

CHAPTER 3

Properties of the Weibull Distribution

3.1 THE WEIBULL CUMULATIVE DISTRIBUTION FUNCTION (CDF), PERCENTILES, MOMENTS, AND HAZARD FUNCTION

There are two forms of the Weibull distribution distinguished by the presence of either two or three parameters. Unlike the normal distribution, the Weibull CDF is expressible in closed form. The CDF of the three parameter version of the Weibull distribution may be written as:

$$F(x) = 1 - \exp\left[-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right];\ x > \gamma. \tag{3.1}$$

γ, the location parameter, is also known as the threshold parameter, or, in life testing applications, as the guarantee time, since failure cannot occur until x exceeds γ. When γ is zero the three-parameter Weibull distribution specializes to the much more widely employed two-parameter version. η, the scale parameter, is known as the characteristic value or, in life testing applications, as the characteristic life, for a reason explained further below. β is called the shape parameter and for $\beta = 1$ the two-parameter Weibull distribution specializes to the exponential distribution. The parameters η and β are positive. γ is generally taken to be positive although there is no mathematical reason that this is necessary. In this chapter we will deal, except where noted, with the two-parameter Weibull model. Other parameterizations of the Weibull are possible. For example, some writers replace $1/\eta^{\beta}$ by λ for the typographical benefit of writing the CDF on a single line. The version we have adopted has the useful advantage that the parameter η is expressed in the same units as the random variable itself.

The reliability function $R(x)$, known more usually as the survivorship function in biomedical applications, expresses the probability that the life of a device or subject will exceed a given value. For the two-parameter Weibull distribution this function has the form:

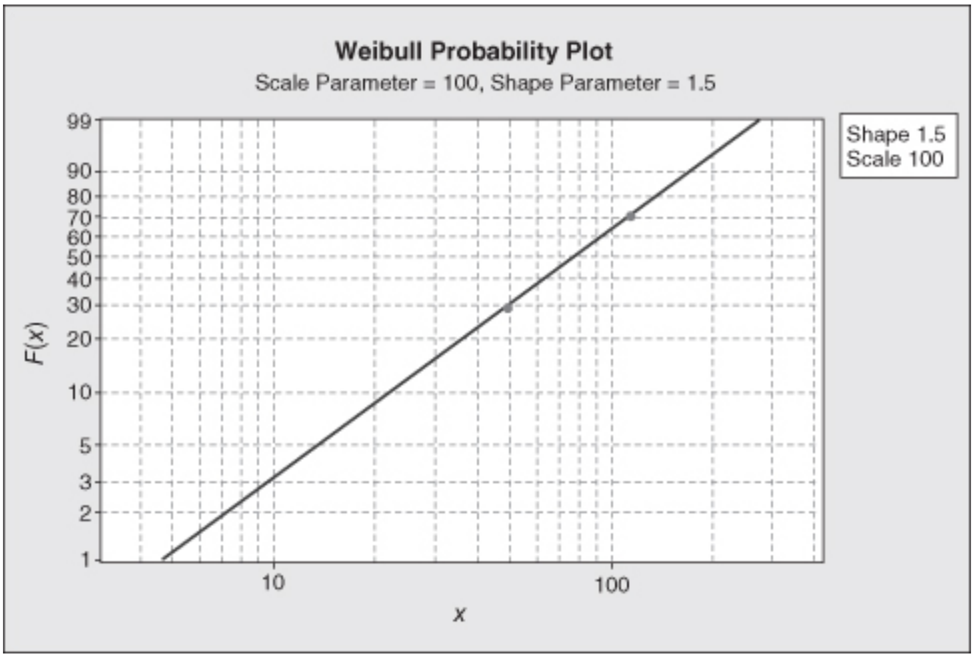
$$R(x) = \text{Prob}[X > x] = exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right];\ x > 0. \tag{3.2}$$

Taking logarithms twice gives

$$\ln\ln\left(\frac{1}{R(x)}\right) = \beta\ln x - \beta\ln\eta. \tag{3.3}$$

On graph paper with a vertical scale ruled proportionally to the values of $\ln\ln(1/R(x))$ and with a logarithmic horizontal scale, the Weibull CDF plots against x as a straight line. Graph paper so constructed is often called Weibull paper. Figure 3.1 shows the line corresponding to the Weibull distribution having $\beta = 1.5$ and $\eta = 100$. This plot can be used to determine $F(x)$ to within graphical accuracy for x values of interest. The principal use of Weibull paper, however, is in the graphical estimation of the Weibull parameters, a topic that we will take up in Chapter 5. With one choice of the scales used in the construction of Weibull paper, the slope of the straight line representation becomes numerically equal to the Weibull shape parameter. In some engineering literature β is therefore known as the Weibull slope. In some other graph paper designs in popular use, an auxiliary scale is provided to set the shape parameter appropriately. See for example Nelson (1967) and Nelson and Thompson (1971).

Figure 3.1 A Weibull probability plot for the population $W(100, 15)$.

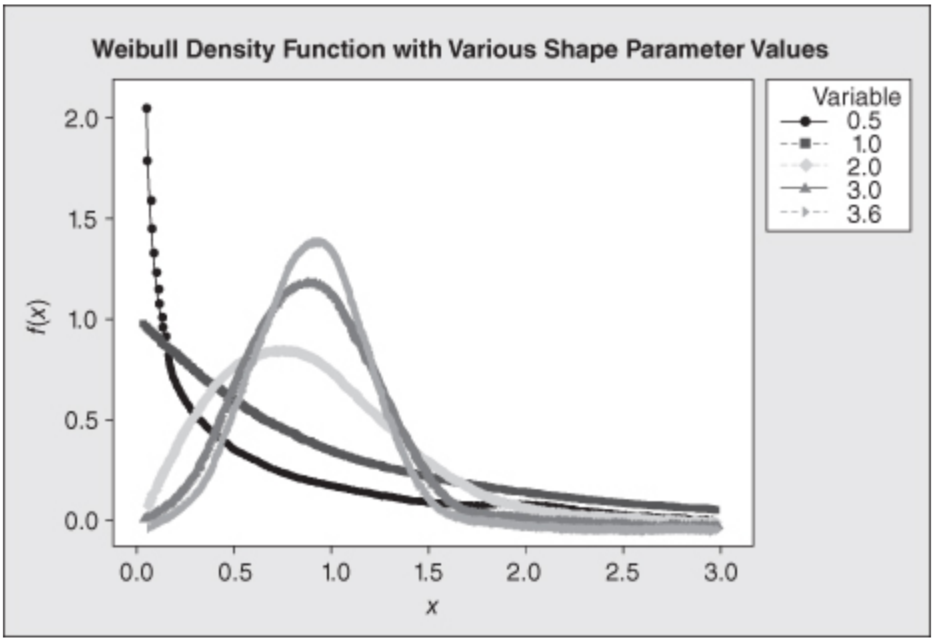


The probability density function (pdf) of the two parameter Weibull distribution is:

$$f(x;\eta,\beta) = \frac{dF(x)}{dx} = \frac{\beta}{\eta}\left[\frac{x}{\eta}\right]^{\beta-1} \cdot exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right]. \tag{3.4}$$

Figure 3.2 is a plot of $f(x)$ for $\eta = 1$ and several values of β. It illustrates the powerful role of β in determining the shape of the Weibull density and suggests visually why the Weibull distribution can be fit to widely diverse kinds of random phenomena.

