CHAPTER 13

STATISTICAL ASPECTS OF FATIGUE

Scatter in fatigue testing or in component fatigue life is a very important consideration in using test data. A variety of factors contribute to scatter. These include inherent material variability (i.e., variations in chemical composition, impurity levels, and discontinuities), variations in heat treatment and manufacturing (i.e., surface finish and hardness variations), variations in specimen or component geometry (i.e., differences in notch radii and weld geometry), and variability from differences in the testing conditions (i.e., environment and test machine alignment variations). In addition, there are sources of uncertainty arising from measured or applied load history variations, as well as from the analytical methods used, such as using the Palmgren-Miner linear damage rule to assess damage for a variable amplitude load history. These variations and uncertainties can result in significant fatigue life variations in the specimen, component, or machine.

Statistical analysis can be used to describe and analyze fatigue properties, as well as to estimate the probability associated with fatigue failure or product life. Such analysis allows us to evaluate component or product reliability quantitatively and to predict service performance for a given margin of safety. Statistical analysis can also be used for the design of experiments such that confounding of sources of variability is avoided and the number of specimen or component tests required for a given reliability and confidence level can be determined.

This chapter first presents basic definitions and concepts in statistical analysis and then discusses commonly used probability distributions for fatigue analysis. Brief discussions of tolerance limits, regression analysis, and reliability analysis are also included. When fatigue life variations are described, the

analysis and discussion presented apply equally to fatigue crack nucleation life, fatigue crack growth life, and total life. For more expanded discussion of these and other related topics, books on probability and statistics such as [1], and books on applications of probability and statistics to fatigue analysis such as [2–5], are recommended.

13.1 DEFINITIONS AND QUANTIFICATION OF DATA SCATTER

A quantity such as fatigue strength or fatigue life that has a statistical variation is called a "stochastic variable," x. Characteristics of such a variable for a population are usually obtained from a small part of the population, called a "sample." The mean or average for a sample size n is defined by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{13.1}$$

The mean value gives a measure of the central value of the sample. Another measure of the central value is the median, which is the middle value in an ordered array of the variable x in the sample. If the sample contains an even number of values, the median is the mean value of the two middle values.

The spread or dispersion of the values in the sample is measured by the sample standard deviation, given by

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{13.2}$$

Standard deviation gives a measure of the magnitude of the variation. Other measures of dispersion in the sample are the variance, given by the square of the standard deviation, S^2, and the coefficient of variation, given by

$$C = \frac{S}{\bar{x}} \tag{13.3}$$

The coefficient of variation normalizes the standard deviation with the mean and is therefore a dimensionless quantity, often given as a percentage. The smaller the standard deviation or the smaller the coefficient of variation, the smaller the variability of the sample values (i.e., x values) and the closer the values in the sample are to the sample mean.

13.2 PROBABILITY DISTRIBUTIONS

Variation of the variable x in its range can be described quantitatively by a probability function, $f(x)$. The probability that the variable x assumes any particular value in the range can be specified by this function. This function is also called the "probability density function" or the "frequency function." The probability that the variable x is less than or equal to a particular value in its range of values is given by the "cumulative probability distribution function." This function usually has a sigmoidal shape. The important statistical distributions often used in fatigue and durability analysis are the normal, log-normal, and Weibull distributions. These will now be discussed.

13.2.1 Normal and Log-Normal Distributions

If variations of the variable x are symmetric with respect to the mean, the probability density function, $f(x)$, is represented by a normal or Gaussian distribution expressed as

$$f(x) = \frac{1}{S\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{S}\right)^2} \tag{13.4}$$

The normal distribution is described in terms of the mean, $\bar{x}$, and standard deviation, S, and has a bell-shaped curve, as shown in Fig. 13.1*a*. The area under the curve in Fig. 13.1*a* is unity. If a set of data conforms to a normal distribution, then 68.3 percent of the data fall within $\pm S$ from the mean, 95.5 percent of the data fall within $\pm 2S$ from the mean, and 99.7 percent of the data fall within $\pm 3S$ from the mean. The cumulative frequency function for this distribution is shown in Fig. 13.1*b* and is given by

$$F(x) = \int f(x)\, dx \tag{13.5}$$

Sinclair and Dolan [6] performed an extensive statistical fatigue study involving 174 so-called identical highly polished, unnotched 7075-T6 aluminum alloy specimens. They used six different alternating stress levels under fully reversed ($S_m = 0$) conditions. Figure 13.2 shows histograms of the fatigue life distribution of 57 specimens tested at a 207 MPa (30 ksi) stress level. A stepped histogram can be replaced by a frequency distribution curve (dashed lines in Fig. 13.2). When normalized to the unit area, this becomes the probability density function. The curve on the left, Fig. 13.2*a*, which is based on cycles to failure, is skewed and hence is not a normal or Gaussian distribution. The curve on the right, Fig. 13.2*b*, is based on the logarithm of cycles to failure and reasonably approximates a normal or Gaussian distribution. This is called a "log-normal distribution," and its probability density function is the same as that in Eq. 13.4, with $x = \log N_f$. Figure 13.3 includes all 174 tests plotted on log-normal probability paper for each value of S_a. If the data are truly log-normal, the data points for each value of S_a will be on a straight line. As seen, a log-normal distribution for the actual data region at each stress level appears reasonable. Based on these and many other statistical test results, a log-normal distribution of fatigue life is often assumed in fatigue design. From the probability distribution functions for each stress level, a family of S–N curves at different probabilities of failure can then be constructed. Such a plot for the fatigue data in Fig. 13.3 is shown in Fig. 13.4.

