STATISTICA, anno LXII, n. 1, 2002

ON A LINEAR METHOD IN BOOTSTRAP CONFIDENCE INTERVALS

Andrea Pallini

1. INTRODUCTION

Theoretical work has established that the bootstrap (Efron, 1979) can be applied to the construction of non-parametric confidence intervals in various ways. Bootstrap percentile-t confidence intervals may behave better than classical asymptotic confidence intervals (in terms of coverage error), for relatively small sample sizes n. Bootstrap percentile-t confidence intervals require asymptotically pivotal quantities, where the asymptotic variance is estimated from each bootstrap resample. Further enhancements may be produced by asymptotic corrections and adjustments of specific terms in the empirical Edgeworth expansion, which rigorously explains the asymptotic behavior of bootstrap percentile-t confidence intervals, as n -. See DiCiccio and Romano (1988), Hall (1992a), chapter 3, Efron and Tibshirani (1993), chapters 13, 14 and 22, Shao and Tu (1995), chapter 4, DiCiccio and Efron (1996) and Davison and Hinkley (1997), chapter 5.

Focussing on smooth functions of means (cf. Bhattacharya and Ghosh, 1978; and Hall, 1992a, chapters 2 and 3), most of these corrections and adjustments are difficult to implement, beyond the univariate mean example. One needs accurate estimates of the bootstrap mean, variance, symmetry and kurtosis of asymptotically pivotal and non-pivotal quantities, where the numerator and denominator are typically functions of several univariate means. A powerful linear approximation method for asymptotically pivotal and non- pivotal quantities can be a good answer to overcoming these difficulties. The classical delta method (cf. Rao, 1973, chapter 6; and Serfling, 1980, chapter 3) considers only the first term in Taylor expansions of smooth functions, which is generally of order $O_p(1)$ (with asymptotically pivotal quantities) or $O_p(n^{-1/2})$ (with asymptotically non-pivotal quantities) and cannot be used to improve (in terms of coverage error) over bootstrap percentile-t confidence intervals. The linear method we wish to study in section 2 below is based on an approximating rule, which simply rewrites a smooth function as a sum of n independent and identically distributed smooth functions, with an error of order $O_p(n^{-3/2})$ or $O_p(n^{-2})$. This set of n smooth functions can be imagined as an original random sample to be resampled, from which it is simple to obtain the bootstrap mean, variance, symmetry and kurtosis of their sample mean, which is indeed the basic linear approximation.

Let $X = \{V,, ..., V,)$ be a random sample drawn from an unknown d-variate random variable V with unknown population distribution function F and mean $p = E(\mathbf{V}_1)$. Let us construct a two-sided α-level bootstrap confidence interval for a real-valued population characteristic $\theta = g(\mu)$ of interest, where $g : \mathbb{R}^d \rightarrow R'$ is a smooth function and α E (0,1). The confidence interval is to be constructed from the bootstrap distribution of the sample estimate $\hat{\theta} = g(\bar{\mathbf{v}})$, where $\bar{\mathbf{v}} = n^{-1}\sum_{i=1}^{n} \mathbf{V}_i$. Relevant examples are provided and studied in section *3*.

Let $\mathbf{u}^{(i)}$ be the I-th component of any vector $\mathbf{u}$ E $\mathbb{R}^d$. We set $\mu_{s1 \quad s_j} = E\{(\mathbf{V}_1 - \mu)^{(s_1)}$ .. $(\mathbf{V}_1 - \mu)^{(s_j)}\}$, $g_{,, \quad s_j}(\mu) = (\partial^j/\partial \mathbf{u}^{(s_1)}$ .. $\partial \mathbf{u}^{(s_j)}) g(\mathbf{u}) \mid_{\mathbf{u} = \mu}$. The asymptotic variance of $n^{1/2}\,\hat{8}$ is

$$\sigma^2 = \sum_{s_1=1}^{d} \sum_{s_2=1}^{d} g_{s_1}(\mu) g_{s_2}(\mu) \mu_{s_1 s_2} . \tag{1}$$

We write $\hat{\sigma}^2$ to indicate a natural estimate of σ^2, which follows from (1) by substituting p with $\bar{\mathbf{v}}$, and $p_{,,,,}$ with its sample counterpart $\hat{\mu}_{s_1 s_2}$. A classical confidence interval may be obtained from the asymptotic Normal $N(0, 1)$ distribution of the quantity

$$T_n = \frac{n^{1/2}\{g(\bar{\mathbf{v}}) - g(\mu)\}}{\hat{\sigma}} \tag{2}$$

Let $X'' = \{\mathbf{V}_1^*, ..., \mathbf{V}_n^*\}$ be a generic bootstrap resample, drawn with replacement from the original sample X. We denote by $\bar{\mathbf{v}}^*$ and $\hat{\sigma}^*$ the bootstrap version of $\bar{\mathbf{v}}$ and 6, computed from $X*$, not from X. Define $T_n^* = n^{1/2}\{g(\bar{\mathbf{v}}^*) - g(\bar{\mathbf{v}})\}/\hat{\sigma}^*$ to be the bootstrap version of *(2)*.Two-sided bootstrap percentile-t confidence intervals with nominal coverage a are defined as

