(4)

AD-A207 378

David Taylor Research Center

Bethesda, Maryland 20084-5000

DTRC-SME-88/63 January 1989

Ship Materials Engineering Department

Research and Development Report

ESTIMATION OF LOWER BOUND PROPERTIES FROM MATERIAL TEST DATA

by

M. T. Kirk

Approved for public release; distribution unlimited

DTRC-SME-88/63 Estimation of Lower Bound Properties from Material Test Data

MAJOR DTRC TECHNICAL COMPONENTS

CODE 011	DIRECTOR OF TECHNOLOGY, PLANS AND ASSESSMENT
12	SHIP SYSTEMS INTEGRATION DEPARTMENT
14	SHIP ELECTROMAGNETIC SIGNATURES DEPARTMENT
15	SHIP HYDROMECHANICS DEPARTMENT
16	AVIATION DEPARTMENT
17	SHIP STRUCTURES AND PROTECTION DEPARTMENT
18	COMPUTATION, MATHEMATICS & LOGISTICS DEPARTMENT
19	SHIP ACOUSTICS DEPARTMENT
27	PROPULSION AND AUXILIARY SYSTEMS DEPARTMENT
28	SHIP MATERIALS ENGINEERING DEPARTMENT

DTRC ISSUES THREE TYPES OF REPORTS:

1. **DTRC reports, a formal series,** contain information of permanent technical value. They carry a consecutive numerical identification regardless of their classification or the originating department.

2. **Departmental reports, a semiformal series,** contain information of a preliminary, temporary, or proprietary nature or of limited interest or significance. They carry a departmental alphanumerical identification.

3. **Technical memoranda, an informal series,** contain technical documentation of limited use and interest. They are primarily working papers intended for internal use. They carry an identifying number which indicates their type and the numerical code of the originating department. Any distribution outside DTRC must be approved by the head of the originating department on a case-by-case basis.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for Public Release, Distribution Unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S) DTRC/SME-88-63	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION DTRC 6b. OFFICE SYMBOL (If applicable) Code 2814	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Bethesda, MD 20084-5000	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION DTRC 8b. OFFICE SYMBOL (If applicable) 011.5	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 62234N PROJECT NO: TASK NO.: RS345S50 WORK UNIT ACCESSION NO: DN507603

11. TITLE (Include Security Classification)

(U) ESTIMATION OF LOWER BOND PROPERTIES FROM MATERIAL TEST DATA

12. PERSONAL AUTHOR(S) M.T. Kirk

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
R&D	FROM 9/87 TO 1/89	January 1989	24

16. SUPPLEMENTARY NOTATION

1-2814-198-20, MA1.6/2

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Lower Bond, Statistical Evaluation, Monte Carlo, Material Characterization.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Monte Carlo simulations were performed to determine how the accuracy of lower bound values estimated from experimental data is influenced by sample size, required confidence level, and assumed statistical mode. Population distributions having different degrees of skewness, selected to bracket those expected in actual experimental data, were studied. For nearly every case considered, lower bound estimates calculated using Log-Normal statistics were more accurate than estimates calculated using either Normal or Weibull statistics. It was demonstrated that testing more than three samples per condition can greatly reduce the error associated with the lower bound estimate. However, after the thirteenth sample, no additional sample will reduce the lower bound estimation error by more than 2.5% for all statistical distribution/ confidence level combinations considered. When applied to material properties for which the population distribution has been established by previous testing, it was demonstrated that a Monte Carlo simulation can be used to assess the maximum expected lower bound estimation error as a function of sample size and confidence level. (SEE BACK)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL M.T. Kirk	22b. TELEPHONE (Include Area Code) (301) 267-3755 22c. OFFICE SYMBOL Code 2814

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

This information can be used to determine the minimum number of specimens needed to obtain a lower bound estimate of acceptable accuracy when sampling a known population.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Monte Carlo simulations were performed to determine how the accuracy of lower bound values estimated from experimental data is influenced by sample size, required confidence level, and assumed statistical model. Population distributions having different degrees of skewness, selected to bracket those expected in actual experimental data, were studied. For nearly every case considered, lower bound estimates calculated using Log-Normal statistics were more accurate than estimates calculated using either Normal or Weibull statistics. It was demonstrated that testing more than three samples per condition can greatly reduce the error associated with the lower bound estimate. However, after the twelfth sample, no additional sample will reduce the lower bound estimation error by more than 2.5% for all statistical distribution / confidence level combinations considered. When applied to material properties for which the population distribution has been established by previous testing, it was demonstrated that a Monte Carlo simulation can be used to assess the maximum expected lower bound estimation error as a function of sample size and confidence level. This information can be used to determine the minimum number of specimens needed to obtain a lower bound estimate of acceptable accuracy when sampling a known population.

ADMINISTRATIVE INFORMATION

This report was prepared as part of the Surface Ship and Craft Materials Block under the sponsorship of Mr. I. Caplan (DTRC 011.5). This effort was performed at this Center under Program Element 62234N, Task Area RS345S50, Work Unit 1-2814-198-20. The work was performed under the supervision of Mr. T.W. Montemarano. This report satisfies milestone MA1.6/2

INTRODUCTION

For either engineering or research and development purposes, it is often necessary to determine the properties of a material (e.g. strength, toughness) using small experimental data sets. Lower bound properties, estimated from these data, can then be used to conservatively assess the fitness of a component for continued service. The accuracy of this estimated lower bound depends on the variability of the property, the confidence required of the estimate, and the amount of experimental data available.

Although material specifications and surveillance programs frequently base component acceptance or rejection on the lowest of three experimental datum [1], there is no established relationship between this value and the actual lower bound. Jutla and Garwood [2] demonstrated that, if nothing is known a priori regarding the sampled population, the lowest of three data falls, with 90% confidence, below only 46% the entire population, indicating that this value is not a very accurate lower bound measure. As shown in Figure 1 [2], these results also indicate that the lowest measured value approximates a 90% confidence level lower bound value only for rather large samples (greater than 24 values). Any alternative to estimating the lower bound with the lowest measured value involves a statistical evaluation of the data. By making assumptions regarding the population distribution sampled by experimental data, statistical models allow the available data to be extrapolated, or interpolated, to establish a lower bound value.

Three statistical models commonly used to analyze material data are the Normal statistical model, the Log-Normal statistical model, and the Weibull statistical model. While a lower bound can be estimated using any of these models, the different characteristics of each, illustrated in Figure 2, cause these estimates to depend on the model used to make the estimate. Unfortunately, there is no straightforward way to assess the accuracy of these various lower bound estimates. In recent work, Doig [3] used a Monte Carlo simulation to determine the accuracy with which lower bound estimates can be made based on limited data using Normal and Weibull statistical models. This work indicated that a Weibull model gives more accurate 95% confidence lower bound estimates than does a Normal model for a variety of population distributions.

OBJECTIVES

The objectives of this study are as follows:

1. To determine what statistical model, of Normal, Log-Normal, and Weibull, provides the most accurate lower bound estimate for different sample sizes and confidence levels.

2. To determine at what point additional sampling fails to substantially reduce the error of the estimated lower bound value.

3. To demonstrate how a Monte Carlo analysis can be used to assess the maximum lower bound estimation error when samples are drawn from a known population.

To achieve these objectives, the procedure suggested by Doig was employed. Data was drawn from populations having different degrees of skewness, these having been selected to bracket those commonly observed in actual experimental data. The first two objectives were addressed by performing Monte Carlo simulations of a random sampling process using data from these populations. An experimentally determined population was analyzed in a similar manner to meet the third objective.

