Biometrika (1999), **86,** 1, *pp.* 107–118

Printed in Great Britain

Nonparametric likelihood ratio confidence intervals

By STEPHEN M. S. LEE
Department of Statistics, The University of Hong Kong, Pokfulam Road, Hong Kong
smslee@hkusua.hku.hk

AND G. ALASTAIR YOUNG
Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, U.K.
g.a.young@statslab.cam.ac.uk

Summary

We consider construction of two-sided nonparametric confidence intervals in a smooth function model setting. A nonparametric likelihood approach based on Stein's least favourable family is proposed as an alternative to empirical likelihood. The approach enjoys the same asymptotic properties as empirical likelihood, but is analytically and computationally less cumbersome. The simplicity of the method allows us to propose and analyse asymptotic and bootstrapping techniques as a means of reducing coverage error to levels comparable with those obtained by more computationally-intensive techniques such as the iterated bootstrap. A simulation study confirms that coverage error may be substantially reduced by simple analytic adjustment of the nonparametric likelihood interval and that bootstrapping the distribution of the nonparametric likelihood ratio results in very desirable coverage accuracy.

Some key words: Bootstrap; Coverage; Empirical likelihood; Least favourable family; Nonparametric likelihood.

1. Introduction

Let $X_1, \ldots, X_n$ be independent, identically distributed from an unknown d-variate distribution function F. Let $\mu_F = E_F(X)$. Suppose we wish to construct a two-sided α-level confidence interval for $\theta_F = g(\mu_F)$, where g is some real-valued smooth function defined on $\mathbb{R}^d$. This smooth function model, introduced by Bhattacharya & Ghosh (1978), provides a general framework in which analytical calculations on the behaviour of confidence interval procedures are feasible; see, for example, Hall (1992).

Empirical likelihood, proposed by Owen (1988, 1990), provides a likelihood-based alternative to bootstrap procedures for construction of nonparametric confidence intervals. The empirical likelihood approach dispenses with the need for extensive Monte Carlo simulation, as typically required by bootstrap approaches, requiring instead a numerical optimisation. A two-sided interval constructed using empirical likelihood has an $O(n^{-1})$ coverage error (Hall & La Scala, 1990). DiCiccio & Romano (1989) show that error can be reduced to $O(n^{-2})$ by adjusting the mean and variance of the signed root of the empirical loglikelihood ratio, while DiCiccio, Hall & Romano (1991) show that the empirical likelihood confidence interval can be Bartlett-corrected to achieve similar coverage accuracy.

A description of empirical likelihood and its theoretical properties is provided by Hall & La Scala (1990), who show that bootstrapping the distribution of the empirical likelihood ratio produces effects on coverage accuracy in the case of a two-sided interval resembling those from Bartlett correction.

DiCiccio & Romano (1990) consider several approaches, including empirical likelihood, to construction of nonparametric confidence intervals. Essentially, each approach consists of formulating a least favourable family and considering the associated parametric problem. Different least favourable families yield alternative formulations of nonparametric likelihood, all with similar coverage accuracy. The present paper considers a computationally attractive formulation of nonparametric likelihood based on the least favourable family of Stein (1956).

We show that higher-order correction to the confidence interval can be achieved by various techniques, including asymptotic adjustment to the chi-squared percentiles, bootstrapping the nonparametric likelihood ratio and direct substitution of truncated asymptotic expansions for the confidence interval endpoints. Two-sided confidence intervals constructed using the above techniques have coverage errors of order $O(n^{-2})$. They inherit the advantages of empirical likelihood, such as having a data-driven shape, which are lacking in many refined bootstrap confidence intervals of the same order of coverage error, such as the percentile-t^2 technique described in Hall (1992, p. 110). Moreover, since they do not require nested levels of bootstrap sampling, the procedures we described are computationally much less intensive than iterated bootstrap methods, which offer coverage accuracy of the same order in the two-sided case. Simulation results reported in § 4 show that bootstrapping the nonparametric likelihood ratio results in a very accurate interval which compares favourably with the iterated bootstrap method.

Our focus is on construction of two-sided confidence intervals. The proposed correction techniques will generally not be effective in reducing error of one-sided confidence limits formed from the signed root of the nonparametric loglikelihood ratio.

Section 2 reviews the empirical likelihood approach. Our proposed methods are detailed in § 3. A simulation study is reported in § 4, where empirical comparisons with the bootstrap percentile and iterated bootstrap percentile method intervals are also given. Section 5 provides discussion. Technical details are given in the Appendix.

2. Empirical likelihood

We observe a random sample $\mathscr{X}=(X_1,\ldots,X_n)$ from an unknown distribution F. Assume a multinomial model which places a mass p_i on X_i such that $\sum_{i=1}^n p_i=1$. Then we may define a likelihood of $\mathscr{X}$ to be $\prod_{i=1}^n p_i$. It is easily shown that this likelihood is maximised at $p_i=n^{-1}$ for all i. Define, for any $\theta\in\mathbb{R}$,

$$\mathscr{P}_{n,\theta}=\left\{p=(p_1,\ldots,p_n):p_i\geqslant 0,\ \sum_i p_i=1,\ g\left(\sum_i p_iX_i\right)=\theta\right\}.$$

Then the empirical likelihood $L(\theta)$ is defined to be the profile likelihood, given by

$$L(\theta)=\max_{p\in\mathscr{P}_{n,\theta}}\prod_{i=1}^n p_i.$$

Owen (1998) shows that the empirical loglikelihood ratio, namely

$$-2\log\max_{p\in\mathscr{P}_{n,\theta_F}}\prod_{i=1}^n np_i,$$

has a limiting chi-squared distribution and thus obeys Wilks' theorem. Empirical likelihood may therefore be used to construct nonparametric confidence limits which are asymptotically correct.

Construction of an empirical likelihood confidence interval typically involves complicated numerical computation, necessary to solve a constrained optimisation problem. Algorithms for solving different versions of the optimisation problem have been proposed. Owen (1990) suggests a nested procedure to profile out nuisance parameters in the case $d > 1$, and Hall & La Scala (1990) transformed the problem into one of finding turning points of θ subject to constraints on the p_i. Both approaches require nontrivial and problem-specific formulation of a system of equations. Solution of this system involves numerical computations that become considerably more sophisticated with more complicated parameters.

