On the Relative Accuracy of Certain Bootstrap Procedures

Andrey Feuerverger
University of Toronto, Canada

John Robinson
University of Sydney, Australia

Augustine Wong
York University, Toronto, Canada

June 29, 1998

Running title: Relative Bootstrap Accuracy

Abstract

We show the second order relative accuracy, on bounded sets, of the studentized bootstrap, exponentially tilted bootstrap and nonparametric likelihood tilted bootstrap, for means and smooth functions of means. We also consider the relative errors for larger deviations. Our method exploits certain connections between Edgeworth and saddlepoint approximations to simplify the computations.

Résumé

Les auteurs étudient la précision relative du deuxième ordre, sur des ensembles bornés, de variantes studentisée, inclinée exponentiellement et inclinée par vraisemblance non paramétrique de la méthode d'auto-amorçage pour la moyenne et les fonctions lisses d'icelle. Ils examinent en outre le comportement de l'erreur relative pour de grandes déviations. Leurs calculs sont facilités par certaines relations existant entre l'approximation de Edgeworth et la méthode du point de selle.

Key words and phrases: Bootstrap approximation, Edgeworth expansions, Empirical likelihood; Exponential tilting; Saddlepoint approximation; Studentizing.

AMS Subject Classifications: 62G10, 62G09, 62G20.

1 INTRODUCTION AND SUMMARY

Let $X_1, \ldots, X_n$ be identically and independently distributed with distribution function F, and set $EX_1 = \mu = \mu(F)$ and $Var(X_1) = \sigma^2$. Suppose we are interested in testing the hypothesis $H_0 : \mu = \mu_0$. The exact p-value for the usual t-test can be taken to be

$$p = P\left(\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq \frac{\bar{x} - \mu_0}{s/\sqrt{n}}\middle| \mu(F) = \mu_0\right) \tag{1.1}$$

where $\bar{x} = \sum x_j/n$ is the observed value of the sample mean $\bar{X} = \sum X_j/n$, and $s^2 = \sum(x_j - \bar{x})^2/n$ is the observed value of the sample variance $S^2 = \sum(X_j - \bar{X})^2/n$. Since F is unknown, this probability cannot be determined, but the bootstrap may be used to approximate it. To do this, $X_1^*, \ldots, X_n^*$ are randomly sampled from the empirical distribution function F_n. Note that the mean and variance of observations from this distribution are $E^*X_1^* = \bar{x}$ and $Var^*X_1^* = s^2$. The Studentized bootstrap p-value is then defined as

$$p_s^* = P\left(\frac{\bar{X}^* - \bar{x}}{S^*/\sqrt{n}} \geq \frac{\bar{x} - \mu_0}{s/\sqrt{n}}\middle| F_n\right) \tag{1.2}$$

where $\bar{X}^* = \sum X_j^*/n$ and $S^{*2} = \sum(X_j^* - \bar{X}^*)^2/n$ are the mean and variance of the bootstrap sample. (Note that the convenient choice of divisor n instead of $n-1$ in our definitions of sample variance has no effect since this divisor appears on both sides of the inequalities in (1.1) and (1.2).)

The question considered here is: how close is p_s^* to p? Hall (1988) shows that $p_s^* - p = O_P(1/n)$, and Jing, Feuerverger and Robinson (1994) show that $p_s^*/p = 1 + O_P\left[\{(\sqrt{n}a)^3 \vee 1\}/n\right]$, where $a = (\bar{x} - \mu_0)/s$. The importance of this second result is apparent by noting that, particularly for small sample size, the absolute error may not be small compared with the p-value when the p-value is small. The relative error makes plain that even small p-values are well approximated by the bootstrap methods. The inclusion of the term $\sqrt{n}a$, which is about 2 when the p-value is about .025, shows that the accuracy cannot be sustained to large deviations and that the result is only useful for $\sqrt{n}a$ of smaller order than $n^{1/3}$. We do not consider coverage errors as in Hall (1988) since there are a number of technical difficulties involved in using saddlepoint approximations for these which are not yet resolved. For simplicity we keep our focus, in much of this Introduction, on the simplest problem as described above in terms of a univariate mean. However we generalize this below to univariate statistics that can be written as smooth functions of multivariate means.