An analysis of Fig. 13.3 reveals that less scatter occurred at the higher stress levels, as indicated by the steeper slopes. At the highest stress level the life varied from about 1.5×10^4 to 2×10^4 cycles, or a factor of less than 2. At the lowest stress level the life varied from about 2×10^6 to 7×10^7 cycles, or a factor of about 35. Factors of 100 in life are not uncommon for very low stress level fatigue tests. Scatter is usually greater in unnotched, polished, specimens than in notched or cracked specimens. As discussed previously, scatter can be attributed to testing techniques, specimen preparation, and variations in the material. The greater scatter at low stress levels in these

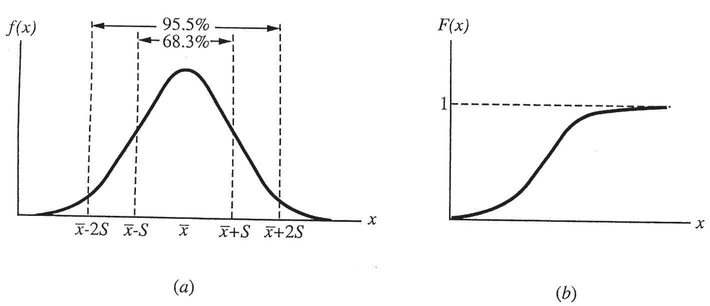

Figure 13.1 Plots of normal distribution. (*a*) Probability density function. (*b*) Cumulative distribution function.

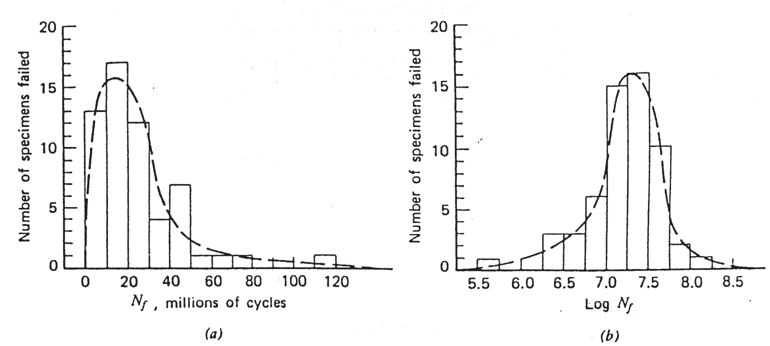

Figure 13.2 Histograms showing fatigue life distributions for 57 specimens of 7075-T6 aluminum alloy tested at 207 MPa (30 ksi) [6] (reprinted with permission of the American Society of Mechanical Engineers).

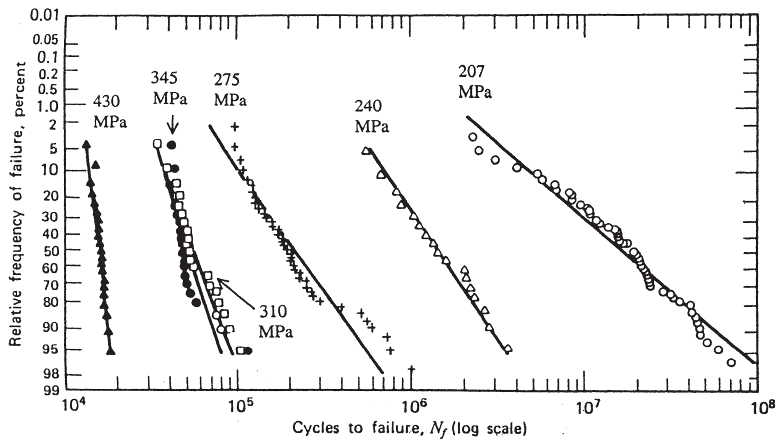

Figure 13.3 Log-normal probability plot at different stress levels for 7075-T6 aluminum alloy [6] (reprinted with permission of the American Society of Mechanical Engineers).

