United States
Department of
Agriculture

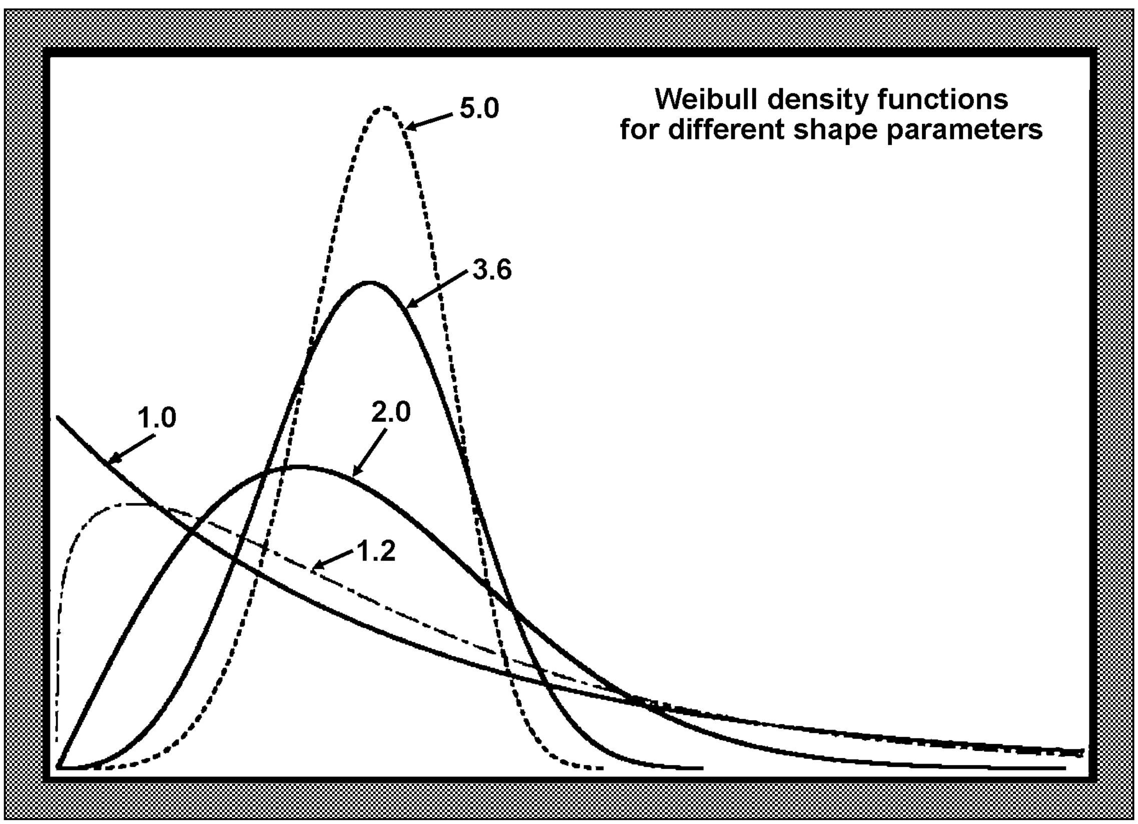
Forest Service

Forest
Products
Laboratory

Research
Paper
FPL–RP–606

Confidence Intervals for Predicting Lumber Strength Properties Based on Ratios of Percentiles From Two Weibull Populations

Richard A. Johnson
James W. Evans
David W. Green



Abstract

Ratios of strength properties of lumber are commonly used to calculate property values for standards. Although originally proposed in terms of means, ratios are being applied without regard to position in the distribution. It is now known that lumber strength properties are generally not normally distributed. Therefore, nonparametric methods are often used to derive property values. In some situations, estimating properties based on a parametric estimate is required. For these situations, the three-parameter Weibull distribution looks promising. To use this approach, procedures for estimating confidence intervals for ratios of percentiles from two Weibull populations are needed. In this study, we employed the large sample properties of maximum likelihood estimators to obtain a confidence interval for the ratio of 100α-th percentiles from two different three-parameter Weibull distributions. The coverage probabilities were investigated by a computer simulation study. We concluded that the procedure has considerable promise, but many questions remain to be answered.

Keywords: three-parameter Weibull, confidence intervals, ratio of percentiles

August 2003

Johnson, Richard A.; Evans, James W.; Green, David W. 2003. Confidence intervals for predicting lumber strength properties based on ratios of percentiles from two Weibull populations. Res. Pap. FPL-RP-606. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 8 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.

