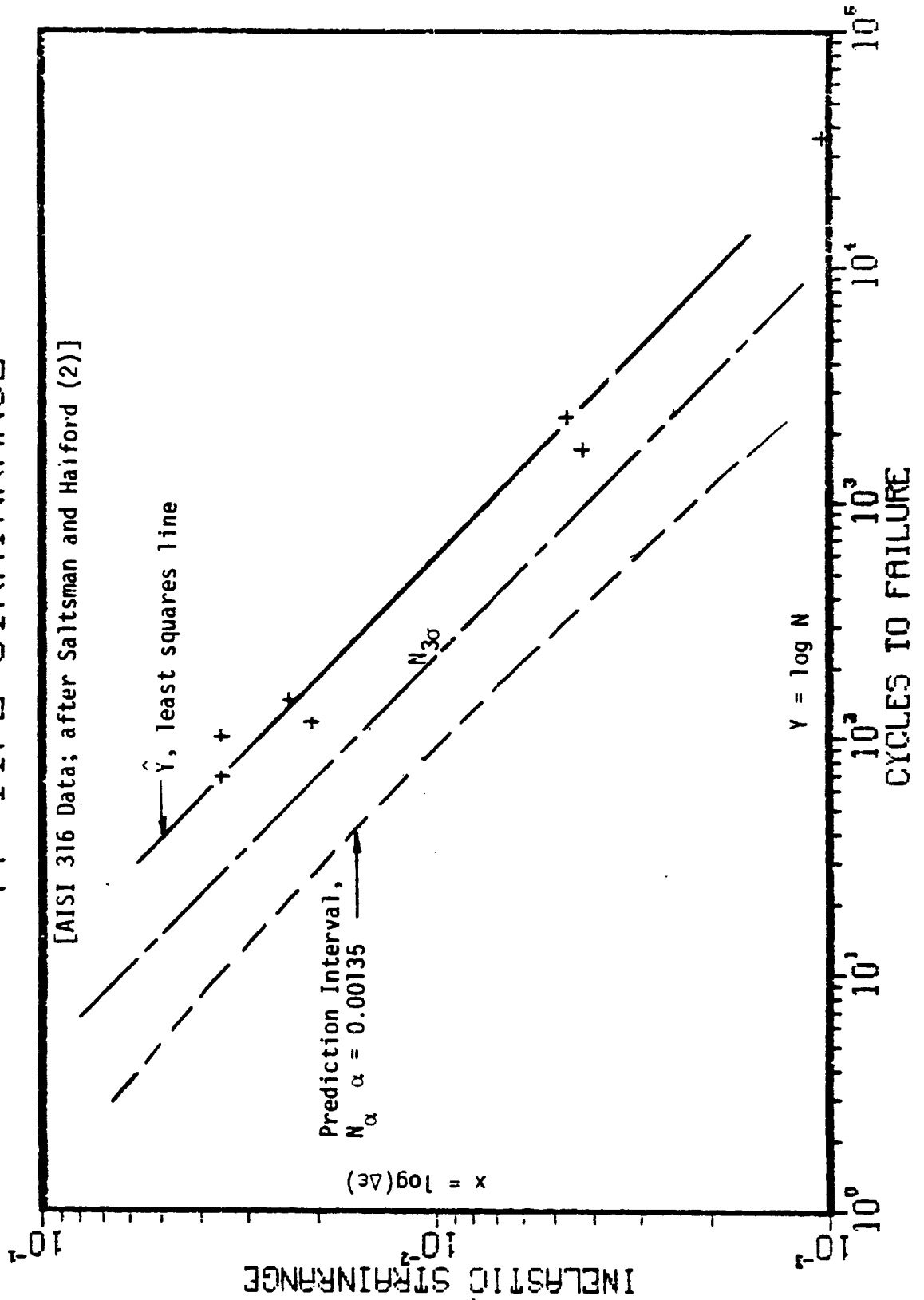
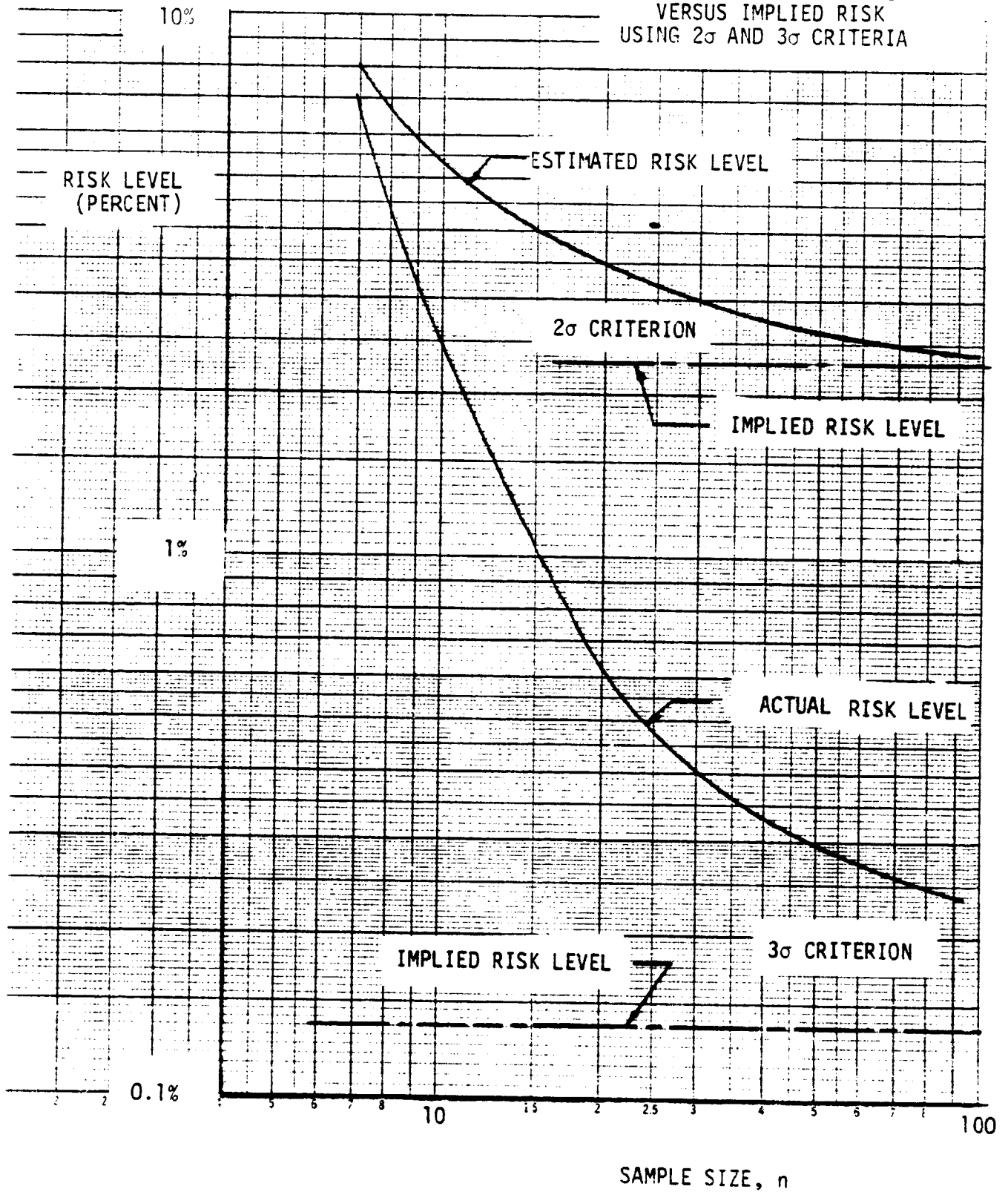


FIG. 4.5

COMPARISON OF ACTUAL  
VERSUS IMPLIED RISK  
USING  $2\sigma$  AND  $3\sigma$  CRITERIA





#### 4.6 Implied Risk in Fatigue Design Curves of the ASME Boiler and Pressure Vessel Code

Fatigue evaluation methods for pressure vessels are included in Section III, Division I and Section VIII, Division 2 of the ASME Boiler and Pressure Vessel Code (4, 5). Both sections contain fatigue design curves.

The fatigue design curves were developed by applying factors of 2 on the median strain range and 20 on the median life (6).

The median curve is the Langer model (7), having the form

$$S = BN^{-1/2} + S_e \quad (4.12)$$

where  $S$  is defined as

$$S = 1/2(\Delta\epsilon)E \quad (4.13)$$

$E$  being the elastic modulus.  $K$  and  $S_e$  are the parameters to be determined from the data.  $S_e$  can be interpreted as the endurance limit. The design curve is actually a lower bound envelope. See Figure 4.6. Past experience indicates that these factors satisfactorily account for possible size effects, data scatter, surface-finish influences, moderate environmental effects and time history effects on fatigue resistance\*.

The factor of 20 on life dominates the design curve for cycle lives typically less than 20,000 (6) as suggested by Figure 4.6. It is of interest to examine the actual risk implied by such a criterion.

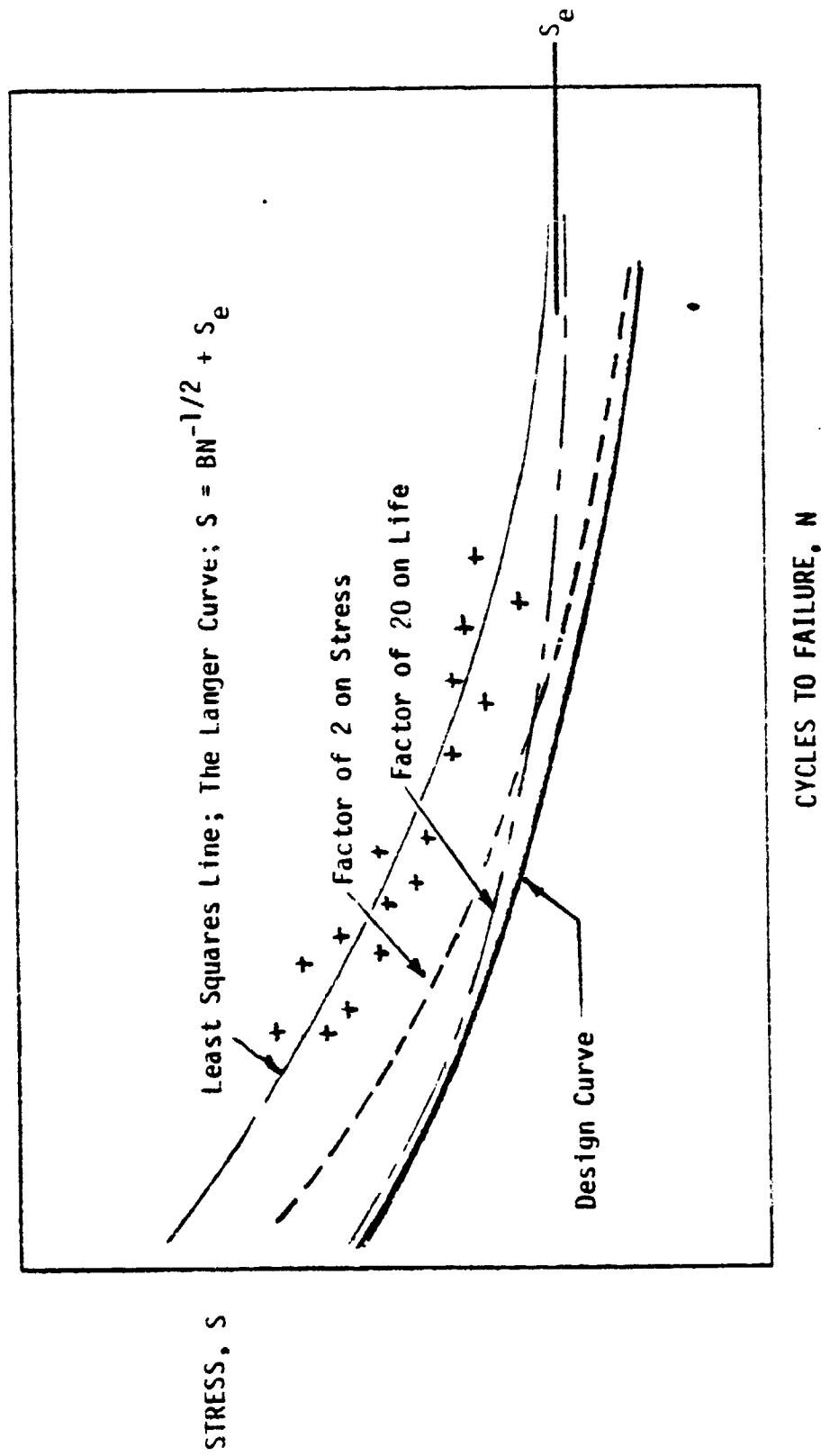
Figure 4.7 illustrates how the design curve,  $N_D$  and risk level  $\alpha$  are related. Note that  $\alpha$  depends on the amount of scatter in  $N$ , but  $N_D$  does not. For a narrow distribution of  $N|\Delta\epsilon$ ,  $\alpha$  is small and vice versa.

The relationship between  $\alpha$  and the coefficient of variation of  $N$ , denoted as  $C_N$ , can be derived as follows

\* For an interesting commentary on the ASME fatigue curves, see the article "Safety Factors in Fatigue Design; Arbitrary or Rational", by J.T. Fong and J.H. Smith in Critical Issues in Materials and Mechanical Engineering, ASME, 1981.

Figure 4.6

How the Design Curve is Established in the ASME



# 2 OF 3

N82-14531

UNCLAS

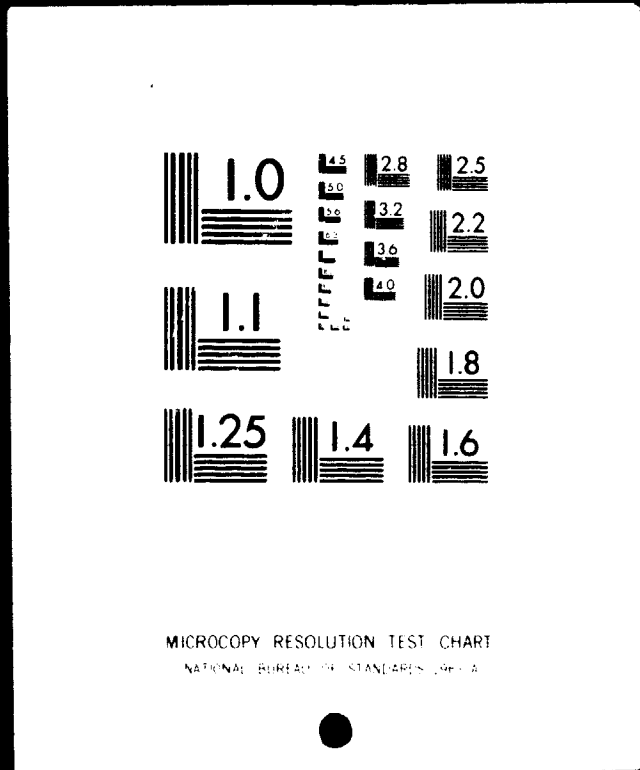
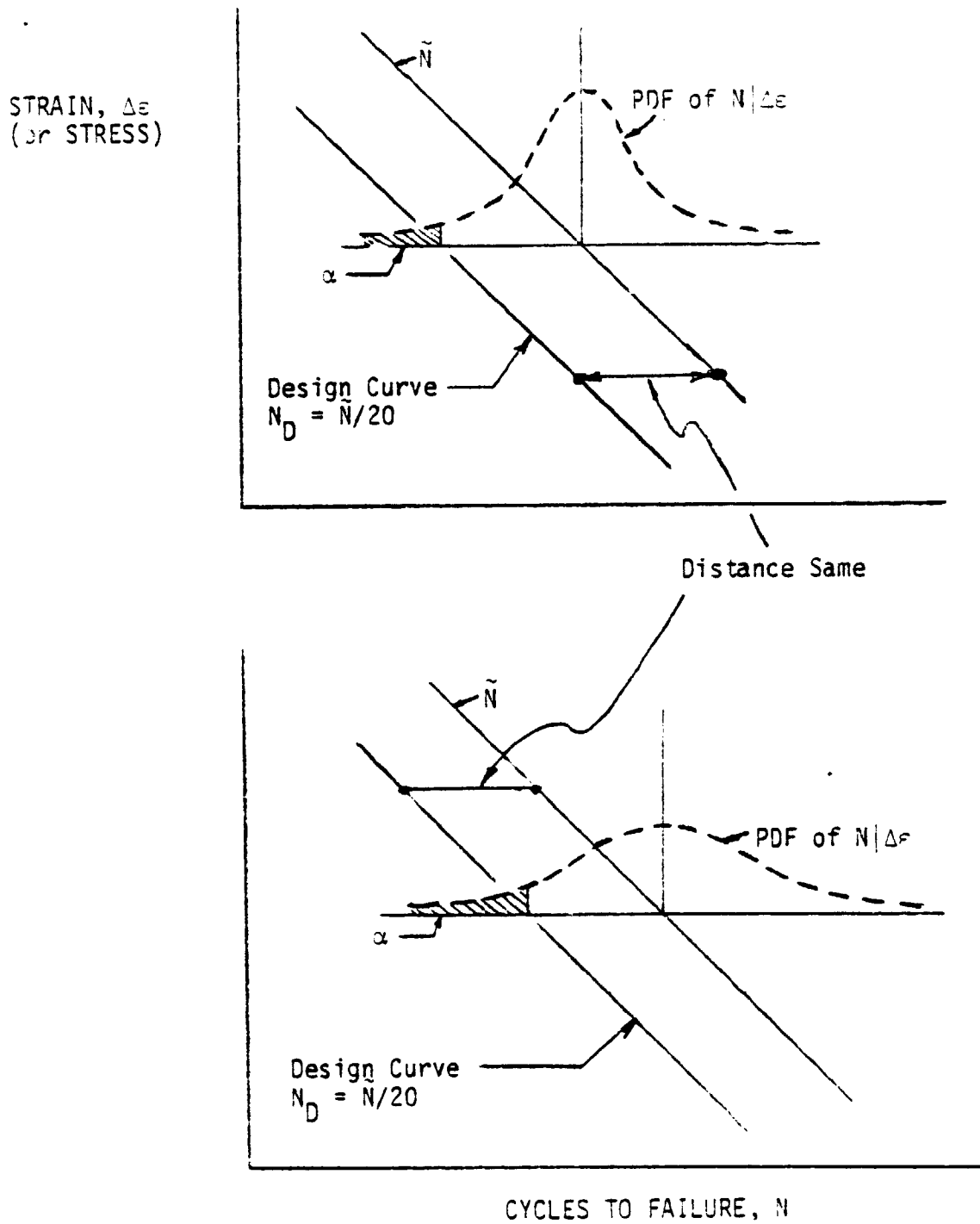


Figure 4.7

Risk Level as Defined by the  $\tilde{N}/20$  Curve for Two Different Distributions of  $N$



1-2

$$\begin{aligned}\alpha &= P[N < N_D] \\ &= P[N < N/20]\end{aligned}\tag{4.14}$$

Assuming that  $N$  has a lognormal distribution (see Appendix 1), it follows that

$$\alpha = \Phi(-\log(20)/\sigma_{\log N})\tag{4.15}$$

where

$$\sigma_{\log N} = \sqrt{0.434 \log(1 + C_N^2)}\tag{4.16}$$

and where  $\Phi(\cdot)$  is the standard normal distribution function.

In Figure 4.8 the risk level  $\alpha$  is plotted as a function of  $C_N$  for values of  $C_N$  commonly observed in fatigue test data. Smooth specimen low cycle data typically have  $C_N$  from 0.30 to 0.70, but welded joint data for example can have values of  $C_N$  as high as 1.5.

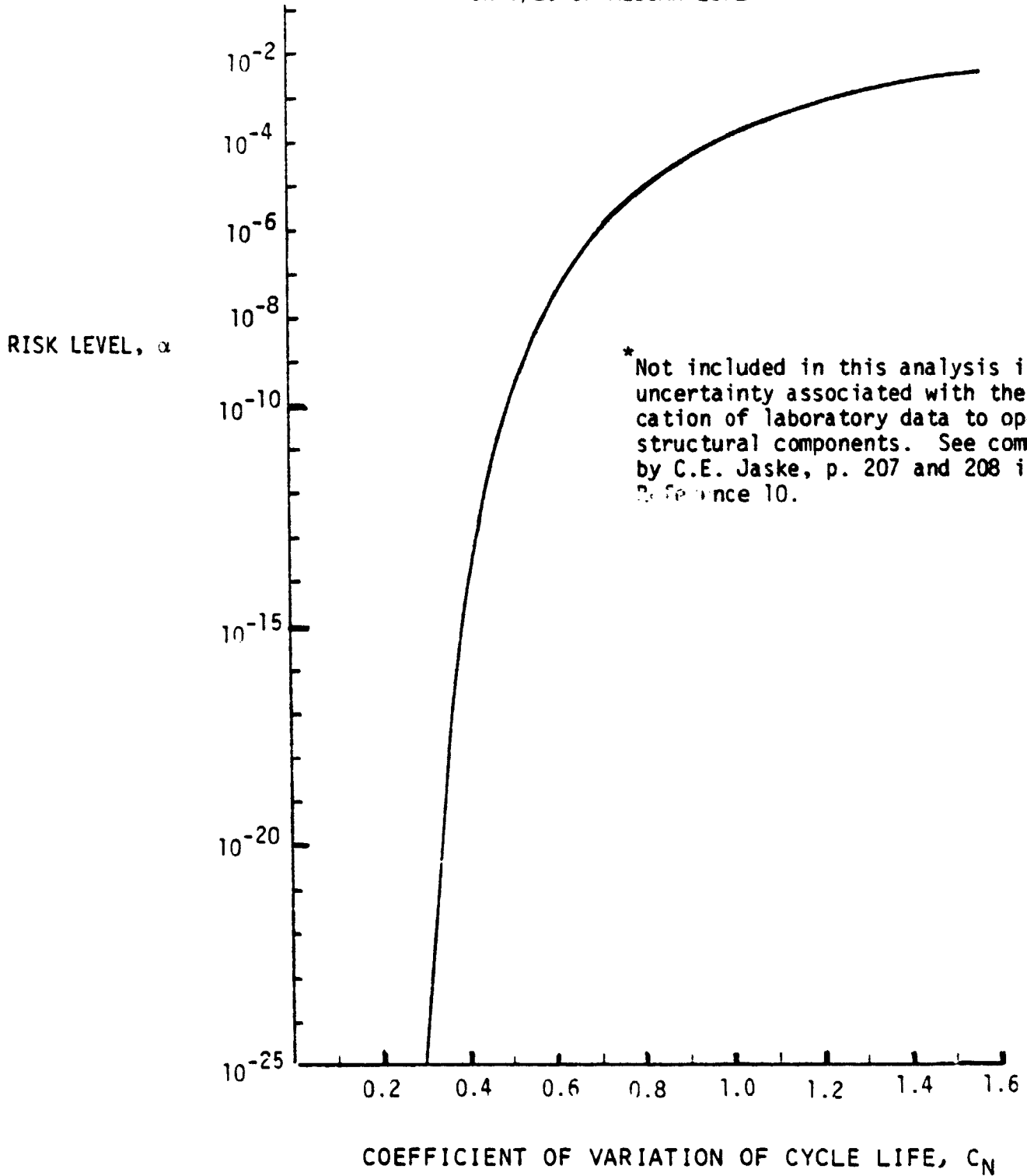
In summary, the plot of Figure 4.8 clearly shows that application of the "factor of 20 on life" curve leads to inconsistent levels of risk.

#### 4.7 The Tolerance Interval Used to Establish a Design S-N Curve

Examples above illustrated how certain rules for drawing design curves can lead to inconsistent levels of risk between data sets. Two methods are presented herein for constructing design curves which are statistically consistent. They are a) the tolerance interval method described in this section, and b) the equivalent prediction interval (EPI) described in Section 4.8. The EPI method can also provide appropriate statistics for reliability analysis.

FIG. 4.8

RISK LEVEL AS A FUNCTION OF  
COEFFICIENT OF VARIATION OF  
LIFE FOR A DESIGN CURVE BASED  
ON 1/20 OF MEDIAN LIFE



The tolerance limit is established for a single variable as described in Chapter 2. However, it may be possible to use the tolerance limit to define a design curve from SN data. It is necessary, however, to assume the existence of failure trajectories, lines parallel to the least squares line are illustrated in Figure 4.9. Thus SN data can be translated into an "equivalent" random sample of  $N$  for any  $S$ . The standard deviation of  $Y = \log N$  is constant, with the mean of  $Y$  defined by the least squares line.

The tolerance limit at any stress level will lie to the left of the least squares line  $\hat{Y}$ , a distance of  $K_{\alpha, \gamma} s$ . Thus the design curve is

$$N_D = \hat{Y} - K_{\alpha, \gamma} s \quad (4.17)$$

An illustration of the design curve based on a tolerance interval is shown in Figure 4.10.

The PP data of Reference 2 is used to provide an example of a design curve based on the tolerance interval. The data is shown in Figure 4.11. Assume that the decision has been made to define the design curve as the line below which no more than  $\alpha=1\%$  of the population is expected to fall with confidence  $\gamma=95\%$ ;  $n=7$ . From Table 2.10  $K_{\alpha, \gamma} = 4.64$ . The design curve is

$$\begin{aligned} \hat{Y} - K_{\alpha, \gamma} s &= \hat{Y} - 4.64(0.1427) \\ &= \hat{Y} - 0.662 \end{aligned} \quad (4.18)$$

The scatter factor is

$$\begin{aligned} z &= 10^{K_{\alpha, \gamma} s} = 10^{0.662} \\ &= 4.59 \end{aligned} \quad (4.19)$$

In summary, the concept of the tolerance interval can be used to establish rational and consistent design SN curves. It is required to assume that failure trajectories define material behavior. Moreover, the tolerance interval does

Figure 4.9

The Use of Failure Trajectories to Establish an Equivalent Sample of N at any Stress Level

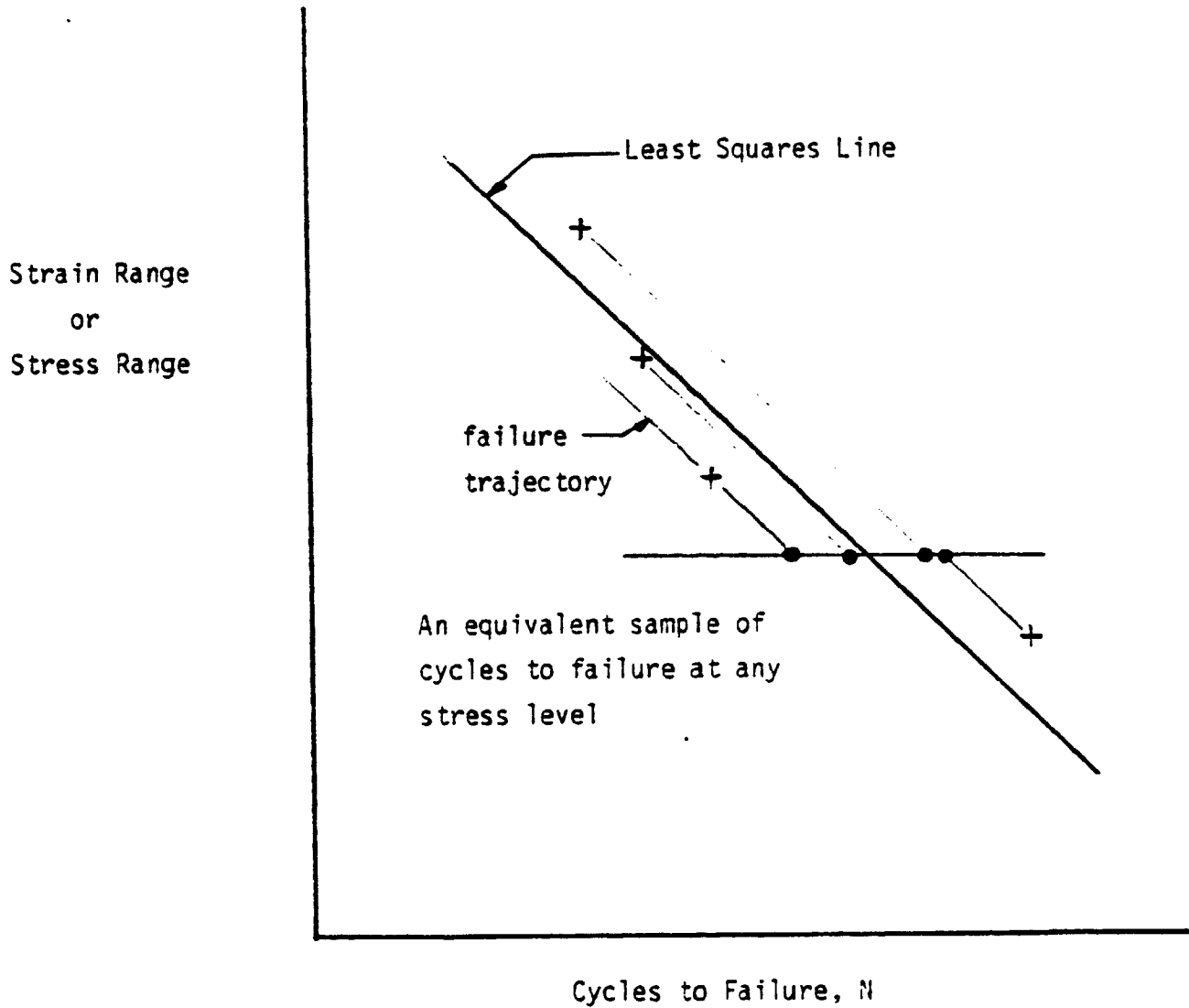




Figure 4.10

Design Curve for SN Data

Based on Tolerance Interval

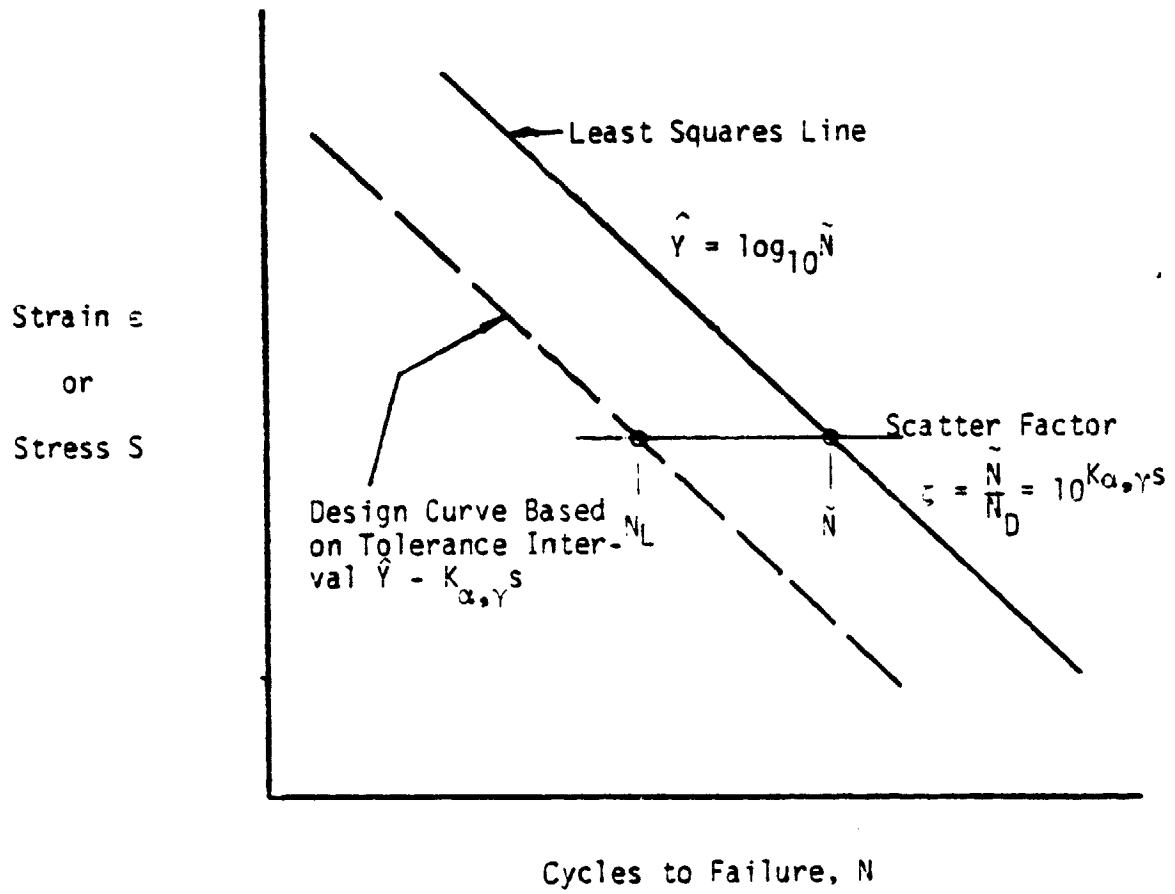
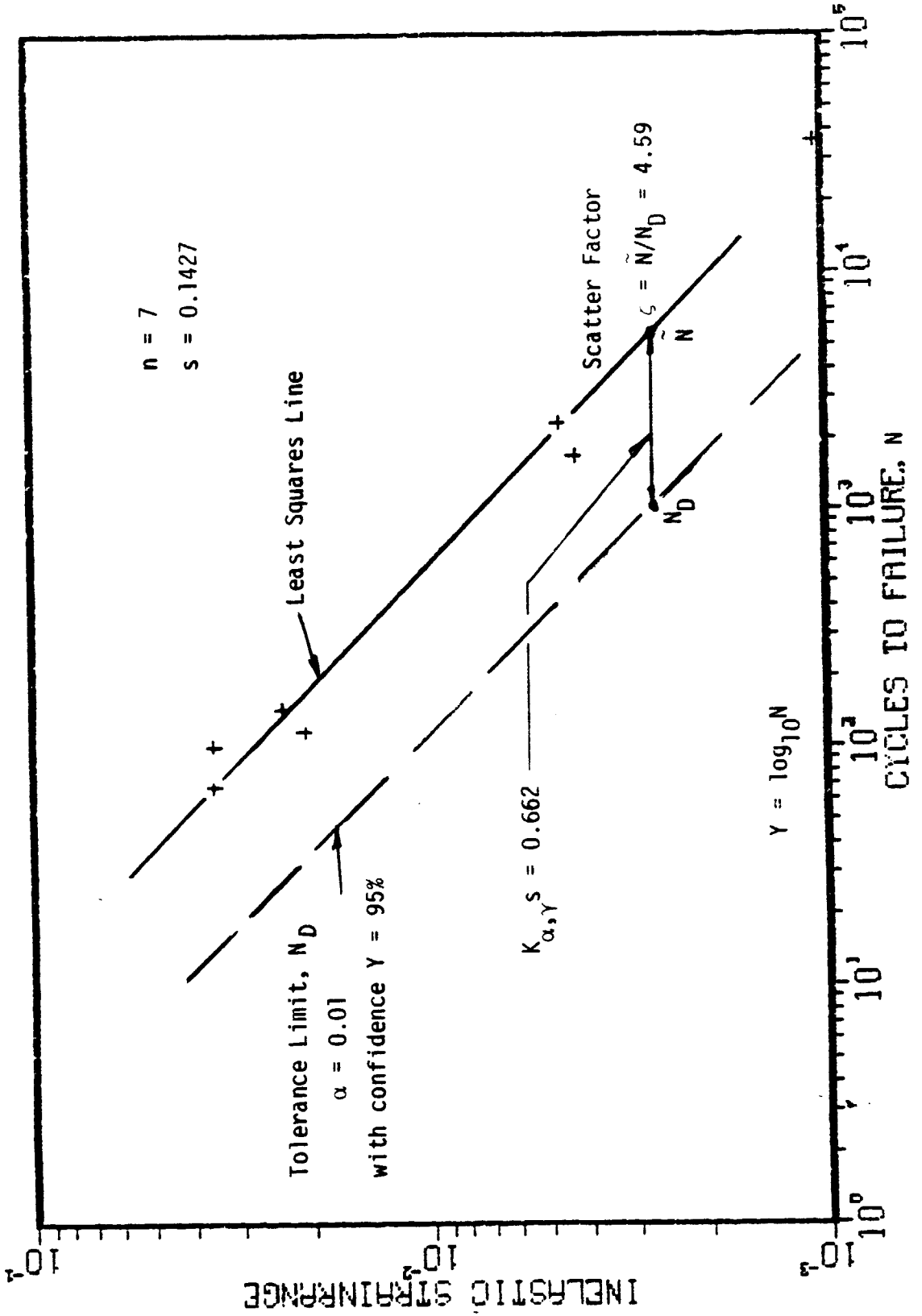


Figure 4.11

Use of the Tolerance Interval to Establish Design Curve [after Saltzman and Halford (2)]

# PP TYPE STRAINRANGE



not directly provide the information for a probabilistic design approach. These restrictions do not apply to the use of the equivalent prediction interval described in the next section.

#### 4.8 A Consistent Method for Characterizing Linear Fatigue Data - The Equivalent Prediction Interval (EPI)

The goal of the analysis described in this section is to develop a procedure of fatigue data analysis having the following properties: (a) scatter in data is represented, (b) consideration given to the statistical distributions of the estimators, (c) probability estimates from the model are consistent and accurate. (d) appropriate statistics for a probabilistic design analysis are provided and (e) the model is easy to use. The equivalent prediction interval proposed herein satisfies these requirements.

Assume that the strain-life relationship is of the form

$$N = K(\Delta\varepsilon)^m \quad (4.20)$$

For a given  $\Delta\varepsilon$ ,  $K$  and  $m$  (and therefore  $N$ ) are random variables by virtue of the scatter in the data. The value of  $K$  in Equation 4.20 is now equivalent to  $K_0 K$  in Equation 4.5.

In a design procedure, two strategies exist.

1. A deterministic approach in which a "design curve" is established on the safe side of the data. Assume that  $Y|X$  is normal with standard deviation of  $s$ . For example, to establish a line below which only 1% of the points would be expected to fall, let  $\alpha=.01$  and  $z_{1-\alpha}=2.33$  from normal tables. The "design curve" is given as

$$\log N_D = \hat{Y} - z_{1-\alpha} s \quad (4.21)$$

$N_D$  is shown in Figure 4.12 (see also Figure 4. ). Use of this as a design curve is subject to the pitfalls as described in Section 4.5 and Section 4.6.

But this approach insures a lower 1% bound only in the limit as  $n \rightarrow \infty$ . For small  $n$ , typical of fatigue tests, this method is unconservative as described below.

2. A probabilistic approach in which  $\hat{a}$  and  $\hat{b}$  (and therefore  $m$  and  $K$ ) are treated as random variables. Standard probabilistic design methods are employed to insure structural integrity. But  $\hat{a}$  and  $\hat{b}$  are bivariate normal (3), and therefore  $m$  and  $K$  appear in the failure function as correlated random variables. The problem becomes complicated.

The distribution of a future observation, defined by the prediction interval is described in Section 4.5. The lower 1% prediction interval is also shown in Figure 4.12. Again, we see that use of the  $N_D$  curve is misleading in terms of the implied risk.

It would be difficult to incorporate the notion of a curved prediction interval into a design algorithm. However, it is observed that the equal probability  $N_\alpha$  curve is relatively flat. This suggests that prediction intervals can be approximated as straight lines, called herein, equivalent prediction intervals, and abbreviated EPI.

Development of the EPI concept is described by Wirsching and Hsieh (8) and summarized in following discussions.