3. Nonparametric likelihood approaches based on Stein's least favourable family

Denote by $x^{(i)}$ the ith component of the vector $x \in \mathbb{R}^d$. Assume g is continuously differentiable up to some sufficiently high order for the purpose of our analytical calculations. Denote by $g_{r_1 \ldots r_k}$ the partial derivative $\partial^k g / \partial x^{(r_1)} \ldots \partial x^{(r_k)}$. Let also $\hat{\theta} = g(\bar{X})$ and $\hat{g}_{r_1 \ldots r_k} = g_{r_1 \ldots r_k}(\bar{X})$, where $\bar{X} = n^{-1} \sum_{i=1}^n X_i$ denotes the sample mean.

The least favourable family is chosen to pass through $(n^{-1}, \ldots, n^{-1})$ in the n-dimensional simplex $\{(p_1, \ldots, p_n) : p_i \geqslant 0,\ \sum_i p_i = 1\}$ in a direction fixed by $(U_1, \ldots, U_n)$, where $U_i = \sum_{r=1}^d \hat{g}_r (X_i - \bar{X})^{(r)}$. It is least favourable in the sense that the construction yields the same Fisher information for θ_F as does the original multiparameter formulation. The inferential problem is therefore not made artificially easier by reduction to a single-parameter problem.

The multinomial model used in defining empirical likelihood is reduced via Stein's construction to a single-parameter model in which

$$p_i = p_i(\theta) = e^{t_\theta U_i} \Big/ \sum_{j=1}^n e^{t_\theta U_j},$$

where t_θ is chosen to satisfy

$$g\left\{\sum_i p_i(\theta) X_i\right\} = \theta. \tag{1}$$

The corresponding nonparametric loglikelihood ratio function is then

$$R(\theta) = -2 \log \prod_{i=1}^n n p_i(\theta).$$

DiCiccio & Romano (1990) find that $R(\theta_F)$ has also a limiting chi-squared distribution with one degree of freedom. An approximate α-level two-sided confidence interval, asymptotically correct with $O(n^{-1})$ coverage error, can then be constructed as

$$\mathscr{I}_{UL} = \{\theta : R(\theta) \leqslant \hat{c}_\alpha^{UL}\},$$

where $\hat{c}_\alpha^{UL}$ denotes the αth quantile of χ_1^2.

In contrast with the empirical loglikelihood ratio, computation of $R(\theta)$ can be handled in a straightforward manner. The explicit formulae for the p_i enjoy a basic structure, irrespective of the complexity of g. Further, $(U_1, \ldots, U_n)$ do not depend on θ. Calculation of p_i requires only solution of one equation in one unknown variable: this can be done by simple numerical algorithms such as the Newton method.

Our main contribution in this paper is to propose three methods of correcting $\mathscr{I}_{UL}$ in order to reduce the coverage error to $O(n^{-2})$. We denote by ϕ and Φ the standard normal density and distribution functions respectively. The first method adjusts $\hat{c}_\alpha^{UL}$ to

$$\hat{c}_\alpha^{CL} = \{z_\xi + n^{-1}\hat{\eta}_2(z_\xi)\}^2, \tag{2}$$

where $\xi = \frac{1}{2}(1+\alpha)$, $z_\xi = \Phi^{-1}(\xi)$ and $\hat{\eta}_2$ is an odd polynomial to be specified, resulting in a corrected nonparametric likelihood confidence interval $\mathscr{I}_{CL}$. The detailed definition of $\hat{\eta}_2$ is given by equation (A5) in the Appendix. The second method expands the true quantile of $R(\theta_F)$ in an asymptotic series and derives explicit asymptotic expansions for the interval endpoints up to order $O_p(n^{-2})$; see equations (A2) and (A3). An asymptotic nonparametric likelihood confidence interval $\mathscr{I}_{AL}$ is then formed by substituting sample quantities into the asymptotic expansions of the endpoints. Its detailed derivation is given in the Appendix.

We stress again the very straightforward computations necessary for construction of $\mathscr{I}_{CL}$ and $\mathscr{I}_{AL}$. The analytical calculations required for evaluation of the asymptotic expressions involve only partial differentiation of the smooth function g. These calculations can be handled conveniently by exact derivative evaluation routines which require no symbolic computation. The key idea is that within the smooth function model derivatives are required only of functions which are compositions of certain basic functions. They may therefore be computed by repeated use of the chain rule of differentiation and numerical evaluation of the values and derivatives of a suite of such basic functions. Lee & Young (1995) provide details of the way in which such calculations may be packed for automatic and general use.

The third method of correcting $\mathscr{I}_{UL}$ estimates the true quantile of $R(\theta_F)$ from its bootstrap distribution. One level of bootstrap resampling is required by this method. We denote by $\mathscr{I}_{BL}$ the resulting bootstrap nonparametric likelihood confidence interval and by $\hat{c}_\alpha^{BL}$ the αth bootstrap quantile of $R(\theta_F)$.

Note that the endpoints of $\mathscr{I}_{CL}$, $\mathscr{I}_{AL}$ and $\mathscr{I}_{BL}$ are equivalent up to and including the $O_p(n^{-2})$ term in their asymptotic expansions, and therefore have coverages typically differing at order $O(n^{-2})$.

The following propositions state the order of coverage error for the four nonparametric likelihood confidence intervals, $\mathscr{I}_{UL}$, $\mathscr{I}_{CL}$, $\mathscr{I}_{AL}$ and $\mathscr{I}_{BL}$. The basic assumptions here are those made in Hall (1988) for the smooth function model. Briefly, we require that F have moments of sufficiently high order, that g be continuously differentiable up to a sufficiently high order, and that Cramér's condition hold for F.