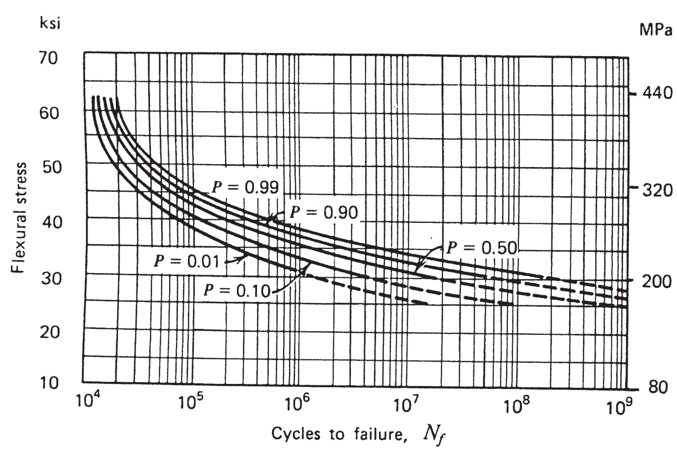

Figure 13.4 S–N curves for different probabilities of failure for specimens of 7075-T6 aluminum alloy [6] (reprinted with permission of the American Society of Mechanical Engineers).

smooth, unnotched specimens can be attributed to the greater percentage of life needed to nucleate small microcracks and then macrocracks. At higher stress levels a greater percentage of the fatigue life involves growth of macrocracks. Tests involving only fatigue crack growth under constant amplitude conditions usually show scatter factors of 2 or 3 or less for identical tests. Thus, the greatest scatter in fatigue involves the nucleation of microcracks and small macrocracks. In notched specimens and components, cracks form more quickly, and subsequently a greater proportion of the total fatigue life involves crack growth that has less scatter.

13.2.2 Weibull Distributions

Weibull distributions are often used in preference to the log-normal distribution to analyze probability aspects of fatigue results. Weibull developed this engineering approach [7] and applied it to the analysis of fatigue test results [8]. Both two- and three-parameter Weibull distribution functions exist, but the two-parameter function is most frequently used in fatigue design and testing. It assumes that the minimum life, N_{fo}, of a population is zero, while the three-parameter function defines a finite minimum life other than zero. The three-parameter Weibull model is

$$F(N_f) = 1 - e^{-\left(\frac{N_f - N_{fo}}{\theta - N_{fo}}\right)^b} \tag{13.6}$$

where

$F(N_f)$ = fraction failed in time or cycles, N_f
N_{fo} = minimum time or cycles to failure
θ = characteristic life (time or cycles when 63.2 percent have failed)
b = Weibull slope or shape parameter

The terms N_{fo}, θ, and b are three Weibull parameters. The two-parameter Weibull model has $N_{fo} = 0$ and hence

$$F(N_f) = 1 - e^{-\left(\frac{N_f}{\theta}\right)^b} \tag{13.7}$$

The slope, b, gives a measure of the shape or skewness of the distribution. Two-parameter Weibull distributions for several values of b are shown in Fig. 13.5 [9]. For b between 3.3 and 3.5, the Weibull distribution function is approximately normal or Gaussian, and for $b = 1$ it is exponential. The coefficient of variation (standard deviation/mean) is approximately $C \approx 1/b$ for the two-parameter Weibull distribution. For values of b typical of fatigue, i.e., between 3 and 6, the error from this approximation is about 10 to 15 percent. It should be noted that here we have used $x = N_f$, since Weibull

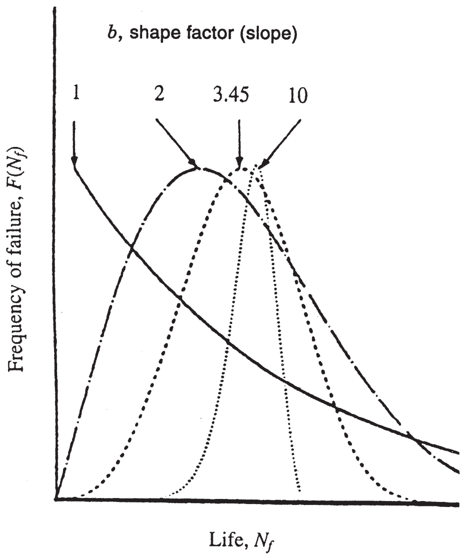

Figure 13.5 Two-parameter Weibull distribution for different values of shape parameter b [9].

distributions are often used for statistical treatment of fatigue life. If a quantity other than fatigue life is being considered, x can be used as the variable in Eqs. 13.6 and 13.7, similar to Eqs. 13.4 and 13.5 for normal or log-normal distributions.

On Weibull probability paper, percent failed is plotted against time or cycles to failure. In general, to plot the data on Weibull probability paper, the array of n data points must first be ordered or ranked from smallest to largest (i.e., $i = 1, 2, 3, \ldots, n$). Next, a plotting position in terms of percent failure is determined for each data point. This can be done by dividing each data point rank by $n + 1$, $i/(n + 1)$. An alternative plotting position is recommended in [10], which is independent of the shape of the distribution and is given by $(i - 0.3)/(n + 0.4)$. This plotting position reduces the bias in estimation of the standard deviation from small samples. Once a plotting position is determined, each data point is plotted at its proper position on the probability paper. This procedure is illustrated by the example problem in Section 13.6 for a two-parameter Weibull probability distribution.