Figure 3.2 Weibull pdf for various values of the shape parameter β.



3.1.1 Hazard Function

From Equation 2.49 in Chapter 2 the hazard function for the two-parameter Weibull distribution is:

$$\lambda(x) = \frac{f(x)}{R(x)} = \frac{\beta x^{\beta-1}}{\eta^{\beta}}. \tag{3.5}$$

When $\beta = 1$, the hazard function is constant signifying that the likelihood of failure is unaffected by age. In this case the Weibull reduces to the exponential distribution. The constant failure rate is $\lambda = 1/\eta$. For $\beta > 1$ the hazard increases with age, while for $\beta < 1$, the hazard decreases with age. The ability to account for increasing, decreasing, and constant failure rate behavior underscores the great flexibility of the Weibull distribution.

As indicated by Equation 2.50, the reliability function is related to the hazard rate by:

$$R(x) = \exp\left[-\int_0^x \lambda(x)\,dx\right] = \exp[-\Lambda(x)]. \tag{3.6}$$

$\Lambda(x)$ is the integral of the hazard function from 0 to x, and is known as the cumulative hazard function. Integrating Equation 3.5 gives:

$$\Lambda(x) = \left(\frac{x}{\eta}\right)^{\beta}. \tag{3.7}$$

Using Equation 2.52, one may express the average failure rate over an interval from x_1 to x_2 as:

$$\bar{\lambda} = \frac{\int_{x_1}^{x_2} \lambda(x)\,dx}{x_2 - x_2} = \frac{\Lambda(x_2) - \Lambda(x_1)}{x_2 - x_1} = \frac{x_2^{\beta} - x_1^{\beta}}{\eta^{\beta}[x_2 - x_1]}. \tag{3.8}$$

Example

The life of a product follows a Weibull distribution with a shape parameter of 1.5 and a scale parameter of 1000 hours. Compute the instantaneous failure rate at 500 hours and the average failure rate over the time interval from 500 to 1500 hours:

$$\lambda(500) = \frac{1.5 \cdot (500)^{1.5-1}}{(1000)^{1.5}} = 1.061 \times 10^{-3} = 1.061/1000 \text{ hours.}$$

The average failure rate from 500 to 1500 hours is:

$$\bar{\lambda} = \frac{(1500)^{1.5} - (500)^{1.5}}{1500 - 1000} = 1.484/1000 \text{ hours.}$$

If a limited mission life is at issue one could consider the approximation of replacing a Weibull model by an exponential model having a constant failure rate equal to the Weibull average failure rate over the mission life. This approximation could facilitate further calculations and is reasonable if the failure rates at the beginning and end of the interval in question are not greatly different.

3.1.2 The Mode

The mode of a distribution is the value of x at which the probability density function is largest. For a β value of 1.0 or less, the mode of the Weibull distribution occurs at $x = 0$. For $\beta \geq 1.0$ the mode occurs at:

$$x_m = \eta\left(\frac{\beta - 1}{\beta}\right)^{1/\beta} \tag{3.9}$$

For large values of the shape parameter the mode approaches the scale parameter η.

3.1.3 Quantiles

The p-th quantile of the two-parameter Weibull distribution, found as the solution of $F(x_p) = p$, is:

$$x_p = \left\{\ln\left(\frac{1}{1-p}\right)\right\}^{1/\beta} \cdot \eta. \tag{3.10}$$

It is convenient to define

$$k_p \equiv \ln\left(\frac{1}{(1-p)}\right). \tag{3.11}$$

so that the p-th quantile may be written,

$$x_p = \eta k_p^{1/\beta}. \tag{3.12}$$

To graphical accuracy one may read quantiles from a Weibull plot by setting $F(x) = 100p$ on the ordinate and reading the corresponding abscissa.

Using Equation 3.12 in Equation 3.2, one may express the CDF in terms of β and a general quantile, that is,

$$F(x) = 1 - exp\left[-k_p\left(\frac{x}{x_p}\right)^{\beta}\right]. \tag{3.13}$$

This parameterization is convenient when a product is rated by the value of a specific quantile. The bearing industry, for example, rates its products by the tenth quantile $x_{0.10}$ and manufacturers' catalogs contain factors whereby $x_{0.10}$ may be computed for a given loading. It is therefore convenient for bearing engineers to calculate reliabilities directly in terms of $x_{0.10}$ rather than having. to first convert x_{010} to η.

An equivalent alternative form is:

$$F(x) = 1 - (1-p)^{\left(\frac{x}{x_p}\right)^{\beta}}. \tag{3.14}$$

For $p = 1 - 1/e = 0.632$, $x_p = \eta$, regardless of the value of β. Since all other percentiles depend on both η and β, $x_{0.632}$ or η is called the characteristic value or, in life testing applications, the characteristic life.

The ratio of two quantiles, say x_p and x_q, may be found using Equation 3.10 as:

$$\frac{x_p}{x_q} = \left[\frac{\ln(1-p)}{\ln(1-\mathrm{q})}\right]^{1/\beta}. \tag{3.15}$$

3.1.4 Moments

The expected value of x raised to an integer power, k, is conveniently expressible as:

$$E(X^k) = \int_0^{\infty} x^k dF(x) = \eta^k \Gamma\left(\frac{k}{\beta}+1\right) = \eta^k B_k. \tag{3.16}$$

where,

$$B_k = \Gamma\left(\frac{k}{\beta}+1\right). \tag{3.17}$$

And $\Gamma(\cdot)$ is the gamma function of applied mathematics defined as:

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt. \tag{3.18}$$

The gamma function is widely tabulated and otherwise available in computing software. Excel contains a function that computes the logarithm of the gamma function. $E(X^k)$ is called the k-th raw moment. The terminology *raw moment* distinguishes it from the k-th *central moment* defined as $E(x - \mu)^k$.

The mean $\mu = E(x)$ may therefore be written,

$$\mu = \eta B_1. \tag{3.19}$$

In life testing applications the mean is often designated MTTF for mean time to failure. It is sometimes called MTBF for mean time between failures, although most writers reserve this term for repairable systems. Bearing application engineers, knowing the computed value of $x_{0.10}$ for their product may, on occasion, encounter requests from reliability analysts of a client company to provide the MTTF for their bearings. They may use Equation 3.19 after first computing η in terms of $x_{0.10}$ using Equation 3.10:

$$\eta = \frac{x_{0.10}}{(-\ln(0.90))^{1/\beta}}. \tag{3.20}$$

The second raw moment is,

$$E(X^2) = \eta^2 B_2. \tag{3.21}$$

The variance σ^2 is thus expressible as:

$$\sigma^2 = E(X^2) - \mu^2 = \eta^2[B_2 - B_1^2]. \tag{3.22}$$

Values of $B_1 \equiv \Gamma\left(\frac{1}{\beta}+1\right)$ and $B_2 - B_1^2 = \Gamma(2/\beta+1) - \Gamma^2(1/\beta+1)$ taken from Abramowitz and Stegun (1964) are listed in Table 3.1. It is seen from Table 3.1 that for a fixed value of η, σ^2 decreases with increasing β. β itself is therefore often used directly to characterize the dispersion of the Weibull distribution. As $\beta \to \infty$ the variance approaches 0 and the mean approaches η. Thus, the Weibull distribution with a very large shape parameter value acts like a constant equal to its scale parameter. Referring to the Weibull plot in Figure 3.1 as β increases, the line pivots about the point $F(x) = 0.632$, $x = \eta$, as it approaches the vertical.