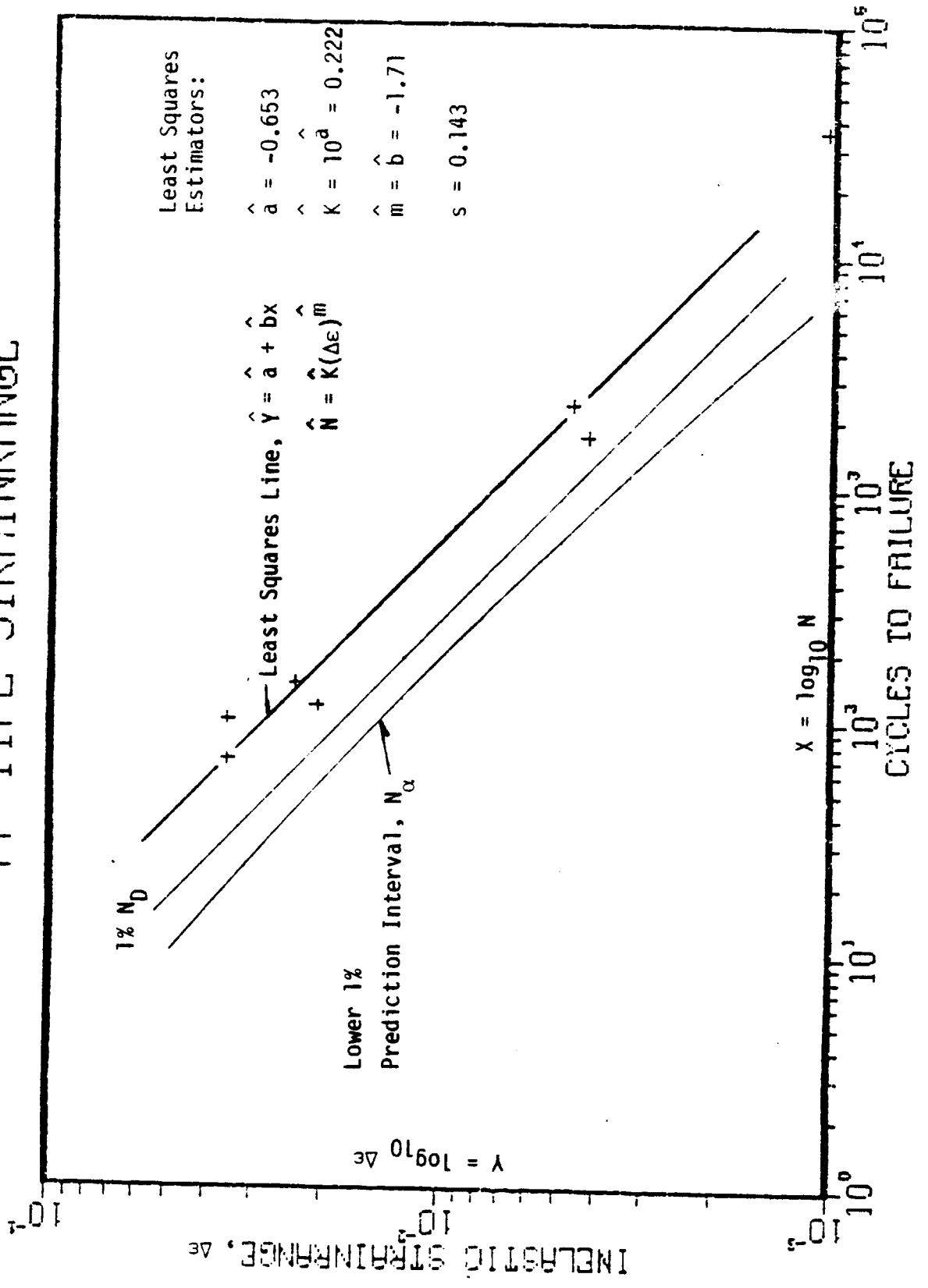
Define an equivalent constant standard deviation of  $Y$  as

$$\sigma_0 = s g(n, \alpha) \quad (4.22)$$

Figure 4.12

The 1%  $N_D$  Curve and the Lower 1% Prediction Interval [after Saltzman and Halford (2)]

# PP TYPE STRAINRANGE



where

$$g(n, \alpha) = \exp[A(\alpha) \{ \ln n \}^{-B(\alpha)}] \quad (4.23)$$

$$A(\alpha) = 1.55 \left[ \frac{1}{2} \ln \left( \frac{2-\alpha}{\alpha} \right) \right]^{1.12} \quad (4.24)$$

$$B(\alpha) = 3.32 - 1.7\alpha$$

$$6 \leq n \leq 50; 0.01 \leq \alpha \leq 0.15$$

Basically the idea is to use the model of Equation 4.2, but with an expanded value of  $s$  (the value of  $\sigma_0$  defined below) so that the  $N_D$  curve is shifted to the left to match the  $N_\alpha$  curve.

$g(n, \alpha) \geq 1$  is in essence, an adjustment factor to  $s$  to account for the fact that there is uncertainty in the estimates of  $a$  and  $b$  and  $s$ . For convenience, a chart of  $g(n, \alpha)$  is given in Figure 4.13.

The model, suggested by the above discussion, is as follows:

1. Let  $m = \hat{b}$  be a constant.
2. Assume that all of the uncertainty due to scatter in the data is accounted for in  $\hat{a}$  (and therefore  $K$ ) by considering the  $y$  intercept as a random variable.
3. Therefore, let the empirical relationship be

$$Y = a_0 + \hat{b}x \quad (4.25)$$

where  $a_0$  has a normal distribution with mean  $\hat{a}$  and standard deviation  $\sigma_0$ .

The consequences of such a model are

1.  $Y|X$  has a normal distribution. (Thus  $N$  given  $\Delta\varepsilon$  has a lognormal distribution)
2. The mean value of  $Y|X$  is  $\hat{a} + \hat{b}x$ . (Thus the estimate of the median of  $N$  is  $\hat{N} = \hat{K}(\Delta\varepsilon)^m$ )
3. The standard deviation of  $Y|X$  is  $\sigma_0$  (and is not a function of  $X$ ).
4.  $\hat{a}_0 = \log K$  is normal and  $K$  is lognormal. The median  $\tilde{K}$  and coefficient of variation  $C_K$  of  $K$  can be obtained from the lognormal (base 10) forms

$g(\alpha, n)$

1.7

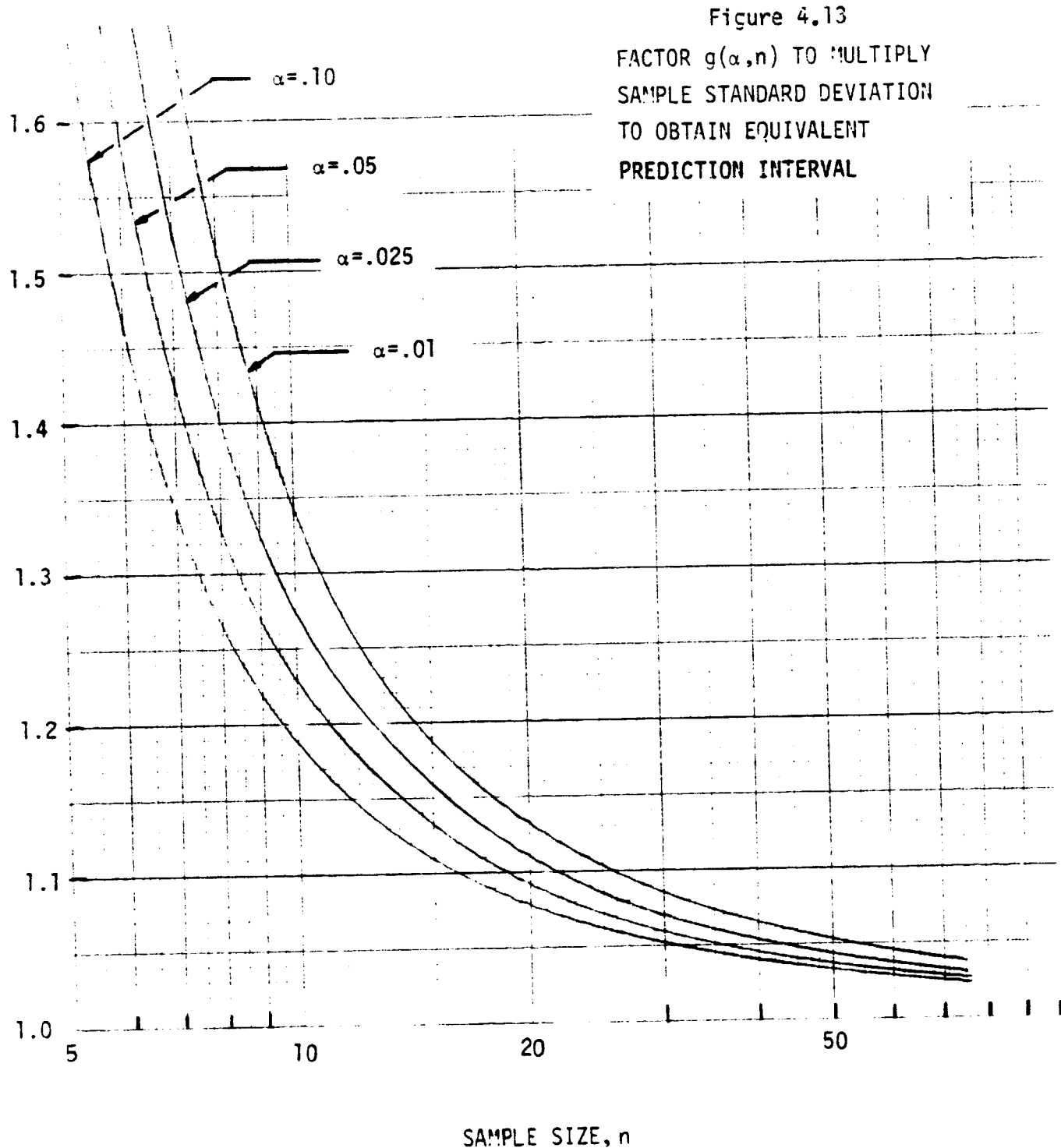


Figure 4.13  
FACTOR  $g(\alpha, n)$  TO MULTIPLY  
SAMPLE STANDARD DEVIATION  
TO OBTAIN EQUIVALENT  
PREDICTION INTERVAL

$$\hat{K} = 10^{\hat{a}} \quad (4.26)$$

$$C_K = \sqrt{10^{(\sigma_0^2 / .434)} - 1} \quad (4.27)$$

**EXAMPLE** Given the fatigue data ( $n = 7$ ) as illustrated in Figure 4.12

define a design S-N line which is estimated to be on the safe side of 99% of the data. This line is to be the  $\alpha = 0.01$  EPI. The basic data is summarized in Table 3.1. The estimators are presented in Figure 4.12.

Using Equation 4.23 with  $n = 7$ ,  $\alpha = 0.01$

$$A(\alpha) = 4.64 \quad B(\alpha) = 3.30 \quad g(n, \alpha) = 1.67 \quad (4.28)$$

Thus, the equivalent standard deviation is,

$$\begin{aligned} \sigma_0 &= g \cdot s \\ &= (1.67)(0.143) = 0.239 \end{aligned} \quad (4.29)$$

**Design Curve** The 1% EPI is given as  $N^*$ , where

$$\log N^* = \hat{Y} - z_{\alpha} \sigma_0 \quad (4.30)$$

This EPI could be used as the design curve in the conventional approach. The EPI is shown in Figure 4.14 along with the prediction interval  $N_{\alpha}$  which the EPI approximates.

Figure 4.14 suggests that the EPI is a reasonable approximation to  $N_{\alpha}$  the prediction interval. As the sample size becomes larger,  $N_{\alpha}$  becomes flatter and the EPI becomes an even better approximation [7].

**Probabilistic Format** The data will be analyzed in a format which is convenient for probabilistic design procedures.

The fatigue equation will contain  $m$  and  $K$ . Using the method described above,  $m$  is a constant and equal to

$$\hat{m} = \hat{b} = -1.71 \quad (4.31)$$

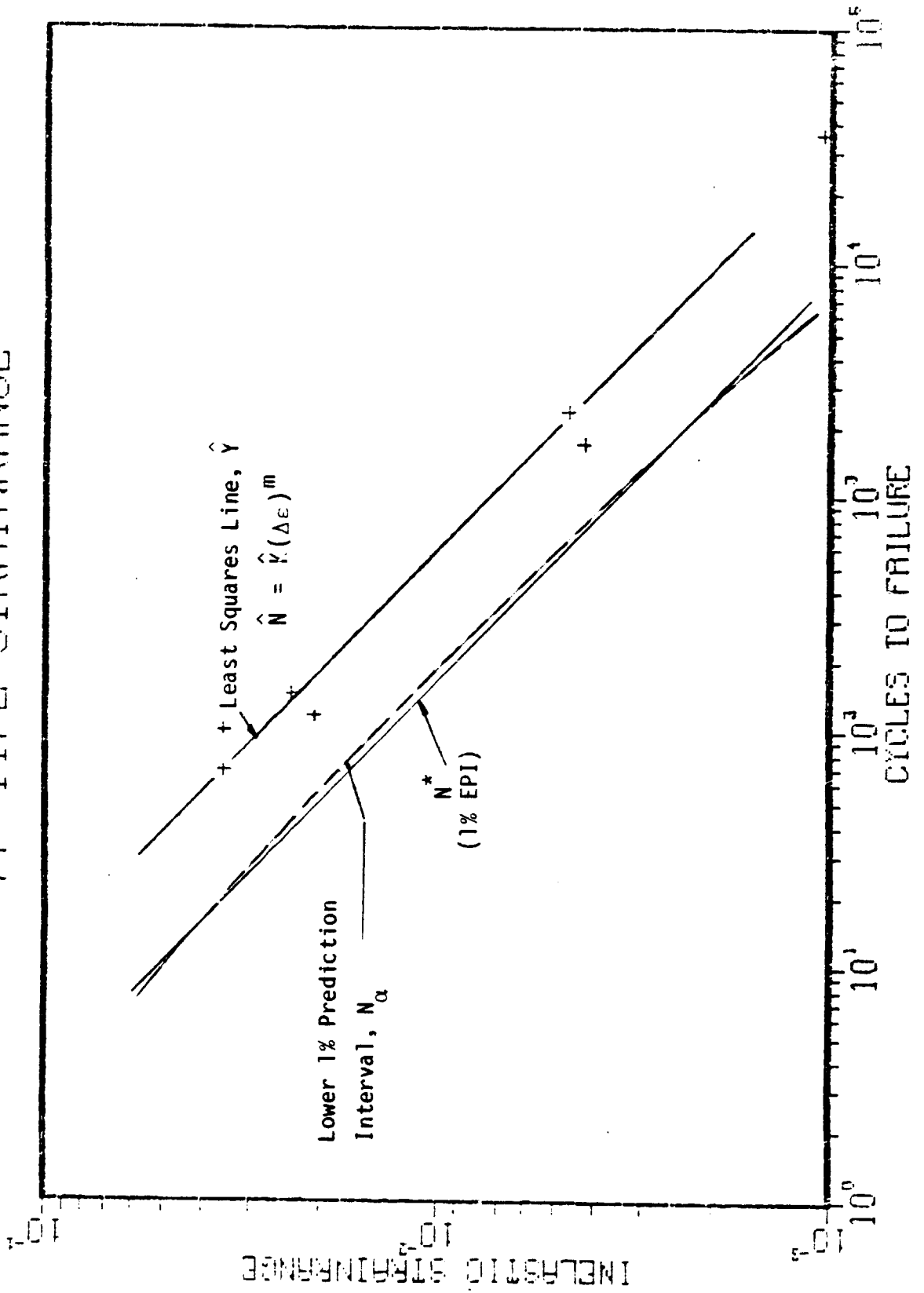


Figure 4.14

Comparison of the Lower 1% Prediction Interval with the Equivalent Prediction Interval

[after Saltzman and Halford (2)]

# PP TYPE STRAINRANGE



K will be lognormal with a median value of (Equation 4.26)

$$\tilde{K} = 10^{-.6530} = .222 \quad (4.32)$$

K will have a coefficient of variation of (Equation 4.33) based on  $\alpha = 0.01$

$$C_K = \sqrt{10^{(.239)^2 / .434} - 1} = 0.595 \quad (4.33)$$

where  $\sigma_0$  was obtained from Equation 4.27.

In summary, Chapter 4 discussed the problem of establishing a design curve and/or appropriate statistics for the case where strain-life data has a linear trend (on a log-log basis) and uniform scatter. Analysis associated with more complicated strain life models are treated in Chapter 5.

#### 4.9 References for Chapter 4

1. Manson, S.S., Halford, G.R. and Hirschberg, M.H., "Creep-Fatigue Analysis by Strainrange Partitioning", NASA TM X-67838.
2. Saltsman, J.F., and Halford, G.R., "Application of Strainrange Partitioning to the Prediction of Creep Fatigue Lives of AISI Types 304 and 316 Stainless Steel", ASME Paper No. 76-PVP-26, 1976.
3. Mood, A.M., and Graybill, F.A., Introduction to the Theory of Statistics, McGraw-Hill, 1963.
4. "Section III, Division I - Nuclear Power Plant Components, Subsection NA-General Requirements", ASME Boiler and Pressure Vessel Code, ASME, New York, 1974.
5. "Section VIII, Division II - Alternative Rules", ASME Boiler and Pressure Vessel Code, ASME, New York, 1974.
6. Jaske, C.E. and O'Donnell, W.J., "Fatigue Design Criteria for Pressure Vessel Alloys", Journal of Pressure Vessel Technology, ASME, Nov. 1977.
7. Langer, B.F., "Design of Pressure Vessels for Low Cycle Fatigue", Journal Basic Engineering, ASME, Vol. 84, No. 3, Sept. 1962.
8. Wirsching, P.H. and Hsieh, S., "Linear Model in Probabilistic Fatigue Design", Journal of the Engineering Mechanics Division, ASCE, Vol. 106 No. EM6, Dec. 1980.
9. Bowker, A.H. and Lieberman, G.J., Engineering Statistics, 2nd Edition, 1972.
10. Critical Issues in Materials and Mechanical Engineering, J.T. Fong, editor, ASME, 1981.

## Chapter 5 STRAIN-LIFE DATA ANALYSIS; SOME ADVANCED TOPICS

### 5.1 Preliminary Remarks

Chapter 4 discussed how to establish a design curve, and provide parameters for a probabilistic design format for strain-life data which has a linear trend throughout the range of interest, and has a constant standard deviation of  $\log N$  given  $\log S$  ( $\sigma_{\log N}$ ). This chapter will describe analysis procedures for other types of SN data.

1. The SN data has a linear trend, but scatter increases with decreasing stress (or strain)
2. The general strain life relationship which is now commonly used in design (1), i.e., for the local strain approach
3. A distribution of stress given life  $N$ , in the high cycle range where a stress endurance limit is assumed

### 5.2 Linear SN Curve With Variable Scatter

Consider the fatigue data of Figure 5.1. Visual examination leads one to the conclusion that while there is a linear trend to the data, it does appear that the scatter in cycles to failure  $N$  increases with decreasing strain  $\epsilon$  (or stress). Data having this characteristic is said to be "heteroscedastic". This behavior is fairly common in fatigue data.

The following analysis procedure is proposed to characterize the data for design purposes. The data of Figure 5.1 is used as an example.

1. It is assumed that (a) the data follows a linear relationship, (b) the distribution of  $Y = \log N$  given  $X = \log \Delta\epsilon$  is normal. The least squares line  $\hat{Y} = \hat{a} + \hat{b}x$ , is obtained from the data  $(X_i, Y_i)$ ,  $i = 1, n$ . It is an estimate of the mean of  $Y$  for a given  $X$  and therefore the median of  $N$  for a given  $\Delta\epsilon$ . The least squares line is shown in Figure 5.2.

Figure 5.1  
Strain-Life Data Indicating Increased Scatter at Lower Strain Levels

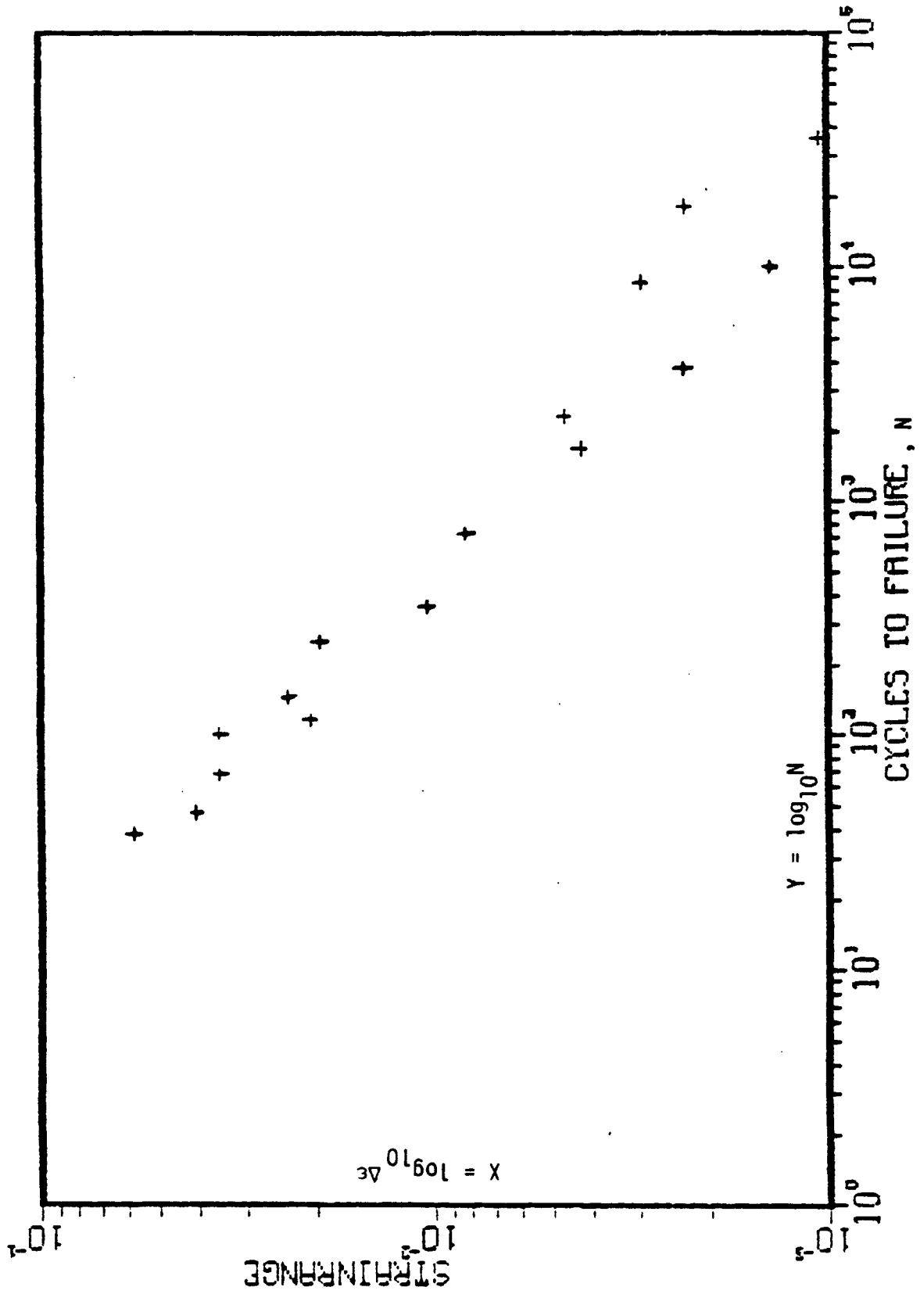
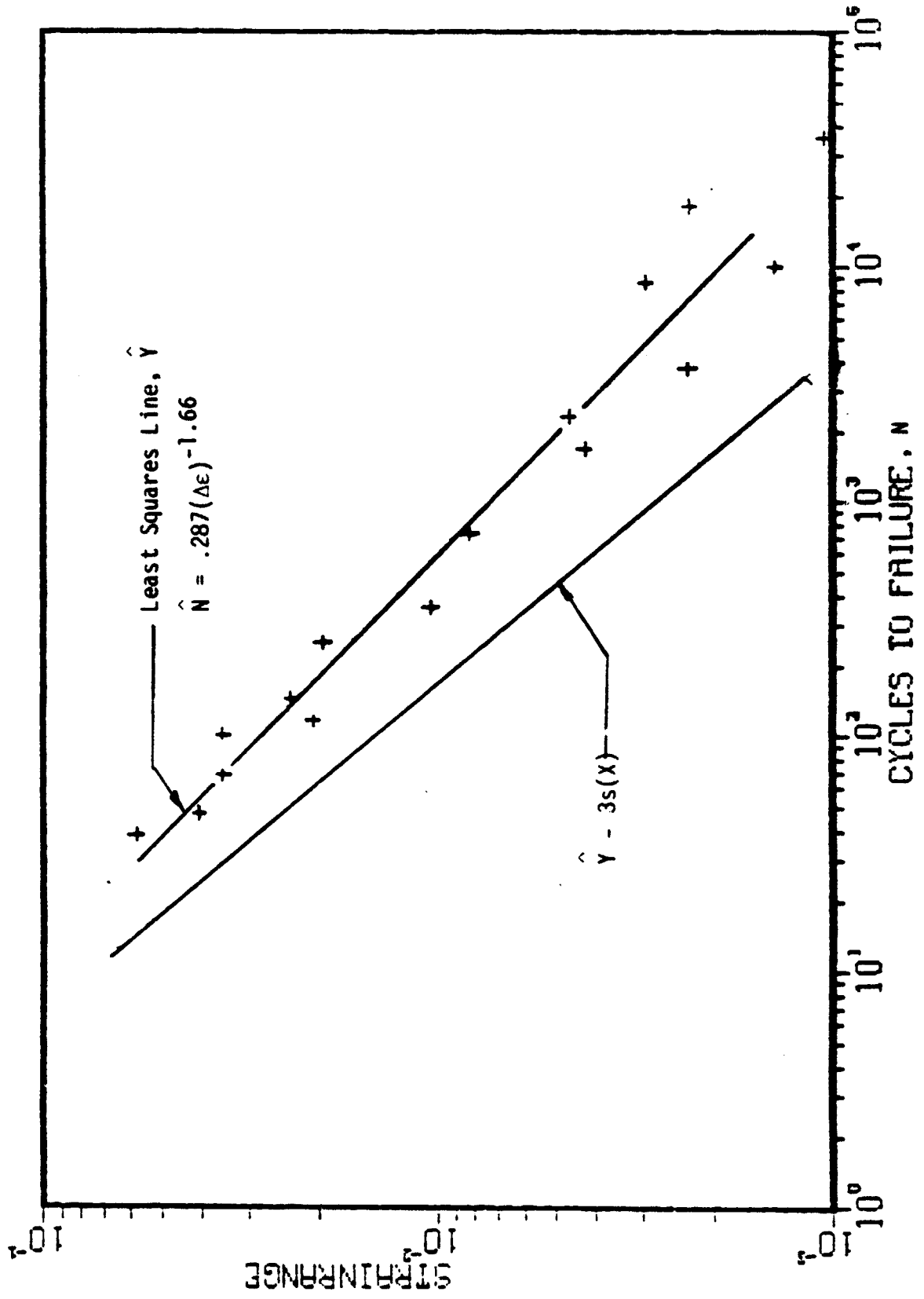


Figure 5.2

The Least Squares Line and the "3 $\sigma$ " Line



2. Define the deviations as  $d_i = (\hat{Y} - Y_i)$ . Compute,

$$s = \sqrt{\frac{1}{n-2} \sum_{i=1}^n d_i^2} \quad (5.1)$$

as an estimate of  $\sigma_{\text{avg}}$ , the average standard deviation of  $Y$  given  $X$ . For the data of Figure 5.1,  $s = 0.194$ .

3. Assume that  $|d|$ , the absolute value of the deviation, has a linear relationship with  $X$ . Determine the least squares line  $|\hat{d}(x)| = \hat{c} + \hat{e}x$ . This is shown on Figure 5.3.
4. Assume that the standard deviation of  $Y$  given  $X$  has a linear relationship to  $X$ . It is denoted as  $\sigma(x)$ . The estimate of  $\sigma(x)$  is denoted as  $s(x)$ . Assume that  $s(x)$  is proportional to  $|\hat{d}(x)|$ . The scale of  $s(x)$  is established by defining  $s$  (the estimate of  $\sigma_{\text{avg}}$ ) as the value of  $s(x)$  corresponding to the mean value of  $X$ , denoted as  $\bar{X}$ . In the example  $\bar{X} = -2.05$ . Both  $\bar{X}$  and  $s$  are shown in Figure 5.3. The scale of  $s(x)$  is established by noting that  $|\hat{d}(x)|$  and  $s(x)$  are zero at the same point.

To use this information for design, several of the techniques described previously can be used. As examples,

1. A design curve defined as the mean minus  $M$  standard deviations can be easily constructed from Figure 5.3. The  $M = 3$  curve shown in Figure 5.2 illustrates the heteroscedastic nature of the data
2. Assume that fatigue behavior can be described by failure trajectories which are flared from the least squares line in proportion to  $|\hat{d}(x)|$ . An example is shown in Figure 5.4. Then, for example, a lower tolerance limit can be used as a design curve as shown. The values of  $P = 0.01$  and  $\gamma = 0.95$  are those of the "A - values" in MIL-HDBK-5.

Figure 5.3

$|\hat{d}|$  and  $s(x)$  for the Data of Figure 5.1

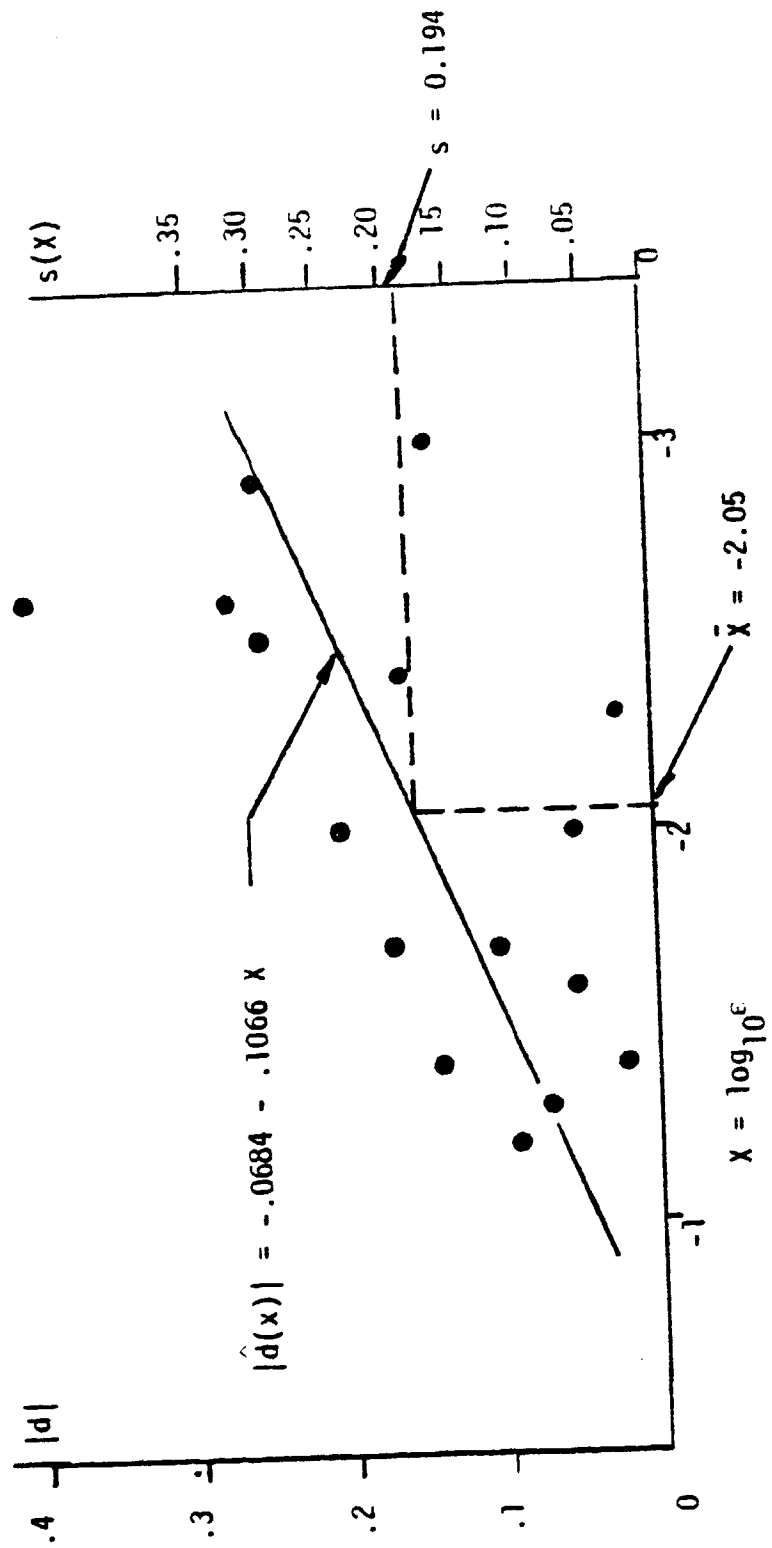
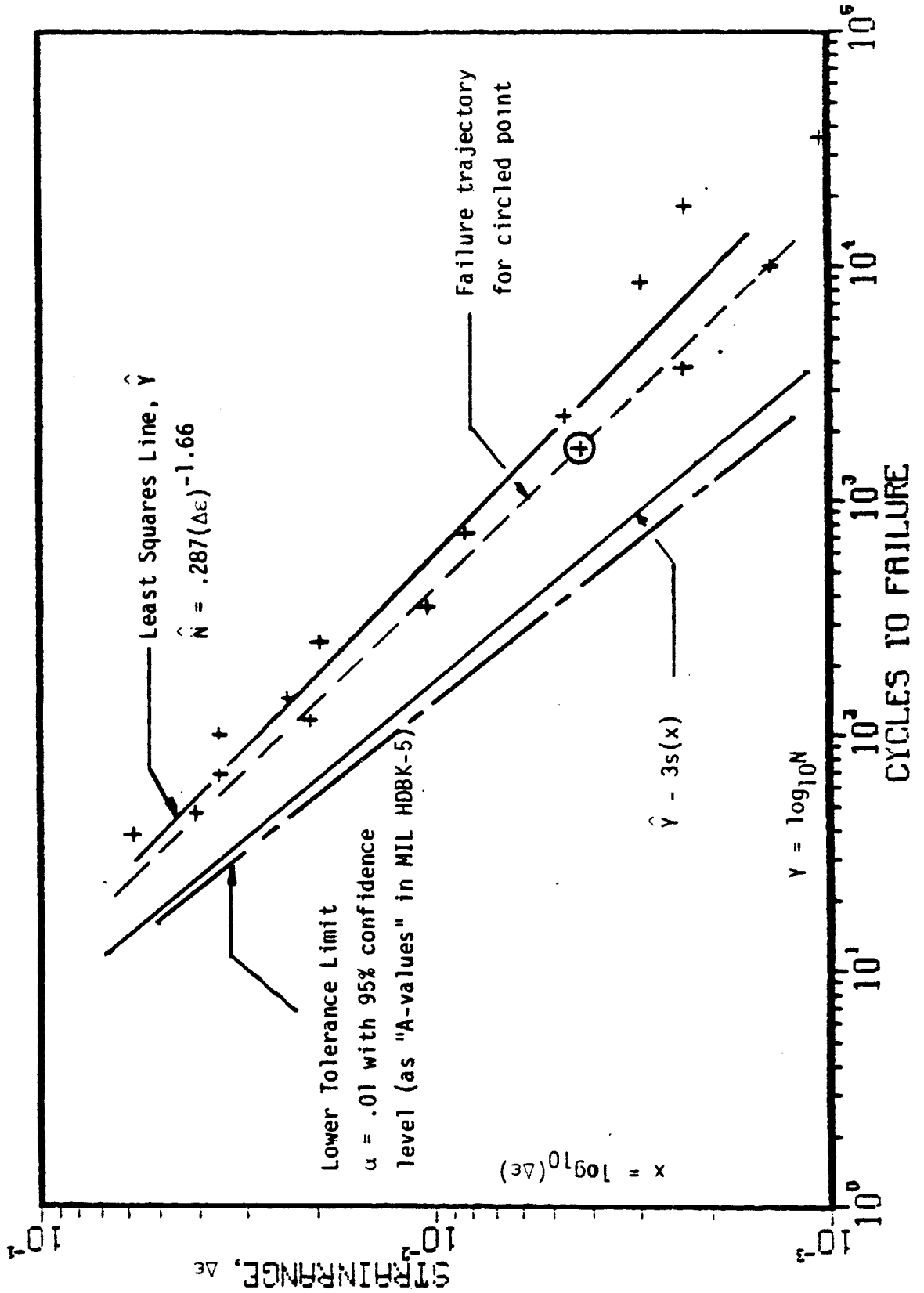


Figure 5.4

Least Squares Line and Design Curve for Heteroscedastic Data of Figure 5.1





3. For a reliability format, it is necessary to have the least squares line, and the standard deviation. The standard deviation has the functional form,

$$\sigma(x) = A + BX \quad (5.2)$$

For example, the estimate of  $\sigma(x)$  is, from Figure 5.3,

$$s(x) = -0.0884 - 0.1377x \quad (5.3)$$

This representation of  $\sigma(x)$  significantly complicates reliability analysis. However, a simplification using an equivalent standard deviation  $\sigma_0$  can be made. Assume that a variable amplitude loading, consisting of  $k$  discrete strain amplitude levels,  $\epsilon_j$ ,  $j = 1, k$  are applied. The fraction of time that  $\epsilon_j$  is applied is known to be  $f_j$ . An average, or equivalent standard deviation can be defined as a weighted average,

$$\sigma_0 = \sum_{i=1}^k f_i \sigma(X_i) \quad (5.4)$$

where  $X_j = \log \epsilon_j$ . But the right hand side is, by definition, the expected value of  $\sigma(x)$ , denoted as  $E[\sigma(x)]$ . Substituting Equation 5.2,

$$\sigma_0 = E[\sigma(x)] = E[A + BX]$$

$$\sigma_0 = A + B \cdot E[\log \epsilon] \quad (5.5)$$

To find  $E[\log \epsilon]$  it is necessary to evaluate

$$E[\log \epsilon] = \sum_{i=1}^k f_i \log \epsilon_i \quad (5.6)$$

and if the strain amplitude levels are continuous with probability density function  $f(\varepsilon)$ , then

$$E[\log \varepsilon] = \int_0^{\infty} (\log \varepsilon) f(\varepsilon) d\varepsilon \quad (5.7)$$

In summary, the value of  $\sigma_0$  as computed by Equation 5.5 can be used in a reliability analysis. The least squares,  $\hat{N}S^m = \hat{K}$  defines the median value of  $K$ , and the coefficient of variation of  $K$  is, (using lognormal relationships)

$$C_K = \sqrt{10^{(\sigma_0^2/0.434)} - 1} \quad (5.8)$$

### 5.3 Statistical Analysis of the General Strain-Life Relationships

Fatigue tests can include cyclic plasticity and accompanying low cycle fatigue. A stable hysteresis loop is recorded for each specimen as shown in Figure 5.5. The figure also defines strain types. Strain reversals to failure,  $2N$ , is recorded for each specimen. The general strain life relationship is commonly used to describe fatigue strength for the modern local strain approach to fatigue. General references include References 2 and 3.

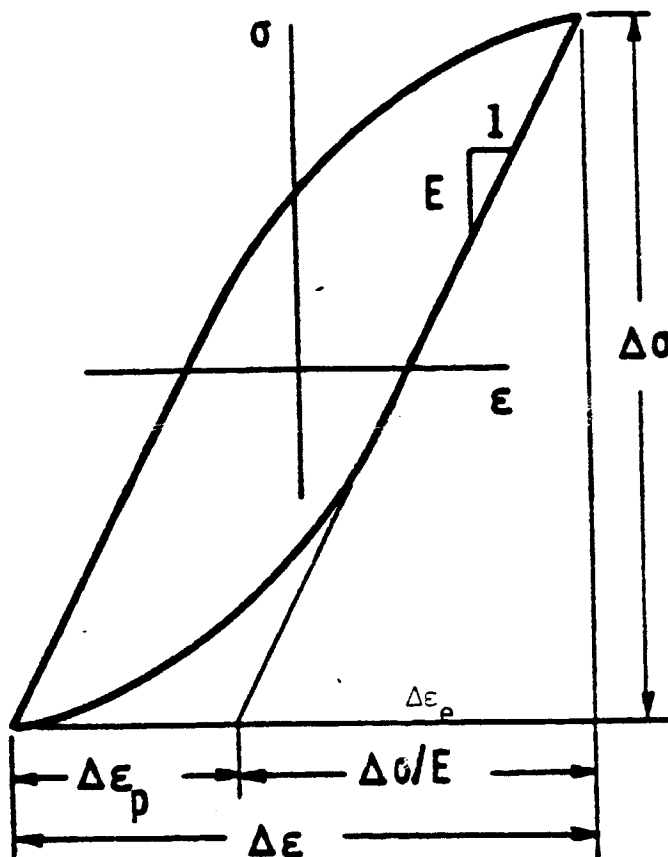
An example is presented in this discussion. The data as well as the analyses are given in Table 5.1.