Figure 13.6 is a three-parameter Weibull distribution plot of 1814 fatigue tests of mild steel thin sheet specimens tested at one constant amplitude condition [11]. Data shown represent only partial data from the extensive test program. Several points have been labeled for better clarity, indicating the number of specimens that failed at that life. The minimum life, N_{fo}, of 80 987

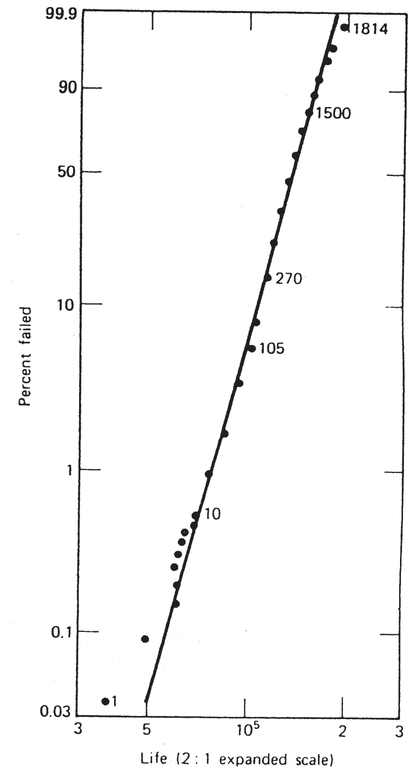

Figure 13.6 Three-parameter Weibull plot for 1814 specimens of mild steel [11].

cycles must be added to the lives in Fig. 13.6 to obtain the total life. On Weibull paper, most data can be plotted as straight lines by judicious choice of N_{fo}. However, to use a three-parameter Weibull distribution, we must be certain that a non-zero minimum life exists below which fatigue failure is unlikely and that all of the data are from a homogeneous sample (i.e., with the same failure mode, the same testing system, the same environment, etc.). Also, as with all distribution function models, the tails at either end may deviate from the model. This is seen in Fig. 13.6. Additional numerical data on scatter in fatigue and its treatment by statistical methods are given in [12].

13.2.3 Estimating Low Probabilities of Failure

The data in Fig. 13.3 for 7075-T6 aluminum are quite comprehensive, yet they do not provide adequate engineering information for fatigue life decisions based on real-life desired probability of failure. Data do not extend past 2 percent probability of failure. Engineers are most interested in 0.01 percent probability of failure or less. This means that extrapolation is needed, for which no accurate method or mathematical justification exists. Only engineering judgment justifies extrapolation. For instance, in Fig. 13.3, if the six stress level curves were extrapolated to desirable low probabilities of failure, they would intersect, which implies that a higher stress level would give longer fatigue life than a lower stress level, which is unreasonable. Great caution must be used in extrapolating fatigue data to low probabilities of failure.

To be able to estimate low probabilities of failure, Abelkis [13] pooled data from more than 6600 specimens of aluminum alloys tested in 1180 groups or samples. He derived a three-term exponential equation:

$$F(z) = A_1 e^{s_1 z} + A_2 e^{s_2 z} + A_3 e^{s_3 z} \qquad \text{for } z < 0 \tag{13.8}$$

where $z = [\log N_f - \log(\text{median } N_f)]/S$
N_f = number of cycles to failure
S = standard deviation of $\log N_f$
F = probability of failure (F = 50 percent for $z = 0$)

The coefficients and exponents are

$$A_1 = 1.687\sqrt{S} \qquad s_1 = 1.3 + 0.86\sqrt{S}$$
$$A_2 = 0.015 \qquad s_2 = 0.28 + 0.44\sqrt{S}$$
$$A_3 = 0.485 - 1.687\sqrt{S} \qquad s_3 = 1.09 + 2.16\sqrt{S}$$

Near the mean, this distribution is almost the same as the log-normal distribution, but for low probabilities of failure it predicts a much lower life than the log-normal distribution. For example, Abelkis shows test data from 2103 specimens in 375 groups. The standard deviation of $\log N_f$ is $S = 0.175$. Both the log-normal distribution and his distribution show that 16 percent of specimens failed at $z = -1$ or at a life $1/10^{0.175} = 67$ percent of the median life. Table 13.1 shows values that apply for lower fractions of failure. Use of the log-normal distribution would have predicted five times longer life than Abelkis found for the first five failures of 10 000 specimens. The formula and data from Abelkis [13] are quoted not to recommend them as the best distribution function, but as an example of a function that is more realistic than the log-normal function for estimating low probabilities of failure.

13.3 TOLERANCE LIMITS

As discussed in Section 13.1, statistical variations are usually obtained from a random sample, which is often a small part of the population. As a result, the sample parameters such as its mean and variance are not the same as the population mean and variance. Obviously the larger the sample size, the better the quality of the estimate for the population. Tolerance limits are used to better estimate population parameters from sample parameters.

Specifying a confidence level provides a quantitative measure of uncertainty or confidence. It is possible to determine a factor, k, by which a probability of survival, p, with a confidence level, γ, for a sample size, n, can be predicted. The expected value of the variable, x, is then given by

$$x = \bar{x} - k\,S \tag{13.9}$$

where $\bar{x}$ and S are the sample mean and standard deviation, respectively. The factor k is called the "one-sided tolerance limit factor," and its values, based on normal distributions for various sample sizes, confidence levels, and probabilities of survival, are listed in Table 13.2.