MONTE CARLO SIMULATION

Kleijnen [4] discussed how Monte Carlo simulations are used to determine parameters that describe a stochastic variable's distribution (e.g. mean, variance, lower bound, upper bound). During a simulation, samples are randomly drawn from the population being studied. The simulation thus imitates the process of characterizing a lot of material using data from mechanical test specimens (e.g. Charpy V-Notch, Compact Tension, Tensile) removed from the lot. The population distribution used in a Monte Carlo simulation can either be derived from experimental data, or based on a population distribution equation.

In this study, four populations, having shapes ranging from skewed left to skewed right, were studied. These populations are shown in Figure 3. The Monte Carlo simulations, shown schematically in Figure 4, were conducted as follows:

1. A sample of n values were randomly drawn from the population being studied.
2. A b% confidence lower bound value was estimated from this sample, using Normal, Log-Normal, and Weibull statistical models.
3. Steps 1 and 2 were repeated 1,000 times to determine:
   a. The range of predicted b% lower bound estimates expected for each statistical model.
   b. The maximum lower bound estimation error, $|E_{max}|$, as defined in Figure 4.

This process was repeated for each distribution for values of n (sample size) ranging from 3 to 31 at b = 90%, 95%, and 99% confidence levels. Descriptions of how lower bound estimates are made using Normal, Log-Normal, and Weibull statistics can be found in references [3,5-6].

RESULTS AND DISCUSSION

ACCURACY OF STATISTICALLY ESTIMATED LOWER BOUND VALUES

Figure 5 shows typical results from these analyses. When the sampled distribution was approximately symmetric or skewed right (Figure 5a), the ranges of all three lower bound estimates converged to the true lower bound as the sample size increased. However, when a skewed left distribution was sampled (Figure 5b), only Weibull and Log-Normal lower bound estimates converged to the true lower bound value. In this case, the Normal lower bound estimates remained negatively biased even for large sample sizes. This bias occurred due to the symmetry assumed by a Normal statistical model. Figure 5 also shows that lower bounds estimated from small samples depend significantly on the statistical model used to make the estimate. In particular, the Normal statistical model estimated negative lower bounds, even when all of the values in the sample were positive. This occurred because the existence of a finite lower bound is not assumed by the Normal statistical model.

To rank these statistical models by lower bound estimation accuracy, the normalized maximum estimation error; $|E_{max}|$/Standard Deviation , $|E_{max}|$ having been defined in Figure 4; was computed for each distribution / confidence level combination. Normalizing the errors in this manner facilitates comparison of estimation errors for different distributions on a common scale. These data, presented in Figure 6, show that Normal statistics estimated the least accurate lower bounds in every instance, especially when the sampled distribution was heavily skewed left. Of the other two statistical models, Log-Normal lower bound estimates were typically either more accurate or nearly as accurate as Weibull estimates. The one major exception to this trend occurred for samples of 6 or fewer values drawn from a distribution that was heavily skewed left. In this case, Weibull estimated lower bounds were more accurate than Log-Normal estimated lower bounds for all confidence levels considered. However, this exception is sufficiently restricted that lower bounds calculated using Log-Normal statistics would be expected to be the most accurate when sampling from an unknown population.

Figure 6 only shows the results of the Monte Carlo simulation for the 95% confidence level; the trends for 90% and 99% confidence levels being essentially the same. Figure 7 compares the maximum lower bound estimation error for these confidence levels to the maximum estimation error at the 95% confidence level. In this figure, y-axis ratios near unity indicate that the accuracy of the lower bound estimate is not sensitive to confidence level. Thus, these data indicate that the accuracy of Log-Normal lower bound estimates are the least sensitive to confidence level, while Normal lower bound estimates are the most sensitive. There is, however, a general trend in Figure 7 of increasing lower bound estimation error with increasing confidence level for all three statistical models, implying that high confidence lower bound estimates are more difficult to make accurately than low confidence lower bound estimates. This occurs because, generally speaking, lower bound estimates are made using a formula of the following type:

$$\text{Estimated Lower Bound} = \text{Estimated Average} - \beta\cdot(\text{Estimated Standard Deviation})$$

From this formula, it follows that the error in the estimated lower bound is the error in the estimated average plus β times the error in the estimated standard deviation. The β value depends on the statistical model used to evaluate the data. It increases with both decreasing sample size and with increasing confidence level, making β quite large for high confidence lower bound estimates based on small samples. Thus, errors observed in high confidence lower bound estimates based on small samples are large not only due to the errors in the estimated average and standard deviation from which they are calculated, but also due to the large β values inherent to this type of estimate.

LOG-NORMAL LOWER BOUND ESTIMATES

It was demonstrated above that, in most cases, Log-Normal lower bound estimates are both more accurate and less sensitive to confidence level than either Normal or Weibull lower bound estimates. In this section, the effect of sample size and confidence level on Log-Normal lower bound estimates are examined in further detail.

In experimental studies, three replicate tests are often performed to establish trends with varying test conditions. While this degree of replication is typically sufficient for these purposes, the data presented in Figure 6 indicate that lower bounds calculated from such a small sample could be in error by between 29% and 163% of the standard deviation, depending upon the distribution sampled. In other situations, where such inaccuracy is unacceptable due to the dire consequences of structural failure, additional data would be required to improve the lower bound estimation accuracy. Figure

8 shows that these additional data considerably reduce the lower bound estimation error, the degree of error reduction not being strongly effected by either the confidence level or by the sampled distribution. In all cases, the first few additional values produce the greatest error reduction. The data presented in Figure 8 can be used to assess when the achieved error reduction fails to justify the cost of conducting additional experiments. While this 'break even' point depends on the ultimate application of the data, it would be logical to terminate data collection when the amount of error reduction expected by obtaining the next sample becomes small.

For general guidance in designing experimental test programs, it is useful to note from Figure 8 that after the twelfth sample is obtained, no additional sample will reduce the lower bound estimation error by more than 2.5% for all statistical distribution / confidence level combinations considered. However, this observation should be considered with the fact that the 95% confidence Log-Normal lower bound estimate calculated from a sample having twelve values may be in error by 10% to 47% of the distribution standard deviation, as indicated in Figure 6. Thus, samples of twelve values do not guarantee the accuracy of the estimated lower bound; rather, large increases in sample size beyond twelve appear to be needed to substantially improve the lower bound estimation accuracy.

APPLICATION OF MONTE CARLO SIMULATION TO ACTUAL DATA

When considerable experience exists with a particular material, the results of a Monte Carlo simulation can be used to full advantage. One instance where such detailed data exists is for Charpy V-Notch (CVN) tests at +30°F of a high strength steel where fracture is by microvoid coalescence. Figure 9 shows a histogram constructed from the results of 1559 CVN tests performed on this material. The probability distribution used in the Monte Carlo simulation was based on these data.

The results of this analysis are presented in Figure 10. In this figure, the maximum lower bound estimation error was expressed as a percent of the true lower bound, rather than as a certain number of standard deviations, because the numerical values of the true lower bounds were known from the data shown in Figure 9. These results indicate that accurate lower bound estimates having high confidence levels cannot be obtained with only three data values in this particular situation. Further, these data demonstrate that collecting more than twelve samples does not significantly reduce the maximum lower bound estimation error, as was predicted in the previous section. Information of this type can be used to determine the minimum number of specimens needed to obtain a lower bound estimate of acceptable accuracy when sampling from a known population.

SUMMARY AND CONCLUSIONS

This study examined the influence of sample size, confidence level, and statistical model on the accuracy with which lower bound values can be estimated from experimental data. Based on Monte Carlo simulations using mathematically and experimentally derived probability distributions, the following conclusions may be drawn:

1. In situations where the statistical distribution of the quantity being sampled is not known, lower bound estimates made using Log-Normal statistics are generally more accurate and less sensitive to confidence level than those made using either Normal or Weibull statistics for sample sizes between 3 and 31 and confidence levels between 90% and 99%.

