한국통계학회논문집 제3권 1호
1996년 4월 pp. 87-99

Bootstrap Confidence Intervals for Reliability in 1-way ANOVA Random Model

Dal Ho Kim[1] and Jang Sik Cho[2]

Abstract

We construct bootstrap confidence intervals for reliability, $R=P\{X>Y\}$, where X and Y are independent normal random variables. One way ANOVA random effect models are assumed for the populations of X and Y, where standard deviations σ_x and σ_y are unequal. We investigate the accuracy of the proposed bootstrap confidence intervals and classical confidence interval via Monte Carlo simulation. Results indicate that proposed bootstrap confidence intervals work better than classical confidence interval for small sample and/or large value of R.

1. Introduction

One way random effect model has been widely used in a variety of areas when the number of batches in a population is large. In statistical quality control and reliability analysis sometimes one is interested in $P\{X>Y\}$. For example, an engineer would like to compare failure times (X and Y) of automobile batteries of types A and B, respectively. He randomly selects n_1 batteries from each of l_1 batches of type A and n_2 from each of l_2 batches of type B batteries. Then he tests the batteries under a specific temperature and records failure times. The aim of the engineer is to find a confidence interval for the reliability, $R=P\{X>Y\}$.

Reisser and Guttman(1986) obtained approximate confidence interval for R in stress strength model with normal distribution. Guttman, Johnson, Bhattacharyya and Reisser(1988) found approximate confidence interval for R in stress strength model with explanatory variables. Aminzadeh(1991) derived approximate confidence interval based on the asymptotic normal distribution for the reliability under random effect model. Since the true distribution of

1) Full-time Lecturer, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.
2) Lecturer, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.

the estimator for R is often skewed and biased for a small sample and/or large value of R, the interval based on the asymptotic normal distribution may deteriorate the accuracy. We will use the bootstrap method to rectify these problems. Efron(1979) initially introduced the bootstrap method to assign the accuracy for an estimator. To construct approximate confidence interval for an estimator, Efron(1981, 1982, 1987) and Hall(1988) proposed the percentile method, the bias correct method(BC method), the bias correct acceleration method (BCa method), and the percentile- t method, etc.

In this paper, we consider the problem of derivation of confidence intervals for reliability under 1-way random effect ANOVA. Specifically, we derive a large sample property for the bootstrap estimator of R and propose approximate bootstrap confidence intervals for R based on percentile, BC, BCa and percentile- t methods. Also we investigate the accuracy of the proposed bootstrap confidence intervals and confidence interval based on Aminzadeh(1991)'s method via Monte Carlo simulation. In particular, we observe the accuracy of these intervals for small sample and/or large value of R.

2. Consistency for Bootstrap Estimator

We assume that n_1 measurements from each of l_1 batches of population 1 and n_2 from each of l_2 batches of population 2 are selected. Let μ_x and μ_y are overall means for populations 1 and 2. And let A and B are batch effects for populations 1 and 2. Then 1-way random effect models for X and Y are defined as follows:

$$X_{ij} = \mu_x + A_j + e_{ij}, \quad i=1,2,\cdots,n_1, \quad j=1,2,\cdots,l_1 \tag{2.1}$$

and

$$Y_{qr} = \mu_y + B_r + \varepsilon_{qr}, \quad q=1,2,\cdots,n_2, \quad r=1,2,\cdots,l_2, \tag{2.2}$$

where $A_j, e_{ij}, B_r, \varepsilon_{qr}$ are stochastically independent normal random variables with means zero and standard deviations, $\sigma_A, \sigma_e, \sigma_B, \sigma_\varepsilon$, respectively.

From (2.1) and (2.2) we can see that X_{ij} and Y_{qr} have normal distribution with means μ_x, μ_y and variances $\sigma_x^2 = \sigma_A^2 + \sigma_e^2$, $\sigma_y^2 = \sigma_B^2 + \sigma_\varepsilon^2$, respectively. Then the reliability is computed as $R = \Phi(\delta)$, where Φ denotes the cumulative distribution function of a standard normal random variable and $\delta = \frac{\mu_x - \mu_y}{\sqrt{\sigma_x^2 + \sigma_y^2}}$. Let $\underline{X} = (X_{11}, X_{12}, \cdots, X_{n_1 l_1})$ and $\underline{Y} =$