The Coffin-Manson rule for the plastic strain-life relationship is

$$\varepsilon_{pa} = \frac{\Delta \varepsilon_p}{2} = \varepsilon_f' (2N)^c \quad (5.9)$$

The empirical constants  $\varepsilon_f'$  (fatigue ductility coefficient) and  $c$  (fatigue ductility exponent) are established by a least squares line through the data as illustrated in Figure 5.6.

Figure 5.5  
Stable Hysteresis Loop



$\Delta\epsilon$  = total strain range

$\Delta\epsilon_p$  = plastic strain range

$\epsilon_{pa} = \Delta\epsilon_p/2$  = plastic strain amplitude

$\Delta\epsilon_e$  = elastic strain range

$\epsilon_{ea} = \frac{\Delta\epsilon_e}{2}$  = elastic strain amplitude

Table 5.1

Strain-Life Data for Example Problem\*

SAMPLE SIZE • 9	PLASTIC STRAINRANGE	ELASTIC STRAINRANGE	CYCLES TO FAILURE
.4020E-01	.3540E-01	.4800E-02	.3350E+03
.1980E-01	.1560E-01	.4200E-02	.1320E+04
.1600E-01	.1180E-01	.4200E-02	.1300E+04
.1580E-01	.1120E-01	.4600E-02	.1200E+04
.1420E-01	.1020E-01	.4000E-02	.2805E+04
.9600E-02	.5800E-02	.3800E-02	.4400E+04
.7800E-02	.4400E-02	.3600E-02	.1105E+05
.7600E-02	.4000E-02	.3600E-02	.9050E+04
.5000E-02	.2200E-02	.2800E-02	.6000E+05

MODULUS OF ELASTICITY • 29500.

FATIGUE STRENGTH COEFFICIENT, SF • 148.76

FATIGUE STRENGTH EXPONENT, B = -.1075

FATIGUE DUCTILITY COEFFICIENT, EPSILON • .4912

FATIGUE DUCTILITY EXPONENT, C = -.5397

STANDARD DEVIATION (ELASTIC STRAINS) = .0143\*\*

STANDARD DEVIATION (PLASTIC STRAINS) = .0084

STANDARD DEVIATION (TOTAL STRAINS) = .0568

COV BASED ON 5 PERCENT EQUIVALENT PREDICTION INTERVAL

FATIGUE STRENGTH COEFFICIENT COV = .0420

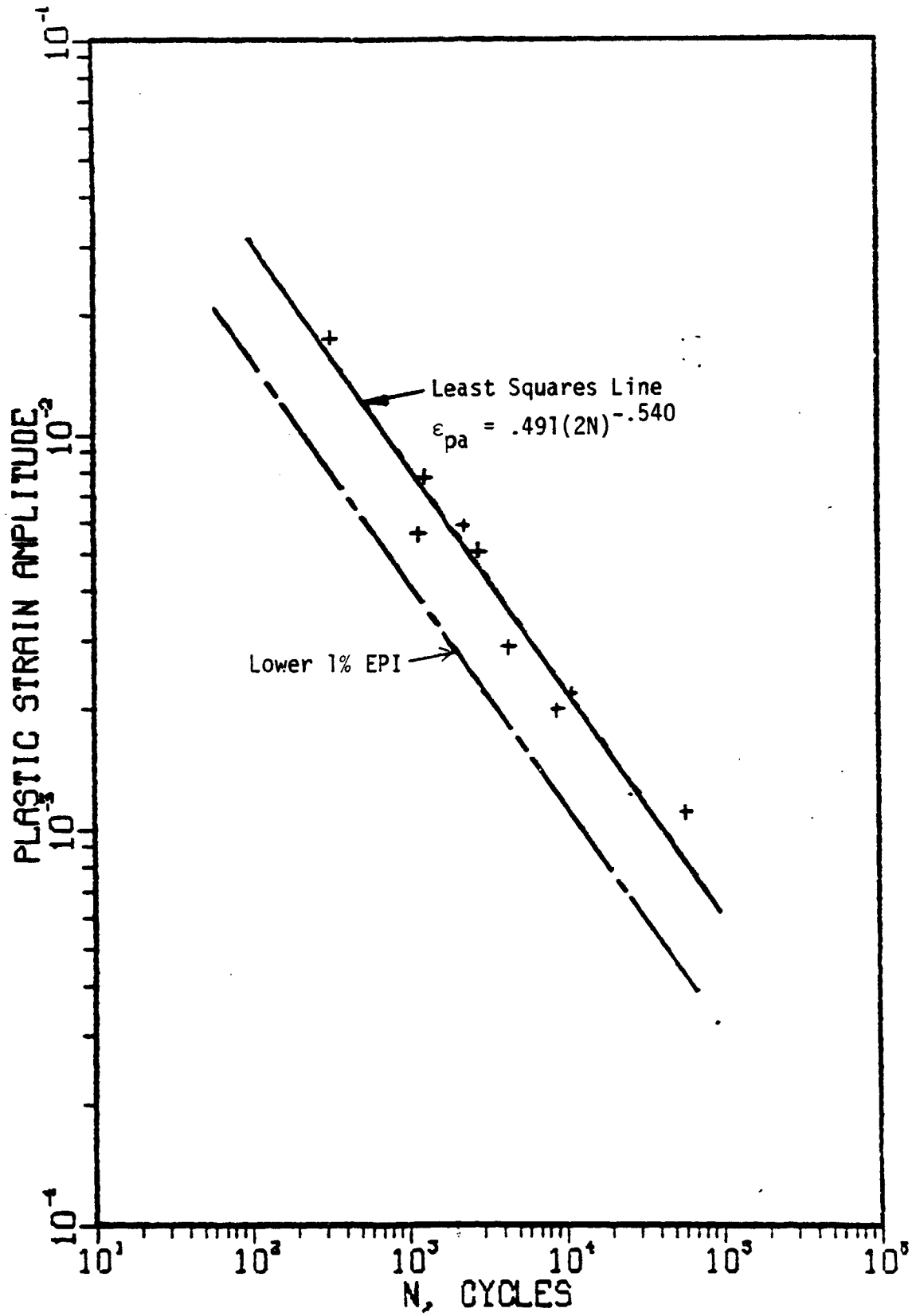
FATIGUE DUCTILITY COEFFICIENT COV = .2635

\* This analysis was based on a least squares analysis in which cycles to failure is considered the independent variable, i.e.  $x = \log N$  and  $Y = \log \epsilon_a$

\*\* Estimates of the standard deviation of  $\log \epsilon$  (vertical deviations)

Figure 5.6

Plastic Strain-Life Curve for Example Problem



The Basquin equation for the elastic strain-life relationship is, (4)

$$\epsilon_{ea} = \frac{\Delta \epsilon_e}{2} = \frac{S'_f}{E} (2N)^b \quad (5.10)$$

Again, the empirical constants  $S'_f$  and  $b$  are established by a least squares method. See Figure 5.7.

The total strain life curve is obtained by adding the elastic and plastic strains.

$$\begin{aligned} \epsilon_a &= \frac{\Delta \epsilon}{2} = \epsilon_{ea} + \epsilon_{pa} \\ &= \frac{\sigma'_f}{E} (2N)^b + \epsilon'_f (2N)^c \end{aligned} \quad (5.11)$$

In this example, the total strain-life curve is shown on Figure 5.8.

The "standard deviations" in Table 5.1 were computed as follows. Consider a failure point  $(\epsilon_i, N_i)$ . Let  $x_i = \log_{10} \epsilon_i$  and  $\hat{x}_i = \log_{10} \hat{\epsilon}_i$  where  $\hat{\epsilon}_i$  is the value strain from the least squares line at  $N_i$ . Let

$$d_i = x_i - \hat{x}_i \quad (5.12)$$

The sample standard deviation, an estimate of  $\sigma_{\log \epsilon}$ , is given as

$$s = \left[ \frac{1}{n-2} \sum_{i=1}^n d_i^2 \right]^{1/2} \quad (5.13)$$

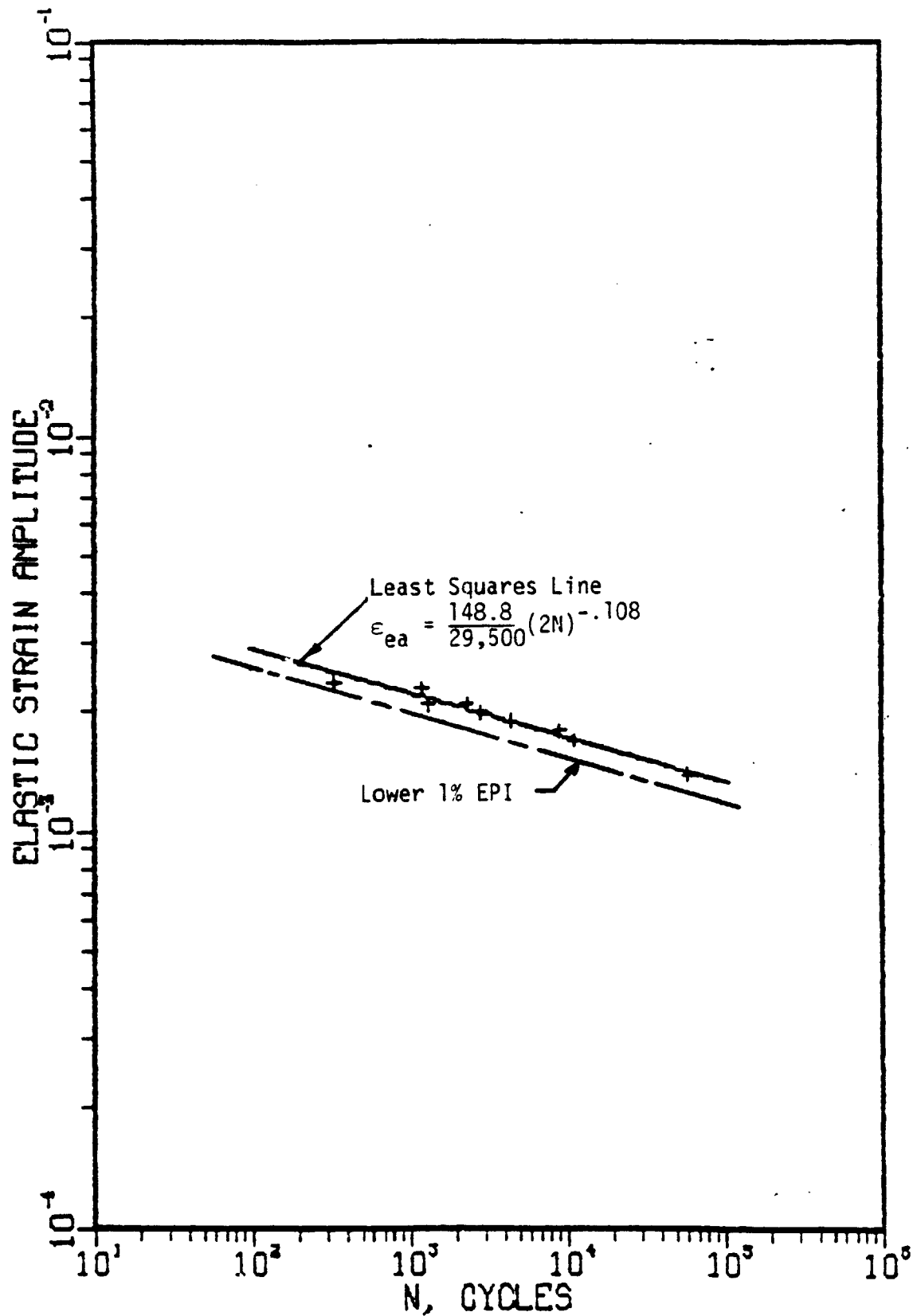
Values of  $s$ , so computed, are given in Table 5.1, not only for the plastic and elastic strain curves, but also for the total strain curve.

For design purposes, it may be required to define a design curve on the safe (lower) side of the data and/or specify statistics of the parameters. The latter would be necessary if a complete fatigue reliability analysis is to be performed.

At this time only preliminary studies of the statistical character of

Figure 5.7

Elastic Strain-Life Curve for Example Problem



strain-life data have been completed. Analyses of 14 sets of fatigue data of steels and nickel base superalloys of various types have suggested the following conclusions.

1. There is no statistical relationship between the plastic and elastic strain amplitude for a failed specimen. One might suspect one of two possibilities (a) a specimen having high ductility would have low strength, ... good plastic strength, poor elastic strength, or (b) a defect causes poor strength in both plastic and elastic strain. No evidence in these preliminary studies suggests either phenomena is dominant.
2. Scatter in cycle life ( $N$ ) increases with decreasing strain level ( $\epsilon_a$ ). This behavior is typical of fatigue data.
3. Scatter in fatigue strength ( $\epsilon_a$ ) decreases with increasing life ( $N$ ). Experimentalists have suggested to the author that the standard deviation of  $\log \epsilon_a$  is constant (not a function of  $N$ ). No evidence that this is fact is found in the present study.

The above conclusions are preliminary, and work on this problem is continuing.

Based on the observations as described above, the following procedures characterizing the data are presented as being reasonable.

1. The lower  $\alpha\%$  EPI curves associated with the plastic and elastic strain life data can be established as described in Chapter 4. The 1% EPI curves are shown in Figures 5.6 and 5.7. Then the total 1% EPI curve of Figure 5.8 is obtained by adding (vertically) the two.
2. A lower tolerance limit can be established in the same way, i.e. by first using conventional methods (see Chapter 4) for  $\epsilon_{pa}^{-N}$  and  $\epsilon_{ea}^{-N}$  curves and then adding the two. The result is shown in Figure 5.9.



Figure 5.8

Strain-Life Curve for Example Problem

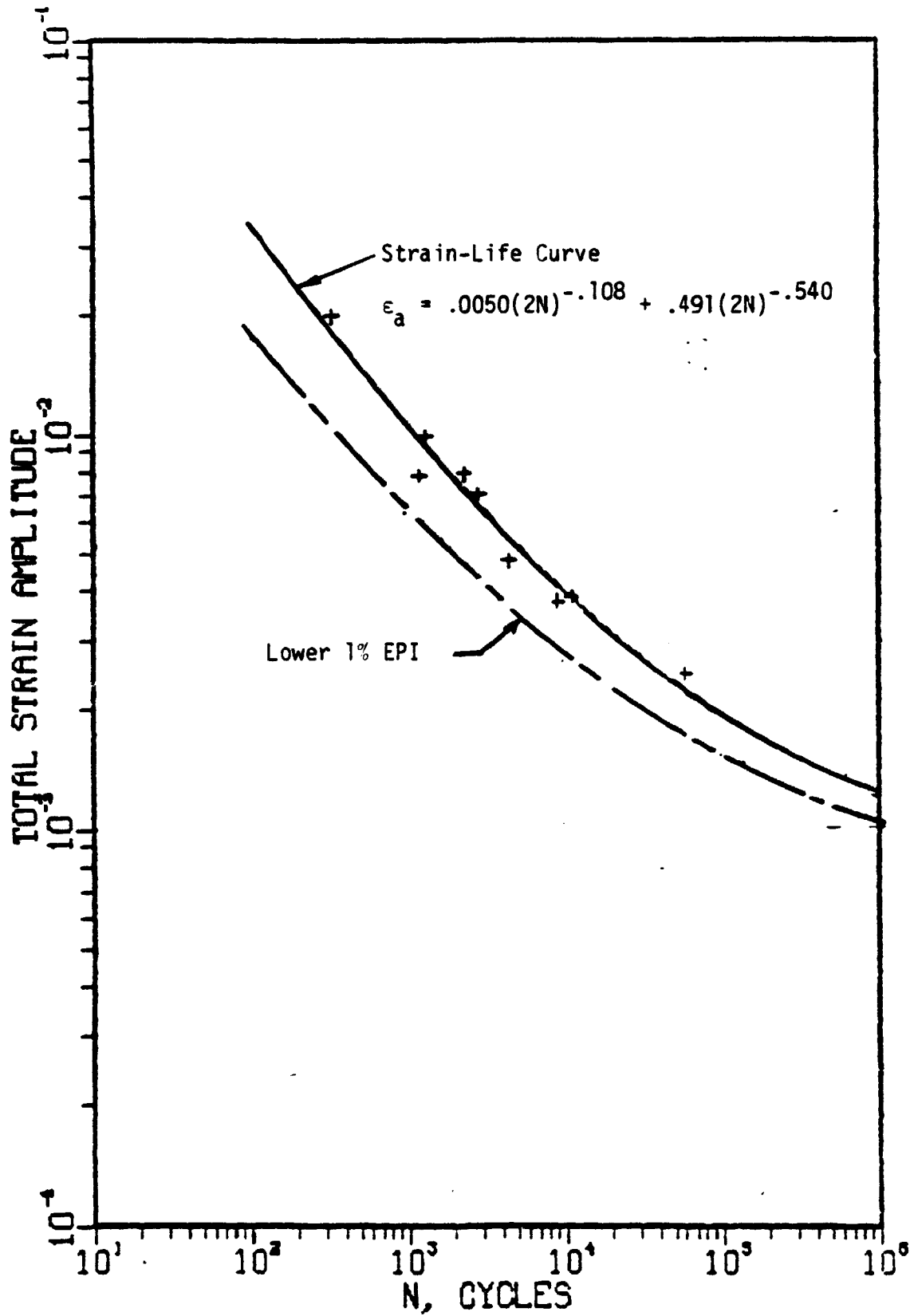
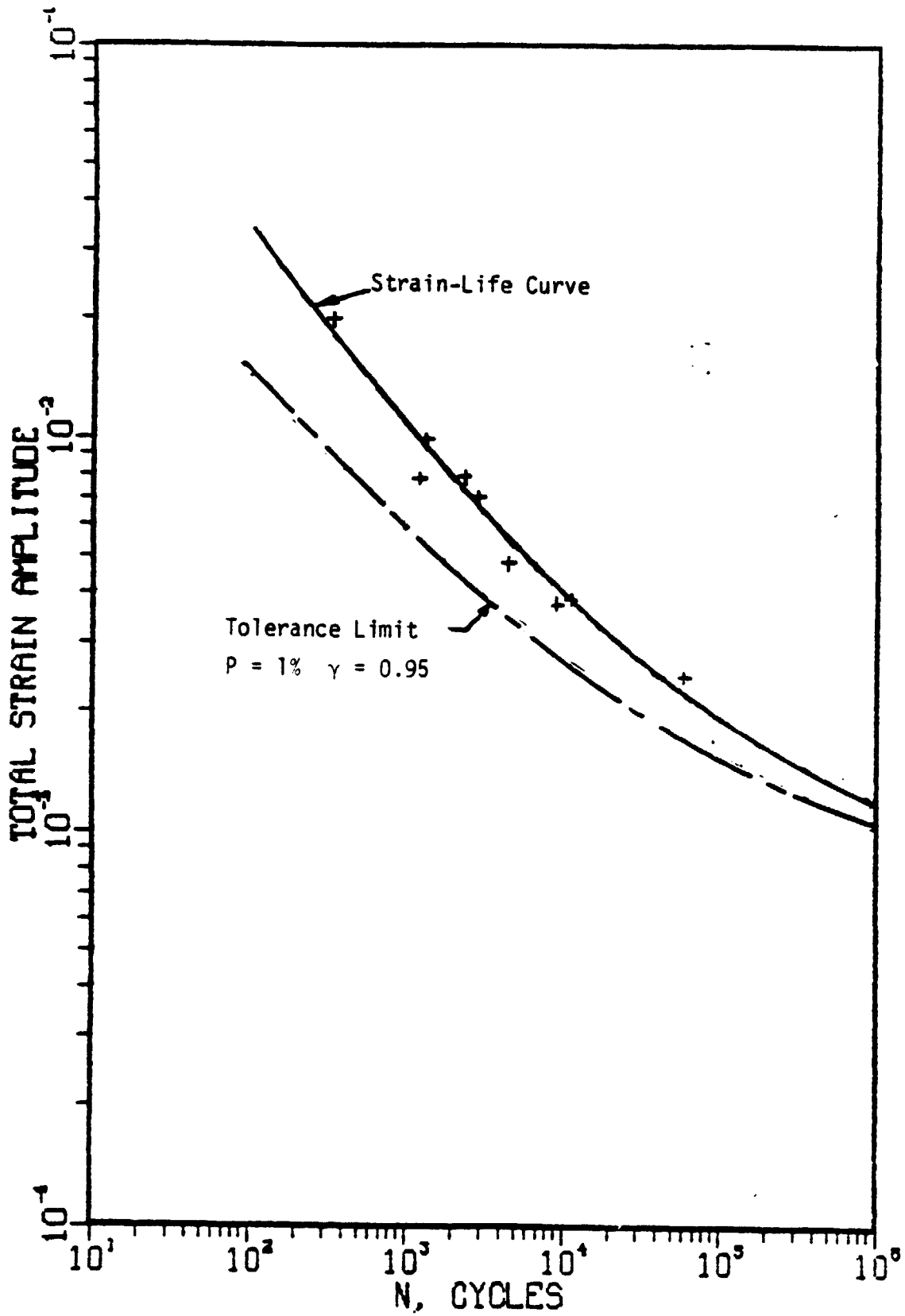


Figure 5.9

Strain-Life Curve and Tolerance Limit for Example Problem



3. For a reliability analysis, statistics of  $s'_f$ ,  $\epsilon'_f$ ,  $b$  and  $c$  are required. Assume  $b$  and  $c$  are constants and equal to the least squares estimators. Assume that  $s'_f$  and  $\epsilon'_f$  are independent random variables having lognormal distributions. In the example of Table 5.1, the statistics required for a reliability analysis are as follows.

Table 5.2

Statistics for Reliability Analyses for Example Problem

	Median	Coefficient of Variation
$s'_f$	148.8 ksi	0.042*
$\epsilon'_f$	.491	0.263*
$b$	-.108	0**
$c$	-.540	0**

\*\* based on a 5% EPI, these values would be only slightly different from the 1% EPI values given in Figures 5.6, 5.7 and 5.8

\* variable not random

## 5.4 A Model for the Distribution of Fatigue Strengths at High Cycle Lives

### 5.4.1 Preliminary Remarks

The number of load cycles experienced during the lifetime of a wide variety of structural and mechanical components is in the range from  $10^7$  to  $10^{10}$  cycles. When metal fatigue is considered to be the mode of failure, it is necessary for the designer to know the fatigue strength (the maximum amplitude of the oscillatory stress) for a given cycle life  $N$ .

Fatigue testing is however expensive, partly because of the length of time that it takes to apply these long cycle lives. An accelerated test method which has been used successfully involves loading of a specimen at a relatively high stress level, with relatively short cycle lives, and then extrapolating the data to the higher cycle lives.

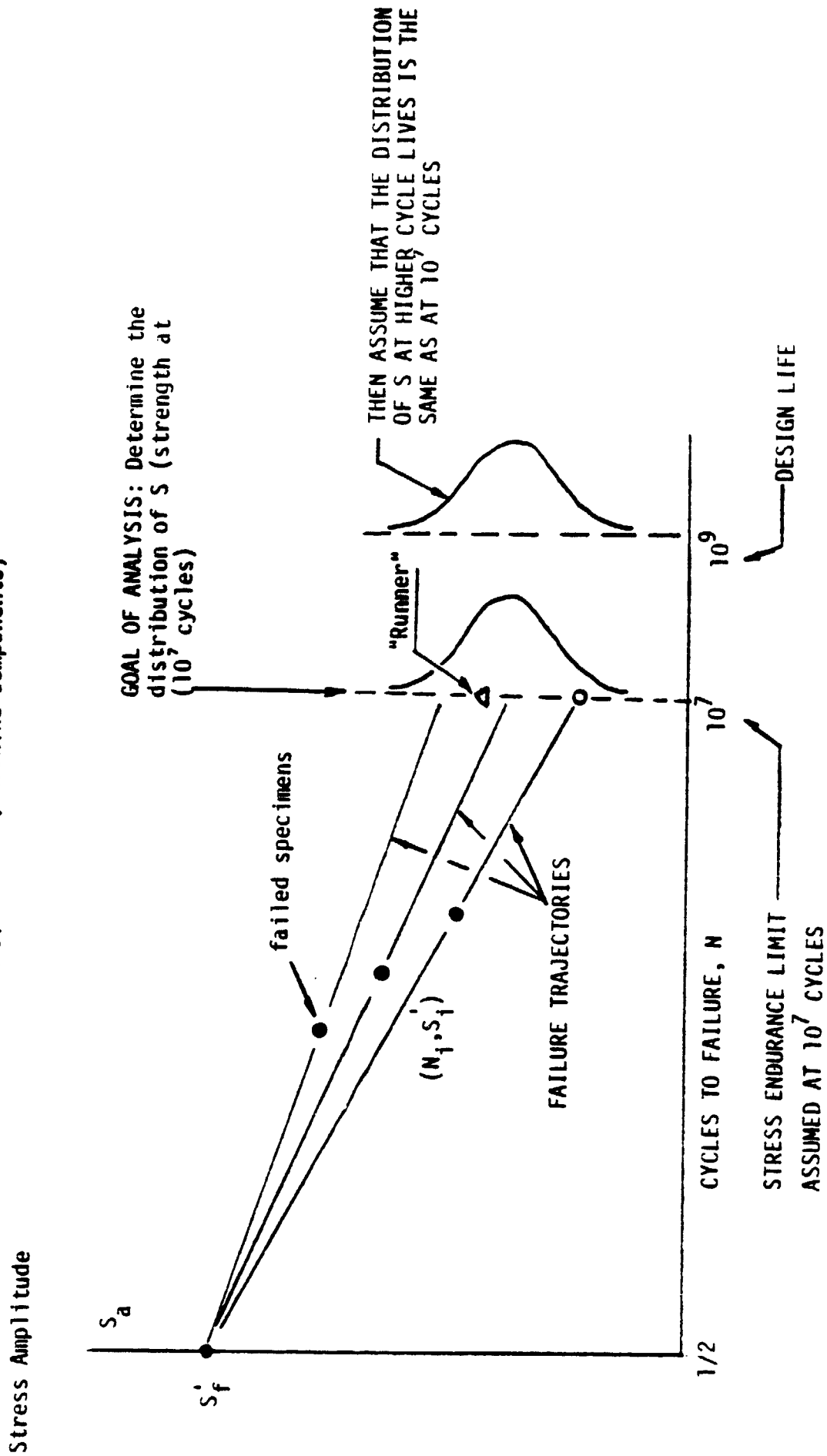
Described as follows is a study of an analytical method for the statistical distribution of fatigue strengths at high cycle lives based on data at relatively low cycle lives.

### 5.4.2 Basic Assumptions and Description of the Problem

Consider a constant amplitude high cycle fatigue test in which pairs of data  $(S_i', N_i)$ ,  $i = 1, n$  are collected.  $N_i$  (the dependent variable) is the cycles to failure associated with stress amplitude  $S_i'$  (the independent variable) and  $n$  is the sample size. Data from a hypothetical test are shown in Figure 5.10. An endurance limit is assumed as illustrated by the horizontal segment of this S-N curve. It is assumed that the endurance limit occurs at  $10^7$  cycles. Furthermore, it is assumed that the distribution of  $S$  at a given  $N$  at  $10^7$  cycles will be the same for any higher life cycle as shown in Figure 5.10.

Figure 5.10

A MODEL FOR HIGH CYCLE FATIGUE  
(Typical Use; Machine Components)



In general, a fatigue test plan specifies a termination of the loading on a component which does not fail before a given life. This specimen is called a "runner" and the data points are shown by the triangles of Figure 5.10 in which the test was suspended at  $10^7$  cycles. This data must be included in the analysis.

To obtain a random sample of  $S$ , now defined as a random variable denoting the fatigue strength at  $10^7$  cycles, the following procedure is used. It is assumed that the fatigue strength for a given specimen is defined by a straight line (failure trajectory) connecting the fatigue strength coefficient at  $s_f$  1/2 cycle and  $(N_i, s_i)$  as shown in Figure 5.10. The data point (solid point in Figure 5.10) is projected to the life cycle where the endurance limit occurs, which is assumed here to be  $10^7$  cycles. The sample point (circle in Figure 5.10) is denoted as  $S_i$ . By such a scheme one can obtain a random sample of  $s$ .

#### 5.4.3 Analysis of Suspended Fatigue Test Data

Incomplete failure data consisting of levels on failed components and unfailed components are called "multiply censored." This suspended data can be analyzed utilizing the "median rank concept" and suspended items approach [5]. The median ranks extracted from this approach will be used to establish the empirical distribution function of fatigue strengths.

Lipson and Sheth have described a method for obtaining the empirical distribution function from a random sample which includes suspended items [5]. This method is a combination of a modified sudden death approach and Johnson's concept of median ranking [6]. The method involves an adjustment of the order

number based upon the position of the suspended items. Step-by-step instructions are illustrated by the following:

1. Suppose that the failure data on  $n$  specimens consist of the failure stresses (at  $10^7$  cycles for the failed units and the stress levels for the unfailed units (see Table 5.3 and Figure 5.11)). Order the failed units in the sample from the smallest to largest failure stresses as shown in Column 3.

2. Obtain the number of suspended items which precede each failed unit (Column 2).

3. Determining the "new increment" of each failed unit by using the formula

$$\text{New Increment} = \frac{(n + 1) - \text{previous failure order number}}{1 + \text{number of items following present suspended set}}$$

The new increment is recalculated each time a suspension is encountered in the ordered stress table (see Column 4).

4. Calculate the order number for each failed unit. This calculation is done by simple addition of the last order number and the new increment (Column 5).

5. Obtain the median rank (empirical distribution function) of the order number for each failed unit by the formula (Column 6).

The empirical distribution function  $F(S_{(j)})$  of Column 6 is used as a basis for selecting a suitable statistical model. The method of least squares, as described in Sections 2.5, 2.6, and 2.7, can be used. The correlation parameter for the three distributions considered are

Table 5.3  
Analysis of Suspended Fatigue Data

EXAMPLE: Given a sample of fatigue failure stress at  $10^7$  cycles (in ksi);  
n = 13; the data are ordered.

11.7 | 12.2 | 12.5 | 12.7 | 12.8 | 13.3 | 13.8 | 14.0\* | 14.0\* | 14.0\* | 14.4 | 14.7 | 15.0\*

\*Specimen did not fail at  $10^7$  cycles.

Organize data as shown below:

(1) Failure Number	(2) Suspended Items Preceding Failure	(3) Stress $S(j)$	(4) New Increment**	(5) Order Number** (j)	(6) Median Rank $F(S(j)) = \frac{j - 0.3}{n + 0.4}$
1	0	11.7	1.00	1.00	0.052
2	0	12.2	1.00	2.00	0.127
3	0	12.5	1.00	3.00	0.201
4	0	12.7	1.00	4.00	0.276
5	0	12.8	1.00	5.00	0.351
6	0	13.3	1.00	6.00	0.425
7	0	13.8	1.00	7.00	0.500
8	3	14.4	1.75	8.75	0.631
9	0	14.7	1.75	10.50	0.761

\*\*An example of how the new increment and order number are calculated:

The eight failure is preceded by three suspended items. Therefore, to find the increment, as shown in Step 3 above,

$$I = \frac{(13 + 1) - 7}{1 + 3} = 1.75$$

Thus, the new order number is  $j = 7.00 + 1.75 = 8.75$

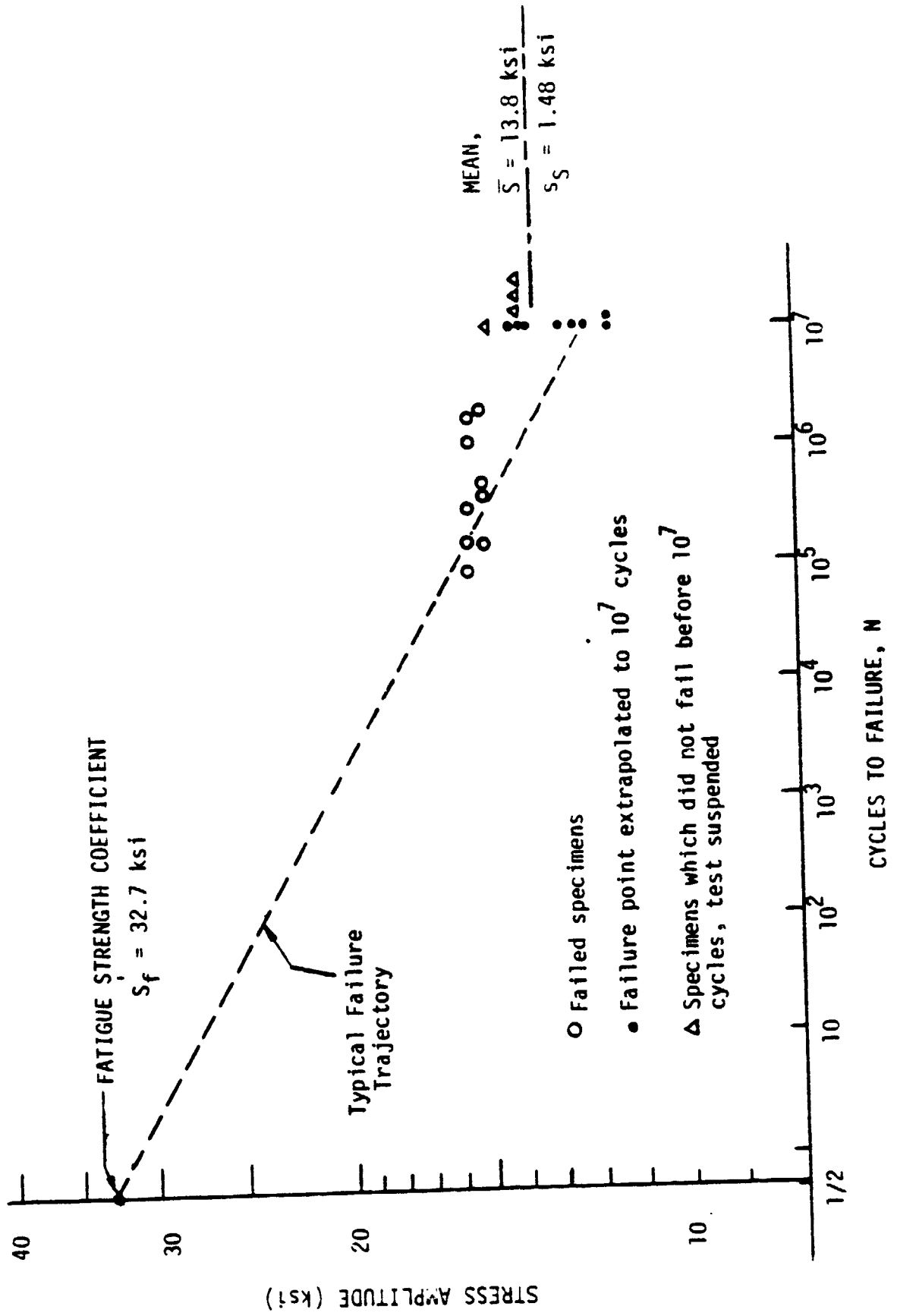
The remaining order numbers are determined by using the same procedure and

$$F = \frac{8.75 - 0.3}{13 + .4} = 0.631$$



Figure 5.11

S-N DIAGRAM FOR EXAMPLE OF DATA EXTRAPOLATED INTO HIGH CYCLE RANGE



Distribution Family	Correlation Parameter, $\epsilon^*$
Normal	0.2010
Lognormal	0.1809
Weibull	0.2738

The lognormal is considered to have the best fit because it has the smallest  $\epsilon$ . The probability plot is given in Figure 5.12. The line is fit by the least squares method and the distribution parameters on the figure are computed from the least squares line (using Equation 2.15 and 2.16).

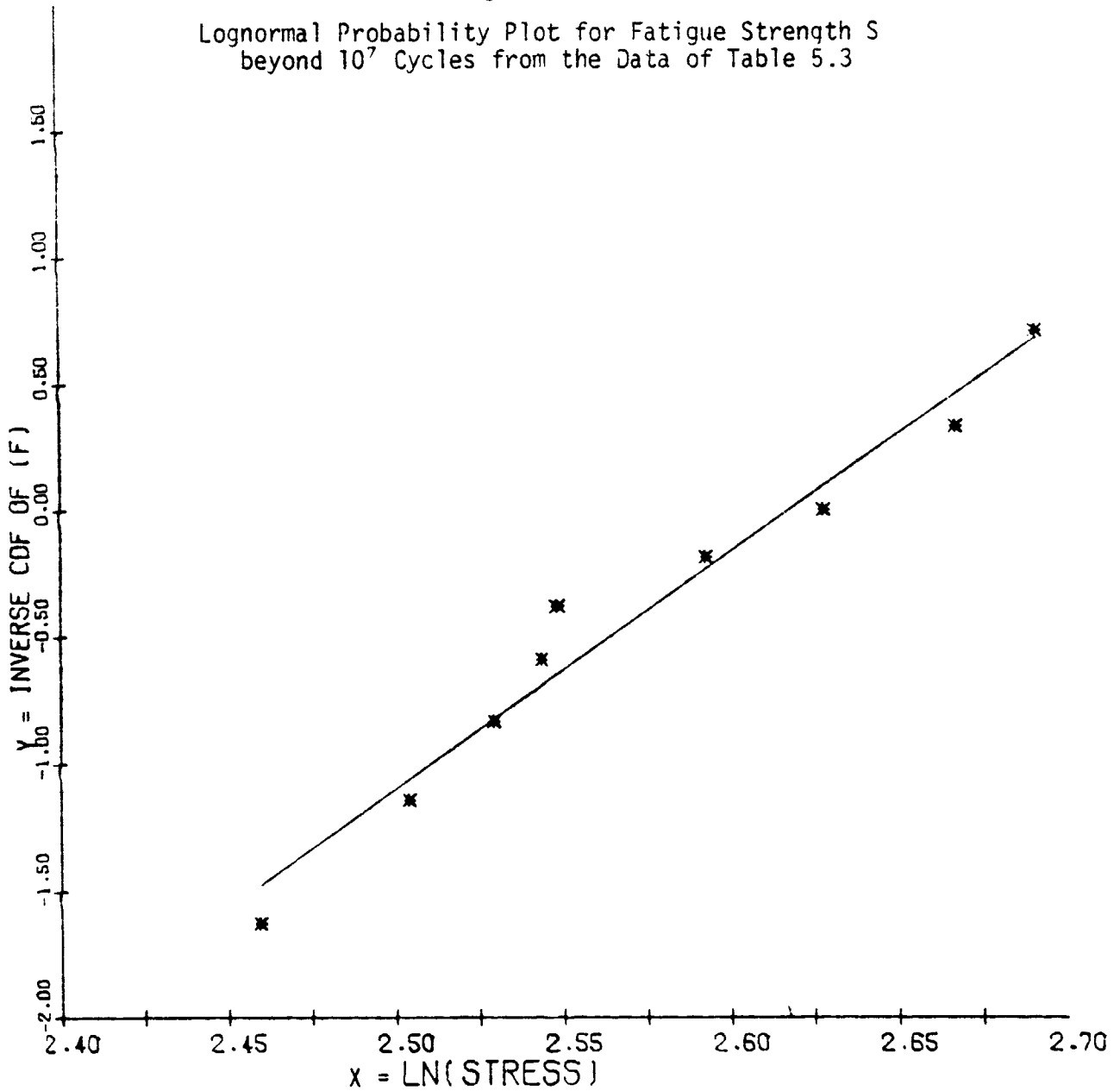
In summary, for example, the distribution of fatigue strength  $S$  given cycle life  $N$  in the high cycle region (beyond  $10^7$  cycles) can be modelled as a lognormal distribution. The parameters are  $\mu_X = 2.67$  and  $\sigma_X = 0.107$ , where  $X = \ln S$ . Using the relationships of Appendix 1, the sample mean and standard deviation of  $S$  is  $\bar{S} = 13.8$  ksi and  $s_S = 1.48$ , respectively. Full consideration was given to "runners" in the data analysis.

Statistical methods of estimating the fatigue strength at a given cycle life are available. Such methods, summarized by Reemsnyder (7), Lipson and Sheth (5), and by Collins (8) include the survival method, the staircase method, and the Prot method. Because these methods typically require long cycle lives and relatively large sample sizes, testing tends to become expensive.