$$J = (\hat{\theta} - n^{-1/2}\hat{\sigma}\,\hat{v}_\beta, \hat{\theta} + n^{-1/2}\hat{\sigma}\,\hat{v}_\beta), \tag{3}$$

where $\beta = (1 + \alpha)/2$ and $\hat{v}_\beta$ is the bootstrap yuantile, which is (cf. Hall, *1986)* the exact solution to the equation $P(T_n^* \leq u \mid X) = \beta$. Bootstrap percentile-t confidence intervals *(3)*typically require Monte Carlo simulation of a sufficiently large number m of independent resamples. Ignoring simulation error, their coverage error is known to be of order n^{-1}, that is $P(8 \mathrm{E} J) - \alpha = O(n^{-1})$.

Better two-sided bootstrap confidence intervals include Edgeworth-corrected intervals, intervals constructed after removing skewness by a suitable transformation, bias-corrected, accelerated-bias-corrected confidence intervals and short confidence intervals. See section 4. Specifically oriented techniques have been proposed for simplifying the use of these and other asymptotic bootstrap confidence intervals. See, for instance, DiCiccio and Efron *(1992)*and Lee and Young *(1995)*.Here, we aim to obtain a simple and more flexible way of constructing the aforementioned bootstrap confidence intervals. In section 4, we start from the new linear approximation method for asymptotically pivotal and non-pivotal smooth functions of means proposed in section 2. Bootstrap cumulants are easily approximated, without needing Monte Carlo simulation. The linear approximation method reproduces achievable coverage errors of the asymptotic bootstrap confidence intervals in terms of order, which can be from $O(n^{-1})$ to $O(n^{-2})$, according to definitions in section 4. In section 5, we discuss results of a simulation study on empirical coverage probability with the examples in section 3.

2. A LINEAR APPROXIMATION METHOD

In order to construct bootstrap confidence intervals, we need to represent asymptotically pivotal quantity T_n as a unique smooth function of means which is a function of both the numerator and denominator 6 in (2). Domain of T_n generally becomes a vector of means, obtained by extending every random vector V, in sample X, $i = 1, \ldots, n$, from the dimension d to a specific dimension e, where e > d. In particular, we need to define vectors X, with d components $\mathbf{X}_i^{(s_1)} = \mathbf{V}_i^{(s_1)}$, $s_1 = 1, \ldots, d$, and the remaining e - d components, which equate specific transformations of V,, $i = 1, \ldots, n$ (cf. Bhattacharya and Ghosh, 1978; and Hall, 1992a, chapters 2 and 3). See examples in section 3 below. For brevity, we still use X to indicate the original 'sample' consisting of vectors X, with dimension e, $i = 1, \ldots, n$, that is $\mathcal{X} = \{$X,, ..., X,$\}$. Accordingly, we set p = E($\mathbf{X}_1$) and p,, $_{s_j}$ = E$\{(\mathbf{X}_1 - \mu)^{(s_1)} \ldots$ (X, $- \mu)^{(s_j)}\}$. With this notation, we assume that there exists a known smooth function of means h, which defines the asymptotic variance σ^2 as $\sigma^2 = h(\mu)^2$. The natural estimate $\hat{\sigma}^2$ of σ^2 may be written as 6' $= h(\bar{\mathbf{x}})^2$.

Let A : R' $\rightarrow \mathbb{R}^1$ be the smooth function satisfying $A(\mu) = 0$, defined as $A(\mathbf{u}) = \{g(\mathbf{u}) - g(\mu)\}/h(\mathbf{u})$, where **0** is the zero element in R^c and **u** E R'. Asymptotically pivotal quantity T_n given by (2) may be written as

$$T_n = n^{1/2} A(\bar{\mathbf{x}}) \,. \tag{4}$$

We assume the following conditions.

(C1). Function A has bounded continuous derivatives of order r_1 in a neighborhood of p, where $r_1 \geq 4$.

We may write partial derivatives of A as

$$a_{s_1 \ldots s_j}(\mu) = \left. \frac{\partial^j A(\mathbf{u})}{\partial \mathbf{u}^{(s_1)} \ldots \partial \mathbf{u}^{(s_j)}} \right|_{\mathbf{u}=\mu} , \tag{5}$$

where $s_j = 1, \ldots,$ e, and $j = 1, \ldots, r_1$. It is easy to see that

$$\sum_{s_1=1}^{e} \sum_{s_2=1}^{e} a_{s_1}(\mu)\, a_{s_2}(\mu)\, \mu_{s_1 s_2} = 1. \tag{6}$$

The sample distribution of T_n is asymptotically Normal $N(0, 1)$, as $n \rightarrow \infty$ (cf. Kao, 1973, section 6a; and Serfling, 1980, section 3.3). We also assume that the distribution of T_n admits an Edgeworth expansion. We let $\| \mathbf{Y}_1 \| = \{(\mathbf{Y}_1^{(1)})^2 + \ldots + (\mathbf{Y}_1^{(c)})^2\}^{1/2}$. Briefly, what follows is required.