TABLE 13.1 Expected Fractions of Median Life for Various Probabilities of Failure

	Expected z and Percent of Median Life			
	From Log-Normal		From Abelkis	
Fraction Failed, Percent	z	Fraction of Med. Life, Percent	z	Fraction of Med. Life, Percent
2	−2	45	−2.2	41
0.2	−2.9	31	−4.7	15
0.05	−3.3	26	−7.3	5

TABLE 13.2 Values of Factor k for One-Sided Tolerance Limits Assuming Normal Distribution [12,14]

	γ = 50%			γ = 90%			γ = 95%		
n	p = 90%	99%	99.9%	90%	99%	99.9%	90%	99%	99.9%
4	1.42	2.60	3.46	3.19	5.44	7.13	4.16	7.04	9.21
6	1.36	2.48	3.30	2.49	4.24	5.56	3.01	5.06	6.61
8	1.34	2.44	3.24	2.22	3.78	4.95	2.58	4.35	5.69
10	1.32	2.41	3.21	2.07	3.53	4.63	2.35	3.98	5.20
20	1.30	2.37	3.14	1.77	3.05	4.01	1.93	3.30	4.32
50				1.56	2.73	3.60	1.65	2.86	3.77
100				1.47	2.60	3.44	1.53	2.68	3.54
500				1.36	2.44	3.24	1.39	2.48	3.28
∞				1.28	2.33	3.09	1.28	2.33	3.09

Example Ultimate tensile strength values obtained from eight specimens of a material were measured to be 616, 669, 649, 600, 658, 629, 684, and 639 MPa. Determine the tensile strength for 99 percent reliability with 95 percent confidence level.

The mean tensile strength is found to be $\bar{x}$ = 643 MPa, and the standard deviation is calculated from Eq. 13.2 to be S = 27.8 MPa. Assuming normal distribution, with p = 99 percent, γ = 95 percent, and n = 8, Table 13.2 gives k = 4.35. The predicted tensile strength is then calculated from Eq. 13.9 to be

$$x = \bar{x} - k\,S = 643 - 4.35\,(27.8) = 522 \text{ MPa}$$

We can also predict the value of variable x to be within the interval $(\bar{x} - k\,S)$ and $(\bar{x} + k\,S)$ with a specific probability and confidence level. In this case, the factor k used is for two-sided tolerance limits, and its values are different from the one-sided tolerance limits listed in Table 13.2. Values of k for two-sided tolerance limits can be found in [14]. Based on these values, tolerance limits for the variable x can be established.

Tolerance limits can also be established for a Weibull distribution. Lower and upper tolerance limits for the two-parameter Weibull distribution are given by

$$\text{Lower limit} = F(N_f) - k \tag{13.10}$$

$$\text{Upper limit} = F(N_f) + k \tag{13.11}$$

where k is a function of the sample size, n, given in Fig. 13.7 for a 90 percent confidence level. Establishing the lower and upper tolerance intervals for the two-parameter Weibull distribution is illustrated by the example problem in Section 13.6.

13.4 REGRESSION ANALYSIS OF FATIGUE DATA

Regression analysis is used to obtain a curve that best fits a set of data points. In fatigue data analysis, a linear or linearized regression is often employed using a least squares fit, where the square of the deviations of the data points from the straight line is minimized. The equation of the straight line can be expressed by

$$y = a + b\,x \tag{13.12}$$

where x is the independent variable and y is the dependent variable. Fitting constants a and b are the regression coefficients, where b is the slope of the line and a is the y intercept. Such linear regression is often used for analysis of S–N, ε–N, and da/dN–ΔK fatigue data and curves. For example, the coefficients A and B in Basquin's log-log straight line S–N relationship (Eq. 4.7) are found by least squares fit of $(\log S_a)$ versus $(\log N_f)$ data. It should be recognized that since the stress amplitude, S_a, is the input or controlled parameter in S–N tests, it is the independent variable. Fatigue life, N_f, is considered to be the response or observed parameter in the test and is therefore the dependent variable. For S–N data, Eq. 13.12 is then written as

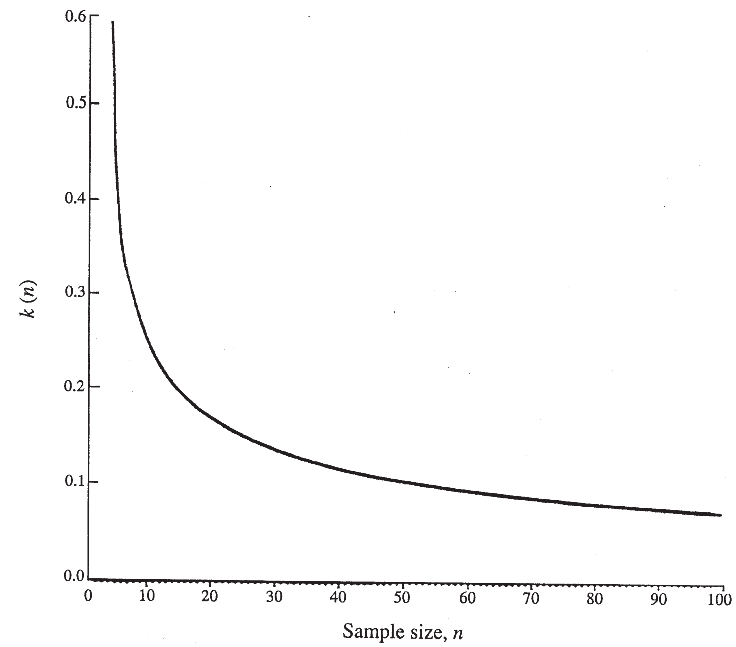

Figure 13.7 Variation of factor k with sample size n for a 90 percent confidence level for the Weibull distribution [9].