\* See Equation 2.29

Figure 5.12

Lognormal Probability Plot for Fatigue Strength S  
beyond  $10^7$  Cycles from the Data of Table 5.3



Estimated Parameters:

Mean of S = 13.15

Standard Deviation of S = 1.01

$\mu_x = 2.62$

$\sigma_x = 0.107$

## 5.5 References for Chapter 5

1. Dowling, N.E., Brose, W.R. and Wilson, W.K., "Notched Member Fatigue Life Predictions by the Local Strain Approach", Fatigue Under Complex Loading; Analyses and Experiments, AE-6, SAE, 1977.
2. Fatigue Design Handbook, AE-4, SAE, 1968.
3. Fatigue Under Complex Loading: Analyses and Experiments, AE-6, SAE, 1977.
4. Fuchs, H.O., and Stephens, R.I., Metal Fatigue in Engineering, 1980.
5. Lipson, C. and Sheth, N.J., Statistical Design and Analysis of Engineering Experiments, McGraw-Hill, 1973.
6. Johnson, L.G., "The Median Ranks of Sample Values in Their Population with an Application to Certain Fatigue Studies", Industrial Mathematics, Vol. 2, 1951.
7. Reemsnyder, H.S., "Procurement and Analysis of Structural Fatigue Data", Journal of the Structural Division, ASCE, Vol. 95, No. ST 7, July 1969.
8. Collins, J.A., Failure of Materials in Mechanical Design, Wiley, 1981.

## CHAPTER 6 LIFE PREDICTION EQUATIONS USED FOR FATIGUE DESIGN

### 6.1 Preliminary Remarks

Engineering models used by designers for fatigue failure avoidance usually employ the concept of "damage". Summarized in this chapter are some of the various forms of damage in current use. It is shown that the concept of damage can be employed for variable amplitude loading using both a characteristic S-N curve approach and under certain assumptions the fracture mechanics approach. Moreover the concept is used also for strainrange partitioning in which damage is identified not according to stress level but with each of four different types of inelastic strains.

The limit state is achieved when damage  $D$  attains a critical level  $\Delta$ . Experimental results in which damage at failure  $\Delta$  is measured are summarized herein. Both variable amplitude loading and strainrange partitioning results are presented as statistical distributions of  $\Delta$ .

### 6.2 The Concept of Damage

Assume that the SN curve for a component subjected to constant amplitude loading is as shown in Figure 6.1 and that no statistical scatter exists. Blocks of constant amplitude loads are now applied as shown. Fatigue damage  $D$  associated with this variable amplitude process is defined as

$$D = \sum_{i=1}^k \frac{n_i}{N(S_i)} \quad (6.1)$$

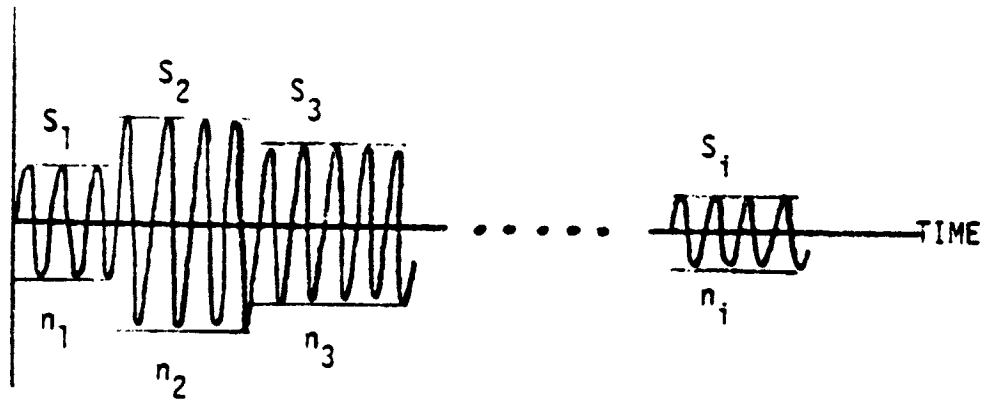
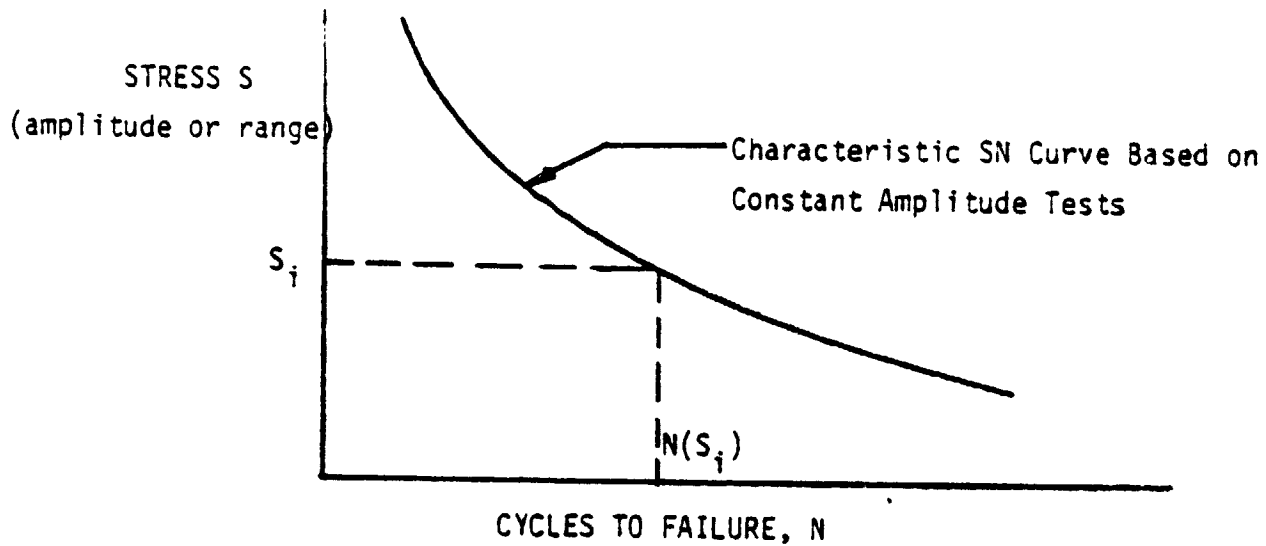
where  $n_i$ , and  $N(S_i)$  and  $k$  are defined on Figure 6.1.

More generally as a component accumulates a load history in service, damage can be expressed as a function of time  $\tau$  and several other design factors denoted by the vector  $\underline{U}$ , which relate to the loads as well as to the SN curve. Thus

$$\text{DAMAGE} \equiv D(n, \underline{U}) \text{ or } D(\tau, \underline{U}) \quad (6.2)$$

Figure 6.1

Basic Terms Used to Define Damage Under Variable Amplitude Stresses



$k$  = number of blocks of loading

$S_i$  = stress level (amplitude or range)

$n_i$  = number of cycles applied at stress level  $S_i$

$N(S_i)$  = number of cycles to failure at stress level  $S_i$

Miner's rule states that failure occurs when  $D$ , a monotonically increasing function of time, equals unity. A more general definition is provided in Section 6.5. The following two sections summarize some damage models in current use.

### 6.3 Some Expressions for Damage Used by Designers for Variable Amplitude Loading

Following are examples of expressions for fatigue damage under variable amplitude (random) loading. The model to be used depends upon how the load history and the SN relationship are specified.

**CASE 1** Often field measurements result in a variable amplitude stress range histogram as shown in Figure 6.2. The fraction of time that the stress range is at level  $S_i$  is  $f_i$ . Thus  $n_i$  can be written as a fraction of the total number of applied cycles,  $n$

$$n_i = f_i n \quad (6.3)$$

and the damage at cycle life  $n$  is,

$$D = n \sum_{i=1}^{\infty} \frac{f_i}{N(S_i)} \quad (6.4)$$

**CASE 2** Assume a linear S-N curve,  $N = KS^{-m}$ . Note that the value of  $K$  depends upon whether stress amplitude or range is considered. Equation 6.4 becomes

$$D = \frac{n}{K} \sum_{i=1}^{\infty} f_i S_i^m \quad (6.5)$$

But the summation is recognized as the expected value of  $S^m$  for a discrete random variable (1).

$$E(S^m) = \sum_{i=1}^{\infty} f_i S_i^m \quad (6.6)$$

Thus 
$$D = \frac{n}{K} E(S^m) \quad (6.7)$$

Figure 6.2

Probability Density Function for Fatigue Stress Amplitudes (or Ranges);  
Discrete Model

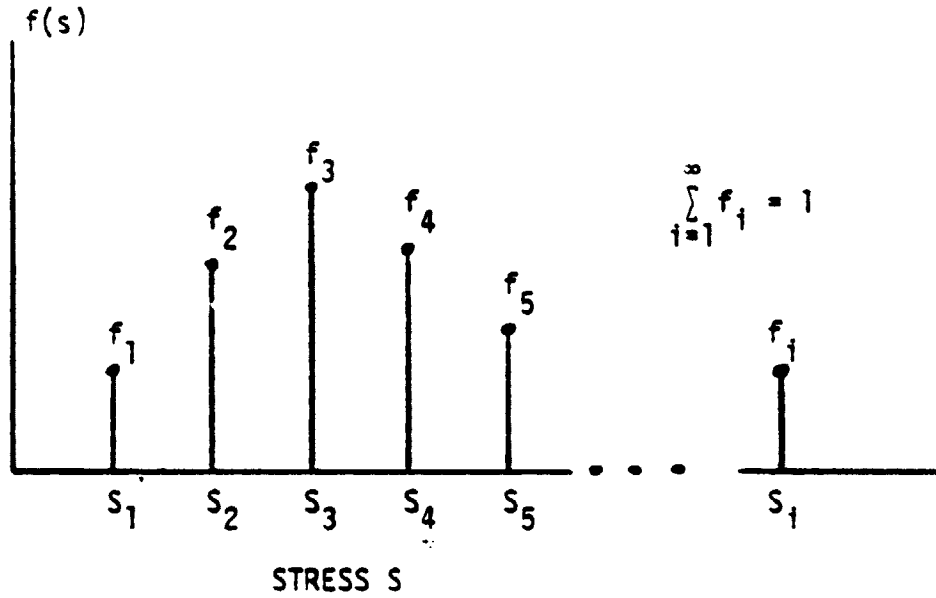
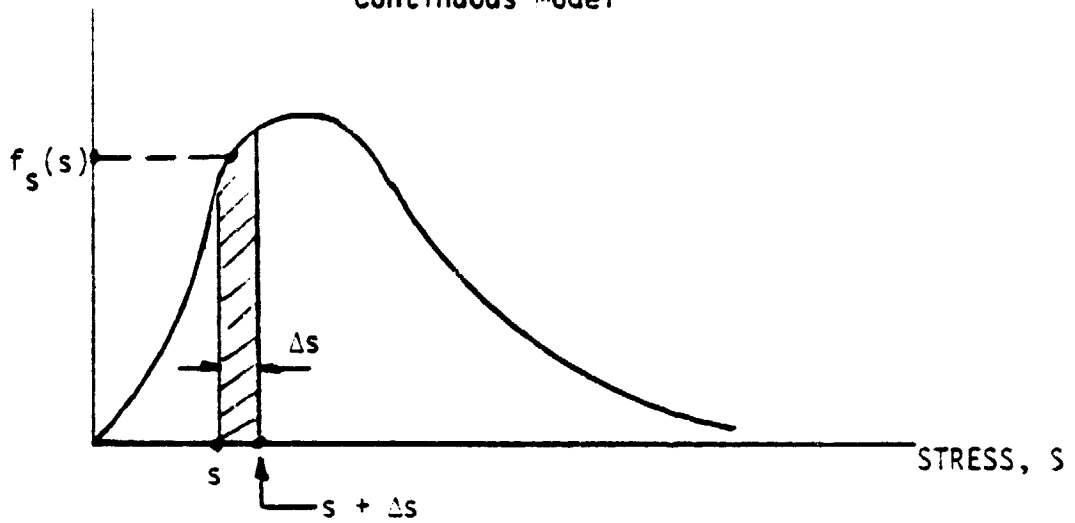


Figure 6.3

Probability Density Function for Fatigue Stress Amplitudes (or Ranges);  
Continuous Model





Note that in the special case of constant amplitude loading  $E(S^m) = S^m$ , and Equation 6.7 becomes

$$D = \frac{n}{K} S^m \quad (6.8)$$

An equivalent constant amplitude stress can be derived by comparing Equation 6.7 and Equation 6.8

$$S_e = [E(S^m)]^{1/m} \quad (6.9)$$

This stress is sometimes called "Miner's stress". Assuming Miner's rule works, the value of  $S_e$  can be entered into a constant amplitude S-N curve to predict life  $N$  under variable amplitude loading.

CASE 3 Assume that the distribution of stress ranges  $S$  can be treated as a continuous random variable, the pdf of which is shown in Figure 6.3. The fraction of cycles at stress level  $s$  in the interval  $(s, s + \Delta s)$  is

$$f_i \cong f_S(s)\Delta s \quad (6.10)$$

Combining Equation 6.1 and 6.3 with Equation 6.10 to obtain the continuous equivalent of Equation 6.4,

$$D = n \sum_{i=1}^k \frac{f_S(s)\Delta s}{N(s)} \quad (6.11)$$

In the limit as  $\Delta s \rightarrow 0$ ,

$$D \cong n \int_0^{\infty} \frac{f_S(s)ds}{N(s)} \quad (6.12)$$

CASE 4 When  $S$  is continuous and it can be assumed that the S-N curve is linear of the form  $NS^m = K$ , the expression for damage becomes

$$D = \frac{n}{K} \int_0^{\infty} s^m f_S(s) ds \quad (6.13)$$

Note that the integral is by definition the expected value of  $S^m$ , denoted as  $E(S^m)$ .

Thus, as in the discrete case,

$$D = \frac{n}{K} E(S^m) \quad (6.14)$$

Any statistical model for  $S$  can be used, but it is common to use the Weibull, having a distribution function of the form

$$F_S(s) = 1 - \exp[-(s/\delta)^\xi] \quad (6.15)$$

where the parameters  $\xi$  and  $\delta$  must be determined from an analysis of the stress data. For the Weibull

$$E(S^m) = \delta^m \Gamma\left(\frac{m}{\xi} + 1\right) \quad (6.16)$$

In the special case where  $\xi = 2$ , the Weibull reduces to the Rayleigh distribution. This is an important case because the Rayleigh is the distribution of peaks or ranges in a stationary narrow band gaussian process having RMS value of  $\sigma$ .

$$\delta = \begin{cases} (\sqrt{2}\sigma) & \text{if stress amplitudes are considered} \\ (2\sqrt{2}\sigma) & \text{if stress ranges are considered} \end{cases}$$

Thus we get the familiar form of Miles' equation for damage under  $n$  cycles of stress of a narrow band process in the case where  $K$  is based on amplitude (2).

$$D = (n/K)(\sqrt{2}\sigma)^m \Gamma(m/2 + 1) \quad (6.17)$$

#### 6.4 Damage Using a Fracture Mechanics Approach to Fatigue

Consider the Paris equation for fatigue crack propagation under a sequence of constant amplitude stress ranges,  $S$

$$da/dn = C(\Delta\kappa)^m \quad (6.18)$$

where  $C$  and  $m$  are empirical constants,  $a$  is the crack length and  $n$  is cycles.  $\Delta\kappa$  is the range of stress intensity factor

$$\Delta\kappa = YS\sqrt{a} \quad (6.19)$$

where  $Y$  is the geometry factor.

An expression for cycles to failure can be derived by integrating Equation 6.18,  $a$  from  $a_0$  (the initial crack size) to  $a_c$  (the critical crack length), and  $n$  from 0 to  $N$  (the cycles to failure). Assuming that  $Y$  is constant (not a function of  $a$ ) and  $a_c \gg a_0$ , it follows that

$$NS^m = \frac{1}{a_0^{m/2 - 1} (m/2 - 1) CY^m} \equiv K \quad (6.20)$$

But this form is identical to the form of the SN curve used in the previous section. The right hand side is identified with  $K$ , an empirical constant.

Consider a random load sequence of  $S_i$ ,  $i = 1, n$ . If sequence effects were not important, it may be reasonable to use Miner's rule as above, and express damage as

$$D = \frac{n}{K} E(S^m) \quad (6.21)$$

For this to be a useful model, there should be a reasonable distribution of  $S_i$ 's, e.g. no grouping of the large and small amplitudes. There should be no large

overloads to cause significant crack retardation. Moreover, it must be assumed that the mean stress associated with each cycle does not influence the constants, C and m.

As described in Section 6.3, a Miner's stress  $S_e$  can be defined by Equation 6.9. Then an equivalent stress intensity factor can be defined as

$$\Delta K_e = Y S_e \sqrt{a} \quad (6.22)$$

where  $S_e = [E(S^m)]^{1/m}$ .

If the model is valid, a  $da/dn - \Delta K$  curve observed from random loading using Equation 6.22 should produce the same results as the constant amplitude curve.

#### 6.5 Damage Expressions for High Temperature Low Cycle Fatigue

*The Taira Law:* A damage expression which includes the effects of fatigue and creep is\*

$$D = \sum_{i=1}^k \left( \frac{n_i}{N_i} \right) + \sum_{i=1}^q \left( \frac{t_i}{T_i} \right) \quad (6.23)$$

where definitions of  $n_i$  and  $N_i$  are given in Section 6.2 (see Figure 6.1);  $t_i$  is the time duration of the load condition,  $i$ , and  $T_i$  is the time to rupture at load condition  $i$ . There are  $q$  discrete conditions of creep.

This model is easy for designers to use, but it does not account for synergistic effects of fatigue and creep.

*Strainrange Partitioning (SRP):* The method of strainrange partitioning (SRP) developed by Manson, Halford, and Hirschberg [3] is well documented in their report as well as the later reports of Hirschberg and Halford [4] and Bernstein [5].

\*A combination of Miner's rule and Robinson's law.

Four different types of strainranges can be defined as the basic building blocks for any conceivable hysteresis loop:

1. PP -- tensile plasticity reversed by compressive plasticity
2. CP -- tensile creep reversed by compressive plasticity
3. PC -- tensile plasticity reversed by compressive creep
4. CC -- tensile creep reversed by compressive creep.

First strain-life relationships for each of the strainrange types are established by experiment

$$N_i = A_i (\Delta \epsilon)_i^{a_i} \quad i = PP, CP, PC, CC \quad (6.24)$$

Given a hysteresis loop for fatigue stress at a point (obtained from stress analysis), the fraction of each strainrange type  $f_i$ , a component of the total inelastic strainrange is identified using an algorithm as described in References 3 and 4. For example,  $f_{pp} = \Delta \epsilon_{pp} / \Delta \epsilon_{in}$  where  $\Delta \epsilon_{pp}$  is PP strainrange and  $\Delta \epsilon_{in}$  is the total inelastic strainrange.

$$\Delta \epsilon_{in} = \sum_{i=1}^4 \Delta \epsilon_i \quad \sum_{i=1}^4 f_i = 1 \quad i = PP, CP, PC, CC \quad (6.25)$$

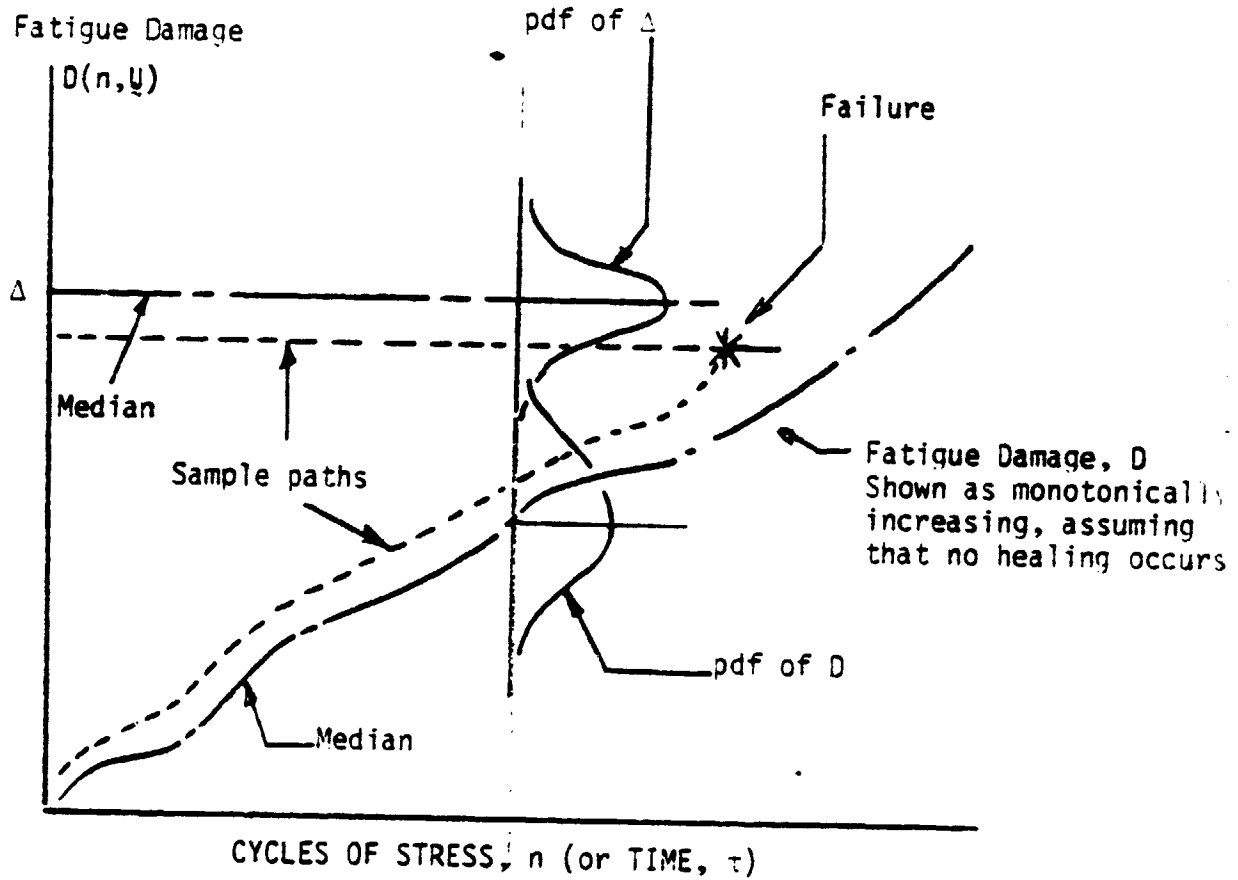
Damage after application of  $n$  cycles of the constant amplitude strain cycle, is defined as

$$D = n \sum_{i=1}^4 (f_i / N_i) \quad i = PP, CP, PC, CC \quad (6.26)$$

## 6.6 Damage at Failure--Some General Remarks

Assuming that no healing occurs in the component, fatigue damage will be a monotonically increasing function of  $n$  (or  $\tau$ ) as shown in Figure 6.4. Damage  $D$  will increase until it reaches a failure level  $\Delta$ .

Figure 6.4  
 Damage as a Function of Time\*



\* A more advanced model using fracture mechanics methods, considers strength as a process which monotonically decreases: See "Reliability Analysis of Fatigue Sensitive Aircraft Structures Under Random Loading and Periodic Inspection" by J.N. Yang and W.J. Trapp, AFML-TR-74-29, February 1974.

Note that if only one stress level is applied, failure occurs when  $D = \Delta = 1$ . Without any knowledge of the physical process of fatigue, the obvious assumption to make is that in general failure occurs when  $D = \Delta = 1$  for any sequence of stresses. Indeed this is the Palmgren-Miner hypothesis, or "Miner's Rule" as it is often called [6,7]. But recent studies have suggested that for design purposes it is useful to model  $\Delta$  as a random variable [8].

Consider a hypothetical fatigue test in which the SN behavior is deterministic as shown in Figure 6.1. Damage at failure denoted as  $\Delta_i$  is measured for  $m$  specimens subjected to different load histories as shown in Figure 6.5. If Miner's rule is exact, all  $\Delta_i = 1$ . But Miner's rule is a simple model used to describe a complicated phenomena. As a result, primarily of sequence effects, one should expect statistical scatter in the sample  $\Delta_i$ ,  $i = 1, m$ . Therefore damage at failure  $\Delta$  should in general be treated as a random variable.

In practice of course, observed values of  $\Delta_i$  will also contain scatter by virtue of basic variability in material behavior. To measure the statistics on  $\Delta$ , somehow material variability effects must be removed.

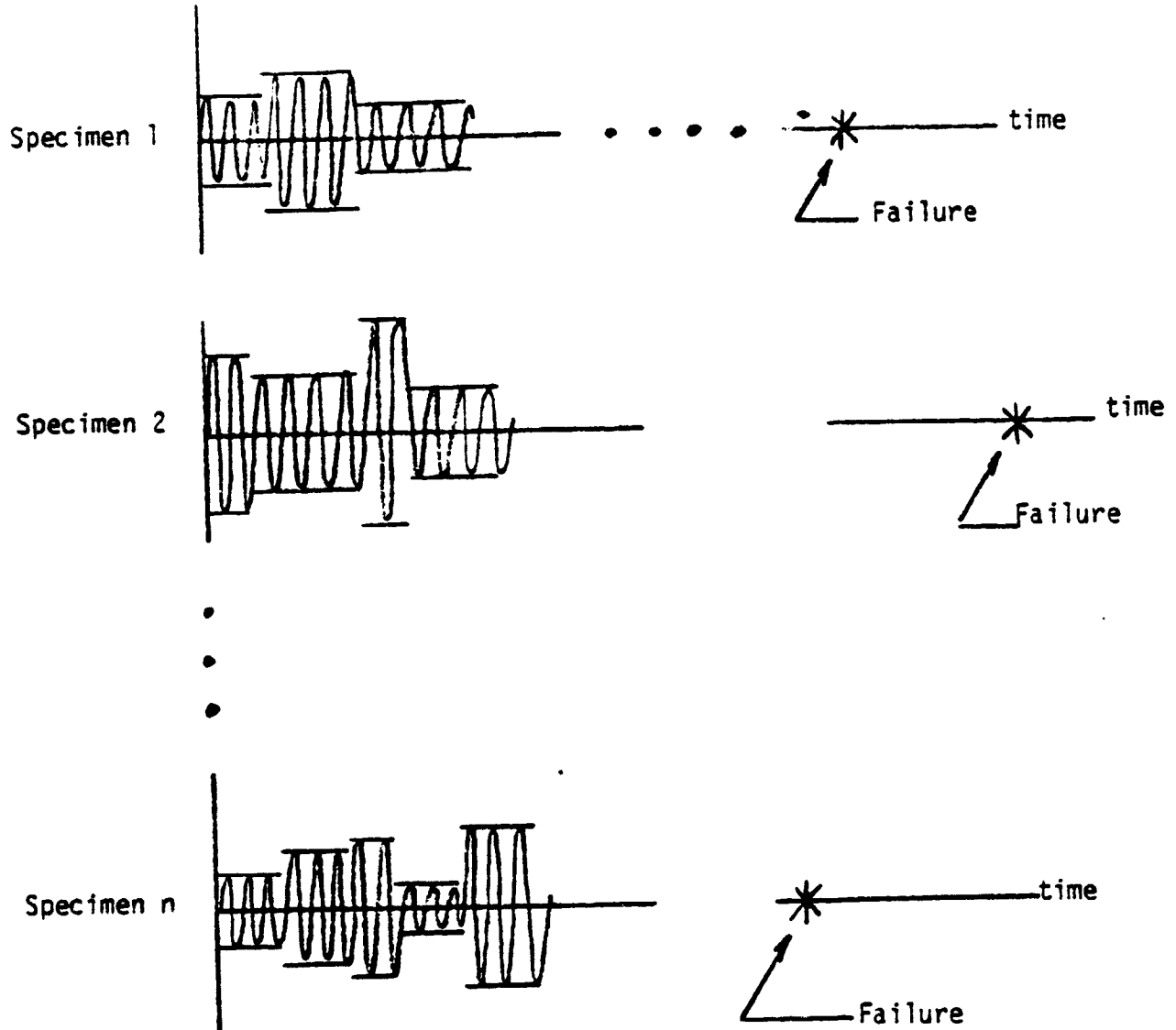
### 6.7 Fatigue Design Relationships

The design life  $N_0$  in cycles (or  $T_0$  in time) is specified. Three equivalent formulations are commonly used for design purposes.

*Formulation 1;* Damage at the design life  $N_0$  is denoted as  $D(N_0, U)$ . For a model in which the designer assumes that  $\Delta$  and  $U$  are constants, the condition for a safe design at life  $N_0$  is

$$D(N_0, U) < \Delta \quad (6.27)$$

Figure 6.5  
Fatigue Stress Sequences for Several Specimens in  
a Hypothetical Fatigue Test





To insure a safe design given statistical scatter and uncertainties, the designer can use a) a design curve on the safe side of SN data, as well as "safe-side" values for all  $U_i$  and/or b) a reduced value of  $\Delta$ , e.g.,  $\Delta = 0.30$ , rather than the  $\Delta = 1$  of Miner's rule.

Assume that  $\Delta$  and  $\underline{U}$  are treated as random variables. The density function for each at a given life is shown in Figure 6.4. The probability of failure at life  $n$  can be written as,

$$p_f = P[\Delta < D(N_0, \underline{U})] \quad (6.28)$$

Example: Consider the damage expression when stress is a stationary narrow band process (Equation 6.17). The probability of failure is

$$p_f = P[\Delta \leq (N_0/K)(\sqrt{2}\sigma)^m \Gamma(m/2 + 1)] \quad 6.29$$

where in general  $\Delta$ ,  $K$ ,  $\sigma$  and  $m$  could be random variables. The problem of evaluating  $p_f$  is described in Chapter 7.

*Formulation 2:* Let  $N$  be a random variable denoting cycles to failure. Thus  $D(N, \underline{U})$  is damage at failure,  $\Delta$

$$\Delta = D(N, \underline{U}) \quad (6.30)$$

Inverting the expression,

$$N = N(\Delta, \underline{U}) \quad (6.31)$$

Assuming that  $\Delta$  and  $\underline{U}$  are constants, the condition for a safe design is

$$N_0 < N \quad (\text{or } T_0 < T) \quad (6.32)$$

Given statistical scatter and uncertainty, a safe design can be insured by using a) a design curve on the safe side of the SN data, and/or b) conservative (high) values of stresses, and/or c) a value of  $N_0$  larger than the expected service life  $N_S$ . For example,  $N_0 = 2N_S$  has been used [9].\*

In a reliability approach, the probability of failure can be written as,

$$P_f = P[N < N_0] \quad (6.33)$$

$$= P[N(\Delta, \underline{U}) < N_0]$$

in which  $\Delta$  and  $\underline{U}$  are treated as random variables. Figure 6.6 illustrates the pdf of  $N$ , as well as  $P_f$ .

Example: Consider the damage expression of Equation 6.17 with  $n = N$  and  $D = \Delta$ . Solving for  $N$  and substituting into Equation 6.34,

$$p_f = P[\Delta K / (\sqrt{2}\sigma)^m \Gamma(m/2 + 1) \leq N_0] \quad (6.35)$$

where in general  $\Delta$ ,  $K$ ,  $\sigma$  and  $m$  would be considered as random variables.

*Formulation 3.* Another approach commonly used requires knowing the distribution of fatigue strength  $S$  at a given design life  $N_0$ . The pdf of  $S|N_0$  is illustrated in Figure 6.7. Miner's stress  $S_e$  (Equation 6.9) can be compared to strength. Miner's stress will be a function of random design factors  $\underline{U}$ .

In the special case of constant amplitude stress at level  $S_0$ ,  $S_e = S_0$ .

If the SN curve is given  $S = S(N)$ , then the condition for a safe design is

$$S_e < S(N_0)$$

\* applied to an S-N curve already on the safe side of the data

Figure 6.6

Statistical Distribution of Cycles to Failure  $N$  at Stress Level,  $S_e$

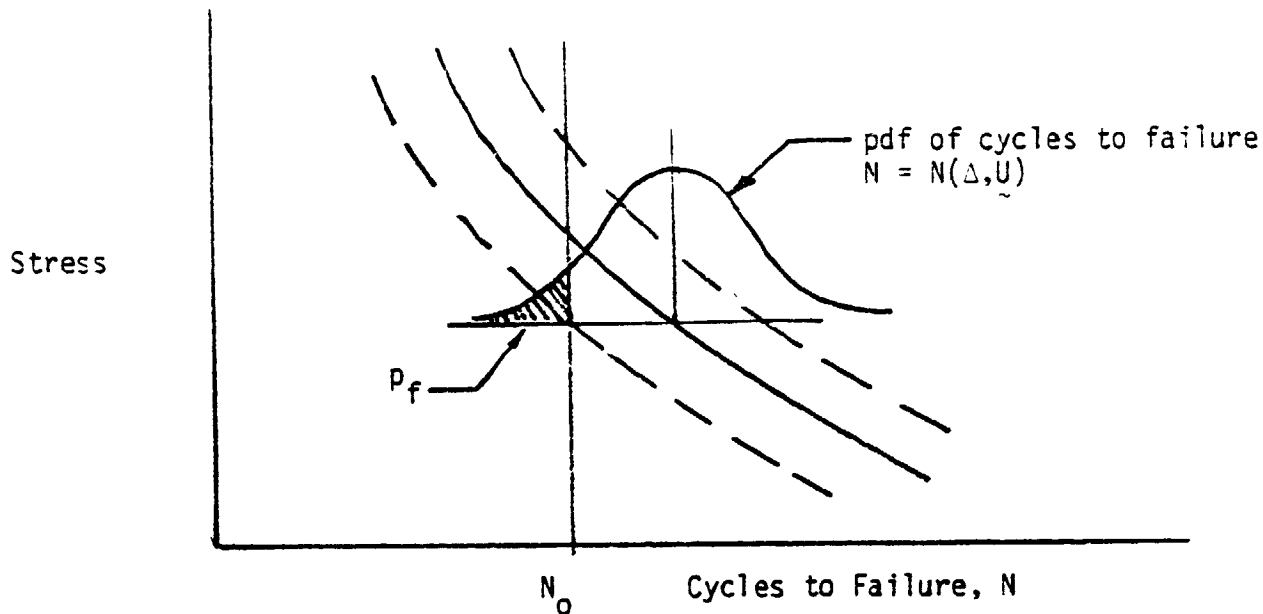
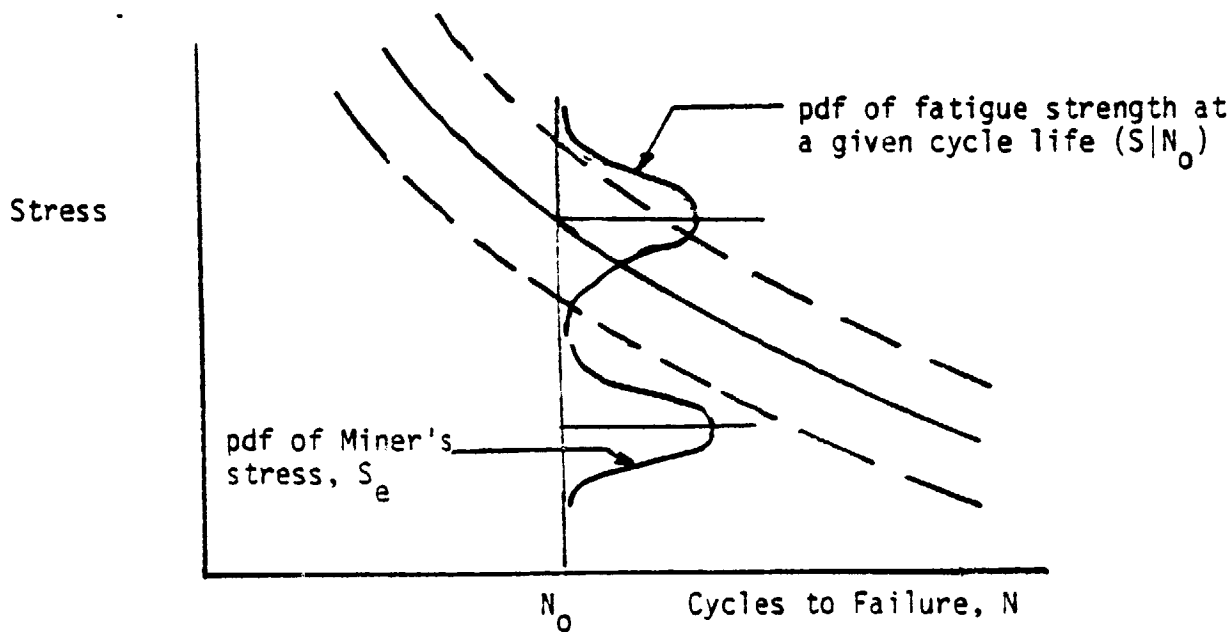


Figure 6.7

Distribution of Fatigue Strength and Stress at Design Life  $N_0$



For example, if the S-N curve is  $NS^m = K$ , then a safe design occurs when  $S_e < (K/N_0)^{1/m}$ . But, in general,  $S_e$  will be a random variable as suggested by the pdf of Figure 6.7. The probability of failure is

$$p_f = P[S < S_e(U)] \quad (6.36)$$

Example: Miner's stress for the case of a stationary narrow band Gaussian process is obtained by combining Equation 6.9 and 6.16

$$S_e = \sqrt{2[\Gamma(m/2 + 1)]}^{1/m} \sigma \quad (6.37)$$

Assume that fatigue strength at  $N_0$  is given as  $N_0 S^m = K$ . Thus the probability of failure is

$$p_f = P[(K/N_0)^{1/m} \leq \sqrt{2[\Gamma(m/2 + 1)]}^{1/m} \sigma] \quad (6.38)$$

In general,  $K$ ,  $\sigma$  and  $m$  will be random variables.

Note that this form is similar to the two above examples with the exception that  $\Delta$  is missing from this form. It's influence could be accounted for by a variable which multiplies  $S_e$ . Such a variable would describe the uncertainty associated with using Miner's stress as a characteristic stress.

### 6.8 Statistical Considerations of the Palmgren-Miner Fatigue Index

The Palmgren-Miner linear damage accumulation rule (referred to herein as "Miner's Rule") for estimating variable amplitude fatigue life is easy to use and therefore widely employed in design procedures. But because fatigue is a complicated process involving many factors, Miner's rule, a simplified description of fatigue, does not provide consistently accurate predictions.

There is a need to quantify the performance of Miner's rule to provide guidance to designers.

A statistical study of the results of fatigue experiments of several investigators was performed by Wirsching, Yao, and Stahl, and summarized in Reference 8. A summary of this study is presented here.

Composite statistical models of fatigue damage at failure (denoted as  $\Delta$ ) were developed. The lognormal distribution was used in the statistical analysis of  $\Delta$ . Such a study of the statistical variability of a relatively simple theoretical model is necessary and desirable for the development of probability-based fatigue design codes.

Define  $D$  as fatigue damage as computed by a linear damage rule. Consider a fatigue test performed by an investigator. Let  $r$  be the number of specimens in the test. Let  $\Delta_p$  be the value of  $D$  at failure for the  $p$ th specimen. To evaluate  $\Delta_p$ ,

$$\Delta_p = \frac{\text{Cycles to failure for the } p\text{th specimen}}{\text{Cycles to failure predicted by the PM rule}} \quad (6.39)$$

A sample from an investigator would be  $\Delta_p$ ;  $p = 1, r$ , with the tests in general being performed at different stress levels.

Because of the scatter observed in the data, it is suggested that the quantity  $\Delta$  be represented as a random variable. In general, failure can be defined as the event ( $D > \Delta$ ). This is a generalization of Miner's rule which states that failure occurs when  $D = 1$ . The probability of failure is given by

$$P_f = P(D > \Delta) \quad (6.40)$$

Assume that  $\Delta$  has a lognormal distribution with mean  $\mu_\Delta$ , standard deviation  $\sigma_\Delta$ , and a coefficient of variation (C.O.V.)  $C_\Delta = \sigma_\Delta/\mu_\Delta$ .

Test data on damage at failure  $\Delta$  obtained by several investigators is summarized by the distribution functions of Figure 6.8 and in Table 6.1. Following is commentary on this data summary and how it was developed.