$(Y_{11}, Y_{12}, \cdots, Y_{n_2 l_2})$ be vectors of measurements for X and Y, respectively. And let $N_1 = n_1 l_1$ and $N_2 = n_2 l_2$. By Aminzadeh(1991), the estimator $\hat{R}$ of R is given by

$$\hat{R} = \Phi(\hat{\delta}) = \Phi\left(\frac{\overline{X}_{..} - \overline{Y}_{..}}{\sqrt{\widehat{\sigma_x}^2 + \widehat{\sigma_y}^2}} \right), \tag{2.3}$$

where $\overline{X}_{..} = \frac{1}{N_1} \sum_{i=1}^{n_1} \sum_{j=1}^{l_1} X_{ij}$, $\overline{Y}_{..} = \frac{1}{N_2} \sum_{q=1}^{n_2} \sum_{r=1}^{l_2} Y_{qr}$, $\widehat{\sigma_x}^2 = \frac{(n_1 - 1) S_e^2 + S_A^2}{n_1}$ and $\widehat{\sigma_y}^2 = \frac{(n_2 - 1) S_\varepsilon^2 + S_B^2}{n_2}$. Note that S_e^2, S_ε^2 and S_A^2, S_B^2 are mean squares within and between batches for populations 1 and 2, respectively.

Related to the Aminzadeh's procedure is the bootstrap procedure which is a resampling scheme that one attempts to learn the sampling properties of a statistic by recomputing its value on the basis of a new sample realized from the original one. The bootstrap procedure provides confidence interval estimates by using the plug-in principle for R. The bootstrap procedure for the construction of bootstrap estimators for R can be described as follows:

(1) Compute the plug-in estimates of μ_x, μ_y, σ_x^2 and σ_y^2 given by $\overline{X}_{..}$, $\overline{Y}_{..}$, $S_x^2 = \frac{1}{N_1} \sum_{i=1}^{n_1} \sum_{j=1}^{l_1} (X_{ij} - \overline{X}_{..})^2$ and $S_y^2 = \frac{1}{N_2} \sum_{q=1}^{n_2} \sum_{r=1}^{l_2} (Y_{qr} - \overline{Y}_{..})^2$ from $\underline{X}$ and $\underline{Y}$ respectively.

(2) Construct the sampling distribution $\hat{F}$ and $\hat{G}$ (from $\underline{X}$ and $\underline{Y}$) based on $\overline{X}_{..}$, $\overline{Y}_{..}$, S_x^2 and S_y^2, respectively. That is, $\hat{F} \sim N(\overline{X}_{..}, S_x^2)$ and $\hat{G} \sim N(\overline{Y}_{..}, S_y^2)$.

(3) Generate B random samples of size N_1 and N_2 from fixed $\hat{F}$ and $\hat{G}$, respectively. The corresponding samples called the *bootstrap samples* are denoted by $\underline{X}^{*b} = (X_{11}^{*b}, X_{12}^{*b}, \cdots, X_{n_1 l_1}^{*b})$ and $\underline{Y}^{*b} = (Y_{11}^{*b}, Y_{12}^{*b}, \cdots, Y_{n_2 l_2}^{*b})$, $b = 1, 2, \cdots, B$. As the bootstrap samples for any b, we use $\underline{X}^*$ and $\underline{Y}^*$ instead of $\underline{X}^{*b}$ and $\underline{Y}^{*b}$, respectively.

(4) Compute $\hat{R}^{*b} = \Phi(\hat{\delta}^{*b})$, where

$$\hat{\delta}^{*b} = \frac{\overline{X}_{..}^{*b} - \overline{Y}_{..}^{*b}}{\sqrt{S_x^{2*b} + S_y^{2*b}}}, \quad \overline{X}_{..}^{*b} = \frac{1}{N_1} \sum_{i=1}^{n_1} \sum_{j=1}^{l_1} X_{ij}^{*b}, \quad \overline{Y}_{..}^{*b} = \frac{1}{N_2} \sum_{q=1}^{n_2} \sum_{r=1}^{l_2} Y_{qr}^{*b},$$

$$S_x^{2*b} = \frac{1}{N_1} \sum_{i=1}^{n_1} \sum_{j=1}^{l_1} (X_{ij}^{*b} - \overline{X}_{..}^{*b})^2 \text{ and } S_y^{2*b} = \frac{1}{N_2} \sum_{q=1}^{n_2} \sum_{r=1}^{l_2} (Y_{qr}^{*b} - \overline{Y}_{..}^{*b})^2.$$

We call

$\overline{X}_{..}^{*b}$, $\overline{Y}_{..}^{*b}$, S_x^{2*b}, S_y^{2*b} and $\hat{R}^{*b}$ by *bootstrap estimators* for $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ and R, respectively. As the bootstrap estimators for any b, we use $\overline{X}_{..}^{*}$, $\overline{Y}_{..}^{*}$, S_x^{2*}, S_y^{2*} and $\hat{R}^{*}$ instead of $\overline{X}_{..}^{*b}$, $\overline{Y}_{..}^{*b}$, S_x^{2*b}, S_y^{2*b} and $\hat{R}^{*b}$, respectively.

