
Volume 140
Issue 7
July 2010
ISSN 0378-3758
ELSEVIER
journal of statistical planning and inference
For Contents see page iii
Available online at www.sciencedirect.com
ScienceDirect

Journal of Statistical Planning and Inference 140 (2010) 1754–1764

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Confidence limits for stress–strength reliability involving Weibull models

K. Krishnamoorthy *, Yin Lin

Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504, USA

A R T I C L E I N F O

Article history:
Received 20 December 2008
Received in revised form
26 December 2009
Accepted 29 December 2009
Available online 7 January 2010

Keywords:
Confidence limits
Coverage probability
Extreme-value distribution
One-sided tolerance limits
ROC curve
Type I censored samples

A B S T R A C T

The problem of interval estimation of the stress–strength reliability involving two independent Weibull distributions is considered. An interval estimation procedure based on the generalized variable (GV) approach is given when the shape parameters are unknown and arbitrary. The coverage probabilities of the GV approach are evaluated by Monte Carlo simulation. Simulation studies show that the proposed generalized variable approach is very satisfactory even for small samples. For the case of equal shape parameter, it is shown that the generalized confidence limits are exact. Some available asymptotic methods for the case of equal shape parameter are described and their coverage probabilities are evaluated using Monte Carlo simulation. Simulation studies indicate that no asymptotic approach based on the likelihood method is satisfactory even for large samples. Applicability of the GV approach for censored samples is also discussed. The results are illustrated using an example.

1. Introduction

The stress–strength reliability problem involves two independent random variables X_1 and X_2, where X_2 represents the strength variable of a component, and X_1 represents the stress variable to which the component is subjected. If $X_2 \leq X_1$, then either the component fails or the system that uses the component may malfunction. Hall (1984) provided an example of a system application where the breakdown voltage X_2 of a capacitor must exceed the voltage output X_1 of a transverter (power supply) in order for a component to work properly. Guttman et al. (1988) presented a rocket-motor experiment data where X_2 represents the chamber burst strength and X_1 represents the operating pressure. The reliability parameter R of the system or a component can be expressed as $R = P(X_2 > X_1)$. This problem of estimating $P(X_2 > X_1)$ also arises in receiver operating characteristic (ROC) curve analysis. The ROC curve is the plot of points $(P(X_1 \leq a), P(X_2 \leq a))$ for every real number a. Bamber (1975) noted that the area above the ROC curve is $P(X_2 > X_1)$. This area is used to measure the difference between two populations. The area above the ROC curve is commonly used to judge how accurately a test (such as treatment or diagnostic procedure) differentiate two populations (such as treatment and control groups). For more details on relevance of the parameter R on ROC curve analysis, see Swets (1996) and Reiser (2000). The parameter R also arises in many other applications. Wolfe and Hogg (1971) have introduced R as a general measure of difference, Hauck et al. (2000) have considered its usefulness in clinical trial applications.

As noted by McCool (1991), several authors suggested that in some cases the difference between two populations is more naturally characterized by $P(X_2 > X_1)$ than by the difference of their locations. For instance, if X_1 represents a patient's

* Corresponding author. Tel.: +1 337 482 5283; fax: +1 337 482 5346.
E-mail address: krishna@louisiana.edu (K. Krishnamoorthy).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2009.12.028

survival time if treated with drug A and X_2 represents the same if treated with drug B, then drug B could be preferred to drug A if $R = P(X_2 > X_1) > \frac{1}{2}$. A lower confidence bound on R can be used to check if R is greater than $\frac{1}{2}$. For a good exposition of this stress–strength reliability problem and other applications, we refer the readers to the book by Kotz et al. (2003).

The stress–strength reliability problem has been well addressed for the case of normally distributed stress and strength variables (see, Hall, 1984; Reiser and Guttman, 1986; Weerahandi and Johnson, 1992; Guo and Krishnamoorthy, 2004). For other distributions such as gamma and two-parameter exponential distributions, see Basu (1981), Constantine et al. (1989, 1990), Krishnamoorthy et al. (2007, 2008) and the references therein. However, only limited results are available in the case where X_1 and X_2 are independent Weibull random variables. McCool (1991) considered this stress–strength reliability problem for the Weibull case, and presented an interval estimation procedure. He has also provided table values that can be used to find confidence limits for R when the sample sizes are equal and the shape parameters are equal. Assuming that the shape parameters are equal, Mukherjee and Maiti (1998) developed interval estimation procedure on the basis of asymptotic normality of the maximum likelihood estimator (MLE) of R. They also provided interval estimation procedures based on variance stabilizing transformations such as logit and arc sine.

In this article, we propose a *generalized variable* (GV) approach to develop inferential procedures for the reliability parameter R. The concept of *generalized p-value* was introduced by Tsui and Weerahandi (1989) and that of *generalized confidence intervals* by Weerahandi (1993). The GV approach is useful to develop a so called *generalized pivotal quantity* (GPQ) which is used to construct confidence intervals for a parametric function of interest. Unlike the ordinary pivotal quantity, a GPQ is a function of observed statistics and random variables whose distributions are free of unknown parameters. This GV approach has been used successfully to address several complex problems such as estimating log-normal mean and for constructing tolerance limits in one-way random model and some mixed models; see Krishnamoorthy and Mathew (2003, 2004), Liao et al. (2005), and Chapters 4–6 of Krishnamoorthy and Mathew (2009). Hannig et al. (2006) have noted that the generalized variable procedures are a special case of fiducial inference procedures, and are asymptotically exact in many situations. For a good exposition of the GV approach along with numerous applications, see the books by Weerahandi (1995, 2004).

To assess the stress–strength reliability $R = P(X_2 > X_1)$, we assume that the stress variable $X_1 \sim \text{Weibull}(b_1, c_1)$ independently of the strength variable $X_2 \sim \text{Weibull}(b_2, c_2)$, where all the parameters are unknown. To express stress–strength reliability in terms of the parameters, we first note that the probability density function (pdf) of a Weibull distribution with the scale parameter b and the shape parameter c is given by

$$f(x|b,c) = \frac{c}{b}\left(\frac{x}{b}\right)^{c-1}\exp\left\{-\left[\frac{x}{b}\right]^{c}\right\}, \quad x>0, \ b>0, \ c>0. \tag{1}$$

The cumulative distribution function (cdf) is $1-\exp(-(x/b)^c)$. Using these pdf and cdf, it is easy to check that the stress–strength reliability is given by

$$R = P(X_2 > X_1) = \frac{c_1}{b_1}\int_0^\infty e^{-(x_1/b_2)^{c_2}}\left(\frac{x_1}{b_1}\right)^{c_1-1} e^{-(x_1/b_1)^{c_1}}\,dx_1. \tag{2}$$

The integral in (2) can be evaluated analytically as an infinite series expression (see Kotz et al., 2003, p. 53) involving gamma functions, but we found the above integral was easier to evaluate numerically (see Section 4) than the infinite series. A simple expression for R can be obtained if $c_1 = c_2$. In this case, the reliability can be expressed as

$$R_e = \frac{b_2^c}{b_2^c + b_1^c}, \tag{3}$$

where c is the unknown common shape parameter.

The rest of the article is organized as follows. In the following section, we give likelihood equations to find the MLEs, some distributional results concerning the MLEs, and the GPQs for the Weibull parameters following the approach of Krishnamoorthy et al. (2009). In Section 3, we first provide the generalized confidence limits when the shape parameters are unknown and arbitrary. For the special case of equal shape parameter, we describe the GV method, the asymptotic likelihood methods by Mukherjee and Maiti (1998) and the asymptotic approach by McCool (1991). We also show that the GV method is exact for making inference on the reliability R_e defined in (3). The coverage probabilities of the asymptotic methods are evaluated using Monte Carlo simulation in Section 4. An illustrative example involving real data is given in Section 5, and some concluding remarks are given in Section 6.