The sample mean of  $\Delta$  obtained by the  $i$ th investigator is

$$\bar{\Delta}_i = \frac{1}{r} \sum_{p=1}^r \Delta_p \quad (6.41)$$

and is an estimate of  $\mu_{\Delta}$ . The sample standard deviation is

$$s_i = \left[ \frac{1}{r-1} \sum_{p=1}^r (\Delta_p - \bar{\Delta}_i)^2 \right]^{1/2} \quad (6.42)$$

and is an estimate of  $\sigma_{\Delta}$ . The estimate of the C.O.V. and median of the data of the  $i$ th investigator are, respectively,

$$C_i = s_i / \bar{\Delta}_i \quad (6.43)$$

$$\tilde{\Delta}_i = \bar{\Delta}_i / \sqrt{1 + C_i^2} \quad (6.44)$$

The assumption that  $\Delta$  is lognormal is based on observations of the empirical distribution functions of Figure 6.8.

### 6.8.1 The SAE Fatigue Program

The SAE Cumulative Fatigue Damage Test Program was an effort to exchange information between experts in universities and industry on methods of testing and life estimation [18]. The load histories defined were considered to be typical of those experienced by components of automotive vehicles. Test specimens of RQC-100\* and Man-Ten\*\* steel were compact tension specimens with

---

\*Bethlehem Steel; roller quenched and tempered to 100 ksi minimum yield strength.

\*\*U. S. Steel; hot rolled to 50 ksi minimum yield strength.

Figure 6.8

Empirical Distribution Functions of  
Damage at Failure from Various  
Investigators (Key to the numbered  
curves is given in Table 6.1

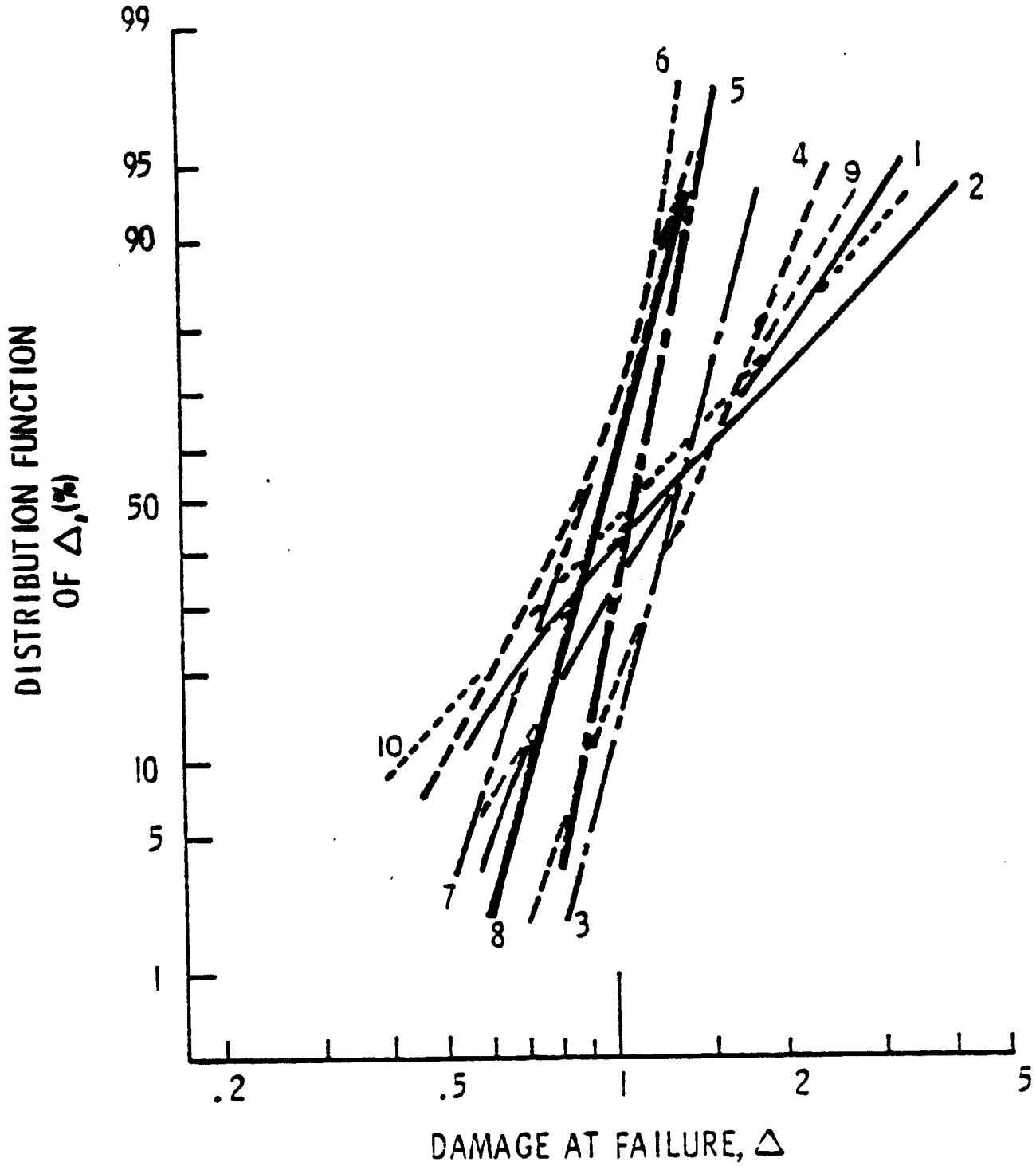


Table 6.1

Estimates of Statistics of Damage at Failure,  $\Delta$ Based on a Lognormal Model for  $\Delta$ .  
(n = number of specimens)

	Mean of $\Delta$ $\bar{\Delta}_i$	Median of $\Delta$ $\bar{\Delta}$	Standard Deviation of $\Delta$ $s_i$	Coefficient of Variation of $\Delta$ $C_i$
1. Grichlow et al. (10); an amalgamation of test data of many configurations and materials of flight vehicle structures including full scale P-51 and C-46 airplane wing test (n=266)	1.53	1.30	0.962	0.627
2. Schijve (11); amalgamation of results from 19 reports; all specimens 2024 and 7075 aluminum	1.72	1.23	1.68	0.980
3. Richart and Newmark (12); axially loaded large (3 1/2 x 1/2 in) specimens of ASTM-A7 (n=31)	1.27 (1.79)*	1.23	0.341	0.269
4. Richart and Newmark (12); rotating beam small specimens (dia. = 0.160 in) of ASTM-A7 (n=29)	1.47 (2.06)*	1.39	0.519	0.353
5. Topper, Sandor, and Morrow (13); strain controlled tests on 2024-T3 aluminum (n=18)	1.15	1.14	0.186	0.161
6. Dowling (14); 2024-T4 specimens; mostly in the elastic range, but some with large plastic strains; rainflow method used to count cycles (n=83)	0.363	0.323	0.271	0.314
7. Topper and Sandor (15); strain controlled tests on 2024-T4 aluminum with tensile and compressive mean stresses (n=11)	0.336	0.309	0.219	0.262
8. Miner (7); tests run at two or more stress levels on 245-T Alcad (n=11)	0.980	0.349	0.251	0.256
9. Schilling (16); welded cover plated beam specimens of A36 and ASTM A514 under simulated random bridge loading (n=36)	1.44	1.25	0.78	0.54
10. SAE Fatigue Design & Evaluation Committee; tests on Man-Ten & RQC-100; notch specimen with cyclic plasticity at notch root; See Table III; (n=54)	1.46	1.09	1.30	0.389
Swanson (17); an amalgamation of random fatigue test data (n=671); Values based on an assumption that $C_{\Delta} = 0.50$ , and $F_{\Delta}(1.0) = 0.47$	1.21	1.04	0.726	0.50

\*Estimates of  $\Delta_i$  using a "minimum S-N curve"



a keyhole notch. Failure was said to occur at crack initiation, which was arbitrarily defined as a crack of 2.54 mm (0.10 inch). A total of 54 specimens were tested.

Results of several investigators in the SAE program, using various analysis methods to predict cycles to failure, all using the PM rule, are presented in Table 6.2. Data from both materials are assumed to be from the same population. No attempt was made to judge the quality of the various analysis methods in computing the average values at the bottom of Table 6.2. These average values were used in Table 6.1.

### 6.8.2 A Statistical Composite of $\Delta$

The goal of this study is to provide information to designers. To synthesize, for design purposes, the data of Table 6.1, statistics of composite data from two or more investigators, the following forms were used.

$$\bar{\Delta} = \frac{1}{n} \sum_{i=1}^m n_i \bar{\Delta}_i \quad (6.45)$$

$$s_{\Delta}^2 = \frac{1}{n-1} \sum_{i=1}^m [(n_i - 1)s_i^2 + n_i(\bar{\Delta}_i - \bar{\Delta})^2] \quad (6.46)$$

where  $\bar{\Delta}$  is the sample mean of  $\Delta$ ,  $\bar{\Delta}_i$  is the sample mean obtained by the  $i$ th investigator (as given in Table 6.1),  $n_i$  is the sample size of the  $i$ th investigator,  $m$  is the number of investigators,  $s_{\Delta}^2$  is the sample variance of  $\Delta$ ,  $s_i^2$  is the sample variance obtained by the  $i$ th investigator using  $\bar{\Delta}_i$  (this is what is given in Table 6.1), and  $n$  is the total number of observations by all investigators. Table 6.3 and Figure 6.9 provide summaries (using Equations 6.45 and 6.46) of the analysis of the data listed in Table 6.1.

Table 6.2

Summary Statistics on  $\Delta$  From  
SAE Cumulative Fatigue Damage  
Test Program (18)

	$\mu_{\Delta}$	$\bar{\Delta}$	$\sigma_{\Delta}$	$C_{\Delta}$
<u>Socie and Morrow (19)</u>				
Nominal strain life analysis	.885	.568	1.06	1.19
Component calibration (incl mean)	1.17	.891	1.00	.855
Neuber notch analysis	.90	.693	.745	.828
Component calibration	1.21	.884	1.14	.939
<u>Dowling, Brose, Wilson (20)</u>				
Base line case	1.20	.964	.898	.746
Base line case (Smith mean par.)	1.15	.922	.848	.739
Load strain extrapolation of elastic analysis	.687	.505	.632	.920
Load strain calibration from Neuber's rule	2.56	2.08	1.82	.713
Load strain calibration from finite element analysis	1.23	.967	.955	.779
Constant amplitude load life	.677	.561	.457	.674
<u>Landgraf, Richards, La Pointe (21)</u>				
Neuber notch analysis	2.79	2.22	2.11	.756
Load local strain analysis	2.69	2.16	2.01	.748
Strain based analysis	1.93	1.51	1.53	.794
<u>Brose (22)</u>				
Incl mean stress and overstrain effects	1.43	.970	1.56	1.09
AVERAGE VALUES	1.46	1.09	1.30	.889

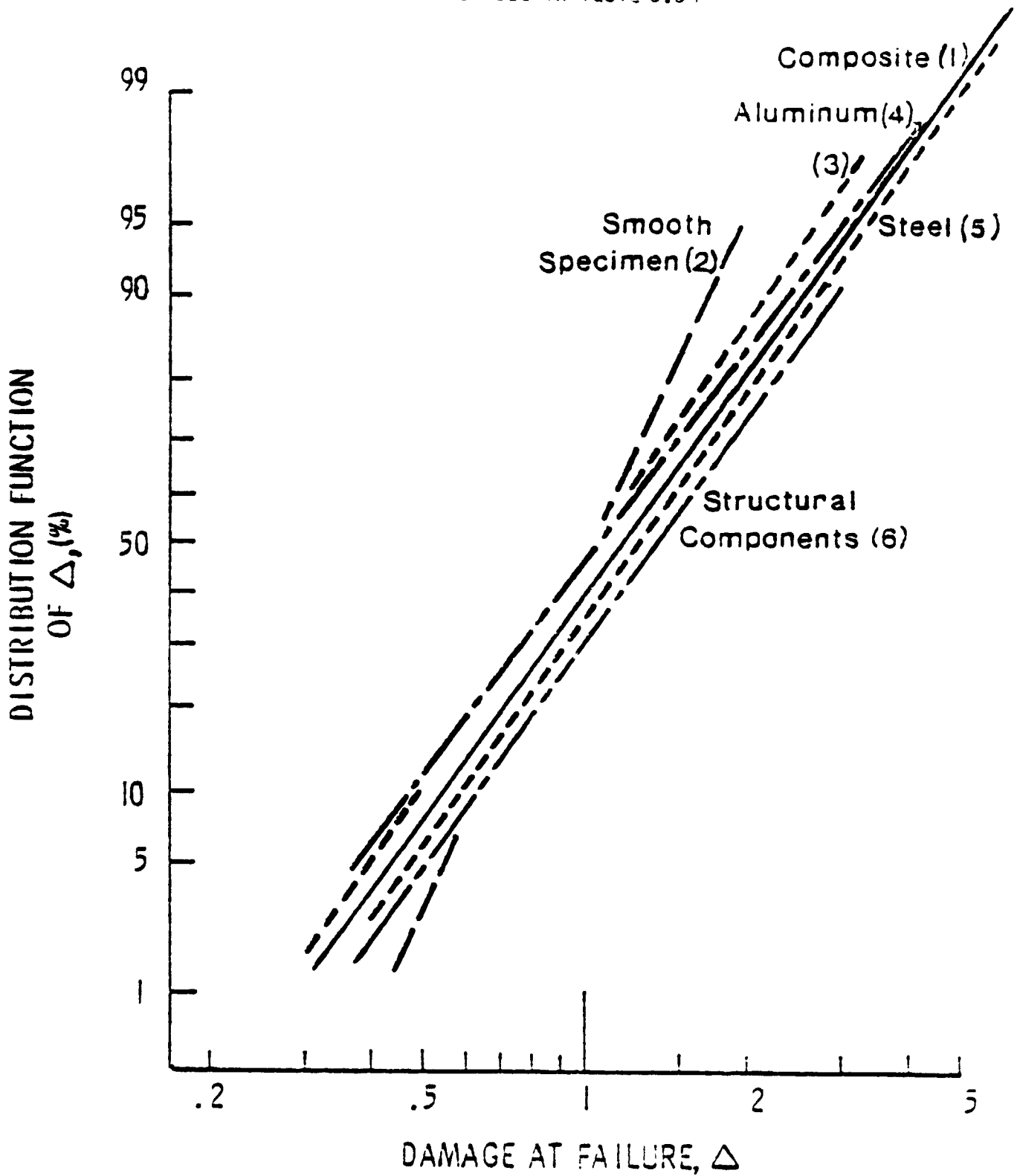
TABLE 6.3

A Statistical Summary of Models  
of Composite Data

	Mean $\mu_{\Delta}$	Median $\tilde{\Delta}$	C.O.V. ( $C_{\Delta}$ )	% less than 1.0
1. Composite (all data; n=537)	1.35	1.13	0.643	42
2. Smooth specimen composite (all data except Crichlow Schilling, and SAE; n=183)	1.06	0.997	0.380	51
3. Composite (all data except Crichlow's 266 points; n=271)	1.17	1.01	0.612	49
4. Aluminum (all data; n=389)	1.33	1.11	0.650	43
5. Steel (all data; n=148)	1.37	1.16	0.638	40
6. Full scale structural components (Crichlow & Schilling's data, n=300)	1.50	1.27	0.620	34

Figure 6.9

Empirical Distribution Functions of Statistical  
Composites of Damage at Failure  
(Models are described in Table 6.3)



### 6.8.3 Analysis of the Variability in $\Delta$

For design purposes, it would be desirable to quantify the contributions of various factors to the total variability in  $\Delta$ . However, much of the detail regarding experimental procedure is lost in the data of Figure 6.6 and Table 6.1. Nevertheless, this data as a whole can be interpreted as an amalgamation of experiences and, as such, implicitly includes effects of many of the relevant factors. Values of coefficients of variation for material strength, e.g., yield strength of steel, ultimate strength of composites, etc. are typically less than 0.15 and, therefore, the values of  $C_i$  listed in Tables 6.1 and 6.2 suggest a relatively large scatter in the data.

Assuming the  $\Delta$  is the product of several random factors whose distributions are lognormal, it can be shown that

$$1 + C_{\Delta}^2 = \prod_{i=1}^j (1 + C_i^2) \quad (6.47)$$

where  $j$  is the number of factors which influence  $\Delta$ , and  $C_i$  is the coefficient of variation of the  $i$ th factor. It is assumed that all factors are statistically independent.

Often available is statistical information on fatigue variability inherent in the material, described by  $C_N$ . Suppose that this is the only factor for which statistical information is available. Then Equation 6.47 reduces to

$$1 + C_{\Delta}^2 = (1 + C_N^2) (1 + C_0^2) \quad (6.48)$$

where  $C_0$  is the coefficient of variation of all of the other factors.

In general, it is difficult to separate the effect of material variability from statistics on  $\Delta$  using the method described above. In one example, the writer is suggesting that reasonable statistics for  $\Delta$  when material variability is considered elsewhere is

$$\text{Median, } \Delta = 1.00$$

$$\text{Cov, } C_{\Delta} = 0.30$$

with  $\Delta$  being lognormal [8].

### 6.9 Statistical Considerations of Damage at Failure for Strainrange Partitioning

Statistical analysis of damage at failure  $\Delta$  for some strainrange partitioning (SRP) test data is presented in this section. This exercise was intended to provide necessary information for a total reliability analysis. Furthermore, statistics on  $\Delta$  provide a quantitative measure of the performance of SRP.

Fatigue damage after  $n$  cycles is (see Section 6.5 for definition of terms)

$$D = n \sum_{i=1}^4 (f_i / N_i) \quad i = \text{PP, CP, PC, CC} \quad (6.49)$$

Note that, in general,  $f_i$  and  $N_i$  are random variables so that  $D$  also is a random variable.

The SRP model defines a "predicted life" (cycles to failure)  $N_p$ . In Equation 6.49 let  $D = 1$  and  $n = N_p$  at failure; choose the median, or best estimate, values  $f_i$  and  $N_i$ . Then the predicted life is

$$\frac{1}{N_p} = \sum_i (\tilde{f}_i / \tilde{N}_i) \quad i = \text{PP, CP, PC, CC} \quad (6.50)$$

Consider a hypothetical experiment where several specimens are tested, having the same  $N_p$  but different proportions of strainrange types  $f_i$ , and therefore different lives  $N_i$ . The cycles to failure  $N_f$  for each specimen is recorded. The scatter in the results reflects the uncertainty in a) the life relationship,  $N_i$ , i.e., scatter in  $\Delta\epsilon$ - $N$  data, for each strainrange type, b) the process of computing the partitioned strainranges,  $f_i$ , and c) in the performance of the linear damage model. Thus,  $N_f$  is a random variable depending upon these three fundamental sources of uncertainty.

Let damage at failure be a random variable denoted as  $\Delta_0$ . This is the random variable which describes the uncertainty in the performance of the linear damage model. At failure  $D = \Delta_0$  and  $n = N_f$ . An expression for  $N_f$  can be written as (from Equation 6.49)

$$N_f = \Delta_0 / \sum_i (f_i / N_i) \quad (6.51)$$

$\Delta_0$ ,  $f_i$ , and  $N_i$  ( $i = 1, 4$ ) are random variables representing the three sources of uncertainty mentioned above. The distribution of  $\Delta_0$  reflects only the inaccuracies of the SRP model. Unfortunately, there is no convenient way to measure  $\Delta_0$ . However, it is possible from test data to obtain some information on SRP performance.

Equation 6.51 can be reformulated using the median (and constant) values  $\tilde{f}_i$  and  $\tilde{N}_i$ .

$$N_f = \Delta / \sum (\tilde{f}_i / \tilde{N}_i) \quad (6.52)$$

Now  $\Delta$  includes the variability of  $f_i$  and  $N_i$ , as well as model uncertainty. However,  $\Delta$  can be measured directly from fatigue data. Substituting Equation 6.50, it follows that

$$\Delta = N_f/N_p \quad (6.53)$$

Given a single test specimen, the value of  $\Delta_i$  is

$$\Delta_i = N_{OBS(i)}/N_{PRE(i)} \quad (6.54)$$

where  $N_{OBS(i)}$  = observed cycles to failure and  $N_{PRE(i)}$  = predicted cycles to failure. These values are available in the literature for SRP tests of various types. Table 6.4 summarizes the results of statistical analyses of several tests. For a given sample,  $\Delta_i$ ,  $i = 1, r$ , the statistics were obtained using Equation 6.41 through 6.44.

Following is a brief description of the differences in baseline and verification tests as noted in Table 6.4.

Baseline Tests - Refer to the high temperature low cycle tests used directly in the establishment of the four SRP inelastic strainrange versus life ( $\Delta\epsilon_{in} - N$ ) relationships.

Verification Tests - Refers to the non-baseline high temperature low cycle fatigue tests and results used to check how well the established SRP relationships can be used to predict cyclic lives. A primary requisite of a verification test is that it should contain some feature or complexity not present in the baseline test.

Finally it should be noted again that the scatter in  $\Delta$  as reported by Table 6.4 includes uncertainties due to  $f_i$  and  $N_i$  as well as uncertainties in the SRP model. It is generally desirable to separate these uncertainties ( to obtain the distribution of  $\Delta_0$ ) in order to perform a reliability analysis. No attempt is made to do so in this study.



Table 6.4

A Statistical Summary of SRP Data on Damage at Failure  $\Delta$ 

	Sample Size, n	Sample Mean	Sample Std. Dev.	Median	COV
AISI 316; 705C (1300F) Baseline tests Ref: Saltsman and Halford (23)	25	1.25	1.11	.933	.891
AISI 316 Verification tests Ref: Saltsman and Halford (23)	66	2.87	5.81	1.27	2.03
AISI 304 Verification tests Ref: Saltsman and Halford (23)	110	1.65	1.89	1.08	1.15
AISI 304 and 316 All verification test points Ref: Saltsman and Halford (23)	176	3.02	6.96	1.21	2.30
AISI 316 PP + PC + CC data only Ref: Saltsman and Halford (23)	9	1.86	2.15	1.21	1.16
AISI 316 PP + CP + CC data only Ref: Saltsman and Halford (23)	8	12.4	17.0	7.29	1.37
AISI 316 PP + CP data only Ref: Saltsman and Halford (23)	38	2.67	5.09	1.24	1.91
AISI 316 PP + PC data only Ref: Saltsman and Halford (23)	5	2.91	3.89	1.74	1.34
AISI 304 PP + CP data only Ref: Saltsman and Halford (23)	92	1.71	2.03	1.10	1.19

...Table 6.4 continued

	Sample Size, n	Mean	Std. Dev.	Median	COV
AISI 304 PP + PC data only Ref: Saltsman and Halford (23)	9	1.54	1.16	1.23	.750
1Cr-1Mo-.25V; 540C Normalized and tempered Ref: Saltsman and Halford (23)	26	1.00	.409	.928	.408
1Cr-1Mo-.25V; 485C Normalized and tempered Ref: Saltsman and Halford (23)	12	1.27	.293	1.24	.231
2.25Cr-1Mo; 540C Annealed Ref: Saltsman and Halford (23)	17	1.21	.241	1.19	.198
2.25Cr-1Mo; 540C Normalized and tempered Ref: Saltsman and Halford (23)	17	0.820	.190	.799	.232
2.25Cr-1Mo; 485C Quenched and tempered Ref: Saltsman and Halford (23)	12	0.615	.222	.579	.360
AISI 304; 650C Solution annealed Ref: Saltsman and Halford (23)	15	0.899	.367	.832	.408
AISI 304; 565C Solution annealed Ref: Saltsman and Halford (23)	13	0.467	.185	.436	.395
Rene 95; 1200F Baseline Tests Ref: Hyzak and Bernstein (24)	42	1.25	.830	1.04	.663
Rene 95 Verification Tests Ref: Hyzak and Bernstein (24)	13	0.920	.479	.816	.520
Rene 95 Validation Tests Ref: Hyzak and Bernstein (24)	16	1.13	.327	1.08	.290

#### 6.10 Statistical Distribution of $\Delta$

The least squares program for distribution analysis as described in Section 2.5 was used to analyze the data on  $\Delta$ . The results are summarized in Table 6.5.

The model which best fits the data for each sample is indicated by the box in Table 6.5. The results show clearly that the lognormal is generally the best fit model for the samples considered. It should be noted that this implies that it is reasonable to assume that cycles to failure is lognormally distributed.

#### 6.11 Summary Comments on Chapter 6

Provided in this chapter were equations, using the concept of damage, which are used by designers to predict fatigue failure under both variable and constant amplitude stresses. Statistical information on the distribution of damage at failure was also presented. Summary statistics provide a quantitative description of the performance of Miner's rule. Statistics were also provided to describe the behavior of the SRP model.

Given that uncertainty exists in all of the fatigue design factors, the question now is how to use the damage equations to make decisions which will produce a high reliability product. Chapter 7 will summarize elements of modern structural and mechanical reliability theory, and Chapter 8 will address specifically the problem of fatigue reliability.

Table 6.5

A Summary of the Best Fit Distribution for SRP Data on Damage  
at Failure Using the Least Squares Method

		$\epsilon = \sqrt{1 - \rho^2}^*$			
		NORMAL	LOGNORMAL	WEIBULL	
	Sample Size, n				
AISI 316; 705C (1300F) Baseline tests Ref: Saltsman and Halford (23)		25	.551	.197	.280
AISI 316 Verification tests Ref: Saltsman and Halford (23)		66	.731	.108	.222
AISI 304 Verification tests Ref: Saltsman and Halford (23)		110	.553	.124	.178
AISI 304 and 316 All verification test points Ref: Saltsman and Halford (23)		176	.772	.147	.237
AISI 316 PP + PC + CC data only Ref: Saltsman and Halford (23)		9	.484	.179	.146
AISI 316 PP + CP + CC data only Ref: Saltsman and Halford (23)		8	.485	.332	.320
AISI 316 PP + CP data only Ref: Saltsman and Halford (23)		38	.715	.141	.251
AISI 316 PP + PC data only Ref: Saltsman and Halford (23)		5	.578	.282	.370
AISI 304 PP + CP data only Ref: Saltsman and Halford (23)		92	.566	.152	.204

\* smallest value of  $\epsilon$  indicates that distribution which best fits the data;  
the box indicates the best fit.

...Table 6.5 continued

	Sample Size, n	NORMAL	LOGNORMAL	WEIBULL
AISI 304 PP + PC data only Ref: Saltsman and Halford (23)	9	.304	.423	.365
1Cr-1Mo-.25V; 540C Normalized and tempered Ref: Saltsman and Halford (23)	26	.313	.228	.257
1Cr-1Mo-.25V; 485C Normalized and tempered Ref: Saltsman and Halford (23)	12	.260	.417	.296
2.25Cr-1Mo; 540C Annealed Ref: Saltsman and Halford (23)	17	.247	.285	.291
2.25Cr-1Mo; 540C Normalized and tempered Ref: Saltsman and Halford (23)	17	.167	.245	.188
2.25Cr-1Mo; 485C Quenched and tempered Ref: Saltsman and Halford (23)	12	.272	.231	.238
AISI 304; 650C Solution annealed Ref: Saltsman and Halford (23)	15	.420	.265	.434
AISI 304; 565C Solution annealed Ref: Saltsman and Halford (23)	13	.221	.212	.269
Rene 95; 1200F Baseline Tests Ref: Hyzak and Bernstein (24)	42	.622	.286	.421
Rene 95 Verification Tests Ref: Hyzak and Bernstein (24)	13	.218	.414	.268
Rene 95 Validation Tests Ref: Hyzak and Bernstein (24)	16	.171	.109	.191

## 6.12 References for Chapter 6

1. Hines, W. W. and Montgomery, D. C., Probability and Statistics in Engineering and Management Science, 2nd Ed., Wiley, 1980.
2. Miles, J. W., "On Structural Fatigue under Random Loading," Journal of the Aeronautical Sciences, Vol. 21, Nov. 1954.
3. Manson, S. S., Halford, G. R., and Hirschberg, M. H., "Creep-Fatigue Analysis by Strainrange Partitioning," NASA TM X-67838, 1971.
4. Hirschberg, M. H. and Halford, G. R., "Use of Strainrange Partitioning Predict High Temperature Low Cycle Fatigue Life," NASA TN D-8072, 1976.
5. Bernstein, H. L., "An Evaluation of Four Current Models to Predict the Creep-Fatigue Interaction of Rene 92.," AMFL-TR-79-4075, Wright-Patterson AFB, Ohio, June 1979.
6. Palmgren, A., "Die Lebensdauer von Kugellagern," Zeitschrift des Vereines, Deutsches Ingenieure, Vol. 68, No. 14, 1924.
7. Miner, M. A., "Cumulative Damage in Fatigue," Transactions in the ASME, Vol. 67, 1945.
8. Wirsching, P. H., "Fatigue Reliability in Welded Joints of Offshore Structures," International Journal of Fatigue, April 1980.
9. API Recommended Practice for Planning Designing and Constructing Fixed Offshore Platforms, API RP2A, American Petroleum Institute, Dallas, TX, Jan 1981.
10. Crichlow, W. J., et al., "An Engineering Evaluation of Methods for the Prediction of Fatigue Life in Airframe Structures," Report No. ASD-TR-61-434, Wright-Patterson AFB, Ohio, 1962.
11. Schijve, J., "Estimation of Fatigue Performance of Aircraft Structures," ASTM STP 338, American Society for Testing and Materials, 1962.
12. Richart, F. E. and Newmark, N. M., "A Hypothesis for the Determination of Cumulative Damage in Fatigue," Proceedings of the ASTM, Vol. 48, 1948.
13. Topper, T. H., Sandor, B. I. and Morrow, J. D., "Cumulative Fatigue Damage under Cyclic Strain Control," Journal of Materials, Vol. 4, No. 1, March 1969.
14. Dowling, N. E., "Fatigue Failure Predictions for Complicated Stress-Strain Histories," Journal of Materials, Vol. 7, No. 1, March 1972.

References for Chapter 6 - continued

15. Topper, T. H. and Sandor, B. I., "Effects of Mean Stress and Pre-strain on Fatigue Damage Summation," Effects of Environment and Complex Load History on Fatigue Life, ASTM STP 462, 1970.
16. Schilling, C. G., et al., "Fatigue of Welded Steel Bridge Members under Variable-Amplitude Loadings," Research Results Digest No. 60, Highway Research Board, April 1974.
17. Swanson, S. R., "Random Load Fatigue Testing: A State of the Art Survey," Materials Research and Standards, Vol. 8, No. 4, April 1968.
18. Fatigue under Complex Loading, SAE Publication AE-6, 1977.
19. Socie, D. F. and Morrow, J. D., "Review of Contemporary Approaches to Fatigue Damage Analysis," FCP Report No. 24, University of Illinois, Dec 1976.
20. Dowling, N. E., Brose, W. R., and Wilson, W. K., "Notched Member Fatigue Life Predictions by the Local Strain Approach," Fatigue under Complex Loading, SAE Publication AE-6, 1977.
21. Landgraf, R. W. Richards, F. D. and LaPointe, N. R., "Fatigue Life Predictions of a Notched Member under Complex Load Histories," Fatigue under Complex Loading, SAE Publication AE-6, 1977.
22. Brose, W. R., "Fatigue Life Predictions for a Notched Plate with Analysis of Mean Stress and Overload Effects," Fatigue under Complex Loading, SAE Publication AE-6, 1977.
23. Saltsman, J. F. and Halford, G. R., "Application of Strainrange Partitioning to Prediction of Creep Fatigue Lives of AISI Types 304 and 316 Stainless Steel, ASME Paper No. 76-PVP-26, 1976.
24. Hyzak, J. M. and Bernstein, H. L., "An Analysis of the Low Cycle Fatigue Behavior of the Superalloy Rene 95 by Strainrange Partitioning," AFML-TR-78-174.

## Chapter 7 A REVIEW OF MODERN PROBABILISTIC DESIGN THEORY

### 7.1 Introduction

An example of a fatigue reliability problem is provided in Table 7.1. The design factors are treated as random variables reflecting uncertainties due to statistical scatter in observations as well as assumptions made in the analysis. The probability of failure is computed by solving a probability problem. If  $p_f$  is less than  $p_0$ , a predetermined maximum allowable risk, the design is considered to be safe.

It is assumed that the condition for failure is described by an algebraic expression. For example, from Table 7.1, the condition for fatigue failure is

$$YK_0[Z(6QL/bh^2)]^{-m} \leq N_0$$

Each term on the left hand side could be treated as a random variable. The general goal of design is to select design variables so that the probability of occurrence e.g. the above event, is acceptably small over the design life.

It is the goal of this chapter to provide a concise summary of the tools available for reliability analysis, i.e. computing the reliability or the probability of failure of an index thereof. A number of methods are reviewed. No specific recommendation is made regarding a preferred approach. Indeed a reliability analysis must be tailored to the design procedure, the data available, and the way in which the results are to be used.

The following summary of modern probabilistic design theory is based on the articles by Hasofer and Lind (1), Ditlevsen and Skov (2), and the reports of Ellingwood, Galambos, MacGregor, and Cornell (3), Comite Europeen du Beton No. 112 (particularly the commentary by Rackwitz) (4), and CIRIA No. 63 (5). Many of the operations described are application of elementary probability theory. Therefore

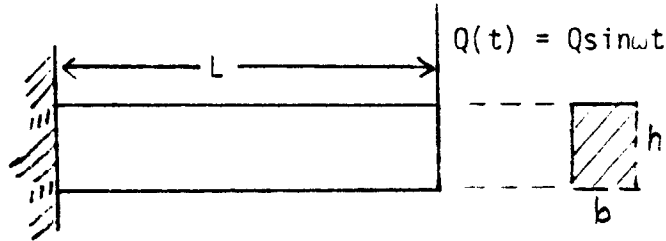


Table 7.1

A SUMMARY OF A FATIGUE RELIABILITY PROBLEM

GIVEN:

(a) Stress



$$\text{Stress Amplitude, } S_0 = \frac{Mc}{I} = \frac{6QL}{bh^2}$$

(b) Fatigue Strength of Material;  $NS^m = K$

where  $N$  = number of cycles to failure

(c) Design Life,  $N_0$

(d) Design factors ( $Q, L, b, h, K, m$ )

are in general, random variables.

(e) Non-statistical uncertainty exists

i Due to the effects of manufacturing and assembly operations and use in service

$$K = YK_0$$

ii Due to assumptions made to compute fatigue stresses

$$S = ZS_0$$

...  $Y$  and  $Z$  are random variables which describe this "professional" uncertainty

FIND The probability of failure,  $p_f$

FORMULATION The event of failure is  $N \leq N_0$ . The probability of failure is

$$p_f = P[N \leq N_0]$$

Substituting into the above expression for  $N$ ,

$$p_f = P\{YK_0[Z(6QL/bh^2)]^{-m} \leq N_0\}$$

We are left with what is, in general, a difficult probability problem. Moreover we must also define a safe level of risk  $p_0$ . The design is acceptable if  $p_f \leq p_0$ .

the reader is encouraged to have available a good elementary text in probability theory and statistics, e.g. Meyer (6), Hines and Montgomery (7), Bowker and Lieberman (8), etc.

## 7.2 Basic Definitions

A basic assumption in the following discussions is that the failure condition can be written as an algebraic function,

$$R(\underline{X}) \leq S(\underline{Y}) \quad (7.1)$$

where R and S, functions of several random design factors  $\underline{X}$  and  $\underline{Y}$ , represent strength and stress functions respectively. The assignment of factors to R and S is often arbitrary. The tilde underneath the variable indicates that the quantity is a vector (or multivalued). The design factors  $\underline{U} = (\underline{X}, \underline{Y})$  can be separated into the known random variables  $\underline{V}$  and the design parameters to be determined  $\underline{A}$ , e.g., cross sectional area, sectional modulus, etc. Thus  $\underline{U} = (\underline{V}, \underline{A})$ . Following are the basic assumptions regarding  $\underline{U}$ .

It will be assumed for simplicity that the standard deviation  $\sigma_i = 0$  for each  $A_i$ . Geometric variances are usually small compared to others. Furthermore, it is assumed that all  $U_i$  are independent. If two or more design factors are correlated, then a transformation to uncorrelated variables may be made (1).

Define the failure surface or limit state as

$$g(\underline{U}) = R(\underline{X}) - S(\underline{Y}) = 0 \quad (7.2)$$

The failure surface is the boundary between the safe and the failed regions in design parameter space. In general, there can be several limit states for a component, e.g., yield, buckling, fatigue, etc.

The failure function,  $g$ , is defined as

$$g(\underline{U}) = R(\underline{X}) - S(\underline{Y}) \quad (7.3)$$

and the failure condition (or failure event) occurs when  $S(\underline{Y})$  exceeds  $R(\underline{X})$ . Thus the failure condition is,

$$g(\underline{U}) \leq 0 \quad (7.3)$$

For example, Figure 7.1 shows a single element, single load structure with random design factors  $R$  and  $S$ . Here  $R$  and  $S$  are considered the basic variables. The overlap in pdf's suggests that it is possible to select  $R$  and  $S$  at random such that  $R < S$  and failure occurs.

Design parameter space for this case is shown in Figure 7.2. The limit state is  $R = S$ . The failure condition is  $R - S \leq 0$ , or simply  $R < S$ .

### 7.3 The Full Distributional Approach

It is assumed that the exact statistical distribution of each design factor  $U_i$  is known.

The probability of failure (or risk)  $p_f$  is defined as the probability of the failure event,

$$p_f = P(g(\underline{U}) \leq 0) \quad (7.4)$$

For example, if  $R$  and  $S$  are the basic variables, the probability of failure is  $p_f = P(R - S \leq 0)$  or  $P(R \leq S)$ .

In general,  $p_f$  can be computed as (3)

$$p_f = \int_{\Omega} f_{\underline{U}}(u_1, \dots, u_k) du_1 \dots du_k \quad (7.5)$$

where  $\Omega$  is the region of  $\underline{U}$  where  $g(\underline{U}) \leq 0$ ,  $f_{\underline{U}}$  is the joint pdf of  $\underline{U}$ , and  $k$  is

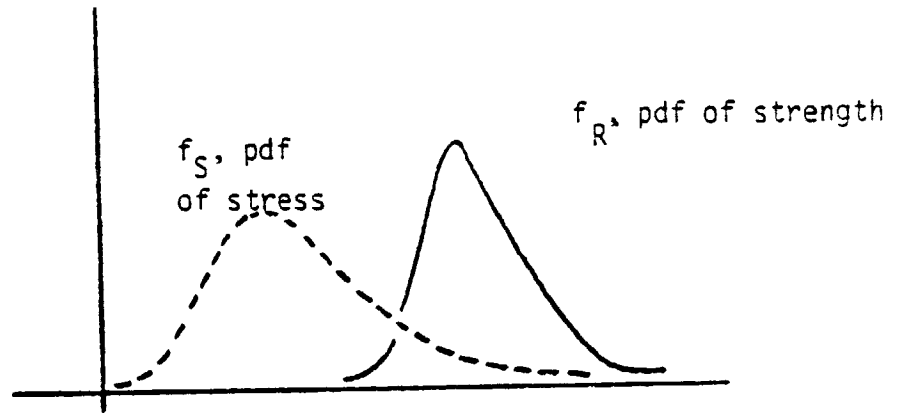
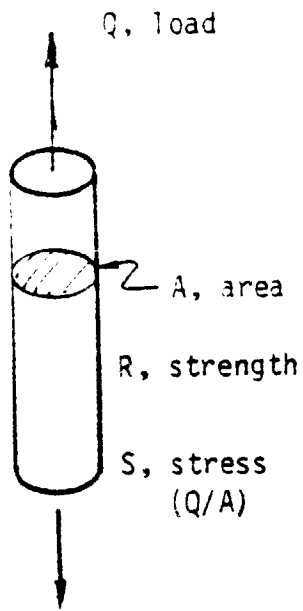


Figure 7.1 Probability Density Functions (pdf) of Stress and Strength

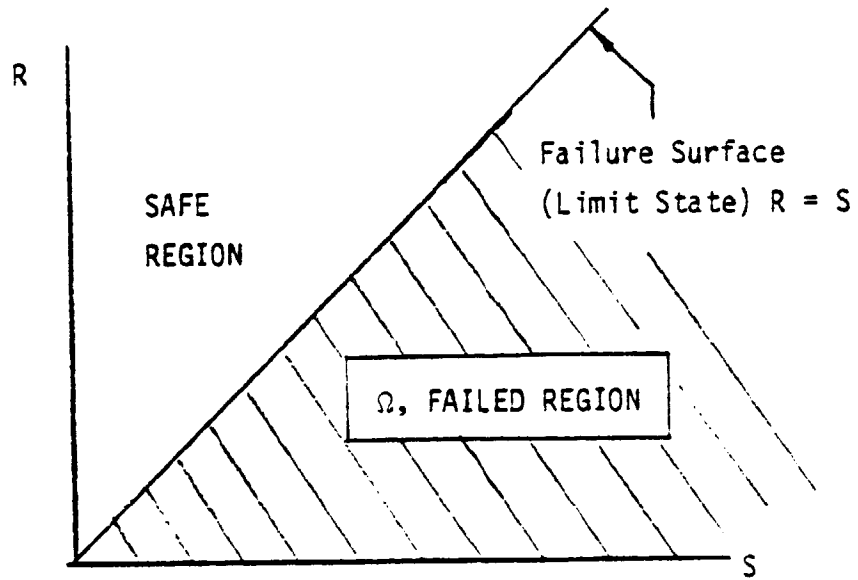


Figure 7.2 Design Parameter Space when R and S are the Basic Variables.

the number of design factors. In the special case where the two basic variables are (R,S) and statistically independent, it can be shown that

$$\begin{aligned}
 p_f &= 1 - \int_0^{\infty} f_R(r)F_S(r) dr \\
 &= \int_0^{\infty} f_S(s)F_R(s) ds
 \end{aligned}
 \tag{7.6}$$

where  $f(\cdot)$  = probability density function and  $F(\cdot)$  = cumulative distribution function. The subscripts denote the random variables.