Table 3.1 Values of $B_1 = \Gamma(1/\beta + 1) =$ and $\mathbf{B_2 - B_1^2 = \Gamma(2/\beta+1) - \Gamma^2(1/\beta+1)}$ as a Function of the Shape Parameter β

β	B_1	$B_2 - B_1^2$
1.0	1.0000	1.0000
1.1	0.9649	0.7714
1.2	0.9407	0.6197
1.3	0.9336	0.5133
1.4	0.9114	0.4351
1.5	0.9027	0.3757
1.6	0.8966	0.3292
1.7	0.8922	0.2919
1.8	0.8893	0.2614
1.9	0.8874	0.2360
2.0	0.8862	0.2146
2.5	0.8873	0.1441
3.0	0.8930	0.1053
3.5	0.8997	0.0811
4.0	0.9064	0.0647
5.0	0.9182	0.0442

For $\beta = 1$ the Weibull distribution becomes the single-parameter exponential distribution and its mean is the scale parameter η while its variance is η^2. As already noted its failure rate becomes a constant equal to $1/\eta$.

The coefficient of variation is:

$$cv = \frac{\sigma}{\mu} = [B_2 - B_1^2]^{0.5} / B_1. \tag{3.23}$$

The coefficient of variation depends only upon the shape parameter β. Once β is specified the standard deviation is set as a fixed fraction of the mean. The coefficient of variation for the exponential distribution is 1.0.

The skewness, sk is defined as,

$$sk = \frac{\mu_3}{\sigma^3} \tag{3.24}$$

μ_3 is the third moment about the mean $E(x - \mu)^3$. Expressing all terms in the numerator and denominator in terms of B_i gives:

$$sk = \frac{B_3 - 3B_1B_2 + 2B_1^2}{[B_2 - B_1^2]^{3/2}}. \tag{3.25}$$

Skewness decreases with β, becoming zero at about $\beta = 3.6$, at which value the Weibull distribution is a reasonable approximation to the normal. This does not mean that the Weibull can approximate any normal distribution. Once β is specified to be 3.6 in order to achieve zero skewness, the standard deviation becomes a fixed fraction of the mean as shown by the coefficient of variation expression in Equation 3.23, that is, $\sigma = (0.308)\mu$. Only normal distributions for which the standard deviation and the mean are in this proportion will be well approximated by the Weibull. Dubey (1967) has studied the comparative behavior of Weibull and, normal distributions in some detail.

Examples

1. An item is randomly drawn from a two-parameter Weibull population having a shape parameter β = 1.5 and a scale parameter η = 100.0 hours. What is the probability that the item fails before achieving a life of x = 25 hours?

From Equation 3.2,

$$\begin{aligned}\text{Prob}[life < 25.0] &= 1 - e^{-\left(\frac{25}{100}\right)^{1.5}} \\ &= 1 - e^{-0.125} = 0.118.\end{aligned}$$

This result may be confirmed to graphical accuracy from the Weibull plot in Figure 3.1. Figure 3.3 shows the pdf with the calculated probability shown shaded.

2. Compute the tenth percentile $x_{0.10}$ for this distribution. From Equation 3.10,

$$x_{0.10} = [0.10536]^{\left(\frac{1}{1.5}\right)} \cdot 100.0 = 22.31.$$

Entering the ordinate of Figure 3.1 at $F(x)$ = 10% will confirm this computation to within graphical accuracy.

3. Compute the ratio of the 10th and 50th percentiles for this distribution. Taking p = 0.50 and q = 0.10 in Equation 3.14, the ratio is:

$$\frac{x_{0.50}}{x_{0.10}} = \left(\frac{\ln(1-0.50)}{\ln(1-0.10)}\right)^{\frac{1}{1.5}} = 3.51.$$

The value of the median, or $x_{0.50}$ is therefore 3.51 × 22.31 = 78.3. The median is shown on the pdf in Figure 3.4.

Note that, like the coefficient of variation, the ratio of the 50th and the 10th percentiles depends only on the shape parameter β. Bearing engineers sometimes state that $x_{0.50}$ is about 5 times $x_{0.10}$. This claim is literally true only if β = 1.17, which is fairly typical of the shape parameter values commonly reported for ball bearings. For β = 10, $x_{0.50}$ = 1.2 $x_{0.10}$.

4. Calculate the MTTF and the variance for the Weibull population of example 1.

From Table 3.1 for β = 1.5, B_1 = 0.9027 and $B_2 - B_1^2 = 0.3757$.

From Equation 3.19,

$$\mu = E(X) = MTTF = 100.0 \times 0.9027 = 90.27.$$

From Equation 3.22,

$$\sigma^2 = (100)^2 \times 0.3757 = 3757.0.$$

Figure 3.3 Weibull density showing $P[X < 25]$.

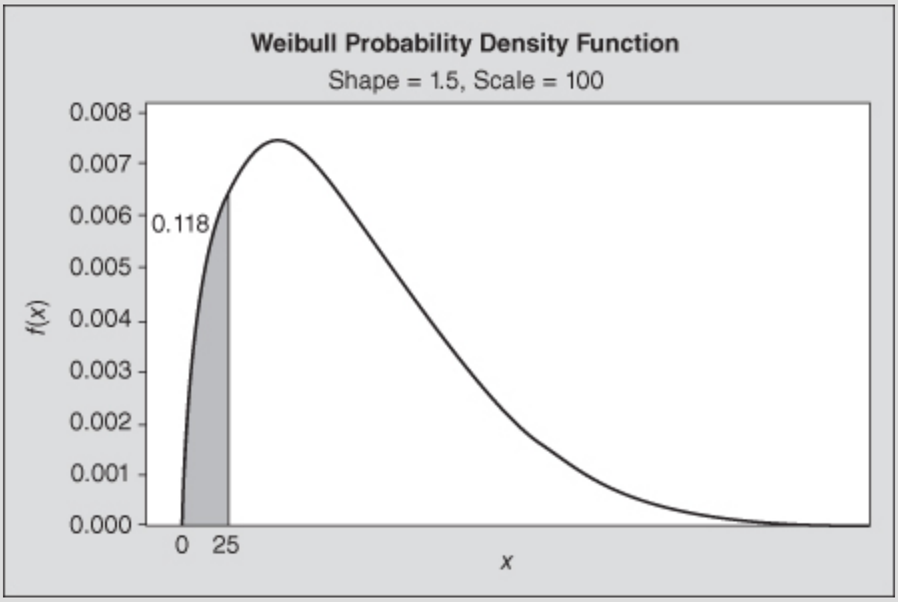
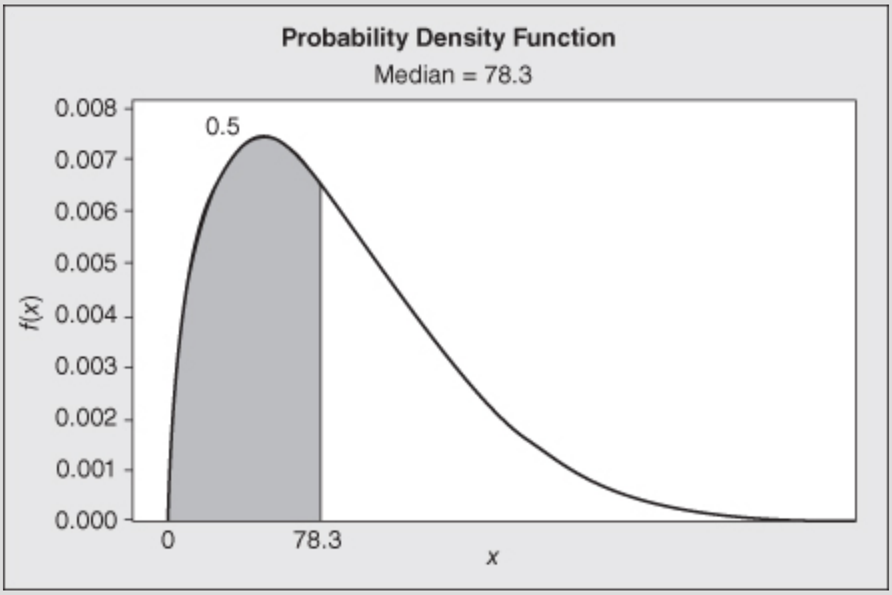


Figure 3.4 The Weibull median $x_{0.50}$.