The following theorem is related to the weak law of large numbers and the convergence in distribution of bootstrap estimator $\hat{R}^{*} = \Phi(\hat{\delta}^{*})$.

Theorem. For given $\underline{X}$ and $\underline{Y}$ from the population 1 and 2, suppose that $\underline{X}^{*}$ and $\underline{Y}^{*}$ are the bootstrap samples of sizes N_1 and N_2 from the sample distribution function $\hat{F}$ and $\hat{G}$. Then the bootstrap estimator $\hat{R}^{*}$ is a consistent estimator of R.

Proof. For arbitrary positive ε,

$$P(\mid \overline{X}_{..}^{*} - \overline{X}_{..} \mid \geq \varepsilon) \leq \frac{E(\overline{X}_{..}^{*} - \overline{X}_{..})^2}{\varepsilon^2}$$

$$= \frac{E[E[(\overline{X}_{..}^{*} - \overline{X}_{..})^2 \mid \underline{X}\,]]}{\varepsilon^2}$$

$$= \frac{E(S_x^2)}{N_1 \varepsilon^2}$$

$$= \frac{(N_1 - 1)\sigma^2}{(\varepsilon N_1)^2} \rightarrow 0, \text{ as } N_1 \rightarrow \infty.$$

Also,

$$P(\mid S_x^{2*} - S_x^2 \mid \geq \varepsilon) \leq \frac{E(S_x^{2*} - S_x^2)^2}{\varepsilon^2}$$

$$= \frac{E[E[(S_x^{2*} - S_x^2)^2 \mid \underline{X}]]}{\varepsilon^2}$$

$$= \frac{2N_1 - 1}{(\varepsilon N_1)^2} E(S_x^4)$$

$$= \frac{(2N_1 - 1)(N_1^2 - 1)\sigma_x^4}{\varepsilon^2 N_1^4} \rightarrow 0, \text{ as } N_1 \rightarrow 0.$$

Therefore, $\overline{X}_{..}^{*}$ and S_x^{2*} converge in probability to μ_x and σ_x^2, respectively. Similarly, $\overline{Y}_{..}^{*}$ and S_y^{2*} converge in probability to μ_y and σ_y^2, respectively. Hence, $\hat{\delta}^{*}=\dfrac{\overline{X}_{..}^{*}-\overline{Y}_{..}^{*}}{\sqrt{S_x^{2*}+S_y^{2*}}}$ is a consistent estimator of δ. Since Φ is continuous function, $\hat{R}^{*}$ is a consistent estimator of R.

Note that the asymptotic distribution of $\hat{R}^{*}$ and $\hat{R}$ are the same under the assumptions of theorem.

3. Bootstrap Confidence Intervals for Reliability

In this section we construct approximate bootstrap confidence intervals for R. All confidence intervals are two-sided and equal-tailed with confidence level $100(1-2\alpha)\%$.

Before deriving bootstrap confidence intervals, we consider approximate confidence interval based on the normal approximation. Aminzadeh(1991) proved that $\hat{\delta}$ has asymptotic normal distribution with mean δ and variance $\sigma_{\delta}{}^{2}=\left(\dfrac{\sigma_x^2+\sigma_y^2}{K}+\dfrac{\delta^2}{2Q}\right)^{-1}$, where

$$K=\frac{(n_1-1)\sigma_A^2+\sigma_x^2}{N_1}+\frac{(n_2-1)\sigma_B^2+\sigma_y^2}{N_2},\quad Q=\frac{(\sigma_x^2+\sigma_y^2)^2}{f^{-1}\sigma_x^4+g^{-1}\sigma_y^4},$$

$$f=\frac{(k_1^2+1)(l_1-1)N_1}{N_1(k_1+n_1^{-1})^2+(1-n_1^{-1})(l_1-1)},\quad g=\frac{(k_2+1)^2(l_2-1)N_2}{N_2(k_2+n_2^{-1})^2+(1-n_2)(l_2-1)},\quad k_1=\frac{\sigma_A^2}{\sigma_e^2}\ \text{and}$$

$k_2=\dfrac{\sigma_B^2}{\sigma_\varepsilon^2}$. The asymptotic variance of $\hat{\delta}$ is estimated by $\widehat{\sigma_{\delta}}^{2}=\left(\dfrac{\widehat{\sigma_x^2}+\widehat{\sigma_y^2}}{\hat{K}}+\dfrac{\hat{\delta}^2}{2\hat{Q}}\right)^{-1}$, where $\hat{K}$ and $\hat{Q}$ are computed by using $\widehat{\sigma_x^2}$, $\widehat{\sigma_y^2}$, $\widehat{\sigma_A^2}=\dfrac{(S_A^2-S_e^2)}{n_1}$, $\widehat{\sigma_B^2}=\dfrac{(S_b^2-S_\varepsilon^2)}{n_2}$, $\hat{k}_1=\dfrac{(s_A^2/S_e^2-1)}{n_1}$ and $\hat{k}_2=\dfrac{(S_B^2/S_\varepsilon^2-1)}{n_2}$ instead of σ_x^2, σ_y^2, σ_A^2, σ_B^2, k_1 and k_2, respectively.