*Examples of Closed Form Solutions*

(a) Only one random variable. If  $S = S_0$ , a deterministic real number, then Equation 7.6 becomes,

$$p_f = P(R \leq S_0) = F_R(S_0) \tag{7.7}$$

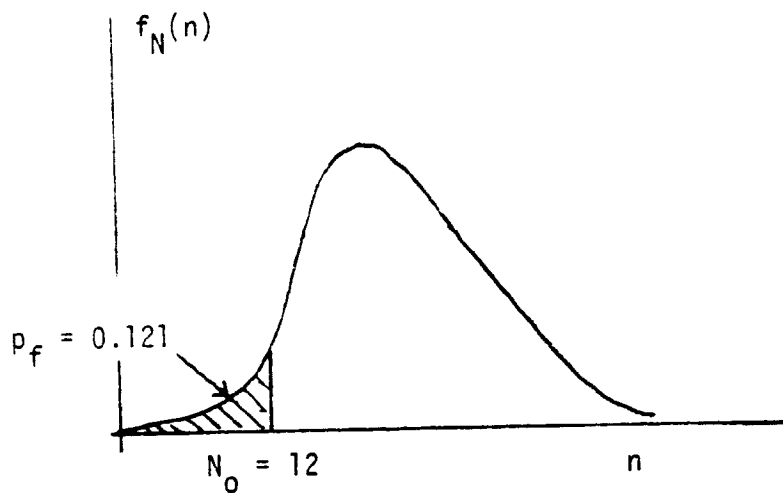
and if  $R = R_0$ , deterministic with only S random,

$$p_f = P(R_0 < S) = 1 - F_S(R_0) \tag{7.8}$$

Example: Cycles to failure N of a component is known to have a Weibull distribution with parameters  $\alpha = 3.58$  and  $\beta = 21.25$  thou-cycles. N is a "strength" variable. The design life ("stress") is given as  $N_0 = 12$  thou-cycles. Compute the probability of failure, i.e. if a member is selected at random, what is the probability that its life will be less than 12 thou-cycles?

$$\begin{aligned}
 p_f &= (N < N_0) = F_R(N_0) \\
 &= 1 - \exp[-(12/21.25)^{3.58}] \\
 &= 0.121
 \end{aligned}
 \tag{7.9}$$

The pdf of N and p is illustrated in the following figure.



(B) The Normal Format: Assume all  $U_i$  have normal distributions and that the limit state has the linear form. Let

$$Z \equiv g(\underline{U}) = A + \sum_{i=1}^k B_i U_i \quad (7.10)$$

where  $A$  and  $B_i$  are constants.

From elementary probability theory (e.g. Reference 7) the random variable  $Z$  will also have a normal distribution. The mean and standard deviation of  $Z$  are

$$\mu_Z = A + \sum_{i=1}^k B_i \mu_i \quad \sigma_Z = \sqrt{\sum_{i=1}^k B_i^2 \sigma_i^2} \quad (7.11)$$

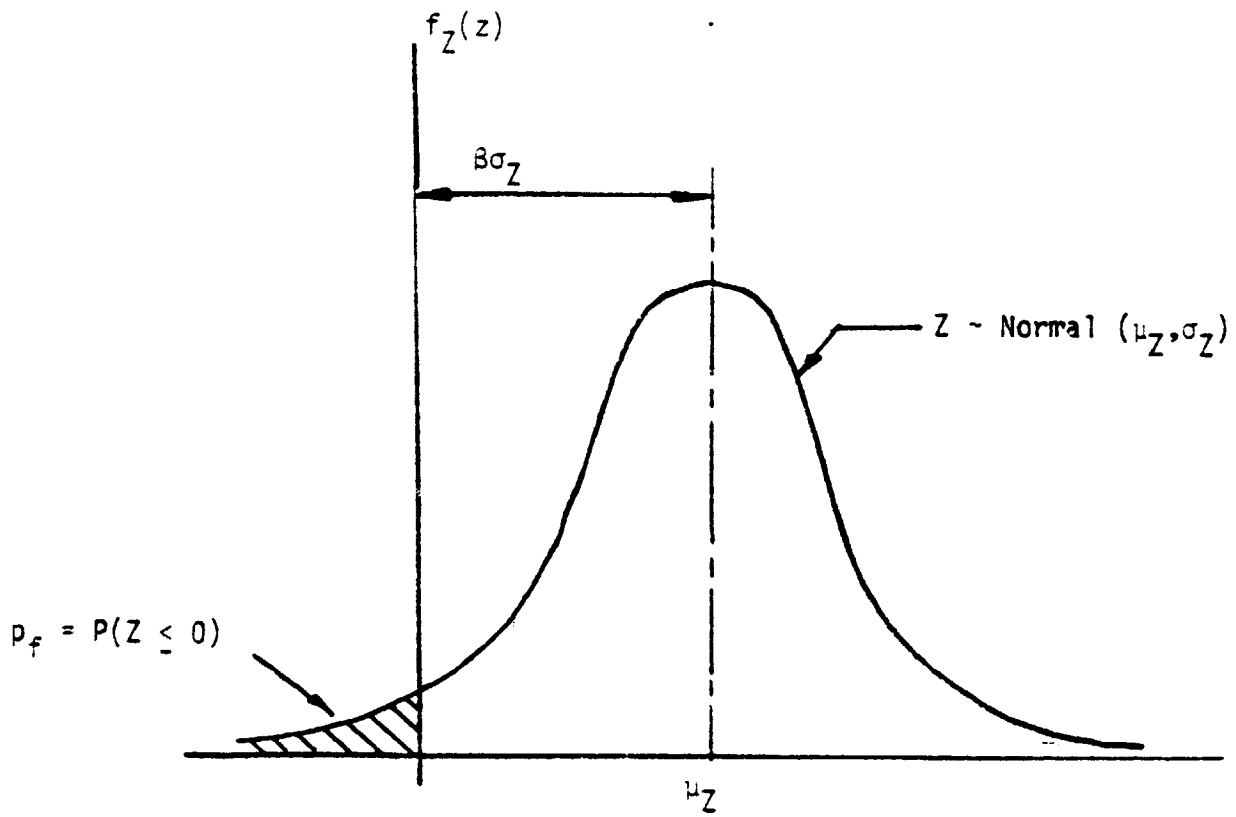
The failure condition is  $Z \leq 0$ . The probability of failure is

$$p_f = P(Z \leq 0) \quad (7.12)$$

The probability density function of  $Z$  is shown in Figure 7.3. Using standard methods of calculating probabilities for a normal variate,

$$p_f = P(Z \leq 0) = P\left\{\frac{Z - \mu_Z}{\sigma_Z} \leq \frac{\mu_Z}{\sigma_Z}\right\} \quad (7.13)$$

Figure 7.3  
The Probability Density Function for Z



$$p_f = \Phi(-\beta)$$

where  $\Phi$  is the standard normal cdf, and

$$\beta = \mu_Z / \sigma_Z \quad (7.14)$$

$\beta$  is called the safety index.  $\beta$  is related to  $p_f$  as shown in Figure 7.4. Note from both Figure 7.3 and 7.4 that the safety index  $\beta$  plays the same role in probabilistic design as does the factor of safety in conventional design, i.e., the larger is  $\beta$ , the smaller is the risk, or  $p_f$ .

**Example of the Normal Format:** Consider the tensile bar of Figure 7.1. Stress,  $S = Q/A$  where  $Q$  is a random load variable having mean  $\mu_Q$  and standard deviation  $\sigma_Q$ . The strength of the bar  $R$  is considered as a random variable having mean and standard deviation of  $\mu_R$  and  $\sigma_R$ . Compute the probability of failure.

The mean and standard deviation of  $Z$  is, (from Equation 7.11)

$$\mu_Z = \mu_R - \mu_Q/A \quad (7.15)$$

$$\sigma_Z = \sqrt{\sigma_R^2 + (\sigma_Q/A)^2} \quad (7.16)$$

Then,

$$p_f = \Phi(-\beta) \quad (7.17)$$

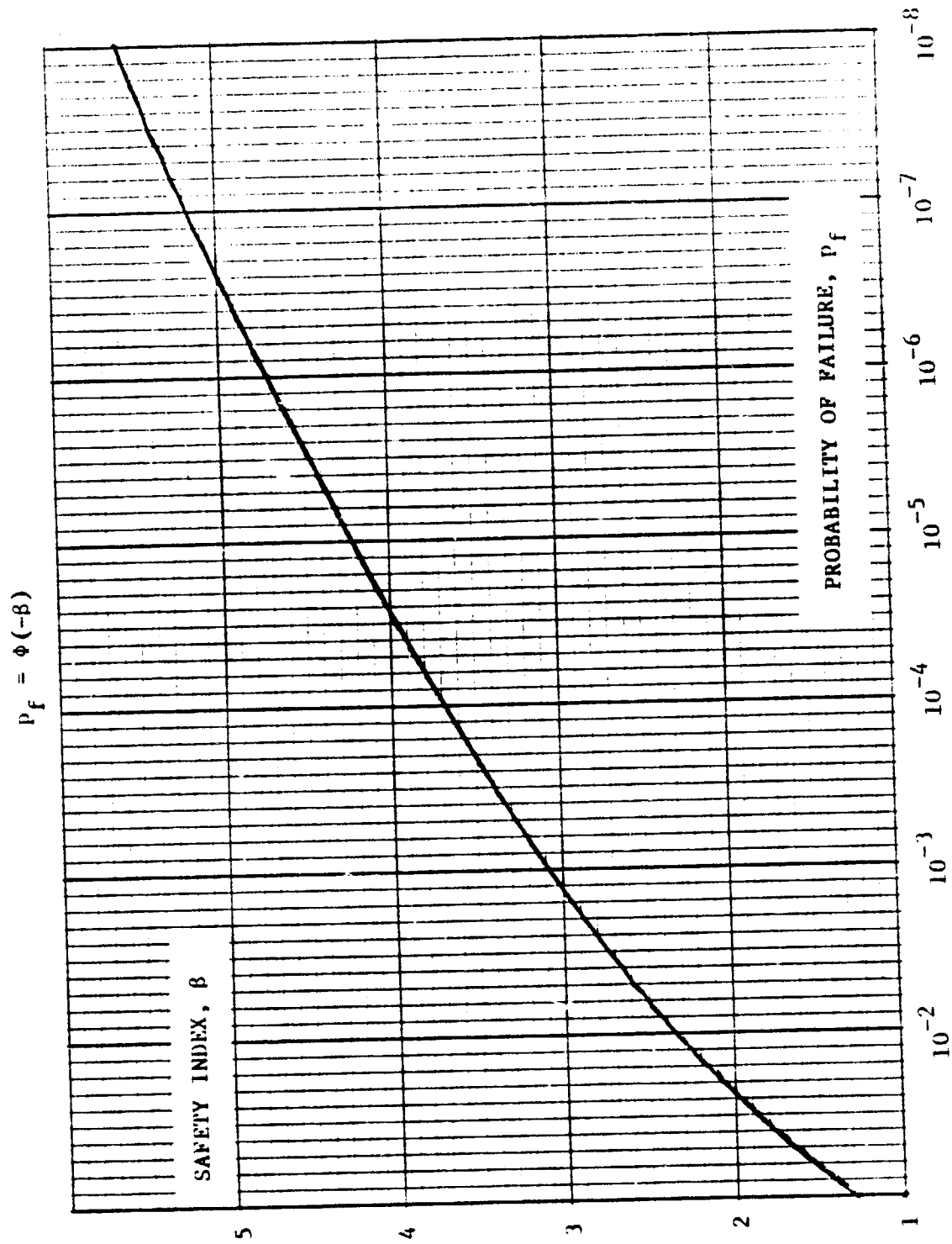
where  $\beta = \mu_Z / \sigma_Z$ .

(c) The Lognormal Format. The lognormal distribution plays an important role in probabilistic design and in particular for problems which involve fatigue. Properties of the lognormal distribution are summarized in Table 2.5 with a more complete discussion in Appendix 1.

Consider a restatement of the limit state and a redefinition of  $g$ . Let,



Figure 7.4  
 Failure Probability as a Function  
 of Safety Index,  $\beta$



$$g(\underline{U}) = R(\underline{X})/S(\underline{Y}) = 1 \quad (7.18)$$

Now the failure condition is  $g \leq 1$ . Assume that (a) all  $U_i$  have lognormal distributions, and (b)  $g(\underline{U})$  is a multiplicative function of the design factors

$$g(\underline{U}) = \prod_{i=1}^K B U_i^{a_i} \quad (7.19)$$

where  $B$  and all  $a_i$  are constants. Let  $Z = \ln g$ .

$$Z = \ln B + \sum_{i=1}^K a_i \ln U_i \quad (7.20)$$

Because  $U_i$  is lognormal, it follows that  $\ln U_i$  is normal. Then this format becomes identical to the normal format.

The failure condition is  $g \leq 1$  or, as above,  $Z \leq 0$ . The  $p_f$  is given by Equation 7.17 with  $\beta = \mu_Z/\sigma_Z$ , and

$$\mu_Z = \tilde{Z} = \ln \tilde{g} = \ln \left[ \prod_{i=1}^K B \tilde{U}_i^{a_i} \right] \quad (7.21)$$

$$\sigma_Z^2 = \ln \left[ \prod_{i=1}^K (1 + C_i^2)^{a_i^2} \right]$$

where the tilde indicates a median value and  $C_i$  is the coefficient of variation (cov) of  $U_i$ . For a random variable  $X$ , the median and cov in terms of the mean  $\mu_X$  and standard deviation  $\sigma_X$  are respectively,

$$\tilde{X} = \mu_X / \sqrt{1 + C_X^2} \quad C_X = \sigma_X / \mu_X \quad (7.22)$$

See Appendix 1 for additional detail.

The lognormal format is particularly useful in fatigue design because some of the important equations tend to be multiplicative.

Example: The limit state for fatigue under a stationary gaussian process having RMS,  $\sigma$  is (see Equation 6.29)

$$\Delta = \frac{A\sigma^m}{K} \quad (7.23)$$

where  $A = N_0(\sqrt{2})^n \Gamma(m/2 + 1)$ , a constant.  $\Delta$ ,  $K$  and  $\sigma$  are random variables. Failure occurs when the right hand side (stress) exceeds the left hand side (strength)

$$g(\underline{U}) = \frac{\Delta K}{A\sigma^m} \quad (7.24)$$

failure occurs when  $g \leq 1$ . The mean value of  $Z$  is,

$$\mu_Z = \ln(\tilde{\Delta K}/A\tilde{\sigma}^m) \quad (7.25)$$

where the tilde indicates median values. Note that for the lognormal, the mean  $\mu_X$  and the median  $\tilde{X}$  of the random variable  $X$  are related by  $\tilde{X} = \mu_X/\sqrt{1 + C_X^2}$ .

Also

$$\sigma_Z = [\ln\{(1 + C_K^2)(1 + C_\Delta^2)(1 + C_\sigma^2)^m\}]^{1/2} \quad (7.26)$$

Then  $\beta$  and  $p_f$  are determined by Equation 7.13 and 7.14.

#### 7.4 The Full Distributional Approach: Monte Carlo Methods

Consider a complicated design problem having a limit state, for example,

$$Z = g(\underline{U}) = (AX^a/Y^b) + CU^c \ln V + K = 0 \quad (7.27)$$

in which  $A$ ,  $C$ ,  $a$ ,  $b$ ,  $c$ , and  $K$  are constants and  $X$ ,  $Y$ ,  $U$  and  $V$  are random variables having a Weibull, normal, lognormal and gamma distribution respectively. Evaluation of  $p_f$  using Equation 7.5 is a practical impossibility.

Monte Carlo methods, however, can be used to obtain approximate solutions to complicated probabilistic design problems as described above. Estimates of the

distributions of complicated functions of random variables can be used to evaluate risk.

Monte Carlo methods are widely used, and only a demonstration will be presented as follows:

Example: Cycles to failure in low cycle fatigue for AISI 316 stainless steel is given as

$$N = \Delta G(\epsilon)^{-1.71} \quad (7.28)$$

where  $\Delta$ ,  $G$  and  $\epsilon$  are lognormally distributed random variables having the following values for the median and coefficient of variation.

	Median	COV
$\Delta$	1.00	0.30
$G$	0.222	0.593
$\epsilon$	0.015	0.30

The design life is given as  $N_0 = 40$  cycles. Determine the probability of failure.

The limit state is  $N = N_0$ , and the event of failure is  $(N \leq N_0)$ . Thus

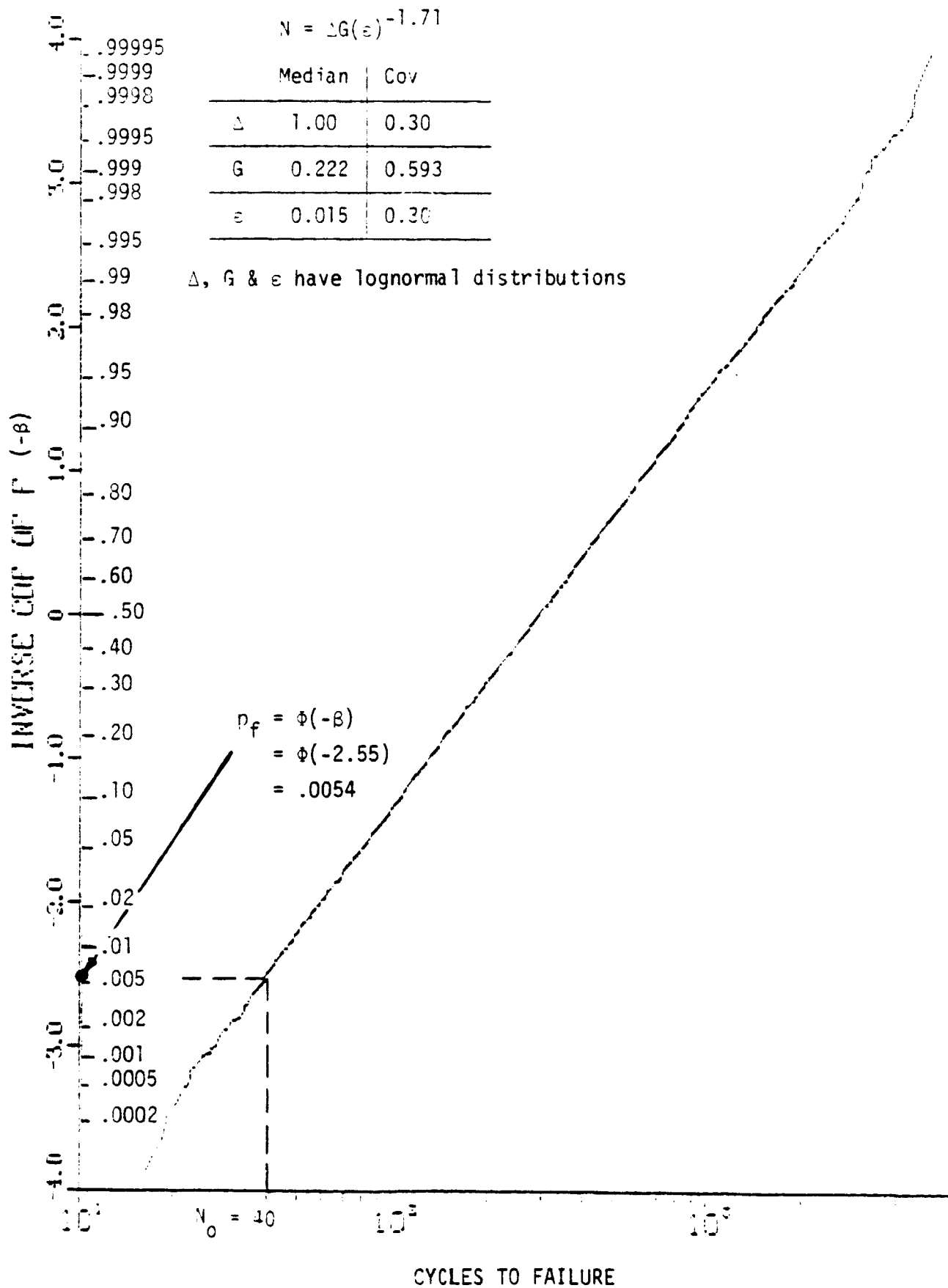
$$p_f = P(N \leq N_0) \quad (7.29)$$

An estimate of the cdf of  $N$  can be made by first sampling  $K$  values of  $\Delta$ ,  $G$  and  $\epsilon$  at random from the lognormal. These values are used to compute a sample of size  $J$  of  $N$  from Equation 7.28. Statistical analysis of  $N$  can include establishing the empirical cdf. For the above example the empirical cdf is automatically plotted as shown in Figure 7.5. The plot shows an estimate of  $p_f \approx 0.0054$ .

A closed form solution to the above problem is available (See Appendix 1 ). However, if for example  $\Delta$  had a lognormal,  $K$  had a Weibull and  $\epsilon$  had a normal, establishing the distribution of  $N$  by analytical methods would be a practical impossibility. Nevertheless, the approximate Monte Carlo solution is no more difficult than that of the above example.

Figure 7.5

The Empirical Distribution of N Plotted on Lognormal Probability Paper



## 7.5 First Order Second Moment Methods

In the full distributional approach described above, it is assumed that each design factor has a known distribution. However in practice, these probability laws are seldom known precisely because of a general scarcity of data. In some cases only the mean and standard deviation may be known with any confidence and even if the distributions of all design factors are known, it is usually impractical to evaluate Equation 7.5.

The difficulties above have motivated the development of first order second moment (FOSM) reliability methods. The random variables are characterized by their mean and standard deviation.

The basic measure of reliability is the safety index  $\beta = \mu_Z / \sigma_Z$ . From examination of Figure 7.6 it is clear that  $\beta$  as a measure of safety can be employed for any  $g(\underline{U})$  and any distribution of each  $U_i$ . However  $Z$  has a normal distribution and Equation 7.13 for  $p_f$  applies only for the linear form, Equation 7.10.

Determination of  $\mu_Z$  and  $\sigma_Z$  where  $Z = g(\underline{U})$ , a complicated function, is in general very difficult. However these terms are relatively easy to evaluate if  $g(\underline{U})$  can be linearized by a Taylor's series expansion about  $\underline{U}^*$ . Excluding higher order terms,

$$Z \approx g(\underline{U}^*) + \sum_{i=1}^n \left( \frac{\partial g}{\partial U_i} \right)_{\underline{U}^*} (U_i - U_i^*) \quad (7.30)$$

It is now a relatively simple matter to compute  $\mu_Z$  and  $\sigma_Z$ , but a key consideration is selection of the appropriate linearizing point  $\underline{U}^*$ .

*Mean Value Methods.* In earlier structural reliability studies, the point  $\underline{U}^*$  was set equal to the mean value  $\underline{\mu}$ . Assuming that all  $U_i$  are independent, it can be shown that

$$\begin{aligned} \mu_Z &\approx g(\underline{\mu}) \\ \sigma_Z &\approx \left[ \sum \left( \frac{\partial g}{\partial U_i} \right)_{\underline{\mu}}^2 \sigma_i^2 \right]^{1/2} \end{aligned} \quad (7.31)$$

Figure 7.6

Relationship with Safety Index to Risk and to  $\mu_Z$  and  $\sigma_Z$

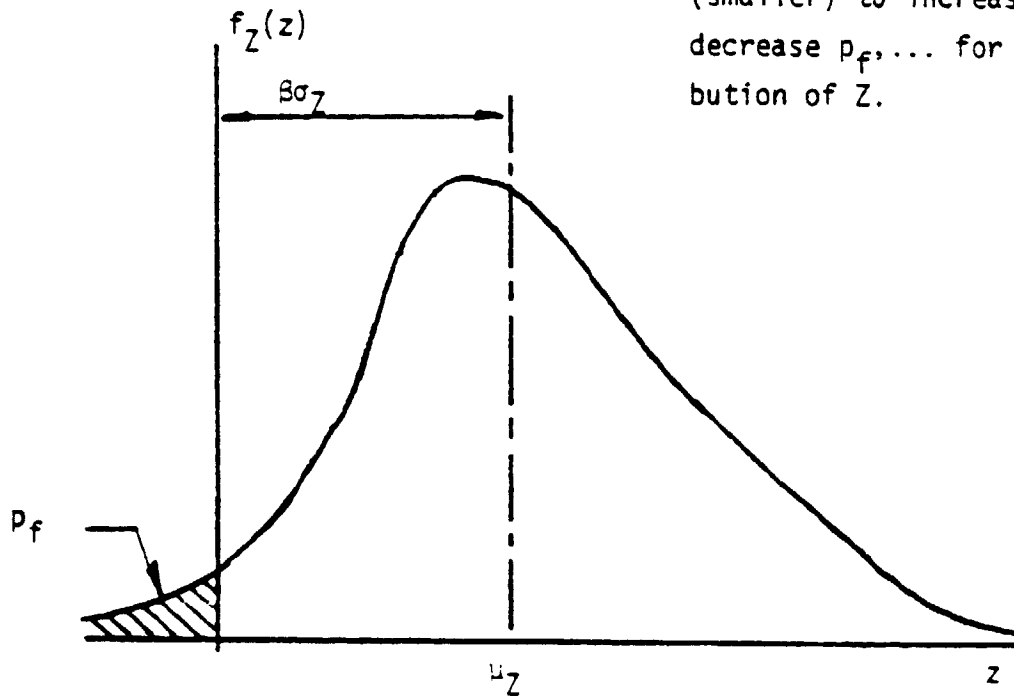
$$\text{Failure} \rightarrow Z = g(\underline{U}) \leq 0$$

Mean and Standard Deviation of Z;

$$\mu_Z = E[g(\underline{U})] \quad \sigma_Z = \sqrt{\text{Var}[g(\underline{U})]}$$

Safety Index;

$$\beta = \mu_Z / \sigma_Z$$



Changes in  $\mu_Z$  (larger) and  $\sigma_Z$  (smaller) to increase  $\beta$ , will decrease  $p_f$ , ... for any distribution of Z.

where  $\sigma_i$  = the standard deviation of  $U_i$ .

Example: Consider the tension bar of Figure 7.7. Compute the safety index.

$$Z = g(U) = g(R,A) = R - Q/A \quad (7.32)$$

Note that the function  $g$  is formulated so that the limit state is  $Z = 0$  and the failure condition is  $Z < 0$ .

$$\mu_Z = g(\mu_R, \mu_A) = \mu_R - Q/\mu_A \quad (7.33)$$

$$\sigma_Z^2 = \left(\frac{\partial g}{\partial R}\right)^2 \sigma_R^2 + \left(\frac{\partial g}{\partial A}\right)^2 \sigma_A^2$$

with the partials evaluated at the means.

$$\sigma_Z^2 = \sigma_R^2 + (Q/\mu_A^2)^2 \sigma_A^2 \quad (7.34)$$

Substituting the values from Figure 7.7  $\mu_Z = 26.3$  and  $\sigma_Z = 6.45$ . Thus

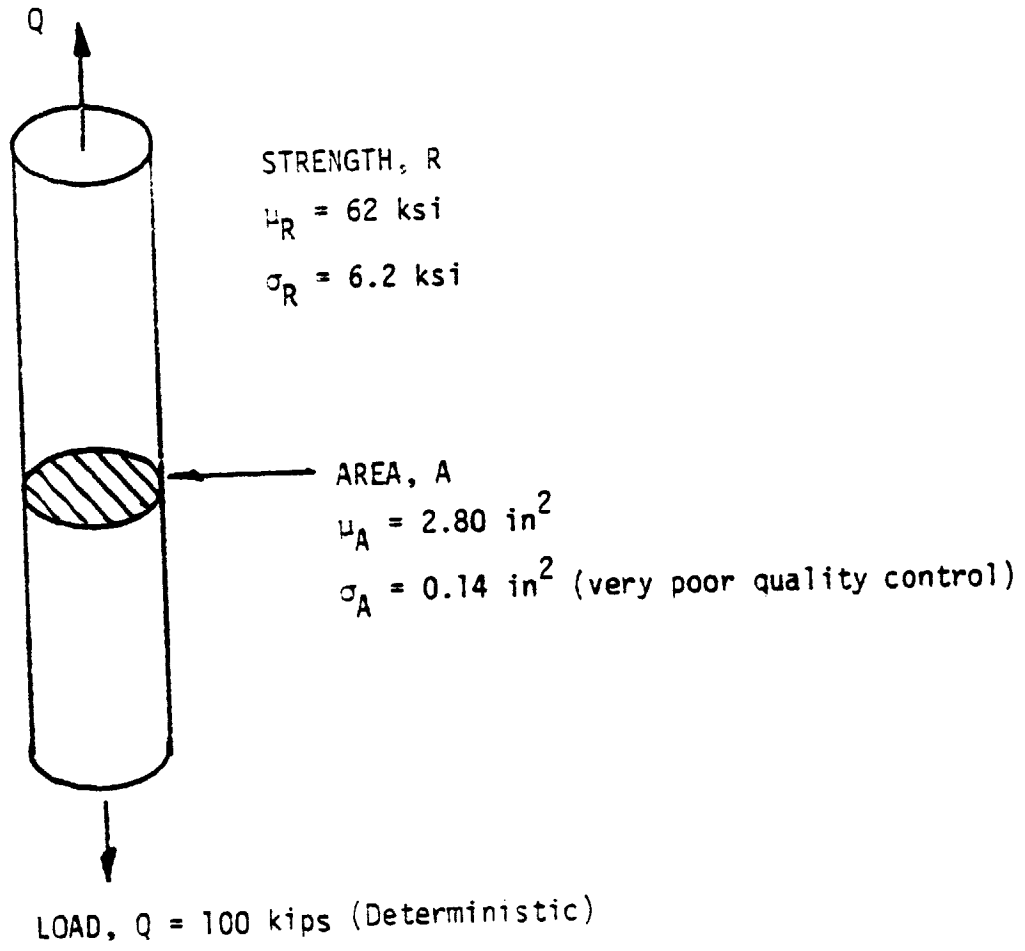
$$\beta = \mu_Z/\sigma_Z = 4.07 \quad (7.35)$$

Evaluation of  $\beta$  for this case was relatively easy. Even for more complicated limit state functions the operations are not difficult. On the other hand there are significant limitations and criticism of mean value first order second moment methods which have focused on the following:

- (a) In this approach, the measure of safety is  $\beta$ , which in some cases gives weak information on the probability of failure,
- (b) Information on distributions, if available, cannot be included in a logical way,
- (c) The linear approximation of  $g(U)$  at the mean of the various variables



Figure 7.7  
DATA FOR EXAMPLE PROBLEM



Limit State:

Formulation 1

$$Z = R - Q/A = 0$$

Formulation 2

$$Z = RA - Q = 0$$

appears to be too inaccurate in the face of severe non-linearities of this function in many design situations, i.e., the higher order terms in the Taylor's series expansion are important.

- (d) Equation 7.31 for  $\mu_Z$  and  $\sigma_Z$  is valid only for relatively small variances (typically for  $C_i < 0.15$ ). In the fatigue problem most design factors have "large" variances, ... again the higher order terms are important
- (e) and the most important, the method fails to be invariant in that  $\beta$  depends upon the mechanical formulation of the problem.

The lack of invariance problem can be illustrated using the above example.

Note that an equally valid statement of the failure condition is

$$Z = RA - Q < 0 \quad (7.36)$$

Using procedures described above,

$$\mu_Z = \mu_R \mu_A - Q \quad (7.37)$$

$$\sigma_Z^2 = \mu_A^2 \sigma_R^2 + \mu_R^2 \sigma_A^2$$

Substituting values from Figure 7.7,  $\mu_Z = 73.6$  and  $\sigma_Z = 19.4$  and  $\beta = 3.79$ .

Different values of  $\beta$  are obtained from two equally valid statements of the failure condition. It is principally this problem with the mean value method that led to the development of the generalized safety index (1).

*Advanced First Order Second Moment Methods.* It is now recognized that  $\beta$  will be invariant to formulation if the linearization point  $\underline{U}^*$  is chosen as a point on the failure surface. The following discussion describes how to compute the generalized safety index and to establish the appropriate  $\underline{U}^*$ .

Hasofer and Lind have proposed a general approach which is exact and invariant to mechanical formulation (1). As an example, consider R and S as the basic

variables. In Figure 7.8 the probability of failure  $p_f$  will be the volume under the joint density function in region  $\Omega$ . Perform a transformation; define the reduced variables:

$$r = (R - \mu_R)/\sigma_R \quad S = (S - \mu_S)/\sigma_S \quad (7.38)$$

$r$  and  $s$  have means equal to zero and standard deviations equal to one.

Figure 7.9 shows the space of reduced variables. Note that again  $p_f$  is the volume under the joint density function in  $\Omega$ . Thus the distance  $\Delta \equiv OP$  is a measure of reliability. As the failure surface moves to the right,  $p_f$  gets smaller and  $\Delta$  gets larger. Therefore  $\Delta$  is a measure of reliability. Furthermore from geometric considerations

$$\Delta = (\mu_R - \mu_S) / \sqrt{\sigma_R^2 + \sigma_S^2} \quad (7.39)$$

which is exactly the same as  $\beta$  obtained from Equation 7.31 for  $g = R - S$ .

The design point is defined as the point on the failure surface closest to the origin of the reduced coordinates. In the above example it would be the point  $(s^*, r^*)$  as shown in Figure 7.9. The design point in the basic variables is then determined by inverting the original expression, e.g.,  $R^* = r^* \sigma_R + \mu_R$  from Equation 7.38.

Define the generalized safety index,  $\beta$  as the minimum distance from the origin of the reduced variables to the failure surface

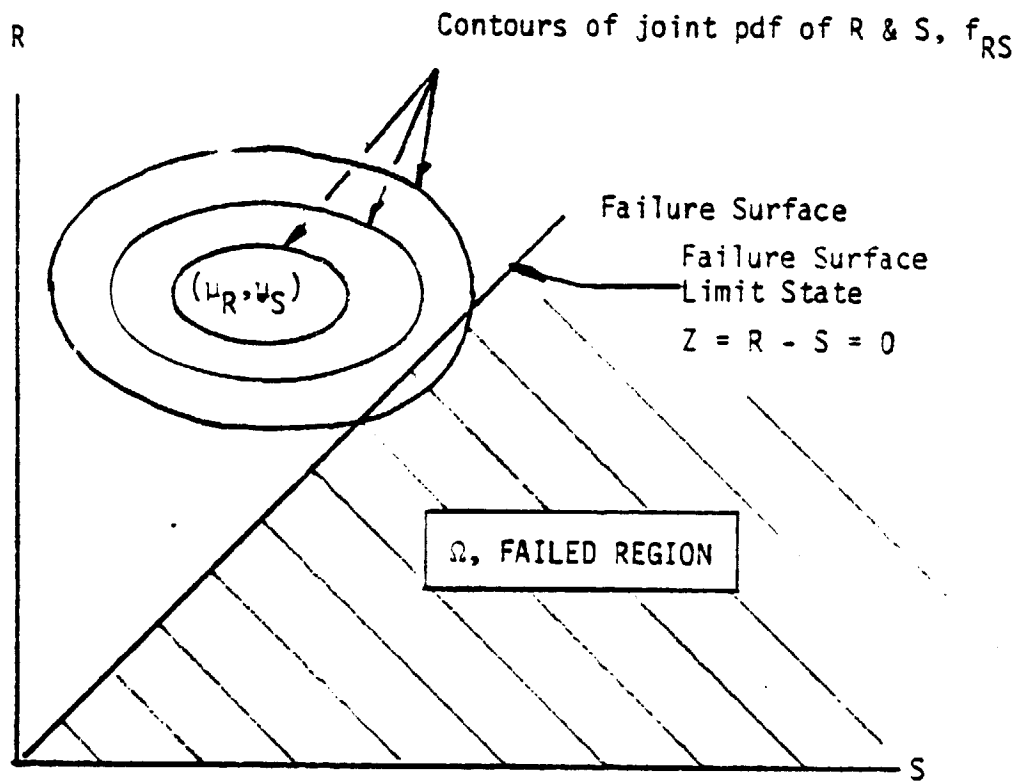
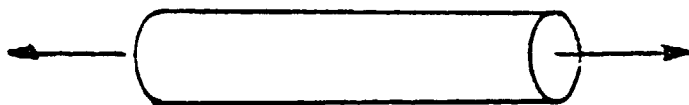
$$\beta = \min \{ \sqrt{u_1^2 + u_2^2 + \dots + u_n^2} \} \quad (7.40)$$

Subject to  $g(U) = 0$ , where the reduced variables in terms of  $U_i$  are

$$u_i = (U_i - \mu_i) / \sigma_i \quad (7.41)$$

Figure 7.8

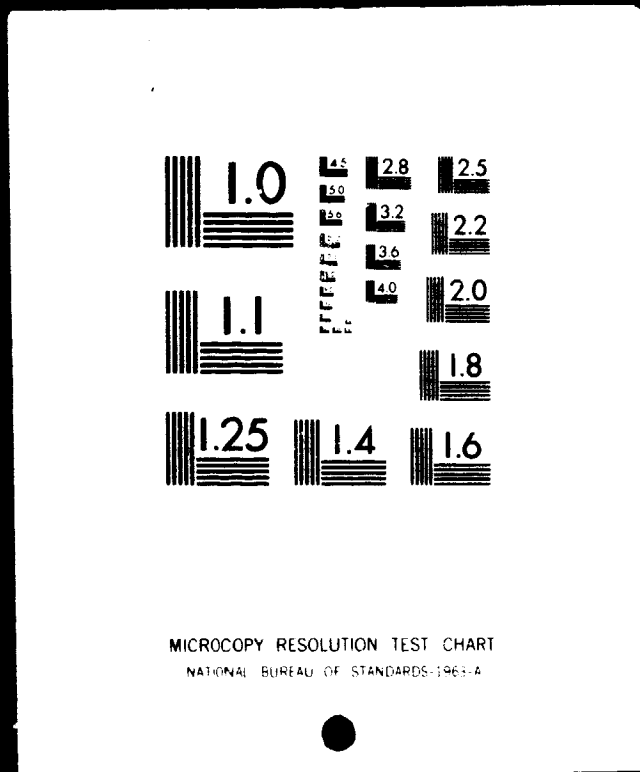
Parameter Space When R and S are the Basic Variables and the Limit State is  $Z = R - S = 0$



# 3 OF 3

N82-14531

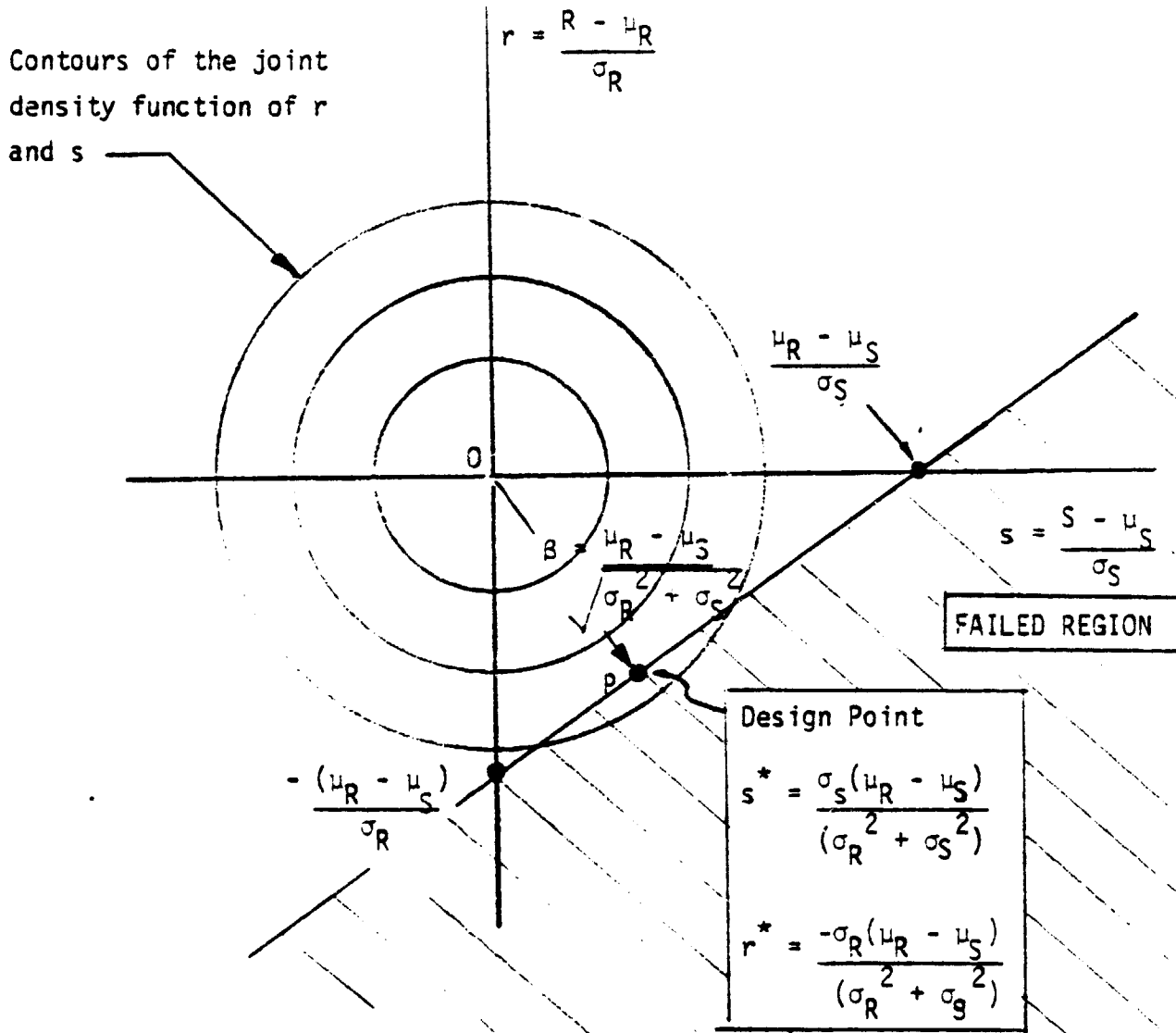
UNCLAS



ORIGINAL PAGE IS  
OF POOR QUALITY

Figure 7.8

Parameter Space when R and S are the Basic Variables and  
the Limit State is  $Z = R - S = 0$



C-3

Such a procedure will produce a design algorithm which is invariant to formulation because the failure surface is the same regardless of how it was written. Figure 7.10 illustrates the failure surface in two dimensions.  $U_i$  are the reduced variables.