2. Testing more than three specimens does not linearly decrease the error associated with the estimated lower bound value; the most significant error reductions being achieved by the first few additional specimens tested. The amount of error reduction achieved by additional testing does not depend strongly on either the distribution sampled or on the confidence level of the lower bound estimate. It was determined that, after the twelfth experiment, no additional experiment will reduce the lower bound estimation error by more than 2.5% for all statistical distribution / confidence level combinations considered.

3. A Monte Carlo simulation can be used to assess the maximum expected lower bound estimation error as a function of sample size and confidence level, provided that the characteristics of the population have been established by previous testing. The results of this type of analysis can be used to determine the minimum sample size needed to obtain a lower bound estimate of acceptable accuracy.

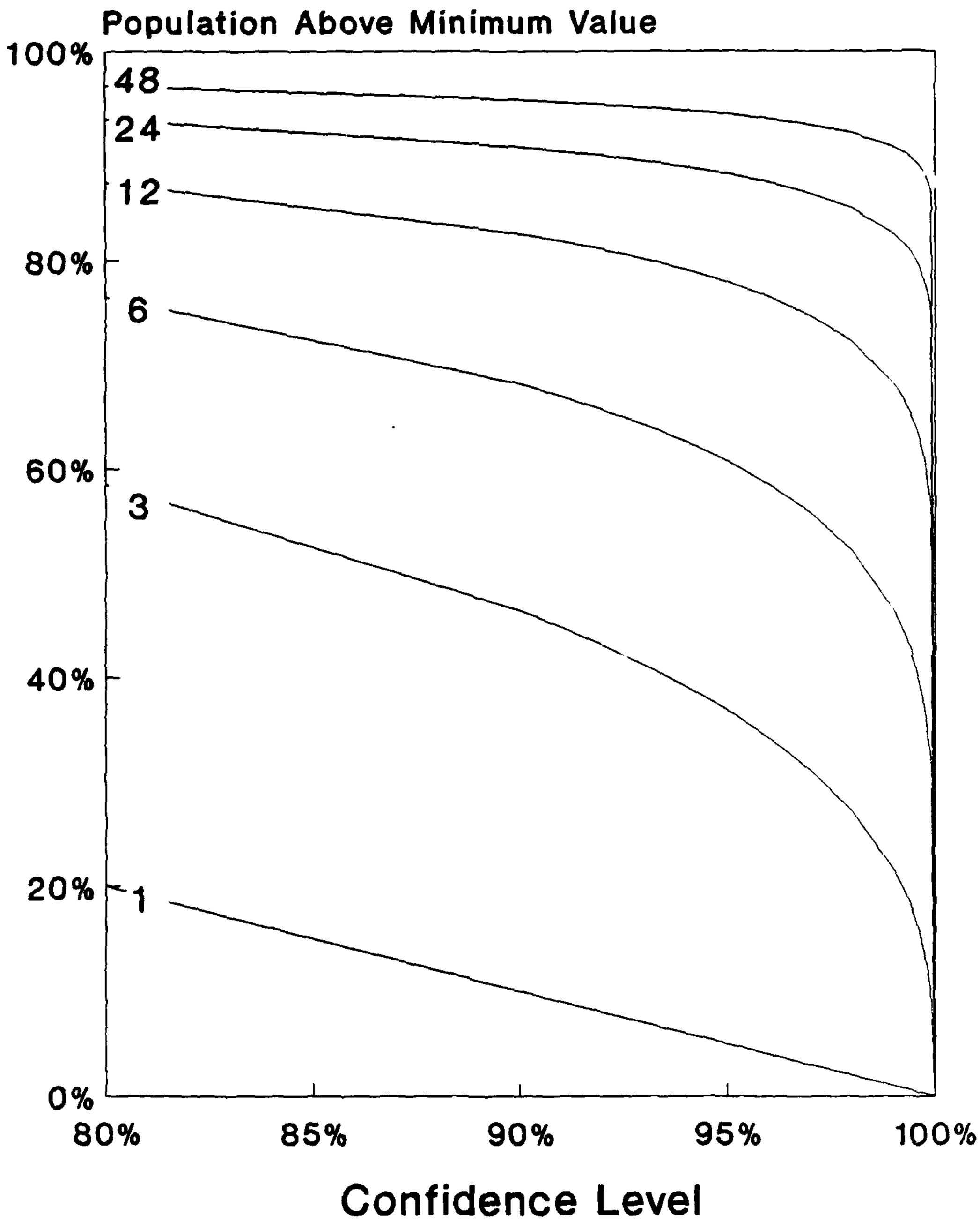


Figure 1: Relation of the minimum of n samples (n shown on figure) to the rest of the population at various confidence levels, after ref. [2].