Hence, $100(1-2\alpha)\%$ confidence interval for R is given by

$$[\Phi(\hat{\delta}+z^{(\alpha)}\cdot\widehat{\sigma_\delta}),\ \Phi(\hat{\delta}+z^{(1-\alpha)}\cdot\widehat{\sigma_\delta})]\quad, \tag{3.1}$$

where $z^{(\alpha)}$ is the $100\cdot\alpha$ percentile of standard normal distribution.

3.1. Percentile method

The confidence interval by the bootstrap percentile method(percentile interval) is obtained by percentiles of the empirical bootstrap distribution of $\hat{R}^*$. Let $\hat{H}^*$ be the empirical cumulative distribution function of $\hat{R}^*$. Then it is constructed by $\hat{H}^*(s)=\frac{1}{B}\sum_{b=1}^{B} I(\hat{R}^{*b}\le s)$, where s is arbitrary real value and $I(\cdot)$ is an indicator function. And let $\hat{H}^{*^{-1}}(\alpha)$ be the 100α empirical percentile of $\hat{R}^*$ given by

$$\hat{H}^{*^{-1}}(\alpha)=\inf\{s:\ \hat{H}^*(s)\ \ge\ \alpha\}. \tag{3.2}$$

That is, $\hat{H}^{*^{-1}}(\alpha)$ is the $B\alpha$th value in the ordered list of the B replications of $\hat{R}^{*b}$. If $B\alpha$ is not an integer, we can take the largest integer that less than or equal to $(B+1)\alpha$. Then $100(1-2\alpha)\%$ percentile interval for R is approximated by

$$(\hat{H}^{*^{-1}}(\alpha)\ ,\ \hat{H}^{*^{-1}}(1-\alpha)). \tag{3.3}$$

3.2 Bias correct method

The BC method adjusts a possible bias in estimating R. The bias correction is given by

$$\widehat{z_0}\ =\ \Phi^{-1}(\hat{H}^*(\hat{R}))\ =\ \Phi^{-1}\left[\frac{1}{B}\sum_{b=1}^{B} I(\hat{R}^{*b}\le\hat{R})\right], \tag{3.4}$$

where $\Phi^{-1}(\cdot)$ indicates the inverse function of the standard normal cumulative distribution function. That is, $\widehat{z_0}$ is the discrepancy between the medians of $\hat{R}^*$ and $\hat{R}$ in normal unit. Therefore, we have $100(1-2\alpha)\%$ approximate BC interval for R given by

$$(\hat{H}^{*^{-1}}(\alpha_1)\ ,\ \hat{H}^{*^{-1}}(\alpha_2)), \tag{3.5}$$

where $\alpha_1=\Phi(2\,\widehat{z_0}+z^{(\alpha)})$ and $\alpha_2=\Phi(2\,\widehat{z_0}+z^{(1-\alpha)})$.

3.3 Bias correct acceleration method

The BCa method corrects both the bias and standard error for $\hat{R}$. The confidence interval by BCa method(BCa interval) requires to calculate the bias-correction constant $\widehat{z_0}$ and the acceleration constant $\hat{a}$. In fact, the bias-correction constant $\widehat{z_0}$ is the same as that of BC method. And $\hat{a}$, measured on a normalized scale, refers to the rate of change of the standard error of $\hat{R}$ with respect to the true reliability R.

For the parametric bootstrap method, all calculations relate only to the sufficient statistic $\overline{X}_{..}, S_x^2, \overline{Y}_{..}$ and S_y^2 for μ_x, σ_x^2, μ_y and σ_y^2, respectively. Of course, $\overline{X}_{..}$, S_x^2, $\overline{Y}_{..}$ and S_y^2 are distributed $N(\mu_x, \sigma_x^2/N_1)$, $(\sigma_x^2/N_1)\cdot\chi^2(N_1-1)$, $N(\mu_y, \sigma_y^2/N_2)$ and $(\sigma_y^2/N_2)\cdot\chi^2(N_2-1)$, respectively. Also, $\overline{X}_{..}$, S_x^2, $\overline{Y}_{..}$ and S_y^2 are stochastically independent. Let $\hat{\underline{\eta}}' = (\overline{X}_{..}, S_x^2, \overline{Y}_{..}, S_y^2)$ and $\underline{\eta}' = (\mu_x, \sigma_x^2, \mu_y, \sigma_y^2)$. Then the joint probability density function of $\hat{\underline{\eta}}'$ can be written as