(C2). $E(\| \mathbf{Y}_1 \|^{r_2}) < \infty$, for $r_2 \geq r_1 + 2$.

(C3). $\lim_{\|\mathbf{u}\| \to \infty} \sup | E\{\exp(i\mathbf{u}^T \mathbf{Y}_1)\} | < 1$, $\mathbf{u} \in \mathbb{R}^d$

Condition (C3) holds if F has a non-degenerate absolutely continuous component. Details for the validity of the Edgeworth expansion, under conditions (C1), (C2) and Cramer's condition (C3), are fully explained in Hall (1992a), chapter 5. See also Bhattacharya and Ghosh (1978), Theorem 2, and Petrov (1995), Theorem 5.18.

We denote by Φ and ϕ the distribution function and the density function of a standard Normal $N(0, 1)$ variate.

In Hall (1992a), Theorem 2.1, it is shown that the j-th cumulant κ_j of T_n has the power expansion

$$\kappa_j = n^{-(j-2)/2}\{\kappa_{j1} + n^{-1}\kappa_{j2} + n^{-2}\kappa_{j3} + O(n^{-3})\} , \tag{7}$$

where the constants κ_{ji} depend on distribution function F, for i = 1, 2, 3, ..., with $\kappa_{11} = 0$ and $\kappa_{21} = 1$.

The Edgeworth expansion of the distribution of T_n is

$$P(T_n \leq u) = \Phi(u) + n^{-1/2}q_1(u)\, \phi(u) + n^{-1}q_2(u)\, \phi(u) + O(n^{-3/2}) , \tag{8}$$

where $u \in \mathbb{R}^1$ and

$$q_1(u) = -\kappa_{12} - \kappa_{31}(u^2 - 1)/6 , \tag{9}$$

$$q_2(u) = -(\kappa_{12}^2 + \kappa_{22})\, u/2 - (4\kappa_{12}\kappa_{31} + \kappa_{41})\, u(u^2 - 3)/24$$
$$-\kappa_{31}^2 u(u^4 - 10u^2 + 15)/72 . \tag{10}$$

Polynomials $q_j(u)$ in (8), $u \in \mathbb{R}^1$, are of degree at most $3j - 1$. They are an odd or even function according to whether j is even or odd, respectively.

2.1. The *linear approximation*

A preliminary standardization for location of original sample observations is necessary for the linear approximation method we want to propose. It is convenient to take $\mathbf{Y}_i = \mathbf{X}_i - \mu$, $i = 1, ..., n$. We let $\bar{\mathbf{y}} = n^{-1}\sum_{i=1}^{n} \mathbf{Y}_i$. Observe that, trivially, $g(\bar{\mathbf{y}} + \mu) = g(\bar{\mathbf{x}})$ and $h(\bar{\mathbf{y}} + \mu) = h(\bar{\mathbf{x}})$. Asymptotically pivotal quantity T_n given by (4) can be rewritten as

$$T_n = n^{1/2}A(\bar{\mathbf{y}} + \mu) . \tag{11}$$

For every $i = 1, ..., n$, we define $A_i = A(n^{-1}\mathbf{Y}_i + \mu) = \{G_i - g(\mu)\}/H_i$, where $G_i = g(n^{-1}\mathbf{Y}_i + \mu)$ and $H_i = h(n^{-1}\mathbf{Y}_i + \mu)$ is so that $H_i > 0$. Under condition (C1), a linear approximation to T_n given by (11) then is

$$W_n = n^{1/2}\sum_{i=1}^{n} A_i . \tag{12}$$

Quantity W_n approximates to T_n with an error of order $n^{-3/2}$,

$$T_n = W_n + O_p(n^{-3/2})\,. \tag{13}$$

Standardizing the original observations in sample X for location p characterizes an invariance property which theoretically justifies the use of W_n as a convenient approximation. An asymptotically pivotal smooth function of means can be written as a sum (mean) of n independent and identically distributed smooth functions, with an error of order $O_p(n^{-3/2})$. Keeping in mind that $E(\bar{y}) = E(n^{-1}\mathbf{Y}_1) = 0$, linear approximation W_n is obtained using Taylor expansion of function $T_n = n^{1/2}A(\bar{y} + p)$ around p and approximating to it with the sum of n Taylor expansions of functions A, $= A(n^{-1}\mathbf{Y}_i + \mu)$ around p, i = 1, . ., n. In Appendix 7.1, (12) and (13) are proved with details.