2. MLEs and pivotal quantities

Let $x_{11}, \ldots, x_{1n_1}$ be a sample of observations on the strength variable X_1, and let $x_{21}, \ldots, x_{2n_2}$ be a sample of observations on the stress variable X_2. Recall that $X_1 \sim \text{Weibull}(b_1, c_1)$ independently of $X_2 \sim \text{Weibull}(b_2, c_2)$.

2.1. MLEs for the parameters and pivotal quantities

The MLEs were first derived by Cohen (1965), and they are as follows. The MLE $\hat{c}_i$ of c_i is the solution to the equation

$$\frac{1}{\hat{c}_i}-\frac{\sum_{j=1}^{n_i} x_{ij}\hat{c}_i \ln(x_{ij})}{\sum_{j=1}^{n_i} x_{ij}^{\hat{c}_i}}+\frac{1}{n_i}\sum_{j=1}^{n_i}\ln(x_{ij})=0, \hat{b}_i=\left(\sum_{j=1}^{n_i} x_{ij}^{\hat{c}_i n_i}\right)^{1/\hat{c}_i}, \quad i=1,2. \tag{4}$$

Thoman et al. (1969) showed that the distributions of

$$\frac{\hat{c}_i}{c_i} \quad \text{and} \quad \hat{c}_i \ln\left(\frac{\hat{b}_i}{b_i}\right), \quad i=1,2, \tag{5}$$

do not depend on any parameters, and so they are pivotal quantities. As a consequence, we see that

$$\frac{\hat{c}_i}{c_i} \sim \hat{c}_i^*, \quad \text{and} \quad \hat{c}_i \ln\left(\frac{\hat{b}_i}{b_i}\right) \sim \hat{c}_i^* \ln(\hat{b}_i^*), \tag{6}$$

where the notation "$\sim$" means "distributed as", and $\hat{c}_i^*$ and $\hat{b}_i^*$ are the "MLEs" based on a sample $x_{i1},\ldots,x_{in_i}$ from a Weibull$(1,1)$ distribution. That is, $\hat{c}_i^*$ and $\hat{b}_i^*$ are the solutions of the Eq. (4) with $x_{i1},\ldots,x_{in_i}$ being a sample from a Weibull$(1,1)$ distribution. Thus, empirical distributions of these quantities in (6) can be obtained by generating independent samples from a Weibull$(1,1)$ distribution.

As the GV procedures that will be considered in the sequel heavily depend on empirical distributions of the MLEs, it is worth noting an iterative method of solving the likelihood equation (4). Let $y_1,\ldots,y_n$ be a sample of observations from a Weibull(b,c) distribution. Let $z_i=\ln(y_i),\ i=1,\ldots,n$. Menon (1963) showed that the estimator

$$\hat{c}_y=\frac{\pi}{\sqrt{6}}\left(\frac{\sum_{i=1}^{n}(z_i-\bar{z})^2}{n-1}\right)^{-1/2} \tag{7}$$

is asymptotically unbiased with $N(c,1.1c^2/n)$ distribution. Using this unbiased estimator as an initial value, the Newton–Raphson iterative method can be applied to find the root of the Eq. (4). For an algorithm to compute the MLEs, see Thoman et al. (1969).

2.2. MLEs when the shape parameters are equal and pivotal quantities

When the shape parameters are equal, we can write the log-likelihood function as

$$\ln L=(n_1+n_2)\ln c-c(n_1\ln b_1+n_2\ln b_2)+(c-1)\left(\sum_{j=1}^{n_1}\ln(x_{1j})+\sum_{j=1}^{n_2}\ln(x_{2j})\right)-\frac{\sum_{j=1}^{n_1}(x_{1j})^c}{b_1^c}-\frac{\sum_{j=1}^{n_2}(x_{2j})^c}{b_2^c}. \tag{8}$$

Following the lines of Schafer and Sheffield (1976), it can be easily shown that the MLE $\hat{c}$ of the common shape parameter is the solution of the equation

$$\hat{c}^{-1}=\left[\frac{n_1}{n_1+n_2}\frac{\sum_{j=1}^{n_1}x_{1j}^{\hat{c}}\ln(x_{1j})}{\sum_{j=1}^{n_1}x_{1j}^{\hat{c}}}+\frac{n_2}{n_1+n_2}\frac{\sum_{j=1}^{n_2}x_{2j}^{\hat{c}}\ln(x_{2j})}{\sum_{j=1}^{n_2}x_{2j}^{\hat{c}}}\right]-\frac{\sum_{j=1}^{n_1}\ln(x_{1j})+\sum_{j=1}^{n_2}\ln(x_{2j})}{n_1+n_2}, \tag{9}$$

and the MLEs of the scale parameters are given by

$$\hat{b}_1=\left(\frac{\sum_{j=1}^{n_1}x_{1j}^{\hat{c}}}{n_1}\right)^{1/\hat{c}} \quad \text{and} \quad \hat{b}_2=\left(\frac{\sum_{j=1}^{n_2}x_{2j}^{\hat{c}}}{n_2}\right)^{1/\hat{c}}. \tag{10}$$

The above equations are generalizations of the log-likelihood equations for the case $n_1=n_2$ given in Schafer and Sheffield (1976). Using the results for the one-sample case by Thoman et al. (1969), Schafer and Sheffield argued that the distributions of

$$\frac{\hat{c}}{c}, \quad \hat{c}\ln\left(\frac{\hat{b}_1}{b_1}\right) \quad \text{and} \quad \hat{c}\ln\left(\frac{\hat{b}_2}{b_2}\right) \tag{11}$$

do not depend on any parameters, and so they are pivotal quantities. As a consequence, we see that

$$\frac{\hat{c}}{c}\sim\hat{c}^*, \quad \hat{c}\ln\left(\frac{\hat{b}_1}{b_1}\right)\sim\hat{c}^*\ln(\hat{b}_1^*) \quad \text{and} \quad \hat{c}\ln\left(\frac{\hat{b}_2}{b_2}\right)\sim\hat{c}^*\ln(\hat{b}_2^*), \tag{12}$$

where $\hat{c}^*$, $\hat{b}_1^*$ and $\hat{b}_2^*$ are the "MLEs" based on independent samples from a Weibull(1, 1) distribution. That is, $\hat{c}^*$, $\hat{b}_1^*$ and $\hat{b}_2^*$ are the solutions of the Eqs. (9) and (10) with $x_{11},\ldots,x_{1n_1}$ and $x_{21},\ldots,x_{2n_2}$ being independent samples from a Weibull(1, 1) distribution.

To find the root of the Eq. (9), the Newton–Raphson method can be used. To obtain a starting value for the root finding method, let $\hat{c}_{iu}$ be the estimator (7) based on the sample $x_{i1},\ldots,x_{in_i}$ from a Weibull(b_i, c_i) distribution, $i=1,2$. Then $\hat{c}_u=(n_1\hat{c}_{1u}+n_2\hat{c}_{2u})/(n_1+n_2)$ can be used as a starting value for the Newton–Raphson iterative method. The following algorithm can be used to compute the MLEs and it can coded in any programming language such as Fortran and C.

Algorithm 1.

