BETTER NONPARAMETRIC BOOTSTRAP CONFIDENCE INTERVALS FOR THE CORRELATION COEFFICIENT

Peter Hall
Michael A. Martin

The Australian National University

William R. Schucany
Southern Methodist University

SMU/DS/TR-221

May 1988

BETTER NONPARAMETRIC BOOTSTRAP CONFIDENCE INTERVALS FOR THE CORRELATION COEFFICIENT

Peter Hall, Michael A. Martin and William R. Schucany

SUMMARY. We respond to criticism levelled at bootstrap confidence intervals for the correlation coefficient by recent authors by arguing that in the correlation coefficient case, non-standard methods should be employed. We propose two such methods. The first is a bootstrap coverage correction algorithm using iterated bootstrap techniques (Hall, 1986; Beran, 1987a) applied to ordinary percentile method intervals (Efron, 1979), giving intervals with high coverage accuracy and stable lengths and endpoints. The simulation study carried out for this method gives results for sample sizes 8, 10, and 12 in three parent populations. The second technique involves the construction of percentile-t bootstrap confidence intervals for a transformed correlation coefficient, followed by an inversion of the transformation, to obtain "transformed percentile-t" intervals for the correlation coefficient. In particular, Fisher's z-transformation is used, and nonparametric delta method and jackknife variance estimates are used to Studentize the transformed correlation coefficient, with the jackknife-Studentized transformed percentile-t interval yielding the better coverage accuracy, in general. Percentile-t intervals constructed without first using the transformation perform very poorly, having large expected lengths and erratically fluctuating endpoints. The simulation study illustrating this technique gives results for sample sizes 10, 15, and 20 in four parent populations. Our techniques provide confidence intervals for the correlation coefficient which have good coverage accuracy (unlike ordinary percentile intervals), and stable lengths and endpoints (unlike ordinary percentile-t intervals).

KEY WORDS: coverage, Fisher's z, jackknife, nonparametric delta method, percentile method, percentile-t method, prepivoting.

Peter Hall is Reader in the Department of Statistics, Australian National University, Canberra, Australia, 2601, Michael Martin is Research Scholar in the Department of Statistics, Australian National University, Canberra, Australia, 2601, and William Schucany is Professor in the Department of Statistical Science, Southern Methodist University, Dallas, TX 75275. This research is partially supported by ONR Contract N00014-85-K-0340, and was carried out while Professor Schucany was visiting the Australian National University.

1. INTRODUCTION

The problem of constructing bootstrap confidence intervals for the correlation coefficient has received considerable recent attention. One very popular bootstrap algorithm, the percentile method, has been criticized because of coverage inaccuracy (see Rasmussen, 1987 and Lunneborg, 1985). Efron (1981) has pointed out that another favored technique, the percentile-t method, fails strikingly when used with the correlation coefficient. Efron employed a jackknife variance estimate to standardize for scale. We have repeated Efron's experiment, using both the jackknife and the nonparametric delta method variance estimates to Studentize, with almost identical results. The percentile-t bootstrap method founders in both cases because the lengths of these confidence intervals fluctuate erratically, their large expectations and standard deviations reflecting the inaccuracy of these particular variance estimates in small samples.

In the present paper, we argue that the correlation coefficient is better treated by specialized means. The percentile-t method, which is becoming accepted as an accurate and standard tool for simpler problems (Bickel and Freedman, 1981; Beran, 1987a, 1987b, 1987c; Schenker, 1987; Hall, 1987, 1988; Singh, 1987), is apparently less suitable for more complex statistics such as the correlation coefficient. We propose two practical and accurate alternative methods. One is based on coverage-correction of the percentile method and is discussed in Section 2; the other is an application of percentile-t to a transformed correlation coefficient, and is treated in Section 3. Both promise to be useful in other complex interval construction problems. The performance of each method is illustrated in a simulation study, which shows that each type of interval has good coverage accuracy (unlike ordinary percentile intervals), and good stability properties (unlike ordinary percentile-t intervals).