Confidence Intervals for Predicting Lumber Strength Properties Based on Ratios of Percentiles From Two Weibull Populations

Richard A. Johnson, Statistician
Department of Statistics, University of Wisconsin, Madison, Wisconsin

James W. Evans, Supervisory Mathematical Statistician
David W. Green, Supervisory Research General Engineer
Forest Products Laboratory, Madison, Wisconsin

Introduction

The ratio of two property estimates is commonly used in engineering design codes to establish allowable properties. For example, dry–green ratios (the ratio of small clear specimens dried to 12% moisture content to matched green specimens) are given in ASTM D2555 for a number of timber species (ASTM 1997b) and may be used to calculate allowable properties. Also, an E/G ratio of 16 is assumed in ASTM D2915 (ASTM 1997c) for adjusting the flexural modulus of elasticity (MOE) based on any span to depth ratio and several loading modes. As a third example, ASTM D245 specifies "strength ratios in tension parallel to grain are 55% of the corresponding bending strength ratios" (ASTM 1997a). Strength ratios are the strength that wood with a defect (like knots) is expected to have compared with the strength of a clear piece of wood. So ASTM D245 is saying that a defect lowering the bending strength to 80% of a clear piece (a strength ratio of 80%) would lower the tensile strength to 44% (0.80 times 0.55) of the clear wood specimen bending strength.

Ratios often are used when it is not practical, or perhaps not possible, to conduct tests for all combinations of factors (such as grades, sizes, test modes, environmental conditions) that may affect allowable properties. In wood engineering, the usual practice has been to conduct extensive tests for one combination of factors and to develop ratios for adjusting allowable properties from the measured combination of factors to a different set of factors. A primary example of this problem is estimating tensile strength parallel to the grain. Due to experimental difficulties in determining the strength of wood stressed in tension parallel to the grain, relatively little data exist on which to base allowable tensile properties. Thus, tensile strength has historically been estimated as a percentage of bending strength (Galligan and others 1979). A similar approach is taken in estimating the strength of lumber at various moisture content levels (ASTM 1997a, b).

Most of these ratios were originally established by analysis of mean trends. However, they are being applied without regard to position in the distribution. Recent studies have begun to focus on the ratio of properties at other percentile levels, especially the fifth percentile. These include studies of the effect of rate of loading on tensile strength (Gerhards and others 1984), the effect of moisture content on flexural strength (Aplin and others 1986; McLain and others 1984), and the effect of redrying on the strength of CCA-treated lumber (Barnes and Mitchell 1984).

It was once commonly assumed that lumber strength properties were normally distributed. It is now generally recognized that lumber strength properties are not normally distributed. Usually, nonparametric methods are now used for deriving allowable lumber properties (ASTM 1997c). In some situations, however, a parametric estimate of lumber properties for reliability-based design is required. Based on empirical evidence (Aplin and others 1986, Bodig 1977, Hoyle and others 1979, McLain and others 1984, Pierce 1976, Warren 1973), it is the three-parameter Weibull that emerges as a serious candidate. To use this parametric approach to estimate ratios of lumber properties, it is necessary to develop procedures for estimating confidence intervals for ratios of percentiles from two Weibull populations.

Typically the variability of ratio estimators can be very large. An underestimate of the property in the denominator and a corresponding overestimate of the property in the numerator of the ratio can result in a large estimate of the ratio. Correspondingly, an overestimate in the denominator and an underestimate in the numerator can produce a small estimate of a ratio. Since many decisions regarding ASTM standards are based on confidence intervals associated with an estimate of the ratio of properties, it is important to develop the best confidence limits on this ratio. The objective of this paper was to develop and evaluate procedures to create confidence intervals that have approximately 90% or 95% nominal coverage for the ratio of percentiles from two different three-parameter Weibull populations. In an earlier study, Johnson and Haskell (1984) investigated large sample tolerance bounds and confidence intervals for percentiles from a single three-parameter Weibull distribution. In the Procedures section of this paper, we develop three large sample confidence interval procedures based on the earlier work of Johnson and Haskell and one procedure based on order statistics. Using simulation techniques, we then evaluate the performance of a large sample approximation to the distribution of an estimate of a percentile ratio.

Procedures

Method 1

Let us assume we have two populations, both of which follow a Weibull distribution. Let $X_1, \ldots, X_{n_1}$ be a sample of size n_1 from the first Weibull population, which has the cumulative distribution function

$$F(x) = 1 - \exp[-\{(x - c_1)/b_1\}^{a_1}], \quad x \geq c_1$$

with probability density function $f(x)$ and 100α-th percentile given by

$$\xi_{1\alpha} = c_1 + b_1[-\ln(1-\alpha)]^{1/a_1}$$

Let $Y_1, \ldots, Y_{n_2}$ be a sample of size n_2 from the second Weibull population, which has the cumulative distribution function

$$G(y) = 1 - \exp[-\{(y - c_2)/b_2\}^{a_2}], \quad y \geq c_2$$

with probability density function $g(y)$ and 100α-th percentile given by

$$\xi_{2\alpha} = c_2 + b_2[-\ln(1-\alpha)]^{1/a_2}$$

Then the ratio of the population 100α-th percentiles is given by

$$\frac{\xi_{1\alpha}}{\xi_{2\alpha}} = \frac{c_1 + b_1[-\ln(1-\alpha)]^{1/a_1}}{c_2 + b_2[-\ln(1-\alpha)]^{1/a_2}}$$

Let

$$\left(\hat{a}_1, \hat{b}_1, \hat{c}_1\right)$$

and

$$\left(\hat{a}_2, \hat{b}_2, \hat{c}_2\right)$$

be maximum likelihood estimates obtained from the two samples.