1. For a given sample $x_{11},\ldots,x_{1n_1}$ from Weibull(b_1, c) and a sample $x_{21},\ldots,x_{2n_2}$ from Weibull(b_2, c), set
   $y_{1i}=\ln(x_{1i}),\ i=1,\ldots,n_1$ and $y_{2j}=\ln(x_{2j}), j=1,\ldots,n_2$; set $q_1=n_1/(n_1+n_2)$ and $q_2=1-q_1$
2. Compute $s_{11}=\sum_{i=1}^{n_1} y_{1i}$ and $s_{12}=\sum_{i=1}^{n_2} y_{2i}$
   $\overline{y}_1=s_{11}/n_1$; $\overline{y}_2=s_{12}/n_2$;
   $s_1^2=\sum_{i=1}^{n_1}(y_{1i}-\overline{y}_1)^2$ and $s_2^2=\sum_{i=1}^{n_1}(y_{2i}-\overline{y}_2)^2$;
   $\hat{c}_{1u}=\pi/\sqrt{6}\sqrt{n_1-1}/s_1$; $\hat{c}_{2u}=\pi/\sqrt{6}\sqrt{n_2-1}/s_2$;
   $c_u=q_1\hat{c}_{1u}+q_2\hat{c}_{2u}$
3. For $j=1$ to number of iterations
   $s_{21}=\sum_{i=1}^{n_1} x_1^{c_u}$; $s_{22}=\sum_{i=1}^{n_1} x_2^{c_u}$
   $s_{31}=\sum_{i=1}^{n_1} x_{1i}^{c_u}y_{1i}$; $s_{32}=\sum_{i=1}^{n_2} x_2^{c_u}y_{2i}$
   $s_{41}=\sum_{i=1}^{n_1} x_{1i}^{c_u}y_{1i}^2$; $s_{42}=\sum_{i=1}^{n_2} x_2^{c_u}y_{2i}^2$
   $f=1/c_u+q_1s_{11}/n_1-q_1s_{31}/s_{21}+q_2s_{12}/n_2-q_2s_{32}/s_{22}$
   if (abs(f) $\leq$ error tolerance) return $c_{\text{mle}}=c_u$
   $f'=1/c_u^2+q_1(s_{21}s_{41}-s_{31}^2)/s_{21}^2+q_2(s_{22}s_{42}-s_{32}^2)/s_{22}^2$
   $c_u=c_u+f/f'$
   (end j loop)
4. $b_{1\text{mle}}=(\sum_{i=1}^{n_1} x_{1i}^{c_{\text{mle}}}/n_1)^{1/c_{\text{mle}}}$; $b_{2\text{mle}}=(\sum_{i=1}^{n_2} x_{2i}^{c_{\text{mle}}}/n_2)^{1/c_{\text{mle}}}$

In most cases, the above algorithm converges in 10 or less number of iterations when the error tolerance is 10^{-6}.

3. Confidence limits

We shall now describe various methods of obtaining confidence limits for the stress–strength reliability parameter R.

3.1. Generalized confidence limits for R

The GPQs for the parameters b_1, b_2, c_1 and c_2 are given in Krishnamoorthy et al. (2009), and they can be expressed as follows. Let $(\hat{c}_{10},\hat{b}_{10},\hat{c}_{20},\hat{b}_{20})$ be an observed value of the MLE $(\hat{c}_1,\hat{b}_1,\hat{c}_2,\hat{b}_2)$. Let G_θ denote the GPQ for θ. Then

$$G_{c_i}=\frac{c_i}{\hat{c}_i}\hat{c}_{i0}=\frac{\hat{c}_{i0}}{\hat{c}_i^*}\quad\text{and}\quad G_{b_i}=\left(\frac{b_i}{\hat{b}_i}\right)^{\hat{c}_i/\hat{c}_{i0}}\hat{b}_{i0}=\left(\frac{1}{\hat{b}_i^*}\right)^{\hat{c}_i^*/\hat{c}_{i0}}\hat{b}_{i0},\quad i=1,2, \tag{13}$$

where $\hat{c}_i^*$ and $\hat{b}_i^*$ are as defined in (12). In general, a GPQ for a parameter should satisfy two conditions. Let us show that these conditions are satisfied by G_c. First, the value of G_{c_i} at $\hat{c}_i=\hat{c}_{i0}$ should be c_i; it is clear from (13) that $G_{c_i}=c_i$ at $\hat{c}_i=\hat{c}_{i0}$. Second, for a given $\hat{c}_{i0}$, the distribution of G_{c_i} is free of parameters; this condition also holds because the distribution c_i^* is free of any parameters. Using (6), it is easy to see that G_{b_i} also satisfies similar conditions. Unlike the pivotal quantities, the appropriate percentiles of a GPQ themselves form a confidence interval (CI) for the corresponding parameter. For example, $(G_{b_1;\alpha/2},G_{b_1;1-\alpha/2})$, where $G_{b_1;\alpha}$ is the 100 α th percentile of G_{b_1}, is a $1-\alpha$ generalized CI for b_1.

A useful feature of the GV approach is that a GPQ for a function of b_i' s and c_i' s can be obtained by simply plugging their GPQs in the function. Specifically, a GPQ for a function $h(b_i,c_i)$ is given by $h(G_{b_i},G_{c_i})$. This feature enables us to find a GPQ for the stress–strength reliability by substitution.

Let $R(b_1,c_1,b_2,c_2)$ denote the stress–strength reliability. Then, a GPQ for the reliability is given by

$$G_R=R(G_{b_1},G_{c_1},G_{b_2},G_{c_2}). \tag{14}$$

Notice that, for a given $(\hat{b}_{10},\hat{c}_{10},\hat{b}_{20},\hat{c}_{20})$, the distribution of G_R does not depend on any unknown parameters, and so Monte Carlo simulation can be used to estimate the percentiles of G_R. For example, an estimate of 100 α th percentile of G_R is a $1-\alpha$ lower confidence limit for R. The percentiles of G_R can be estimated using the following algorithm.

Algorithm 2.

1. For given samples, compute the MLEs $\hat{c}_{10}$, $\hat{b}_{10}$, $\hat{b}_{20}$ and $\hat{c}_{20}$.
2. Compute the MLEs $(\hat{b}_i^*, \hat{c}_i^*)$ based on a simulated sample of size n_i from Weibull(1, 1) distribution, $i=1,2$.
3. Compute the GPQs $G_{b_1}, G_{c_1}, G_{b_2}$ and G_{c_2} using (13).
4. Compute $R(G_{b_1}, G_{c_1}, G_{b_2}, G_{c_2})$ using a numerical integration procedure.
5. Repeat the steps 2–4 a large number of times, say, M times.

The 100αth percentiles of these M generated $R(G_{b_1}, G_{c_1}, G_{b_2}, G_{c_2})$' s is a $1-\alpha$ lower confidence limit for R. In order to get consistent results regardless of the values of seed used for random number generation, we recommend simulation consisting of at least $M=10,000$ runs. The precision of a Monte Carlo estimate based on 10,000 simulation runs seems to be very satisfactory (see Section 3.2).

3.2. *Confidence limits when the shape parameters are equal*

3.2.1. *A GPQ for the reliability parameter when $c_1=c_2=c$*

The GPQs for the parameters b_1, b_2 and c can be obtained using the distributional results in (12), and they are

$$G_c = \frac{c}{\hat{c}}\hat{c}_0 = \frac{\hat{c}_0}{\hat{c}^*} \quad \text{and} \quad G_{b_i} = \left(\frac{b_i}{\hat{b}_i}\right)^{\hat{c}/\hat{c}_0} \hat{b}_{i0} = \left(\frac{1}{\hat{b}_i^*}\right)^{\hat{c}^*/\hat{c}_0} \hat{b}_{i0}, \quad i=1,2, \tag{15}$$

where the MLEs $\hat{c}$, $\hat{b}_1$ and $\hat{b}_2$ are defined in (9) and (10), $(\hat{b}_{10}, \hat{b}_{20}, \hat{c}_0)$ is an observed value of $(\hat{b}_1, \hat{b}_2, \hat{c})$ and $(\hat{b}_1^*, \hat{b}_2^*, \hat{c}^*)$ is as defined in (12).