PROPOSITION 1 (*DiCiccio & Romano*, 1990). *Under Hall's smooth function model,*

$$\mathrm{pr}_F(\theta_F \in \mathscr{I}_{UL}) = \alpha + O(n^{-1}).$$

PROPOSITION 2. *Under Hall's smooth function model,*

$$\mathrm{pr}_F(\theta_F \in \mathscr{I}) = \alpha + O(n^{-2}),$$

for $\mathscr{I} = \mathscr{I}_{CL}$, $\mathscr{I}_{AL}$ *or* $\mathscr{I}_{BL}$.

To prove Proposition 2 it suffices to consider only $\mathscr{I}_{AL}$, for the reasons given earlier. The coverage error of $\mathscr{I}_{AL}$ follows immediately from the derivation given in the Appendix. We can easily see from the proof in the Appendix that the one-sided counterparts of $\mathscr{I}_{UL}$, $\mathscr{I}_{CL}$, $\mathscr{I}_{BL}$ and $\mathscr{I}_{AL}$ constructed using the signed root of $R(\theta)$ all have coverage errors of order $O(n^{-\frac{1}{2}})$. Thus our proposed correction techniques are not effective in reducing coverage error in the one-sided case.

Numerical procedures will usually be necessary for computation of either $R(\theta)$ or the explicit endpoints in construction of $\mathscr{I}_{UL}$, $\mathscr{I}_{CL}$ and $\mathscr{I}_{BL}$. One level of bootstrap resampling is further required by $\mathscr{I}_{BL}$ to approximate bootstrap quantiles. A little algebra shows that explicit endpoints I_L and I_U for these intervals are given by

$$I_L \text{ (or } I_U) = g\left(e^{-\hat{c}_\alpha/2n} \sum_i X_i e^{tU_i} \Big/ n \right), \qquad (3)$$

where $\hat{c}_\alpha$ stands for $\hat{c}_\alpha^{UL}$, $\hat{c}_\alpha^{CL}$ or $\hat{c}_\alpha^{BL}$, and t satisfies

$$\sum_i e^{tU_i} = ne^{\hat{c}_\alpha/2n}. \qquad (4)$$

Note that (4) typically admits two distinct solutions for t, which may be obtained using a Newton method and setting the initial values for t to be $n^{-\frac{1}{2}} z_\xi/\hat{\sigma}$ and $-n^{-\frac{1}{2}} z_\xi/\hat{\sigma}$, where $\hat{\sigma} = (\sum_i U_i^2/n)^{\frac{1}{2}}$. Substitution of the two distinct solutions for t into (3) yields I_L and I_U.

Note that the $p_i(\theta)$ in (1) are parameterised in terms of t_θ, so an alternative to root-finding to calculate the explicit interval endpoints would be the numerically simpler procedure of calculating these p_i for a range of values of t_θ and using spline methods to read off approximate interval end-points. Davison & Hinkley (1997, p. 504) suggest how Poisson regression models might also be used to perform the root-finding.

Construction of $\mathscr{I}_{AL}$ is extremely straightforward, requiring only substitution of sample moments into known formulae.

The arguments of Hall & Martin (1988) show that a two-sided iterated bootstrap confidence interval based on a pivotal statistic has a coverage error of order $O(n^{-2})$. Construction of the iterated interval does, however, require two computationally expensive nested levels of resampling. Proposition 2 shows that $\mathscr{I}_{BL}$ achieves the same order of coverage error as the iterated interval, while requiring only one level of bootstrap resampling. The small price paid for avoiding a level of bootstrap sampling is the more sophisticated numerical calculations needed for computation of $R(\theta)$.

4. Simulation study

4·1. *General framework*

We conducted a simulation study to examine coverage accuracy of the various nonparametric likelihood confidence intervals discussed in this paper. The percentile method interval $\mathscr{I}_P$ (Efron, 1982, § 10.4) and the iterated bootstrap percentile method interval $\mathscr{I}_F$ were included in the study for comparison. The latter, which enjoys theoretical $O(n^{-2})$ coverage error and is seen from previous empirical studies to be exceptionally accurate in practice, provides an appropriate standard.

Two cases were considered, with the parameter of interest θ_F the variance and correlation coefficient respectively. Results were obtained for four different nominal coverage levels, 0·80, 0·90, 0·95 and 0·99. For $\alpha = 0{\cdot}80$, 0·90 and 0·95, the coverage probability was estimated using 1600 Monte Carlo random samples, whereas 5000 were used for $\alpha = 0{\cdot}99$. From each random sample 1000 bootstrap resamples were drawn to construct $\mathscr{I}_{BL}$ and $\mathscr{I}_P$, and 1000 inner level resamples were further drawn from each bootstrap resample to construct $\mathscr{I}_F$. In some cases the numbers of Monte Carlo and bootstrap samples taken to study the coverage of $\mathscr{I}_F$ were reduced to lessen the computational burden.

To estimate the coverages of $\mathscr{I}_{UL}$, $\mathscr{I}_{CL}$ and $\mathscr{I}_{BL}$, we computed directly $R(\theta_F)$ and its approximate αth quantile $\hat{c}_\alpha$, whose formulation depends on the particular interval in question. The coverage probability $\text{pr}_F\{R(\theta_F) \leqslant \hat{c}_\alpha\}$ was then approximated by averaging over the random samples. To compute $R(\theta_F)$ we used the Newton method to solve the equation $g(\sum_i X_i e^{tU_i}/\sum_j e^{tU_j}) = \theta_F$ for t.

The percentile-t^2 method of constructing two-sided confidence intervals is known to have $O(n^{-2})$ coverage error and its construction is computationally simpler than both $\mathscr{I}_{BL}$ and $\mathscr{I}_F$ (Hall, 1992, p. 110). In this simulation study, we also investigated how the percentile-t^2 intervals, constructed using the same bootstrap resamples that produced $\mathscr{I}_{BL}$ and $\mathscr{I}_P$, compare with the other methods for the case $\alpha = 0{\cdot}90$.