$$f_{\underline{\eta}}(\hat{\underline{\eta}}') = f_0(\hat{\underline{\eta}}')\exp[g_0(\hat{\underline{\eta}}', \underline{\eta}') - \Psi_0(\underline{\eta}')], \tag{3.6}$$

where

$$f_0(\hat{\underline{\eta}}') = \left[2\pi\Gamma\left(\frac{N_1-1}{2}\right)\cdot\Gamma\left(\frac{N_2-1}{2}\right)\cdot 2^{\frac{N_1+N_2-2}{2}}\right]^{-1}\cdot N_1^{N_1/2}N_2^{N_2/2},$$

$$g_0(\underline{\eta}', \hat{\underline{\eta}}') = \frac{2N_1\mu_x\overline{X}_{..} - N_1\overline{X}_{..}^{\,2} - N_1S_x^2}{2\sigma_x^2} + \frac{2N_2\mu_y\overline{Y}_{..} - N_2\overline{Y}_{..}^{\,2} - N_2S_y^2}{2\sigma_y^2}$$

$$+\frac{N_1-3}{2}\cdot\log(S_x^2) + \frac{N_2-3}{2}\cdot\log(S_y^2)$$

and $$\Psi_0(\underline{\eta}') = \frac{N_1\mu_x^2}{2\sigma_x^2} + \frac{N_1}{2}\log(\sigma_x^2) + \frac{N_2\mu_y^2}{2\sigma_y^2} + \frac{N_2}{2}\log(\sigma_y^2).$$

For multiparameter family case, we will find $\hat{a}$ by following Stein's construction(1956). That is, we replace the multiparameter family $\Im = \{f_{\underline{\eta}}(Z)\}$ by the least favorable one parameter family $\hat{\Im} = \{\widetilde{f_{\underline{\eta}}(Z)} \equiv f_{\hat{\underline{\eta}}+\lambda\hat{\underline{\omega}}}(Z)\}$, where $Z = (X, Y)$. Then we first obtain $\hat{\underline{\omega}}$ such that the least favorable direction at $\underline{\eta} = \hat{\underline{\eta}}$ is defined to be $\hat{\underline{\omega}} = (\ddot{\ell}_{\hat{\underline{\eta}}})^{-1}\hat{\nabla}_{\hat{\underline{\eta}}'}$ where $\ddot{\ell}_{\hat{\underline{\eta}}}$ is Fisher information matrix and $\hat{\nabla}_{\hat{\underline{\eta}}}$ is the gradient of δ given by $\hat{\nabla}_{\hat{\underline{\eta}}} = \frac{\partial\delta}{\partial\underline{\eta}}\Big|_{\underline{\eta}=\hat{\underline{\eta}}}$. By some algebraic calculation, we have

$$\ddot{\ell}_{\hat{\underline{\eta}}} = \begin{pmatrix} N_1/S_x^2 & 0 & 0 & 0 \\ 0 & N_2/S_y^2 & 0 & 0 \\ 0 & 0 & N_1/(2S_x^2) & 0 \\ 0 & 0 & 0 & N_2/(2S_y^2) \end{pmatrix}$$

and

$$\hat{\nabla}_{\hat{\underline{\eta}}} = \begin{pmatrix} \frac{1}{\sqrt{S_x^2+S_y^2}} \\ -\frac{1}{\sqrt{S_x^2+S_y^2}} \\ -\frac{\overline{X}_{..}-\overline{Y}_{..}}{2(S_x^2+S_y^2)^{3/2}} \\ -\frac{\overline{X}_{..}-\overline{Y}_{..}}{2(S_x^2+S_y^2)^{3/2}} \end{pmatrix}.$$

Hence, we have $\hat{\underline{\omega}}' = (W_1, W_2, W_3, W_4)$, where

$W_1 = \frac{S_x^2}{N_1\sqrt{S_x^2+S_y^2}}$, $W_2 = -\frac{S_y^2}{N_2\sqrt{S_x^2+S_y^2}}$, $W_3 = -\frac{S_x^4(\overline{X}_{..}-\overline{Y}_{..})}{N_1(S_x^2+S_y^2)^{3/2}}$ and