Recall that the stress–strength reliability in (3) when $c_1=c_2$ can be written as $R_e = 1/(1+\eta)$, where $\eta=(b_1/b_2)^c$. As R_e is a one-one function of η, it is enough to find a confidence limit for η. Towards this, we note that a GPQ for η can be obtained by replacing the parameters by their GPQs, and is given by

$$G_\eta = \left(\frac{G_{b_1}}{G_{b_2}}\right)^{G_c} = \left(\frac{\hat{b}_2^*}{\hat{b}_1^*}\right)\left(\frac{\hat{b}_{10}}{\hat{b}_{20}}\right)^{\hat{c}_0/\hat{c}^*} = \left(\frac{\hat{b}_2^*}{\hat{b}_1^*}\right)\hat{\eta}_0^{1/\hat{c}^*}, \tag{16}$$

where (G_{b_1}, G_{b_2}, G_c) and $(\hat{c}_0, \hat{b}_{10}, \hat{b}_{20})$ are given in (15) and

$$\hat{\eta}_0 = \left(\frac{\hat{b}_{10}}{\hat{b}_{20}}\right)^{\hat{c}_0}.$$

For a given $\hat{\eta}_0$, the distribution of G_η does not depend on any unknown parameters, and so Monte Carlo simulation can be used to estimate the percentiles of G_η. If $G_{\eta;p}$ denotes the p th quantile of G_η, then $(1+G_{\eta;1-\alpha})^{-1}$ is a $1-\alpha$ lower confidence limit for the stress–strength reliability R_e in (3). The following algorithm can be used to compute the percentiles of G_η.

Algorithm 3.

1. For given data sets, compute the MLEs $\hat{c}_0$, $\hat{b}_{10}$ and $\hat{b}_{20}$ using Algorithm 1.
2. Generate independent samples $x_{11}, \ldots, x_{1n_1}$ and $x_{21}, \ldots, x_{2n_2}$ from Weibull(1, 1).
3. Compute the MLEs $\hat{c}^*$, $\hat{b}_1^*$ and $\hat{b}_2^*$ using the samples in step 2 and Algorithm 1.
4. Compute the GPQ G_η using (16).
5. Repeat the steps 2–4 for a large number of times, say, 10,000.

The 100pth percentile of these 10,000 simulated values of G_η is a Monte Carlo estimate of $G_{\eta,p}$.

The above generalized CIs for η are acceptance regions of an exact test, and so they are exact. To see this, consider testing

$$H_0 : \eta \le \eta_0 \text{vs. } H_a : \eta > \eta_0, \tag{17}$$

based on the MLE $(\hat{b}_1/\hat{b}_2)^{\hat{c}}$ of η. Let $(\hat{b}_{10}/\hat{b}_{20})^{\hat{c}_0}$ be an observed value of $(\hat{b}_1/\hat{b}_2)^{\hat{c}}$. The p-value for testing above hypotheses is given by

$$\sup_{H_0} P\left[\left(\frac{\hat{b}_1}{\hat{b}_2}\right)^{\hat{c}} \ge \left(\frac{\hat{b}_{10}}{\hat{b}_{20}}\right)^{\hat{c}_0}\right] = \sup_{H_0} P\left[\left(\frac{b_1}{b_2}\right)^{\hat{c}}\left(\frac{\hat{b}_1/b_1}{\hat{b}_2/b_2}\right)^{\hat{c}} \ge \left(\frac{\hat{b}_{10}}{\hat{b}_{20}}\right)^{\hat{c}_0}\right] = \sup_{H_0} P\left[\eta^{\hat{c}^*}\left(\frac{\hat{b}_1^*}{\hat{b}_2^*}\right)^{\hat{c}^*} \ge \left(\frac{\hat{b}_{10}}{\hat{b}_{20}}\right)^{\hat{c}_0}\right] \quad \text{[using (12)]}$$
$$= P\left[\eta_0^{\hat{c}^*}\left(\frac{\hat{b}_1^*}{\hat{b}_2^*}\right)^{\hat{c}^*} \ge \left(\frac{\hat{b}_{10}}{\hat{b}_{20}}\right)^{\hat{c}_0}\right] = P\left[\hat{\eta}_0^{1/\hat{c}^*}\left(\frac{\hat{b}_2^*}{\hat{b}_1^*}\right) \le \eta_0\right] = P(G_\eta \le \eta_0), \tag{18}$$

where G_η is given in (16). Note that the above p-value does not depend on any unknown parameters, and so it can be computed using Monte Carlo simulation. Furthermore, it is clear from Eq. (1) of (18) and *probability integral transform* that

Table 1
95% lower confidence limits for the Weibull stress–strength reliability R_e when the shape parameters are equal; $\hat{R}_e =$ MLE of R_e.

$\hat{R}_e$	$n_1 = n_2$									
	8	10	12	15	18	20	25	30	40	50
95% lower confidence limits										
0.80	0.577	0.607	0.629	0.654	0.669	0.677	0.694	0.704	0.720	0.730
0.81	0.587	0.621	0.642	0.664	0.682	0.688	0.705	0.716	0.732	0.741
0.82	0.601	0.632	0.654	0.677	0.693	0.701	0.718	0.728	0.743	0.752
0.84	0.628	0.658	0.681	0.704	0.718	0.727	0.742	0.752	0.766	0.777
0.85	0.640	0.670	0.693	0.715	0.733	0.740	0.754	0.764	0.779	0.788
0.86	0.654	0.684	0.706	0.730	0.745	0.752	0.768	0.777	0.791	0.800
0.87	0.666	0.699	0.721	0.744	0.759	0.766	0.780	0.791	0.804	0.812
0.89	0.695	0.728	0.748	0.771	0.785	0.794	0.807	0.816	0.829	0.838
0.90	0.716	0.744	0.765	0.787	0.801	0.806	0.821	0.831	0.843	0.850
0.91	0.729	0.761	0.781	0.801	0.816	0.823	0.835	0.845	0.856	0.863
0.92	0.748	0.779	0.798	0.818	0.831	0.838	0.850	0.859	0.869	0.876
0.93	0.768	0.796	0.816	0.835	0.847	0.853	0.865	0.873	0.883	0.889
0.94	0.790	0.815	0.834	0.851	0.863	0.869	0.881	0.888	0.897	0.903
0.95	0.810	0.835	0.853	0.870	0.881	0.886	0.897	0.904	0.912	0.917
0.96	0.834	0.858	0.873	0.889	0.899	0.905	0.913	0.920	0.927	0.932
0.97	0.857	0.881	0.896	0.910	0.919	0.924	0.931	0.936	0.943	0.947
0.98	0.891	0.909	0.922	0.933	0.940	0.944	0.950	0.954	0.959	0.962

the above p-value is a realization of a uniform $(0,1)$ random variable. So the test that rejects H_0 in (17) whenever the above p-value is less than a nominal level α, or equivalently, when $G_{\eta;\alpha}$ is greater than η_0, is an exact test. Thus, the generalized confidence limits for R_e based on G_η in (16) are exact.

As the reliability is commonly assessed by a lower confidence limit, we estimated 95% lower confidence limits for R_e using Monte Carlo simulation (using Algorithm 3) consisting of 10,000 runs, and presented them in Table 1. These table values are given for equal sample size n ranging from 8 to 50, and the values of $\hat{R}_e$ in the range 0.8 to 0.98. Note that the GV approach is also applicable for unequal sample sizes; we chose equal sample sizes for convenience. These table values are provided so that a user can compare the results of his/her program with those reported in Table 1.