In the single sample setting, Johnson and Haskell (1984) derived a large sample approximation for confidence intervals of percentiles from a Weibull distribution. Their 100(1 – γ)% confidence interval for the 100(1 – α) population percentile is

$$\hat{\xi}_\alpha - z_{\frac{\gamma}{2}} \hat{\sigma}_\alpha / \sqrt{n} \leq \xi_\alpha \leq \hat{\xi}_\alpha + z_{\frac{\gamma}{2}} \hat{\sigma}_\alpha / \sqrt{n}$$

if $\left(\hat{a}_1, \hat{b}_1, \hat{c}_1\right)$ occurs at the interior of Ω, where Ω is the parameter space (see Johnson and Haskell (1984) for a more thorough discussion of the parameter space) and

$$\sigma_\alpha^2 = \left(\frac{\partial \xi_\alpha}{\partial a}, \frac{\partial \xi_\alpha}{\partial b}, \frac{\partial \xi_\alpha}{\partial c}\right) \mathbf{I}^{-1} \begin{pmatrix} \frac{\partial \xi_\alpha}{\partial a} \\ \frac{\partial \xi_\alpha}{\partial b} \\ \frac{\partial \xi_\alpha}{\partial c} \end{pmatrix}$$

where

$$\left(\frac{\partial \xi_\alpha}{\partial a}, \frac{\partial \xi_\alpha}{\partial b}, \frac{\partial \xi_\alpha}{\partial c}\right) = \left(-\frac{b}{a^2}[-\ln(1-\alpha)]^{1/a} \ln[-\ln(1-\alpha)], [-\ln(1-\alpha)]^{1/a},\ 1\right)$$

and $\mathbf{I}$ is the observed Fisher information matrix. The information matrix can be estimated by

$$\hat{\mathbf{I}}_n = \left[-\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2}{\partial w_j \partial w_k} \ln f(x_i; \hat{a}, \hat{b}, \hat{c})\right]_{j,k=1,2,3}$$

where $(w_1, w_2, w_3) = (a, b, c)$. The second-order partial derivatives of the log-likelihood going into this estimated information matrix are

$$-\hat{I}_{11} = \frac{-1}{a^2} - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - c}{b}\right)^a \ln^2\left(\frac{x_i - c}{b}\right)$$

$$-\hat{I}_{22} = \frac{a}{b^2} - \frac{a(a+1)}{b^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - c}{b}\right)^a$$

$$-\hat{I}_{33} = -(a-1)\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(x_i-c)^2} - \frac{a(a-1)}{b^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-c}{b}\right)^{a-2}$$

$$-\hat{I}_{12} = -\frac{1}{b} + \frac{1}{b}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-c}{b}\right)^{a}$$

$$+\frac{a}{b}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-c}{b}\right)^{a}\ln\left(\frac{x_i-c}{b}\right)$$

$$-\hat{I}_{13} = -\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(x_i-c)} + \frac{1}{b}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-c}{b}\right)^{a-1}$$

$$+\frac{a}{b}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-c}{b}\right)^{a-1}\ln\left(\frac{x_i-c}{b}\right)$$

$$-\hat{I}_{23} = -\frac{a^2}{b^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-c}{b}\right)^{a-1}$$

So we can estimate the variance by

$$\hat{\sigma}_\alpha^2 = \left(\frac{\partial\xi_\alpha}{\partial a}, \frac{\partial\xi_\alpha}{\partial b}, \frac{\partial\xi_\alpha}{\partial c}\right)\Bigg|_{(\hat{a},\hat{b},\hat{c})} \hat{\mathbf{I}}^{-1} \left.\begin{pmatrix}\dfrac{\partial\xi_\alpha}{\partial a}\\ \dfrac{\partial\xi_\alpha}{\partial b}\\ \dfrac{\partial\xi_\alpha}{\partial c}\end{pmatrix}\right|_{(\hat{a},\hat{b},\hat{c})}$$

If the estimated shape parameter has value 1, placing it on the boundary of the parameter space, Johnson and Haskell (1984) recommend using a modified maximum likelihood solution for the two-parameter exponential distribution. In this case, we take

$$\hat{\xi}_\alpha = X_{(1)} + \hat{b}[-\ln(1-\alpha) - (n-1)^{-1}(1+\ln(1-\alpha))]$$

and its estimated variance (Johnson and Haskell 1984) to obtain the confidence interval

$$\hat{\xi}_\alpha - z_{\frac{\gamma}{2}}\hat{b}\left\{\left[-\ln(1-\alpha) - (n-1)^{-1}\left(1+\ln(1-\alpha)\right)\right]^2 \times\left[\frac{1}{n}-\frac{1}{n^2}\right]+\frac{1}{n^2}\right\}^{1/2}$$

$$\le \xi_\alpha \le \hat{\xi}_\alpha + z_{\frac{\gamma}{2}}\hat{b}\left\{\left[-\ln(1-\alpha) - (n-1)^{-1}\left(1+\ln(1-\alpha)\right)\right]^2 \times\left[\frac{1}{n}-\frac{1}{n^2}\right]+\frac{1}{n^2}\right\}^{1/2}$$

