JOURNAL OF MATERIALS SCIENCE 14 (1979) 1080–1084

On estimating the Weibull modulus for a brittle material

K. TRUSTRUM, A. DE S. JAYATILAKA*
School of Mathematical and Physical Sciences, and School of Engineering and Applied Sciences, University of Sussex, Brighton, UK

Common methods of estimating the Weibull modulus are surveyed. Computer simulation is used to obtain the statistical properties of different estimators. Most estimators are shown to be biased and their respective adjustment factors, for a range of experimentally feasible sample sizes, are given.

1. Introduction

The tensile strength of a brittle material under uniform loading is found both theoretically and practically to follow the Weibull distribution, with probability of failure at stress σ given by

$$P_\mathrm{f} = 1 - \exp\left\{-V\left(\frac{\sigma - \sigma_\mathrm{u}}{\sigma_0}\right)^m\right\}, \qquad (1)$$

where m is a parameter known as the Weibull modulus, V is the volume of the material, σ_0 is a normalizing factor and σ_u is the stress below which there is zero probability of failure. The Weibull modulus, which was previously considered to be an empirical constant, is given a physical meaning by Jayatilaka and Trustrum [1]. It is shown to be related to the flaw size distribution of a brittle material. Thus m is an important material parameter which characterizes the "brittleness" of a material.

The mean of the Distribution 1 is

$$\bar{\sigma} = \sigma_\mathrm{u} + \frac{\sigma_0 \Gamma[1 + (1/m)]}{V^{1/m}}, \qquad (2)$$

where Γ is the gamma function, so Equation 1 can be written in the form

$$P_\mathrm{f} = 1 - \exp\left\{-\left[\Gamma\left(1 + \frac{1}{m}\right)\left(\frac{\sigma - \sigma_\mathrm{u}}{\bar{\sigma} - \sigma_\mathrm{u}}\right)\right]^m\right\}. \qquad (3)$$

It follows from Equation 2 that the mean failure stresses, $\bar{\sigma}_1$ and $\bar{\sigma}_2$, of two specimens of the same material with respective volumes V_1 and V_2 are related by the formula

$$\frac{\bar{\sigma}_1 - \bar{\sigma}_\mathrm{u}}{\bar{\sigma}_2 - \bar{\sigma}_\mathrm{u}} = \left(\frac{V_2}{V_1}\right)^{1/m}. \qquad (4)$$

This is a useful result as the mean failure stress for a large volume V_2 can be estimated from the observed mean failure stress of a sample of specimens of smaller volume V_1, provided that m and σ_u are known for the material. So one problem is how to estimate m and σ_u for a material, given a set of failure stresses $\sigma_1, \sigma_2, \ldots, \sigma_n$ for nominally identical specimens, and a related problem is how to choose the sample size n.

It must be mentioned that many workers have estimated the parameters in the Weibull distribution, but usually they appear to have overlooked the possible inaccuracies which can arise as a result of the sample size or the method of estimation chosen. In this paper some common methods of estimation are analysed.

2. Estimation of m and σ_u

The method of maximum likelihood was used to estimate m and σ_u for a set of 32 experimentally observed flexural strengths of reaction-sintered silicon nitride specimens, tested in three-point bending; the observed values varied between 137.5 and 192.0 MN m^{-2}. Maximum likelihood estimators are approximately normally distributed for large samples and are asymptotically unbiased and minimum variance estimators.

*Now at the Faculty of Engineering, University of Sri Lanka, Pradeniya, Sri Lanka.

The likelihood function for the observed failure stresses $\sigma_1, \sigma_2, \ldots, \sigma_{32}$ is

$$L = f(\sigma_1)f(\sigma_2)\ldots f(\sigma_{32}), \quad (5)$$

where $f(\sigma) = \mathrm{d}P_f/\mathrm{d}\sigma$ is the probability density function, and the maximum likelihood estimates of m, σ_u and $\bar{\sigma}$ are the values of m, σ_u and $\bar{\sigma}$ which maximize L. These values can be found by search methods or by the Newton–Raphson method.

Table I gives the maximum likelihood estimates, m_L and $\bar{\sigma}_L$, for selected values of σ_u. Allowing σ_u to vary as well, the maximum of L occurred at $m = 4.04$, $\sigma_u = 118.5$ and $\bar{\sigma} = 167.16$, which is also included in Table I.

The relationship between σ_u and m_L is observed to be closely linear and this follows from the result that maximum likelihood estimators are asymptotically multivariate normal, so that the expected value of m_L is a linear function of σ_u and $\bar{\sigma}_L$, and $\bar{\sigma}_L$ is almost constant, as would be expected.

Also shown in Table I are the safety factors, $\bar{\sigma}_L/\sigma$, where σ is calculated from Equation 3 for a given value of P_f. The safety factors can be seen to decrease as σ_u increases.

TABLE I Maximum likelihood estimates, m_L and $\bar{\sigma}_L$, for given σ_u, and safety factors, $\bar{\sigma}_L/\sigma$, for $P_f = 10^{-2}, 10^{-4}, 10^{-6}$.

σ_u	m_L	$\bar{\sigma}_L$	$\bar{\sigma}_L/\sigma$		
			$P_f = 10^{-2}$	$P_f = 10^{-4}$	$P_f = 10^{-6}$
0	13.86	166.98	1.34	1.87	2.61
20	12.22	167.00	1.33	1.81	2.40
40	10.56	167.00	1.32	1.75	2.20
60	8.96	167.07	1.31	1.66	1.99
80	7.30	167.11	1.29	1.57	1.78
100	5.63	167.16	1.27	1.46	1.57
118.5	4.04	167.16	1.23	1.35	1.39
130	2.99	167.08	1.20	1.27	1.28

In Fig. 1, the Weibull Distribution 3 is plotted for the cases $\sigma_u = 0$, $m = 13.86$, $\bar{\sigma} = 166.98$ and $\sigma_u = 118.5$, $m = 4.04$, $\bar{\sigma} = 167.16$ (the overall maximum likelihood estimates). Also shown for comparison is the empirical distribution function, $S(\sigma)$, which is the proportion of the observed σ_1, $\sigma_2, \ldots, \sigma_{32}$ which are less than or equal to σ. A

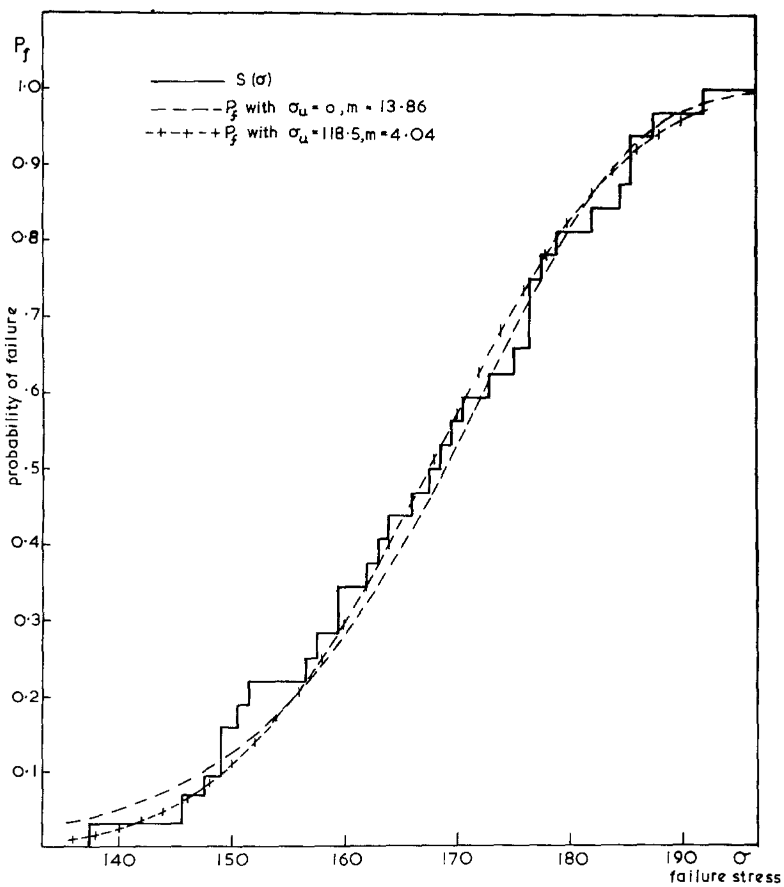

Figure 1 The graphs of $P_f(\sigma)$ given by Equation 3 and the empirical distribution function, $S(\sigma)$.

statistic which tests the goodness of fit of a distribution to the observations is the Kolmogorov–Smirnov statistic,

$$D = \underset{\sigma}{\text{maximum}}\,|P_f(\sigma) - S(\sigma)|. \qquad (6)$$

For the Weibull distribution with $\sigma_u = 0$, $D = 0.075$ and for $\sigma_u = 118.5$, $D = 0.089$, so using this criterion, the Weibull distribution with $\sigma_u = 0$ gives the closer fit. Whatever criterion is used there is clearly not much to choose between the two distributions as regards their goodness of fit to the data. Also the safety factors for $\sigma_u = 0$ are higher than those for other values of σ_u, and σ_u for a material could depend on the volume V. Thus we recommend that for brittle materials $\sigma_u = 0$, both in the interest of safety and for consistency in the definition of the Weibull modulus.

3. Methods of estimating m

In this and the following sections σ_u is assumed to be zero. Several different methods of estimating m have been proposed and some of these methods are discussed.

Putting $\sigma_u = 0$, Equation 1 can be put into the form

$$y = \ln\ln\left(\frac{1}{1-P_f}\right) = \ln\frac{V}{\sigma_0^m} + m\ln\sigma, \qquad (7)$$

so that the graph of y against $\ln\sigma$ is a straight line with slope m. Suppose we have a random sample of the observed failure stresses of n nominally identical specimens subjected to uniaxial tension and that the stresses are ordered so that

$$\sigma_1 \leqslant \sigma_2 \leqslant \ldots \leqslant \sigma_n.$$

Then it can be shown that the expected value of $P_f(\sigma_i)$ is $i/(n+1)$, so one method of estimating m is to use the method of least squares on the linear model

$$\begin{aligned} y_i &= \ln\ln\left\{\frac{1}{1-[i/(n+1)]}\right\} \\ &= \ln\frac{V}{\sigma_0^m} + m\ln\sigma_i + \epsilon_i, \end{aligned} \qquad (8)$$

where ϵ_i is the error arising from the difference between the observed value of $P_f(\sigma_i)$ and its expected value. The least squares estimate of m is

$$m_S = \frac{\sum_i y_i \ln\sigma_i - (1/n)\sum_i y_i \sum_i \ln\sigma_i}{\sum_i (\ln\sigma_i)^2 - (1/n)(\sum_i \ln\sigma_i)^2} \qquad (9)$$

The properties of the estimator m_S are difficult to obtain analytically as the ϵ_i do not satisfy the usual assumptions of the linear model, which are that the errors are uncorrelated with zero mean and constant variance. Under these assumptions the least squares estimator is the minimum variance, linear, unbiased estimator.

Another plotting position for $P_f(\sigma_i)$ is $(i-\frac{1}{2})/n$, which is the average value of the empirical density function, $S(\sigma)$, before and after the jump at σ_i.

A second estimate of m from the Linear Model 8 is

$$m_B = \frac{\sum_{i=1}^{n/3} y_i - \sum_{i=2n/3}^{n} y_i}{\sum_{i=1}^{n/3} \ln\sigma_i - \sum_{i=2n/3}^{n} \ln\sigma_i} \qquad (10)$$

which is obtained by dividing the observations $(\ln\sigma_i, y_i)$ into three equal groups according to the size of σ_i and then m_B is given by the slope of the line joining the centroids of the two extreme groups. This estimate is easily calculated and is recommended for use when the usual assumptions of the linear model are not satisfied [2].

The method of maximum likelihood can also be used to estimate m and although this estimate has optimal properties, a computer is recommended for its calculation. To obtain the equation for the maximum likelihood estimate, m_L, we write,

$$P_f = 1 - \exp(-b\sigma^m), \qquad (11)$$

where $b = \{\Gamma[1+(1/m)]/\bar{\sigma}\}^m$. The probability density function

$$f(\sigma) = \frac{dP_f}{d\sigma} = bm\sigma^{m-1}\exp(-b\sigma^m), \qquad (12)$$

and it follows from Equation 5 that

$$\begin{aligned} \ln L &= n\ln m + n\ln b \\ &\quad + (m-1)\sum_i \ln\sigma_i - b\sum_i \sigma_i^m. \end{aligned} \qquad (13)$$

Differentiating $\ln L$ with respect to m and b and equating the partial derivatives to zero gives the maximum likelihood estimates, m_L and b_L, which satisfy

$$\frac{n}{m_L} + \sum_i \ln\sigma_i - b_L\sum_i \sigma_i^{m_L}\ln\sigma_i = 0, \qquad (14)$$

$$\frac{n}{b_L} - \sum_i \sigma_i^{m_L} = 0. \qquad (15)$$

Eliminating b_L between Equations 14 and 15 gives the following equation for m_L, which can be solved by the Newton–Raphson method:

$$\frac{n}{m_L} - n \frac{\sum_i \sigma_i^{m_L} \ln \sigma_i}{\sum_i \sigma_i^{m_L}} + \sum \ln \sigma_i = 0. \quad (16)$$

Another method of estimating m is the method of moments, which is a plausible method, but the estimates obtained do not, in general have many optimal statistical properties. This method equates the sample moments with the distribution moments, so for the Weibull Distribution 11, the method of moments estimates, m_M and b_M, are given by solving

$$\bar{\sigma} = b^{-1/m} \, \Gamma\left(1 + \frac{1}{m}\right), \quad (17)$$

$$s^2 = \frac{\Sigma(\sigma_i - \bar{\sigma})^2}{n-1} = b^{-2/m}\left[\Gamma\left(1 + \frac{2}{m}\right) - \Gamma^2\left(1 + \frac{1}{m}\right)\right], \quad (18)$$

where s^2 is the sample variance. Eliminating b gives the following equation for m_M,

$$\frac{\{\Gamma[1 + (2/m)] - \Gamma^2[1 + (1/m)]\}^{1/2}}{\Gamma[1 + (1/m)]} = \frac{s}{\bar{\sigma}}, \quad (19)$$

where $s/\bar{\sigma}$ is known as the coefficient of variation. The solutions for m_M can be obtained graphically or by curve fitting, and it can be shown that

$$|m_M - 1.267\frac{\bar{\sigma}}{s} + 0.526| < 0.02$$

for

$$4 \leqslant m_M \leqslant 14. \quad (20)$$

A variety of other methods of estimating m have been proposed, many of them based on order statistics, however, most of them are more complicated and none of the estimators appear to have smaller variance than the maximum likelihood estimator for $n > 30$.

4. Properties of estimators

The statistical properties of the estimators described in Section 3 are difficult to obtain analytically, so simulation was used. Equation 3 can be expressed in the form

$$\sigma = \left[\ln\left(\frac{1}{1 - P_f}\right)\right]^{1/m} \Gamma\left(1 + \frac{1}{m}\right) \quad (21)$$

on putting $\sigma_u = 0$ and $\bar{\sigma} = 1$. Random samples, $\sigma_1, \sigma_2, \ldots, \sigma_n$, were then generated by substituting random numbers in the range 0 to 1 for $1 - P_f$. Care needs to be taken in the choice of method used to generate random numbers, as some of the computer programs for generating random numbers are not satisfactory.

Fortunately, apart from the method of moments estimator, Bain and Antle [3] and Thoman *et al.* [4] have shown that only one value of m needs to be considered for each estimator to obtain its bias and standard error for all values of m. If we put $x = b\sigma^m$ into Equation 11, then x is an exponential random variable with distribution function $1 - e^{-x}$. Also substituting $x = b\sigma^m$ into Equations 9, 10 and 16 gives

$$\frac{m_S}{m} = \frac{\sum_i y_i \ln x_i - (1/n) \sum_i y_i \sum_i \ln x_i}{\sum_i (\ln x_i)^2 - (1/n)(\Sigma \ln x_i)^2}, \quad (22)$$

$$\frac{m_B}{m} = \frac{\sum_{i=1}^{n/3} y_i - \sum_{i=2n/3}^{n} y_i}{\sum_{i=1}^{n/3} \ln x_i - \sum_{i=2n/3}^{n} \ln x_i}, \quad (23)$$

and

$$\frac{n}{m_L/m} - n \frac{\sum_i x_i^{m_L/m} \ln x_i}{\sum_i x_i^{m_L/m}} + \sum_i \ln x_i = 0, \quad (24)$$

respectively, so that the distributions of the estimators m_S/m, m_B/m and m_L/m only depend on the distribution of x and hence are independent of m and b. This property does not hold for m_M/m.

Random samples of sizes $n = 10$, 20, 30, 40 and 50 were simulated from the Weibull Distribution 3 with the parameters $m = 5$, $\sigma_u = 0$, and $\bar{\sigma} = 1$, and for each n, 2600 samples were simulated. The estimates m_S, m_B, m_L and m_M were calculated for each sample using the Equations 9, 10, 16 and 20, and in the case of m_S and m_B the plotting positions $i/(n+1)$ and $(i - \frac{1}{2})/n$ for $P_f(\sigma_i)$ were both used. Table II shows the mean values of m_S/m, m_B/m, m_L/m and m_M/m calculated from 2600 estimates. For an unbiased estimator the mean value should be close to 1. The estimated standard errors of the distributions of m_S/m, m_B/m, m_L/m and m_M/m are shown in brackets.

The Cramer–Rao lower bound for the variance of unbiased estimators of m based on random samples of size n is $0.603m^2/n$ [4], and in the last column of Table II the values of $\sqrt{(0.608/n)}$ are given for comparison with the other standard

TABLE II The estimated means and standard errors of the distributions of the estimators of m for sample sizes n = 10, 20, 30, 40 and 50, obtained from Equations 9, 10, 16 and 20

n	m_S/m		m_B/m		m_L/m	m_M/m	$\sqrt{(0.608/n)}$
	$P_f = \frac{i}{n+1}$	$P_f = \frac{i-\frac{1}{2}}{n}$	$P_f = \frac{i}{n+1}$	$P_f = \frac{i-\frac{1}{2}}{n}$			
10	0.867	1.055	0.944	1.129	1.165	1.102	
	(0.28)	(0.34)	(0.30)	(0.36)	(0.34)	(0.35)	(0.25)
20	0.894	1.109	0.952	1.063	1.078	1.049	
	(0.20)	(0.23)	(0.20)	(0.23)	(0.20)	(0.21)	(0.17)
30	0.910	1.008	0.961	1.039	1.048	1.031	
	(0.17)	(0.19)	(0.16)	(0.17)	(0.16)	(0.17)	(0.14)
40	0.920	1.002	0.964	1.028	1.035	1.021	
	(0.15)	(0.16)	(0.14)	(0.15)	(0.13)	(0.14)	(0.12)
50	0.924	0.996	0.966	1.020	1.025	1.014	
	(0.14)	(0.14)	(0.12)	(0.13)	(0.12)	(0.12)	(0.11)

errors. The results for m_S/m ($n = 10$, 20 and 30) agree closely with those reported by Bain and Antle [3] and those for m_L/m ($n = 10, 20, 30, 40$ and 50) are close to those given by Thoman *et al.* [4].

5. Discussion of results

In designing an experiment to find the Weibull modulus, m, of a brittle material, the experimenter has to decide on how many specimens to use, which involves time and expense, and having observed a sample of failure stresses, he has to choose a method to estimate m. The lower bound of the standard error of the estimate is inversely proportional to the square root of sample size, n, so n has to be increased by a factor of 4 to halve the standard error. Also the standard error is not negligible, for example when $m = 10$ and $n = 50$ the standard error of any unbiased estimate is at least 1.1, which should be taken into account when computing safety factors. As can be seen from Table II, most of the estimators are biased, though the bias is on the whole much smaller than the standard error.

Taking into consideration the ease of calculation of the esimate as well as its bias and standard error, we would recommend the use of the least squares estimate, m_S, with a sample size of about 40 and the plotting position $P_f = (i - \frac{1}{2})/n$. However for sample sizes over 20, the estimate with the smallest standard error is the maximum likelihood estimate.

Table II can be used to adjust an estimate for its bias, for example if $m_L = 11.5$ and $n = 40$, then m should be estimated by $11.5/1.035 = 11.1$ with standard error $11.1 \times 0.13 = 1.4$. Another example is provided by the data described in Section 2 for which $n = 32$, $m_L = 13.86$ and $m_S = 13.16$ [plotting position $P_f = i/(n+1)$]. So using linear interpolation to estimate the bias, the adjusted estimates and their standard errors, shown in brackets, are $13.86/1.029 = 13.5$ (2.0) and $13.16/0.912 = 14.4$ (2.4), respectively, which shows how large the errors can be.

For the least squares estimator, m_S, the plotting position $P_f = (i - \frac{1}{2})/n$ is preferable, as it is the less biased and the coefficients of variation for the two plotting positions are almost equal. There is little to choose between the two plotting positions for the estimator, m_B, and in general m_B is preferable to m_S in that it has a smaller standard error and is less biased. Although the method of moments estimator, m_M, is easy to compute, using graphical or curve fitting methods, and has relatively small bias and standard error, it suffers from the disadvantage that the values given in Table II only hold for $m = 5$.

References

1. A. DE S. JAYATILAKA and K. TRUSTRUM, *J. Mater. Sci.* **12** (1977) 1426.
2. M. G. KENDALL and A. STUART, "The Advanced Theory of Statistics", Vol. 2 (Griffin, London, 1973) p. 420.
3. L. J. BAIN and C. E. ANTLE, *Technometrics* **9** (1967) 621.
4. D. R. THOMAN, L. J. BAIN and C. E. ANTLE, *ibid.* **11** (1969) 445.

Received 9 June and accepted 21 July 1978.