The tilted bootstrap involves sampling $X^*_{\beta 1}, \ldots, X^*_{\beta n}$ from $x_1, \ldots, x_n$ with probabilities

$$q_{\beta j} = \frac{e^{\beta x_j}}{\sum e^{\beta x_j}},$$

for $j = 1, \ldots, n$; that is we are sampling from $F_{\beta n}$ where

$$dF_{\beta n}(x) = e^{-\{K_n(\beta) - \beta x\}} dF_n(x),$$

and where

$$K_n(\beta) = \log \frac{1}{n} \sum e^{\beta x_j}.$$

For testing $H_0 : \mu = \mu_0$, we choose $\beta \equiv \beta(\mu_0)$ to be such that $EX^*_{\beta 1} = \mu_0$; that is

$$\mu_0 = K'_n(\beta) = \frac{\sum x_j e^{\beta x_j}}{\sum e^{\beta x_j}},$$

and then define the tilted bootstrap p-value to be

$$p^*_t = P(\bar{X}^*_\beta \geq \bar{x} | F_{\beta n})$$

where $\bar{X}^*_\beta$ is the mean of the $X^*_{\beta j}$'s. This is intended to be an approximate p-value for our hypothesis test. Here $F_{\beta n}$ is considered to be the best that we can do to approximate the hypothesized distribution. This may then be compared to the p-value for the Studentized bootstrap. DiCiccio and Romano (1990) essentially show that $p^*_t - p^*_s = O_P(1/n)$. We show that $p^*_t / p^*_s = 1 + O_P[\{(\sqrt{n}a)^3 \vee 1\}/n]$. We can obtain the same result for the case where the bootstrap sampling has probabilities in common with those of empirical likelihood.

In Section 2, we consider more general statistics that can be written as smooth functions of means and which include those above as a special case. We show that the relative error of the bootstrap of the Studentized versions of such statistics can be obtained by exploiting the relationship of the saddlepoint approximations to the Edgeworth expansions and using the results of Hall (1988). Then we consider the exponentially tilted bootstrap for such general statistics and show that their relative error is of the same form as that obtained in the case of the mean by again exploiting the results on Edgeworth expansions obtained by DiCiccio and Romano (1990). We also consider the two bootstrap sampling schemes from the families $\mathbf{F}_2$ and $\mathbf{F}_3$ of these authors and show that the same relative accuracy results hold for these methods.

In Section 3 we consider a form of tilting which is slightly different from those in DiCiccio and Romano (1990); this is derived by maximizing a non-parametric likelihood derived from an exponential family supported on the observed values, subject to a constraint given by defining the function of the mean as the parameter of interest, and so we have called it tilting to a profile likelihood.

Section 4 gives the results of a small numerical study comparing the various bootstrap and saddlepoint approximations discussed here. In general we find the saddlepoint approximations to be somewhat superior to the bootstrap approximations; they also required less computing time by a large factor. Further, although the tilted and studentized means have asymptotically the same order of error, the studentized mean gives consistently better results in the tail.

Finally, Section 5 contains some technical details giving an indication of the proof needed to obtain the Barndorff-Neilsen saddlepoint approximation in the general case for the exponentially tilted bootstrap.

2 RELATIVE ERROR FOR SMOOTH FUNCTIONS OF MEANS

Let $X_1, \ldots, X_n$ be a sample of random vectors in $\mathbb{R}^k$ from a distribution F in a family $\mathcal{F}$. Frequently, more general parameters $\theta(F)$ may be written as smooth functions $g(\mu)$ of the vector of expectations $\mu = EX_1$, and estimates $\theta(F_n)$ may be written as $g(\bar{X})$. A Studentized version of this (in the notation of Hall 1992) is given by

$$A(\bar{X}) = \frac{g(\bar{X}) - g(\mu)}{h(\bar{X})}$$

where

$$h(\bar{X}) = \sqrt{g'(\bar{X})^T \ \hat{\Sigma} \ g'(\bar{X})}$$

for $\hat{\Sigma} = n^{-1} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^T$, and we take $g'(\bar{X})$ as a vector of derivatives in $\mathbb{R}^k$. Note that we can assume the sums of squares and products arising in $\hat{\Sigma}$ to be elements of the original vector of means $\bar{X}$. We wish to test $H_0 : g(\mu) = \theta$.

2.1 STUDENTIZED STATISTICS

The Edgeworth expansion arguments of Hall (1988) show that if $a = A(\bar{x})$,

$$p = P\left(A(\bar{X}) \geq a\right) = 1 - \Phi(\sqrt{n}a) - \frac{1}{\sqrt{n}} q_1(\sqrt{n}a)\varphi(\sqrt{n}a) + O(1/n) \qquad (2.1)$$

where q_1 is a polynomial of degree 2 with coefficients depending on the cumulants of X_1 and on the derivatives of A. Further

$$p^* = P\left\{A(\bar{X}^*) \geq a\right\} = 1 - \Phi(\sqrt{n}a) - \frac{1}{\sqrt{n}} \hat{q}_1(\sqrt{n}a)\varphi(\sqrt{n}a) + O_P(1/n), \qquad (2.2)$$