2. COVERAGE-CORRECTED PERCENTILE METHOD

Coverage correction by bootstrap iteration has been studied before (Hall,1986; Beran, 1987a; Hall and Martin, 1988), but as a device for improving on methods which already have reasonable coverage accuracy. In particular, it is usually advocated for percentile-t confidence intervals. Our approach in this paper is different. Acknowledging that the percentile method has good stability properties - that is, successive bootstrap estimates of percentile method critical points do not fluctuate erratically - we adjust the interval so as to improve its much-criticized coverage accuracy. By so doing, we manage to retain the interval's stability, yet greatly enhance its coverage performance. In simpler problems, such as estimation of a mean, it makes little sense to coverage-correct a percentile-method interval, when the percentile-t interval has superior coverage accuracy to start with, and almost as good stability.

To explain our coverage correction method, we first show how to construct a basic $(1-\alpha)$-level percentile interval (Efron, 1979). Let θ be the unknown parameter (correlation coefficient, in our case), and let $\hat{\theta}$ be its estimate based on an n-sample, $\mathcal{X}$. Let $\mathcal{X}^*$ be an n-sample (called a resample) drawn with replacement from $\mathcal{X}$, and let $\hat{\theta}^*$ be the version of $\hat{\theta}$ computed using $\mathcal{X}^*$ instead of $\mathcal{X}$. By repeated resampling from $\mathcal{X}$, find $\hat{x}_\alpha$ and $\hat{y}_\alpha$ such that

$$P(\hat{\theta}^* \leq \hat{x}_\alpha \mid \mathcal{X}) = \alpha/2 = P(\hat{\theta}^* \geq \hat{y}_\alpha \mid \mathcal{X}).$$

The equal-tailed percentile interval is $I(\mathcal{X}) \equiv [\hat{x}_\alpha, \hat{y}_\alpha]$. Other types of percentile method interval are discussed by Hall (1988).

There are at least two ways of coverage correcting this interval by bootstrap iteration. One operates at the level of the quantile (Hall, 1986), the other by adjustment of nominal coverage (Beran, 1987a).(See also Loh, 1987.) In practice they give almost identical numerical results. For the former interval, define $\hat{x}_\alpha(t) \equiv \hat{x}_\alpha - t$, and $\hat{y}_\alpha(t) \equiv \hat{y}_\alpha + t$; for the latter, put $\hat{x}_\alpha(t) \equiv \hat{x}_{\alpha-t}$, and $\hat{y}_\alpha(t) \equiv \hat{y}_{\alpha-t}$. The aim is to choose t (usually small and non-zero) so as to correct for coverage error. The first step is to estimate coverage, as follows. Let $I_t(\mathcal{X}) \equiv [\hat{x}_\alpha(t), \hat{y}_\alpha(t)]$.

(When t=0, $I_t(\mathcal{X})$ is just the percentile interval, $I(\mathcal{X})$.) Our estimate of the true coverage, $\pi(t)$, of $I_t(\mathcal{X})$ is

$$\hat{\pi}(t) \equiv P\{\ \hat{\theta} \in I_t(\mathcal{X}^*) \mid \mathcal{X}\ \},$$

which we obtain by simulation. Usually, $\hat{\pi}(t)$ will be close to the nominal coverage, $1-\alpha$, when t is close to zero. Choose $t = \hat{t}$ to solve the equation, $\hat{\pi}(t) = 1-\alpha$. Then $I_{\hat{t}}(\mathcal{X})$ is the coverage-corrected interval.

It may be shown as in Hall (1988) that the basic percentile interval has coverage error of order n^{-1}, and that the coverage-corrected interval has coverage error of order n^{-2}; see Hall and Martin (1988). However, it could be unwise to rely too heavily on this type of asymptotic result when dealing with the correlation coefficient in small samples, especially given that the highly erratic behavior of percentile-t intervals is certainly not apparent from Edgeworth expansions.

In practice, the iterated percentile method for constructing confidence intervals is relatively computationally expensive, when compared with ordinary percentile or percentile-t methods. In particular, if at each resampling operation, B resamples are drawn, then the time taken to construct an iterated percentile method interval is roughly proportional to B^2 compared with B for an ordinary percentile or percentile-t interval. Depending on the coverage accuracy required, it is our experience that one iteration is usually enough to substantially improve coverage accuracy to a reasonable level.