3.2 THE MINIMA OF WEIBULL SAMPLES

We denote the fact that X follows the two-parameter Weibull distribution by the notation $X \sim W(\eta, \beta)$. Given a random sample of size n drawn randomly from $W(\eta, \beta)$, the random variable,

$$Y = \min(x_1, x_2, \ldots x_n). \quad (3.26)$$

follows $W(\eta/n^{1/\beta}, \beta)$. Suppose for example you had 10,000 samples of size 5 drawn from $W(100, 2)$ and you sorted each sample and found the 10,000 smallest values in each sample. The distribution of those 10,000 minima would act like a large sample drawn from $W(100/\sqrt{5}, 2)$. The scale parameter would be reduced to 44.72 but the shape parameter would be the same, $\beta = 2$. The two populations would appear as parallel lines on a Weibull plot with the distribution corresponding to the minima appearing above the original or parent distribution.

It is readily shown that the distribution of Y is as claimed. If the smallest value in a sample exceeds some value y, then *every* member of the sample must exceed y. The probability that a Weibull random variable exceeds y is $1 - F(y)$, where $F(y)$ is the Weibull CDF. Assuming independence of the observations in a sample, the probability that every sample observation exceeds y is then $[1 - F(y)]^n$.

Let Y denote the random variable representing the minimum of a sample of size n and having CDF $G(y)$, then,

$$P(Y > y) = 1 - G(y) = [1 - F(y)]^n = \left[exp\left(-\left(\frac{y}{\eta}\right)^\beta\right)\right]^n. \tag{3.27}$$

Thus, the CDF of the distribution of minima is:

$$G(y) = 1 - exp\left[-\left(\frac{y}{\eta / n^{1/\beta}}\right)^\beta\right]. \tag{3.28}$$

That the Weibull is preserved under minimization reflects the role of the Weibull distribution as a limiting distribution of smallest extremes. This fact is often used to justify the Weibull model as appropriate to failure mechanisms governed by a “weakest link” behavior. This property will be discussed later in this chapter. The distribution of Weibull minima is the basis for a testing strategy called sudden death testing espoused by Johnson (1964) and discussed further in Section 7.6. It is also related to the life of series systems made of n identical Weibull elements and discussed in conjunction with system reliability in Chapter 7.

The mean, variance, and percentiles for the distribution of the minimum are computed in the same way as for any Weibull distribution after using $\eta/n^{1/\beta}$ in lieu of η.

3.3 TRANSFORMATIONS

3.3.1 The Power Transformation

If a random variable is Weibull distributed, that is, $X \sim W(\eta, \beta)$, we now show that the transformed variable $Y = aX^c$, where a and c are positive constants, is also Weibull distributed. Let $G(y)$ denote the CDF of Y.

$$\begin{aligned} G(y) &= \text{Prob}[Y < y] = \text{Prob}\left[X < \left(\frac{y}{a}\right)^{\frac{1}{c}}\right] = F\left(\left(\frac{y}{a}\right)^{\frac{1}{c}}\right) \\ &= 1 - \exp\left[-\left(\frac{\left(\frac{y}{a}\right)^{\frac{1}{c}}}{\eta}\right)^\beta\right]. \end{aligned} \tag{3.29}$$

Thus, $Y \sim W(a\eta^c, \beta/c)$.

This transformation for the Weibull is reminiscent of the linear transformation $Y = a + bX$ for the normal. In both cases the distribution of the transformed variable is in the same family as the untransformed variable, but with modified parameters. For the normal if $b = 1$, Y has the same variance as X but the mean is modified by an additive amount "a." In the Weibull case for $c = 1$, the Weibull distribution of Y has the same shape parameter as X but the scale parameter η is modified by the multiplicative factor a.

Choosing $a = 1$ and $c = \beta$ transforms any Weibull distribution to the exponential ($\beta = 1$) with mean value η. Suitably choosing a and c, one can transform one Weibull variable to any other, different, Weibull distribution having a more "convenient" shape parameter. Nelson (1994) made use of this fact to transform an exponential random variable (Weibull with $\beta = 1.0$) to a Weibull distribution having a shape parameter of $\beta = 3.6$ and for which the skewness is zero, using $a = 1$ and $c = 1/3.6 = 0.277$. As noted previously when the shape parameter is 3.6 the Weibull distribution is roughly symmetrical about its mean, and in a control chart of a Weibull variable thus transformed, one may reasonably employ auxiliary tests such as runs above and below the mean in assessing whether the process is in a state of control.

A purely multiplicative transformation, $Y = aX,(c = 1)$, does not affect the shape parameter and multiplies the scale parameter by the same factor. The practical benefit is that one may freely change units provided the change is purely multiplicative. If X in inches is $W(1, 3)$, then Y in millimeters will be $W(25.4, 3)$. If, however, temperature in centigrade follows the two-parameter Weibull, the temperature in Fahrenheit will not follow the two-parameter Weibull distribution since the transformation between the two scales is not purely multiplicative.

In the design and analysis of experiments with normally distributed response, it is customary to hypothesize that the effect of experimental factors is to change the mean by an additive amount while leaving the variance unchanged. With a Weibull response variable, the analogous assumption would be that external factors such as stress produce a multiplicative effect on the Weibull scale parameter while leaving the shape parameter unchanged. Such a model has been substantiated for bearing endurance life by Lieblein and Zelen (1956). In view of Equation 3.22, this implies heteroscedasticity; that is, the variance will increase with the mean response although β remains the same. The analysis of a one-way layout under the assumption of a constant shape parameter and multiplicative effects on the scale parameter was considered by McCool (1979) and extended to a two-way layout in McCool (1993). This topic is discussed further in Chapters 8 and 11.