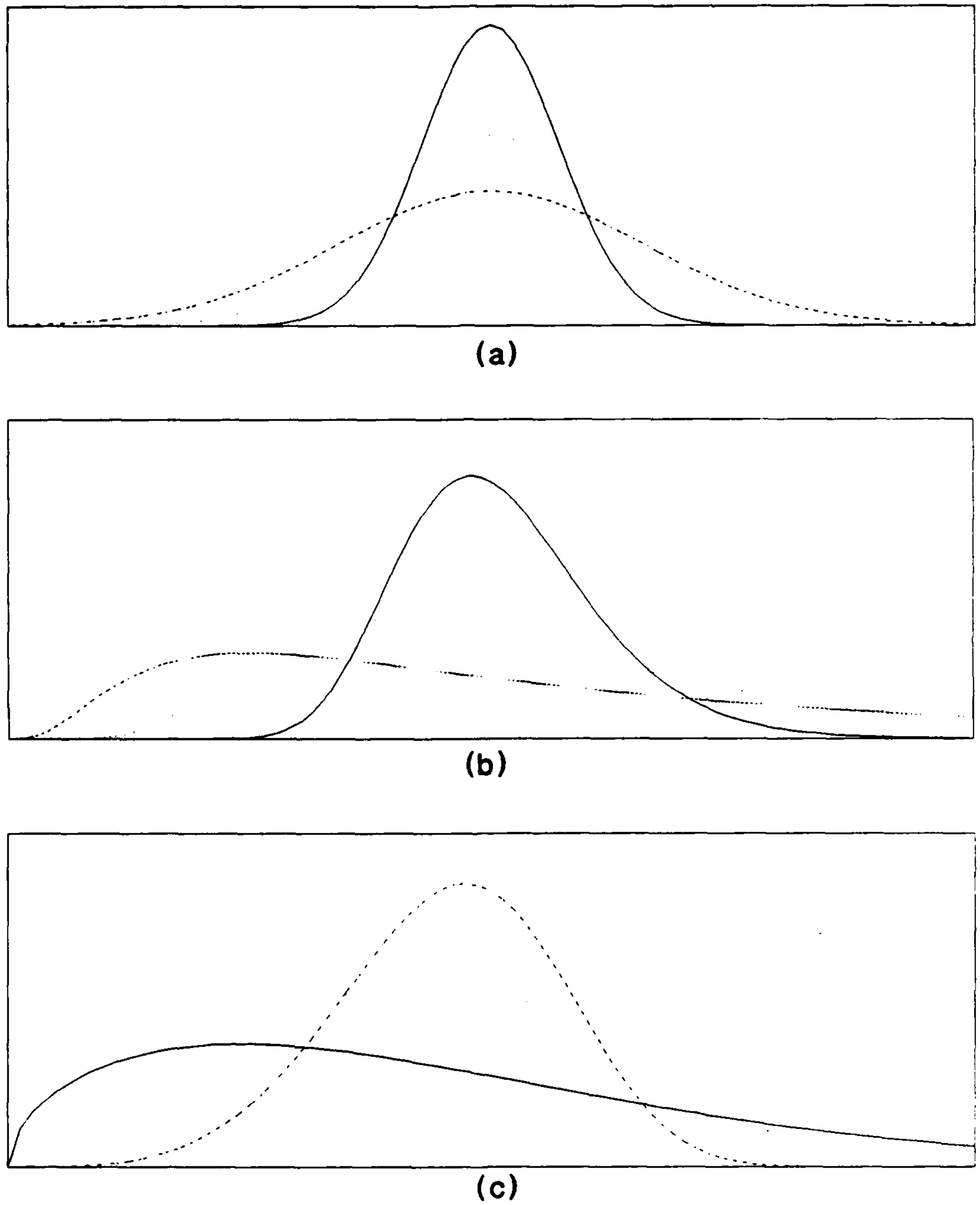


Figure 2: Probability distribution functions drawn from (a) Normal, (b) Log-Normal, and (c) Weibull statistical models. The three curves on each graph show the different shapes each model can produce.

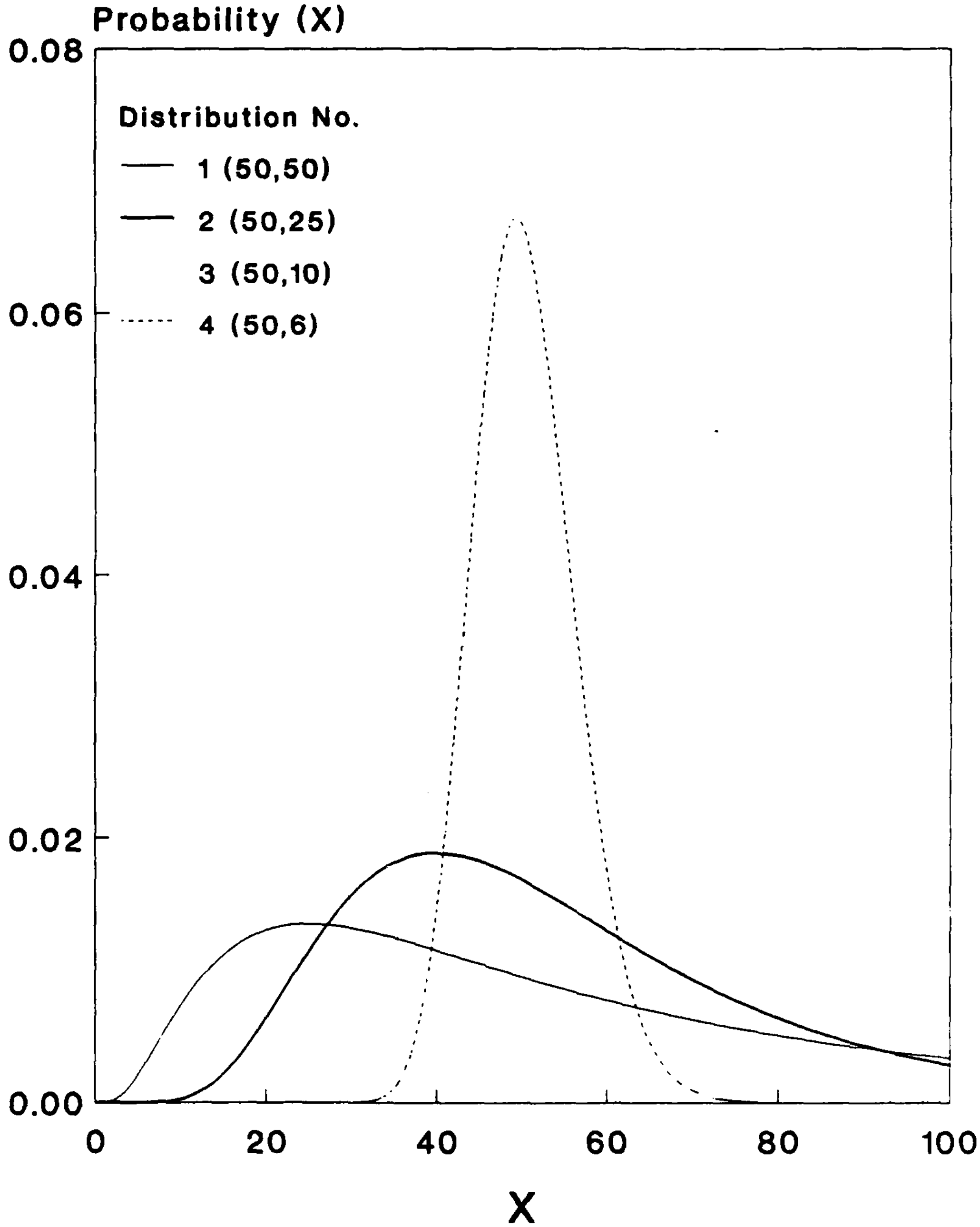


Figure 3: Probability distributions sampled in this study; the numbers in parenthesis are the distribution median and standard deviation, respectively.