To judge the precision of the Monte Carlo estimates in Table 1, we shall use the method described in Dudewicz and van der Meulen (1984). Let $P_1, \ldots, P_N$ be simulated values of G_η. Let $P_{(1)} < \cdots < P_{(N)}$ be the ordered values. Let $[x]$ denote the largest integer less than or equal x. The $100p$ th percentile, that is, $P_{([Np])}$, is a Monte Carlo estimate of the p quantile $G_{\eta;p}$. Define

$$r = [z_{\alpha/2}\sqrt{Np(1-p)} + Np + 0.5] \quad \text{and} \quad s = [z_{1-\alpha/2}\sqrt{Np(1-p)} + Np + 1 + 0.5],$$

where z_p is the p quantile of a standard normal distribution. The interval $(P_{(r)}, P_{(s)})$ is a CI for $G_{\eta;p}$ with confidence at least $1-\alpha$. The width of the CI can be used to assess the precision of a Monte Carlo estimate. For the present problem, $N = 10,000$ and $p = 0.05$. To judge the accuracy of the Monte Carlo estimates in Table 1, we computed 95% CIs for a few cases as follows: Note that $r = 457$ and $s = 544$. For the case of $(n, \hat{R}_e) = (8, 0.80)$, the Monte Carlo estimate of $G_{\eta;0.05}$ is 0.577 and the CI is (0.572, 0.583); for $(n, \hat{R}_e) = (15, 0.85)$, the point estimate is 0.715 and the CI is (0.712, 0.720); for $(n, \hat{R}_e) = (20, 0.81)$, the point estimate is 0.688 and the CI is (0.685, 0.690); and for $(n, \hat{R}_e) = (20, 0.90)$, the point estimate is 0.806 and the CI is (0.804, 0.809). Note that for each case considered, the absolute differences between the point estimate and the endpoints of the corresponding CI are no more than 0.006.

3.2.2. Asymptotic likelihood methods

To find an asymptotic variance of the MLE $\hat{R}_e = \hat{b}_2^{\hat{c}}/(\hat{b}_1^{\hat{c}} + \hat{b}_2^{\hat{c}})$, we shall first write the Fisher information matrix $V = (v_{ij})$, where the partial derivatives v_{ij}' s given in Mukherjee and Maiti (1998) are

$$v_{11} = -\frac{\partial^2 \ln L}{\partial b_1^2} = -\frac{n_1 c}{b_1^2} + \frac{c(c+1)}{b_1^{c+2}} \sum_{j=1}^{n_1} x_{1j}^c,$$

$$v_{22} = -\frac{\partial^2 \ln L}{\partial b_2^2} = -\frac{n_2 c}{b_2^2} + \frac{c(c+1)}{b_2^{c+2}} \sum_{j=1}^{n_2} x_{2j}^c,$$

$$v_{33} = -\frac{\partial^2 \ln L}{\partial c^2} = \frac{n_1 + n_2}{c^2} + \frac{1}{b_1^c} \sum_{j=1}^{n_1} x_{1j}^c \left(\ln \frac{x_{1j}}{b_1}\right)^2 + \frac{1}{b_2^c} \sum_{j=1}^{n_2} x_{2j}^c \left(\ln \frac{x_{2j}}{b_2}\right)^2,$$

$$v_{12} = 0,$$

$$\nu_{13} = -\frac{\partial^2 \ln L}{\partial b_1 \partial c} = \frac{n_1}{b_1} - \frac{1}{b_1^{c+1}}\left\{\sum_{i=1}^{n_1} x_{1j}^c + c\sum_{i=1}^{n_1} x_{1j}^c\left(\ln\frac{x_{1j}}{b_1}\right)\right\},$$

$$\nu_{23} = -\frac{\partial^2 \ln L}{\partial b_2 \partial c} = \frac{n_2}{b_2} - \frac{1}{b_2^{c+1}}\left\{\sum_{j=1}^{n_2} x_{2j}^c + c\sum_{j=1}^{n_2} x_{2j}^c\left(\ln\frac{x_{2j}}{b_2}\right)\right\},$$

and $\nu_{ji} = \nu_{ij}$. An approximate estimate of the variance-covariance matrix of $(\hat{b}_1, \hat{b}_2, \hat{c})$ is $V^{-1}|_{\hat{b}_1,\hat{b}_2,\hat{c}}$. In order to find an approximate estimate of the variance of $\hat{R}_e$ using the *delta method*, let

$$G' = \left(\frac{\partial R_e}{\partial b_1}, \frac{\partial R_e}{\partial b_2}, \frac{\partial R_e}{\partial c}\right),$$

where

$$\frac{\partial R_e}{\partial b_1} = -\frac{c b_1^{(c-1)} b_2^c}{(b_1^c + b_2^c)^2}, \quad \frac{\partial R_e}{\partial b_2} = \frac{c b_1^c b_2^{c-1}}{(b_1^c + b_2^c)^2}, \quad \text{and} \quad \frac{\partial R_e}{\partial c} = \frac{(b_1 b_2)^c \ln(b_2/b_1)}{(b_1^c + b_2^c)^2}.$$

Then an approximate estimate of Var $(\hat{R}_e)$ is given by

$$\hat{\text{V}}\text{ar}(\hat{R}_e) \simeq [G'V^{-1}G]|_{\hat{b}_1,\hat{b}_2,\hat{c}}. \tag{19}$$

Thus, $(\hat{R}_e - R_e)/\sqrt{\hat{\text{V}}\text{ar}(\hat{R}_e)} \sim N(0,1)$ asymptotically. This result yields an approximate confidence interval for R_e as

$$\hat{R}_e \pm z_{1-\alpha/2}\sqrt{\hat{\text{V}}\text{ar}(\hat{R}_e)}. \tag{20}$$

Instead of approximating $\hat{R}_e$ by a normal distribution, Mukherjee and Maiti (1998) also considered some normalizing transformation $g(\hat{R}_e)$, whose approximate variance can be obtained by the delta method as

$$\text{Var } g(\hat{R}_e) \simeq (g'(R_e))^2 \text{Var}(\hat{R}_e).$$

Mukherjee and Maiti (1998) considered the following two transformations:

3.2.3. *Logit transformation*

Let $g(\hat{R}_e) = \ln(\hat{R}_e/1-\hat{R}_e)$, with $g'(\hat{R}_e) = 1/\hat{R}_e(1-\hat{R}_e)$. The $1-\alpha$ CI for $g(R_e)$ is given by

$$\ln\left(\frac{\hat{R}_e}{1-\hat{R}_e}\right) \pm z_{1-\alpha/2}\frac{\sqrt{\hat{\text{V}}\text{ar}(\hat{R}_e)}}{\hat{R}_e(1-\hat{R}_e).} \tag{21}$$

If (L,U) denotes the above CI, then the CI for R_e is $(e^L(1+e^L)^{-1}, e^U(1+e^U)^{-1})$.

3.2.4. *Arc sine transformation*

Let $g(\hat{R}_e) = \sin^{-1}(\sqrt{\hat{R}_e})$. Noting that $g'(\hat{R}_e) = \dfrac{1}{2\sqrt{\hat{R}_e(1-\hat{R}_e)}}$, a $1-\alpha$ CI for $g(\hat{R}_e)$ is obtained as

$$\sin^{-1}(\sqrt{\hat{R}_e}) \pm z_{1-\alpha/2}\sqrt{\frac{\text{Var}(\hat{R}_e)}{4\hat{R}_e(1-\hat{R}_e)}}. \tag{22}$$

Thus, an approximate $1-\alpha$ CI for R_e is given by $(\sin^2(L), \sin^2(U))$, where (L,U) is the CI given in (22).