The essential point is that function A,, for every $i = 1, \ldots$, n, has partial derivatives a,, $_{s_j}(\mu)$ of order j, which equate partial derivatives *(5)*of function A of the same order, for every $j = 1, \ldots$, r,. From another theoretical perspective, compared with $T_n = n^{1/2}A(\bar{y} + p)$, linear approximation $W_n = n^{1/2}\sum_{i-1}^{n} A_i$ defines by transformation A, a different coordinate system with $A(n^{-1}\mathbf{u} + p)$-coordinates instead of $(n^{-1}\mathbf{u} + p)$-coordinates, u E R^c, without varying j-th order partial derivatives a,, $_{s_j}(\mu)$, $j = 1, \ldots, r_1$.

Recalling (13), it is easy to see that W_n as T_n is asymptotically Normal $N(0, 1)$, as $n \to \infty$ (cf. Rao, 1973, section 6a; and Serfling, 1980, section 3 3). It follows that

$$P(T_n \le u) = P(W_n \le u) + O(n^{-3/2})\,, \tag{14}$$

u E $\mathbb{R}^1$ (cf. Khattacharya and Ghosh, 1978; and Hall, 1992a, section 2.7); in particular, Edgeworth expansions of T_n and W_n disagree only in terms of order $O(n^{-3/2})$ or smaller. From (14), it follows that approximation W_n given by (12) thus entails an expansion parallel to (7), with the same constants K,,, κ_{12}, κ_{21}, κ_{22}, κ_{31} and κ_{41}. See Appendix 7.2, for a different proof.

3. SOME EXAMPLES

3.1. *Univariate mean*

The random sample is $X = \{\mathbf{X}_1, \ldots$, X,}, where vector observations X, have dimension c = 2, that is $\mathbf{X}_i = (V_i, V_i^2)^T$, for $i = 1, \ldots$, n. Here, the random variable under study is the first component $\mathbf{X}_1^{(1)} = V_1$, which has dimension $d = 1$, while the second component $\mathbf{X}_1^{(2)} = V_1^2$ helps to define the sample variance of V_1. The asymptotically pivotal mean T_n of the random variable V_1, function of means $\bar{\mathbf{x}}^{(1)} = n^{-1}\sum_{i=1}^{n}\mathbf{X}_i^{(1)}$ and $\bar{\mathbf{x}}^{(2)} = n^{-1}\sum_{i=1}^{n}\mathbf{X}_i^{(2)}$, is $T_n = n^{1/2}(\bar{\mathbf{x}}^{(1)} - \mu^{(1)})\{\bar{\mathbf{x}}^{(2)} - (\bar{\mathbf{x}}^{(1)})^2\}^{-1/2}$, where $g(\bar{\mathbf{x}}) = \bar{\mathbf{x}}^{(1)}$, $g(\mu) = \mu^{(1)} = E(\mathbf{X}_1^{(1)})$, and $h(\bar{\mathbf{x}}) =$ {E'" - %(")')]i'. We may also write T_n as T„ $= n^{1/2}A(\bar{y} + p)$,

$$T_n = \frac{n^{1/2}\{(\bar{y}^{(1)} + \mu^{(1)}) - \mu^{(1)}\}}{\{(\bar{y}^{(2)} + \mu^{(2)}) - (\bar{y}^{(1)} + \mu^{(1)})^2\}^{1/2}}\,,$$

where $\mu^{(2)} = E(\mathbf{X}_1^{(2)})$. Approximation $W_n = n^{1/2}\sum_{i=1}^{n} A_i$ is defined by

$$A_i = \frac{(n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)}) - \mu^{(1)}}{\{(n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)}) - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}^{1/2}}$$

$$= \frac{n^{-1}\mathbf{Y}_i^{(1)}}{\{(n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)}) - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}^{1/2}}.$$

3.2. *Univariate* variance

The random sample is $X = \{X_1, ..., X_n\}$, where $\mathbf{X}_i = (V_i, V_i^2, V_i^3, V_i^4)^T$. Here, dimension e = 4, while dimension d = 2; $\mathbf{X}_1^{(1)} = V_i$ and $\mathbf{X}_1^{(2)} = V_i^2$ are the components which define the variance of the univariate random variable V_1. Recalling (4) and $A(\bar{y} + \mu) = \{g(\bar{y} + \mu) - g(\mu)\}/h(\bar{y} + \mu)$, the asymptotically pivotal variance $T_n = n^{1/2}A(\bar{y} + \mu)$ of the random variable V_1 is defined by

$$g(\bar{y} + \mu) = \bar{y}^{(2)} + \mu^{(2)} - (\bar{y}^{(1)} + \mu^{(1)})^2\}, \qquad g(\mu) = \mu^{(2)} - (\mu^{(1)})^2,$$

$$h(\bar{y} + \mu) = \{\bar{y}^{(4)} + \mu^{(4)} - (\bar{y}^{(2)} + \mu^{(2)})^2 - 4(\bar{y}^{(1)} + \mu^{(1)})(\bar{y}^{(3)} + \mu^{(3)})$$
$$+ 8(\bar{y}^{(1)} + \mu^{(1)})^2(\bar{y}^{(2)} + \mu^{(2)}) - 4(\bar{y}^{(1)} + \mu^{(1)})^4\}^{1/2}.$$