3.3.2 The Logarithmic Transformation

The natural logarithm of a Weibull variable follows the type I distribution of smallest extremes, (cf. Gumbel 1958). There are many advantages of working with the logarithmically transformed Weibull variable as we shall discover. Under the logarithmic transformation

(3.30) $$Z = \ln(X)$$

Z will be less than an arbitrary value z, as long as X is less than e^z. Therefore, the CDF $G(z)$ may be found as:

(3.31) $$G(z) = \text{Prob}[Z < z] = F(e^z) = 1 - exp\left(-\left[\frac{e^z}{\eta}\right]^\beta\right).$$

Introducing the new parameters $\delta = \ln \eta$ and $\xi = 1/\beta$, $G(z)$ may be written in the following form:

(3.32) $$G(z) = 1 - exp\left\{-exp\left[\frac{z-\delta}{\xi}\right]\right\}.$$

In this form the distribution has a location parameter δ and scale parameter ξ analogous to the two parameters of the normal distribution. It is clear from Equation 3.32 why this distribution is sometimes called the doubly exponential distribution. As with the normal, it is useful to define a standardized variable Y as:

$$Y = \frac{(Z-\delta)}{\xi}. \tag{3.33}$$

Y follows the extreme value distribution with location parameter 0 and scale parameter 1. The p-th quantile of Y is:

$$y_p = \ln\ln\left\{\frac{1}{(1-p)}\right\} = \ln(k_p). \tag{3.34}$$

The corresponding quantile for Z is

$$z_p = \delta + \xi y_p. \tag{3.35}$$

The same quantile for the Weibull variate is computed by retransforming:

$$x_p = \exp(z_p). \tag{3.36}$$

The mean of the standardized variate is the negative of Euler's constant $\gamma = 0.57721$, so in terms of the Weibull parameters,

$$E(Z) = \ln\eta - \frac{\gamma}{\beta}. \tag{3.37}$$

The variance of the standardized variable Y is $\pi^2/6$ so that

$$\sigma_y^2 = \frac{\sigma_z^2}{\xi^2} = \frac{\pi^2}{6} \tag{3.38}$$

and thus the variance of $Z = \ln X$ is

$$\sigma_z^2 = \frac{\pi^2}{6\beta^2}. \tag{3.39}$$

Equations 3.37 and 3.39 show that a multiplicative factor applied to the scale parameter, that is, replacing η by $a\eta$ has no effect on the variance of $Z = \ln X$ and adds an amount $\ln(a)$ to its expected value. Thus, a multiplicative Weibull model for the effects of one or more factors, combined with the assumption of a constant shape parameter, will result in data which when logarithmically transformed will follow an additive model with homogeneous variance. Except for normality, such data obey the principal assumptions of the analysis of variance. A reasonable approximate analysis of multifactor Weibull data applicable if censoring is absent would be to apply the analysis of variance to the logarithms of the observations. An exact approach for multifactor Weibull data is discussed in Chapter 11.

Menon (1963) recommends estimating the Weibull parameters by the method of moments applied to the logarithms of the data. Menon's approach and recent extensions to it are discussed in Section 5.3.1.

3.4 THE CONDITIONAL WEIBULL DISTRIBUTION

The distribution of a Weibull variate conditional on $X \geq x_0$ has been termed the conditional Weibull distribution by Aroian (1965). It is termed the truncated Weibull distribution by Harlow (1989). It is of particular use when the Weibull is used as a lifetime model although other applications are possible.

Let $A = \{x_0 < X < x\}$ and $B = \{X > x_0\}$. Then $P\,[A|B]$ is the CDF of X conditional on $X > x_0$:

$$F(x|x_0) = \frac{F(x) - F(x_0)}{1 - F(x_0)} = 1 - \exp\left[-\left(\frac{x}{\eta}\right)^{\beta} + \left(\frac{x_0}{\eta}\right)^{\beta}\right]; x > x_0. \tag{3.40}$$

It is useful to define y as life lived beyond x_0, that is, $y = x - x_0$ is the residual life after having run for a period of x_0. This is the life as measured by a customer who has received a product that was "burned in" for a period of time x_0. To do so we replace x by $y + x_0$ in Equation 3.40.

$$F(y|x_0) = 1 - exp\left[-\left(\frac{y + x_0}{\eta}\right)^{\beta} + \left(\frac{x_0}{\eta}\right)^{\beta}\right]. \tag{3.41}$$

This expression shows that when $\beta = 1.0$ the conditional exponential is the same as the unconditional exponential. The run-in or burn-in period has had no effect. This is a manifestation of the so-called memorylessness property of the exponential distribution. For $\beta \neq 1$ the conditional Weibull does not have the Weibull form.

A run-in time that is too short increases the risk of failure under warranty when the burned-in product is delivered to the customer. An overly long run-in time increases the cost due to the loss of product failing during run-in and the running costs. Determining the optimum run-in duration to minimize the total cost is considered in Section 4.9.

A simpler, equivalent way of expressing Equation 3.41 is in terms of the reliability function, namely:

$$R(y|x_0) = \frac{R(y + x_0)}{R(x_0)} = exp\left[-\left(\frac{y + x_0}{\eta}\right)^{\beta} + \left(\frac{x_0}{\eta}\right)^{\beta}\right]. \tag{3.42}$$

For $x_0 = 0$ this reduces to the ordinary, unconditional, reliability function.

The p-th quantile of the conditional distribution y_p may be found by setting $F(y|x_0)$ equal to p and solving for y. It is expressible in terms of the p quantile of the original, unconditional, distribution as,

$$y_p = x_p\left[1 + \left(\frac{x_0}{x_p}\right)^{\beta}\right]^{1/\beta} - x_0. \tag{3.43}$$

When equipment is "burned in," that is, subject to a period of operation of length x_0 to eliminate early failures (infant mortality), the customer's shipment is drawn from the population of survivors. Equation 3.43 gives the 100p-th percentile as measured from when the customer puts the item into service.

When $\beta = 1.0$, $y_p = x_p$, signifying that no deleterious effect of aging has occurred. For $\beta > 1.0$, $y_p < x_p$; that is, aging occurs. For $\beta < 1$, $y_p > x_p$, indicating that the survivors of the run-in period are superior to the population as a whole.

The mean remaining life, also called the mean residual life, is obtained by integrating the conditional reliability function. It is a function of the run-in time x_0:

$$MRL(x_0) = \int_0^\infty R(y|x_0)dy. \quad (3.44)$$

Since the reliability function for $x_0 = 0$ is the unconditional reliability function, MRL(0) = MTTF.

MRL(x_0) may be computed in terms of the incomplete gamma function (cf.Leemis, 1995)

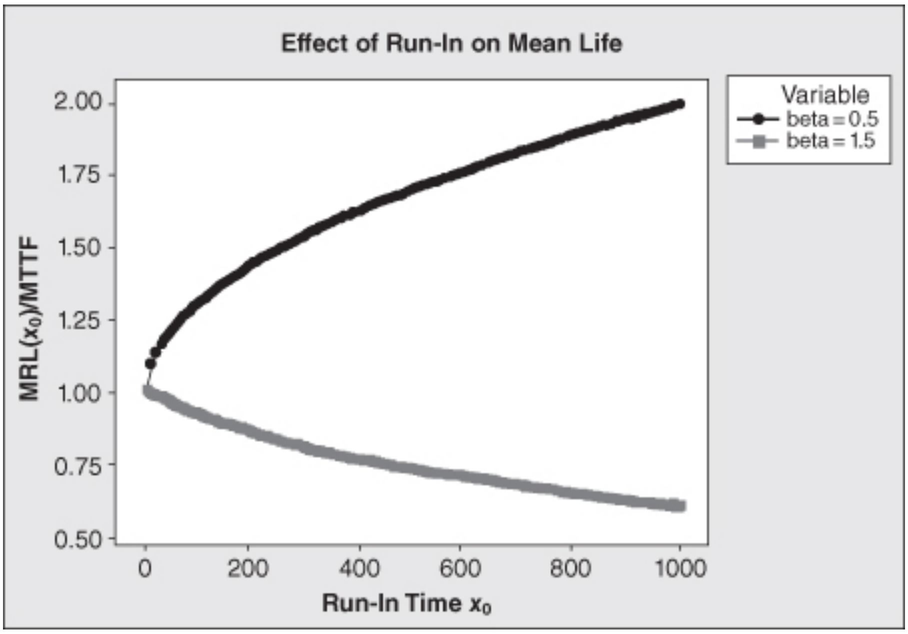
$$MRL(x_0) = \frac{\eta}{\beta} exp\left[\left(\frac{x_0}{\eta}\right)^\beta\right]\Gamma\left[\frac{1}{\beta}, \left(\frac{x_0}{\eta}\right)^\beta\right] \quad (3.45)$$