Confidence limits based on asymptotic normality: McCool (1991) proposed an approximate CI for R_e assuming approximate normality for $\ln(\hat{\eta})$. To describe this CI, let $V = \hat{c}/c$, $T = \hat{c}(\ln(\hat{b}_1/b_1) - \ln(\hat{b}_2/b_2))$ and $\hat{x}_0 = \ln(\hat{\eta}_0)$. Furthermore, let $\mu_V = E(V)$, $\sigma_V^2 = \text{Var}(V)$ and $\sigma_T^2 = \text{Var}(T)$. Recall that V and T are pivotal quantities, and so the expectation and variances can be estimated using Monte Carlo simulation. In terms of these quantities, an approximate $1-\alpha$ two-sided CI for $\ln(\eta)$ is given by

$$\frac{\hat{x}_0\mu_V \pm z_{\alpha/2}\sqrt{\sigma_V^2(\hat{x}_0^2 - z_{\alpha/2}^2\sigma_T^2) + \mu_V^2\sigma_T^2}}{\mu_V^2 - z_{\alpha/2}^2\sigma_V^2}, \tag{23}$$

where z_α denotes the α quantile of the standard normal distribution. Note that the above CI for $\ln(\eta)$ can be easily transformed to obtain a CI for the reliability parameter R_e in (3). It should be noted that the asymptotic CI in (23) involves quantities that have to be estimated using Monte Carlo simulation. McCool provided values of μ_V, σ_T^2 and σ_V^2 for various sample sizes. Our simulation evaluation of these quantities indicated that many of the reported values in McCool's (1991) Table I are in error. For instance, when the common sample size $n=20$, our Monte Carlo estimate of $(\mu_V, \sigma_V, \sigma_T)$ is $(1.0555, 0.1392, 0.3397)$ whereas the reported value in McCool's Table I is $(1.0518, 0.1407, 0.2219)$; for $n=10$, ours is $(1.1127, 0.2202, 0.5251)$ and McCool's table value is $(1.1125, 0.2406, 0.4059)$.

Remark 1. A reviewer has pointed out that, in some applications, it is desired to estimate $R_t = P(X_2 > X_1 + t)$, where t is a known positive number. When $c_1 = c_2 = c$, it can be easily checked that

$$R_t = \frac{c}{b_1}\int_t^{\infty} e^{-(z/b_2)^c}\left(\frac{z-t}{b_1}\right)^{c-1} e^{-((z-t)/b_1)^c}\,dz, \quad (24)$$

and the integral can be evaluated numerically. A GPQ for R_t, say, G_{R_t} can be obtained by replacing the parameters by their GPQs. Monte Carlo method can be used to find the percentiles of G_{R_t}. Alternatively, a confidence limit for R_t can be obtained as follows. Note that

$$G_{R_t} = E_Z(e^{-(Z/G_{b_2})^{G_c}} | G_{b_1}, G_{b_2}, G_c),$$

where Z has a three-parameter Weibull(t, G_{b_1}, G_c) distribution with the known location parameter t, the scale parameter G_{b_1} and the shape parameter G_c. A two-stage simulation (two nested "do loops"), one for finding the expectation given (G_{b_1}, G_{b_2}, G_c) and another for estimating the percentiles of G_{R_t}, can be used to find confidence limits for R_t.

4. Coverage studies

To judge the accuracy of the methods considered in the preceeding sections, we estimated the coverage probabilities of 95% one-sided confidence limits for R in (2) using Monte Carlo method. The simulation for estimating the coverage probabilities of the generalized CIs was carried out as follows. For a given (b_1, c_1, b_2, c_2), we generated 1,000 pairs of samples, one from Weibull(b_1, c_1) and another from Weibull(b_2, c_2), and computed the MLEs $\hat{b}_{10}$, $\hat{c}_{10}$, $\hat{b}_{20}$ and $\hat{c}_{20}$. For each set of these MLEs, we computed 95% one-sided limits for the reliability R using Algorithm 2 with 1,000 repetitions. To evaluate the GPQ in (14), we used the IMSL (International Mathematics and Statistics Library) subroutine QDAGI. The proportion of 1,000 one-sided CIs that include R is a Monte Carlo estimate of the coverage probability. To estimate the coverage probability of the asymptotic methods, we used Monte Carlo simulation consisting of 10,000 runs.

The estimated coverage probabilities of 95% one-sided confidence limits for R defined in (2) are given in Table 2 for some sample sizes. It is easy to check that the generalized limits are scale invariant, and so for coverage studies we can take $b_2 = 1$ and $0 < b_1 \le 1$ without loss of generality. The parameters are chosen so that the reliability R ranges from 0.5 to 0.98. Note that R is the reliability parameter without any assumption on the shape parameters. We observe from Table 2 that the estimated coverage probabilities are very close to the nominal level 0.95 for all the cases considered. Thus, the coverage studies indicate that the GV approach is satisfactory even for small samples, and it can be safely used for applications.

The estimated coverage probabilities of 95% one-sided limits when $c_1 = c_2 = c$ are given in Table 3 for $n_1 = n_2 = 20, 30, 40$ and 60, and the values of c ranging from 0.5 to 5. We again note that the estimation procedures that we compare are scale invariant, and so without loss of generality, we can take $b_2 = 1$ and $0 < b_1 \le 1$ for comparison studies. The asymptotic likelihood one-sided lower limits are liberal for some values (c, b_1, b_2) and the upper limits are conservative for the same case. In other words, one-sided limits over cover in one tail, and under cover in the other tail. These under coverage and over coverage hold even for sample sizes as large as 60. McCool's confidence limits are satisfactory but their performance is depending on the common shape parameter c. In particular, for large values of c, the coverage probabilities of lower confidence limits are higher than the nominal level while those of upper tolerance limits are slightly smaller than the nominal level.

Table 2
Monte Carlo estimates of coverage probabilities of 95% one-sided confidence limits when c_1 and c_2 are arbitrary; L—coverage probabilities of lower limits; U-coverage probabilities of upper limits.

(c_2, b_1, c_1)	R	$n_1 = n_2 = 10$ L(U)	$n_1 = 10, n_2 = 15$ L(U)	$n_1 = n_2 = 20$ L(U)
$b_2 = 1$				
(1,1,1)	0.50	0.96(0.94)	0.95(0.94)	0.95(0.95)
(3,1,1)	0.57	0.95(0.95)	0.95(0.95)	0.95(0.95)
(1,0.7,1)	0.59	0.94(0.96)	0.94(0.96)	0.95(0.95)
(2,0.8,1)	0.62	0.95(0.94)	0.95(0.95)	0.95(0.95)
(5,0.8,1)	0.67	0.95(0.95)	0.95(0.95)	0.95(0.95)
(2,0.6,1)	0.71	0.95(0.94)	0.95(0.95)	0.95(0.95)
(3,0.6,1)	0.74	0.95(0.94)	0.95(0.95)	0.95(0.95)
(2,0.4,1)	0.81	0.95(0.95)	0.94(0.94)	0.95(0.95)
(5,0.5,1)	0.83	0.95(0.93)	0.95(0.95)	0.95(0.95)
(4,0.3,7)	0.87	0.95(0.95)	0.96(0.94)	0.95(0.95)
(4,0.3,0.8)	0.89	0.95(0.95)	0.96(0.94)	0.95(0.95)
(8,0.4,1)	0.90	0.95(0.94)	0.95(0.95)	0.95(0.95)
(4,0.3,1)	0.93	0.95(0.95)	0.96(0.95)	0.95(0.95)
(5,0.2,1)	0.98	0.95(0.95)	0.94(0.95)	0.95(0.95)

Table 3
Coverage probabilities of 95% one-sided confidence limits for stress–strength reliability; L—Coverage probabilities of lower limits; U—coverage probabilities of upper limits.