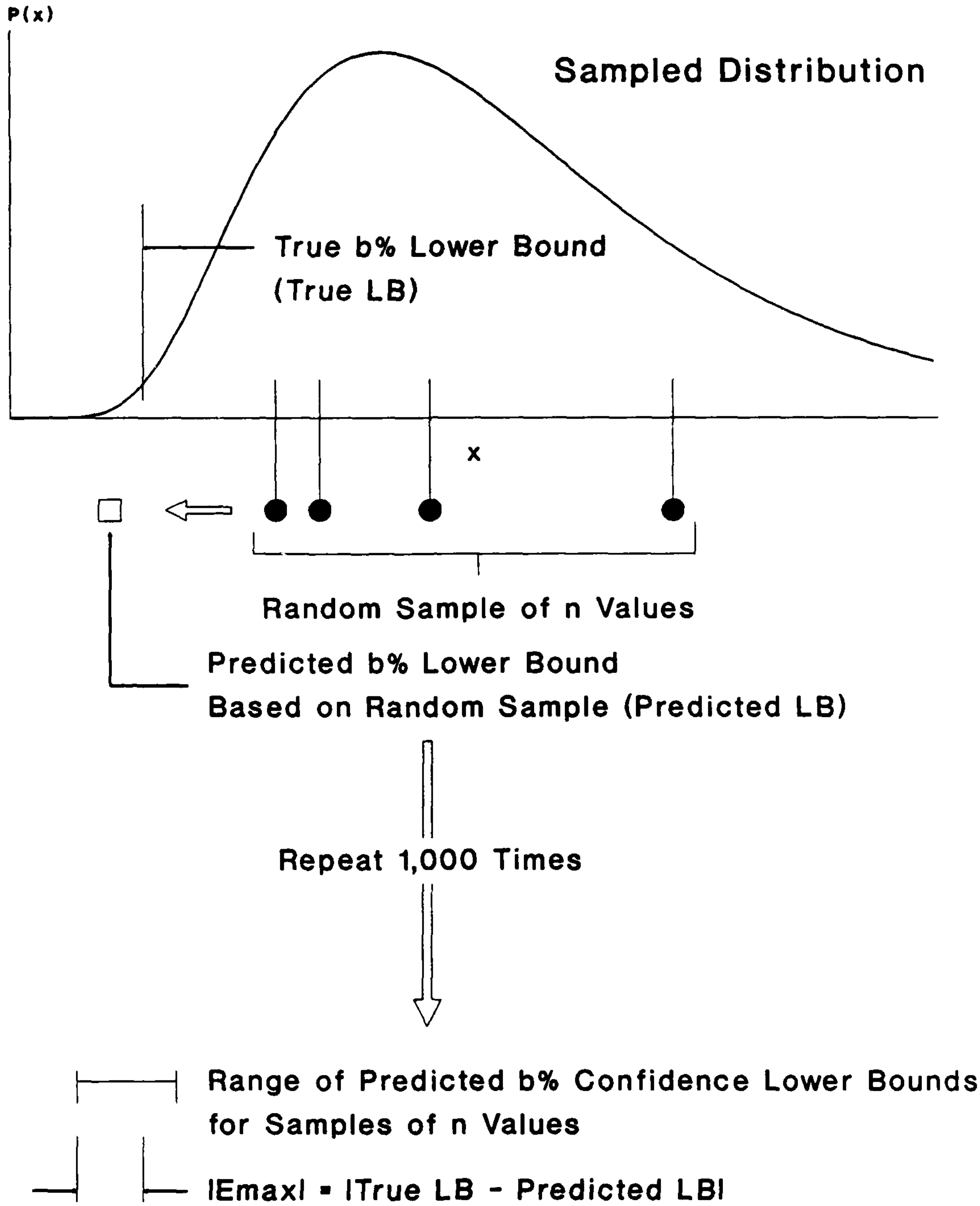
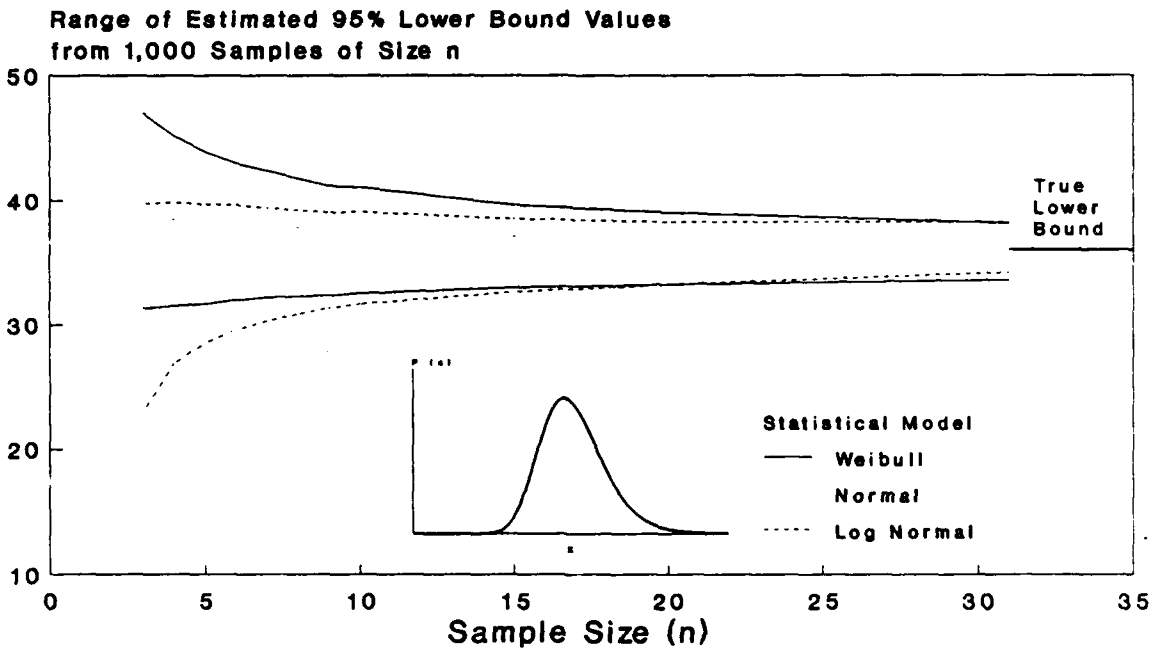
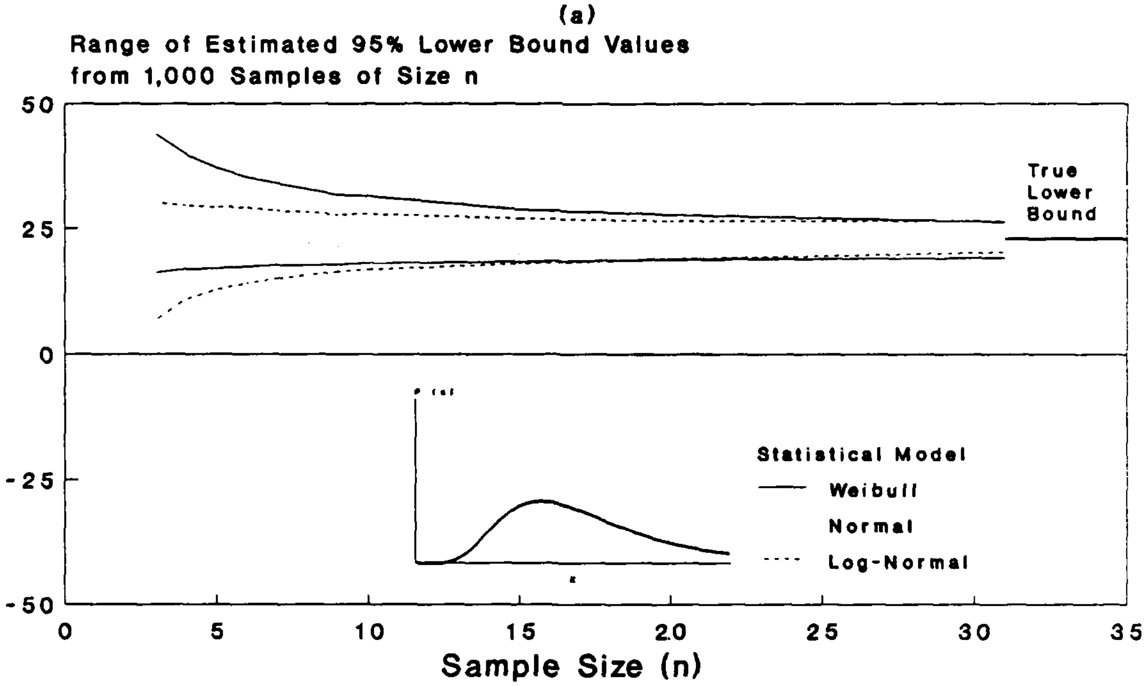


Figure 4: Schematic of Monte Carlo simulation process for determining maximum lower bound estimation error.