We conclude this section with a simulation study of iterated percentile method intervals for the correlation coefficient, ρ. Table 2.1 contains a summary of simulations from three parent populations, using sample sizes n = 8, 10, and 12. For simulation 1, the sample, $\mathcal{X} \equiv \{(X_1,Y_1),\ldots,(X_n,Y_n)\}$, is drawn from a bivariate folded normal population, denoted by

$|N(0,I)|$, with $\rho = 0$; for simulation 2, $\rho = 0.5$, with the sample, $\mathcal{X}$, being drawn such that $X_i = Z_{1i} + W_i$, and $Y_i = Z_{2i} + W_i$, where Z_{1i}, Z_{2i}, and W_i are distributed as independent folded normal variates, for i=1,...,n; and for simulation 3, the sample, $\mathcal{X}$, is drawn from a bivariate lognormal population with $\rho = (1+e^{1/2})^{-1} \approx 0.37754$. In each case, adjustment was made directly to nominal coverage, B=299 resamples were drawn in each resampling operation, and each entry was based on 1000 samples for each value of n. The simulation study indicates that one bootstrap iteration does improve coverage accuracy dramatically. Especially significant is the performance of the iterated percentile method when the parent population is bivariate lognormal. In this instance, the iterated percentile method interval has much better coverage accuracy than the percentile-t methods considered in Section 3. In fact, coverage of the iterated percentile method intervals is correct; that is, not significantly different from the nominal coverage in each of the nine combinations of sample size and parent populations considered. What is more, this coverage correction is realized without unreasonable increase in average interval length over the ordinary percentile method.

TABLE 2.1 NEAR HERE

---

3. TRANSFORMED PERCENTILE-T METHOD

The percentile-t method of constructing confidence intervals has been discussed widely of late, usually applied to statistics for which there is available a reasonably stable variance estimate (see, for example, Bickel and Freedman, 1981; Beran, 1987a, 1987b, 1987c; Hall, 1987, 1988; Singh, 1987). For instance, it is known (Hall, 1988) that the percentile-t method works well when constructing confidence intervals for the Studentized mean. The performance of percentile-t intervals for a parameter, θ, depends largely on how well we are able to estimate the variance of the estimator, $\hat{\theta}$. In the case of the correlation coefficient, the use of standard estimates of variance, such as the jackknife estimate (Miller, 1974), or the nonparametric delta method estimate (Efron and Gong, 1983, p.40), to Studentize, results in percentile-t intervals with erratically varying lengths and endpoints. We propose to 1) transform the correlation coefficient to a statistic for which there is a reasonably stable variance estimate, 2) use the percentile-t method to construct a (1-α)-level confidence interval for the transformed correlation coefficient, and then 3) invert into a (1-α)-level confidence interval for the correlation coefficient.

First recall the construction of a basic (1-α)-level percentile-t interval. Let θ be the unknown parameter, and let $\hat{\theta}$ and $\hat{\sigma}^2$ be estimates of θ and the variance of $\hat{\theta}$, respectively, based on an n-sample, $\mathcal{X}$. Let $\mathcal{X}^*$ be a same-size resample, drawn with replacement from $\mathcal{X}$, and let $\hat{\theta}^*$ and $\hat{\sigma}^{2*}$ be versions of $\hat{\theta}$ and $\hat{\sigma}^2$ computed using $\mathcal{X}^*$ rather than $\mathcal{X}$. By repeated resampling from $\mathcal{X}$, find $\hat{s}_\alpha$ and $\hat{t}_\alpha$ such that

$$P\{(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^* \le \hat{s}_\alpha \mid \mathcal{X}\} = \alpha/2 = P\{(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^* \ge \hat{t}_\alpha \mid \mathcal{X}\}.$$

The equal-tailed percentile-t interval is $J(\mathcal{X}) \equiv [\,\hat{\theta} - \hat{\sigma}\hat{t}_\alpha,\ \hat{\theta} - \hat{\sigma}\hat{s}_\alpha]$.