where

$$A(\bar{X}^*) = \frac{g(\bar{X}^*) - g(\bar{x})}{h(\bar{X}^*)}$$

and $\hat{q}_1$ has coefficients exactly as for q_1 except that the cumulants are of X_1^* and so differ from those of X_1 by $O_P(n^{-1/2})$. It therefore follows that $p_s^* - p = O_P(1/n)$. It is worthwhile noticing that the bounding constants involved in the order term in (2.1) depend on the fourth moments of the distribution F and we note that even if a is random the order term is fixed. The equivalent bounds in (2.2) depend on the sample moments. The set on which these are bounded and on which hold other regularity conditions considered in Jing and Robinson (1994) has probability tending to 1. Thus we use the O_P notation in this case.

We now use these results to establish a relative error for the general studentized bootstrap. The Barndorff-Nielsen (1986) form of the tail area approximation for p follows from Theorem 3 of Jing and Robinson (1994); under conditions stated there

$$p = \left[1 - \Phi\{\sqrt{n}w^*(a)\}\right]\{1 + O(1/n)\}, \tag{2.3}$$

where

$$w^*(a) = \hat{w}(a) - \frac{\log \psi\{\hat{w}(a)\}}{n\hat{w}(a)}, \tag{2.4}$$

and $\hat{w}(a)$, $\psi\{\hat{w}(a)\}$ are as defined in Theorem 3 of Jing and Robinson (1994). Now $\hat{w}(a)$ and $\psi\{\hat{w}(a)\}$ are analytic functions of a in a neighbourhood of the origin. Therefore we can write

$$\hat{w}(a) = A_0 + A_1 a + A_2 a^2 + A_3 a^3 + A_4 a^4 + O(a^5) \tag{2.5}$$

and

$$\psi\{\hat{w}(a)\} = B_0 + B_1 a + B_2 a^2 + B_3 a^3 + B_4 a^4 + O(a^5), \tag{2.6}$$

where the coefficients A_j and B_j do not depend on n but are determined as nice functions of g, of the underlying cumulant generating function, and of their derivatives. In Feuerverger, Robinson and Wong (1998) we establish that $A_0 = 0$, $A_1 = 1$ and $B_0 = 1$. In the same way

$$p_s^* = [1 - \Phi\{\sqrt{n}w_s^*(a)\}]\{1 + O_P(1/n)\}$$

where

$$w_s^*(a) = \hat{w}_s(a) - \frac{\log \psi_s\{\hat{w}_s(a)\}}{n\hat{w}_s(a)}$$

and where all these terms differ from the corresponding ones in (2.3) and (2.4) only by the use here of the empirical cumulant generating function

$$K_n(t_1, t_2) = \log E \exp\left(t_1 \bar{X}^* + t_2 \overline{X^{*2}}\right)$$

instead of the population cumulant generating function. Thus the corresponding coefficients for the series expansions (2.5) and (2.6) depend on sample rather than population cumulants and so differ by quantities of $O_P(n^{-1/2})$.

Note that $\varphi(x)/\{1-\Phi(x)\} < 1+x,\ x > 0$, which implies that for u, v positive,

$$\frac{|\Phi(v) - \Phi(u)|}{1-\Phi(u)} \leq (1+u)|u-v|e^{u|u-v|}.$$

Consequently, for $\sqrt{n}a = o(n^{1/3})$, use of the Mean Value Theorem with some value of u between $w^*(a)$ and $w_s^*(a)$ gives

$$\frac{p_s^*}{p} = \frac{1-\Phi\{\sqrt{n}w_s^*(a)\}}{1-\Phi\{\sqrt{n}w^*(a)\}}\{1+O_P(1/n)\} = 1 + O_P\left\{\frac{(\sqrt{n}a)^3 \vee 1}{n}\right\}.$$

2.2 BOOTSTRAP TILTING

The general exponentially tilted bootstrap is based on sampling from the original sample $x_1, \ldots, x_n$, with probabilities

$$q_{\beta j} = \frac{e^{\beta \cdot x_j}}{\sum_{j=1}^n e^{\beta \cdot x_j}},$$

where $\beta \equiv \beta(\theta)$ is chosen so that if $K_n(\beta) = \log \frac{1}{n} \sum_{j=1}^n e^{\beta \cdot x_j}$, then