If we choose our confidence levels for each of two percentile estimates (which we denote by $(1-\gamma_1)$ and $(1-\gamma_2)$) such that $(1-\gamma_1)(1-\gamma_2) = (1-\gamma)$, then by independence

$$P\left[\frac{\hat{\xi}_{1\alpha} - z_{\frac{\gamma_1}{2}}\hat{\sigma}_{1\alpha}/\sqrt{n_1}}{\hat{\xi}_{2\alpha} + z_{\frac{\gamma_2}{2}}\hat{\sigma}_{2\alpha}/\sqrt{n_2}} \le \frac{\xi_{1\alpha}}{\xi_{2\alpha}} \le \frac{\hat{\xi}_{1\alpha} + z_{\frac{\gamma_1}{2}}\hat{\sigma}_{1\alpha}/\sqrt{n_1}}{\hat{\xi}_{2\alpha} - z_{\frac{\gamma_2}{2}}\hat{\sigma}_{2\alpha}/\sqrt{n_2}}\right] \ge 1-\gamma$$

when neither shape parameter estimate equals 1. The inequalities of the ratios will hold unless $\hat{\xi}_{2\alpha} - z_{\gamma_2/2}\hat{\sigma}_{2\alpha}/\sqrt{n_2}$ was negative. Since all values of a random variable with a Weibull distribution are assumed to be greater than or equal to the location parameter ($c \ge 0$), any negative estimate of the lower confidence limit of $\xi_{2\alpha}$ can be replaced by 0, which would maintain the inequality. If either or both shape parameter estimates equal 1, then the modified maximum likelihood estimate and confidence interval based on the two-parameter exponential distribution can be used.

Since any combination of γ_1 and γ_2 that meet our restriction can be chosen, the procedure above can lead to a variety of confidence intervals for the same two sets of data. We could produce a "shortest" confidence interval by searching for the values of γ_1 and γ_2 that produce the shortest confidence interval. However, it is not clear that we would achieve the desired coverage with a shortest confidence interval and with the coverage dependent on the estimates of $\xi_{1\alpha}$, $\xi_{2\alpha}$, $\sigma_{1\alpha}$, and $\sigma_{2\alpha}$ as well as γ_1 and γ_2. It is also not clear that this procedure would lead to a simple confidence interval for our ratio.

Method 2

Instead, it might be better to employ maximum likelihood estimators to obtain a large sample approximation to the distribution of

$$\hat{\xi}_{1\alpha}/\hat{\xi}_{2\alpha}$$

From a Taylor expansion, we see that

$$\frac{\hat{\xi}_{1\alpha}}{\hat{\xi}_{2\alpha}} = \frac{\hat{c}_1 + \hat{b}_1\left[-\ln(1-\alpha)\right]^{1/\hat{a}_1}}{\hat{c}_2 + \hat{b}_2\left[-\ln(1-\alpha)\right]^{1/\hat{a}_2}}$$

$$\approx \frac{\xi_{1\alpha}}{\xi_{2\alpha}} + \frac{1}{\hat{\xi}_{2\alpha}}\left(\hat{\xi}_{1\alpha}-\xi_{1\alpha}\right) - \frac{\xi_{1\alpha}}{\xi_{2\alpha}^2}\left(\hat{\xi}_{2\alpha}-\xi_{2\alpha}\right) + \text{higher order terms}$$

This suggests the large sample approximation

$$\widehat{\mathrm{Var}}\left(\frac{\hat{\xi}_{1\alpha}}{\hat{\xi}_{2\alpha}}\right) = \frac{1}{\hat{\xi}_{2\alpha}^2}\widehat{\mathrm{Var}}(\hat{\xi}_{1\alpha}) + \frac{\hat{\xi}_{1\alpha}^2}{\hat{\xi}_{2\alpha}^4}\widehat{\mathrm{Var}}(\hat{\xi}_{2\alpha})$$

Then a large sample approximate $(1-\gamma)$ 100% confidence interval for $\xi_{1\alpha}/\xi_{2\alpha}$ is given by

$$\frac{\hat{\xi}_{1\alpha}}{\hat{\xi}_{2\alpha}} \pm z_{\frac{\gamma}{2}}\sqrt{\frac{1}{\hat{\xi}_{2\alpha}^2}(\hat{\sigma}_{1\alpha}^2/n_1) + \frac{\hat{\xi}_{1\alpha}^2}{\hat{\xi}_{2\alpha}^4}(\hat{\sigma}_{2\alpha}^2/n_2)}$$

Method 3

Note that a similar procedure could be obtained by considering

$$\ln(\hat{\xi}_{1\alpha} / \hat{\xi}_{2\alpha}) = \ln(\xi_{1\alpha} / \xi_{2\alpha}) + \frac{1}{\xi_{1\alpha}}(\hat{\xi}_{1\alpha} - \xi_{1\alpha})$$
$$- \frac{1}{\xi_{2\alpha}}(\hat{\xi}_{2\alpha} - \xi_{2\alpha}) \; + \text{ higher order terms}$$