$W_4 = -\frac{S_y^4(\overline{X}_{..}-\overline{Y}_{..})}{N_2(S_x^2+S_y^2)^{3/2}}$. Following Efron(1987), $\hat{a}$ can be obtained by

$$\hat{a} = \frac{1}{6}\cdot\frac{\hat{\Psi}^{(3)}(0)}{(\hat{\Psi}^{(2)}(0))^{(3/2)}}, \tag{3.7}$$

where $\hat{\Psi}^{(j)}(0) = \frac{\partial^j \Psi_0(\hat{\underline{\eta}}+\lambda\hat{\underline{\omega}})}{\partial\lambda^j}\big|_{\lambda=0}$. By calculating $\hat{\Psi}^{(j)}(\cdot)$ and $\hat{\underline{\omega}}$, we can obtain $\hat{a}$ by

$$\hat{a} = \frac{1}{6}\cdot\frac{s_3+s_4}{(s_1+s_2)^{3/2}}, \tag{3.8}$$

where

$$s_1=\frac{N_1}{2}\left[-\frac{W_3^2}{S_x^4}+\frac{2(W_1S_x^2)^2-4W_1W_3\,\overline{X}_{..}S_x^2+2(W_3\,\overline{X}_{..})^2}{S_x^{\,6}}\right],$$

$$s_2=\frac{N_2}{2}\left[-\frac{W_4^2}{S_y^{\,4}}+\frac{2(W_2S_y^2)^2-4W_2W_4\,\overline{Y}_{..}S_y^2+2(W_4\,\overline{Y}_{..})^2}{S_y^{\,6}}\right],$$

$$s_3=\frac{N_1}{2}\left[-\frac{2\,W_3^{\,3}}{S_x^{\,6}}+\frac{-6(W_1S_x^2)^2W_3+12W_1\,W_3^{\,2}\,\overline{X}_{..}S_x^2-6\,W_3^{\,3}\,\overline{X}_{..}^{\,2}}{S_x^{\,8}}\right]$$

and

$$s_4=\frac{N_2}{2}\left[-\frac{2\,W_4^{\,3}}{S_y^{\,6}}+\frac{-6(W_2S_y^2)^2W_4+12W_2\,W_4^{\,2}\,\overline{Y}_{..}S_y^2-6\,W_4^{\,3}\,\overline{Y}_{..}^{\,2}}{S_y^{\,8}}\right].$$

Therefore, we have $100(1-2\alpha)\%$ approximate *BCa* interval for R by

$$[\Phi(\hat{H}^{*^{-1}}(\alpha_3)),\ \Phi(\hat{H}^{*^{-1}}(\alpha_4))], \tag{3.9}$$

where $\alpha_3=\Phi\left[\hat{z}_0+\frac{\hat{z}_0+z^{(\alpha)}}{1-\hat{a}(\hat{z}_0+z^{(\alpha)})}\right]$, $\alpha_4=\Phi\left[\hat{z}_0+\frac{\hat{z}_0+z^{(1-\alpha)}}{1-\hat{a}(\hat{z}_0+z^{(1-\alpha)})}\right]$, and $\hat{z}_0$ is the same as that of *BC* method.

3.4 Percentile- t method

The confidence interval by the bootstrap percentile- t method(percentile- t interval) is constructed by using the bootstrap distribution of an approximately pivotal quantity for $\hat{\delta}$ instead of the bootstrap distribution of $\hat{\delta}$. We define an approximate bootstrap pivotal quantity for $\hat{\delta}$ by

$$\hat{\delta}^*_{STUD}=\frac{\hat{\delta}^*-\hat{\delta}}{\widehat{\sigma_\delta}^*}, \tag{3.10}$$

where $\widehat{\sigma_\delta}^*$ is the bootstrap estimator of σ_δ, that is,

$$\widehat{\sigma_\delta^{2*}}=\left(\frac{\widehat{\sigma_x}^{2*}+\widehat{\sigma_y}^{2*}}{\hat{K}^*}+\frac{\hat{\delta}^{2*}}{2\,\hat{Q}^*}\right)^{-1}, \tag{3.11}$$

where $\widehat{\sigma_x}^{2*}$, $\widehat{\sigma_y}^{2*}$, $\hat{K}^*$, $\hat{\delta}^{2*}$ and $\hat{Q}^*$ are the bootstrap version of $\widehat{\sigma_x}^2$, $\widehat{\sigma_y}^2$, $\hat{K}$, $\hat{\delta}^2$

and $\widehat{Q}$, respectively. We compute the empirical distribution function $\widehat{H}^*_{STUD}$ of $\widehat{\delta}^*_{STUD}$ by

$$\widehat{H}^*_{STUD}(s) = \frac{1}{B}\sum_{b=1}^{B} I(\ \widehat{\delta}^{*b}_{STUD} \le s), \tag{3.12}$$

for all s. Let $\widehat{H}^{*^{-1}}_{STUD}$ denote 100α empirical percentile of $\widehat{\delta}^*_{STUD}$. And we compute $\widehat{H}^{*^{-1}}_{STUD}(\alpha)$ by

$$\widehat{H}^{*^{-1}}_{STUD}(\alpha) = \inf\{s : \ \widehat{H}^*_{STUD}(s) \ge \alpha\}. \tag{3.13}$$