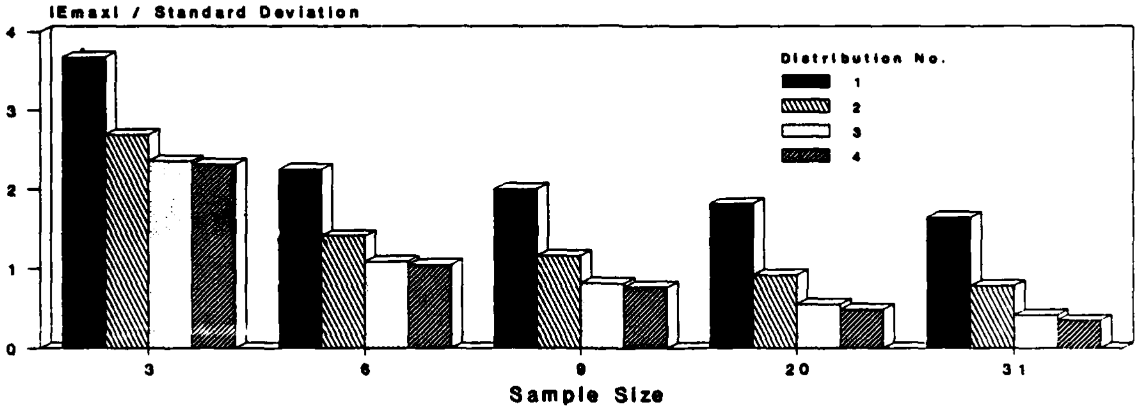


(a)

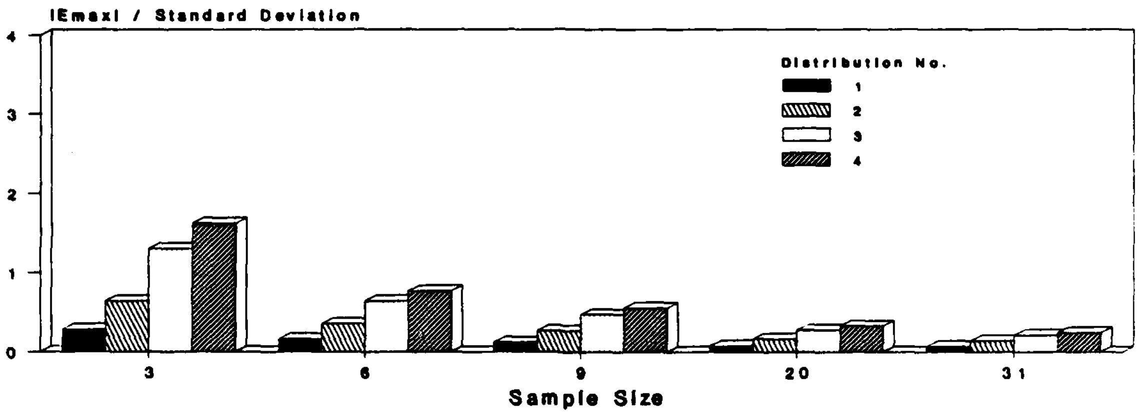


(b)

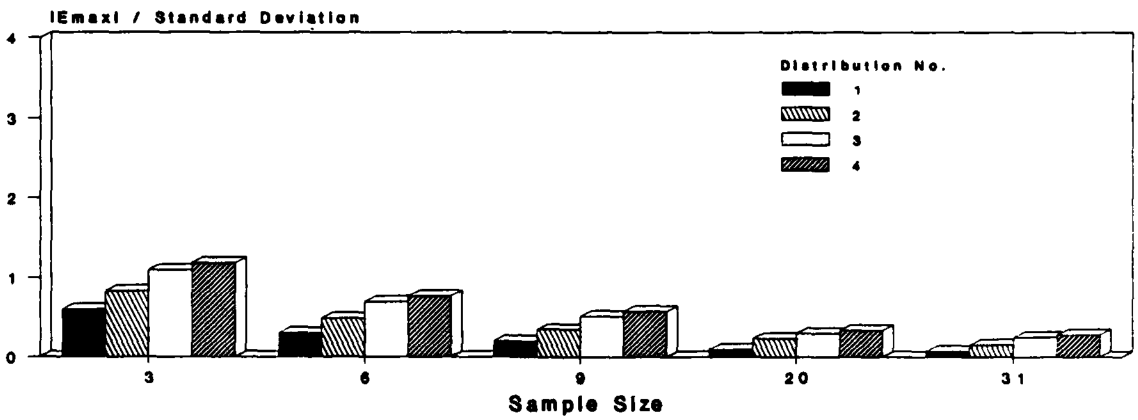
Figure 5: Results of Monte Carlo analysis for (a) an approximately symmetric distribution, and for (b) a distribution that is skewed left. The sampled distributions are shown on each figure.



(a)

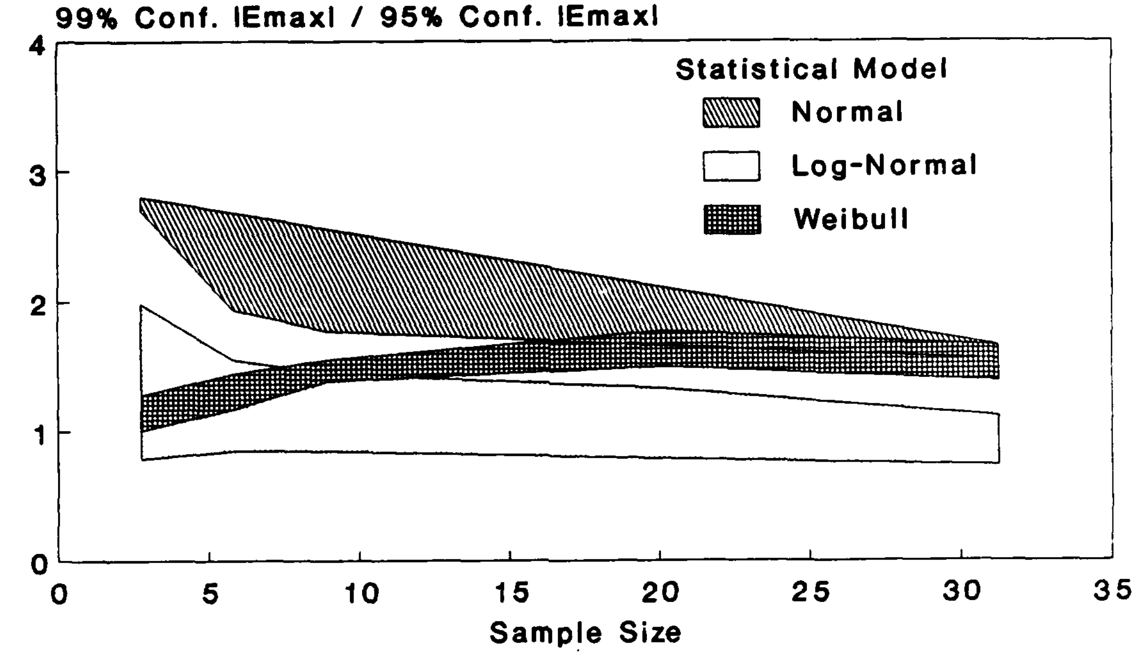


(b)

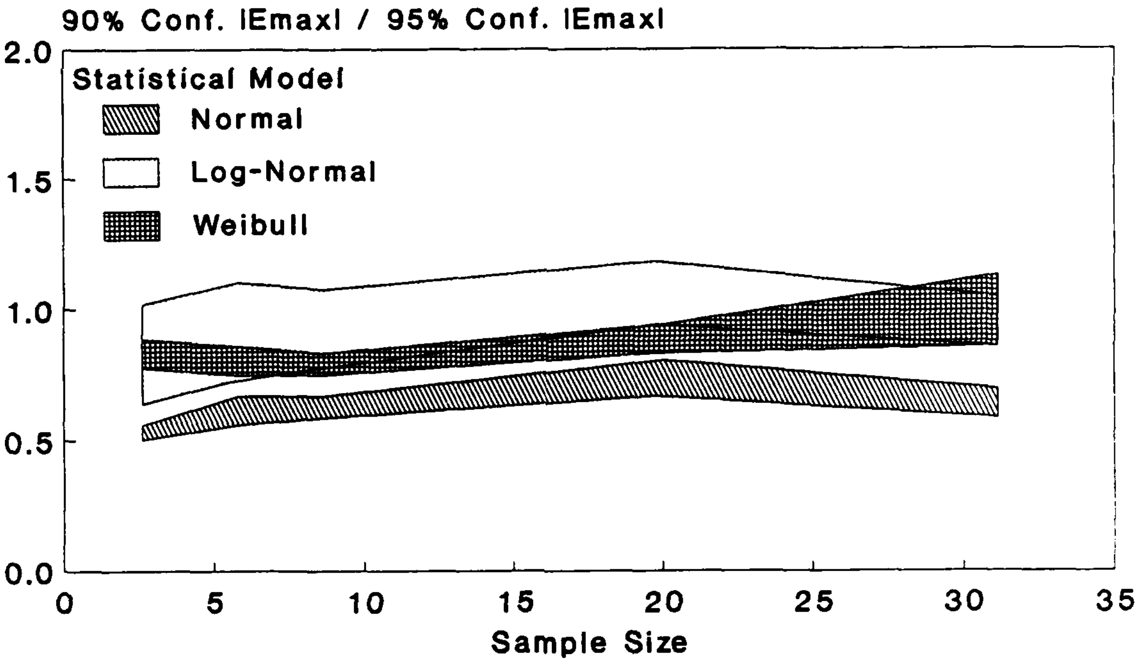


(c)

Figure 6: Maximum error of (a) Normal, (b) Log-Normal, and (c) Weibull estimated 95% confidence lower bounds as determined by Monte Carlo simulated sampling of the probability distributions shown in Figure 3.