If we exponentiate the resulting interval for $\ln(\xi_{1\alpha}/\xi_{2\alpha})$, we can get a confidence interval that would be expected to work well if the normal distribution is a good approximation to

$$\ln(\hat{\xi}_{i\alpha}), \quad i = 1, 2$$

Method 4

Finally, we can get a conservative confidence interval based on order statistics. If the sample size is sufficiently large, it is well known that r_1 and s_1 can be selected so that the order statistics $X_{(r_1)}$ and $X_{(s_1)}$ satisfy

$$P[X_{(r_1)} < \xi_{1\alpha} < X_{(s_1)}] \geq 1 - \gamma_1$$

If $Y_{(r_2)}$ and $Y_{(s_2)}$ satisfy

$$P[Y_{(r_2)} < \xi_{2\alpha} < Y_{(s_2)}] \geq 1 - \gamma_2$$

where $(1 - \gamma_1)(1 - \gamma_2) \geq 1 - \gamma$, then

$$1 - \gamma \leq P[X_{(r_1)} < \xi_{1\alpha} < X_{(s_1)}, Y_{(r_2)} < \xi_{2\alpha} < Y_{(s_2)}]$$
$$\leq P\left[\frac{X_{(r_1)}}{Y_{(s_2)}} < \frac{\xi_{1\alpha}}{\xi_{2\alpha}} < \frac{X_{(s_1)}}{Y_{(r_2)}}\right]$$

Since random variables with a Weibull distribution assume values greater than or equal to the location parameter, which is greater than or equal to zero, all of the order statistics will be greater than or equal to zero, which maintains the inequality. So the interval

$$\frac{X_{(r_1)}}{Y_{(s_2)}}, \frac{X_{(s_1)}}{Y_{(r_2)}}$$

has probability at least $1 - \gamma$ of covering $\xi_{1\alpha}/\xi_{2\alpha}$

Although four order statistics are involved in the confidence limits, it is possible to express the exact coverage probability in terms of a single numerical calculation.

$$P\left[\frac{X_{(r_1)}}{Y_{(s_2)}} < \frac{\xi_{1\alpha}}{\xi_{2\alpha}} < \frac{X_{(s_1)}}{Y_{(r_2)}}\right] = P\left[\frac{X_{(r_1)}}{Y_{(s_2)}} < \frac{\xi_{1\alpha}}{\xi_{2\alpha}}\right] - P\left[\frac{X_{(s_1)}}{Y_{(r_2)}} < \frac{\xi_{1\alpha}}{\xi_{2\alpha}}\right]$$

Since $X_{(r_1)}$ and $Y_{(s_2)}$ are independent,

$$P\left[\frac{X_{(r_1)}}{Y_{(s_2)}} < z \middle| Y_{(s_2)} = y\right] = P[X_{(r_1)} < zy]$$

So

$$P\left[\frac{X_{(r_1)}}{Y_{(s_2)}} < z\right] = \int_0^\infty P[X_{(r_1)} < zy] g_{s_2}(y)\, dy$$

where

$$P[X_{(r_1)} < zy] = \sum_{j=r_1}^{n_1} \binom{n_1}{j} F^j(zy)[1 - F(zy)]^{n_1 - j}$$

and

$$g_{s_2}(y) = \frac{n_2!}{(s_2 - 1)!(n_2 - s_2)!} G^{s_2 - 1}(y) g(y) [1 - G(y)]^{n_2 - s_2}$$

where X has cumulative distribution F and Y has cumulative distribution G. The result can be simplified further by writing

$$F^j = (F - 1 + 1)^j = \sum_{i=0}^{j} \binom{j}{i} (-1)^i (1 - F)^i$$

Thus, given the true parameters for the two Weibull distributions, we could calculate an exact coverage probability for our confidence interval.

Computer Simulation

Of the four methods of developing confidence intervals for ratios of Weibull percentiles, method 2 appears to offer the most promise of giving a unique solution that is not encumbered by having to pick two significance levels whose product is the significance level we want for our interval. To evaluate the statistical properties of this large sample confidence interval procedure using a simulation study, we generated an ordered sample of uniform random numbers, using IMSL (1987) subroutines and then converted these to Weibull variates from a specified Weibull population. Each value of these observations could, for instance, represent the modulus of rupture for a hypothetical piece of lumber. A second sample was generated in a similar manner from another Weibull distribution. Maximum likelihood estimates of the Weibull parameters and corresponding percentile estimates for the 2nd, 5th, 10th, 30th, 50th, 70th, 90th, 95th, and 98th percentiles were calculated for each sample. The 5th percentile was of primary interest since design values for strength properties are often 5th percentile estimates. The two sets of maximum likelihood estimators were then combined according to the confidence interval procedure detailed in the Procedures section.