where $\Gamma(a, x)$ denotes the incomplete gamma function defined as:

$$\Gamma(a, x) = \int_x^\infty t^{a-1}e^{-t}dt. \quad (3.46)$$

The mean residual life increases with x_0 for $\beta < 1$ and decreases with x_0 for $\beta > 1$. Figure 3.5 is a plot of the mean residual life divided by MTTF versus x_0 for two cases: $\beta = 0.5$ and $\beta = 1.5$. In both cases $\eta = 1000$. The figure shows that when $\beta < 1$ the MRL increases with burn-in, while for $\beta > 1.0$ the MRL decreases with burn-in.

Figure 3.5 Mean residual life versus run-in time for two values of β.



Examples

1. A product's life in hours follows the distribution *W*(30E6, 1.5). Compute the first percentile. If the product is run for 2E6 hours and the survivors are sold, what is the first percentile of the surviving population from the point of view of the new owner?

From Equation 3.10 the 0.01 quantile, $x_{0.01}$ is:

$$x_{0.01} = 30E6 \cdot [-\ln(1-0.01)]^{1/1.5} = 6.51E4 \text{ hours.}$$

From Equation 3.43 the 0.01 quantile post run-in, $y_{0.01}$, is:

$$y_{0.10} = 6.51E4 \cdot \left[1 + \left(\frac{2E6}{6.51E4}\right)^{1.5}\right]^{1/1.5} - 2.0 = 7.857E3 \text{ hours.}$$

2. Consider an item that fails in accordance with a Weibull distribution with scale parameter $\eta = 240$ months and shape parameter $\beta = 0.6$. (Recall that the Weibull has a decreasing failure rate when $\beta < 1$.) The reliability at $t = 2$ months without burn-in is given by:

$$R(2) = \exp - \left(\frac{2}{240}\right)^{0.6} = 0.945.$$

Now calculate the reliability at 2 months in the population of survivors of a burn-in period of 3 months. Using Equation 3.42 results in:

$$R(2 \mid 3) = \exp - \left(\frac{2+3}{240}\right)^{0.6} \Big/ \exp - \left(\frac{3}{240}\right)^{0.6} = \frac{0.9066}{0.9304} = 0.9745.$$

Thus the customer would see a sample from a population having reliability of 97.45% rather than 94.5% which would be the reliability without burn-in.

What fraction of the population would fail to survive the burn-in period?

$$\text{Prob}\,(life < 3\; months) = F(2) = 1 - \exp - \left(\frac{3}{240}\right)^{0.6} = 0.0696 \; or \; 6.96\%.$$

The MTTF and MRL(3) are:

$$MTTF = 361.1$$

$$MRL(3) = 385.0.$$

3.5 QUANTILES FOR ORDER STATISTICS OF A WEIBULL SAMPLE

Order statistics are the individual values in a random sample, after they are sorted in ascending sequence; for example, the first order statistic is the smallest value in the sample, the second order statistic is the second smallest value, and so on. We have already seen that the first order statistic in Weibull samples varies from sample to sample in accordance with a Weibull distribution having a reduced scale parameter.

Order statistics are random variables varying from sample to sample in accordance with a probability distribution that depends on the underlying parent distribution, the sample size, and the order number under consideration. Consider a large number of samples of size n randomly drawn from a distribution whose CDF is $F(x)$. The r-th order statistic $x_{r,n}$ will vary from sample to sample. Now transform the values of $x_{r,n}$, the r-th ordered member of each sample by computing $F(x_{r,n})$ where $F(x)$ is the CDF of the distribution from which the samples were drawn. $F(x_{r,n})$ will now vary from sample to sample since $x_{r,n}$ varies from sample to sample. Amazingly the distribution of $F(x_{r,n})$ in repeated samples is the same regardless of the population from which the samples were taken, and it is known as the beta distribution. The beta distribution has two parameters customarily denoted a and b. $F(x_{r,n})$ follows the beta distribution with $a = r$ and $b = n - r + 1$. Let $X(p, a, b)$ denote the p-th quantile of the beta distribution having parameters a, b. In this notation a 90% probability interval for $F(x_{r,n})$ may be expressed as :

$$X(0.05, r, n-r+1) < F(x_{r,n}) < X(0.95, r, n-r+1). \tag{3.47}$$

Values of $X(p, a, b)$ are tabled for various values of p, a, and b by Harter (1964). Commercial statistical software such as Minitab can provide percentage points of the beta distribution. Some values of these percentage points applicable for a sample size $n = 5$ and $r = 1(1)5$ are listed in Table 3.2. In the literature dealing with probability plotting, these values are known as the 5% and 95% ranks. The median value of $F(x_{r,n})$ is called the median rank and is denoted as $X(0.50, r, n - r + 1)$ in our notation. More extensive tables of the 5%, 50%, and 95% ranks are given in tables provided by Kapur and Lamberson (1977) and others. In plotting life data for the purpose of estimating the Weibull parameters, some software includes the plotting of the 5% and 95% ranks as the basis for approximate confidence limits. This is discussed further in Chapter 5.

Table 3.2 Values of $X(0.05, r, 6 - r)$ and $X(0.95, r, 6 - r)$ for $n = 5$

r	$X(0.05, r, 6-r)$	$X(0.95, r, 6-r)$	$X(0.50, r, 6-r)$	$\frac{r-0.3}{n+0.4}$
1	0.01021	0.45072	0.12945	0.1296
2	0.07644	0.65741	0.31381	0.3148
3	0.18926	0.81074	0.50000	0.50
4	0.34259	0.92356	0.68619	0.6852
5	0.54928	0.98979	0.87055	0.8704

The median ranks are often recommended as plotting positions for graphical estimation of the Weibull parameters. The following approximation to the median ranks was proposed by Benard and Bos-Levenbach (1953) and is in wide use:

$$X(0.50, r, n-r+1) \approx \frac{r-0.3}{n+0.4}. \tag{3.48}$$

Values of this approximation for $n = 5$ are shown in column 5 of Table 3.2. It is clear that they are indistinguishable from the exact values to within graphical accuracy.

It is sometimes useful to be able to calculate an interval in which a given order statistic will fall with high probability. For example, if one wished to obtain five failed specimens for metallurgical investigation, and put 10 specimens on test, the waiting time to completion of the test is the fifth order statistic in the sample of size 10, designated $x_{5,10}$.

A 90% probability interval for the r-th order statistic in a sample of size n from any distribution may be calculated by taking the inverse of $F(x)$ in Equation 3.47 as follows,

$$F^{-1}[X(0.05, r, n-r+1)] < x_{r,n} < F^{-1}[X(0.95, r, n-r+1)]. \quad (3.49)$$

where $F^{-1}(\cdot)$ is the inverse of the parent distribution function.

Calculation of the above interval may be carried out graphically on a probability plot by entering the values of $X(p, a, b)$ on the probability ordinate and reading the interval values as the associated abscissa values (cf. McCool 1969).