$$\log N_f = a + b\,(\log S_a) \tag{13.13}$$

The fitting constants a and b in Eq. 13.13 are related to the coefficient A and exponent B in Basquin's Eq. 4.7 by $a = -(1/B)\log A$ and $b = 1/B$. Similarly, linear regressions are used to obtain the strain–life fatigue properties σ_f' and b, and ε_f' and c in Eqs. 5.15 and 5.16, respectively. Examples of such a regression for $\Delta\sigma/2$ versus $2N_f$ and $\Delta\varepsilon_p/2$ versus $2N_f$ for a 4340 steel are shown in Figs. 5.12*b* and 5.12*c*, respectively. Similar linear regression analysis is used for da/dN versus ΔK data in the Paris equation regime to find the coefficient A and slope n in Eq. 6.19 or other similar equations for fatigue crack growth rates.

The degree of correlation between two variables such as x and y in Eq. 13.12 can be quantified by the correlation coefficient, r. The correlation coefficient has a range of values between -1 and $+1$. A value of $r = 0$ indicates no correlation, whereas $r = \pm 1$ indicates perfect correlation. A negative value of r indicates that the regression line has a negative slope. ASTM Standard E739 provides additional details on statistical analysis of linear or linearized stress–life (S–N) and strain–life (ε–N) fatigue data [15].

13.5 RELIABILITY ANALYSIS

In most cases, the service load spectra for a component or structure are probabilistic rather than deterministic. Therefore, statistical analysis is often required to obtain the stress distribution. In addition, as discussed in previous sections, the fatigue strength of the component or structure has a statistical distribution due to the many sources of variability. Schematic representations of probability density functions for both service loading and fatigue strength are shown in Fig. 13.8. If there were no overlap between the two distributions, failure would not occur. However, there is often an overlap between the two distributions, as shown in Fig. 13.8, indicating the possibility of fatigue failure. The overlap can increase as the strength decreases due to accumulation of damage with continued use of the component or structure. Reliability analysis provides an analytical tool to quantify the probability of failure due to the overlapping of the two distributions.

The reliability or probability of survival, R, is obtained from

$$R = \int_0^\infty F_1(x)\, F_2(x)\, dx \tag{13.14}$$

where $F_1(x)$ and $F_2(x)$ are the cumulative density functions of the service stress and fatigue strength, respectively. A detailed presentation and discussion of reliability analysis is beyond the scope of this book, and the interested reader is referred to books such as [16–18].

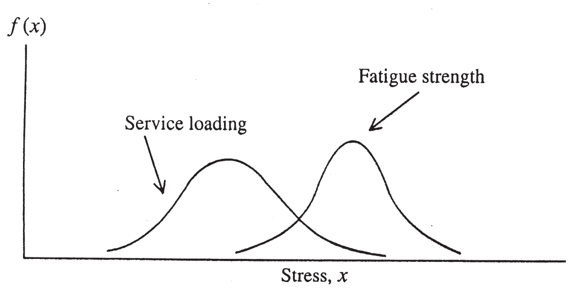

Figure 13.8 Probability distributions for service loading and fatigue strength.

13.6 EXAMPLE PROBLEM USING THE WEIBULL DISTRIBUTION*

A durability test on 10 units resulted in failures at 140, 90, 190, 220, 270, 200, 115, 170, 260, and 330 hours. Assume the life distribution to be a two-parameter Weibull. Determine (a) the slope, b, and characteristic life, θ, for the distribution, (b) the median and B_{10} lives, (c) the percentage of the population that would be expected to fail in 300 hours with 50 percent confidence, and (d) the 90 percent tolerance interval.

The data are first rank ordered in ascending order, and the plotting position for percent failure is determined from $[(i - 0.3)/(n + 0.4)]$ 100 percent, as listed below

Rank, i	Life to Failure (hours)	Plotting Position (%)
1	90	6.7
2	115	16.2
3	140	25.9
4	170	35.5
5	190	45.2
6	200	54.8
7	220	64.5
8	260	74.1
9	270	83.8
10	330	93.3

Life to failure versus percent failure data are then plotted on the Weibull probability plot shown in Fig. 13.9. A straight line is passed through the data.