$$K_n'\{\beta(\theta)\} = \mu(\theta),$$

and where $\mu(\theta)$ and $\lambda(\theta)$ are obtained from the equations

$$g\{\mu(\theta)\} = \theta \tag{2.7}$$

and

$$\beta(\theta) = \lambda(\theta) g'\{\mu(\theta)\}. \tag{2.8}$$

These weights are obtained by minimizing the forward Kullback-Leibler distance between the true distribution and the tilted empirical distribution subject to (2.7) using λ as a Lagrange multiplier as in the treatment of the $\mathbf{F}_2$ family by DiCiccio and Romano (1990). Let $X^*_{\theta 1}, \ldots, X^*_{\theta n}$ be a random sample obtained in this way, let $\bar{X}^*_\theta$ be the corresponding sample mean, and let $T^*_\theta = g(\bar{X}^*_\theta)$ be the bootstrap estimate. Then an approximation to the p-value for a test of $H_0 : g(\mu) = \theta$ is given by

$$p^*_t = P\{\ T^*_\theta \geq g(\bar{x})\ \} = P\left\{\frac{g(\bar{X}^*_\theta) - g(\mu)}{h(\bar{x})} \geq a\right\},$$

where $a = \{g(\bar{x}) - g(\mu)\}/h(\bar{x})$.

Using the methods of DiCiccio and Romano (1990), who dealt with Cornish-Fisher rather than Edgeworth expansions, we can show

$$p^*_t = 1 - \Phi(\sqrt{n}a) - \frac{1}{\sqrt{n}}\hat{q}_1(\sqrt{n}a)\varphi(\sqrt{n}a) + O_P(1),$$

so that

$$p^*_t - p = O_P(1/n),$$

showing that the p-value for the exponentially tilted bootstrap is second order correct.

By extending an idea of Davison and Hinkley (1988), we can obtain a saddlepoint approximation to p^*_t using results of Jing and Robinson (1994). The cumulative generating function of $X^*_{\theta 1}$ is

$$K_{\beta(\theta)n}(t) = \log \frac{\sum_{j=1}^n e^{(\beta(\theta)+t)\cdot x_j}}{\sum_{j=1}^n e^{\beta(\theta)\cdot x_j}} = K_n\{\beta(\theta) + t\} - K_n\{\beta(\theta)\}.$$

Then using the results of Section 4.4 and Theorem 3 of Jing and Robinson (1994), as explained in Section 5 below, we solve to obtain $t(a)$, $y(a)$, $\eta(a)$ from

$$K'_{\beta(\theta)n}\{t(a)\} = K'_n\{\beta(\theta) + t(a)\} = y(a), \tag{2.9}$$

$$g\{y(a)\} = a, \tag{2.10}$$

$$t(a) = \eta(a) \cdot g'\{y(a)\}. \tag{2.11}$$

Then the approximation is

$$p_t^* = \left[1 - \Phi\{\sqrt{n} w_t^*(a)\}\right] \{1 + O_P(1/n)\}$$

where

$$w_t^*(a) = \hat{w}(a) - \frac{\log \psi\{\hat{w}(a)\}}{n\hat{w}(a)}$$

for $\hat{w}(a)$ and $\psi\{\hat{w}(a)\}$ defined as in Section 4.4 of Jing and Robinson (1994), where the quantities involved are all derived from the cumulant generating function $K_{\beta(\theta)n}(t)$. As in Section 2, the functions $\hat{w}(a)$ and $\psi\{\hat{w}(a)\}$ are analytic functions of a and so again, for $\sqrt{n}a = o(n^{1/3})$,

$$\frac{p_t^*}{p_s^*} = 1 + O_P\left\{\frac{(\sqrt{n}a)^3 \vee 1}{n}\right\}.$$

A similar result may be obtained for the family $\mathbf{F}_2$ of DiCiccio and Romano (1990). This gives weights related to those of empirical likelihood, from the equations

$$q_j = \frac{1}{n + \lambda(\theta) g'\{\mu(\theta)\} \cdot \{x_j - \mu(\theta)\}}, \tag{2.12}$$

$$g\{\mu(\theta)\} = \theta, \tag{2.13}$$

$$\mu(\theta) = \sum q_j x_j \tag{2.14}$$

and $\sum q_j = 1$. Again we obtain bootstrap samples based on these weights. These are shown in DiCiccio and Romano (1990) to give second order correct results by Edgeworth methods. So, since here we can again get a saddlepoint approximation of the Barndorff-Nielsen form, we can repeat the calculations for $\mathbf{F}_1$ to obtain the same result on relative errors. Thus we define

$$K_{\theta n}(t) = \sum q_j e^{t \cdot x_j}$$

and then solve to obtain $t(a)$, $y(a)$ and $\eta(a)$ from $K'_{\theta n}(t(a)) = y(a)$, $g\{y(a)\} = a$ and $t(a) = \eta(a) \cdot g'\{y(a)\}$.