4·2. *Variance example*

In this example we considered four different underlying distributions, the standard normal, the folded standard normal, the double exponential and the log-normal distributions. Three sample sizes, $n = 20, 35, 100$, were examined. Construction of $\mathscr{I}_F$ for $n = 100$ is especially computationally demanding. In this case therefore the coverage probability of $\mathscr{I}_F$ was estimated from 1600 Monte Carlo random samples for all four nominal levels, and each interval was constructed using 1000 outer and 100 inner level resamples.

The simulation results for $n = 20$ are summarised in Fig. 1, where coverage error is plotted against the four nominal coverage levels. The results are qualitatively very similar for the other two sample sizes. It is clear that $\mathscr{I}_{BL}$ and $\mathscr{I}_F$ are considerably more accurate than the other intervals, and that $\mathscr{I}_{BL}$ compares very favourably with the computationally much more intensive $\mathscr{I}_F$, especially for higher nominal levels. The percentile interval $\mathscr{I}_P$, which like $\mathscr{I}_{BL}$ requires one level of resampling, generally has the greatest coverage error. The intervals $\mathscr{I}_{CL}$ and $\mathscr{I}_{AL}$ have very similar coverages and both correct $\mathscr{I}_{UL}$ to a certain extent, but not as dramatically as does $\mathscr{I}_{BL}$.

Table 1 compares the coverage probabilities of $\mathscr{I}_{BL}$, $\mathscr{I}_F$ and the percentile-t^2 interval

Table 1: *Variance example. Estimated coverage probabilities (in %) of $\mathscr{I}_{BL}$, $\mathscr{I}_F$ and the percentile-t^2 interval, for* $\alpha = 0{\cdot}90$. (a) *Normal data* $N(0, 1)$, (b) *folded normal data* $|N(0, 1)|$, (c) *double exponential data* $\frac{1}{2}\exp(-|x|)$, (d) *log-normal data* $\exp\{N(0, 1)\}$

Interval	$n = 20$	$n = 35$	$n = 100$	$n = 20$	$n = 35$	$n = 100$
	(a) Normal			(b) Folded normal		
$\mathscr{I}_{BL}$	88·6	87·3	89·2	82·7	84·8	87·0
$\mathscr{I}_F$	89·2	88·6	89·6	84·9	85·9	87·5
Percentile-t^2	86·4	86·8	88·7	83·5	84·6	87·4
	(c) Double exponential			(d) Log-normal		
$\mathscr{I}_{BL}$	82·6	87·1	88·7	67·9	69·9	75·3
$\mathscr{I}_F$	83·9	86·7	89·3	58·0	67·4	74·8
Percentile-t^2	82·5	86·1	88·5	66·5	69·5	72·8

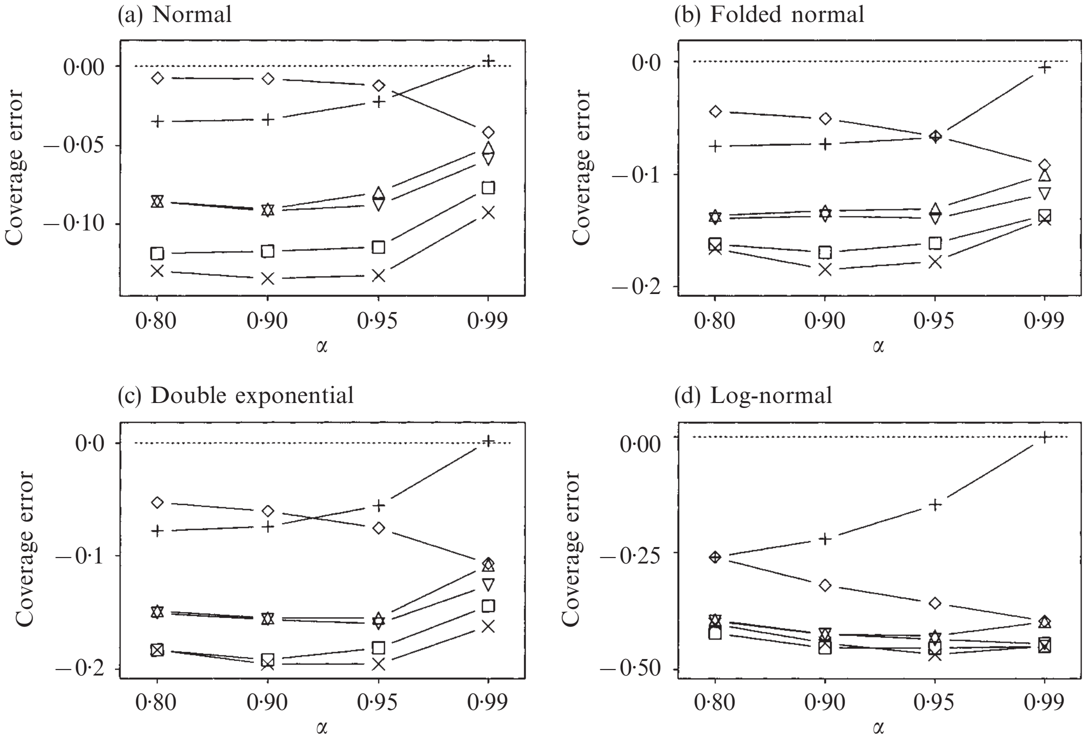

Fig. 1: Variance example. Estimated coverage errors of: □ $\mathscr{I}_{UL}$, ▽ $\mathscr{I}_{CL}$, △ $\mathscr{I}_{AL}$, + $\mathscr{I}_{BL}$, × $\mathscr{I}_{P}$ and ◇ $\mathscr{I}_{F}$, for $n=20$ and $\alpha=0{\cdot}80$, 0·90, 0·95, 0·99. (a) Normal data, (b) folded normal data, (c) double exponential data, (d) log-normal data.

for $\alpha=0{\cdot}90$. We observe that the percentile-t^2 method is generally slightly inferior to $\mathscr{I}_{BL}$ and $\mathscr{I}_{F}$.

4·3. *Correlation coefficient example*

We consider in this example six bivariate distributions, detailed specifications of which can be found in Lee & Young (1995, § 5). Sample sizes considered were $n=15, 20, 30, 50$. The iterated interval $\mathscr{I}_{F}$ was constructed using 1000 outer and 100 inner level bootstrap resamples and its coverage probability was estimated from 1600 Monte Carlo random samples.