It may also be performed analytically by using the mathematical expression for the inverse of the distribution function. For the two-parameter Weibull distribution this yields,

$$\eta\left[\ln\left(\frac{1}{1-X(0.05, r, n-r+1)}\right)\right]^{1/\beta} < x_{r,n} < \eta\left[\ln\left(\frac{1}{1-X(0.95, r, n-r+1)}\right)\right]^{1/\beta}. \quad (3.50)$$

Example

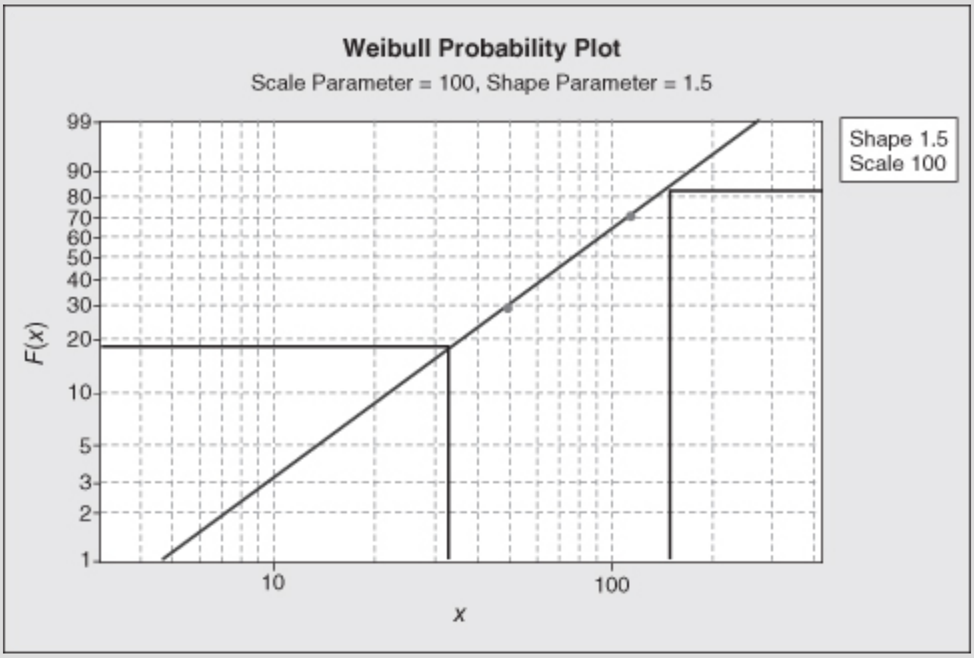
Calculate a 90% interval for the third-order statistic in a sample of size 5 drawn from the two-parameter Weibull population having $\beta = 1.5$ and $\eta = 100$.

Using Equation 3.50 and the 5th and 95th percentiles $X(0.05, 3, 3) = 0.18926$ and $X(0.95, 3, 3) = 0.81074$ gives the following interval:

$$35 < x_{3,5} < 140.$$

These calculations may be roughly corroborated by entering the ordinate of the Weibull plot of the population at these two percentile values (after multiplying by 100) and reading the corresponding abscissae. See Figure 3.6:

Figure 3.6 Graphical determination of probability limits for $x_{3,5}$.



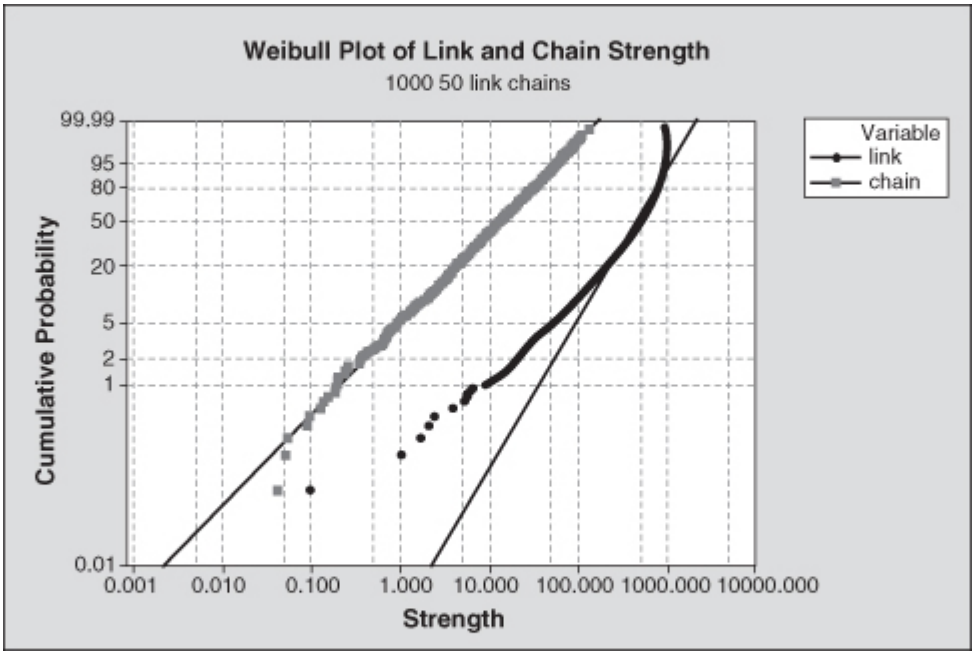
Calculations of this type may be useful for rough estimation of the uncertainty in the time needed to complete a life test. To perform the calculation it is necessary to guess at what the life test results will reveal regarding the parameter values. The evident contradiction is that if the parameters were known there would be no need to conduct the life test.

3.5.1 The Weakest Link Phenomenon

Let X be an arbitrary positive random variable describing the strength of a population of chain links. Randomly assemble chains each consisting of n of these links. The strength of a chain will now be equal to the strength of its weakest link. If in the vicinity of $X = 0$, the CDF of the link strength distribution behaves like a function of the form $F(x) = cx^{\beta}$, then for a sufficiently large number of links, the chain strength will vary from chain to chain in accordance with a Weibull distribution with shape parameter β. This is shown in a somewhat general form by Epstein (1960) in a very readable survey of extreme value theory.

We will demonstrate the truth of this assertion by a simulation exercise. Let link strength X be uniformly distributed $U(0, 1000)$. We know from Equation 2.60 in Section 2.10 that the CDF is $F(x) = x/1000$, so that β, the exponent of x is 1.0. We will randomly sample 1000 sets of 50 observations of $U(0, 1000)$ to represent 1000 chains of 50 links per chain. We will then determine the minimum of each set of 50 to represent the strength of each of the 1000 chains. The 1000 values of chain strength should follow at least an approximate straight line on a Weibull grid. The values of link strength, arbitrarily represented by the 1000 values of the first link in each chain, should not. Figure 3.7 shows the two plots:

Figure 3.7 Weibull plots of chain and link strength.



It is clear that except for some outliers at the lowest percentiles, the chain strength follows the Weibull distribution while the link strength does not. The weakest link argument is often used to justify the adoption of the Weibull to model strength or lifetime. If an item can fail at a large number of potential weak spots such as voids or inclusions the weakest of these will dominate and failure time or strength will vary from item to item in accordance with a Weibull distribution.