Nine runs were used to evaluate the coverage of the large sample confidence intervals for the ratio of the percentiles from two different three-parameter Weibull distributions. Earlier studies (Johnson and Haskell 1983, 1984) indicated that sample sizes of at least 100 are required to obtain relatively small variances and somewhat accurate normal approximations to the percentile estimates in the single population case.

For each run of our simulation, the sample sizes $n_1 = 100$ and $n_2 = 100$ were used. There were 250 replications at each set of conditions. The nine runs were a one-ninth run of a 3×3×3×3 factorial design as shown in Table 1. We selected the particular values as shown in Table 2.

These particular values were selected after examination of the parameter estimates in McLain and others (1984) and Aplin and others (1986), which are typical of lumber industry applications. The particular location/scale ratios were achieved by keeping the location parameters fixed. Specifically, we selected

$$\text{location 1} = \text{location 2} = 0.5$$

which meant the scale parameters associated with the coded values were as shown in Table 3.

Results and Conclusions

Tables 4 and 5 present the empirical coverage proportions and the average length of the intervals for both the approximate 95% and 90% confidence intervals. From Table 4, we see that, at the lower percentiles, the coverage of the 95% intervals is poorest for run 8, run 2, and run 5, in that order. The shape parameter for the first population is 2.75 for all these cases, but this cannot be pinpointed as the single cause because so many other parameters are changing. Because the runs were designed as discussed in section 3 to be a fractional factorial with main effects shape 1, shape 2, location 1/scale 1, and location 2/scale 2 , we can analyze them as such to get an idea of the sensitivity of our coverage probabilities to the parameters. Table 6 gives the mean square errors of the four main effects on the response variable coverage probabilities. The larger the mean square error, the more sensitive the result was to that effect. Table 7 gives the rankings of the effects from Table 6. Table 7 shows that for percentiles in the tail of our distribution, our 95% confidence intervals are most sensitive to the shape parameter of the Weibull distribution in the numerator.

The pattern is not so distinctive for the 90% confidence intervals (Table 5). Most importantly, the proportion of intervals that cover the true ratio of population percentiles is quite close to the nominal value in most cases considered.

Table 1—Parameters for the nine runs

Run	Shape 1	Shape 2	Location 1/ scale 1	Location 2/ scale 2
1	0	0	0	0
2	1	0	2	1
3	2	0	1	2
4	0	1	2	2
5	1	1	1	0
6	2	1	0	1
7	0	2	1	1
8	1	2	0	2
9	2	2	2	0

Table 2—Values selected for the factorial design

	Coded values		
Weibull parameter	0	1	2
Shape 1	2.00	2.75	3.50
Shape 2	2.00	2.75	3.50
Location 1/Scale 1	0.1	0.4	0.7
Location 2/Scale 2	0.1	0.4	0.7

Table 3—Scale parameters obtained with set location parameters

	Coded values for location/scale		
Weibull parameter	0	1	2
Scale 1	5	1.25	0.7142857
Scale 2	5	1.25	0.7142857

In view of the coverages for a single population percentile presented in Johnson and Haskell (1984), this was almost better than could be expected. We again analyzed the coverage probabilities as a fractional factorial design. We see in Table 8 that our 90% confidence intervals are most sensitive to the shape parameter of the Weibull distribution in the numerator.

What should we conclude from all this? This is a very small study, but it does indicate that the procedure has considerable promise. A great deal more work should be completed before we can be sure that the promise is a reality.

Table 4—Proportion of 95% intervals that cover the true ratio of percentiles (average length is given as the second entry)[a]

	Population percentiles (alpha)								
Run	0.02	0.05	0.10	0.30	0.50	0.70	0.90	0.95	0.98
1	0.9040	0.9240	0.9440	0.9560	0.9400	0.9480	0.9400	0.9480	0.9520
	0.7979	0.5661	0.4762	0.3696	0.3059	0.2675	0.2845	0.3172	0.3634
2	0.8920	0.9160	0.9280	0.9080	0.9240	0.9320	0.9560	0.9680	0.9680
	0.3020	0.2352	0.2063	0.1672	0.1383	0.1184	0.1201	0.1310	0.1467
3	0.9280	0.9240	0.9320	0.9520	0.9600	0.9480	0.9480	0.9520	0.9560
	0.4686	0.3626	0.3173	0.2738	0.2393	0.2134	0.2248	0.2486	0.2822
4	0.9280	0.9400	0.9400	0.9360	0.9440	0.9240	0.9240	0.9320	0.9280
	0.2245	0.1805	0.1671	0.1634	0.1572	0.1577	0.1912	0.2265	0.2765
5	0.9040	0.9160	0.9240	0.9360	0.9480	0.9560	0.9560	0.9540	0.9600
	0.2617	0.1611	0.1194	0.0812	0.0650	0.0551	0.0565	0.0627	0.0719
6	0.9480	0.9480	0.9480	0.9480	0.9560	0.9440	0.9520	0.9600	0.9520
	1.3807	0.9966	0.8130	0.6359	0.5442	0.4817	0.5111	0.5744	0.6675
7	0.9320	0.9520	0.9400	0.9480	0.9520	0.9560	0.9480	0.9600	0.9520
	0.2820	0.2138	0.1918	0.1838	0.1773	0.1769	0.2201	0.2628	0.3237
8	0.8760	0.9080	0.9280	0.9480	0.9640	0.9680	0.9520	0.9440	0.9520
	1.2538	0.9864	0.8744	0.7951	0.7438	0.7153	0.8433	0.9895	1.2015
9	0.8920	0.9240	0.9360	0.9440	0.9520	0.9560	0.9600	0.9560	0.9520
	0.1568	0.0938	0.0667	0.0427	0.0337	0.0281	0.0281	0.0311	0.0357