Figure 2 displays the coverage errors of the various intervals, for the case $n=15$. Conclusions are quite similar to those in the variance example, except that $\mathscr{I}_{P}$ now displays a much better performance, even in comparison with the theoretically more accurate $\mathscr{I}_{CL}$ and $\mathscr{I}_{AL}$. Such discrepancies do, however, narrow as sample size increases. The intervals $\mathscr{I}_{CL}$ and $\mathscr{I}_{AL}$ still make consistent corrections to $\mathscr{I}_{UL}$, but have an overall performance inferior to that of the resampling-based intervals $\mathscr{I}_{BL}$, $\mathscr{I}_{F}$ and $\mathscr{I}_{P}$.

As in the variance example, we also constructed percentile-t^2 intervals for $\alpha=0{\cdot}90$ and the coverage probabilities are tabulated in Table 2 together with those of $\mathscr{I}_{BL}$ and $\mathscr{I}_{F}$. All three intervals have very similar performance with the percentile-t^2 method being marginally inferior in the log-normal cases.

We note that the likelihood-based intervals are not in general transformation-respecting. That the percentile method interval $\mathscr{I}_{P}$ is transformation-respecting and therefore benefits

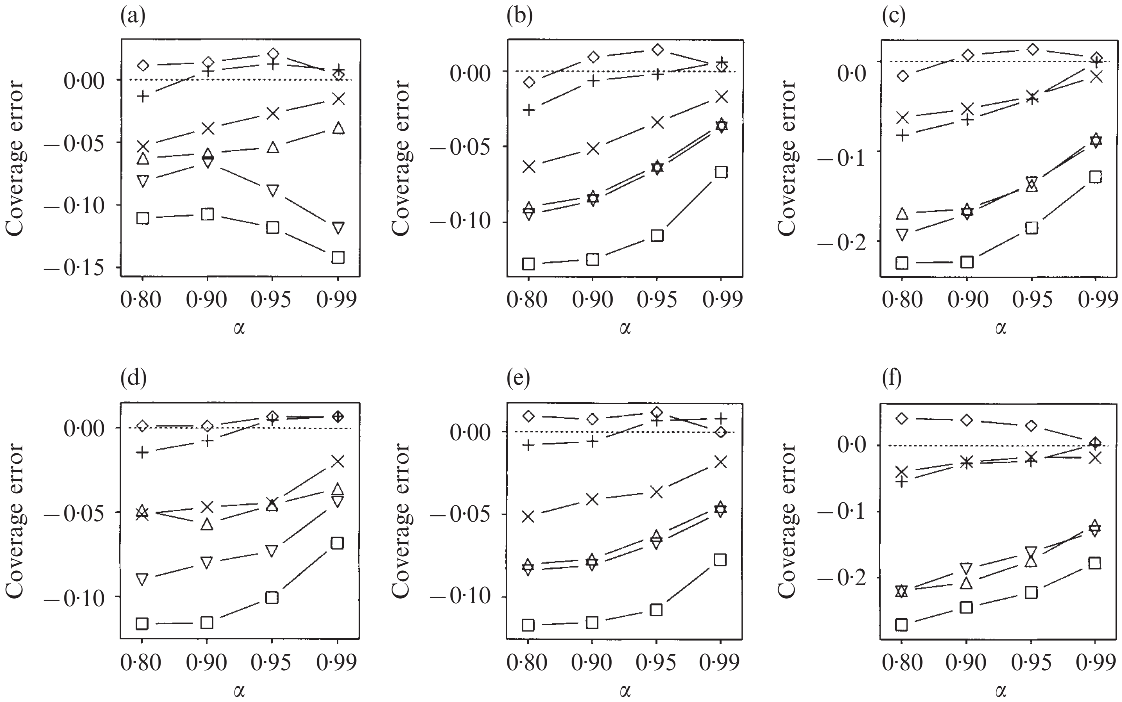

Fig. 2: Correlation coefficient example. Estimated coverage errors of: □ $\mathscr{I}_{UL}$, ▽ $\mathscr{I}_{CL}$, △ $\mathscr{I}_{AL}$, + $\mathscr{I}_{BL}$, × $\mathscr{I}_{P}$ and ◇ $\mathscr{I}_{F}$, for $n = 15$ and $\alpha = 0{\cdot}80, 0{\cdot}90, 0{\cdot}95, 0{\cdot}99$. (a) Normal data, $\theta = 0{\cdot}0$; (b) folded normal data, $\theta = 0{\cdot}0$; (c) log-normal data, $\theta = 0{\cdot}0$; (d) normal data, $\theta = 0{\cdot}5$; (e) folded normal data, $\theta = 0{\cdot}5$; (f) log-normal data, $\theta = 0{\cdot}378$.

Table 2: *Correlation coefficient* (θ) *example. Estimated coverage probabilities* (*in* %) *of* $\mathscr{I}_{BL}$, $\mathscr{I}_{F}$ *and the percentile-*t^2 *interval, for* $\alpha = 0{\cdot}90$. (a) *Folded normal data,* $\theta = 0$; (b) *folded normal data,* $\theta = 0{\cdot}5$; (c) *normal data,* $\theta = 0$; (d) *normal data,* $\theta = 0{\cdot}5$; (e) *log-normal data,* $\theta = 0$; (f) *log-normal data,* $\theta = 0{\cdot}378$