Now consider the transformation of the correlation coefficient, ρ. Fisher's (1921) transformation of the sample correlation coefficient, $\hat{\rho}$, assuming the underlying population is bivariate normal, has variance almost independent of ρ. Let $Z \equiv \log\{(1+\hat{\rho})/(1-\hat{\rho})\}/2$. Then, for even fairly small samples from a bivariate normal population, Z is approximately normally distributed with mean $\xi \equiv \log\{(1+\rho)/(1-\rho)\}/2$ and variance $(n-3)^{-1}$; see Kendall and Stuart

(1977, vol.2, p.312). Efron (1982) and Efron and Gong (1983) suggest that, under the assumption that samples are drawn from a bivariate normal population, the percentile method can be used to construct good confidence intervals for ρ since it is invariant under monotone transformations, and the z-transformed correlation coefficient has almost constant variance. However, this does not resolve the problem of what to do when the underlying population is not bivariate normal. We suggest that Fisher's z-transformation can still be used when the parent population is non-normal, that a percentile-t method be used to construct a $(1-\alpha)$-level confidence interval for ξ, and this interval transformed back into a $(1-\alpha)$-level confidence interval for ρ.

To construct a percentile-t interval for ξ, one needs an estimate of the variance of Z. Simple calculations show that, asymptotically, $n^{1/2}(Z-\xi)$ is distributed with mean zero and variance

$$\rho^2(1-\rho^2)^{-2}\{\mu_{22}\mu_{11}^{-2} + (\mu_{40}\mu_{20}^{-2} + \mu_{04}\mu_{02}^{-2} + 2\mu_{22}\mu_{20}^{-1}\mu_{02}^{-1})/4 - (\mu_{31}\mu_{20}^{-1}\mu_{11}^{-1} + \mu_{13}\mu_{02}^{-1}\mu_{11}^{-1})\}, \qquad (3.1)$$

provided the bivariate moments μ_{ij}, for $i,j=0,\ldots,4$, exist. (When the underlying population is bivariate normal, the expression (3.1) equals one.) The obvious plug-in estimate of (3.1) is just the nonparametric delta method estimate of the variance of the transformed correlation coefficient. Unfortunately, the nonparametric delta method does not provide a very stable estimate of (3.1) in small samples, because of the difficulty in estimating high-order moments accurately. Miller(1974, Section 5.3) advocates using the jackknife estimate of the variance of the transformed statistic. He notes that "transformations are needed to keep the jackknife on scale and thus prevent distortion of the results". Therefore, we also investigate the use of the jackknife estimate of the variance of Z in constructing percentile-t intervals for the transformed correlation coefficient.

A particularly attractive feature of Fisher's z-transformation is that, since it is one-to-one, strictly increasing and maps [-1,1] onto the real line, all the intervals obtained for ρ are constrained to lie within [-1,1]. We have found that constructing percentile-t intervals for ρ without first transforming the correlation coefficient results in occasional extremely long intervals that *contain* [-1,1]. In principle, rather than using Fisher's z-transformation, we could have used almost any one-to-one, strictly increasing function mapping [-1,1] onto the real line.

We conclude this section with a simulation study comparing the performance of four percentile-t methods. Four sets of simulations were carried out using different parent populations. In the first, the sample, $\mathcal{X} \equiv \{(X_1,Y_1),...,(X_n,Y_n)\}$, was drawn from a bivariate normal population with $\rho = 0$; in the second, $\mathcal{X}$ was drawn from a bivariate folded normal population with $\rho = 0$; in the third, $\mathcal{X}$ was drawn such that $X_i = Z_{1i} + W_i$, and $Y_i = Z_{2i} + W_i$, where Z_{1i}, Z_{2i}, and W_i are distributed as independent folded normal variates, for i=1,...,n, resulting in $\rho = 0.5$; and in the fourth, $\mathcal{X}$ was drawn from a bivariate lognormal distribution with $\rho = (1+e^{1/2})^{-1} \approx 0.37754$. The five types of interval constructed were as follows: (i) Ordinary percentile-t interval Studentizing with the nonparametric delta method estimate of the variance of $\hat{\rho}$; (ii) As in (i), except Studentizing with the jackknife estimate of the variance of $\hat{\rho}$; (iii) Ordinary Percentile method interval; (iv) Transformed percentile-t interval, involving Fisher's z, Studentizing with the nonparametric delta method estimate of the variance of Z (sample moments in (3.1)); and (v) As in (iv), except Studentizing with the jackknife estimate of the variance of Z. Throughout, each entry is based on 1000 samples for each of the sample sizes n = 10, 15, and 20, with B=299 resamples being drawn at each resampling operation. Results are given for nominal coverages 90%, 95%, and 99%.