This also holds for the family $\mathbf{F}_3$ of DiCiccio and Romano (1990). The only difference between this and the $\mathbf{F}_1$ case considered earlier is that we must replace (2.8) by

$$\beta(\theta) = \lambda(\theta) g'(\bar{x}).$$

The saddlepoint approximations and results on the relative error then follow in the same way as before, given the equivalent results for absolute errors obtained from Edgeworth methods by DiCiccio and Romano (1990).

3 TILTING TO A PROFILE LIKELIHOOD OF AN EXPONENTIAL FAMILY

Consider the exponentially tilted family with probabilities

$$q_j = \frac{1}{n} \exp\left[\beta(\mu) x_j - K_n\{\beta(\mu)\}\right]$$

on x_j for $j = 1, \ldots, n$ where

$$K_n'\{\beta(\mu)\} = \mu. \tag{3.1}$$

If we consider the log-likelihood

$$L(\mu) = n\left[\beta(\mu) \cdot \bar{x} - K_n\{\beta(\mu)\}\right] - n \log n \ ,$$

then to make inferences on $g(\mu) = \theta$, we find the maximum of this log-likelihood,

$$L\{\mu(\theta)\} = \sup\{L(\mu) : g(\mu) = \theta\}$$

using the method of Lagrange multipliers. Then $\mu(\theta)$ is the maximum likelihood estimate of μ subject to $g(\mu) = \theta$, obtained as the solution of

$$n\beta'\{\mu(\theta)\}\{\bar{x} - \mu(\theta)\} + \lambda(\theta) g'\{\mu(\theta)\} = 0 \tag{3.2}$$

and

$$g\{\mu(\theta)\} = \theta, \tag{3.3}$$

where we note that $\beta'(\mu) = 1/K_n''\{\beta(\mu)\}$.

Notice that (3.2) differs from (2.8). Again we may sample $X_{\beta 1}^*, \ldots, X_{\beta n}^*$ from the tilted family where $\beta \equiv \beta\{\mu(\theta)\}$ is obtained from (3.1) - (3.3). If $\beta(\theta)$ is the value defined in Section 2.2 then

$$\beta - \beta(\theta) = O_P(1/n)$$

so the Edgeworth expansion for

$$p^*_{p\ell} = P\left\{\frac{g(\bar{X}^*_\beta) - \theta}{h(\bar{x})} \geq a\right\}$$

is exactly that obtained in Section 2.2 for the families $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$. It follows that

$$p^*_{p\ell} - p = O_P(1/n).$$

Further, we can proceed to obtain a Barndorff-Nielsen form of the saddlepoint approximation for $p^*_{p\ell}$ and in the same way as in Section 2.2 we can show that

$$\frac{p^*_{p\ell}}{p} = 1 + O_P\left\{\frac{(\sqrt{n}a)^3 \vee 1}{n}\right\}.$$

4 NUMERICAL STUDY

In this section we describe the results of a numerical study to illustrate and compare the methods that have been discussed, and in particular to demonstrate the closeness of the various p-value estimates along with their saddlepoint approximations. We used the S-Plus statistical package in our work (see Becker et. al. 1988) augmented by C programming code to solve the Lagrangian equations (2.9)-(2.11) and (2.12)-(2.14). In particular, we examined two different statistical contexts, using two different distributions in each case.

The first context involves testing $H_0 : \mu = 0$ for the mean in the family (a) $F \sim N(\mu, 1)$ and also in the mixture family (b) $F \sim .9 \times N(\mu, 1) + .1 \times N(\mu, 10^2)$. Samples of size $n = 10, 20, 40$ were examined. Tables 1 and 2 respectively provide the results (ordered according to increasing p-values) of ten trials conducted at the sample size $n = 20$. In order to obtain a reasonable spread of p-values near the usual value of interest 0.05, we used $\mu = 0.4$ in case (a) and $\mu = 0.7$ in case (b). In each of these twenty trials, 100,000 Monte Carlo trials were used to compute the bootstrap p-values appearing in the columns labeled 'Studentized Bootstrap', 'Tilted Bootstrap', and 'Emp. Lik. Bootstrap'. The p-values obtained by using the saddlepoint method via exponential tilting and empirical likelihood are given in the columns labeled 'SP Tilting' and 'SP Emp. Lik.' The columns labeled 'Monte Carlo' are based on (1.1) using the true sampling distributions (a) and (b) respectively. In Table 1 the true p-value (which is easily computed in this case using the t-distribution with 19 degrees of freedom) is given in the column labeled 'Exact'. The standard errors shown in parentheses are based on the usual Binomial standard error formula $\sqrt{\theta(1-\theta)/100,000}$ where θ is the accompanying proportion. Examination of these standard errors shows that the use of 100,000 Monte Carlo simulations is adequate here. The results for case (a) show the methods to be fairly comparable, with the saddlepoint methods consistently being slightly better than the bootstrap methods. The same pattern holds also for case (b). Numerically, the bootstrap methods were much more time consuming than the saddlepoint methods, by a factor of about 500.