Interval	$n = 15$	$n = 20$	$n = 30$	$n = 50$	$n = 15$	$n = 20$	$n = 30$	$n = 50$
	(a) Folded normal, $\theta = 0$				(b) Folded normal, $\theta = 0{\cdot}5$			
$\mathscr{I}_{BL}$	89·4	88·6	88·6	86·6	89·4	88·8	89·1	87·6
$\mathscr{I}_{F}$	90·9	90·6	89·9	88·8	90·8	90·7	90·8	90·3
Percentile-t^2	88·4	88·1	88·4	86·9	89·2	88·6	89·3	88·6
	(c) Normal, $\theta = 0$				(d) Normal, $\theta = 0{\cdot}5$			
$\mathscr{I}_{BL}$	90·7	90·3	89·9	90·6	89·3	90·5	89·7	89·5
$\mathscr{I}_{F}$	91·4	91·3	91·3	91·9	90·1	91·1	90·9	90·6
Percentile-t^2	89·7	89·9	89·1	89·9	88·3	90·1	89·0	89·9
	(e) Log-normal, $\theta = 0$				(f) Log-normal, $\theta = 0{\cdot}378$			
$\mathscr{I}_{BL}$	83·5	83·9	82·8	85·7	87·3	86·3	86·3	85·5
$\mathscr{I}_{F}$	90·8	90·0	88·6	88·4	93·8	92·2	92·6	91·7
Percentile-t^2	82·7	82·6	82·0	83·8	85·5	84·6	85·4	84·1

automatically from variance stabilisation may account for its high accuracy. We repeated the simulation exercise for $\alpha = 0{\cdot}90$ to study the effects of Fisher's transformation on the likelihood-based intervals, but found very little change in their estimated coverages.

5. Discussion

As is evident from its theoretical properties and empirical performance, the bootstrap nonparametric likelihood interval $\mathscr{I}_{BL}$ stands as a strong candidate among bootstrap intervals which have accurate coverage. A similar or even marginally better coverage accuracy compared to the iterated bootstrap interval can be achieved using only one level of resampling.

The nonparametric likelihood method requires numerical computation of $R(\theta)$ and the endpoints of the interval, which typically involves numerical algorithms such as the Newton method. To investigate how nonconvergence of the Newton method might affect the proper functioning of the nonparametric likelihood procedures, we recorded the frequencies of non-convergence suffered by the likelihood-based intervals in the simulation study discussed above. We found that the nonconvergence frequencies did not exceed 4·7% and 0·72% in the variance and correlation coefficient examples respectively, and that iteration failures arose much less frequently for larger sample sizes. This suggests that failures were mainly caused by nonexistence of feasible solutions to the equation (1), as might occur if no weighting of sample data could result in the desired θ. The restriction from a multinomial model to a single-parameter family limits the scope of the feasible p_i's, and may therefore result in more frequent iteration failures. It would be of interest to compare $\mathscr{I}_{BL}$ with the empirical likelihood interval in this respect.

The intervals $\mathscr{I}_{CL}$ and $\mathscr{I}_{AL}$ are found to correct the naive nonparametric likelihood interval $\mathscr{I}_{UL}$ significantly. However, they do not appear to approximate $\mathscr{I}_{BL}$ particularly accurately, although they have the same order of coverage error. This is in contrast to the findings of Lee & Young (1995), who observe that asymptotic approximations to the iterated bootstrap interval $\mathscr{I}_{F}$, constructed by the same techniques as are the intervals $\mathscr{I}_{CL}$ and $\mathscr{I}_{AL}$, and requiring less or even no resampling, so yielding very significant computational gain, provide coverage errors which are not significantly greater than those of $\mathscr{I}_{F}$. In the context of the current paper, we judge that the arguments for use of $\mathscr{I}_{AL}$ and $\mathscr{I}_{CL}$ are not compelling: the interval $\mathscr{I}_{BL}$ requires just one level of bootstrap sampling, but yields noticeable improvements in terms of coverage accuracy.

It is evident that information on the orders of coverage error alone is inadequate to discriminate effectively between alternative procedures. Lee & Young (1997) observe that explicit evaluation of the leading term in expansion of coverage error may be useful in discriminating qualitatively between different methods of confidence interval construction, although it may not be as effective in predicting the actual coverage error. We illustrate this point by evaluating explicitly the leading terms in expansions of theoretical coverage errors in the variance example of § 4. The explicit formulae are obtained by expanding the calculations shown in the Appendix to include the next higher-order terms in all expansions, and then using the techniques of Lee & Young (1997). Table 3 lists the figures for intervals which have $O(n^{-2})$ coverage errors, based on a nominal level of 0·90. It is found, contrary to our empirical findings, that the percentile-t^2 interval has the smallest leading term in asymptotic coverage error expansion, and that $\mathscr{I}_{BL}$ is asymptotically less accurate than $\mathscr{I}_{F}$. The two approximate likelihood intervals $\mathscr{I}_{AL}$ and $\mathscr{I}_{CL}$ are the least

Table 3: *Variance example. Leading terms in asymptotic expansions of coverage errors of $\mathscr{I}_{BL}$, $\mathscr{I}_{AL}$, $\mathscr{I}_{CL}$, $\mathscr{I}_F$ and the percentile-t^2 interval, for* $\alpha = 0{\cdot}90$

Interval	Standard normal ($\times n^{-2}$)	Folded normal ($\times 10^3 n^{-2}$)	Double exponential ($\times 10^4 n^{-2}$)	Log-normal ($\times 10^{20} n^{-2}$)
$\mathscr{I}_{BL}$	−177·6	−1·88	−2·15	−7·09
$\mathscr{I}_{AL}$	−347·3	−2·85	−2·22	−5·49
$\mathscr{I}_{CL}$	−340·7	−2·77	−2·14	−5·28
$\mathscr{I}_F$	−149·9	−1·37	−1·24	−2·49
Percentile-t^2	−36·6	−0·19	0·08	2·07

accurate, except for the double exponential and log-normal cases where they are comparable to $\mathscr{I}_{BL}$.

Acknowledgement

S. M. S. Lee was supported by a grant funded by the Committee on Research and Conference Grants, the University of Hong Kong.

Appendix

Derivations

This appendix serves two purposes. First, it provides explicit definitions of the intervals $\mathscr{I}_{AL}$ and $\mathscr{I}_{CL}$. It also derives asymptotic expansions for the endpoints of the nonparametric likelihood intervals. The proof of Proposition 2 then follows by a straightforward application of Proposition 3.1 of Hall (1992).