Approximation $W_n = n^{1/2}\sum_{i=1}^{n} A_i$ is obtained by $A_i = \{G_i - g(\mu)\}/H_i$, where

$$G_i = n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2,$$

$$H_i = \{n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)} - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2 - 4(n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})(n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)})$$
$$+ 8(n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2(n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)}) - 4(n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^4\}^{1/2}.$$

3.3. Ratio of means

The random sample is $X = \{X_1, ..., \mathbf{X}_n\}$, where vector observations $\mathbf{X}_i$ have dimension e = 5, $\mathbf{X}_i = (\mathbf{V}_i^{(1)}, \mathbf{V}_i^{(2)}, (\mathbf{V}_i^{(1)})^2, (\mathbf{V}_i^{(2)})^2, \mathbf{V}_i^{(1)}\mathbf{V}_i^{(2)})^T$, i = 1, ..., n. Suppose that $\mathbf{V}_1^{(2)}$ ranges in a set of positive values. Note that dimension d = 2; $\mathbf{X}_1^{(1)} = \mathbf{V}_i^{(1)}$ and $\mathbf{X}_1^{(2)} = \mathbf{V}_i^{(2)}$ are the components, which define the ratio of means of the marginals $\mathbf{V}_1^{(1)}$ and $\mathbf{V}_1^{(2)}$ of the bivariate random variable $(\mathbf{V}_1^{(1)}, \mathbf{V}_1^{(2)})^T$. Recalling (4) and $A(\bar{y} + \mu) = \{g(\bar{y} + \mu) - g(\mu)\}/h(\bar{y} + \mu)$, the asymptotically pivotal ratio of means $T_n = n^{1/2}A(\bar{y} + \mu)$ is defined by

$$g(\bar{y} + \mu) = \frac{\bar{y}^{(1)} + \mu^{(1)}}{\bar{y}^{(2)} + \mu^{(2)}}, \qquad g(\mu) = \frac{\mu^{(1)}}{\mu^{(2)}},$$

and $h(\bar{y} + \mu)$, obtained from (1). Variance and covariance estimates $\hat{\mu}_{11}$, $\hat{\mu}_{22}$ and $\hat{\mu}_{12} = \hat{\mu}_{21}$ in (1) can be written as in Example 1. That is,

$\hat{\mu}_{11} = (\bar{y}^{(3)} + \mu^{(3)}) - (\bar{y}^{(1)} + \mu^{(1)})^2$,

$\hat{\mu}_{22} = (\bar{y}^{(4)} + \mu^{(4)}) - (\bar{y}^{(2)} + \mu^{(2)})^2$,

$\hat{\mu}_{12} = (\bar{y}^{(5)} + \mu^{(5)}) - (\bar{y}^{(1)} + \mu^{(1)}) (\bar{y}^{(2)} + \mu^{(2)})$.

Derivatives $g_1(\mu)$ and $g_2(\mu)$ in (1) may be set as

$$g_1(\bar{y} + \mu) = \frac{1}{\bar{y}^{(2)} + \mu^{(2)}}, \quad g_2(\bar{y} + \mu) = -\frac{\bar{y}^{(1)} + \mu^{(1)}}{(\bar{y}^{(2)} + \mu^{(2)})^2}.$$

We also let $\hat{\mu}_{s_1s_2,i} = \hat{\mu}_{s_1s_2}(n^{-1}\mathbf{Y}_i + \mathrm{p})$, after considering $\hat{\mu}_{s_1s_2}$ as any term in a smooth function of means $A(\bar{y} + \mathrm{p})$ to be approximated linearly. In particular,

$$\hat{\mu}_{11,i} = (n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)}) - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2, \tag{15}$$

$$\hat{\mu}_{22,i} = (n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)}) - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2, \tag{16}$$

$$\hat{\mu}_{12,i} = (n^{-1}\mathbf{Y}_i^{(5)} + \mu^{(5)}) - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)}) (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)}) \tag{17}$$

Approximation $W_n = n^{1/2}\sum_{i=1}^{n} A_i$ is defined bp A, = {G, - $g(\mu)\}/H_i$, where

$$G_i = \frac{n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)}}{n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)}},$$

$$H_i = \left\{\sum_{s_1=1}^{e} \sum_{s_2=1}^{e} g_{s_1}(n^{-1}\mathbf{Y}_i + \mu) g_{s_2}(n^{-1}\mathbf{Y}_i + \mu) \hat{\mu}_{s_1s_2,i}\right\}^{1/2}, \tag{18}$$

with $\hat{\mu}_{s_1s_2,i}$ defined in (15), (16), (17), and

$$g_1(n^{-1}\mathbf{Y}_i + \mu) = \frac{1}{n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)}}, \quad g_2(n^{-1}\mathbf{Y}_i + \mu) = -\frac{n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)}}{(n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2}.$$