[a] n_1 = 100, n_2 = 100, 250 replications.

Table 5—Proportion of 90% intervals that cover the true ratio of percentiles (average length is given as the second entry)[a]

	Population percentiles (alpha)								
Run	0.02	0.05	0.10	0.30	0.50	0.70	0.90	0.95	0.98
1	0.8560	0.8680	0.8920	0.9040	0.9040	0.8920	0.8800	0.8880	0.8920
	0.6696	0.4751	0.3997	0.3102	0.2567	0.2245	0.2388	0.2662	0.3050
2	0.8440	0.8520	0.8640	0.8640	0.8680	0.8880	0.9080	0.9000	0.8960
	0.2534	0.1974	0.1731	0.1403	0.1161	0.0993	0.1008	0.1099	0.1231
3	0.8800	0.8560	0.8600	0.8920	0.9000	0.9120	0.9360	0.9320	0.9160
	0.3933	0.3043	0.2663	0.2298	0.2008	0.1791	0.1887	0.2087	0.2368
4	0.8800	0.8800	0.8960	0.9040	0.8960	0.8880	0.8760	0.8800	0.8920
	0.1884	0.1515	0.1402	0.1371	0.1319	0.1307	0.1605	0.1901	0.2320
5	0.8560	0.8640	0.8760	0.8720	0.8920	0.8920	0.9240	0.9240	0.9200
	0.2196	0.1352	0.1002	0.0682	0.0546	0.0463	0.0474	0.0526	0.0604
6	0.8800	0.8960	0.9000	0.9000	0.9200	0.9200	0.9200	0.9040	0.9240
	1.1587	0.8364	0.6823	0.5337	0.4567	0.4043	0.4289	0.4820	0.5601
7	0.8840	0.8840	0.9000	0.8960	0.9160	0.8880	0.9120	0.9160	0.9320
	0.2367	0.1794	0.1610	0.1543	0.1488	0.1484	0.1847	0.2205	0.2716
8	0.8240	0.8600	0.8680	0.9040	0.9000	0.9200	0.9040	0.9080	0.9040
	1.0522	0.8278	0.7338	0.6673	0.6242	0.6003	0.7078	0.8304	1.0084
9	0.8800	0.8640	0.8800	0.8880	0.9160	0.9240	0.9160	0.9000	0.8920
	0.1316	0.0788	0.0560	0.0359	0.0282	0.0236	0.0236	0.0261	0.0300

[a] n_1 = 100, n_2 = 100, 250 replications.

Table 6—Mean square errors for the effects of the Weibull parameters from analysis of the proportion of 95% intervals that cover the true ratio of percentiles[a]

	Population percentiles (alpha)								
	0.02	0.05	0.10	0.30	0.50	0.70	0.90	0.95	0.98
Shape 1	0.00098311	0.00051733	0.00018311	0.00027911	0.00011378	0.00006933	0.00027911	0.00008133	0.00019378
Shape 2	0.00056178	0.00013333	0.00000711	0.00005511	0.00016178	0.00032533	0.00006578	0.00004133	0.00010844
Location 1/ scale 1	0.00023644	0.00001600	0.00004978	0.00036978	0.00017778	0.00025600	0.00001244	0.00001733	0.00003378
Location 2/ scale 2	0.00043378	0.00026133	0.00002311	0.00011378	0.00011911	0.00006933	0.00011378	0.00030000	0.00011911

[a] n_1 = 100, n_2 =100, 250 replications.

Table 7—Relative importance[a] of the effects of the Weibull parameters from analysis of the proportion of 95% intervals that cover the true ratio of percentiles[b]

	Population percentiles (alpha)								
	0.02	0.05	0.10	0.30	0.50	0.70	0.90	0.95	0.98
Shape 1	1	1	1	2	4	3	1	2	1
Shape 2	2	3	4	4	2	1	3	3	3
Location 1/scale 1	4	4	2	1	1	2	4	4	4
Location 2/scale 2	3	2	3	3	3	3	2	1	2

[a] Importance is rated on a scale of 1 to 4, 1 being most important and 4 least.
[b] n_1 = 100, n_2 = 100, 250 replications.