That is, $\widehat{H}^{*^{-1}}_{STUD}(\alpha)$ is the $B\alpha^{\text{th}}$ value in the ordered list of the B replications of $\widehat{\delta}^*_{STUD}$. Then we have $100(1-2\alpha)\%$ approximate percentile-t interval for R by

$$\left[\Phi(\widehat{\delta} + \widehat{\sigma}_{\widehat{\delta}} \cdot \ \widehat{H}^*_{STUD}{}^{-1}(\alpha))\ ,\ \Phi(\widehat{\delta} + \widehat{\sigma}_{\widehat{\delta}} \cdot \ \widehat{H}^*_{STUD}{}^{-1}(1-\alpha))\right]. \tag{3.14}$$

4. Monte Carlo Simulation Studies

To compare the proposed bootstrap confidence interval estimates with the confidence interval estimate based on asymptotic normal distribution, we will compute the results obtained in Section 3. The methods are compared mainly based on coverage probability (CP) and interval length (L). The normal random numbers were generated by IMSL subroutine RNNOF. We use the true reliabilities R=0.3, 0.5, 0.7, 0.9 and batch sizes $l_1 = l_2$=3, 5, 10, 20 with fixed $n_1 = n_2 = 3$. We also use confidence level $(1-2\alpha)$=0.90. For given independent random samples, the approximate confidence intervals for each method were constructed with bootstrap replications B= 1000 times. And the Monte Carlo samplings were repeated 500 times. Table 1 provides the coverage probability (CP) for all cases. Table 2 reports the length (L) of all intervals.

Based on the Monte Carlo study, our conclusions are summarized below.

(1) In the most cases, the proposed bootstrap confidence intervals are better than that of the interval based on asymptotic normal distribution for all R even in small batch size in the sense of CP, which is well illustrated in Figure 1. In particular, BC and BCa intervals work well.

(2) In the most cases, the values of L for all approximate confidence intervals tend to decrease as R deviates from 0.5, which is illustrated in Figure 2. As one might expect, the

value of L for the interval based on Aminzadeh's method is shorter than those of the intervals based on bootstrap methods.

(3) The values of CP for all approximate confidence intervals converge to true coverage level $(1-2\alpha)$ as the batch size increases, which is illustrated in Figure 3. Also, the values of CP for the intervals based on all bootstrap methods are closer to true confidence level than that of the interval based on Aminzadeh's method for most batch sizes. In particular, the values of CP for bootstrap methods are much better than that of the interval based on asymptotic normal distribution for large R value.

(4) An inspection of Figure 4 reveals that the values of L for the approximate intervals based on all methods decrease as batch size increases, and might converge to the true interval length.

References

[1] Aminzadeh, M. S. (1991). Confidence Bounds for Pr(X>Y) in 1-Way ANOVA Random Model, *IEEE*, Vol. 40, NO 5, 537-540.

[2] Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife, *The Annals of Statistics*, 7, 1-26.

[3] Efron, B. (1981). Nonparametric Standard Errors and Confidence Intervals (With Discussion), *Canadian Journal of Statistics*, 9, 139-172.

[4] Efron, B. (1982). The Jackknife, The Bootstrap and Other Resampling Plans, SIAM, Philadelphia, 38.

[5] Efron, B. (1987). Better Bootstrap Confidence intervals(With discussion), *Journal of the American Statistician Association*, 82, 171-200.

[6] Guttman, I., Johnson, R., Bhattacharyya, G. K., and Reisser, B. (1988). Confidence Limits for Stress-Strength Models with explanatory variables, *Technometrics*, 30, 161-168.

[7] Hall, P. (1988a). Theoretical Comparison of Bootstrap Confidence Intervals (With Discussion), *The Annals of Statistics*, 16, 927-953.

[8] Reisse, B. and Guttman, I. (1986). Statistical Inference for Pr(Y<X), *Technometrics*, 28, 253-257.

[9] Stein, C. (1956). Efficient nonparametric testing and estimation, Procedure of the third Berkeley Symposium, University of California Press, 187-196, Berkeley.