The definition of $\mathscr{I}_{CL}$ is obtained from an asymptotic expansion of an exact α-level confidence interval obtained from the nonparametric likelihood ratio R, which we now derive; see equations (A2), (A3) and (A5).

First define $\mu_{i_1\ldots i_r} = E_F\{(X-\mu)^{(i_1)}\ldots(X-\mu)^{(i_r)}\}$ and let $\hat{\mu}_{i_1\ldots i_r}$ be its sample version. Denote by c_α the exact αth quantile of $R(\theta_F)$. The confidence limits of an exact α-level interval can then be obtained by solving $R(\theta) = c_\alpha$ for θ. Under the smooth function model the confidence limits are typically within $O_p(n^{-\frac{1}{2}})$ of the estimate $\hat{\theta}$. It thus suffices to assume $\theta = \hat{\theta} + O_p(n^{-\frac{1}{2}})$.

Define

$$\hat{\sigma}^2 = \sum_{i=1}^n U_i^2/n, \quad \hat{\gamma} = \sum_{i=1}^n U_i^3/n, \quad \hat{\delta} = \sum_{i=1}^n U_i^4/n, \quad \hat{\varepsilon} = \sum_{i=1}^n U_i^5/n,$$

$$\hat{G}_1 = \sum_{r_1\ldots r_4=1}^d \hat{g}_{r_1}\hat{g}_{r_2}\hat{g}_{r_3r_4}\hat{\mu}_{r_1r_3}\hat{\mu}_{r_2r_4}, \quad \hat{G}_2 = \sum_{r_1\ldots r_5=1}^d \hat{g}_{r_1}\hat{g}_{r_2}\hat{g}_{r_3}\hat{g}_{r_4r_5}\hat{\mu}_{r_1r_4}\hat{\mu}_{r_2r_3r_5},$$

$$\hat{G}_3 = \sum_{r_1\ldots r_6=1}^d \hat{g}_{r_1}\hat{g}_{r_2}\hat{g}_{r_3}\hat{g}_{r_4}\hat{g}_{r_5r_6}\hat{\mu}_{r_1r_5}\hat{\mu}_{r_2r_3r_4r_6}, \quad \hat{G}_4 = \sum_{r_1\ldots r_6=1}^d \hat{g}_{r_1}\hat{g}_{r_2}\hat{g}_{r_3}\hat{g}_{r_4}\hat{g}_{r_5r_6}\hat{\mu}_{r_1r_2r_5}\hat{\mu}_{r_3r_4r_6},$$

$$\hat{G}_5 = \sum_{r_1\ldots r_6=1}^d \hat{g}_{r_1}\hat{g}_{r_2}\hat{g}_{r_3}\hat{g}_{r_4r_5r_6}\hat{\mu}_{r_1r_4}\hat{\mu}_{r_2r_5}\hat{\mu}_{r_3r_6}, \quad \hat{G}_6 = \sum_{r_1\ldots r_7=1}^d \hat{g}_{r_1}\hat{g}_{r_2}\hat{g}_{r_3}\hat{g}_{r_4}\hat{g}_{r_5r_6r_7}\hat{\mu}_{r_1r_5}\hat{\mu}_{r_2r_6}\hat{\mu}_{r_3r_4r_7},$$

$$\hat{G}_7 = \sum_{r_1\ldots r_8=1}^d \hat{g}_{r_1}\hat{g}_{r_2}\hat{g}_{r_3}\hat{g}_{r_4}\hat{g}_{r_5r_6r_7r_8}\hat{\mu}_{r_1r_5}\hat{\mu}_{r_2r_6}\hat{\mu}_{r_3r_7}\hat{\mu}_{r_4r_8},$$

$$\hat{C}_3 = -(\tfrac{2}{3}\hat{\gamma} + \hat{G}_1), \quad \hat{C}_4 = \tfrac{1}{4}\hat{\sigma}^{-2}(\hat{\gamma} + \hat{G}_1)(3\hat{\gamma} + 5\hat{G}_1) + \tfrac{3}{4}\hat{\sigma}^4 - \tfrac{1}{4}\hat{\delta} - \hat{G}_2 - \tfrac{1}{3}\hat{G}_5,$$

$$\hat{C}_5 = \tfrac{1}{6}\hat{\sigma}^{-2}\{(5\hat{\gamma} + 6\hat{G}_1)(3\hat{G}_2 + \hat{G}_5) + \hat{\delta}(4\hat{\gamma} + 5\hat{G}_1)\} - \tfrac{1}{4}\hat{\sigma}^{-4}(\hat{\gamma} + \hat{G}_1)^2(4\hat{\gamma} + 7\hat{G}_1)$$
$$- \tfrac{1}{6}\hat{\sigma}^2(8\hat{\gamma} + 9\hat{G}_1) - \tfrac{1}{15}\hat{\varepsilon} - \tfrac{1}{3}\hat{G}_3 - \tfrac{1}{4}\hat{G}_4 - \tfrac{1}{2}\hat{G}_6 - \tfrac{1}{12}\hat{G}_7.$$

By expanding t_θ in a power series in $n^{-\frac{1}{2}}T_\theta = (\hat{\theta} - \theta)/\hat{\sigma}$, we obtain

$$R(\theta) = T_\theta^2 - n^{-1/2}\hat{C}_3 T_\theta^3/\hat{\sigma}^3 + n^{-1}\hat{C}_4 T_\theta^4/\hat{\sigma}^4 - n^{-3/2}\hat{C}_5 T_\theta^5/\hat{\sigma}^5 + O_p(n^{-2}). \quad \text{(A1)}$$

Note that $R(\theta_F)$ is asymptotically chi-squared distributed and so the equation $R(\theta) = c_\alpha$ typically admits two distinct solutions for sufficiently large n. Denote the two solutions by $\tilde{\theta}_L$ and $\tilde{\theta}_U$ such that $\tilde{\theta}_L \leqslant \tilde{\theta}_U$. Then an exact α-level confidence interval for θ is $[\tilde{\theta}_L, \tilde{\theta}_U]$.