3.6 SIMULATING WEIBULL SAMPLES

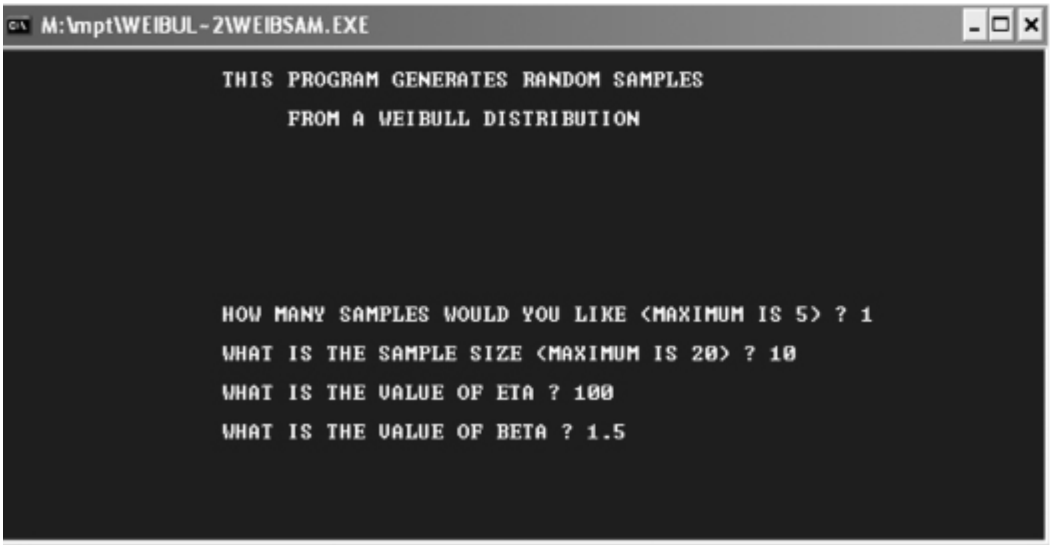
It is useful to be able to generate small samples from a known Weibull population to use for training purposes or for exploring ideas related to data analysis. Large-scale simulations on the order of 10,000 samples are used to develop the distribution of functions needed for confidence intervals and hypothesis tests on the Weibull parameters as discussed in Chapters 5 and 6.

The method described in Section 2.10 works quite well for the Weibull. The CDF $F(x)$ is equated to a value u randomly taken from the distribution which is uniformly distributed over the interval (0, 1). The equation is then solved for x in terms of u and yields:

$$x = \eta(-\ln[u])^{1/\beta} + \gamma. \tag{3.51}$$

This expression uses the fact that the complement $1 - u$ of a uniform random variable is also uniform. Taking a series of n values of the uniform random variable u and applying Equation 3.51 to each produces a sample of size n from the Weibull population with a given set of parameters. As noted in Section 2.10, techniques for generating values that act like samples from a uniform distribution have been widely studied. The values generated are called pseudorandom since a deterministic set of calculations is used to compute them and hence they are not truly random. Excel may be used to generate a set of uniform random values and Equation 3.51 applied to yield Weibull observations. The disk operating system (DOS) program Weibsam may be used to generate small samples from a two-parameter Weibull ($\gamma = 0$). Figure 3.8 shows the input screen with the input needed to generate a single Weibull sample of size 10 from the distribution $W(100, 1.5)$.

Figure 3.8 Input screen for DOS program Weibsam.exe.



The random values thus generated are shown in Figure 3.9.

Figure 3.9 Results screen for DOS progam Weibsam.exe.

REFERENCES

Abramowitz, M. and I. Stegun, eds. 1964. *Handbook of Mathematical Functions*. Applied Mathematics. U.S. Department of Commerce, National Bureau of Standards, Washington, DC.

Aroian, L.A. 1965. Some Properties of the Conditional Weibull Distribution. ASQC Technical Conference Transactions, ASQC Milwaukee, WI.

Benard, A. and E.C. Bos-Levenbach. 1953. The plotting of observations on probability paper (in Dutch). *Statististica Neerlandica* 7: 163–173.

Dubey, S. 1967. Normal and Weibull distributions. *Naval Research Logistics Quarterly* 14: 69–79.

Epstein, B. 1960. Elements of the theory of extreme values. *Technometrics* 2(1): 27–41.

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

Harlow, D.G. 1989. The effect of proof-testing on the Weibull distribution. *Journal of Material Science* 24: 1467–1473.

Harter, H.L. 1964. *New Tables of the Incomplete Gamma-Function Ratio and of the Percentage Points of the Chi-Square and Beta Distributions*. U.S. Govt. Printing Office, Washington, D.C.

Johnson, L.G. 1964. *The Statistical Treatment of Fatigue Experiments*. Elsevier, New York, NY.

Kapur, K.C. and L.R. Lamberson. 1977. *Reliability in Engineering Design*. Wiley, New York, NY.

Leemis, L.M. 1995. *Reliability*. Prentice Hall, Englewood Cliffs, NJ. 07632.

Lieblein, J. and M. Zelen. 1956. Statistical investigation of the fatigue life of deep-groove ball bearings. *Journal of Research of the National Bureau of Standards* 57(5): 273–316.

McCool, J.I. 1969. Graphical determination and uses of order statistics quantiles. *Journal of Quality Technology* 1: 237–241.

McCool, J.I. 1979. Analysis of single classification experiments based on censored samples from the two-parameter Weibull distribution. *Journal of Statistical Planning and Inference* 3(1): 39–68.

McCool, J.I. 1993. The analysis of a two way layout with two parameter Weibull response. PhD Thesis, Department of Statistics, Temple University, Philadelphia, PA, p. 143.

Menon, M.V. 1963. Estimation of the shape and scale parameters of the Weibull distribution. *Technometrics* 5(2): 175–182.

Nelson, L.S. 1967. Weibull probability paper. *Industrial Quality Control* 23: 452–453.

Nelson, L.S. 1994. A control chart for parts-per-million nonconforming items. *Journal of Quality Technology* 26(3): 239–240.

Nelson, W.B. and V.C. Thompson. 1971. Weibull probability papers. *Journal of Quality Technology* 3(2): 45–50.

EXERCISES

1. Let $\eta = 100.0$, $\beta = 1.5$. Find, graphically and/or by computation,

a. Prob[$X < 34$]

b. The value of $x_{0.10}$

c. The value of $x_{0.50}$

d. Prob[$80 < X < 120$]

e. Compute the average failure rate over the interval (80, 120).

2. For what β value is it true that $x_{0.50} = 4.5\ x_{0.10}$?

3. Compute the median of the distribution $W(100, 2)$. Use Weibsam or Excel to generate a sample of 20 values from this distribution. How many are below the population median? How many did you expect to be below the median?

4. The strength in pounds of a population of chain links is $W(200, 2)$. Chains assembled from 20 links randomly selected from this population will break under a load equal to the strength of its weakest link. Compute the mean and median of the chain strength distribution.

5. The radius r in inches of a population of tree trunks is distributed as $W(30, 4)$. Find the mean of the cross-sectional area $A = \pi r^2$ of this population.

6. Under a prescribed set of conditions the life of a product in hours follows $W(1000, 1.5)$. If a random sample of 5 such items are put on test under those conditions, compute a 90% probability interval for the life of the fourth failure in this sample.

7. Compute the median life of the product in problem 6. If the product is run-in for 500 hours what is the median of the population of survivors?

8. A sample of ball bearings is drawn from a Weibull population having $\beta = 1.3$ and a tenth percentile of $x_{0.10} = 10.0$ million revolutions The survivors after running for $x_0 = 5.0$ million revolutions will now have what value of the 10th percentile?

9. Using Equations 3.10 and 3.13 verify Equation 3.14.