c	Method[a]	$n_1=n_2=20$, b_1			$n_1=n_2=30$, b_1		
		0.4 L(U)	0.6 L(U)	0.8 L(U)	0.4 L(U)	0.6 L(U)	0.8 L(U)
$b_2=1.0,\ c_1=c_2=c$							
0.5	1	0.90(0.96)	0.92(0.94)	0.93(0.94)	0.92(0.96)	0.93(0.95)	0.94(0.94)
	2	0.93(0.95)	0.94(0.94)	0.94(0.95)	0.94(0.95)	0.94(0.95)	0.94(0.95)
	3	0.92(0.96)	0.93(0.94)	0.93(0.94)	0.93(0.96)	0.93(0.95)	0.94(0.94)
	4	0.95(0.95)	0.96(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)
1.0	1	0.87(0.98)	0.91(0.95)	0.93(0.94)	0.89(0.97)	0.92(0.96)	0.93(0.94)
	2	0.92(0.96)	0.94(0.95)	0.94(0.94)	0.93(0.95)	0.94(0.95)	0.94(0.94)
	3	0.89(0.97)	0.92(0.95)	0.94(0.94)	0.91(0.96)	0.93(0.95)	0.94(0.94)
	4	0.96(0.94)	0.95(0.94)	0.96(0.95)	0.96(0.94)	0.95(0.95)	0.95(0.95)
2.0	1	0.83(0.99)	0.88(0.98)	0.91(0.95)	0.86(0.99)	0.90(0.97)	0.93(0.95)
	2	0.93(0.96)	0.93(0.96)	0.94(0.95)	0.93(0.96)	0.93(0.96)	0.94(0.95)
	3	0.87(0.98)	0.90(0.97)	0.92(0.95)	0.89(0.98)	0.92(0.97)	0.94(0.95)
	4	0.96(0.94)	0.96(0.94)	0.94(0.94)	0.96(0.94)	0.96(0.94)	0.96(0.94)
3.0	1	0.79(1)	0.86(0.99)	0.91(0.96)	0.82(1)	0.88(0.98)	0.92(0.96)
	2	0.92(0.97)	0.93(0.96)	0.93(0.95)	0.93(0.96)	0.93(0.96)	0.94(0.95)
	3	0.85(0.99)	0.89(0.97)	0.92(0.95)	0.87(0.99)	0.91(0.97)	0.93(0.96)
	4	0.97(0.94)	0.96(0.94)	0.96(0.95)	0.97(0.94)	0.97(0.94)	0.97(0.94)
5.0	1	0.74(1)	0.82(1)	0.89(0.97)	0.78(1)	0.84(0.99)	0.90(0.97)
	2	0.92(0.97)	0.93(0.97)	0.93(0.96)	0.93(0.97)	0.93(0.97)	0.93(0.96)
	3	0.82(1)	0.87(0.98)	0.91(0.96)	0.84(0.99)	0.88(0.98)	0.92(0.96)
	4	0.97(0.94)	0.97(0.94)	0.97(0.94)	0.97(0.94)	0.97(0.94)	0.97(0.94)
		$n_1=n_2=40$, b_1			$n_1=n_2=60$, b_1		
c	Method[a]	0.4 L(U)	0.6 L(U)	0.8 L(U)	0.4 L(U)	0.6 L(U)	0.8 L(U)
0.5	1	0.92(0.96)	0.93(0.95)	0.94(0.94)	0.93(0.96)	0.94(0.95)	0.95(0.95)
	2	0.94(0.95)	0.95(0.95)	0.95(0.95)	0.94(0.95)	0.94(0.95)	0.95(0.95)
	3	0.93(0.96)	0.94(0.95)	0.95(0.95)	0.94(0.96)	0.94(0.95)	0.95(0.95)
	4	0.95(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)
1.0	1	0.91(0.97)	0.92(0.96)	0.94(0.95)	0.91(0.97)	0.93(0.96)	0.94(0.95)
	2	0.94(0.95)	0.94(0.95)	0.94(0.95)	0.94(0.96)	0.94(0.96)	0.95(0.95)
	3	0.92(0.96)	0.93(0.96)	0.94(0.95)	0.92(0.96)	0.93(0.96)	0.94(0.95)
	4	0.95(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)	0.95(0.95)
2.0	1	0.87(0.99)	0.91(0.97)	0.93(0.95)	0.88(0.98)	0.91(0.97)	0.94(0.95)
	2	0.93(0.96)	0.94(0.95)	0.94(0.95)	0.93(0.96)	0.93(0.96)	0.95(0.95)
	3	0.90(0.98)	0.92(0.96)	0.94(0.95)	0.91(0.97)	0.92(0.96)	0.95(0.95)
	4	0.95(0.94)	0.95(0.94)	0.95(0.94)	0.95(0.95)	0.95(0.95)	0.95(0.95)
3.0	1	0.84(1)	0.90(0.98)	0.93(0.96)	0.87(0.99)	0.90(0.98)	0.93(0.96)
	2	0.93(0.96)	0.94(0.96)	0.94(0.95)	0.94(0.96)	0.94(0.96)	0.94(0.95)
	3	0.88(0.98)	0.92(0.97)	0.93(0.95)	0.90(0.98)	0.92(0.97)	0.94(0.95)
	4	0.97(0.94)	0.96(0.94)	0.96(0.94)	0.97(0.94)	0.96(0.94)	0.96(0.94)
5.0	1	0.81(1)	0.86(0.99)	0.91(0.97)	0.83(1)	0.89(0.99)	0.92(0.96)
	2	0.93(0.96)	0.93(0.96)	0.94(0.95)	0.94(0.97)	0.94(0.96)	0.94(0.95)
	3	0.86(0.99)	0.89(0.98)	0.93(0.96)	0.88(0.99)	0.91(0.97)	0.93(0.96)
	4	0.97(0.94)	0.96(0.94)	0.97(0.94)	0.97(0.94)	0.96(0.94)	0.97(0.94)

[a] 1. CI (20) based on asymptotic normality of $\hat{R}$; 2. CI (21) based on logit transformation; 3. CI (22) based on arc sine transformation; 4. McCool's CI in (23).

Table 4
Fatigue voltages (in kilovolts per millimeter) for two types of electric cable insulation.

Type I Insulation, X_1	32.0	35.4	36.2	39.8	41.2	43.3	45.5	46.0	46.2	46.4
	46.5	46.8	47.3	47.3	47.6	49.2	50.4	50.9	52.4	56.3
Type II Insulation, X_2	39.4	45.3	49.2	49.4	51.3	52.0	53.2	53.2	54.9	55.5
	57.1	57.2	57.5	59.2	61.0	62.4	63.8	64.3	67.3	67.7

5. An example

The data are taken from Example 5.4.2 of Lawless (2003), and they represent failure voltage levels of two types of electrical cable insulation when specimens were subjected to an increasing voltage stress in a laboratory test. Twenty specimens of each type were tested and the failure voltages are given in Table 4. It is shown in Lawless (2003) that both samples fit Weibull distributions well. In this example, we are interested in finding the type of insulation that has longer life. Specifically, let X_1 represent the life time of a type I insulation, and let X_2 represent the same for a type II insulation. Then a lower confidence limit for $P(X_2 > X_1)$ with a value greater than 0.5 indicates the superiority of the type II insulation in terms of longevity. The MLEs based on the X_2 sample are $\hat{c}_2 = 9.141$ and $\hat{b}_2 = 59.125$, and those based on the X_1 sample are $\hat{c}_1 = 9.383$ and $\hat{b}_1 = 47.781$. In the following, each lower confidence limit is estimated using Monte Carlo method consisting of 10,000 runs.

As the observed ratio $\hat{c}_{20}/\hat{c}_{10} = 0.974$, the assumption of common shape parameter c seems to be tenable. To estimate $R_e = P(X_2 > X_1)$, we computed the MLEs using (9) and (10) as $\hat{c} = 9.261$, $\hat{b}_2 = 59.161$ and $\hat{b}_1 = 47.753$. Using these MLEs, we obtained $\hat{R}_e = \hat{b}_2^{\hat{c}}/(\hat{b}_1^{\hat{c}} + \hat{b}_2^{\hat{c}}) = 0.879$. The 95% lower confidence limit for R_e, using (16) with 10,000 simulation runs, was obtained as 0.778. That is, the probability that a type II electrical cable insulation lasts longer than a type I electric cable insulation is at least 0.778 with confidence 0.95. To get the 95% lower limit from Table 1, we note $\hat{R}_e \simeq 0.88$, and so we can take the average of the lower limits when $\hat{R}_e = 0.87$ and $\hat{R}_e = 0.89$, which is 0.78.