Another viewpoint of this problem is to consider the system of equations (6),

$$\alpha_i = \frac{\partial g_1 / \partial u_i}{[\sum (\partial g_1 / \partial u_i)^2]^{1/2}} \quad (7.42)$$

$$u_i = -\alpha_i \beta$$

$$g_1(u_i^*) = 0$$

where  $\beta$  is the distance to point  $\underline{u}^*$  on the failure surface  $g_1$  in reduced coordinates.  $\alpha_i$  is the direction cosine of the  $i^{\text{th}}$  coordinate relative to  $\underline{u}^*$ . These systems of equations are solved for the  $\alpha_i$ 's and  $\underline{u}^*$  which minimizes  $\beta$ .

It is also possible to redefine any or all of the basic variables by making a log transformation  $U'_i = \ln U_i$ . Then, the mean and standard deviation of  $U'$  are (assuming the lognormal distribution moments)

$$\mu_{U'} = \ln \mu_U - 1/2 \ln (1 + C_U^2) = \ln \tilde{U} \quad (7.43)$$

$$\sigma_{U'} = \sqrt{\ln (1 + C_U^2)}$$

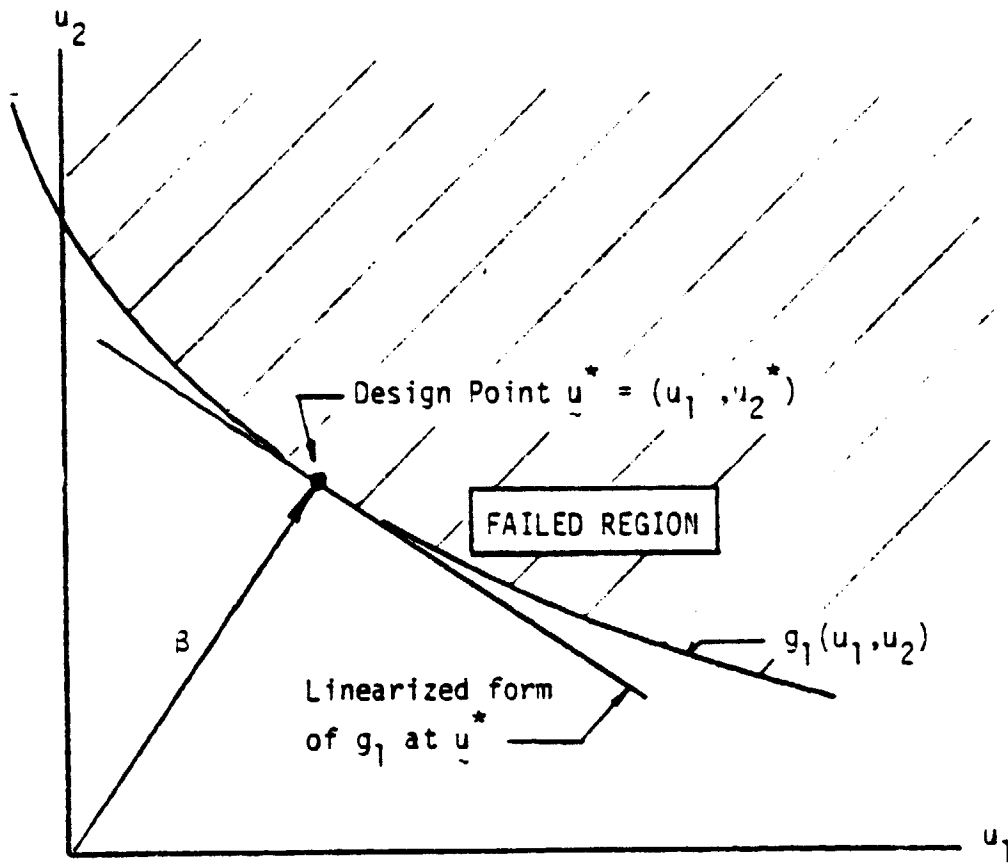
For convenience, the subscript "i" has been dropped. The reduced variables are

$$u = (U' - \mu_{U'}) / \sigma_{U'} \quad (7.44)$$

so that the original variables in terms of the reduced variables are

Figure 7.10

Generalized Safety Index in Two Dimensions





$$U = \exp(\mu_U + u\sigma_U) \quad (7.45)$$

In general the log transformation would be performed on large variance design factors, and in particular would be useful in the fatigue problem.

To compute  $\beta$  one must in general perform an optimization calculation. The author's experience is that most failure surfaces are well behaved and as a result, convergence to an accurate (four place) result using a relatively crude optimization algorithm, requires a negligible amount of computer time, even for several design variables.

Example: Consider the problem of Figure 7.7. The two valid expressions for the failure surface are

$$Z = R - Q/A = 0 \quad (7.46)$$

$$= RA - Q = 0$$

The reduced coordinates are

$$r = (R - \mu_R)/\sigma_R \quad a = (A - \mu_A)/\sigma_A \quad (7.47)$$

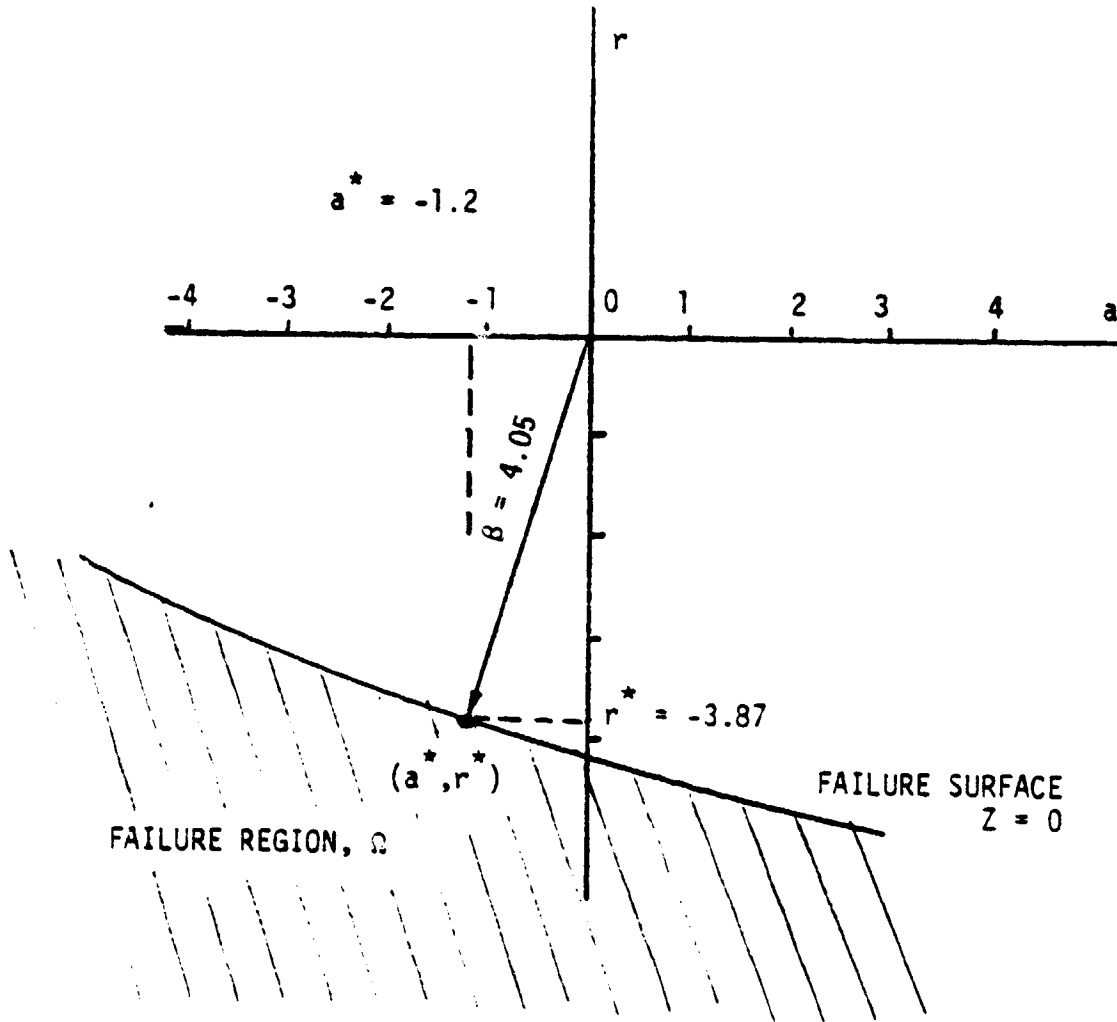
Using the values in Figure 7.7 and substituting for R and A into either of the expressions for Z, the same expression for the limit state results

$$g_1(r,s) = (62 + 6.2r)(2.8 + .14a) - 100 = 0 \quad (7.48)$$

This function is plotted in Figure 7.11. Also shown is  $\beta$ , the minimum distance to the failure surface as well as the design point  $(a^*, r^*)$ . Here the values were determined graphically.

Note that from Figure 7.10, the procedure for determination of  $\beta$  is equivalent to linearizing the limit state equation at  $u^*$  and computing  $\beta$  associated with

Figure 7.11  
GRAPHICAL DETERMINATION OF THE GENERALIZED SAFETY  
INDEX FOR THE PROBLEM OF FIGURE 7.7.



the linearized rather than the original limit state. From Equation 7.30, it can be shown that

$$\mu_Z = E(Z) = \sum_{i=1}^n \left( \frac{\partial Z}{\partial U_i} \right)_{U^*} (\mu_i - U_i^*) \quad (7.49)$$

$$\sigma_Z^2 = E[(Z - \mu_Z)^2] = \sum \left( \frac{\partial Z}{\partial U_i} \right)_{U^*}^2 \sigma_i^2 \quad (7.50)$$

In the above example of Figure 7.11  $R^* = 38.0$ ,  $A^* = 2.632$  and  $\beta = 4.05$  in both formulations.

In summary, the generalized safety index  $\beta$  provides a measure of reliability for which is invariant to the mechanical formation of the limit state. It is necessary to specify only the mean and standard deviation of each design variable. The  $\beta$  so defined contains in general, weak probability information. However, if the variables are normal or lognormal (and the log transformation is used), then  $p_f = \Phi(-\beta)$  is a good approximation to the exact probability of failure as would be computed by Equation 7.5.

## 7.6 Extended Form of the Generalized Safety Index

The procedure described in the previous section gives values of  $\beta$  which may be related to the probability of failure only in those special cases where design factors have normal or lognormal distributions. But many design problems involve random variables having other distributions e.g. Weibull. Computation of the probability of failure in the general case of a non-linear limit state and design variables having arbitrary distributions can be accomplished by a very complicated numerical integration (Equation 7.5) or by Monte Carlo methods. Neither method is practical for general design application.

In 1976 Rackwitz has proposed an extension of the generalized safety index concept (6,7,7,12) in which appropriate distributional information could be incorporated. The basic idea is to transform the non-normal variables into

equivalent normal variables. This transformation may be accomplished by approximating the true distribution of variable  $U_i$  by a normal distribution at the value  $U_i^*$  corresponding to a point on the failure surface. The justification for this is that if the normalization takes place at the point close to that where failure is most likely, (i.e., minimum  $\beta$ ), the estimates of the failure probability obtained by the approximate procedure should approximate the true (but unknown) failure probability quite closely (6).

The mean and standard deviation of the equivalent normal variable are determined such that at the value  $U_i^*$ , the cumulative probability and probability density of the actual and approximating normal variable are equal. Thus,

$$\sigma_i^N = \frac{\phi(\Phi^{-1}[F_i(U_i^*)])}{f_i(U_i^*)} \quad (2.24a)$$

$$\mu_i^N = U_i^* - \Phi^{-1}[F_i(U_i^*)]\sigma_i^N \quad (2.24b)$$

in which  $F_i$  and  $f_i$  = non-normal distribution and density function of  $U_i$  and  $\phi(\cdot)$  is the density function for the standard normal variate. Having determined  $\mu_i^N$  and  $\sigma_i^N$  of the equivalent normal distributions, the solution proceeds as described in Equations 7.40 to 7.42. Inasmuch as the checking point variable  $U_i^*$  changes with each iteration, the parameters  $\mu_i^N$  and  $\sigma_i^N$  must be recomputed during each iteration cycle also. However, since all calculations are performed by computer, this does not materially add to the complexity of the reliability analysis described earlier.

It has been observed that this approximate technique often yields excellent agreement with the exact solution, e.g. as obtained by Equation 7.5 (6.12). However, it has been noted (15) that the checking point may not correspond exactly to the point where the joint probability density is maximum and failure is most likely. However, this procedure does not reduce the error which is due

to the linearization of what may be a generally nonlinear failure boundary at the design point. Unless the failure boundary is highly nonlinear, however, as is the case in some stability problems, this source of error is small compared to the accuracy with which most of the parameters in engineering reliability analysis can be estimated.

A computer code using the Rackwitz algorithm has been written as an activity of this study (called SUPER-HASL). An example which illustrates the efficiency of this algorithm is presented in Section 8.4. In this demonstration, the probability of failure is computed at roughly 1/6 of the cost of Monte Carlo.

SUPER-HASL has only recently been developed. The efficiency and accuracy of the Rackwitz algorithm will be investigated in future studies.

### 7.7 Summary Comments on Target Reliability Levels and Technique

Important questions have not been addressed in the previous discussion. Following are answers to some key questions regarding implementation of reliability methods.

*How Much Confidence Can We Put Into  $p_f$ ?*

A computed value of probability of failure  $p_f$  should be interpreted as an "index" of risk rather than an actuarial prediction.  $p_f$  can be sensitive to the many assumptions used that one may have little confidence in the quality of the point estimate of  $p_f$ . An estimate of  $p_f$  is sometimes called the "notional probability of failure". In practice,  $\beta$  is now more commonly used (See Ref 4).

*What Should be a Target Value of  $\beta$  (or  $p_f$ )?*

The target design value of  $\beta$ , denoted as  $\beta_0$ , depends entirely on the application. Values of  $\beta$  implied in current codes and procedures typically range from 2.0 to 6.0. The value of  $\beta_0 = 3.0$  has been suggested for general construction (4).

To establish  $\beta_0$  the engineer can (a) attempt to evaluate  $\beta$  inherent in current design procedures as a reference, and (b) use an implied relationship between  $\beta$  and  $p_f$  assuming that  $p_f$  accurately describes risk. Ultimately the designer has to use his judgement and experience in making a decision regarding acceptable levels.

*Can We Provide a General Reliability Method That Can be Used For All Cases?*

It is impossible to recommend a general procedure which can be used in all cases because each class of problem seems to have its own unique characteristics. A reliability method must be tailored to the situation. However some general observations can be made from the above discussion. A summary of procedures available for each problem type is provided in Table 7.2.

#### 7.8 References for Chapter 7

1. Freudenthal, A.M, Garrelts, J.M. and Shinozuka, M., "The Analysis of Structural Safety", Journal of the Structural Division, ASCE, Vol. 92, No. ST1, Feb. 1966.
2. Hasofer, A.M. and Lind, N.C., "Exact and Invariant Second Moment Code Format," J. Engr. Mech. Div., ASCE, Bef. 1974.
3. Ditlevsen, O. and Skov, K., "Uncertainty Theoretical Definition of Structural Safety Suggested by the NKB - Safety Committee," DIALOG 2-76, Mexico City, Jan. 1979.
4. Ravindra, M.K. and Galambos, T.V., "Load and Resistance Factor Design for Steel," Journal of the Structural Division, ASCE, Vol. 104, No. ST9, Sept. 1978.
5. Turksova, C. and Grigoriu, M., "Generalized Second Moment Reliability Analysis", ASME, No. 78-PVP-105, 1978.
6. Ellingwood, B., Galambos, T.V., MacGregor, J.G., Cornell, C.A., "Development of a Probability Based Load Criterion for American National Standard A58", NBS SP 577, June 1980.
7. "First Order Reliability Concepts for Design Codes", Bulletin D'Information N° 112, Comite Europeen du Beton, July 1976.
8. Rationalization of Safety and Serviceability Factors in Structural Codes. Report of CIRIA, Construction Industry Research and Information Center, 6 Storey Gate, London SW 1P 3 AU.
9. Meyer, P.L., Introductory Probability and Statistical Applications, Addison-Wesley, 1970.

10. Hines, W.W., and Montgomery, D.C., Probability and Statistics in Engineering and Management Science, 2nd Ed., Wiley, 1980.
11. Bowker, A.H., and Lieberman, G.J., Engineering Statistics, Prentice-Hall, 1972.
12. Rackwitz, R., Fiessler, B., "Notes on Discrete Safety Checking When Using Non-Normal Stochastic Models for Basic Variables", Loads Project Working Session, MIT, Cambridge, June 1976.

Table 7.2  
A SUMMARY OF METHODS OF PROBABILISTIC DESIGN

Form of the Failure Function $Z = g(U)$	Reliability Method of Preference
A stress and strength variable have known distributions Linear (Equation 7.10); all design factors normal	$p_f$ can be computed by numerical integration (Equation 7.5) $Z$ is exactly normal and $p_f$ can be computed exactly (Equation 7.13 and 7.14)
Multiplicative (Equation 7.19); all design factors lognormal	$Z = \ln g$ is exactly normal and $p_f$ can be computed exactly (Equation 7.13 and 7.14)
$g$ is any function; all design factors have known distributions	Monte Carlo method; a relatively easy method for estimating $p$ , but unweildly for a designer; the Rackwitz algorithm may be the best
$g$ is any function; all design factors have known means and variances but the variance is small ( $COV < 0.15$ )	The mean value first order second moment method provides an easy approximate method to evaluate $\beta$ .
$g$ is any function; all design factors have known means and variances; if variances are large, use a log transformation on each factor	The generalized safety index can be computed using a digital computer code (Ref 2,6).
$g$ is any function; some design factors have known distributions but others have only means and variances known.	The generalized safety index can be computed. A digital computer program is required. The method of analysis is not described herein, but is provided in Ref. 6.



## Chapter 8 DEMONSTRATION OF RELIABILITY ANALYSIS TO FATIGUE PROBLEMS

### 8.1 Preliminary Remarks

Each design and safety check problem seems to have its own unique features. Therefore it would be difficult to propose a general reliability format which could be used in all problems.

Three examples, presented in this chapter, illustrate how some of the reliability methods discussed in Chapter 7 can be employed in fatigue problems. The examples treat models used to predict high temperature low cycle fatigue. The first uses strainrange partitioning to define the limit state and the second and third use the local strain approach.

These examples do not exhaust available reliability methods but provide only a sample of the possibilities.

### 8.2 Example 1; Strainrange Partitioning in a Reliability Format.

Consider a component for which high temperature fatigue analysis is performed using strainrange partitioning. Only PP and PC strains are present. The limit state is

$$\Delta = N_0 \left[ \frac{f_{PP}}{G(\Delta\varepsilon)^Y} + \frac{f_{PC}}{H(\Delta\varepsilon)^n} \right] \quad (8.1)$$

See Section 6.5 and 6.9 and Reference 1 for definition of terms. The data for all parameters are given in Table 8.1.

The right hand side of Equation 8.1 is damage at the design life  $N_0$ . Failure occurs if this exceeds  $\Delta$ , the damage at failure. Thus the left hand side can be considered as "strength" and the right hand side as "stress".

The random variables  $\Delta$ ,  $G$ ,  $H$ , and  $\Delta\varepsilon$  are related in such a way that a closed form solution for probability of failure, e.g. using Equation 7.5, would be extremely difficult. In lieu of this, two approximate techniques for reliability analysis are pursued.

Table 8.1  
Data for Example 1

$N_0$	500 cycles, design life
$f_{pp}$	0.70
$f_{PC}$	0.30
$\gamma$	-1.711
$\eta$	-1.188

Random Variables (all assumed to be lognormally distributed)

	Median	Coefficient of Variation
$\Delta$	1.0	0.30
G	.222	0.406
H	1.673	0.393
$\Delta\epsilon$	0.02	0.300

1. The generalized safety index can be computed. A log transformation is performed on all four variables. A digital computer code (HASL) is employed\*. The output is presented in Table 8.2. The computed safety index is  $\beta = 4.08$ . The generalized safety index provides a good approximation to the probability of failure using  $p_f = \Phi(-\beta)$  when the design factors are normal or log-normal, even though the limit state is not linear in the design variables\*\*. Here  $p_f = 2.28 \times 10^{-5}$ . (See Figure 7.4)
2. The Monte Carlo method can be used to estimate the probability of failure. Equation 8.1 can be rewritten as

$$N = \Delta / \left[ \frac{f_{PP}}{G(\Delta\varepsilon)^Y} + \frac{f_{PC}}{H(\Delta\varepsilon)^\eta} \right] = N_0 \quad (8.2)$$

$N$  is a random variable denoting cycles to failure. Failure occurs if  $N \leq N_0$ . The Monte Carlo method (Section 7.3) is used to obtain a random sample of  $N$ , the distribution function of which is plotted on lognormal probability paper in Figure 8.1.

Graphically extrapolating the empirical distribution function to  $N_0 = 500$  cycles, the inverse cdf, equal to  $-\beta$ , indicates that  $\beta \approx 3.9$ . The corresponding  $p_f = \Phi(-3.9) = 4.85 \times 10^{-5}$ .

Considering the degree of extrapolation required for the Monte Carlo method, it is not surprising that there is only fair agreement between the two methods.

\* See Section 7.4 for the description of the generalized safety index.

\*\* Just how good this approximation is, in general, is presently under study.

Table 8.2

Output of HASL<sup>(a)</sup> --A Program for Computing the Generalized Safety Index--for the Example of Table 8.1.

DESIGN VARIABLES

VARIABLE	TRANSFORMATION (a)	MEAN/MEDIAN (b)	STD DEV/COV (c)
DELTA	1	.10000E+01	.30000E+00
G	1	.22200E+00	.40600E+00
EPSLN	1	.20000E-02	.30000E+00
H	1	.16730E+01	.39300E+00

DESIGN POINT

VARIABLE	REDUCED VALUE	BASIC VALUE
DELTA	XR(1) = -2.02318	X(1) = .55215E+00
G	XR(2) = -1.49219	X(2) = .12394E+00
EPSLN	XR(3) = 2.99219	X(3) = .48140E-02
H	XR(4) = -1.16406	X(4) = .10754E+01

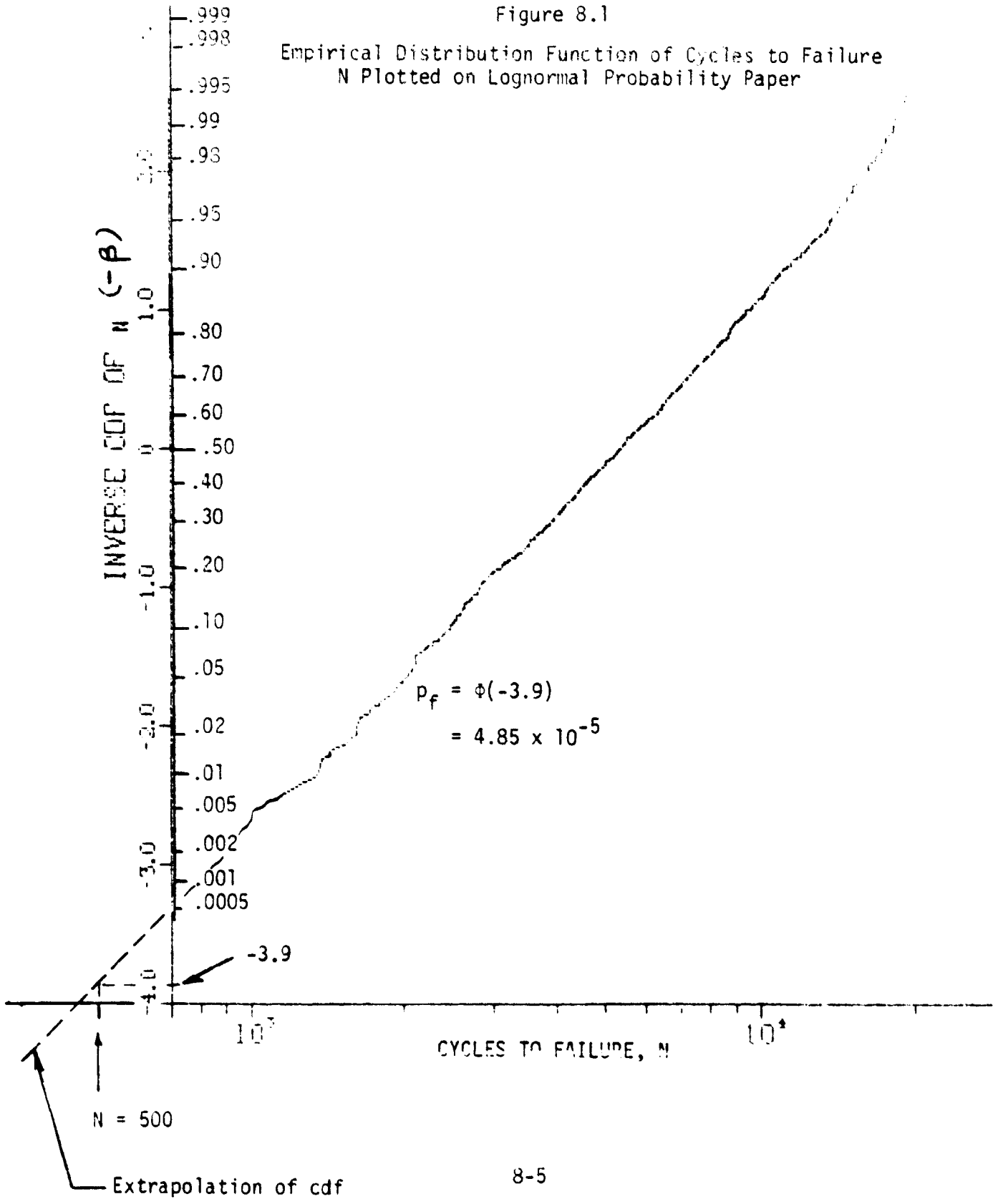
SAFETY INDEX, BETA = 4.07832

NOTES:

- (a) HASL is a digital computer program developed by P. H. Wirsching, University of Arizona, for computing the generalized safety index
- (b) A "1" indicates a log transformation on each variable
- (c) With a log transformation, the listed values are the Median and the COV

Figure 8.1

Empirical Distribution Function of Cycles to Failure  
N Plotted on Lognormal Probability Paper



### 8.3 Example 2: A Local Strain Analysis Problem in a Reliability Format

A fatigue life estimate is required for a turbine disk. It is assumed that the fatigue crack will initiate at a notch in the rim provided for a blade as indicated in Figure 8.2. Stress and thermal analyses based on the definition of a duty cycle (one flight) will provide the normal stress  $S$  as well as the equivalent stress  $\sigma_e = K_t S$  and the temperature at the notch as a function of time. A typical stress and temperature history shown in Figure 8.3 shows that the peak stress will not, in general, occur at the same time as the peak temperature.

Generally speaking, in the event of cyclic plasticity at a notch, the local strain approach to fatigue life prediction (1, 2, 3) would be the method of preference. The most serious limitation to such an approach in dealing with high temperature is that the strain life relationship must be based on a given constant temperature.

Application of the local strain approach to high temperature fatigue problems has considerable appeal. This approach to fatigue life prediction is now commonly used by mechanical designers in many fields. Analysis methods are well developed, including life prediction of components subjected to random loads and having notches which experience cyclic plasticity. While this method cannot explicitly treat the synergistic effects of creep and fatigue, life estimates may be reasonable for those components where creep effects may be small in comparison to fatigue. This may be the case for example for some turbine hot section components where cycle lives of the order of  $10^4$  are anticipated.

Following will be a demonstration of how reliability analysis methods, previously described, can be applied. First, it is necessary to identify sources of uncertainty in the life prediction process.

#### Environment

1. The use of the "duty cycle" to represent operational conditions.
2. The temperature of the hot gases.
3. The use of an equivalent temperature.

#### Material Behavior

1. Effect of strain hardening (or softening).
2. Accuracy in the cyclic stress strain curve used.

Figure 8.2

Notch in Rim of Turbine Disk Where Fatigue Crack  
is Assumed to Initiate

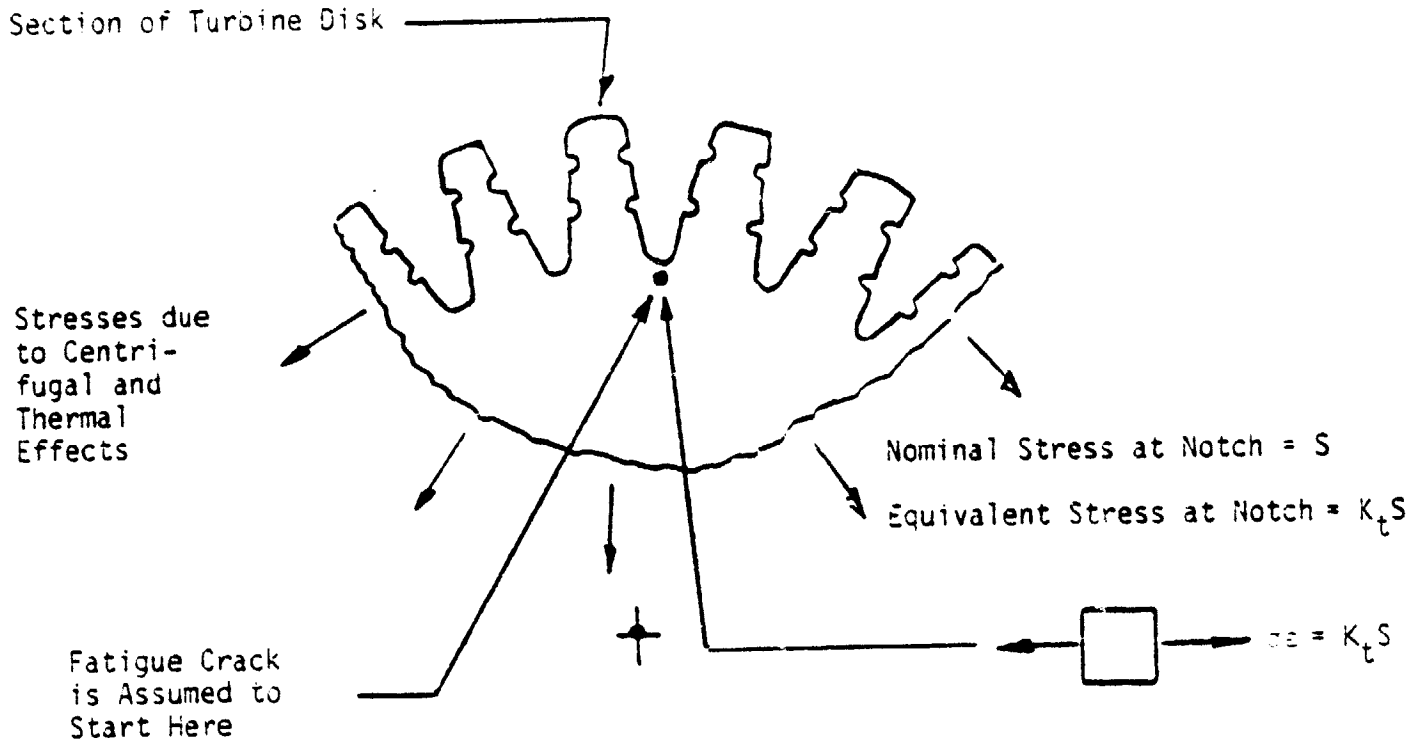
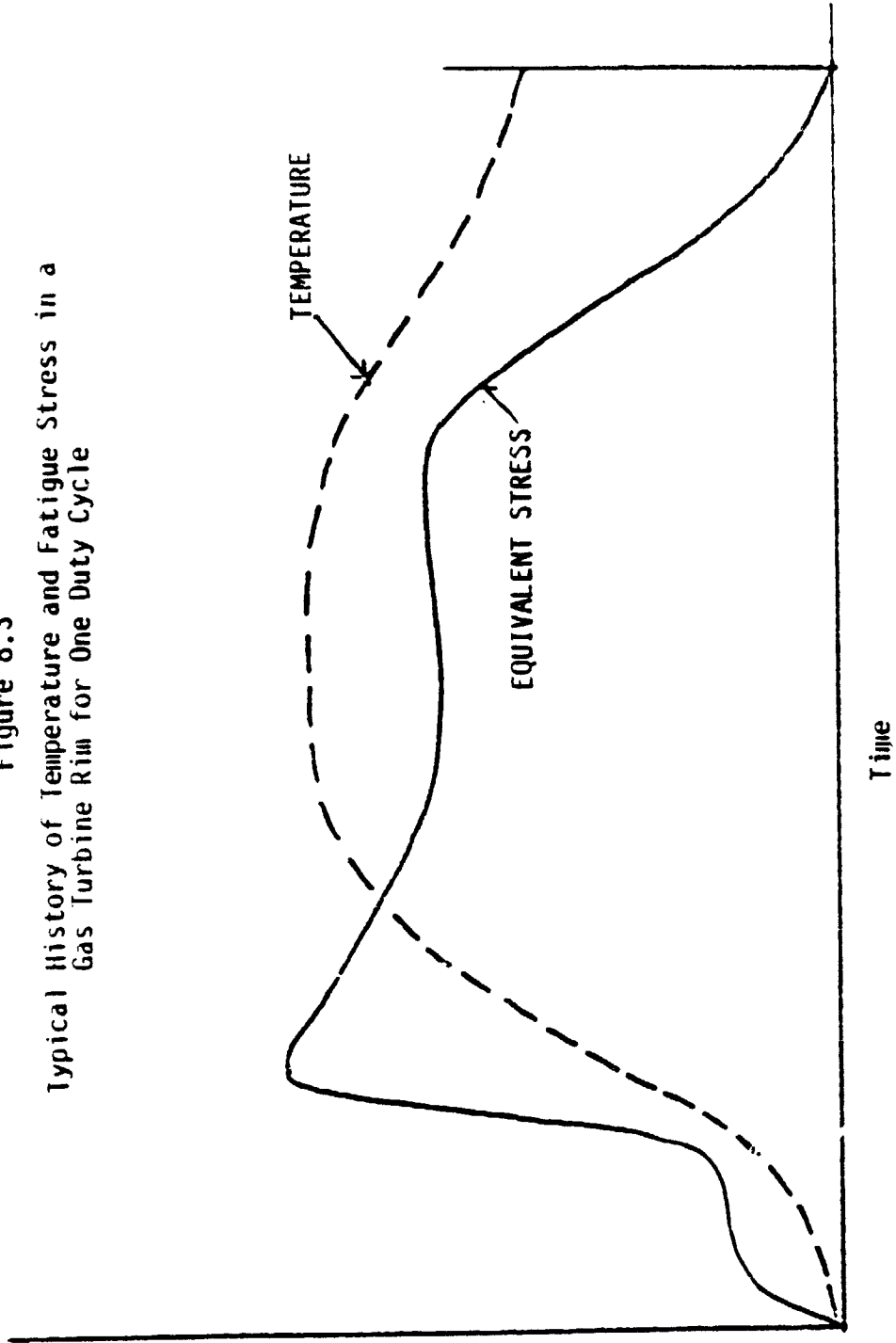


Figure 8.3  
Typical History of Temperature and Fatigue Stress in a  
Gas Turbine Rim for One Duty Cycle





3. The use of Neuber's rule to obtain the nominal stress-notch strain curve.
4. Basic fatigue behavior as evidence by statistical scatter in strain-life data.
5. Effect of surface finish, size effect, grain directions, etc. on the strain-life curve.
6. The assumption to include (or not include) mean stress relaxation.

### Analysis

1. Accuracy of heat transfer analysis.
2. Stress analysis procedures (including thermal stresses) to determine inelastic notch strains or elastic stress concentration factors.
3. Procedure used to establish the stress-strain history at the notch.

It will be assumed in this demonstration that the service nominal stress varies from zero to peak  $S$  as shown in Figure 8.4. The equivalent stress at the notch also varies from zero to a peak of  $\sigma_e = K_t S$  where  $K_t$  is the theoretical stress concentration factor. The equivalent and assumed constant temperature is denoted as  $T$ .

Knowing the cyclic stress strain curve and applying Neuber's rule, the actual stress strain ( $\sigma - \epsilon$ ) curve at the notch can be established using standard methods of local strain analysis (1, 2). As illustrated in Figure 8.5, the resulting  $\sigma - \epsilon$  history so established will define the strain range  $\Delta\epsilon$  and mean stress  $\sigma_m$ , the information needed to compare with the strain-life relationship.