3.4. *Correlation coefficient*

To define an asymptotically pivotal correlation coefficient $T_n = n^{1/2}A(\bar{y} + \mathrm{p})$, a random sample X of vectors X,, i = 1, ..., n, with $e = 17$ components is needed. The $d = 5$ components defining the correlation coefficient are $\mathbf{X}_i^{(1)} = \mathbf{V}_i^{(1)}$, $\mathbf{X}_i^{(2)} = \mathbf{V}_i^{(2)}$, $\mathbf{X}_i^{(3)} = (\mathbf{V}_i^{(1)})^2$, $\mathbf{X}_i^{(4)} = (\mathbf{V}_i^{(2)})^2$, $\mathbf{X}_i^{(5)} = \mathbf{V}_i^{(1)}\mathbf{V}_i^{(2)}$. The $e = 17$ components are $\mathbf{X}_i^{(1)}$ and $\mathbf{X}_i^{(2)}$ and the $d(1 + d)/2$ distinct products $\mathbf{X}_i^{(s_1)}\mathbf{X}_i^{(s_2)}$, s_1, $s_2 = 1$, ..., d. Recalling (4) and $A(\bar{y} + \mathrm{p}) = \{g(\bar{y} + \mu) - g(\mu)\}/h(\bar{y} + \mathrm{p})$, we obtain the asymptotically pivotal correlation coefficient $T_n = n^{1/2}A(\bar{y} + \mathrm{p})$ between the marginals $\mathbf{V}_1^{(1)}$ and $\mathbf{V}_1^{(2)}$ of the bivariate random variable $(\mathbf{V}_1^{(1)}, \mathbf{V}_1^{(2)})^T$ by

$$g(\bar{y} + \mu) = \{\bar{y}^{(5)} + \mu^{(5)} - (\bar{y}^{(1)} + \mu^{(1)}) (\bar{y}^{(2)} + \mu^{(2)})\}$$
$$\cdot \{\bar{y}^{(3)} + \mu^{(3)} - (\bar{y}^{(1)} + \mu^{(1)})^2\}^{-1/2} \cdot \{\bar{y}^{(4)} + \mu^{(4)} - (\bar{y}^{(2)} + \mu^{(2)})^2\}^{-1/2},$$

$$g(\mu) = \frac{\mu^{(5)} - \mu^{(1)}\mu^{(2)}}{\{\mu^{(3)} - (\mu^{(1)})^2\}^{1/2}\{\mu^{(4)} - (\mu^{(2)})^2\}^{1/2}},$$

and $h(\bar{y} + p)$, from (1). Approximation $W_n = n^{1/2}\sum_{i=1}^{n} A_i$ is defined by $A_i = \{G_i - g(\mu)\}/H_i$, where

$$G_i = \{n^{-1}\mathbf{Y}_i^{(5)} + \mu^{(5)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})(n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})\}$$
$$\cdot \{n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}^{-1/2}$$
$$\cdot \{n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)} - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2\}^{-1/2},$$

and H_i given by (18). In particular, definitions of $\hat{\mu}_{s_1 s_2, i}$, $s_1, s_2 = 1, \dots, d$, $d = 5$, may be deduced from (15), (16), (17). Derivatives $g_{s_1}(n^{-1}\mathbf{Y}_i + p)$, $s_1 = 1, \dots, d$, are

$$g_1(n^{-1}\mathbf{Y}_i + \mu) = G_i \cdot (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})$$
$$\cdot \{n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}^{-1}$$
$$- (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)}) \cdot \{n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}^{-1/2}$$
$$\cdot \{n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)} - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2\}^{-1/2},$$

$$g_2(n^{-1}\mathbf{Y}_i + \mu) = G_i \cdot (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})$$
$$\cdot \{n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)} - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2\}^{-1}$$
$$- (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)}) \cdot \{n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}^{-1/2}$$
$$\cdot \{n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)} - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2\}^{-1/2},$$

$$g_3(n^{-1}\mathbf{Y}_i + \mu) = - G_i \cdot [2\{n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}]^{-1},$$

$$g_4(n^{-1}\mathbf{Y}_i + \mu) = - G_i \cdot [2\{n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)} - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2\}]^{-1},$$

$$g_5(n^{-1}\mathbf{Y}_i + \mu) = \{n^{-1}\mathbf{Y}_i^{(3)} + \mu^{(3)} - (n^{-1}\mathbf{Y}_i^{(1)} + \mu^{(1)})^2\}^{-1/2}$$
$$\cdot \{n^{-1}\mathbf{Y}_i^{(4)} + \mu^{(4)} - (n^{-1}\mathbf{Y}_i^{(2)} + \mu^{(2)})^2\}^{-1/2}.$$

With this example, property $H_i > 0$ is always fulfilled, if we let $l(\bar{y} + p) = \{h(\bar{y} + \mu)\}^2$ and, finally, we determine H_i as $H_i = \{l(n^{-1}\mathbf{Y}_i + \mu)\}^{1/2}$.