We remark that for $n = 40$ (not shown) the results are all much closer together, as might be expected, however the saddlepoint methods still consistently outperform the bootstrap methods. For $n = 10$ we observe essentially similar patterns, however in this case a substantial proportion (roughly 15%) of the samples result in failure of the tilting and empirical likelihood methods because the range of the sample may fail to contain the hypothesized mean; in those cases the required Lagrangian will not exist.

The second context which we examined involves testing $H_0 : \rho = 0$ for the correlation coefficient in the case when the means are known to be 0, and

$$\text{(c)} \qquad F \sim N\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right\}$$

(regarded as a member of a 3 parameter family), and also in the mixture family

$$\text{(d)} \qquad F \sim .9 \times N\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right\} + .1 \times N\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 10^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right\}.$$

In these examples, the role of the function $g(\cdot)$ is played by $R = \sum X_j Y_j / \sqrt{\sum X_j^2 \cdot \sum Y_j^2}$

considered as a function of the mean of the vectors (X_jY_j, X_j^2, Y_j^2). We used $\rho = 0.4$ in case (c) and $\rho = 0.75$ in case (d). Tables 3 and 4 show the results for 10 trials with sample sizes $n = 20$, essentially as before, except that the last column labeled 'Profile Tilting' is based on the method described in Section 3. The profile tilting results are close to those obtained by the tilted bootstrap, but neither of these is as accurate as the results obtained by the saddlepoint methods. The overall pattern of the results in the cases (c) and (d) is similar to those of cases (a) and (b). Of course, example (d) is the most difficult of the inference problems considered here, and so the results of Table 4, as might be expected, are substantially less accurate than in the three other cases. Some further simulation results are given in Feuerverger, Robinson and Wong (1998).

5 TECHNICAL DETAILS

We give an indication of the proof needed to obtain the Barndorff-Nielsen saddlepoint approximation in the general case for the exponentially tilted bootstrap (such a proof also covers the approximations of Sections 2.1 and 2.2). This is given in more detail in Jing and Robinson (1994) and follows ideas of Daniels and Young (1992). We give it here to provide a short explanation on why the result holds, in a form which is more readily understood.

Following an exponential tilt of $X_1, \ldots, X_n$, a formal Edgeworth approximation to the density of the tilted mean leads to

$$P\{g(\bar{X}_\beta^*) \geq a\} = \int_{\{y:g(y)=a\}} \frac{e^{-n\Lambda(y)}}{(2\pi/n)^{k/2}\Delta(y)^{1/2}} dy \{1 + O(1/n)\},$$

where

$$\Lambda(y) = \inf_t \{t \cdot y - K_{\beta n}(t)\}$$

and

$$\Delta(y) = \det\left[K''_{\beta n}\{t(y)\}\right].$$

We change variables in the integral from y to (u, v) where $u = g(y)$; the transformation has non-zero Jacobian J (that is it is one to one). Then

$$P\left\{g(\bar{X}_\beta^*) \geq a\right\} = \frac{1}{\sqrt{2\pi/n}} \int_a^\infty e^{-nH(u)} \int_{R^{k-1}} \frac{e^{-n(\Lambda(y)-H(u))}}{(2\pi/n)^{(k-1)/2}\Delta(y)^{1/2}} J(y)\, du\, dv \left\{1 + O(1/n)\right\},$$

where

$$H(u) = \inf\{\Lambda(y) : g(y) = u\} = \inf_{y,\lambda}[\Lambda(y) + \lambda\{g(y) - u\}],$$

using the method of Lagrange multipliers. This minimization leads to equations (2.9) - (2.11). This definition of $H(u)$ forces $\Lambda\{y(u,v)\} - H(u)$ to be approximable in a neighbourhood of $v = 0$ by a quadratic form $v^T Av$. So

$$P\{g(\bar{X}_\beta^*) \geq a\} = \frac{1}{\sqrt{2\pi/n}} \int_a^\infty e^{-nH(u)} G(u)\, du \left\{1 + O(1/n)\right\}$$

where the form of $G(u)$ is given in Jing and Robinson (1994), and need not be given explicitly here. Now setting $w = \sqrt{2H(u)}\,\mathrm{sign}(u)$, we can proceed exactly as in Jing, Feuerverger and Robinson (1994) to obtain the Barndorff-Nielsen formula. Some further details are given in Feuerverger, Robinson and Wong (1998).