Table 8—Relative importance[a] of the effects of the Weibull parameters from analysis of the proportion of 90% intervals that cover the true ratio of percentiles[b]

	Population percentiles (alpha)								
	0.02	0.05	0.10	0.30	0.50	0.70	0.90	0.95	0.98
Shape 1	1	2	1	1	1	1	1	2	4
Shape 2	3	1	2	4	2	2	4	3	3
Location 1/scale 1	2	4	4	2	3	2	2	1	1
Location 2/scale 2	4	3	3	3	4	4	3	4	2

[a] Importance is rated on a scale of 1 to 4, 1 being most important and 4 least.
[b] n_1 = 100, n_2 = 100, 250 replications.

Some of the things that need to be done or questions we should try to answer are as follows:

1. We need to look at a wider range of shape, scale, and location parameters in a way that allows us to look for interactions in the variables. For wood-related applications, shape parameters from 1 to 10 would be a good range. Since the shape parameter is related to the coefficient of variation of data from a two-parameter Weibull, can we model the relationship of variability of the data and the needed sample size to get intervals of a certain width on ratios of percentiles?

2. It would simplify the simulation work if we could show that ratios of parameters (like location to scale) were the important factors.

3. The number of replications in this study was much too small. With the variability that comes with estimating Weibull parameters, we might need 1,000 to 10,000 replications to get four decimal place accuracy on our coverage probabilities. We certainly need to find out how many replications it takes to get highly repeatable results.

4. We need to look at different sample sizes. How good is this procedure when we have 400 in each distribution of 40?

5. We need to look at the design aspects of the problem. If we have 300 green wood specimens, how many should we dry to estimate a dry/green ratio for the material?

6. Can we get a closed form solution for the problem? If we go to a two-parameter Weibull distribution, how does this change things?

7. We also need to look at the other methods for confidence limits. How well does the nonparametric procedure (method 4) work? If we use method 3, how do confidence intervals it produces compare with the method 2 intervals? Is there an easy way in method 1 to pick our confidence levels to get a shortest confidence interval?

8. We could also expand the problem to other distributional forms. Both the normal and lognormal distributions are very useful in wood-related research.

Clearly several questions could be looked at, and the results in this paper imply that the procedure does have promise and should be investigated further.

Literature Cited

Aplin, E.N.; Green, D.W.; Evans, J.W.; Barrett, J.D. 1986. The influence of moisture content on the flexural properties of Douglas Fir dimension lumber. Res. Pap. FPL–RP–475. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 32 p.

ASTM. 1997a. Standard practice for establishing structural grades and related allowable properties for visually-graded lumber. ASTM D245–93. Philadelphia, PA: American Society for Testing and Materials.

ASTM. 1997b. Standard test methods for establishing clear wood strength values. ASTM D2555–96. Philadelphia, PA: American Society for Testing and Materials.

ASTM. 1997c. Standard practice for evaluating allowable properties for grades of structural lumber. ASTM D2915–94. Philadelphia, PA: American Society for Testing and Materials.

Barnes, H.M.; Mitchell, P.H. 1984. Effect of post-treatment drying schedule on the strength of CCA-treated southern pine dimension lumber. Forest Products Journal. 34(6): 29–33.

Bodig, J. 1977. Bending properties of Douglas Fir-Larch and Hem-Fir dimension lumber. Spec. Publ. No. 6888. Fort Collins, CO: Colorado State University, Department of Forestry and Wood Science.

Galligan, W.L.; Gerhards, C.C.; Ethington, R.L. 1979. Evolution of tensile design stresses for lumber. Gen. Tech. Rep. FPL–GTR–28. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.

Gerhards, C.C.; Marx, C.M.; Green, D.W.; Evans, J.W. 1984. Effect of loading rate on tensile strength of Douglas-fir 2 by 6's. Forest Products Journal. 34(4): 23–26.

Hoyle, R.J., Jr.; Galligan, W.L.; Haskell, J.H. 1979. Characterizing lumber properties for truss research. In: Proceedings, Metal plate wood truss conference, 1979 Nov. 13–16, St. Louis, MO. Madison, WI: Forest Prod. Res. Soc.: 32–64.

IMSL, Inc. 1987. User's manual: Fortran subroutine library. Houston, TX: IMSL, Inc.

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

Johnson, R.A.; Haskell, J.H. 1984. An approximate lower tolerance bound for the three-parameter Weibull applied to lumber property characterization. Statistics and Probability Letters. 2: 67–76.

McLain, T.E.; DeBonis, A.L.; Green, D.W.; Wilson, F.J.; Link, C.L. 1984. The influence of moisture content on the flexural properties of southern pine dimension lumber. Res. Pap. FPL–RP–447. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 40 p.

Pierce, C.B. 1976. The Weibull distribution and the determination of its parameters for application to timber strength data. Current Pap. CP/76. Princes Risborough, Aylesbury, Buckinghamshire: Princes Risborough Laboratory. Build. Res. Estab.

Warren, W.G. 1973. Estimation of the exclusion limit for dimension lumber. In: Proceedings, IUFRO Division V meeting, Sept. and Oct. 1973, South Africa, Vol. 2: 1148–1154.