(a) A line parallel to the line through the data and passing through the index point shown in the plot by $\oplus$ intersects the slope scale on top of the figure and gives the Weibull slope, $b = 2.77$. The characteristic life, θ, is found from the line passing through the data at 63.2 percent failure, $\theta = 225$ hours. The Weibull distribution function is then given as

$$F(N_f) = 1 - e^{-\left(\frac{N_f}{225}\right)^{2.77}}$$

* This example problem is courtesy of Dr. H. S. Reemsnyder and is used in the SAEFDE/University of Iowa short course on Fatigue Concepts in Design. Data for this example were supplied by H. R. Jaeckel.

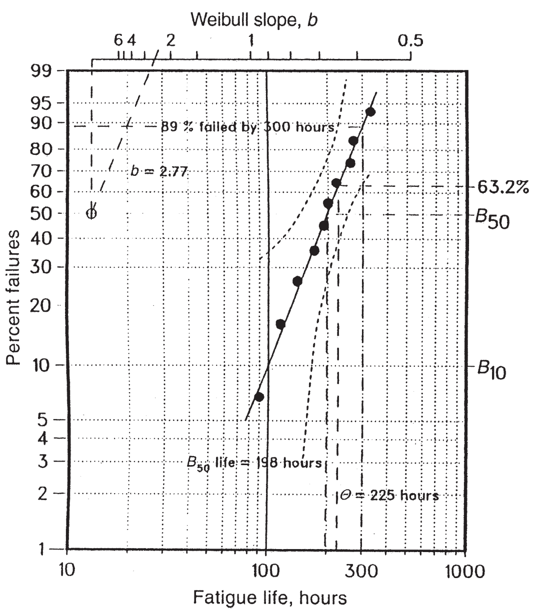

Figure 13.9 Example of two-parameter Weibull plot showing median line and 90 percent tolerance interval [9].

(*b*) The median life is the B_{50} life, which is found from the plot to be $B_{50} = 198$ hours. The B_{10} life is also found from the plot to be $B_{10} =$ 100 hours. This is the life corresponding to a 10 percent failure rate.

(*c*) The line fitted through the data corresponds to a 50 percent confidence level. Therefore, the percentage of the population that would be expected to fail in 300 hours with 50 percent confidence is found from this line for a fatigue life of 300 hours to be 89 percent.

(*d*) The 90 percent tolerance interval for the data is obtained by computing and plotting the lower and upper tolerance limits from Eqs. 13.10 and 13.11, respectively. The value of k for $n = 10$ is found from Fig. 13.7 to be 0.25. Therefore,

$$\text{Lower limit} = F(N_f) - k = 0.75 - e^{-\left(\frac{N_f}{225}\right)^{2.77}}$$

$$\text{Upper limit} = F(N_f) + k = 1.25 - e^{-\left(\frac{N_f}{225}\right)^{2.77}}$$

The lower and upper tolerance limits for various lives are calculated from the above equations, multiplied by 100 to convert them to a percentage, and then plotted, as shown by the dashed curves in Fig. 13.9.

13.7 SUMMARY

Many factors contribute to scatter in fatigue life and properties. Statistical analysis is used to describe scatter quantitatively and to estimate the probability associated with fatigue failure or predicted life. Basic characteristics of data scatter for a sample of data are measured from the sample's mean and standard deviation, and data variation is described by a probability distribution function. Common probability distribution functions in fatigue analysis are the normal, log-normal, and Weibull distributions. The normal distribution assumes symmetric variation about the mean. When $\log(x)$ of a variable x has a normal distribution, the variable x is said to have a log-normal distribution. Distributions of fatigue life data are often skewed, and two- and three-parameter Weibull distributions are most commonly used to analyze probability aspects of fatigue life. At low probabilities of failure the distribution function for fatigue life may differ from the distribution function near the mean. Extrapolations to low probabilities of failure may therefore not be appropriate. Tolerance limits are used to quantify the uncertainty in estimating the population's statistical characteristics from smaller sample statistical parameters. Regression analysis is used to obtain best fit curves to fatigue data, and reliability analysis is used to estimate the probability of failure associated with overlapping of the applied stress and the material or component strength probability distributions.

13.8 DOS AND DON'TS IN DESIGN

1. Don't ignore data scatter in fatigue testing and analysis, as it can result in significant variations in predicted fatigue life.
2. Do use great caution in extrapolating fatigue data to low probabilities of failure, as the probability distribution function for low failure probabilities may be different than for higher failure probabilities.
3. Do use tolerance limit analysis when extrapolating statistical parameters and characteristics from a small sample of data to a much larger population.
4. Do recognize that in regression analysis of fatigue data the fatigue life should be treated as the dependent variable (x), while load, stress, or strain is the independent variable (y).