Table 1. Coverage Probability (CP)

Batch size	R	Aminzadeh	Percentile	BC	BCa	Percentile-t
3	0.3	.7840	.8460	.8980	.8680	.9580
	0.5	.7820	.8180	.8940	.8640	.9440
	0.7	.7800	.8460	.9060	.8700	.9620
	0.9	.6580	.8100	.8720	.8460	.9220
5	0.3	.7840	.8620	.8900	.8780	.9040
	0.5	.8360	.8660	.9060	.8820	.9200
	0.7	.8360	.8840	.9060	.8980	.9440
	0.9	.7480	.8480	.8820	.8780	.8920
10	0.3	.8580	.8920	.9020	.8960	.9120
	0.5	.9140	.9100	.9300	.8980	.9380
	0.7	.8420	.8580	.8820	.8720	.9060
	0.9	.7420	.8520	.8800	.8720	.8880
20	0.3	.8460	.8820	.9040	.8840	.9140
	0.5	.8820	.8760	.8920	.8680	.8960
	0.7	.8820	.8860	.9060	.8900	.9260
	0.9	.7960	.9020	.8980	.9080	.8880

Table 2. Interval Length (L)

Batch size	R	Aminzadeh	Percentile	BC	BCa	Percentile-t
3	0.3	.3240	.3721	.3871	.3792	.5392
	0.5	.3624	.4259	.4322	.4291	.5348
	0.7	.3179	.3765	.3911	.3849	.5303
	0.9	.1751	.2047	.2428	.2222	.5244
5	0.3	.2663	.3025	.3074	.3040	.3550
	0.5	.3068	.3361	.3380	.3375	.3756
	0.7	.2702	.3012	.3071	.3055	.3624
	0.9	.1449	.1715	.1895	.1791	.2945
10	0.3	.1989	.2153	.2171	.2161	.2336
	0.5	.2295	.2397	.2406	.2408	.2514
	0.7	.1988	.2147	.2170	.2162	.2344
	0.9	.0997	.1241	.1312	.1261	.1586
20	0.3	.1440	.1546	.1553	.1557	.1601
	0.5	.1652	.1686	.1690	.1692	.1717
	0.7	.1428	.1537	.1547	.1543	.1588
	0.9	.0735	.0931	.0957	.0934	.1038

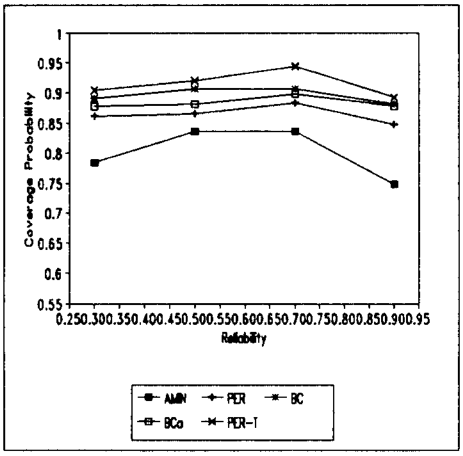

Figure 1 : Coverage probability versus reliability for batch size 5

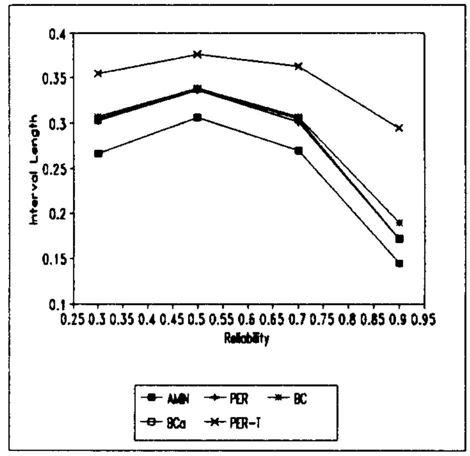

Figure 2 : Interval length versus reliability for batch size 5

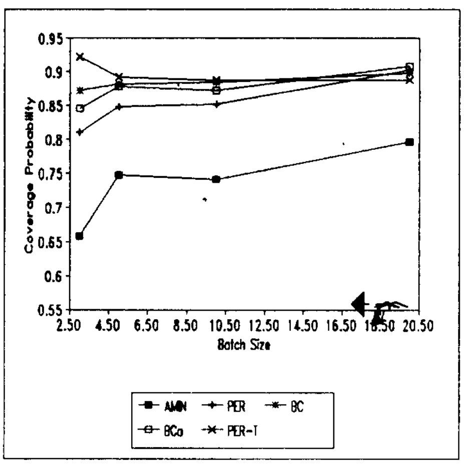

Figure 3 : Coverage probability versus batch size for reliability .9

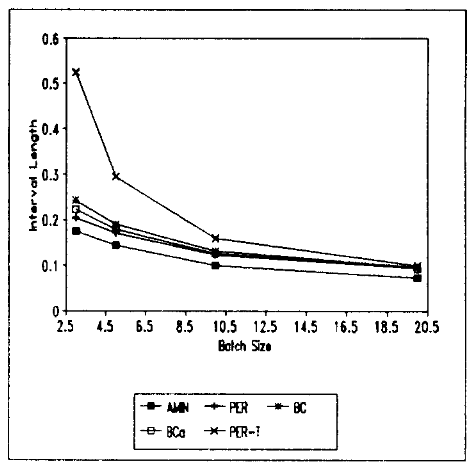

Figure 4 : Interval length versus batch size for reliability .9