4. BOOTSTRAP CONFIDENCE INTERVALS

The linear approximation W_n given by *(12)* is easy to apply with the bootstrap. We consider the original sample X in place of the unknown distribution function F and substitute the means p, $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ with the sample mean $\bar{\mathbf{x}}$ and the bootstrap means $\bar{\mathbf{x}}^*$ and $\bar{\mathbf{y}}^*$, respectively. Denote by W_n^* the bootstrap version of W_n. Recalling (6), it is convenient to write W_n^* as the normalized sum (mean)

$$W_n^* = n^{1/2}\bar{b}^*\,, \tag{19}$$

where $\bar{b}^* = n^{-1}\sum_{i=1}^{n} B_i^*$, $B_i^* = nA_i^*$, with $A_i^* = \{G_i^* - g(\bar{\mathbf{x}})\}/H_i^*$ $G_i^* = g(n^{-1}\mathbf{Y}_i^* + \bar{\mathbf{x}})$ and $H_i^* = h(n^{-1}\mathbf{Y}_i^* + \bar{\mathbf{x}})$. Hereafter, the basic idea is to apply the bootstrap to the set of smooth functions $\{B_1, \ldots, B_n\}$, which are independent and identically distributed random variables. We denote by $\hat{\lambda}_j$ the j-th bootstrap cumulant of B_1^*, obtained by assigning mass n^{-1} to each function B_i, i = 1, ..., n. Since B_1^* is embedded in mean $W_n^* = n^{1/2}\bar{b}^*$, the j-th cumulant $\hat{\kappa}_j$ of W_n^* is such that $\hat{\kappa}_j = n^{-(j-2)/2}\hat{\lambda}_j$, where the j-th cumulant λ_j admits the expansion

$$\hat{\lambda}_j = \hat{\lambda}_{j1} + n^{-1}\hat{\lambda}_{j2} + n^{-2}\hat{\lambda}_{j3} + O_p(n^{-3})\,. \tag{20}$$

Considering (19), it follows that $T_n^* = W_n^* + O_p(n^{-3/2})$ and consequently that $P(T_n^* \le u) = P(W_n^* \le u) + O_p(n^{-3/2})$. Bootstrap estimate $P(T_n^* \le o \mid X)$ admits the empirical Edgeworth expansion

$$P(T_n^* \le u \mid X) = \Phi(u) + n^{-1/2}\hat{q}_1(u)\,\phi(u) + n^{-1}\hat{q}_2(u)\,\phi(u) + O_p(n^{-3/2})\,, \tag{21}$$

where $u \in \mathbb{R}^1$ and $\hat{q}_1(u)$ and $\hat{q}_2(u)$ are the bootstrap counterparts of polynomials $q_1(u)$ and $q_2(u)$, given by *(9)*and (10). We approximate polynomials $\hat{q}_1(u)$ and $\hat{q}_2(u)$, by replacing constants κ_{12}, κ_{22}, κ_{31} and κ_{41} in *(9)*and (10) with $n\hat{\lambda}_1$, $n(\hat{\lambda}_2 - 1)$, $\hat{\lambda}_3$ and $\hat{\lambda}_4$, respectively, with errors of order $O_p(n^{-1})$, and only affecting terms of order $O_p(n^{-3/2})$ or smaller in *(21)*.See Appendix *7.3,* for proofs of (20) and *(21).*

Polynomials $\hat{q}_1(u)$ and $\hat{q}_2(u)$, $u \in \mathbb{R}^1$, can be used to determine various well-known bootstrap refinements on two-sided bootstrap percentile-t confidence intervals (3) with nominal coverage $\alpha \in (0, 1)$. Theoretical coverage errors are guaranteed by straightforward applications of the bootstrap linear approximation W_n^* given by *(19)*.See Appendix *7.4.*

Let z_β be the β-level quantile of the standard Normal $N(0, 1)$ variate, that is $z_\beta = \Phi^{-1}(\beta)$, where $\beta = (1 + \alpha)/2$ and $\alpha \in (0, 1)$.

4.1. Explicit Edgeworth correction

A classical asymptotic confidence interval can be corrected in a standard fashion (cf. Hall, 1992a, section 3.8), by using the bootstrap polynomial $\hat{q}_2(u)$, $u \in \mathbb{R}^1$, in *(21)*.The Edgeworth-corrected (EC) two-sided bootstrap confidence intervals may then be defined as

$$\hat{J}_{\mathrm{EC}} = (\hat{\theta} - n^{-1/2}\hat{\sigma}z_\beta + n^{-3/2}\hat{\sigma}\hat{q}_2(z_\beta),\ \hat{\theta} + n^{-1/2}\hat{\sigma}z_\beta - n^{-3/2}\hat{\sigma}\hat{q}_2(z_\beta))\,. \tag{22}$$

The coverage error of two-sided confidence intervals $\hat{J}_{\mathrm{EC}}$ is shown to be $P(\theta \in \hat{J}_{\mathrm{EC}}) - \alpha = O(n^{-2})$. See Appendix *7.4.*