TABLES 3.1-3.4 NEAR HERE

Throughout Tables 3.1 - 3.4 estimated coverages that are greater than two standard errors from the nominal coverage are marked with an asterisk (*), and those that are greater than ten standard errors from the nominal coverage are marked with a dagger (†). It is clear from Tables 3.1 - 3.4 that in all the cases considered, the ordinary percentile-t intervals perform poorly. Even though they have reasonable coverage, often overcovering, they tend to be very long - some of them contain [-1,1] - and the estimated standard error of interval length is large. In comparison, the percentile method intervals, although undercovering rather badly (statistically significantly in all cases), tend to be shorter, with fairly stable length. Except when the underlying population is bivariate lognormal, the transformed percentile-t intervals perform very well, having significantly better coverage accuracy than the percentile method intervals, while being only slightly longer than them, in general.

When the underlying population is bivariate lognormal (see Table 3.4), the transformed percentile-t intervals that use the nonparametric delta method to estimate the variance of Z fail badly. They suffer from severe undercoverage and are still longer than the percentile method intervals. This is probably due to the instability in small samples of the plug-in estimate of the asymptotic variance, (3.1). The transformed percentile-t intervals that employ a jackknife estimate of the variance of Z have better coverage accuracy than the other transformed percentile-t intervals, although they still undercover moderately, having coverage accuracy about the same as the corresponding percentile method intervals. This is a fairly encouraging result in favor of the jackknife since that method is fairly general, making no explicit use of (3.1). None of the four percentile-t methods considered yielded coverage accuracy as high as the iterated percentile method outlined in Section 2.

Overall, the transformed percentile-t interval which used the jackknife estimate of the variance of Z performed consistently best amongst the intervals considered in this section. However, none of the percentile-t intervals discussed performed particularly well when the underlying population was heavy tailed. In that case, if high coverage accuracy is desired, it would be better to use the iterated percentile method.

REFERENCES

Beran, R. (1987a), "Prepivoting to reduce level error in confidence sets", *Biometrika*, **74**, 457-468.

Beran, R. (1987b), "Prepivoting test statistics: a bootstrap view of the Behrens-Fisher problem, the Bartlett adjustment and nonparametric analogues", *Biometrika*, to appear.

Beran, R. (1987c), Discussion of Wu, C.F.J. (1987), "Jackknife, bootstrap and other resampling methods in regression analysis", *The Annals of Statistics*, **14**, 1295-1298.

Bickel, P. and Freedman, D. (1981), "Some asymptotic theory for the bootstrap", *The Annals of Statistics*, **9**, 1196-1217.

Efron, B. (1979), "Bootstrap methods: another look at the jackknife", *The Annals of Statistics*, **7**, 1-26.

Efron, B. (1981), "Nonparametric standard errors and confidence intervals", *Canadian Journal of Statistics*, **9**, 139-172.

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Phil.

Efron, B. and Gong, G. (1983), "A leisurely look at the bootstrap, the jackknife and cross validation", *American Statistician*, **37**, 36-48.

Fisher, R.A. (1921), "On the probable error of a coefficient of correlation deduced from a small sample", *Metron*. **1**, 3-32.

Hall, P. (1986), "On the bootstrap and confidence intervals", *The Annals of Statistics*, **14**, 1431-1452.

Hall, P. (1987), Discussion of Wu, C.F.J. (1987), "Jackknife, bootstrap and other resampling methods in regression analysis", *The Annals of Statistics*, **14**, 1311-1312.

Hall, P. (1988), "Theoretical comparisons of bootstrap confidence intervals", *The Annals of Statistics*, to appear.

Hall, P. and Martin, M.A. (1988), "On bootstrap resampling and iteration", *Biometrika*., to appear.