(a)



(b)

Figure 7: Comparison of the maximum estimation errors for (a) 99%, and (b) 90% confidence lower bound estimates to that of 95% confidence lower bound estimates for various statistical models and sample sizes.

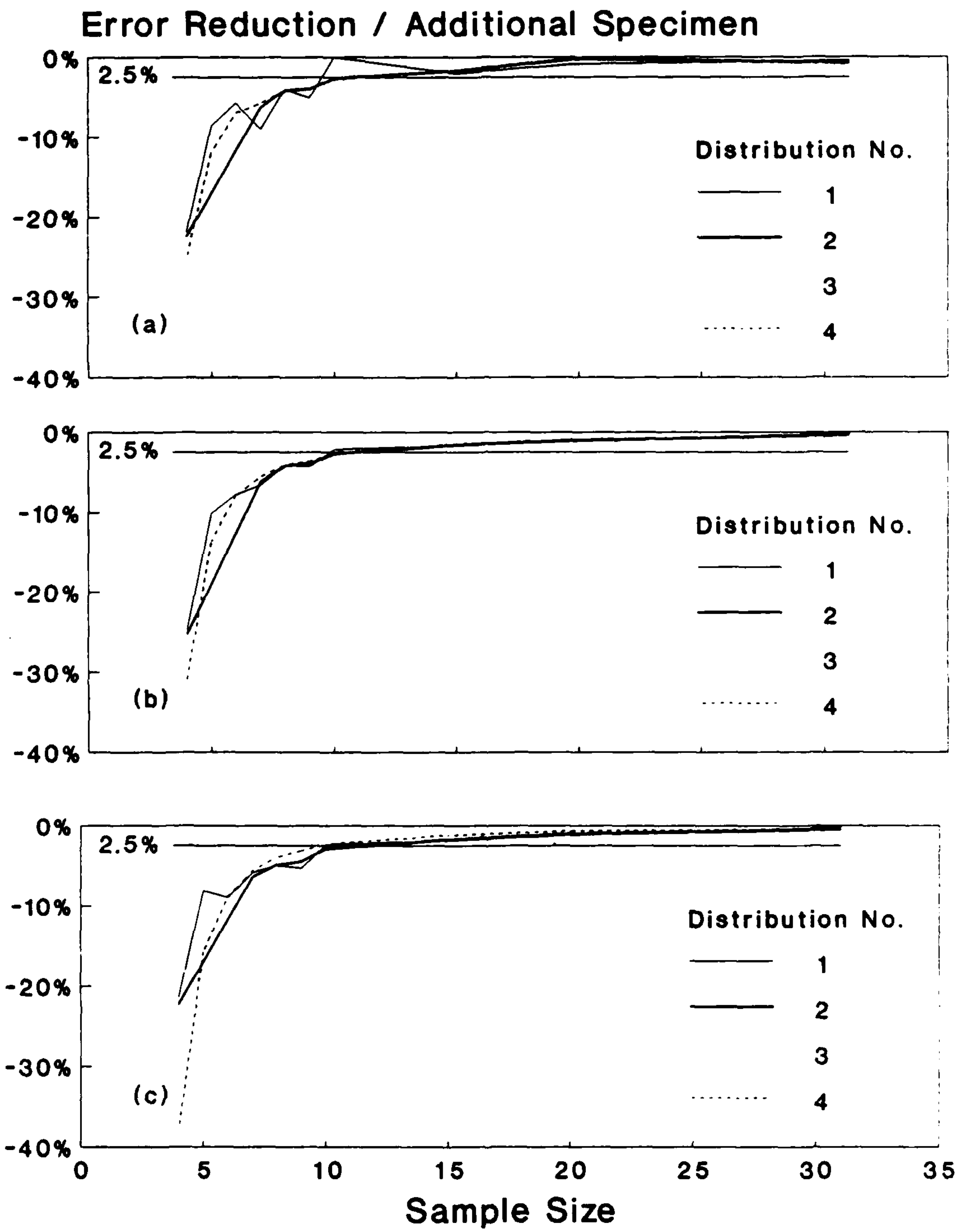


Figure 8: Reduction of the error in Log-Normal estimated lower bounds having 90%, (b) 95%, and (c) 99% confidence with increasing sample size. The sampled distributions are shown in Figure 3.