The strength of the material at a given cycle life  $N$  is defined by the general expression

$$\epsilon_a = B \left[ \frac{(S_f' - \sigma_m)}{E} (2N)^b + \epsilon_f' (2N)^c \right] \quad (8.3)$$

Figure 8.4  
Nominal Stress History at Notch

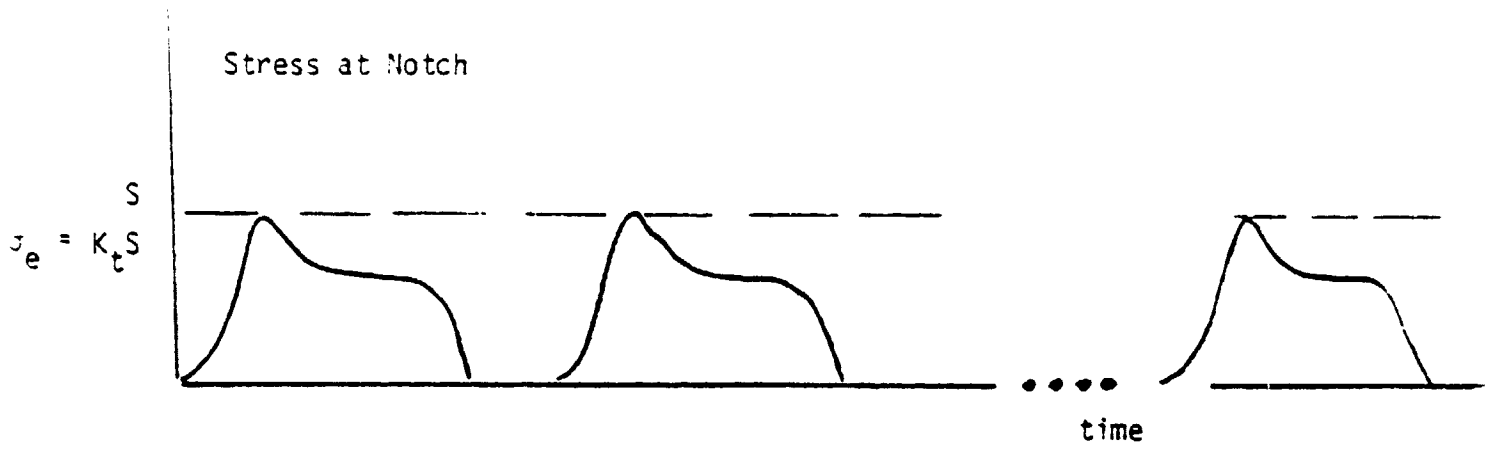
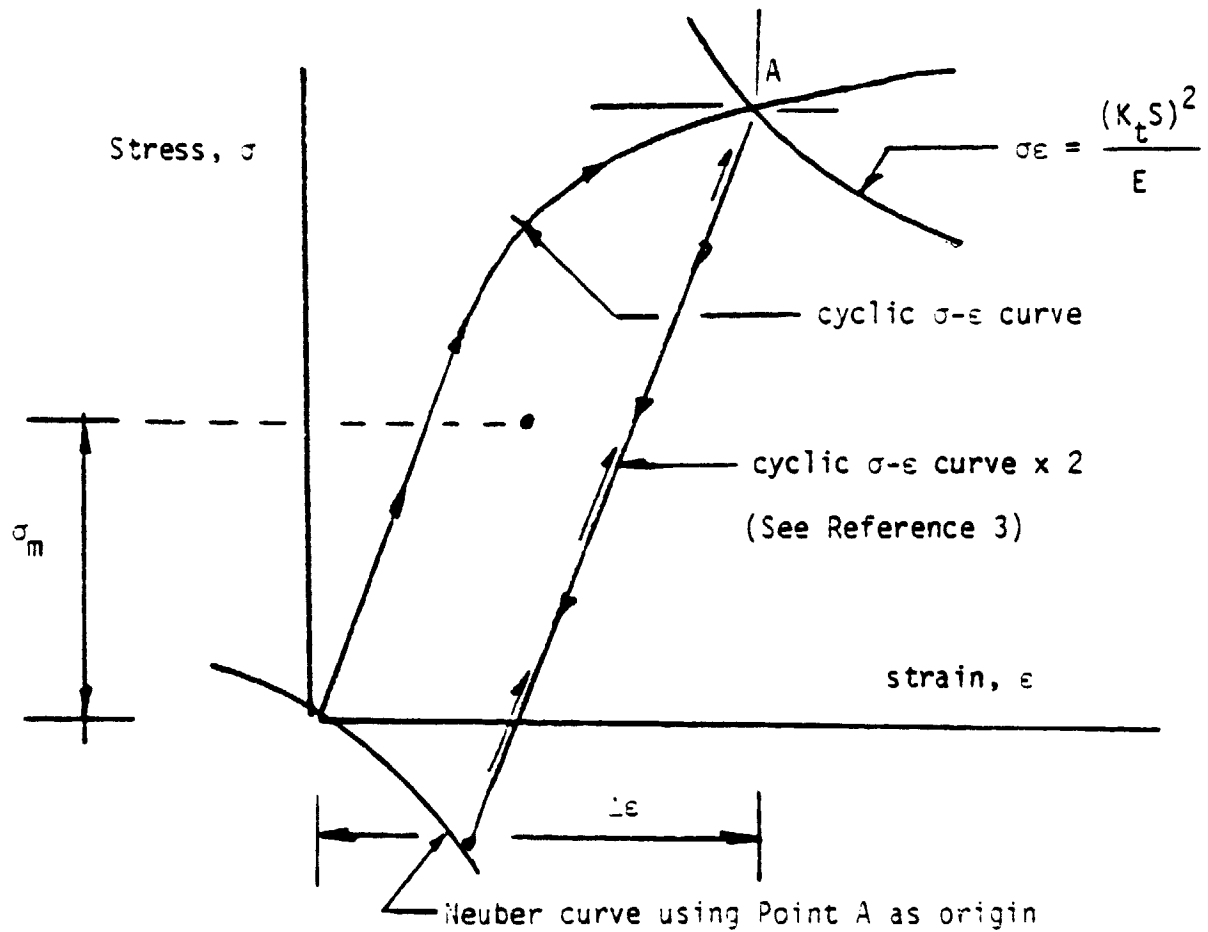


Figure 8.5  
Stress-Strain History at Notch for a Given Temperature T



in which  $\epsilon_a$  is strain amplitude and  $S_f'$ ,  $b$ ,  $\epsilon_f'$  and  $c$  are empirical constants. The term in brackets is the familiar strain-life relationship (4,5).  $B$  is a random variable which accounts for the scatter in fatigue data. Using this form, it is assumed that the standard deviation of  $\log(\text{strain})$  given  $N$  is a constant (See Section 5.3 for further detail).

The mean stress is  $\sigma_m$  and will be a function of both temperature  $T$  and the equivalent stress at the notch  $\sigma_e = K_t S$

$$\sigma_m = \sigma_m(T, K_t, S) \quad (8.4)$$

Similarly the service strain amplitude  $\epsilon_S$  will be a function of  $T$ ,  $K_t$  and  $S$ .

$$\epsilon_S = \Delta\epsilon/2 = \epsilon_S(T, K_t, S) \quad (8.5)$$

Consider operation to a given design life  $N_0$ . Failure is said to occur if "strength" is less than "stress". The event of failure is  $\epsilon_a \leq \epsilon_S$  or

$$B \left\{ \left[ \frac{S_f' - \sigma_m(T, K_t, S)}{E} \right] (2N_0)^b + \epsilon_f' (2N_0)^c \right\} \leq \epsilon_S(T, K_t, S) \quad (8.6)$$

In this example, it is assumed that both  $K_t$  and  $S$  are random variables reflecting uncertainty in analysis procedures. Both  $K_t$  and  $S$  will be the same for each cycle, but there is uncertainty in the value. Temperature  $T$  may be the most important variable but it is assumed to be constant just for the purpose of this demonstration. No theoretical difficulties would be introduced by considering  $T$  as a random variable.

Data for this problem is summarized in Table 8.3. The stress strain history shown in Figure 8.6 obtained using Neuber's rule (see Reference 3) shows that only elastic strains are present after the first cycle. Assuming no mean stress relaxation, the stress strain history will cycle between points B and C as the

Table 8.3

Data for Example Problem

<u>Random Variable</u>	<u>Median</u>	<u>Coefficient of Variation</u>
$K_t$	3	0.06
S	40 ksi (276 MPa)	0.10
$\sigma_e^*$	120 ksi (827 MPa)	0.117
B	1	0.18

Constants

No	5000 cycles
$S_f'$	210 ksi (1448 MPa)
b	-0.07
$\epsilon_f'$	1.20
c	-0.70
E	30,000 ksi (207 GPa)

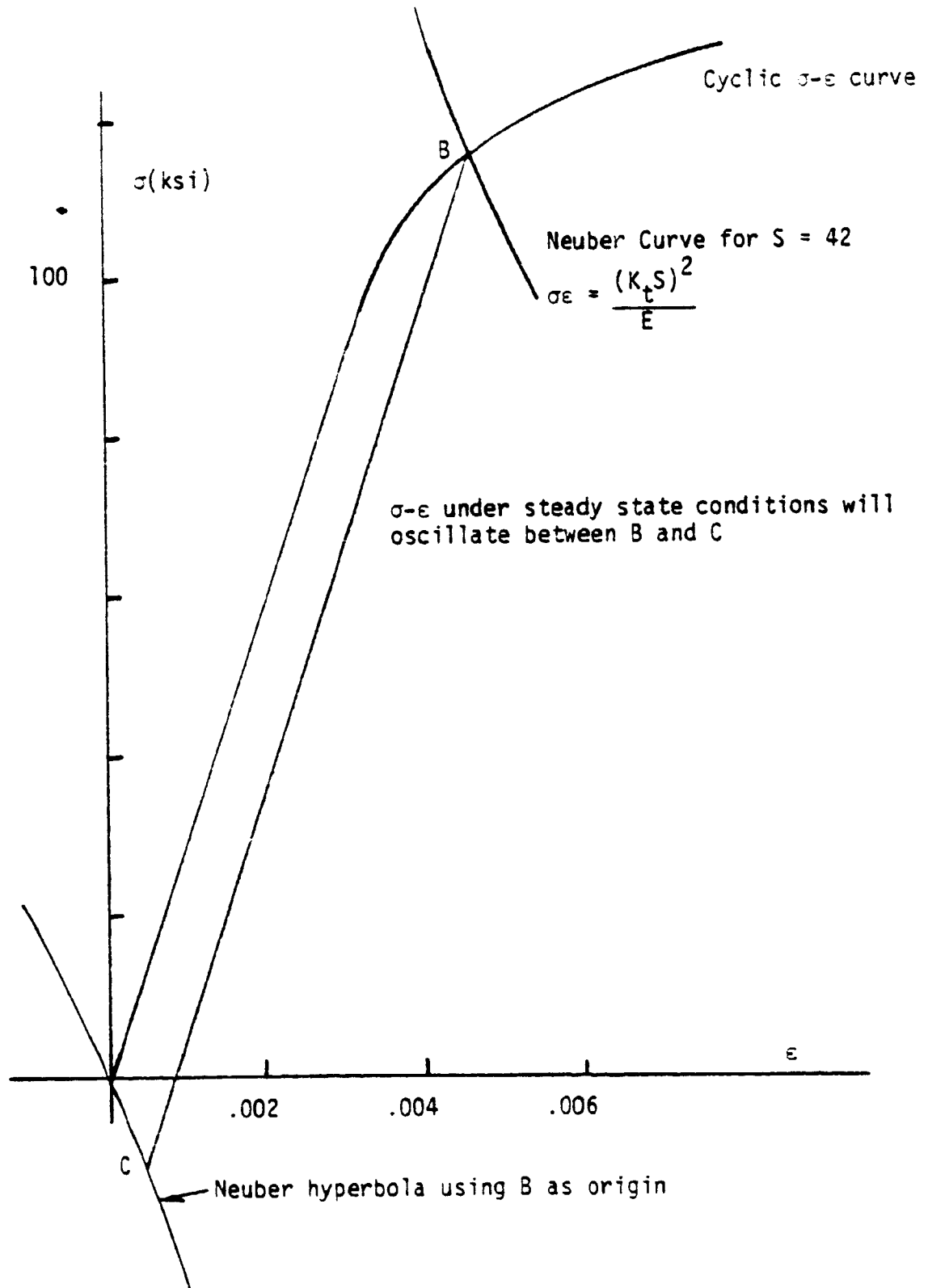
\*From basic properties of lognormal variates

$$\tilde{\sigma}_e = \tilde{K}_t \tilde{S}$$

$$\text{cov}(\sigma_e) = \sqrt{(1 + C_{K_t}^2)(1 + C_S^2) - 1}$$

Figure 8.6

Determination of Actual Stress-Strain History at Notch  
for Nominal Stress,  $S = 42$  ksi



nominal stress cycles between 0 and S. From a graphical analysis, (See Figure 8.6 for various values of S) it is shown that

$$\epsilon_S = \Delta\epsilon/2 = K_t S/2E = K_t S/60,000 \quad (8.7)$$

$$\sigma_m = 52.0 \text{ ksi (359 MPa) for any S}$$

Two methods of reliability analysis are used. In both it is assumed that  $K_t$ , S and B have lognormal distributions.

1. The Monte Carlo method;

Fatigue strength (strain at failure) is defined by Equation 8.3.

Noting that  $\sigma_m$  is constant, it follows upon substitution of the values of Table 8.3 that

$$\epsilon_a = B(0.00467)$$

Service strain is given by Equation 8.7. Failure occurs when

$$\epsilon_a \leq \epsilon_S$$

$$B(0.00467) \leq K_t S/60,000$$

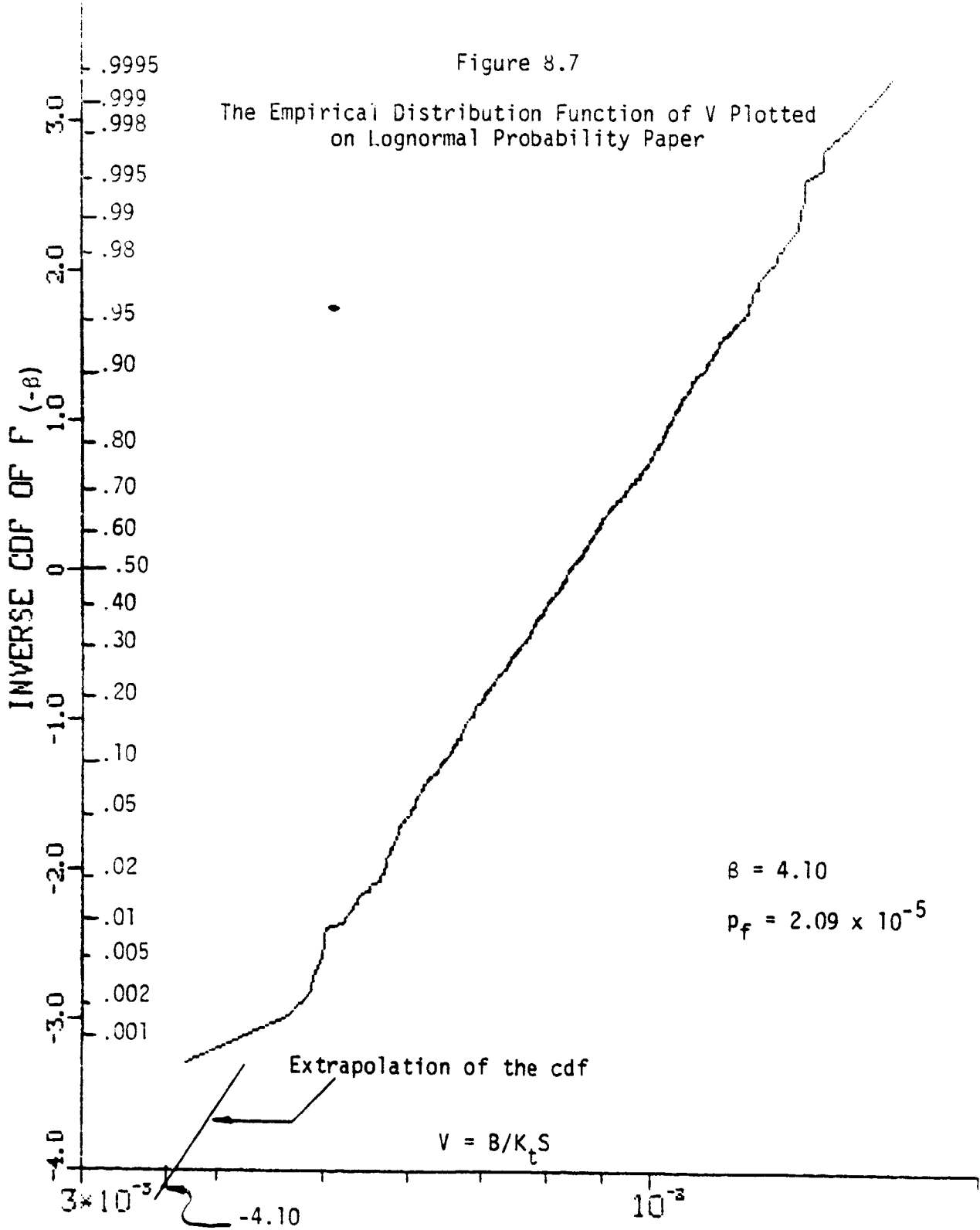
Isolating the random variables on the left hand side, the condition of failure is

$$V \equiv B/K_t S \leq 3.6 \times 10^{-3}$$

Values of  $K_t$ , S and B are selected at random from lognormal distributions having the parameters as given in Table 8.3. Then a sample of V is computed and its empirical distribution function is established as shown in Figure 8.7. The figure shows the estimate of the safety index as 4.10 and the probability of failure as  $p_f = 2.09 \times 10^{-5}$ .

Figure 8.7

The Empirical Distribution Function of V Plotted on Lognormal Probability Paper



2. The generalized safety index;

HASL was used to compute the generalized safety index,  $\beta = 3.97$ . A log transformation was made on all variables (See Section 7.4). The corresponding  $p_f = 3.51 \times 10^{-5}$ . The output is shown in Table 8.4. This problem satisfies the assumptions of the lognormal format (See Section 7.2) and therefore HASL will provide the same result as the closed form expressions (Equations 7.20, 7.21, 7.13 and 7.14). Moreover, the Monte Carlo solution should agree. The slight difference in results illustrates the error that is possible in extrapolating the distribution function into the tail regions.

In both cases an index of structural performance was computed. The decision with regard to safety must now be made by the designer.

As a footnote to this problem, consider the overall reliability of the disk. The estimate of fatigue crack initiation at a given site was  $p_f = 3.51 \times 10^{-5}$ . But assume that there are  $n = 36$  blades (i.e.  $n = 36$  initiation sites). The key question is what is the probability of crack initiation at any site?

This is a standard reliability problem, i.e. a series system where failure of the disk is assumed to occur if failure occurs in any component. System reliability is

$$R_S = \prod_{i=1}^n R_i = \prod_{i=1}^n (1 - p_i)^n \quad (8.8)$$

where  $R_i$  = reliability of  $i^{\text{th}}$  component and  $p_i$  = probability of failure of  $i^{\text{th}}$  component. If  $p_i$  = constant, as it is in this example

$$R_S = (1 - p)^n \quad (8.9)$$

In this example

$$\begin{aligned} R_S &= (1 - 3.51 \times 10^{-5})^{36} \\ &= 0.99874 \\ &= 0.999 \end{aligned}$$



Table 8.4  
 Output of HASL<sup>(a)</sup>, A Program for Computing the Generalized  
 Safety Index for Example 2

DESIGN VARIABLES

VARIABLE	TRANSFORMATION <sup>(b)</sup>	MEAN/MEDIAN <sup>(c)</sup>	STD DEV/COV <sup>(c)</sup>
B	1	.10000E+01	.18000E+00
KT	1	.30000E+01	.60000E-01
S	1	.40000E+02	.10000E+00

DESIGN POINT

VARIABLE	REDUCED VALUE	BASIC VALUE
B	XR(1) = -3.33459	X(1) = .55132E+00
KT	XR(2) = 1.11914	X(2) = .32082E+01
S	XR(3) = 1.86133	X(3) = .48152E+02

SAFETY INDEX, BETA = 3.97860

NOTES:

- (a) HASL is a digital computer program developed by P. H. Wirsching, University of Arizona, for computing the generalized safety index
- (b) A "1" indicates a log transformation on each variable
- (c) With a log transformation, the listed values are the Median and the Cov

Thus we would predict that only about 1 in 1000 of these disks would initiate a fatigue crack at the notch.

#### 8.4 Example 3: A Demonstration of the Efficiency of the Rackwitz Algorithm

The Rackwitz iteration method, as described in Section 7.5, was presented as a method for efficiently computing estimates of probabilities of failure for structural and mechanical components. The statistical distribution of each random variable is specified. The limit state can take any form, although risk estimates are expected to be slightly in error for non-linear cases.

Following is an example which illustrates the efficiency of the Rackwitz algorithm as compared to Monte Carlo. In this example, computer costs for the Monte Carlo solution were six times greater.

Consider a fatigue problem in which the fatigue strength is defined by the general strain-life relationship. Cyclic strain  $\epsilon_s$  at a notch has a constant amplitude (not random), but there is uncertainty regarding its magnitude. Mean stress  $\sigma_m$  also remains constant, but its value is uncertain. The limit state is written as

$$\epsilon_s = \frac{s_f - \sigma_m}{E} (2N)^b + \epsilon_f (2N)^c \quad (8.10)$$

The right hand side of the equation is fatigue strength. All terms have been defined previously in Section 5.3. Failure occurs if the strain (left hand side of Equation 8.10) exceeds the strength.

Values of the design factors are given in Table 8.5. Note that  $\epsilon_s$ ,  $s_f$ ,  $\epsilon_f$ , and  $\sigma_m$  are considered to be random variables. They have different distributions. In general it would be very difficult to compute the probability of failure  $p_f$  using the exact form of Equation 7.5. However two

Table 8.5

Values of Design Factors for Example Problem

Design Life,  $N = 10^4$  cycles

Modulus of Elasticity,  $E = 29,500$  ksi

Fatigue Strength Exponent,  $b = -0.108$

Fatigue Ductility Exponent,  $c = -0.540$

Random Design Factors	Distribution	Mean	Standard Deviation
Fatigue Strength Coefficient, $s'_f$	Lognormal*	148.	.042
Fatigue Ductility Coefficient, $\epsilon'_f$	Lognormal*	.491	.263
Strain Amplitude, $\epsilon_s$	EVD	.002	.0003
Mean Stress, $\sigma_m$	Weibull	20.	2.

\* Values of the mean and standard deviation stated for the lognormal are really median and coefficient of variation respectively.

approximate methods will be employed to estimate  $p_f$ .

The Monte Carlo solution is presented on Figure 8.8. The safety index determined graphically is approximately  $\beta = 2.95$ . The corresponding  $p_f = 0.0016$ .

Solutions using the Rackwitz algorithm using SUPER-HASL is summarized in Table 8.5. Note that the results are essentially the same. In this example, the approximate costs of running the Monte Carlo program and SUPER-HASL were \$4.00 and \$0.70 respectively. Moreover, computer core requirements for SUPER-HASL are very modest. The program can be operated on a small computer.

The quality and efficiency of the Rackwitz algorithm will be the subject of a continuing study.

### 8.5 Summary Comments

The three examples provided in this chapter demonstrate application of fatigue reliability methods. The approach chosen however depends upon the statistical information available as well as the form of the limit state.

This report is considered to be preliminary. Additional examples are being considered and will be included in future reports.

### 8.6 References for Chapter 8

1. Fatigue Under Complex Loading, SAE, No. AE-6, 1977.
2. Dowling, N.E., "Fatigue Life Predictions for Complex Load Versus Time Histories", to appear in Decade of Progress in Pressure Vessels and Piping, ASME.
3. Socie, D.F. and Morrow, J.D.I., "Review of Contemporary Approaches to Fatigue Damage Analysis", FCP Report No. 24, The University of Illinois, 1976.
4. Fatigue Design Handbook, SAE, No. AE-4, 1968.
5. Fuchs, H.O. and Stephens, R.I., Metal Fatigue in Engineering, Wiley, 1980.

SUPER-HASL\* EFFICIENT ITERATIVE SCHEME FOR COMPUTING PROBABILITIES OF FAILURE:  
 R. RACKWITZ, Technische Universität, München, Germany (1976)

<u>INPUT</u>		<u>DESIGN VARIABLES</u>	
VARIABLE		MEAN/MEDIAN	STD DEV/COV
ES( $E_s$ )	EVD	.20000E-02	.30000E-03
SF( $S'_f$ )	LOG	.14800E+03	.42000E-01
SM( $\sigma_m$ )	WEIBULL	.20000E+02	.20000E+01
EF( $\epsilon'_f$ )	LOG	.49120E+00	.26300E+00

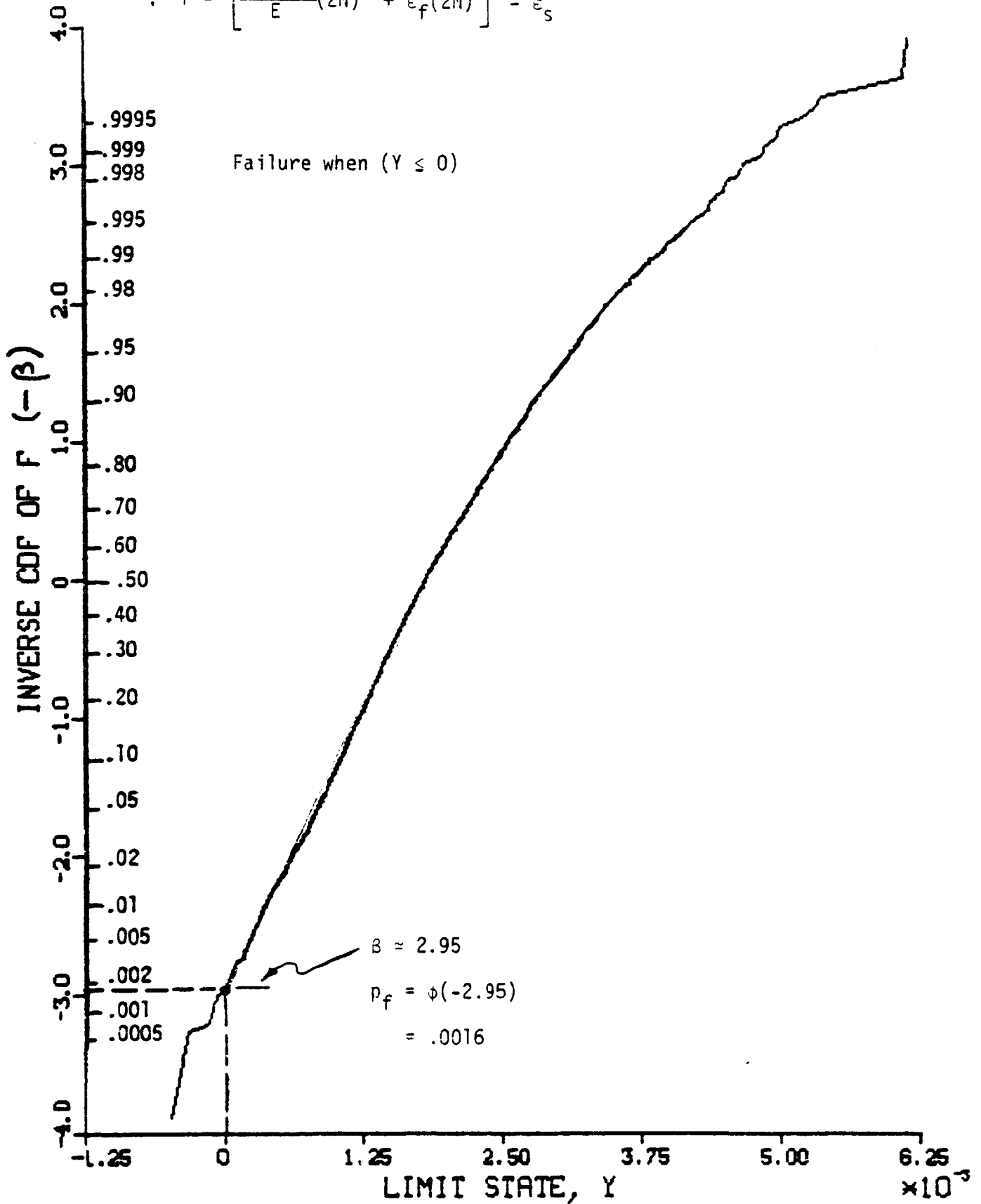
<u>RESULTS</u>	<u>DESIGN POINT</u>	
VARIABLE	REDUCED VALUE	BASIC VALUE
ES	XR(1) = 2.47612	X(1) = .30365E-02
SF	XR(2) = -.26978	X(2) = .14633E+03
SM	XR(3) = .08325	X(3) = .20402E+02
EF	XR(4) = -1.53516	X(4) = .33024E+00

SAFETY INDEX, BETA = 2.92711  
 PROBABILITY OF FAILURE = .001716

\* Computer program developed by Colleen F. Kelly, graduate student in Aerospace and Mechanical Engineering at the University of Arizona

Figure 8.8 Monte Carlo Solution to Example Problem

$$Y = \left[ \frac{S_f' - \sigma_m}{E} (2N)^D + \epsilon_f' (2N)^C \right] - \epsilon_s$$



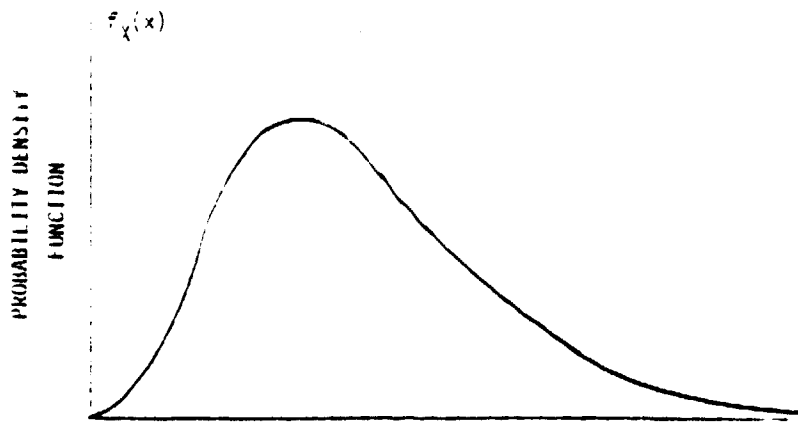
## APPENDIX 1 THE LOGNORMAL DISTRIBUTION AND PROPERTIES OF LOGNORMAL VARIABLES

The lognormal format is proposed herein as a structural reliability model. In this regard Appendix 1 provides, for reference purposes, a detailed description of the lognormal distribution as well as certain properties of lognormal variables that are used for the design equations. Additional information on the lognormal is given by Benjamin and Cornell and Ang and Tang\*.

Given a random variable  $X$ . If  $Y = \ln X$  (or  $\log_{10} X$ ) has a normal distribution, then  $X$  is said to have a "lognormal" distribution. The mean and standard deviation of  $X$  and  $Y$  are  $(\mu_X, \sigma_X)$  and  $(\mu_Y, \sigma_Y)$  respectively. The probability density function (pdf) of  $X$  is

$$f_X(x) = \frac{1}{\sqrt{2\pi} \sigma_Y x} \exp \left[ -\frac{1}{2\sigma_Y^2} (\ln x - \mu_Y)^2 \right] \quad (A1.1)$$

for  $x > 0$



\* Benjamin, J. R. and Cornell, C. A., Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill, 1970.

Ang, A. H. S. and Tang, W. H., Probability Concepts in Engineering Planning and Design, Wiley, 1975.

The "statistical parameters" for the lognormal distribution are  $(\mu_Y, \sigma_Y)$ . The mean and standard deviation of the lognormal variables are  $(\mu_X, \sigma_X)$  however the median of  $X$ , denoted as  $\tilde{X}$  and the coefficient of variation of  $X$ ,  $C_X$ , are often used. Relationships between parameters and moments are

1. Base e

$$\mu_Y = \ln \mu_X - \frac{1}{2} \ln (1 + C_X^2), \text{ or } \mu_Y = \ln \tilde{X} \quad (\text{A1.2})$$

$$\sigma_Y^2 = \ln (1 + C_X^2) \quad (\text{A1.3})$$

$$\mu_X = \exp \left[ \mu_Y + \frac{1}{2} \sigma_Y^2 \right] \quad (\text{A1.4})$$

$$\sigma_X^2 = e^{2\mu_Y} e^{\sigma_Y^2} (e^{\sigma_Y^2} - 1) \quad (\text{A1.5})$$

$$C_X = \sqrt{e^{\sigma_Y^2} - 1}$$

2. Base 10

$$\mu_Y = \log_{10} \mu_X - \frac{1}{2} \log_{10} (1 + C_X^2), \text{ or } \mu_Y = \log_{10} \tilde{X} \quad (\text{A1.6})$$

$$\sigma_Y^2 = 0.434 \log_{10} (1 + C_X^2) \quad (\text{A1.7})$$

$$\mu_X = 10^{\left\{ \mu_Y + \frac{1}{2} \log_{10} (\sigma_Y^2 / 0.434) \right\}} \quad (\text{A1.8})$$

$$C_X = \sqrt{10^{\left( \frac{\sigma_Y^2}{0.434} \right)} - 1} \quad (\text{A1.9})$$

In general, it is more meaningful to use the median (rather than the mean) as a measure of central tendency for any random variable having a large variance. The median of  $X$ , denoted as  $\tilde{X}$ , is defined by the relation  $F_X(\tilde{X}) = 0.50$ , where  $F_X$  is the distribution function.



Thus  $\tilde{X}$  defines the point below which 50% of the population is expected to fall.

The expression for  $\mu_Y$  in Equations A1.2 and A1.6 can be derived from the definition of the median

$$\begin{aligned} 0.50 &= P(X \leq \tilde{X}) = P(\ln X \leq \ln \tilde{X}) \\ &= P(Y \leq \ln \tilde{X}) \\ &= F_Y(\ln \tilde{X}) \end{aligned} \tag{A1.10}$$

But,  $\tilde{Y}$ , the median of  $Y$ , is

$$0.50 = F_Y(\tilde{Y}) \tag{A1.11}$$

and it follows upon comparing Equations A1.10 and A1.11 that

$$\tilde{Y} = \ln \tilde{X} \tag{A1.12}$$

Recall that  $Y$  has a normal distribution and as a result of symmetry, the mean of  $Y$  equals the median,  $\mu_Y = \tilde{Y}$ . Thus

$$\mu_Y = \ln \tilde{X} \tag{A1.13}$$

The mean of  $X$  can be expressed as a function of the median. Equation A1.13 can be written as,

$$\tilde{X} = \exp(\mu_Y) \tag{A1.14}$$

Equations A1.2 can be solved for  $\mu_X$ ,

$$\mu_X = \exp(\mu_Y) \sqrt{1 + C_X^2} \tag{A1.15}$$

Dividing Equation A1.15 by A1.14,

$$\mu_X = \tilde{X} \sqrt{1 + C_X^2} \quad (\text{A1.16})$$

This relationship between  $\tilde{X}$  and  $\mu_X$  proves to be useful because information on  $\mu_X$  and  $C_X$  is often available. But when probability computations are made, it is generally more convenient to use  $\tilde{X}$ .

How to make probability calculations.

$$\begin{aligned} P(X \leq x_0) &= P(\ln X \leq \ln x_0) \\ &= P(Y \leq \ln x_0) \end{aligned} \quad (\text{A1.17})$$

But  $Y$  is normally distributed with mean and standard deviation  $(\mu_Y, \sigma_Y)$ .

Thus,

$$\begin{aligned} P(X \leq x_0) &= P\left(\frac{Y - \mu_Y}{\sigma_Y} \leq \frac{\ln x_0 - \mu_Y}{\sigma_Y}\right) \\ &= \phi\left(\frac{\ln x_0 - \mu_Y}{\sigma_Y}\right) \end{aligned} \quad (\text{A1.18})$$

where  $\phi$  is the standard normal distribution function.

Replacing  $\mu_Y$  with  $\tilde{X}$  using Equation 4.13, an expression which is generally more convenient to use follows,

$$P(X \leq x_0) = \phi\left(\frac{\ln(x_0/\tilde{X})}{\sigma_Y}\right) \quad (\text{A1.19})$$

A similar expression for probabilities is valid when base 10 logs are used.

How to estimate parameters from data. Given a random sample, of size  $n$ , of  $X$ ,

$$\underline{X} = (X_1, X_2, \dots, X_n) \quad (\text{A1.20})$$

the data is converted to a random sample of  $Y$ .

$$\underline{Y} = (Y_1, Y_2, \dots, Y_n) \quad (\text{A1.21})$$

where  $Y_i = \ln X_i$ .

Then the estimates of  $\mu_Y$  and  $\sigma_Y$  are\*,

$$\mu_Y \rightarrow \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (\text{A1.22})$$

$$\sigma_Y^2 \rightarrow s_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (\text{A1.23})$$

Property of lognormal variates useful for design. Consider an expression involving several random variables,

$$V = \frac{AX^{bY^c}}{U^d} \quad (\text{A1.24})$$

where  $X$ ,  $Y$ ,  $U$  are independent lognormal variates.  $b$ ,  $c$ , and  $d$  are constants. It will be shown that  $V$  also has a lognormal distribution. A form for probability calculations on  $V$  will be derived. Taking the log of both sides,

$$\log V = \log A + b \log X + c \log Y - d \log U \quad (\text{A1.25})$$

\*Maximum likelihood estimators

Let,

$$V' = \log V, \quad X' = \log X, \dots \text{etc} \quad (\text{A1.26})$$

Then,

$$V' = \log A + bX' + cY' - dU' \quad (\text{A1.27})$$

Note that  $X'$ ,  $Y'$  and  $U'$  are normal. It follows from the addition property of normal variates that  $V'$  is also normal. Therefore  $V$  has a lognormal distribution.

To make probability calculations, it is necessary to find parameters  $(V, \sigma_{V'})$ . Taking the expected value of both sides of Equation A1.27,

$$\mu_{V'} = \log A + b\mu_{X'} + c\mu_{Y'} - d\mu_{U'} \quad (\text{A1.28})$$

But recall (Equation A1.13) that,  $\mu_{V'} = \log V, \dots$  etc. Substituting for each  $\mu$ , Equation A1.28 becomes

$$\log \bar{V} = \log A + \log \bar{X}^b + \log \bar{Y}^c + \log \bar{U}^{-d} \quad (\text{A1.29})$$

Taking both sides to the "e" power,

$$\bar{V} = \frac{A \bar{X}^b \bar{Y}^c}{\bar{U}^d} \quad (\text{A1.30})$$

Taking the variance of both sides of Equation A1.27,

$$\sigma_{V'}^2 = b^2 \sigma_{X'}^2 + c^2 \sigma_{Y'}^2 + d^2 \sigma_{U'}^2 \quad (\text{A1.31})$$

But recall from Equation A1.3 that  $\sigma_{X'}^2 = \ln(1 + C_X^2) \dots$  etc. It follows that

$$\sigma_{V'}^2 = \ln(1 + C_X^2)b^2 + \ln(1 + C_Y^2)c^2 + \ln(1 + C_U^2)d^2 \quad (\text{A1.32})$$

Then to compute probabilities, consider Equation A1.19,

$$P(V \leq v_0) = \Phi \left[ \frac{\ln(v_0/\bar{V})}{\sigma_{V'}} \right] \quad (\text{A1.33})$$

#### Example of the "Lognormal Format"

Consider a design problem in which time failure of a component, denoted as  $T$ , is a random variable.  $T$  is known to be a function of random design factors  $\Theta$  and  $\Lambda$ . (A description of the physical problem and what these terms mean is not important in this example.)

$$T = \Theta^a / \Lambda^b$$

where  $a$  and  $b$  are constants. It is assumed that the median and COV of  $\Theta$  and  $\Lambda$  are known. The design life of the component is  $T_0$ .

The event of failure of the component is  $(T < T_0)$ , and the probability of failure is,

$$p_f = P(T \leq T_0)$$

At this point, if we can assume that  $\Theta$  and  $\Lambda$  have lognormal distributions, then  $T$  has a lognormal distribution, and from Eq A1.19, (note the reversal of  $T_0$  and  $\bar{T}$ )

$$p_f = \Phi \left[ - \frac{\ln(\bar{T}/T_0)}{\sigma_T} \right]$$

where  $T' = \ln T$ . Define the safety index  $\beta$  as,

$$\beta = \ln (\tilde{T}/T_0) / \sigma_{T'}$$

Then,

$$p_f = \Phi(-\beta)$$

$\beta$  plays the same role in probabilistic design as the factor of safety does in conventional design. To compute  $\beta$ , the parameters must be evaluated using a form of Equations A1.30 and A1.32,

$$\tilde{T} = \tilde{\Theta}^a / \tilde{\Lambda}^b$$

$$\sigma_{T'} = \sigma_{\ln T} = [ \ln\{ (1 + C_{\Theta}^2)^{a^2} (1 + C_{\Lambda}^2)^{b^2} \} ]^{1/2}$$

**END**

**DATE**

**FILMED**

FEB 23 1982