Kendall, M. and Stuart, A. (1977), *The Advanced Theory of Statistics*. Vol.2, 4th Edition. Charles Griffin and Co. London.

Loh, W.-Y. (1987), "Calibrating confidence coefficients", *Journal of the American Statistical Association*, **82**, 155-162.

Lunneborg, C.E. (1985), "Estimating the correlation coefficient - the bootstrap approach", *Psychological Bulletin*, **98**, 209-215.

Miller, R.G. (1974), "The jackknife - a review", *Biometrika*, **61**, 1-17.

Rasmussen, J. (1987), "Estimating correlation coefficients: bootstrap and parametric approaches", *Psychological Bulletin*, **101**, 136-139.

Schenker, N. (1985), "Qualms about bootstrap confidence intervals", *Journal of the American Statistical Association*, **80**, 360-361.

Singh, K. (1987), Discussion of Wu, C.F.J. (1987), "Jackknife, bootstrap and other resampling methods in regression analysis", *The Annals of Statistics*, **14**, 1328-1330.

Table 2.1. Characteristics of ordinary and iterated percentile method bootstrap confidence intervals for the correlation coefficient. Nominal coverage is 90% throughout. Tabulated characteristics are estimated coverage (nominal SE is .01), average length, standard error of length, and average value of upper endpoint.

n		Simulation 1 ($\rho = 0.0$)		Simulation 2 ($\rho = 0.5$)		Simulation 3 ($\rho = 0.3775$)	
		Ordinary percentile method	Iterated percentile method	Ordinary percentile method	Iterated percentile method	Ordinary percentile method	Iterated percentile method
8	cov.	.847	.908	.837	.905	.840	.903
	len.	1.16	1.41	.97	1.25	1.02	1.29
	st.err.	.29	.37	.37	.46	.36	.44
	upper	.55	.68	.84	.89	.86	.91
10	cov.	.844	.905	.846	.899	.840	.896
	len.	1.02	1.21	.84	1.05	.89	1.10
	st.err.	.24	.31	.31	.40	.28	.36
	upper	.51	.60	.82	.86	.83	.87
12	cov.	.863	.910	.841	.893	.848	.898
	len.	.92	1.07	.74	.89	.82	.99
	st.err.	.20	.28	.26	.34	.24	.30
	upper	.46	.53	.79	.82	.80	.84

Table 3.1. Characteristics of four percentile-t method bootstrap confidence intervals for the correlation coefficient. The parent population is bivariate normal with $\rho = 0$. Nominal coverages are 90%, 95% and 99%. Tabulated characteristics are estimated coverage (nominal SE is .01 for 90% intervals, .007 for 95% intervals, and .003 for 99% intervals), average length, standard error of length and average value of upper endpoint. Estimated coverages that are more than two standard errors away from the nominal coverage are marked with a *, and those that are more than ten standard errors from the nominal coverage are marked with a †.

n		Untransformed percentile-t Delta method			Untransformed percentile-t Jackknife			Ordinary Percentile			Transformed percentile-t Delta method			Transformed percentile-t Jackknife		
	α	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99
10	cov.	.94*	.98*	1.0 *	.94*	.98*	1.0 *	.85*	.90*	.98*	.89	.96	.99	.89	.93*	.99
	len.	1.99	3.25	14.7	2.13	3.49	16.8	1.00	1.20	1.57	1.09	1.33	1.70	1.09	1.31	1.68
	st.err.	1.64	2.78	30.9	1.92	3.27	42.3	.24	.26	.25	.33	.33	.25	.36	.38	.27
	upper	.99	1.73	10.4	1.07	1.87	12.1	.50	.58	.75	.55	.68	.87	.54	.65	.86
15	cov.	.93*	.97*	.99	.94*	.97*	.99	.86*	.93*	.98*	.90	.95	.99	.89	.94*	.99
	len.	1.20	1.63	3.37	1.25	1.72	3.65	.82	.97	1.28	.90	1.07	1.41	.87	1.08	1.44
	st.err.	.50	.78	3.10	.57	.92	3.31	.17	.18	.20	.24	.26	.28	.28	.32	.31
	upper	.59	.84	1.92	.62	.89	2.08	.41	.47	.63	.45	.54	.73	.43	.55	.75
20	cov.	.92*	.97*	1.0 *	.93*	.98*	.99	.88*	.92*	.98*	.90	.95	.99	.89	.94	.98
	len.	.92	1.20	2.12	.95	1.24	2.27	.71	.84	1.12	.76	.92	1.23	.77	.92	1.25
	st.err.	.32	.43	1.01	.36	.48	1.18	.13	.14	.17	.19	.21	.26	.23	.27	.29
	upper	.46	.60	1.18	.48	.62	1.27	.36	.40	.54	.38	.45	.64	.40	.47	.65