The results based on the asymptotic likelihood methods are as follows: $\hat{R}_e = 0.879$ and the asymptotic variance estimate $\hat{\mathrm{Var}}(\hat{R}_e) = 0.0018$. Noting that $z_{0.95} = 1.6449$, we computed the asymptotic lower limit (see Eq. (20)) as 0.809. The logit transformation yielded 0.790, and the arc sine transformation produced 0.801. To construct the one-sided lower limit based on McCool's CI in (23), we estimated μ_V, σ_V^2 and σ_T^2 using Monte Carlo simulation as 1.054, 0.0196 and 0.1154, respectively. Using these values, the lower limit was obtained as 0.783. Among all asymptotic limits, McCool's limit is very close to the one based on the GV approach.

We also computed the 95% lower confidence limit for R in (2) (without assuming $c_1 = c_2$) using the GV approach as 0.747. This lower limit was obtained by simulating (14) 10,000 times and evaluating the integral using IMSL subroutine QDAGI. Furthermore, we computed a 95% lower confidence limit for $R_3 = P(X_2 > X_1 + 3)$ following the simulation method in Remark 1, as 0.687. By using simulation and numerical integration of (24), we computed the 95% lower confidence limit for R_3 as 0.684.

6. Concluding remarks

In this article, we proposed a GV approach to set limits for the stress–strength reliability involving two Weibull distributions. The GV method is applicable even when the shape parameters are unknown and arbitrary while other available methods are applicable only when the shape parameters are equal. Furthermore, the proposed GV approach produces satisfactory results even when the sample sizes are small. The GV approach is conceptually simple, and is easy to use. We also note that the procedures are applicable to extreme-value distributions because of the one-one relation between Weibull and extreme-value distributions. In particular, if the samples are from extreme-value distributions, then the procedures in the preceeding sections can be applied to estimate the stress–strength reliability after taking antilogarithmic transformation of samples.

The GV method is also applicable if both samples are type II censored. Specifically, the pivotal quantities based on the MLEs in (12) are also valid when the samples are type II censored, and so GPQs for the parameters can be obtained along the lines for the complete case. To compute the MLEs for censored samples, see Cohen (1965). If the samples are type I censored (time censored), then the pivotal quantities in (12) are no longer valid, and neither are the GPQs. However, Krishnamoorthy et al. (2009) observed that the pivotal quantities in (6) can be used as approximates, and based on them approximate GPQs can be obtained. These authors showed that confidence limits based on these approximate GPQs for a Weibull mean are satisfactory. Furthermore, Lin's (2009) simulation studies indicate that inferential methods based on such approximate GPQs are satisfactory for comparing two Weibull means.

Acknowledgment

The authors are thankful to two reviewers for providing some useful references, comments and suggestions.

References

Bamber, D., 1975. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. Journal of Mathematical Psychology 12, 387–415.

Basu, A.P., 1981. The estimation of $P(X<Y)$ for the distributions useful in life testing. Naval Research Logistic Quarterly 28, 383–392.

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

Constantine, K., Karson, M., Tse, S., 1989. Bootstrapping estimators of $P(Y<X)$ in the gamma case. Journal of Statistical Computation and Simulation 33, 217–231.

Constantine, K., Karson, M., Tse, S., 1990. Confidence interval estimation of $P(Y<X)$ in the gamma case. Communications in Statistics—Simulation and Computation 19, 225–244.

Dudewicz, E.J., van der Meulen, E.C., 1984. On assessing the precision of simulation estimates of percentile points. American Journal of Mathematical and Management Sciences 4, 335–343.

Guo, H., Krishnamoorthy, K., 2004. New approximate inferential methods for the reliability parameter in a stress–strength model: the normal case. Communication in Statistics—Theory and Methods 33, 1715–1731.

Guttman, I., Johnson, R.A., Bhattacharya, G.K., Reiser, B.J., 1988. Confidence limits for stress–strength models with explanatory variables. Technometrics 30, 161–168.

Hall, I.J., 1984. Approximate one-sided tolerance limits for the difference or sum of two independent normal variates. Journal of Quality Technology 16, 15–19.

Hannig, J., Iyer, H.K., Patterson, P., 2006. Fiducial generalized confidence intervals. Journal of the American Statistical Association 101, 254–269.

Hauck, W.W., Hyslop, T., Anderson, S., 2000. Generalized treatment effects for clinical trials. Statistics in Medicine 19, 887–899.

Kotz, S., Lumelskii, Y., Pensky, M., 2003. The Stress–Strength Model and Its Generalizations: Theory and Applications. World Scientific, Singapore.

Krishnamoorthy, K., Mathew, T., 2003. Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals. Journal of Statistical Planning and Inference 115, 103–121.

Krishnamoorthy, K., Mathew, T., 2004. One-sided tolerance limits in balanced and unbalanced one-way random models based on generalized confidence intervals. Technometrics 46, 44–52.

Krishnamoorthy, K., Mathew, T., 2009. Statistical Tolerance Regions: Theory, Applications and Computation. Wiley, Hoboken, NJ.

Krishnamoorthy, K., Mukherjee, S., Guo, H., 2007. Inference on reliability in two-parameter exponential stress–strength models. Metrika 65, 261–273.

Krishnamoorthy, K., Mathew, T., Mukherjee, S., 2008. Normal based methods for a gamma distribution: prediction and tolerance intervals and stress–strength reliability. Technometrics 50, 69–78.

Krishnamoorthy, K., Lin, Y., Xia, Y., 2009. Confidence limits and prediction limits for a Weibull distribution based on the generalized variable approach. Journal of Statistical Planning and Inference 139, 2675–2684.

Lawless, J.F., 2003. Statistical Models and Methods for Lifetime Data. Wiley, New York.

Liao, C.T., Lin, T.Y., Iyer, H.K., 2005. One- and two-sided tolerance intervals for general balanced mixed models and unbalanced one-way random models. Technometrics 47, 323–335.

Lin, Y. 2009. Generalized inference for Weibull distributions. Ph.D. Dissertation, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA, USA.

McCool, J.I., 1991. Inference on $P(Y<X)$ in the Weibull case. Communications in Statistics—Simulation and Computation 20, 129–148.

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

Mukherjee, S.P., Maiti, S.S., 1998. Stress–strength reliability in the Weibull case. Frontiers In Reliability 4, 231-248. Singapore: World Scientific.

Reiser, B.J., Guttman, I., 1986. Statistical inference for Pr $(Y<X)$: The normal case. Technometrics 28, 253–257.

Reiser, B., 2000. Measuring the effectiveness of diagnostic markers in the presence of measurement error through the use of ROC curves. Statistics in Medicine 19, 2115–2129.

Schafer, R., Sheffield, T., 1976. On procedures for comparing two Weibull populations. Technometrics 18, 231–235.

Swets, J.A., 1996. Signal Detection Theory and ROC Analysis in Psychology and Diagnostics. Collected Papers. Lawrence Erlbaum Assoc, New Jersey.

Thoman, D.R., Bain, L.J., Antle, C.E., 1969. Inferences on the parameters of the Weibull distribution. Technometrics 11, 445–446.

Tsui, K.W., Weerahandi, S., 1989. Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. Journal of the American Statistical Association 84, 602–607.

Weerahandi, S., Johnson, R.A., 1992. Testing reliability in a stress–strength model when X and Y are normally distributed. Technometrics 34, 83–91.

Weerahandi, S., 1993. Generalized confidence intervals. Journal of the American Statistical Association 88, 899–905.

Weerahandi, S., 1995. Exact Statistical Methods for Data Analysis. Springer-Verlag, New York.

Weerahandi, S., 2004. Generalized Inference in Repeated Measures: Exact Methods in MANOVA and Mixed Models. Wiley, New York.

Wolfe, A., Hogg, R.V., 1971. On constructing statistics and reporting data. The American Statistician 25, 27–30.