Acknowledgements.

This work was supported by the Natural Sciences and Engineering Research Council of Canada. The second author thanks the Department of Statistics, University of Toronto for research facilities generously provided.

References

Barndorff-Nielsen, O.E. (1986). Inference based on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika*, 73, 307-322.

Becker, R.A., Chambers, J.A. and Wilks, A.R. (1988). *The New S Language.* Wadsworth & Brooks/Cole, Pacific Grove, California.

Daniels, H.E. and Young, G.A. (1991). Saddlepoint approximation for the studentized mean, with an application to the bootstrap. *Biometrika*, 78, 169-179.

Davison, A.C., and Hinkley, D.V. (1988). Saddlepoint approximations in resampling methods. *Biometrika*, 75, 417-431.

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

Feuerverger, A., Robinson, J. and Wong, A. (1998). On the second order relative accuracy of certain bootstrap and saddlepoint approximation procedures. Technical Report No. 9801, Department of Statistics, University of Toronto.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Statist.*, 16, 927-985.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer, New York.

Jing, B.Y., Feuerverger, A., and Robinson, J. (1994). On the bootstrap saddlepoint approximations. *Biometrika*, 81, 211-215.

Jing, B.Y., and Robinson, J. (1994). Saddlepoint approximations for marginal and conditional probabilities of transformed variables. *Ann. Statist.*, 22, 1115-1132.

Table 1. Case (a).

Trial	Exact	Monte Carlo	Studentized Bootstrap	Tilted Bootstrap	Emp. Lik. Bootstrap	SP Tilting	SP Emp. Lik.
1	0.00103	0.00084 (0.00009)	0.00020 (0.0000)	0.00000 (0.0000)	0.00027 (0.0001)	0.00000	0.00012
2	0.00264	0.00256 (0.00016)	0.00107 (0.0001)	0.00043 (0.0001)	0.00229 (0.0002)	0.00038	0.00280
3	0.01021	0.01016 (0.00032)	0.01069 (0.0003)	0.00323 (0.0002)	0.00596 (0.0002)	0.00331	0.00602
4	0.01097	0.01115 (0.00033)	0.01381 (0.0004)	0.00521 (0.0002)	0.00902 (0.0003)	0.00500	0.00911
5	0.04375	0.04381 (0.00065)	0.07612 (0.0008)	0.05571 (0.0007)	0.06388 (0.0008)	0.05500	0.06434
6	0.04854	0.04883 (0.00068)	0.05417 (0.0007)	0.03749 (0.0006)	0.04208 (0.0006)	0.03832	0.04260
7	0.05860	0.05905 (0.00075)	0.06654 (0.0008)	0.04914 (0.0007)	0.05293 (0.0007)	0.04882	0.05288
8	0.06085	0.06091 (0.00076)	0.06702 (0.0008)	0.05505 (0.0007)	0.06728 (0.0008)	0.05442	0.06647
9	0.08656	0.08618 (0.00089)	0.09022 (0.0009)	0.07455 (0.0008)	0.08129 (0.0009)	0.07484	0.08154
10	0.31535	0.31625 (0.00147)	0.31999 (0.0015)	0.31520 (0.0015)	0.31674 (0.0015)	0.31577	0.31644

Table 2. Case (b).

Trial	Monte Carlo	Studentized Bootstrap	Tilted Bootstrap	Emp. Lik. Bootstrap	SP Tilting	SP Emp. Lik.
1	0.00833 (0.00029)	0.00745 (0.0003)	0.00171 (0.0001)	0.00364 (0.0002)	0.00167	0.00369
2	0.01916 (0.00043)	0.00171 (0.0001)	0.00000 (0.0000)	0.00029 (0.0001)	0.00000	0.00027
3	0.02223 (0.00047)	0.07556 (0.0008)	0.05060 (0.0007)	0.05912 (0.0007)	0.05012	0.05919
4	0.03077 (0.00055)	0.03440 (0.0006)	0.02929 (0.0005)	0.04129 (0.0006)	0.02913	0.04239
5	0.03766 (0.00060)	0.00608 (0.0002)	0.00335 (0.0002)	0.02212 (0.0005)	0.00366	0.02217
6	0.04550 (0.00066)	0.01007 (0.0003)	0.00289 (0.0002)	0.01950 (0.0004)	0.00252	0.01910
7	0.07069 (0.00081)	0.10696 (0.0010)	0.08663 (0.0009)	0.09743 (0.0009)	0.08579	0.09645
8	0.10029 (0.00095)	0.10923 (0.0010)	0.09479 (0.0009)	0.10391 (0.0010)	0.09360	0.10243
9	0.18853 (0.00124)	0.24338 (0.0014)	0.20185 (0.0013)	0.21300 (0.0013)	0.20181	0.21053
10	0.42858 (0.00156)	0.46342 (0.0016)	0.45563 (0.0016)	0.45699 (0.0016)	0.45420	0.45428

Table 3. Case (c).

Trial	Monte Carlo	Studentized Bootstrap	Tilted Bootstrap	Emp. Lik. Bootstrap	SP Tilting	SP Emp. Lik.	Profile Tilting
1	0.00113 (0.00011)	0.00541 (0.0002)	0.00014 (0.0000)	0.00041 (0.0001)	0.00100	0.00126	0.00132 (0.0001)
2	0.00315 (0.00018)	0.00600 (0.0002)	0.00041 (0.0001)	0.00150 (0.0001)	0.00023	0.00138	0.00083 (0.0001)
3	0.00716 (0.00027)	0.02056 (0.0004)	0.00100 (0.0001)	0.00145 (0.0001)	0.00198	0.00230	0.00120 (0.0001)
4	0.01023 (0.00032)	0.03799 (0.0006)	0.00690 (0.0003)	0.00923 (0.0003)	0.00785	0.00960	0.00740 (0.0003)
5	0.02159 (0.00046)	0.03231 (0.0006)	0.00891 (0.0003)	0.01040 (0.0003)	0.03218	0.02042	0.01784 (0.0003)
6	0.02814 (0.00052)	0.03036 (0.0005)	0.00540 (0.0002)	0.00821 (0.0003)	0.02600	0.02942	0.0273 (0.0002)
7	0.04198 (0.00063)	0.06618 (0.0008)	0.01820 (0.0004)	0.02055 (0.0004)	0.01079	0.01947	0.01733 (0.0004)
8	0.13191 (0.00107)	0.16010 (0.0012)	0.09242 (0.0009)	0.09326 (0.0009)	0.07041	0.09877	0.09691 (0.0009)
9	0.27801 (0.00142)	0.28315 (0.0014)	0.26297 (0.0014)	0.26490 (0.0014)	0.25140	0.26273	0.26707 (0.0014)
10	0.37931 (0.00153)	0.40125 (0.0015)	0.39212 (0.0015)	0.39079 (0.0015)	0.38109	0.38528	0.38754 (0.0015)

Table 4. Case (d).

Trial	Monte Carlo	Studentized Bootstrap	Tilted Bootstrap	Emp. Lik. Bootstrap	SP Tilting	SP Emp. Lik.	Profile Tilting
1	0.00142 (0.00012)	0.00000 (0.0000)	0.00032 (0.0001)	0.00029 (0.0001)	0.00065	0.00059	0.00037 (0.0001)
2	0.00678 (0.00026)	0.00641 (0.0003)	0.00000 (0.0000)	0.00102 (0.0001)	0.00000	0.00395	0.00000 (0.0000)
3	0.01610 (0.00040)	0.05344 (0.0007)	0.00013 (0.0000)	0.00282 (0.0002)	0.00093	0.00342	0.00109 (0.0001)
4	0.02384 (0.00048)	0.00167 (0.0001)	0.00058 (0.0001)	0.00041 (0.0001)	0.00158	0.00172	0.00163 (0.0001)
5	0.03739 (0.00060)	0.00259 (0.0002)	0.00001 (0.0000)	0.00003 (0.0000)	0.00003	0.00021	0.00001 (0.0000)
6	0.05492 (0.00072)	0.11923 (0.0010)	0.03028 (0.0005)	0.02813 (0.0005)	0.03274	0.04158	0.02946 (0.0005)
7	0.06548 (0.00078)	0.00220 (0.0001)	0.00059 (0.0001)	0.00048 (0.0001)	0.00052	0.00083	0.00052 (0.0001)
8	0.09597 (0.00093)	0.00355 (0.0002)	0.00001 (0.0000)	0.00051 (0.0001)	0.00017	0.00129	0.00071 (0.0001)
9	0.18761 (0.00123)	0.35908 (0.0015)	0.10120 (0.0010)	0.10210 (0.0010)	0.12974	0.13901	0.10510 (0.0010)
10	0.33698 (0.00149)	0.23625 (0.0013)	0.17691 (0.0012)	0.17718 (0.0012)	0.17533	0.19264	0.17713 (0.0012)

Addresses for authors:

Andrey Feuerverger, Department of Statistics, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3.

John Robinson, School of Mathematics and Statistics, The University of Sydney, Sydney, NSW 2006, Australia.

Augustine Wong, Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3.

Ross Building – Room N630

Telephone number: (416) 736-2100 ext 30154

Fax: (416) 736-5773

Email: august@mathstat.yorku.ca