Table 3.2. Characteristics of four percentile-t method bootstrap confidence intervals for the correlation coefficient. The parent populations are independent folded Normals. Nominal coverages are 90%, 95% and 99%. Tabulated characteristics are estimated coverage (nominal SE is .01 for 90% intervals, .007 for 95% intervals, and .003 for 99% intervals), average length, standard error of length and average value of upper endpoint. Estimated coverages that are more than two standard errors away from the nominal coverage are marked with a *, and those that are more than ten standard errors from the nominal coverage are marked with a †.

n		Untransformed percentile-t						Ordinary			Transformed percentile-t					
		Delta method			Jackknife			Percentile			Delta method			Jackknife		
	α	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99
10	cov.	.95*	.98*	1.0 *	.95*	.97*	1.0 *	.85*	.92*	.98*	.92	.95	.99	.89	.95	.99
	len.	1.96	3.08	15.7	2.10	3.31	16.5	1.01	1.21	1.56	1.10	1.30	1.70	1.11	1.34	1.71
	st.err.	1.36	2.54	39.1	1.57	2.79	37.0	.23	.25	.25	.32	.32	.26	.36	.36	.27
	upper	.92	1.55	11.5	1.00	1.68	11.9	.53	.61	.74	.54	.65	.88	.54	.68	.89
15	cov.	.92*	.96	.99	.92*	.96	.99	.86*	.90*	.98*	.90	.94	.99	.90	.94	.99
	len.	1.16	1.64	3.39	1.24	1.79	3.61	.81	.97	1.28	.87	1.05	1.43	.90	1.08	1.46
	st.err.	.60	1.00	2.40	.78	1.43	2.74	.17	.19	.21	.25	.28	.27	.26	.30	.29
	upper	.59	.84	1.96	.64	.93	2.09	.42	.46	.61	.44	.51	.74	.44	.55	.76
20	cov.	.93*	.96	.99	.93*	.97*	1.00*	.88	.91*	.98*	.90	.93*	.99	.90	.93*	.99
	len.	.89	1.20	2.19	.93	1.27	2.30	.70	.84	1.12	.74	.91	1.26	.78	.93	1.27
	st.err.	.32	.49	1.04	.40	.61	1.18	.13	.15	.18	.19	.23	.26	.21	.25	.26
	upper	.46	.63	1.23	.48	.67	1.30	.35	.42	.53	.37	.46	.66	.39	.47	.67

Table 3.3. Characteristics of four percentile-t method bootstrap confidence intervals for the correlation coefficient. The samples $\mathcal{X} \equiv \{X_1, \ldots, X_n\}$ and $\mathcal{Y} = \{Y_1, \ldots, Y_n\}$ are drawn so that $X_i = Z_{1i} + W_i$ and $Y_i = Z_{2i} + W_i$ where Z_{1i}, Z_{2i} and W_i are distributed as independent folded normal variates for $i = 1, \ldots, n$ and $\rho = 0.5$. Nominal coverages are 90%, 95% and 99%. Tabulated characteristics are estimated coverage (nominal SE is .01 for 90% intervals, .007 for 95% intervals, and .003 for 99% intervals), average length, standard error of length and average value of upper endpoint. Estimated coverages that are more than two standard errors away from the nominal coverage are marked with a *, and those that are more than ten standard errors from the nominal coverage are marked with a †.