REFERENCES

1. R. A. Johnson, *Miller and Freund's Probability and Statistics for Engineers,* 5th ed., Prentice-Hall, Upper Saddle River, NJ, 1994.
2. R. R. Little, *Manual on Statistical Planning and Analysis for Fatigue Experiments,* ASTM STP 588, ASTM, West Conshohocken, PA, 1975.
3. L. G. Johnson, *The Statistical Treatment of Fatigue Experiments,* Elsevier Science, New York, 1964.
4. J. L. Bogandoff and F. Kozin, *Probabilistic Models of Cumulative Damage,* John Wiley and Sons, New York, 1985.
5. W. Nelson, *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis,* John Wiley and Sons, New York, 1990.
6. G. M. Sinclair and T. J. Dolan, "Effect of Stress Amplitude on Statistical Variability in Fatigue Life of 75S-T6 Aluminum Alloy," *Trans. ASME,* Vol. 75, 1953, p. 867.
7. W. Weibull, "A Statistical Distribution Function of Wide Applicability," *J. Appl. Mech.,* Vol. 73, 1951, p. 293.
8. W. Weibull, *Fatigue Testing and Analysis of Results,* Pergamon Press, London, 1961.
9. H. S. Reemsnyder, SAEFDE/The University of Iowa Short Course Notes on Fatigue Concepts in Design, 1999.
10. A. Bernard and E. D. Bos-Levenbach, "The Plotting of Observations on Probability Paper," *Statistica,* Vol. 7, 1953, p. 163.
11. W. E. Hering and C. W. Gadd, "Experimental Study of Fatigue Life Variability," Research Publication, GMR-555, General Motors Corp., 1966.
12. *A Guide for Fatigue Testing and the Statistical Analysis of Fatigue Data,* 2nd ed., ASTM STP 91-A, ASTM, West Conshohocken, PA, 1963.
13. P. R. Abelkis, "Fatigue Strength Design and Analysis of Aircraft Structures, Part I: Scatter Factors and Design Charts," AFFDL-TR 66-197, June 1967.
14. R. E. Odeh, D. B. Owen, Z. W. Birnbaum, and L. Fisher, *Pocket Book of Statistical Tables,* Marcel Dekker, New York, 1977.
15. *Standard Practice for Statistical Analysis of Linear or Linearized Stress–Life* (S–N) *and Strain–Life* (ε–N) *Fatigue Data,* ASTM Standard E 739, Vol. 3.01, ASTM, West Conshohocken, PA, 2000, p. 631.
16. N. R. Mann, N. D. Singpurwalla, and R. E. Schafer, *Methods for Statistical Analysis of Reliability and Life Data,* John Wiley and Sons, New York, 1974.
17. W. Nelson, *Applied Life Data Analysis,* John Wiley and Sons, New York, 1982.
18. B. S. Dhillon, *Design Reliability: Fundamentals and Applications,* CRC Press, Boca Raton, FL, 1999.

PROBLEMS

1. The following data from 43 specimens were obtained from rotating bending fatigue tests at one constant stress level and were grouped as follows:

Samples failed, n	3	8	6	10	8	5	3
Cycles to failure, N_f	1600	1900	2200	2500	2800	3100	3400

(*a*) Calculate the mean, median, standard deviation, and coefficient of variation for the fatigue lives. (*b*) Show the data in histograms of fatigue life distribution using both N_f and log N_f. (*c*) Fit the stepped histograms in part (*b*) by frequency distribution curves, assuming normal and log-normal distributions. (*d*) Plot the cumulative histograms and distribution functions. (*e*) Determine B_{10} and B_{50} lives for these data, assuming normal distribution.

2. Fatigue testing of eight similar components at a given load level has produced the following fatigue lives: 14 700, 94 700, 31 700, 27 400, 17 300, 70 100, 21 800, and 39 800. (*a*) Compute the mean life and standard deviation for this load level. (*b*) Plot the data on normal probability paper, log-normal probability paper, and Weibull probability paper. Which probability distribution do the data fit best? (*c*) Obtain the B_{50} life from each distribution in part (*b*).

3. A fatigue life test program on a new product consisted of 10 units subjected to the same load spectrum. The units failed in the following number of hours in ascending order: 75, 100, 130, 150, 185, 200, 210, 240, 265, and 300. Obtain the percent median rank values for the 10 failures. Plot the percent median rank values versus hours to failure on Weibull paper. Draw the best fit (usually a straight line) curve through the data and find (*a*) 10 percent (B_{10}) expected failure life and (*b*) 50 percent (B_{50}) expected failure life. (*c*) What percentage of the population will have failed in 230 hours? (*d*) Consider what the 1.0 or 0.1 percent expected life would be.

4. Plane strain fracture toughness tests of six identical samples from a metal produced the following values of K_{Ic}: 57.2, 54.0, 59.7, 55.5, 62.4, and 58.6 MPa$\sqrt{m}$. Assuming a normal distribution, estimate the plane strain fracture toughness of the metal with 90 percent reliability and a 90 percent confidence level.

5. Fatigue strength, S_f, at 10^6 cycles versus Brinell hardness (HB) for seven grades of SAE 1141 steel are given as follows:

HB	223	277	199	241	217	252	229
S_f (MPa)	286	433	276	342	287	332	296

(*a*) Using linear regression, find the slope, intercept, and correlation coefficient for these data. (*b*) Is there a good correlation between the hardness and fatigue strength of these steel grades? (*c*) What fatigue strength corresponds to a Brinell hardness of 225?

6. Constant amplitude crack growth rate data in the threshold region for a cast aluminum alloy are listed as a function of stress intensity factor range, ΔK: