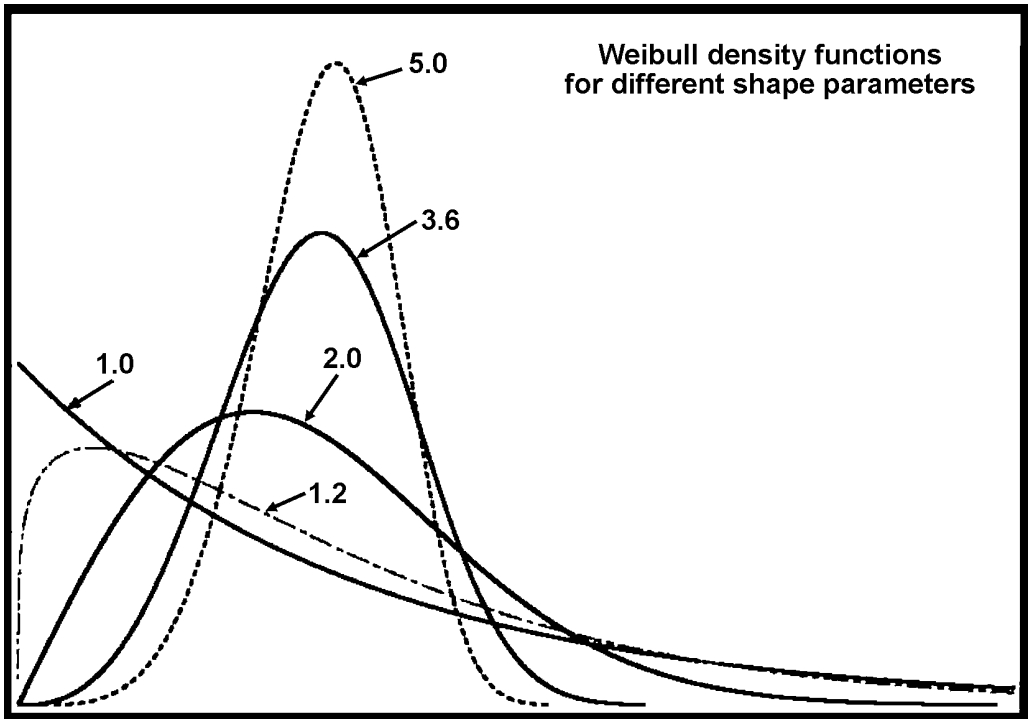
United States Department of Agriculture

Procedures for Estimation of Weibull Parameters

James W. Evans
David E. Kretschmann
David W. Green



Forest Service
Forest Products Laboratory
General Technical Report FPL–GTR–264
February 2019

Abstract

The primary purpose of this publication is to provide an overview of the information in the statistical literature on the different methods developed for fitting a Weibull distribution to an uncensored set of data and on any comparisons between methods that have been studied in the statistics literature. This should help the person using a Weibull distribution to represent a data set realize some advantages and disadvantages of some basic methods. It should also help both in evaluating other studies using different methods of Weibull parameter estimation and in discussions on American Society for Testing and Materials Standard D5457, which appears to allow a choice for the method to estimate the parameters of a Weibull distribution from a data set. Because in D5457 the method to estimate parameters is to some extent optional, the resulting fitted distribution used to derive the reference resistance properties of wood-based materials and structural connections might result in different values developed under the standard. The maximum-likelihood method appears to be the method that should be used as the default, with other methods requiring some type of justification for their use in wood utilization research.

Keywords: Weibull distribution, estimation of parameters, maximum likelihood estimation, regression estimators, simple percentile estimators, method of moments

February 2019

Evans, James W.; Kretschmann, David E.; Green, David W. 2019. Procedures for estimation of Weibull parameters. General Technical Report FPL–GTR–264. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 17 p.

A limited number of free copies of this publication are available to the public from the Forest Products Laboratory, One Gifford Pinchot Drive, Madison, WI 53726-2398. This publication is also available online at www.fpl.fs.fed.us. Laboratory publications are sent to hundreds of libraries in the United States and elsewhere.

The Forest Products Laboratory is maintained in cooperation with the University of Wisconsin.

The use of trade or firm names in this publication is for reader information and does not imply endorsement by the United States Department of Agriculture (USDA) of any product or service.

Contents

1 Introduction 1

2 Background 1

3 Estimation Procedures 1

4 Historical Comparisons of Individual Estimator Types 8

5 Other Methods of Estimating Parameters of the Weibull Distribution 11

6 Discussion 12

7 Conclusion 13

8 References 14

In accordance with Federal civil rights law and U.S. Department of Agriculture (USDA) civil rights regulations and policies, the USDA, its Agencies, offices, and employees, and institutions participating in or administering USDA programs are prohibited from discriminating based on race, color, national origin, religion, sex, gender identity (including gender expression), sexual orientation, disability, age, marital status, family/parental status, income derived from a public assistance program, political beliefs, or reprisal or retaliation for prior civil rights activity, in any program or activity conducted or funded by USDA (not all bases apply to all programs). Remedies and complaint filing deadlines vary by program or incident.

Persons with disabilities who require alternative means of communication for program information (e.g., Braille, large print, audiotape, American Sign Language, etc.) should contact the responsible Agency or USDA's TARGET Center at (202) 720–2600 (voice and TTY) or contact USDA through the Federal Relay Service at (800) 877–8339. Additionally, program information may be made available in languages other than English.

To file a program discrimination complaint, complete the USDA Program Discrimination Complaint Form, AD-3027, found online at http://www.ascr.usda/gov/complaint_filing_cust.html and at any USDA office or write a letter addressed to USDA and provide in the letter all of the information requested in the form. To request a copy of the complaint form, call (866) 632–9992. Submit your completed form or letter to USDA by: (1) mail: U.S. Department of Agriculture, Office of the Assistant Secretary for Civil Rights, 1400 Independence Avenue, SW, Washington, D.C. 20250–9410; (2) fax: (202) 690–7442; or (3) email: program.intake@usda.gov.

USDA is an equal opportunity provider, employer, and lender.

Procedures for Estimation of Weibull Parameters

James W. Evans, Mathematical Statistician
David E. Kretschmann, Research General Engineer (retired[1])
David W. Green, Supervisory Research General Engineer Emeritus
USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin, USA

1 Introduction

Two-parameter and three-parameter Weibull distributions are widely used to represent the strength distribution of structural lumber and engineering-designed wood subassemblies. Current design practices for wood and related material are based on an estimated fifth percentile of the strength distribution. This deterministic method of design is known as allowable stress design (ASD). Although the fifth percentile can be based on a distributional form, such as the Weibull distribution, the real interest in the Weibull distribution is occuring as wood construction practices in the United States and Canada are revised from deterministic procedures to reliability-based design (RBD) procedures. The two-parameter Weibull distribution is the underlying basis of the calculations in load and resistance factor design (LRFD), a subset of RBD that is discussed in American Society for Testing and Materials (ASTM) D5457–04a (ASTM 2009). This design procedure starts with fitting a two-parameter Weibull distribution to either a complete data set or a set of data representing the lower tail of the distribution. The fit can be made using either maximum-likelihood methods or a regression-based estimation procedure. Initially, the standard also considered a method of moments estimate, but this was dropped before the standard was approved. Because of the sensitivity of reliability calculations to the distributional form and, potentially, method of fit, the wood engineering community is interested in looking at (1) evaluating the effect of using a two-parameter instead of a three-parameter Weibull distribution, (2) the effect of estimation method, and (3) the effect of using censored data sets compared with using full data sets. Simulations such as those performed by Durrans and others (1998) encouraged the wood community to address some of these issues. Unfortunately, there is a general lack of knowledge in the wood community of the extensive work available in the statistics literature on these issues, which could lead to needless work in areas in which the information is already available. The primary purpose of this publication is to provide an overview of the basic information in the statistical literature on the different methods developed for fitting a Weibull distribution to an uncensored set of data and on any comparisons between methods that have been studied in the statistics literature. This report is not meant to be a complete summary of all the literature available on the Weibull distribution. It is intended to provide basic information on the more well-known methods for estimating Weibull parameters from data. We start by giving background on the Weibull distribution.

2 Background

The "Weibull" distributional form (Eq. (1)) was first derived through an extreme-value approach by Fisher and Tippett (1928). As noted by Mann (1968), it became known as the Fisher–Tippett Type III distribution of smallest values or as the third asymptotic distribution of smallest (extreme) values. In 1939, a Swedish scientist, Waloddi Weibull (1939a), derived the same distribution in an analysis of breaking strengths using only certain practical requirements. Several examples of its use were given by Weibull (1951). Two other papers by Weibull (1939b, 1952) also used the distribution. Use of the distribution became common in reliability analyses after World War II, and the name Weibull became firmly associated with the distribution.

In its three-parameter form, the Weibull family is represented by the density function,

$$f(x) = ab^{-1}\left\{(x-c)/b\right\}^{a-1}\exp\left[-\left\{(x-c)/b\right\}^{a}\right]$$
$$(x > c; a, b > 0) \qquad (1)$$

where a is the shape parameter, b the scale parameter, and c the location parameter. A more common representation of the Weibull distribution in the wood community is to use the cumulative distribution function,

$$F(x) = 1 - \exp\left[-\left\{(x-c)/b\right\}^{a}\right] \qquad (2)$$

The family of two-parameter Weibull distributions follows from Equation (1) when c = 0.

3 Estimation Procedures

Numerous methods of estimating Weibull parameters have been suggested by many authors. Estimation procedures are generally categorized into one of four major categories: (1) method of moment estimators, (2) linear estimators, such as least-squares-type estimators, (3) estimators based on few order statistics, and (4) maximum-likelihood estimators. There are also a few hybrid estimator systems that combine

[1]now President, American Lumber Standard Committee, Inc., Frederick, Maryland, USA

estimators from more than one category. Zanakis (1979) considered more than 17 combinations of simple Weibull estimators in addition to the maximum-likelihood estimators.

3.1 Method of Moments

The method of moments method of estimation was introduced by Karl Pearson (1894, 1895). The procedure consists of equating as many population moments to sample moments as there are parameters to estimate. Mathematical support for this procedure comes from the principle of moments as discussed in detail in Kendall and Stuart (1969). In essence, this principle says that two distributions that have a finite number of lower moments in common will be approximations of one another. Thus, the distribution of the data is approximated by equating the moments of a distributional form to the data moments.

To see how this could be done with the three-parameter Weibull distribution, and hence the two-parameter Weibull as a special case, let the population kth moments about the origin be given by

$$M_k = \int_{-\infty}^{+\infty} x^k f(x)\,dx \tag{3}$$

For k equal to 1, this is just the expected value or mean of X, denoted $E(X)$. The kth moment about the mean for the population (also called the kth central moment (CM_k)), can be denoted by

$$\mathrm{CM}_k = \int_{-\infty}^{+\infty} (x - M_1)^k f(x)\,dx \tag{4}$$

For a three-parameter Weibull distribution, the first three central moments are

$$\mathrm{CM}_1 = 0 \tag{5}$$

$$\mathrm{CM}_2 = b^2\left(\Gamma_2 - \Gamma_1^{\,2}\right) \tag{6}$$

$$\mathrm{CM}_3 = b^3\left(\Gamma_3 - 3\Gamma_1\Gamma_2 + 2\Gamma_1^{\,3}\right) \tag{7}$$

where

$$\Gamma_j = \Gamma\left[1 + (j/a)\right] \tag{8}$$

and Γ_j is the well-known gamma function. The first three moments about the origin are

$$M_1 = c + b\Gamma_1 \tag{9}$$

$$M_2 = c^2 + b^2\Gamma_2 + 2cb\Gamma_1 \tag{10}$$

$$M_3 = c^3 + b^3\Gamma_3 + 3cb^2\Gamma_2 + 3c^2b\Gamma_1 \tag{11}$$

The corresponding sample moments are

$$m_k = \sum_{i=1}^{n} \frac{(x_i)^k}{n}$$

and

$$\mathrm{cm}_k = \sum_{i=1}^{n} \frac{(x_i - m_1)^k}{n}$$

where n is the number of observations. Note that m_1 is the mean of the observations and cm_2 is the variance of the observations. Johnson and Kotz (1970) and Talreja (1981) pointed out that Equations (6) and (7), along with their corresponding sample moments, are sufficient to estimate the shape and scale parameters of the Weibull distribution. This is done by equating the measure of population skewness given by

$$\frac{\mathrm{CM}_3}{(\mathrm{CM}_2)^{3/2}}$$

which is a function only of a (the shape parameter) to the corresponding sample skewness

$$\frac{\mathrm{cm}_3}{(\mathrm{cm}_2)^{3/2}}$$

Because the population skewness is a function of the shape parameter, finding the value of a that gives the sample skewness requires either a table of a values and the corresponding skewness as is given in Johnson and Kotz (1970) or a computer program that will search for a value of a that gives the sample skewness. Given a, equating CM_2 to cm_2 (the sample variance) gives an estimate of b (the scale parameter). Finally noting that

$$M_1 = c + b\Gamma_1$$

and equating this to m_1 (the mean of the observations) gives an estimate of the location parameter for the three-parameter Weibull. These are sometimes called the method of moments estimates (Johnson and Kotz 1970). Note that it is possible for the estimate of the location parameter c to exceed the smallest observation. Dubey (1966b) proposed an alternative estimator:

$$c = X_{(1)} - \left[b\Gamma_1 / n^{(1/a)}\right] \tag{12}$$

For the two-parameter Weibull distribution, estimation of the shape and scale parameters could be done as before and the location parameter taken as zero. Saylor (1977) proposed an alternative. When $c = 0$ is known, Equations (6) and (8) are sufficient to estimate the shape and scale

parameters. Specifically,

$$\frac{M_1}{\left(\mathrm{CM}_2\right)^{1/2}}$$

is a function of only the shape parameter. When equated to the sample moments

$$\frac{1}{\mathrm{COV}} = \frac{M_1}{\left(\mathrm{CM}_1\right)^{1/2}} = \frac{\text{mean}}{\text{standard deviation}} \tag{13}$$

an estimate of the shape parameter can be obtained. The scale parameter can come from equating the mean (m_1) to the first moment about the origin (M_1).

Note that these choices of moment estimators are not unique. Other choices would also generate estimates of the parameters for either the two- or three-parameter Weibull distribution. However, in general, method of moment estimation has used the smallest moments to provide estimates.

When only a two-parameter Weibull distribution is to be considered, a transformation to the extreme-value distribution is often used to generate moment estimators as follows:

If $X \sim$ Weibull(a,b), i.e., if X has a Weibull distribution given by

$$F(x) = 1 - \exp\left[-(x/b)^a\right]$$

then the variable $T = \ln(X)$ will have an extreme-value distribution with cumulative distribution function

$$F(t) = 1 - \exp\left[-\exp\left\{(t-l)/s\right\}\right] \tag{14}$$

The parameters l and s are location and scale parameters that can be estimated by the first two moments m_1 and cm_2 of the data on this scale (i.e., a mean and variance of the $t_i = \ln(x_i)$ values). The estimation is generally done one of three ways:

Method 1:

$$s = \mathrm{cm}_2^{(1/2)} / S_n$$

and

$$l = m_1 - \left(Y_n / s\right)$$

where S_n and Y_n are numerical constants depending upon the sample size n and are found in Gumbel (1960).

Method 2:

$$s^2 = \left[\mathrm{cm}_2\left(n/n-1\right)\right] / \left[\mathrm{ZM}_2\right]$$

and

$$l = m_1 - \bar{Z}s$$

where

$$Z_i = \ln\left\{\ln\left[1/\left(1-p_i\right)\right]\right\},$$

$$p_i = i/\left(n+1\right),$$

$$\mathrm{ZM}_2 = \sum_{i=1}^{n}\left[Z_i^2 - \left(\sum_{i=1}^{n} Z_i\right)^2 / n\right] / n, \text{ and}$$

$\bar{Z}$ is the mean of the Z_i values.

Mann (1968) noted that a simplified form of method 2 can be derived from the fact that $\bar{Z}$ and ZM_2 are asymptotically equal to Euler's constant $\eta = 0.5772...$ and $\pi/\sqrt{6}$, respectively. The resulting estimates are

Method 3:

$$s = \left(\frac{\sqrt{6}}{\pi}\right)\left\{\mathrm{cm}_2\left[n/n-1\right]\right\}^{1/2}$$

and

$$l = m_i - \eta s$$

These simplified estimates were used in Menon (1963). Given estimates of l and s, estimates of the original Weibull parameters can be obtained from the relationships

$$s = 1/a$$

and

$$l = \ln(b)$$

This approach is also often called the method of moments estimators for the Weibull distribution and will differ slightly in general from the earlier procedures. To help differentiate moment estimators on the extreme-value scale from those on the Weibull scale, some authors (Mann 1968) refer to them as modified moment estimators.

3.2 Least Squares for Two-Parameter Weibull Distributions

The origin of the method of least squares as an estimation technique for Weibull parameters is less clear than that for the method of moments. This is in part because of the fact that several variations of the least squares approach have been proposed for parameter estimation. The general theory of using least squares to obtain estimates of the parameters of a location- and scale-dependent distribution was given by Lloyd (1952, 1962). Although this is different from the approach in the current ASTM standard, it is important to look at the procedure to understand the proposed procedures better. This procedure can be called a weighted least squares procedure.

We will limit our discussion initially to a two-parameter Weibull distribution. Recall that if X has a Weibull distribution given by

$$F(x) = 1 - \exp\left[-(x/b)^a\right]$$

then the variable $T = \ln(X)$ will have an extreme-value distribution with cumulative distribution function

$$F(t) = 1 - \exp\left[-\exp\left\{(t-l)/s\right\}\right] \tag{15}$$

where l and s are location and scale parameters, respectively. Lloyd noted that for any location scale distribution, if we transform the observations by

$$V_r = (T_r - l)/s \tag{16a}$$

to get standardized variables, we get observations of a standardized variable whose distribution is parameter free and thus completely known. Let $V_{(r)}$ denote the rth order statistic from a sample of such standarized variables. Then, its expectation for mean, variance, and covariances can denoted by

$$E\left(V_{(r)}\right) = A_{(r)}$$

$$\mathrm{Var}\left[V_{(r)}\right] = W_{(rr)}$$

and

$$\mathrm{Cov}\left[V_{(i)}, V_{(j)}\right] = W_{(ij)}$$

where i is row and j is column. Then using Equation (16a) and solving for our original order statistics, we can show that

$$E\left[T_{(r)}\right] = l + sA_{(r)} \tag{16b}$$

Equation (16b) is in the format of our usual least squares regression model. If we let D be an $n \times 2$ matrix with ones in the first column and the expected values of the order statistics, $A_{(r)}$, in the second column, then

$$E\left[T_{(r)}\right] = DB$$

where

$$B = \begin{bmatrix} l \\ s \end{bmatrix}$$

is a 2×1 vector of parameters to estimate. Note this is in the form $E(Y) = XB$, which is the usual form for our regression model. From the previous equations, we see

$$\mathrm{Var}\left[T_{(r)}\right] = s^2 W$$

where W is the $n \times n$ variance–covariance matrix with terms $W_{(rr)}$ on the diagonal and $W_{(rs)}$ in the rth row and sth column of the off diagional. Then, the least squares estimate of our parameters B are

$$B_{\text{est}} = \left[d'W^{-1}d\right]^{-1} d'W^{-1}T_{(r)}$$

the ordinary weighted least squares regression. To use it for the extreme-value distribution and hence for the Weibull through transformation, we would take the ordered observations $T_{(r)}$, use the expected value of the order statistics to get the matrix D and, for the distribution in question, find the appropriate weights W. The resulting estimates will be best linear unbiased estimates and thus have very good properties. For example, the estimate of the center of a rectangular distribution from this method will have smaller variance than using the mean of the observations (Lloyd 1952, 1962).

For the extreme-value distribution, and hence the Weibull distribution, however, the appropriate weights were only calculated for very limited sample sizes (generally less than $n = 25$). Some approximating methods existed for generating weights, but these procedures generally stop at $n = 100$ because of computational difficulties.

Recently, several papers have extended this type of procedure. In a recent paper, van Zyl and Schall (2012) proposed using

$$X = \begin{pmatrix} 1 & \log\left(x_{(1)}\right) \\ . & . \\ . & . \\ . & . \\ 1 & \log\left(x_{(n)}\right) \end{pmatrix}$$

$$W = \mathrm{diag}\left(w_1, \ldots, w_n\right)$$

and

$$\hat{\theta} = (X'WX)^{-1} X'Wy$$

where

$$w_r = \frac{n(n-r+1)}{r}\left(\log\left(\frac{n-r+1}{n+1}\right)\right)^2$$

for $r = 1, \ldots n$. Then $\hat{b} = \left(-\hat{\theta}_1 / \hat{a}\right)$ and $\hat{a} = \hat{\theta}_2$. The paper includes simulations where $a = 1.0$ and $b = 1.0$ and then where $a = 1.5$ and $b = 1.0$. These are not typical values for lumber properties. But results for large sample sizes ($n = 100$ and $n = 250$) show very close estimates to the true values.

Several methods exist for obtaining other least-squares-type estimators. Using graphical plotting techniques, Kimball (1960) presented several possible estimators based on the extreme-value scale parameter, which is an inverse of the Weibull shape parameter, including a linear approximation to an unbiased version of the maximum-likelihood estimator of the inverse of the Weibull shape parameter. These methods are based on small complete samples. Other estimators include the "unbiased nearly best linear" estimators of Blom (1958, 1962), a weighted-least-squares estimator of White (1965), and for uncensored data sets only, the estimators of McCool (1965) and Downton (1966). Also similar in some ways to least squares estimates are nonlinear regression estimates, one of which is discussed in Berger and Lawrence (1974).

The methods contained in the ASTM standard are another variation that has been proposed by several authors. That method essentially ignores the weights (i.e., W becomes the identity matrix, and the procedure reduces to ordinary unweighted least squares). This procedure (which we will call ordinary least squares (OLS)) is related to the concept of probability plotting, which is often used for visually assessing the goodness-of-fit of a distribution to a data set. If we plot the ordered observations versus the approximate expected value of the order statistics, the data should lie on a straight line. For the extreme-value distribution, we would plot $T_{(i)}$ versus $\ln\left[-\ln\left(1-p_i\right)\right]$ where $p_i = i/\left(n+1\right)$. Then the slope of the line would be s and the intercept l, which could then be transformed to the Weibull parameter estimates. Alternatively for the two-parameter Weibull, we could also plot $\ln\left(X_{(i)}\right)$ versus $\ln\left[-\ln\left(1-p_i\right)\right]$. The disagreement as to which variable should be considered the independent and which the dependent variable will be discussed later. The slope of this line would be $(1/a)$, and the intercept would be $\ln(b)$.

Using regression procedures to estimate Weibull parameters from this graphical approach is not a new procedure. Miller and Freund (1965) presented the procedure in their textbook without a reference. Gumbel (1954) used it for the extreme-value distribution without a reference. Chernoff and Lieberman (1954), in looking at probability plots for the normal distribution, state the assumption "Let us assume that the visually fitted line is a very good approximation to the line that would be obtained by minimizing the sum of squared deviations (in the x direction) from the line." This assumption is further pursued in Chernoff and Lieberman (1956). Kimball (1960) discusses these two earlier papers and then applies "least squares' theory", i.e., miminizing the sum of squared deviations, in getting estimates for the extreme-value distribution.

Other variations to the basic regression technique include different choices for p_i and the use of robust regression methods, such as L_1 regression, also called absolute deviation regression because we minimize the sum of the absolute values of the deviations. Lawrence and Shier (1981) include L_1 regression in their comparison of estimation techniques, and numerous papers have proposed and evaluated other choices for p_i.

3.3 Simple Percentile Estimators

Zanakis (1979) in a paper doing a simulation study presents some simplified parameter estimates for the three-parameter distribution in a practical way for simulations and scientists. As discussed earlier, if we assume that data come from a Weibull distribution with the cumulative distribution function of Equation (2), then the population percentile x_p given by $p = F(x_p)$ can be determined by

$$x_p = c + b\left[-\ln\left(1-p\right)\right]^{1/a} \tag{17}$$

The paper uses in its simulation an estimate of the percentile of the ith ordered observation using

$$p_i = i/\left(n+1\right)$$

and the corresponding $100p_i$ percent sample percentile t_i by the $y[np_i]$ where [ ] denotes rounding up (if np_i is not an integer). In computer simulations, such rules are needed. There is no evidence to suggest other standard estimates of the percentile of the ith order statistic or different rules related to calculating t_i values would make a difference.

Given three such sample percentiles from the data, which we denote as t_i, t_j, and t_k, which correspond to approximate cumulative probabilities p_i, p_j, and p_k, it is possible to estimate the Weibull parameters by setting up three approximate equations from Equation (17) and then solving system of approximate equations for a, b, and c using

$$\begin{aligned} t_s &= c + b\left[-\ln\left(1-p_s\right)\right]^{1/a} \\ s &= i, j, k \end{aligned} \tag{18}$$

With a little algebra, it follows that for a

$$\frac{t_k - t_j}{t_j - t_i} = \frac{\left[-\ln\left(1-p_k\right)\right]^{1/a} - \left[-\ln\left(1-p_j\right)\right]^{1/a}}{\left[-\ln\left(1-p_j\right)\right]^{1/a} - \left[-\ln\left(1-p_i\right)\right]^{1/a}}$$

If we further choose p_j such that

$$-\ln\left(1-p_j\right) = \left\{\left[-\ln\left(1-p_i\right)\right]\left[-\ln\left(1-p_k\right)\right]\right\}^{1/2} \tag{19}$$

or

$$p_j = 1 - \exp\left\{-\left[\ln\left(1-p_i\right)\ln\left(1-p_k\right)\right]^{1/2}\right\} \tag{20}$$

the estimate of the shape parameter $\hat{a}$ simplifies to

$$\hat{a} = \frac{1}{2}\ln\left(\frac{\left[-\ln\left(1-p_k\right)\right]}{\left[-\ln\left(1-p_i\right)\right]}\right) / \ln\left(\frac{t_k - t_j}{t_j - t_i}\right) \tag{21}$$

Then our estimated scale $\hat{b}$ and location $\hat{c}$ parameters from Equation (18) are

$$\hat{b} = \left(t_s - t_r\right) / \left\{\left[-\ln\left(1-p_s\right)\right]^{1/\hat{a}} - \left[-\ln\left(1-p_r\right)\right]^{1/\hat{a}}\right\} \tag{22}$$

where $\left(r < s\right) \in \left(i, j, k\right)$ and

$$\hat{c} = t_s - \hat{b}\left\{-\ln\left(1-p_s\right)\right\}^{1/\hat{a}} \tag{23}$$

Note, the estimate $\hat{b}$ is simplified if c is known. Then

$$\hat{b} = \left(t_s - c\right) / \left[-\ln\left(1-p_s\right)\right]^{1/\hat{a}}$$
$$s \in (i, j, k) \tag{24}$$

As is the case with any location estimate, it must be less than or equal to the smallest observation, and as with other estimation procedures, there is a possibility that our estimate will not meet this requirement and have to be modified.

Dubey (1967a) studied the two-parameter Weibull simple percentile estimators. When only two parameters need to be estimated, only two percentiles are needed. If we call them t_k and t_i, the equation for the shape parameter simplifies to

$$a = \ln\left(\frac{\left[-\ln\left(1-p_k\right)\right]}{\left[-\ln\left(1-p_i\right)\right]}\right) / \ln\left(\frac{t_k}{t_i}\right) \tag{25}$$

Dubey showed that the asymptotic variance of this estimator is minimized when $p_i = 0.16731$ and $p_k = 0.97366$. This means we should take the sample 17 and 97 percentiles. An estimate of the shape parameter based on these two sample percentiles was shown by Dubey to be 66% efficient compared with the maximum-likelihood estimate of the shape parameter. Normally, 66% efficient is not a good value. But it is simple to calculate when in a field situation and can be used to give an indication of the shape of the distribution. Using these two values of p_i and p_k in Equation (20) would give $p_j = 0.5578$.

Several other estimators based on simple percentiles have been proposed. For any three-parameter Weibull distribution, it is relatively simple to show that the 63rd population percentile is equal to $c + b$ (the sum of the location and scale parameters), i.e.,

$$x_{63} = c + b \tag{26}$$

Thus, for the two-parameter Weibull distribution, the scale parameter can be estimated by the 63rd sample percentile. For the three-parameter Weibull, an estimate of either the scale or location parameter can be inserted into this relationship to get an estimate of the other parameter.

If the shape parameter is small ($0 < a \leq 2$), Dubey (1966a) recommended estimating the location parameter with the smallest order statistic. Dubey (1967b) also proposed a simple estimate of the location parameter based on the first, second, and nth order statistic. This estimator

$$c = \left(X_{(1)}X_{(n)} - X_{(2)}{}^2\right) / \left(X_{(1)} + X_{(n)} - 2X_{(2)}\right) \tag{27}$$

is very close to the first order statistic.

Under the assumption that a good location parameter estimate such as Equation (26) is available, Hassanein (1972) created some asymptotically unbiased shape and scale estimators for which Mann and Fertig (1977) obtained small sample unbiased corrections that were evaluated in Zanakis (1979). Using the fact that the parameters of the extreme-value distribution and the Weibull distribution are related, Bain (1972) proposed for censored samples a simple, unbiased estimator of the scale parameter that is equivalent to the inverse of the shape parameter of the Weibull distribution. These latter estimators had high efficiency for censored samples but zero asymptotic relative efficiency for complete samples. Examples of further use of these estimators with real data has been relatively difficult to find. Englehardt and Bain (1977) developed confidence bounds on Weibull reliability and inference procedures using simplified estimators.

3.4 Maximum-Likelihood Estimation

As pointed out in Harter (1970), the use of the method of maximum likelihood dates back to Gauss (1809), who employed it for some particular problems. However, its general use was first proposed by Fisher (1912). Fisher (1921) began the study of the properties of maximum-likelihood estimators (MLE), which has been continued by numerous researchers. The major justification of maximum-likelihood estimates is usually its large-sample efficiency. Under mild regularity conditions, the MLE of a single parameter from singly censored samples has been shown by Halperin (1952) to be consistent (i.e., converges to the true parameter value as sample size increases), to be asymptotically normally distributed, and to have minimum variance for large samples. Many of these properties can be extended to the case of several parameters and more general censoring.

With all these advantages, one could well ask why not always use maximum-likelihood estimates for the parameters of a Weibull distribution. The problems appear to center around three areas of concern: (1) potential problems in calculating Weibull parameter estimates, (2) the bias of the estimates for small samples, and (3) the possible existence of more efficient and simpler estimates for small sample size. These problems are lessened in many lumber strength properties because the sample size is often quite

large. But, in lumber property studies, many issues can arise from the sampling procedure used.

Efforts to derive MLEs of the parameters of a three-parameter Weibull distribution have received considerable attention in statistical literature. Leone and others (1960) showed the MLE of the scale parameter was a function of the shape and location parameters and the sample values. MLEs for the shape and location were found through an iterative simultaneous solution of two equations involving *a*, *c*, and the sample values. Numerous advances followed, such as allowing various censoring situations (Harter and Moore 1965, 1967; Cohen 1965, 1973; Wingo 1972, 1973). Lemon (1975) discusses some of these earlier advances and then gives a procedure based on the simultaneous solution of two iterative equations that allow both single left and progressive censoring. In getting three-parameter estimates from two equations, Lemon gives two formulas for the location parameter depending if the shape parameter is less than 1 or greater than or equal to 1. Zanakis and Kyparisis (1986) discuss these advances and the computational procedures that have been used to try to solve some of the iteration problems that can develop.

As previously mentioned, MLEs of Weibull parameters have been developed by several researchers. To understand these estimators, it is best to begin with the two-parameter Weibull distribution. In its two-parameter form, the Weibull family is represented by the density function

$$f(x) = ab^{-1}(x/b)^{a-1}\exp\left[-\left((x/b)^{a}\right)\right]$$
$$(x \geq 0; a > 0; b > 0) \qquad (28)$$

where *a* is the shape parameter and *b* the scale parameter.

Consider a random sample of *n* observations. The likelihood function of this sample is

$$L(x_1, \ . \ . \ . \ x_n \ ; a \ , \ b)$$
$$= \prod_{i=1}^{n}(a/b)(x_i/b)^{a-1}\exp\left\{-\left(\frac{x_i}{b}\right)^{a}\right\} \qquad (29)$$

If we take the logarithm of *L* and differentiate with respect to *a* and *b* and then set the derivatives equal to zero, we get two equations:

$$\frac{\partial \ln L}{\partial a} = \frac{n}{a} + \sum_{1}^{n}\ln x_i - \frac{1}{b^a}\sum_{1}^{n}x_i^{\,a}\ln x_i$$

$$\frac{\partial \ln L}{\partial b} = -\frac{n}{b^a} + \frac{1}{b^{2a}}\sum_{1}^{n}x_i^{\,a}$$

Eliminating *b* in the two equations and simplifying gives

$$\left[\frac{\sum_1^n x_i^{\,a}\ln x_i}{\sum_1^n x_i^{\,a}} - \frac{1}{a}\right] = \frac{1}{n}\sum_{1}^{n}\ln x_i$$

which then can be solved through iterative procedures to give an estimate of the shape parameter, denoted by $\hat{a}$. With $\hat{a}$ determined, then

$$\hat{b} = \sum_{1}^{n} x_i^{\,\hat{a}} / n$$

where the symbol (^) denotes that this is an estimate.

Computer routines to produce three-parameter Weibull estimates often show inconsistent convergence. Numerical problems exist, which can make difficult the direct calculation of maximum-likelihood estimates using optimization routines that maximize the likelihood. When the shape parameter is less than or equal to 2, the information matrix of the Weibull distribution has singularities, which means an iterative method that uses second derivatives is likely to be unstable. [This is discussed in Harter (1970) and further in Smith (1985, 1994) with references to Wingo (1973), Rockette and others (1974), Lemon (1975), and Cohen (1975)] In fact, Weibull MLEs are regular if, and only if, at least one of three conditions hold:

(1) $a > 2$,

(2) *c* is known, or

(3) the sample is censored from below (Harter 1970).

(Note: ASTM D5457 uses a two-parameter Weibull and thus meets criteria (2). This means computational problems and problems with asymptotic properties of MLE holding disappear.) For the shape parameter less than or equal to 1, the smallest observation becomes a hyper-efficient estimator of the location parameter (Dubey 1966b), but no true MLE for the location and scale exist. Usually in this case, the smallest observation or some other estimate of the location parameter is used and a pseudo MLE for the shape and scale is produced by assuming *c* is known and reducing the problem to a two-parameter Weibull estimation problem. For $1 < a < 2$, MLEs exist but problems with the information matrix make computational problems possible and make the asymptotic variance–covariance matrix meaningless. For $a = 2$ (the Rayleigh distribution), the determinant of the information matrix is zero and computational problems are possible.

Some concern is periodically expressed about the bias in the MLEs for small samples. It is well known that for small sample sizes, the MLE of the shape parameter can be quite biased. That may or may not be a problem for reliability calculations. For complete samples, Thoman and others (1969) present a table of unbiasing factors for the shape parameter for sample sizes from $N = 5$ to 120. The factors are multiplied by the MLE of the shape parameter to get an unbiased estimate. The following selected entries from the

tables give an indication of the amount of bias for different sample sizes (N):

N	5	6	7	8	10	14	28	44	64	120
Factor	0.669	0.752	0.792	0.820	0.859	0.901	0.951	0.970	0.980	0.990

This bias means that you should multiply the shape parameter by the bias number. The biasing number approaching 1 shows that the estimate is asymptotically consistent (because MLEs are consistent, and for consistent estimators, the bias will approach zero as the sample size increases). Similar tables are given in Billmann and others (1972) in which either 25% or 50% of the largest observations were censored for sample sizes $N = 40, 60, 80, 100, 120$. These results show that the greater the censoring, the greater the bias.

In spite of the potential problem of bias of the shape parameter, two separate studies indicated that functions estimated using estimates of the shape and scale parameters often show little bias. Al-Baidhani and Sinclair (1987) compared several estimators of percentiles of the distribution. Using sample sizes $N = 10, 25$; shape parameter $a = 0.5, 1.0, 2.0, 4.0$; and scale parameter $b = 0.1, 1.0, 10.0, 100.0$, they compared the estimation procedures on the parameter estimates and on their ability to predict the 95th percentile and the 99th percentile. In estimating percentiles, the authors concluded the "MLE is clearly the best method for all the values of (the shape parameter)". Billman and others (1972) looked at the problem of estimating the reliability at time t, which for a two-parameter Weibull distribution is estimating

$$R(t) = \exp\left\{-\left(t/b\right)^a\right\}$$

They concluded "It is fortunate that in spite of skewness, censoring and other difficulties, the MLE of $R(t)$ for reasonable values of $R(t)$ has negligible bias and its variance is very close to the Rao-Cramer lower bound for the variance of a regular unbiased estimator of $R(t)$."

Because in LRFD we are estimating functions of the shape and scale parameters, three options appear possible. We can use the unbiasing factors, assume that the compensation the scale parameter makes to the shape parameter will negate any effect of the bias when estimating functions of both, or do simulations to check if there is a problem.

The potential problem of the possible existence of more efficient and simpler estimates for small sample sizes is not so easily answered. There are certainly some more efficient estimators for small samples than the MLE, as has been shown in numerous studies. However, our evaluation of the literature indicates that they are not method of moment estimators or OLS estimators. More work needs to be done to identify which of these other estimation procedures works well as a set and produces efficient estimators for Weibull distribution characteristics such as the 5th percentile and the functions of the parameters needed in LRFD calculation procedures.

Any discussion of maximum-likelihood estimation of Weibull parameters would be incomplete without at least mentioning the existence of some simple, closed form approximations for MLEs of the parameters of the Weibull distribution. Bain (1972) and Engelhardt and Bain (1973, 1974) developed simple estimators for the parameters of the extreme-value distribution. These relatively efficient, unbiased estimates can easily be converted through the relationship between the extreme-value distribution and the Weibull distribution to provide estimates of the Weibull distribution parameters. Their work showed that the simple estimators were related to maximum-likelihood estimators.

4 Historical Comparisons of Individual Estimator Types

Dubey (1966a) looked at the asymptotic (large sample) efficiency of moment estimators for the Weibull location and scale parameters assuming the shape parameter was known. This efficiency (defined as a ratio of the MLE generalized variance to the moment estimator generalized variance) for shape parameters from $a = 2.1$ to $a = 100$ had values ranging from 12% to 93%. A few specific values are shown in Table 1.

Talreja (1981) generated five random samples of size $n = 10$ from each of two two-parameter Weibull distributions and used the method of moments described first to estimate three-parameter Weibull distribution estimates and two-parameter Weibull distribution estimates. Talreja concluded that the method could not be relied on for estimating all three parameters, especially for his example where the shape

Table 1—A few specific values for shape parameters

Shape	Efficiency of location and scale (%)
2.1	11.98
2.6	54.63
3.1	76.60
3.6	87.18
4.1	91.75
4.6	93.18
5.1	92.96
5.6	91.89
6.1	90.39
6.6	88.72
7.1	87.00
10.0	78.26
20.0	63.96
100.0	51.75

parameter was large ($a = 10$). He also noted that in fitting a two-parameter Weibull, if a three-parameter Weibull distribution was used to generate the data (i.e., the location parameter was not zero), errors in estimating the parameter increased linearly with the true value of the location parameter, to the point that errors were more than 100% when the true location parameter was equal to or greater than the true scale parameter.

Saylor (1977) also used a very small simulation with shape parameters less than 2 to compare method of moment estimators with maximum-likelihood estimates for a three-parameter Weibull. As an estimate of the location parameter, Saylor used Dubey's estimate given in Equation (12) and the estimation procedure based on the mean and coefficient of variation previously described. This required iteration back and forth between the two methods until estimates remained constant. Saylor concluded that the location parameter estimate of Dubey was more efficient than the maximum-likelihood estimate of the location parameter and suggested use of Dubey's estimate, iteratively, with the MLE for scale and shape. He also noted that the moment estimators appeared to be more efficient than MLEs for sample sizes of less than 50. However, he cautioned that further experimentation with other types of generators and larger Monte Carlo sample sizes was needed.

Newby (1984) used the method of moment estimators for the three-parameter Weibull derived using the mean, variance, and skewness. He compared efficiency of the method of moment estimators to the MLEs as a function of the true shape alone. He could only look at shape parameters greater than 2 because the asymptotic information matrix of the maximum-likelihood estimators has singularites when the shape is less than 2 (note this is why the use of iterative methods to get MLEs that are based on second derivatives are often unstable). Newby's results are summarized in the table he produced (Table 2).

Mann (1968) looked at several estimators for the parameters of the two-parameter Weibull distribution including these modified moment estimators. Mann concluded that these estimators gave poor results for estimating s (and hence a) for small as well as large samples compared with other better estimation procedures including maximum likelihood.

Table 2—Joint asymptotic efficiency of estimators for the Weibull

Shape parameter	Efficiency of moment estimators (%)
2.1	3
3.0	41
4.0	71
5.0	85
6.0	89
7.0	89
8.0	88
9.0	86
10.0	83

With such a computationally easy method of estimating parameters available, compared with maximum-likelihood procedures, a natural question is why is a regression procedure not more widely used. There are several reasons. A subtle reason deals with the loss of unbiasedness when unweighted estimates replace weighted estimates. Bias in estimates does not eliminate their use. Statistics is full of biased estimates being used in place of unbiased. The commonly used estimate of a standard deviation is one example. The argument for use of a biased estimate is usually that the bias is small and the variance so much smaller than the unbiased estimate that we will usually be closer to the true value than with the unbiased estimate. This argument is incorporated in the concept of mean square error (MSE) of an estimator. Mathematically,

$$\text{MSE} = \text{Variance} + \text{Bias}^2$$

The square root of the MSE of an estimator is similar to a standard deviation except it incorporates bias. Use of a biased estimator usually involves showing that the MSE or square root of the MSE is smaller for that estimator than for other estimators. Thus, in evaluating least squares estimates of Weibull parameters, we must look at their MSE.

Kimball (1960) looked at the OLS estimator for small samples and different choices of p_i for the extreme-value distribution. His results for s are shown in Table 3.

Thus, the choice of p_i appears important. In the current ASTM standard, $p_i = (i - 0.3)/(n + 0.4)$ is used.

Properties of the OLS procedure for complete samples versus other estimation techniques have been studied by several researchers in addition to Kimball. Berger and Lawrence (1974) compared the OLS estimator with $\ln\left[X_{(i)}\right]$ considered as the fixed points (independent variables) and $\ln\left(-\ln\left(1-p_{(i)}\right)\right)$ as the random variables, as was done in Miller and Freund (1965) and as they said was "common industry practice". This was compared with using nonlinear regression to fit the model on the original scale for sample size $n = 50$, shape parameters 0.5 to 4.0 by 0.5, and scale parameters 2, 4, 6, 8, and 10. 100 replications of each combination were simulated. Berger and Lawrence only considered $p_i = i/(n + 1)$. Their results showed that both procedures produced results with large MSE. They speculated that the reversal of the variables from

Table 3—Kimball (1960) results for *s*

Value of p_i	Bias	Mean square error
1. $i/(n + 1)$	0.2284	0.3011
2. $[i - (3/8)]/[n + (1/4)]$	0.0220	0.1754
3. $[i - (1/2)]/n$	–0.0547	0.1543

considering $\ln\left[X_{(i)}\right]$ as the dependent variable and $\ln\left(-\ln\left(1-p_{(i)}\right)\right)$ as the independent variable contributed to the large variability.

Lawrence and Shier (1981) extended the earlier work of Berger and Lawrence (1974). Both the earlier OLS estimator with $\ln\left[X_{(i)}\right]$ considered as the independent variable and $\ln\left(-\ln\left(1-p_{(i)}\right)\right)$ as the dependent variable and the OLS with the roles reversed (the ASTM procedure) were compared with least absolute deviation (L_1) regression in which the sum of the absolute values of the deviates is minimized. The L_1 regressions considered both roles for the variables as did the OLS estimators. The values of p_i considered were

(1) $(i - 0.5)/n$

(2) $(i - 0.375)/(n + 0.25)$

(3) $(i - 0.3)/(n + 0.4)$ (ASTM procedure)

Sample sizes $N = 20, 30, 40, 50, 100$; shape parameters of 0.5, 1.0, 3.0, 5.0, 10.0; and a scale parameter of 1.0 were considered. Again, 100 replications of each combination were generated. Lawrence and Shier (1981) concluded that using $\ln\left[X_{(i)}\right]$ as the independent variable and $\ln\left(-\ln\left(1-p_{(i)}\right)\right)$ as the dependent variable produced more efficient estimates of the shape parameter. This is how the ASTM procedure is done. They also concluded that L_1 regression was 10% to 30% more efficient in estimating the shape than was the OLS estimator. Estimation results for the scale parameter were mixed with no apparent winner. They also concluded that the number 2 definition of p_i was the most consistent performer.

Engeman and Keefe (1982) compared the simple moment estimates of Menon, maximum-likelihood estimates, ordinary least squares estimates, and the "unbiased nearly best linear" estimates of Blom (1958, 1962) for the two-parameter Weibull for sample size $n = 25$; true shape parameters 0.5, 1, 2, and 4; and true scale parameters 0.1, 1, 10, and 100. They used $p_i = i/(n + 1)$ for the OLS estimator. The number of replications of this simulation were not reported. To standardize the MSEs of the estimators for the different parameterizations, they reported the ratio of the Cramer–Rao lower bounds for the variance of unbiased estimates of the two-parameter Weibull to the observed MSE of the estimator. The Cramer–Rao lower bounds are calculated from

$$\mathrm{Var}(a) = 0.608a^2/n$$

and

$$\mathrm{Var}(b) = 1.109(b/a)^2/n$$

The average relative efficiencies they reported were as follows:

Blom	
OLS	
Menon	
MLE	
Shape parameter a	0.8453
	0.6275
	0.4840
	0.6862
Scale parameter b	1.0207
	0.9903
	1.0187
	1.0483

These results show the relative efficiency of the simple moment estimators to be substantially below those of all other estimators of the shape parameter. For this sample size, the best estimate of shape was that of Blom and for scale the MLE. The MLE was slightly better than the OLS for both parameters.

Shier and Lawrence (1984) compared the OLS estimator and the L_1 estimator with a number of reweighting schemes based on the L_1 residuals. Using $p_i = (i - 0.5)/n$; sample sizes $N = 20, 30, 40, 50, 60$; shape parameters $a = 0.5, 1.0, 3.0, 5.0, 10.0$; and scale parameter $b = 1$, they generated 100 replications for each combination. They again considered $\ln\left[X_{(i)}\right]$ as the independent variable. Their results showed the OLS estimator to be "clearly inferior to all of these five L_1-based procedures. ... On average, the five procedures are some 14% more efficient than the (OLS) procedure".

Al-Baidhani and Sinclair (1987) compared a generalized least squares (GLS) estimate that takes into account the variance–covariance matrix, the OLS estimator, the MLE, a method based on Hazard plotting position, and two mixed methods. Using $p_i = i/(n + 1)$; sample sizes $n = 10, 25$; shape parameter $a = 0.5, 1.0, 2.0, 4.0$; and scale parameter $b = 0.1, 1.0, 10.0, 100.0$, they compared the estimation procedures on the parameter estimates and on their ability to predict the 95th percentile and the 99th percentile. These percentiles were chosen to look at problems engineers might have in predicting the size of floods that occur only once in 100 years. The number of replications was 1,000. Results showed that the highest average relative efficiency for the shape parameter was the GLS estimator, whereas the OLS estimator had the lowest average relative efficiency. For the scale parameter, the MLE was better than the OLS for every case considered and either best or second best in every case compared with all the estimators. In estimating percentiles, the authors concluded that "MLE is clearly the best method for all the values of (the shape parameter)". Looking at

parameter and percentile estimation combined, they concluded "the worst of these methods in every situation was OLS".

Durrans and others (1997), using the two-parameter and three-parameter Weibull distribution on censored data, looked at several methods of estimating quantiles and parameters that would be used in quality control in describing structural lumber failure strengths. They investigated MLEs and regression estimators. On the basis of their simulations, they concluded that "the method of maximum likelihood is nearly universally superior to a regression method . . .". The advantage of regression methods is that they will always lead to a solution, whereas maximum-likelihood methods may not converge in some cases.

When the results of all these studies are taken together, the ASTM standard appears to have several weaknesses. It does not appear to use the most efficient choice of p_i. They are less efficient than other regression type estimators that are easily calculated. The direct comparisons to MLEs showed superiority of the MLE for even a small sample size.

5 Other Methods of Estimating Parameters of the Weibull Distribution

Previously, the most common methods of estimation of Weibull parameters have been discussed. Each of the methods covered has had extensive discussion in statistical literature. There are a few other methods that have been raised that should be discussed. These methods might benefit from more study, but they help provide a broader view than just the most common methods.

5.1 Idea of Mixing Types of Methods

Initially, estimation of Weibull parameters was generally restricted to picking one method of estimation and then using that method to estimate all the parameters of the Weibull distribution. There are several sound reasons for considering this approach that will be covered next. It has already been mentioned that in some cases the estimate of the location parameter for a three-parameter Weibull distribution can be above the lowest order statistic. For some people who are dealing with strength properties of materials, this can be a problem. Many times, computer programs calculating maximum-likelihood estimates of Weibull parameters will substitute the lowest order statistic for the location parameter, which follows Dubey's (1966a) recommendation when the shape parameter is small ($0 < a \leq 2$). When the estimated shape parameter is greater than 2, Dubey's (1967b) estimate in Equation (26) is sometimes used.

The idea of mixing types of estimators was studied in Zanakis (1979). In a simulation study, Zanakis studied how various simple estimators for the shape, scale, and location parameters compared regarding their accuracy in a maximum-likelihood sense. He concludes that the best estimators are Equation (27) for location, Equation (26) for scale where $\hat{c}$ is put in c, and Equation (25) for shape where $p_i = 0.16731$ and $p_k = 0.97366$. He furthermore stated that these three estimates proved to be more accurate than maximum-likleihood estimates for all three parameters, especially when $a < 2$ (Zanakis 1977, 1978).

Some programs used to calculate maximum-likelihood estimates benefit from an initial estimate of the parameters. When this is the case, an initial estimate of the parameters using Zanakis' recommendations might prove useful.

5.2 Minimum Chi-Square Estimation

As an alternative to maximum-likelihood estimation, Berkson (1980) proposed using minimum chi-square estimation in response to a previous publication (Efron 1975). Berkson (1980) suggested that minimum chi-square estimation "yields the same estimating equations as MLE". A spirited discussion followed this paper. Recently, a paper by Barbiero (2016) looked at estimating the parameters of a two-parameter discrete Weibull distribution. Discrete Weibull distributions occur when observations can only be found in categories, such as when measurements must be rounded off into units or a range of units, perhaps because of the inability to get more accurate measurements. The paper looks at least-squares methods and minimum chi-square methods. Using a simulation, Barbiero (2016) concludes "The simulation results indicate that the performance of the three point estimators based on the least-square method is … overall better than that derived through the minimum chi-square method."

5.3 Generalized Least Squares Estimation

Weighted regression and ordinary least squares regression, which were previously covered, are the most common methods of regression used to estimate the parameters of a Weibull distribution. Two other methods will be briefly mentioned in this section and the following section. They are GLS estimation and least absolute deviation estimation. GLS is a variation of regression that is sometimes used when a correlation is expected between the residuals of the model. It was first described by Aitken (1934). The GLS procedure is a more generalized version of a weighted least squares procedure in which the weights are not only on the diagonal of the weighting matrix. Engeman and Keefe (1982) described how to use GLS to estimate the parameters of the two-parameter Weibull distribution. Using a simulation study with very small sample sizes, they found that the GLS estimate of the shape parameter was better than ordinary least squares, maximum likelihood, and the

estimation procedure of Menon (1963). Maximum likelihood provided the best estimate of the scale parameter. Kantar (2015) compared GLS, weighted least squares, and maximum likelihood for the Weibull, log-logistic, and Pareto distributions. Again, a simulation found that GLS provided the best estimate of the shape parameter. Kantar (2015) also provided a review of some other papers using GLS and some real life examples.

5.4 Least Absolute Deviation Estimation

Once regression estimates of Weibull parameters became popular, the logical thing to do was to look at different regression techniques to see if they would solve some potential problems. The use of minimizing the sum of the least absolute deviations

$$S = \sum_{i=1}^{n} \left| y_i - \bar{y}_i \right|$$

This is one of several methods called robust regression because they have favorable properties when in the typical regression analysis, the residuals are not i.i.d. (independently and identically distributed) normally distributed with the same mean and standard deviation. As with GLS, this method is trying to handle a deviation from the assumption of independence of the residuals. This method of least absolute deviation regression, also sometimes called L_1 regression, is trying to handle residuals that might have skewed distributions (i.e., not normally distributed). Lawrence and Shier (1981) compared least square estimates (sometimes called L_2 regression) and least absolute deviation regression estimation of Weibull parameters. They concluded that in some cases L_1 regression can be superior to least squares. Again in 1984, Shier and Lawrence (1984) compared robust regression procedures on the estimation of Weibull parameters. In that paper, they also compared using simulations with several reweighting schemes based on L_1 residuals.

6 Discussion

Initially, estimation of Weibull parameters was generally restricted to picking one method of estimation and then using that method to estimate all the parameters of the Weibull distribution. There are several sound reasons for considering the approach that will be covered next. But the general increase in simple percentile estimates might change this approach as methods become more numerous. You can now mix and match estimators. For example, a simple percentile estimator of the location parameter followed by maximum-likelihood estimates of the shape and scale parameters might sound like a good idea because the maximum-likelihood estimation of the location parameter would often cause nonconvergence of the shape and scale parameters, especially when the shape parameter was near 2 or less than 2. For example, in the U.S. in-grade study of Hem-Fir maximum-likelihood shape parameter values for 13 grade–size combinations of specimens with sample sizes ranging from 20 to 428 at 15% moisture content, the shape parameter estimates for the three-parameter Weibull distribution ranged in value from 1.38 to 4.15 (Evans and Green 1988). There is also the important question of how the estimation procedure might affect the estimate of properties such as the 75% tolerance limit of the fifth percentile of a strength distribution for modulus of rupture, which is used to quantify the strengths of various wood species used in wood construction. Therefore, the problem becomes not only which method of estimating Weibull parameters is best but also which method might provide the best estimate of distributional properties and possibly properties such as confidence limits on these property estimates or goodness-of-fit tests on the fitted distribution. Most earlier work looked at parameter estimation. But, it is time to consider all the benefits that different methods provide.

Looking first at the benefits of using all estimators of the same kind, we consider the case for MLEs for all the parameters of a two-parameter Weibull distribution, which have most of the benefits that a method of estimation can provide. Thoman and others (1969) discussed four major properties that are important:

(1) Exact confidence limits for the parameters are available.

(2) A table of unbiasing factors is available for several sample sizes.

(3) Tests of hypotheses about the parameters and the power of the tests are developed.

(4) Guidelines of when sample sizes are large enough to assume large sample theory are available.

Johnson and Haskell (1983) considered parameter estimates of the three-parameter Weibull and showed the consistency of the shape parameter estimate when the parameter was greater than 1. They also studied the joint distribution of the estimators and determined that estimates of the 5th percentile of the population from the sample 5th percentile required sample sizes of more than 70 to assume that asymptotic normality applied. Evans and others (1989) provided goodness-of-fit tests for both two- and three-parameter fits of Weibull distributions to data. Johnson and others (2003) provided confidence intervals on ratio estimates of Weibull 5th percentiles that would be useful in evaluating dry–green ratios.

Other major individual types of estimators have fewer properties that are discussed in statistical literature. Mixtures of parameter estimator types generally have even fewer known properties, and that can be a problem when trying to discuss what these types of sets of estimators show. That does not mean that they are useless because they can provide important insights into data. However, the lack of

major statistical properties can limit the usefulness of estimates with a mixture of methods for the different parameters when used on a data set.

There are a few papers that have looked at properties that other estimation procedures might have in comparison with MLEs. They generally compare the estimators on issues such as bias, variation, ease of calculation, and other measures that do not include the major properties previously discussed here. It might be useful for people looking at tail properties of a distributional fit to have a study that compares various methods. In wood engineering, some data sets used for grades of visual lumber and machine-stress-rated lumber may be from whole populations in which lower or higher grades were taken out in the process of creating multiple grades.

MLEs, for some combinations of parameter values of the three-parameter Weibull distribution, do not satisfy the regularity conditions, which give estimates of the standard asymptotic properties, resulting in possible inconsistent estimates or even failure to exist. In these cases, some researchers such as Cohen and Whitten (1982) have suggested using some modified estimators. The conditions for which MLEs are regular were previously covered. In cases for which the conditions don't apply, Cohen and Whitten note that the mean μ_x, variance σ_x^2, median Me_x, and the skewness $\alpha_{3:x}$ are

$$\mu_x = c + b\Gamma_1$$

$$\sigma_x^{\ 2} = b^2\left[\Gamma_2 - \Gamma_1^{\ 2}\right] = \mathrm{CM}_2$$

$$\mathrm{Me}_x = c + b\left(\ln 2\right)^{1/a}$$

$$\alpha_{3:x} = \frac{\mathrm{CM}_3}{\left[\mathrm{CM}_2\right]^{3/2}}$$

where Γ_k is defined in Equation (11) and the central moments are defined in Equations (6) and (7). Although the Weibull distribution is usually thought of as being positively skewed, it is really negatively skewed when $a >$ 3.6023494257197 (Cohen 1973). This is basically what was covered in the section on method of moments. Cohen suggests using modified moment estimators and shows through simulation under different conditions that the modified moment estimators may prove useful in some cases. Cohen concludes that for $a \geq 2.2156$ ($\alpha_3 < 0.5$) MLEs perform well with respect to bias and variance. For $a = 4.5$, both MLE and moment estimators perform well with the MLEs taking considerably more time to calculate. (With modern increase in computer power, this is not the case anymore, but in the field, the moment estimators might be useful as initial estimates when computers are not available). When $a < 2.2156$, using modified moment estimators of a may prove useful with respect to bias and error. Cohen offers three different modifications and evaluates them through simulations and ease of calculation.

Gross and Lurie (1977) compared two estimators from Bain and Antle (1967) of the shape parameter with the MLE of the shape for the two-parameter Weibull distribution when $a < 1$. The comparison was based on bias, standard deviation, and root mean square error. This simulation showed that the Bain and Antle estimators were better for small samples and the MLE was superior for larger samples. (For wood engineers, $a < 1$ is unlikely unless there are outliers that have extremely large values.)

Verrill and others (2012) compared regression estimators and MLEs for the parameters of a two-parameter Weibull distribution. The four coefficients of variation considered were 0.1, 0.2, 0.3, and 0.4. The corresponding shape values were 12.154, 5.7974, 3.7138, and 2.6956. They concluded that regression estimators were competitive for sample sizes of 15. But for sample sizes of 30 and larger, the MLE estimates were superior, and they were even better for the shape parameter.

7 Conclusion

Procedures for estimating the parameters of a Weibull distribution from a data set have been widely studied in the statistical literature and are still a major source of study. Any review of such procedures is bound to be incomplete. There are many more studies of the aspects of this problem than can be summarized in a short overview such as this report. However, a general overview of the procedures and their properties as found in the statistical literature can form a background for a more specific comparison of the procedures with uncensored data, which is often encountered in wood utilization research. To summarize a general interpretation of the information presented, the authors would like to stress two points. There are a wide variety of procedures and adaptations of procedures that are available to estimate the parameters of a Weibull distribution. Any comparison of these methods for use on wood engineering has to assume that the data truly has a Weibull distribution, which may not be true for censored data sets such as machine-stress-rated data in which different grades that overlap machine-stress-rated data may have been removed before grading the remaining specimens. Recent research has suggested that these grades of lumber may not follow a Weibull distribution. The second possible conclusion is that, although there is no uniformly "best" method of parameter estimation, the use of maximum-likelihood methods for data that wood scientists are likely to encounter appears to be the method that wood scientists should be using as a default. Other methods should require some type of justification for use in wood utilization research.

Based on this conclusion of the advantage of maximum-likelihood estimates for Weibull data and the distinct possibility that lumber grade data might be a truncated data set, ASTM D5457 appears to have several potential weaknesses that should be studied further. Because the choice for estimating parameters of a Weibull distribution from a data set is optional in D5457 and the resulting fitted distribution is used to derive the reference resistance of wood-based materials and structural connections, there might be an issue in determining the properties of the values developed. This could be an important research topic to further explore. Also, for small sample sizes for tests of glulam beams or wood timbers, other methods that don't have the bias that MLEs can have with small sample sizes may be useful to check on the estimates prescribed in D5457 using plots of the data verses predicted measurements from the distributional estimates.

8 References

Aitken, A.C. 1934. On least-squares and linear combinations of observations. Proceedinmgs of the Royal Society of Edinburgh, 55, 42-48.

Al-Baidhani, F.A.; Sinclair, C.D. 1987. Comparison of methods of estimation of parameters of the Weibull distribution. Communications in Statistics B. Simulation and Computation. 16: 373-384.

ASTM. 2009. D5457-04a. Standard specification for computing reference resistance of wood-based materials and structural connections for load and resistance factor design. In: Annual book of standards. Philadelphia, PA: American Society for Standards and Materials. 5546-5554.

Bain, L.J. 1972. Inferences based on censored sampling from the Weibull or extreme-value distribution. Technometrics. 14: 693-702.

Bain, L.J.; Antle, C.E. 1967. Estimation of parameters in the Weibull distribution. Technometrics. 9(4): 621-627.

Barbiero, A. 2016. Least-squares and minimum chi-square estimation in a discrete Weibull model. Communications in Statistics – Simulation and Computation. 1-21.

Berger, R.W.; Lawrence, K.D. 1974. Estimating Weibull parameters by linear and nonlinear regression. Technometrics. 16: 617-619.

Berkson, J. 1980. Minimum chi-square, not maximum likelihood. The Annals of Statistics. 8(3): 457-487.

Billmann, B.R.; Antle, C.E.; Bain, L.L. 1972. Statistical inference from censored Weibull samples. Technometrics. 14: 831-840.

Blom, G. 1958. Statistical estimates and transformed beta variables. New York: John Wiley & Sons.

Blom, G. 1962. Nearly best linear estimates of location and scale parameters. In: Sarhan, A.E.; Greenberg, B.G., eds. Contributions to order statistics. New York: John Wiley & Sons: 34-46.

Chernoff, H.; Lieberman, G.J. 1954. Use of normal probability paper. Journal of the American Statistical Association. 49: 778-785.

Chernoff, H.; Lieberman, G.J. 1956. The use of generalized probability paper for continuous distributions. Annals of Mathematical Statistics. 27: 806-818.

Cohen, A.C. 1965. Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. Technometrics. 7: 579-588.

Cohen, A.C. 1973. Multi-censored sampling in the three parameter Weibull distribution. Presented to the International Statistical Institute, Vienna, Austria, August 1973.

Cohen, A.C. 1975. Multi-censored sampling in the three parameter Weibull distribution. Technometrics. 17: 347-351.

Cohen, A.C.; Whitten, B. 1982. Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. Communications in Statistics A. Theory and Methods. 11: 2631-2656.

Downton, F. 1966. Linear estimates of parameters in the extreme-value distribution. Technometrics. 8: 3-17.

Dubey, S.D. 1966a. Asymptotic efficiencies of the moment estimators for the parameters of the Weibull laws. Naval Research Logistics Quarterly. 13: 265-288.

Dubey, S.D. 1966b. Hyper-efficient estimators of the location parameter of the Weibull laws. Naval Research Logistics Quarterly. 14: 253-264.

Dubey, S.D. 1967a. Some percentile estimators for Weibull parameters. Technometrics. 9: 119-129.

Dubey, S.D. 1967b. On some permissible estimators of the location parameter of the Weibull and certain other distributions. Technometrics. 9: 293-307.

Durrans, S.R.; Triche, M.H.; Taylor, S.E.; Woeste, F.E. 1997. Parameter and quantile estimation for the distributions of failure strength of structural lumber. Forest Products Journal. 47: 80-88.

Durrans, S.D.; Triche, M.H.; Suddarth, S.K. 1998. Estimation of lower tail quantiles of Weibull probability distributions for lumber strength. Forest Products Journal. 48: 96-101.

Efron, B. 1975. Defining the curvature of a statistical problem (with application to second order effeciency). The Annals of Statistics. 3(6): 1189-1242.

Engelhardt, M.; Bain, L.J. 1973. Some complete and censored sampling results for the Weibull or extreme-value distribution. Technometrics. 15: 541-549.

Englehardt, M.; Bain, L.J. 1974. Some results on point estimation for the two-parameter Weibull or extreme-value disribution. Technometrics. 16: 49-56.

Engelhardt, M.; Bain, L.J. 1977. Simplified statistical procedures for the Weibull or extreme-value distribution. Technometrics. 19: 323-331.

Engeman, R.M.; Keefe, T.J. 1982. On generalized least squares estimation of the Weibull distribution. Communications in Statistics A. Theory and Methods. 11: 2181-2193.

Evans, J.W.; Green, D.W. 1988. Mechanical properties of visually graded dimension lumber: Vol. 3. Hem–Fir. Pub. PB–88–159–405. Springfield, VA: National Technical Information Service. 392 p.

Evans, J.W.; Johnson, R.A.; Green, D.W. 1989. Two and three parameter Weibull goodness-of-fit tests. Res. Pap. FPL-RP-493. Madison, Wisconsin: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.

Fisher, R.A. 1912. On an absolute criterion for fitting frequency curves. Messenger of Mathematics. 41: 155-160.

Fisher, R.A. 1921. On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society of London, Series A. 222: 309-368.

Fisher, R.A.; Tippett, L.H.C. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Cambridge Phil. Soc. 24: 180-190.

Gauss, C.F. 1809. Theoria motus corporum coelestium in sectionibus conicis ambientium. Perthes & Besser, Hamburg. Reprinted in Werke, Band VII, Koniglichen Gessellschaft der Wissenschaften, Gottingen, 1906.

Gross, A.J.; Lurie, D. 1977. Monte carlo comparisons of parameter estimators of the 2-parameter Weibull distribution. IEEE Transactions on Reliability. 26(5): 356-358.

Gumbel, E.J. 1954. Statistical theory of extreme values and some practical applications. National Bureau of Standards, Applied Mathematics Theory, Series 33. Washinton, DC: U.S. Government Printing Office.

Gumbel, E.J. 1960. Statistics of extremes. New York: Columbia University Press.

Halperin, M. 1952. Maximum likelihood estimation in truncated samples. Annals of Mathematical Statistics. 23: 226-238.

Harter, H.L. 1970. Order statistics and their use in testing and estimation, Vol. 2: Estimates based on order statistics of samples from various populations. Washington, DC: U.S. Government Printing Office.

Harter, H.L.; Moore, A.H. 1965. Maximum-likelihood estimation of the parameters of Gamma and Weibull populations from complete and censored samples. Technometrics. 7: 639-643.

Harter, H.L.; Moore, A.H. 1967. Asymptotic variances and covariances of maximum-likelihood estimators, from censored samples, of the parameters of Weibull and gamma populations. Annals of Mathematical Statistics. 38: 557-571.

Hassanein, K.L. 1972. Simulataneous estimation of the parameters of the extreme value distribution by simple quantiles. Technometrics. 14: 63-70.

Johnson, N.L.; Kotz, S. 1970. Continuous univariate distribuions, Vol. 1. Boston, MA: Houghton Mifflin.

Johnson, R.A.; Evans, J.W.; Green, D.W. 2003. Confidence intervals for the ratios of percentiles from two Weibull populations. Res. Pap. FPL-RP-606. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.

Johnson, R.A.; Haskell, J.H. 1983. Sampling properties of estimators of the Weibull distribution of use in the lumber industry. The Canadian Journal of Statistics. 11(2): 155-169.

Kantar, Y.M. 2015. Generalized least squares and weighted least squares estimations for distributional parameters. REVSTAT – Statistical Journal. 13(3): 263-282.

Kendall, M.G.; Stuart, A. 1969. The advanced theory of statistics, Vol. 1. 3rd ed. New York: Hafner Publishing Company. 87 p.

Kimball, B.F. 1960. On the choice of plotting positions on probability paper. Journal of the American Statistical Association. 55: 546-560.

Lawrence, K.D.; Shier, D.R. 1981. A comparison of least squares and least absolute deviation regression models for estimating Weibull parameters. Communications in Statistics B. Simulation and Computation. 10: 315-326.

Lemon, G.H. 1975. Maximum likelihood estimation for the three parameter Weibull distribution based on censored samples. Technometrics. 17: 247-254.

Leone, F.D.; Rutenberg, Y.H.; Topp, C.W. 1960. Order statistics and estimators for the Weibull distribution. Case Statistical Laboratory Publication No. 1026 (AFSOR Report No. TN 60-389) Cleveland: Case Institute of Technology, Statistical Laboratory.

Lloyd, E.H. 1952. Least-squares estimation of location and scale parameters using order statistics. Biometrika. 39: 88-95.

Lloyd, E.H. 1962. Generalized least-squares theorem. In: Sarhan, A.E.; Greenberg, B.G., eds. Contributions to order statistics. New York: John Wiley & Sons: 20-27.

Mann, N.R. 1968. Point and interval estimation procedures for the two-parameter Weibull and extreme-value distributions. Technometrics. 10: 231-256.

Mann, N.R.; Fertig, K.W. 1977. Efficient unbiased quantile estimators for moderate size complete samples from extreme-value and Weibull distributions. Technometrics. 19: 87-93.

McCool, J.I. 1965. The construction of good linear unbiased estimates from the best linear estimates for a smaller sample size. Technometrics. 7: 543-552.

Menon, M.V. 1963. Estimation of the shape and scale parameters of the Weibull distribution. Technometrics. 5: 175-182.

Miller, I.; Freund, J.E. 1965. Probability and statistics for engineers. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Newby, M. 1984. Properties of moment estimators for the 3-parameter Weibull distribution. IEEE Transactions on Reliability. R-33: 192-195.

Pearson, K. 1894. Contributions to the mathematical theory of evolution. II. On the dissection of asymmetrical frequency curves. Philosophical Transactions of the Royal Society of London, Series A, 185, pages unknown. (republished in Karl Pearson's Early Statistical Papers. Cambridge Univ. Press. 1948. pp. 1-41).

Pearson, K. 1895. Contributions to the mathematical theory of evolution. II. Skew variations in homogeneous material. Philosophical Transactions of the Royal Society of London, Series A, 186, 343-414. (republished in Karl Pearson's Early Statistical Papers. Cambridge Univ. Press. 1948. pp. 41-112).

Rockette, H.; Antle, C.; Kimko, L.A. 1974. Maximum likelihood estimation with the Weibull distribution. Journal of the American Statistical Association. 69: 246-249.

Saylor, F. 1977. A monte-carlo comparison of the method-of-moments to the maximum-likelihood estimates of Weibull parameters for CAS data. In: Tsokos and Shimi, eds. Theory and applications of reliability, vol. II. Waltham, MA: Academic Press.

Shier, D.R.; Lawrence, K.D. 1984. A comparison of robust regression techniques for the estimation of Weibull parameters. Communications in Statistics B. Simulation and Computation. 13: 743-750.

Smith, R.L. 1985. Maximum likelihood estimation with the Weibull distribution. Biometrika. 72: 67-90.

Smith, R.L. 1994. Likelihood and modified likelihood estimation for distributions with unknown end points. In: Balakrishnan, N., ed. Recent advances in life-testing and reliability (volume in honor of A. C. Cohen). Boca Raton, FL: CRC Press.

Talreja, R. 1981. Estimation of Weibull parameters for composite material strength and fatigue life data. In: Fatigue of fibrous composite materials. ASTM STP 723. West Conshohocken, PA: American Society for Testing and Materials: 291-311.

Thoman, D.L.; Bain, L.L.; Antle, C.E. 1969. Inferences on the parameters of the Weibull distribution. Technometrics. 11: 445-460.

van Zyl, J.M.; Schall, R. 2012. Parameter estimation through weighted least-squares rank regression with specific reference to the Weibull and Gumbel distributions. Communications in Statistics – Simulation and Computation. 41(9): 1654-1666.

Verrill, S.P.; Evans, J.W.; Kretschmann, D.; Hatfield, C.A. 2012. Small sample properties of asymptotically efficient estimators of the parameters of a bivariate Gaussian-Weibull distribution. Res. Pap. FPL-RP-667. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.

Weibull, W. 1939a. A statistical theory of the strength of materials. Ing. Vetenskaps Akad. Handl. 153: 17ff.

Weibull, W. 1939b. The phenomenon of rupture in solids. Ing. Vetenskaps Akad. Handl. 151: 15ff.

Weibull, W. 1951. A statistical distribution of wide applicability. Journal of Applied Mechanics. 18: 293-297.

Weibull, W. 1952. Statistical design of fatigue experiments. Journal of Applied Mechanics. 19: 109-113.

White, J.S. 1965. Linear estimates for the log Weibull distribution. General Motors Research Publication GMR-481. Warren, MI: General Motors Corporation.

Wingo, D.R. 1972. Maximum likelihood estimation of the parameters of the Weibull distribution by modified quasilinearization. IEEE Transactions on Reliability. R-21: 89-93.

Wingo, D.R. 1973. Solution of the three-parameter Weibull equations by constrained modified quasilinearization. IEEE Transactions on Reliability. R-22: 96-102.

Zanakis, S.H. 1977. Computer experience with some nonlinear optimization algorithms in deriving maximum likelihood estimates for the three-parameter Weibull distribution. In: Neutts, M.F., ed. Algorithmic methods in probability. North Holland/TIMS Studies in Management Sciences. 7: 63-77.

Zanakis, S.H. 1978. Three-parameter Weibull MLE via pattern search with transformations. Unpublished report.

Zanakis, S.H. 1979. A simulation study of some simple estimators for the three-parameter Weibull distribution. Journal of Statistical Computation and Simulation. 9: 101-116.

Zanakis, S.H.; Kyparisis, J. 1986. A review of maximum likelihood estimation methods for the three-parameter Weibull distribution. Journal of Statistical Computation and Simulation. 25: 53-73.