n		Untransformed percentile-t Delta method			Untransformed percentile-t Jackknife			Ordinary Percentile			Transformed percentile-t Delta method			Transformed percentile-t Jackknife		
	α	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99
10	cov.	.91	.98*	.99	.91	.98*	.99	.83*	.92*	.97*	.87*	.95	.99	.90	.94	.99
	len.	1.76	2.81	9.93	1.90	3.05	10.5	.82	1.01	1.41	.99	1.22	1.58	1.01	1.22	1.60
	st.err.	2.44	3.17	18.6	2.42	3.78	19.6	.32	.33	.37	.39	.40	.35	.42	.42	.38
	upper	1.04	1.35	5.35	1.10	1.44	5.61	.82	.86	.92	.80	.86	.95	.80	.86	.95
15	cov.	.92 *	.97*	.99	.92*	.96	.99	.85*	.92*	.97*	.90	.94	.98*	.90	.94	.98*
	len.	.99	1.34	2.65	1.06	1.43	2.87	.66	.79	1.11	.76	.92	1.27	.78	.96	1.30
	st.err.	.57	.75	2.51	.67	.88	3.37	.21	.24	.29	.28	.32	.35	.31	.36	.38
	upper	.85	.96	1.53	.87	1.00	1.64	.77	.80	.86	.76	.81	.90	.77	.81	.91
20	cov.	.91	.96	.99	.91	.96	1.0 *	.86*	.93*	.99	.88	.94	.99	.88	.94	.98
	len.	.77	1.00	1.76	.81	1.06	1.88	.57	.69	.96	.65	.78	1.09	.67	.82	1.12
	st.err.	.33	.44	.90	.38	.51	1.04	.16	.18	.23	.22	.24	.31	.25	.28	.34
	upper	.78	.88	1.16	.80	.90	1.21	.73	.77	.83	.73	.78	.87	.74	.79	.87

Table 3.4. Characteristics of four percentile-t method bootstrap confidence intervals for the correlation coefficient. The parent population is bivariate lognormal with $\rho \approx 0.3775$. Nominal coverages are 90%, 95% and 99%. Tabulated characteristics are estimated coverage (nominal SE is .01 for 90% intervals, .007 for 95% intervals, and .003 for 99% intervals), average length, standard error of length and average value of upper endpoint. Estimated coverages that are more than two standard errors away from the nominal coverage are marked with a *, and those that are more than ten standard errors from the nominal coverage are marked with a †.

n		Untransformed percentile-t Delta method			Untransformed percentile-t Jackknife			Ordinary Percentile			Transformed percentile-t Delta method			Transformed percentile-t Jackknife		
	α	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99	.90	.95	.99
10	cov.	.82*	.91*	.97*	.87*	.94*	.99	.83*	.92*	.98*	.77†	.88*	.96*	.85*	.90*	.97*
	len.	2.24	3.58	12.1	2.82	4.65	15.9	.90	1.08	1.45	.99	1.20	1.56	1.13	1.30	1.65
	st.err.	2.50	4.17	21.1	3.16	5.75	28.8	.28	.30	.33	.40	.40	.36	.44	.43	.36
	upper	.95	1.20	4.16	1.13	1.49	5.33	.83	.89	.94	.72	.78	.87	.77	.80	.90
15	cov.	.80*	.87†	.97*	.87	.93*	.98*	.85*	.92*	.98*	.77†	.84†	.94†	.83*	.90*	.97*
	len.	1.28	1.89	4.10	1.71	2.56	5.70	.75	.89	1.20	.81	.98	1.33	.97	1.16	1.48
	st.err.	.90	1.68	4.39	1.41	2.74	6.47	.21	.22	.24	.32	.36	.39	.43	.42	.39
	upper	.85	.99	1.43	1.02	1.22	1.88	.77	.82	.89	.69	.73	.82	.75	.78	.86
20	cov.	.82*	.88†	.95†	.87*	.92*	.98*	.85*	.93*	.98*	.80†	.86†	.94†	.86*	.90*	.96*
	len.	1.08	1.44	2.50	1.40	1.91	3.50	.66	.78	1.03	.74	.87	1.16	.86	1.03	1.33
	st.err.	.69	1.18	1.86	1.12	1.87	3.15	.18	.18	.20	.30	.34	.36	.38	.42	.40
	upper	.84	.96	1.24	1.00	1.17	1.59	.74	.78	.85	.69	.73	.80	.73	.78	.85