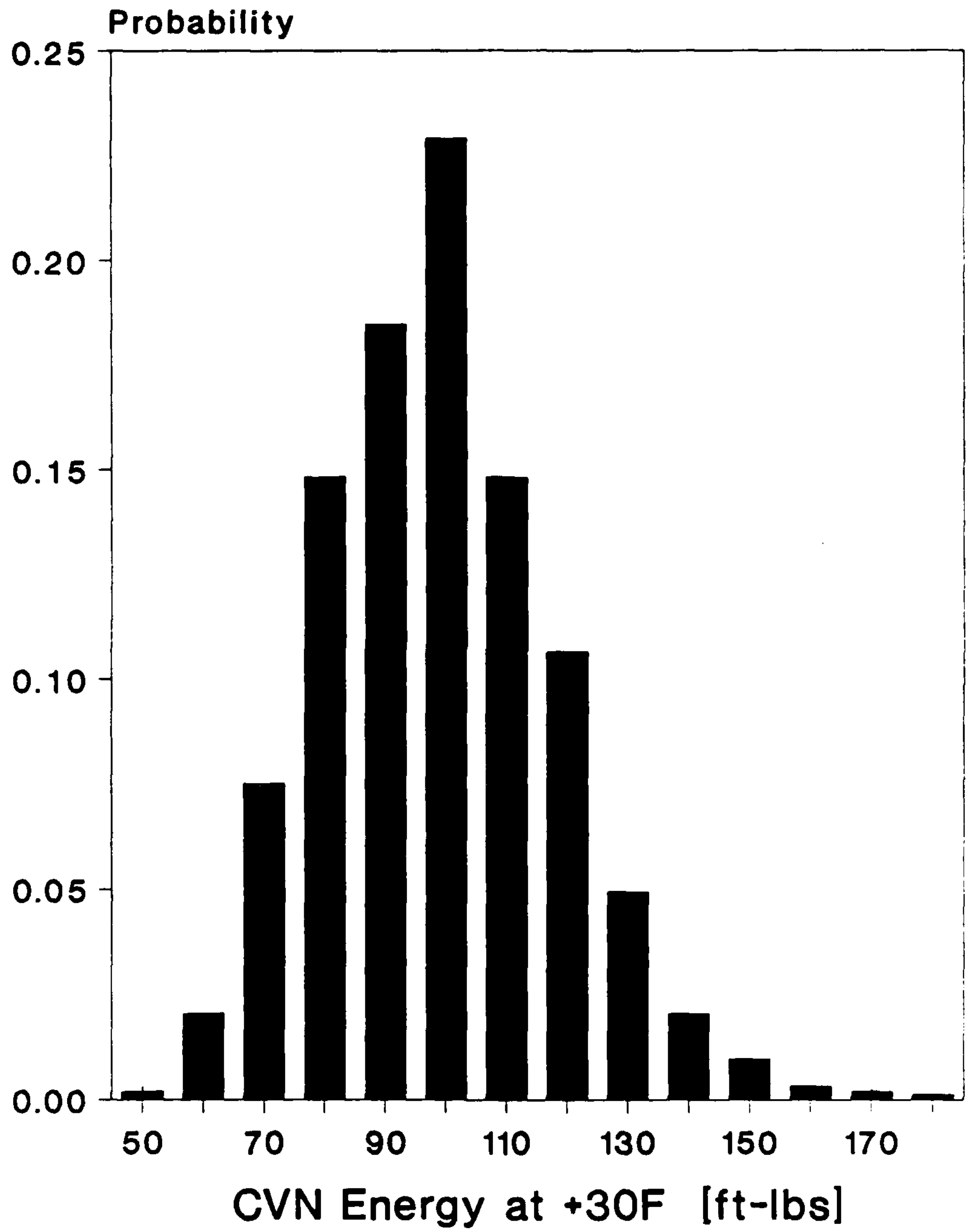


Figure 9: Histogram based on 1559 CVN tests of a high strength steel conducted at +30°F. All fracture surfaces exhibited 100% microvoid coalescence.

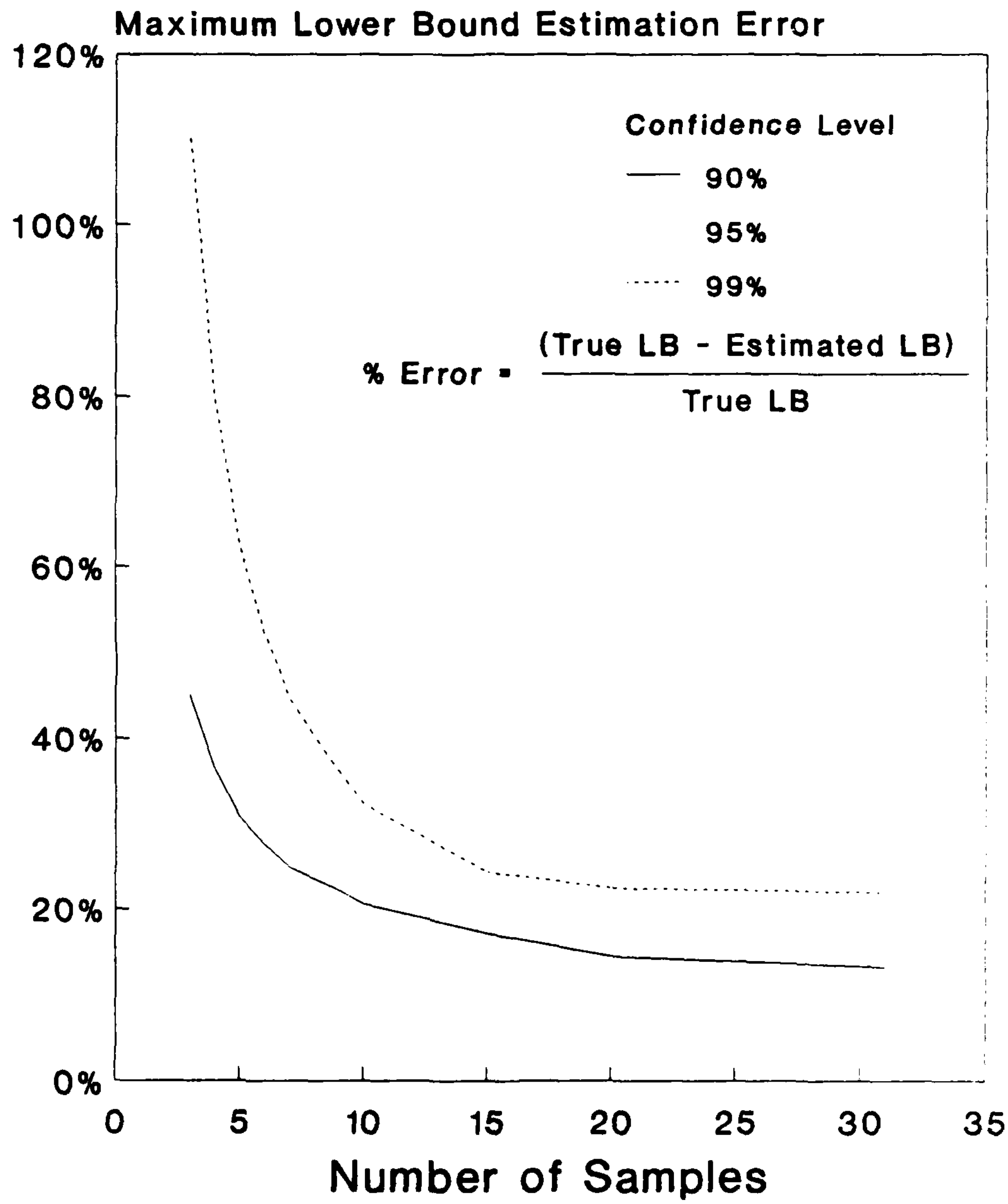


Figure 10: Reduction of Log-Normal lower bound estimation error with increased sample size for the population shown in Figure 9. Percentages on the graph indicate the confidence level associated with the lower bound estimate.

REFERENCES

[1] Military Specification for Steel Plate, Sheet, or Coil, Age-Hardening Alloy, Structural, High Yield Strength (MIL-S-24645(SH)), 4 September 1984.

[2] Jutla T., and Garwood, S.J., "Interpretation of Fracture Toughness Data," *Metal Construction*, pp. 276R-281R, May 1987.

[3] Doig, P., "Evaluation of Lower Bound Fracture Toughness Values using a Weibull Analysis of Single Specimen Data," *Engineering Fracture Mechanics*, Vol. 21, No. 5, pp 963-967, 1985.

[4] Kleijnen, P.P.C., Statistical Techniques in Simulation, Marcel Dekker, Inc., 1974.

[5] Walpole, R.E., and Myers, R.H., Probability and Statistics for Engineers and Scientists, Macmillan Publishing Co., Inc., 1978.

[6] Ang, A.H.S., and Tang, W.H. Probability Concepts in Engineering Planning and Design, Volume I - Basic Principals, John Wiley & Sons, 1975.

INITIAL DISTRIBUTION

OUTSIDE CENTER

8 NAVSEA

- 1 SEA 05M
- 1 SEA 05M2
- 1 SEA 55Y2
- 1 SEA 55Y23
- 2 SEA 55Y3
- 1 SEA 55Y31
- 1 SEA 55W3

1 ONT

- 1 ONT 225

12 DTIC

CENTER DISTRIBUTION

1	011.5	Caplan
1	172	Rockwell
1	173	Beach
1	174	Hansen
1	28	Wacker
1	2801	Crisci
1	2803	Hardy
1	2809	Malec
1	281	Gudas
1	283	Singerman
1	284	Fischer
1	2812	Arora
1	2813	Ferrara
4	2814	Montemarano
20	2814	Kirk
1	2815	Holtzberg
1	2815	DeNale
1	2815	DeLoach
1	522.1	TIC
1	5231	Office Services