Recall that $\xi = (1+\alpha)/2$ and $z_\xi = \Phi^{-1}(\xi)$. Substituting (A1) and writing

$$c_\alpha = \left\{ \sum_{j \geqslant 0} n^{-j/2} \eta_j(z_\xi) \right\}^2$$

for some polynomials η_j, even for odd j and odd for even j, we obtain

$$\tilde{\theta}_U = \hat{\theta} + n^{-1/2}\hat{\sigma}\{\hat{f}_0(z_\xi) + n^{-1/2}\hat{f}_1(z_\xi) + n^{-1}\hat{f}_2(z_\xi) + n^{-3/2}\hat{f}_3(z_\xi) + O_p(n^{-2})\}. \quad \text{(A2)}$$

By symmetry we can put $\eta_1 = \eta_3 \equiv 0$ and get

$$\tilde{\theta}_L = \hat{\theta} + n^{-1/2}\hat{\sigma}\{\hat{f}_0(z_{1-\xi}) + n^{-1/2}\hat{f}_1(z_{1-\xi}) + n^{-1}\hat{f}_2(z_{1-\xi}) + n^{-3/2}\hat{f}_3(z_{1-\xi}) + O_p(n^{-2})\}. \quad \text{(A3)}$$

Here the $\hat{f}_j$ are polynomials, even for odd j and odd for even j, given by

$$\begin{gathered} \hat{f}_0 = \eta_0, \quad \hat{f}_1 = -\tfrac{1}{2}\hat{C}_3\eta_0^2/\hat{\sigma}^3, \quad \hat{f}_2 = \eta_2 + \tfrac{1}{8}(5\hat{C}_3^2 - 4\hat{C}_4\hat{\sigma}^2)\eta_0^3/\hat{\sigma}^6, \\ \hat{f}_3 = -\hat{C}_3\eta_0\eta_2/\hat{\sigma}^3 - \tfrac{1}{2}(2\hat{C}_3^3 - 3\hat{C}_3\hat{C}_4\hat{\sigma}^2 + \hat{C}_5\hat{\sigma}^4)\eta_0^4/\hat{\sigma}^9. \end{gathered} \quad \text{(A4)}$$

If we apply Proposition 3.1 in Hall (1992) and take $\eta_0(z_\xi) = z_\xi$, an asymptotic expansion for the coverage of $[\tilde{\theta}_L, \tilde{\theta}_U]$ is obtained to be $\alpha + n^{-1}\phi(z_\xi)\tilde{f}(z_\xi) + O(n^{-2})$, for some polynomial $\tilde{f}$. Exactness of the interval implies that $\tilde{f} \equiv 0$ and so

$$\eta_2(z_\xi) = -q_2(z_\xi) + \tfrac{1}{8}C_3^2 z_\xi^5/\sigma^6 + \tfrac{1}{2}C_3 z_\xi^2\{z_\xi q_1(z_\xi) - q_1'(z_\xi)\}/\sigma^3 - \tfrac{1}{8}(5C_3^2 - 4C_4\sigma^2)z_\xi^3/\sigma^6 + \tau z_\xi^3, \quad \text{(A5)}$$

where q_1 and q_2 are polynomials defined by

$$\text{pr}_F(T_{\theta_F} \leqslant x) = \Phi(x) + n^{-\frac{1}{2}}q_1(x)\phi(x) + n^{-1}q_2(x)\phi(x) + o(n^{-1}),$$

$\tau = E_F\{n^{\frac{1}{2}}T_{\theta_F}(C_3/\sigma^3 - \hat{C}_3/\hat{\sigma}^3)\}/2 + O(n^{-1})$, and C_j, σ denote the population versions of $\hat{C}_j, \hat{\sigma}$ respectively. The sample version of η_2, denoted by $\hat{\eta}_2$, gives the adjustment (2) for constructing $\mathcal{I}_{CL}$. Similarly, asymptotic expansions for the confidence limits of $\mathcal{I}_{BL}$ are given by expansions (A2) and (A3), with the η_j replaced by their sample versions $\hat{\eta}_j$ in (A4). The resulting expansions, when truncated at the $O_p(n^{-2})$ terms, define the endpoints of $\mathcal{I}_{AL}$.

An application of Proposition 3.1 in Hall (1992) shows that $\mathcal{I}_{AL}$, $\mathcal{I}_{CL}$ and $\mathcal{I}_{BL}$ all have coverage error of order $O_p(n^{-2})$.

References

Bhattacharya, R. N. & Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6**, 435–51.

Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap Methods and their Application.* Cambridge: Cambridge University Press.

DiCiccio, T. J., Hall, P. & Romano, J. P. (1991). Empirical likelihood is Bartlett-correctable. *Ann. Statist.* **19**, 1053–61.

DiCiccio, T. J. & Romano, J. P. (1989). On adjustments based on the signed root of the empirical likelihood ratio statistic. *Biometrika* **76**, 447–56.

DiCiccio, T. J. & Romano, J. P. (1990). Nonparametric confidence limits by resampling methods and least favourable families. *Int. Statist. Rev.* **58**, 59–76.

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

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals (with Discussion). *Ann. Statist.* **16**, 927–85.

HALL, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer.
HALL, P. & LA SCALA, B. (1990). Methodology and algorithms of empirical likelihood. *Int. Statist. Rev.* **58**, 109–27.
HALL, P. & MARTIN, M. A. (1988). On bootstrap resampling and iteration. *Biometrika* **75**, 661–71.
LEE, S. M. S. & YOUNG, G. A. (1995). Asymptotic iterated bootstrap confidence intervals. *Ann. Statist.* **23**, 1301–30.
LEE, S. M. S. & YOUNG, G. A. (1997). Asymptotics and resampling methods. *Comp. Sci. Statist.* **28**, 221–7.
OWEN, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–49.
OWEN, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90–120.
STEIN, C. (1956). Efficient nonparametric testing and estimation. In *Proc. 3rd Berkeley Symp. Math. Statist. Prob.*, **1**, Ed. J. Neyman, pp. 187–95. Berkeley: University of California Press.

[*Received October* 1995. *Revised May* 1998]