4.2. *Transformation to remove skewness*

We start from the linear approximation W_n^* given by (19), to define the transformed statistic suggested in Hall (1992a), section *3.9,* and Hall (1992b). We let $U_n^* = W_n^* - n^{1/2}\hat{\lambda}_1$. Then, $f(U_n^*) = U_n^* - n^{-1/2}(U_n^*)^2(\hat{\lambda}_3/6) + n^{-1}(U_n^*)^3(\hat{\lambda}_3^2/108) + n^{-1/2}(\hat{\lambda}_3/6)$, admits an Edgeworth expansion in which the first term is of order $O_p(n^{-1})$, not $O_p(n^{-1/2})$. Note that equation $f(U_n^*) = u$, $u \in R$, has a unique solution, $f^{-1}(u) = -n^{1/2}(6/\hat{\lambda}_3)[\{1 - (\hat{\lambda}_3/2)(n^{-1/2}u - n^{-1}(\hat{\lambda}_3/6))\}^{1/3} - 1]$. Set $\hat{q}_1'(u) = d\hat{q}_1(u)/du$, $u \in \mathbb{R}^1$. The transformation-based *(TB)* two-sided bootstrap confidence intervals can then be obtained as

$$\hat{J}_{\mathrm{TB}} = (\hat{\theta} - n^{-1/2}\hat{\sigma}f^{-1}(\tilde{v}_\beta) - \hat{\sigma}\hat{\lambda}_1,\ \hat{\theta} - n^{-1/2}\hat{\sigma}f^{-1}(\tilde{v}_{1-\beta}) - \hat{\sigma}\hat{\lambda}_1)\ , \tag{23}$$

where, by Cornish-Fisher inversion of expansion (21), $\tilde{v}_\beta = z_\beta - n^{-1/2}\hat{q}_1(z_\beta) + n^{-1}\{\hat{q}_1(z_\beta)\,\hat{q}_1'(z_\beta) - z_\beta\hat{q}_1(z_\beta)^2/2 - \hat{q}_2(z_\beta)\}$ approximates the bootstrap quantile $\hat{v}_\beta$, so that $\tilde{v}_\beta = \hat{v}_\beta + O_p(n^{-3/2})$. The coverage error is $P(\theta \in \hat{J}_{\mathrm{TB}}) - \alpha = O(n^{-2})$. See Appendix *7.4*.

4.3. *Bias-correction*

The bias-corrected (BC) and the accelerated bias-corrected (BC_a) two-sided bootstrap confidence intervals of Efron *(1987)* can be obtained as follows (cf. Konishi, *1991;* and Hall, 1992a, section 3.10). We write the asymptotically non-pivotal statistic $D_n = g(\bar{\mathbf{x}}) - g(\mu)$ as $D_n = g(\bar{\mathbf{y}} + \mu) - g(\mu)$. Under a condition parallel to *(C1)*, the linear approximation to D_n now is

$$Z_n = \sum_{i=1}^{n}\{G_i - g(\mu)\}\ , \tag{24}$$

where $G_i = g(n^{-1}\mathbf{Y}_i + \mu)$, so that

$$D_n = Z_n + O_p(n^{-2})\ . \tag{25}$$

Proofs of *(24)* and *(25)* are similar to proofs in Appendix *7.1* and Appendix *7.2*. Recalling (1), we denote by $p_1(u)$ and $p_2(u)$ polynomials of the same analytical form as $q_1(u)$ and $q_2(u)$, $u \in \mathbb{R}^1$, given by (9) and (10), except for the fact that they are derived from

$$Q_n = n^{1/2}\frac{Z_n}{\sigma}\ , \tag{26}$$

not from T_n. Under conditions (C2) and (C3), polynomials $p_1(u)$ and $p_2(u)$, $u \in \mathbb{R}^1$, characterize an Edgeworth expansion parallel to expansion (X) (cf. Hall, 1992a, chapter 2). Let us write $Q_n = n^{1/2}\bar{c}/\hat{\sigma}$, where $\hat{\sigma}^2$ estimates asymptotic variance (1), $\bar{c} = n^{-1}\sum_{i=1}^{n} C_i$ and $C_i = n\{G_i' - g(\bar{\mathbf{x}})\}$ with $G_i' = g(n^{-1}\mathbf{Y}_i + \bar{\mathbf{x}})$. We define by $\hat{p}_1(u)$ and $\hat{p}_2(u)$ the bootstrap counterparts of polynomials $p_1(u)$ and $p_2(u)$, $u \in \mathbb{R}^1$. Bootstrap cumulants of C, in $\hat{p}_1(u)$ and $\hat{p}_2(u)$, $u \in \mathbb{R}^1$, can be deduced from Appendix *7.3*.

Let $\hat{u}_\beta$ be the bootstrap quantile, which is the exact solution (cf. Hall, 1986) to the equation $P(Q_n^* \le u_\beta | X) = \beta$. Set $\hat{p}_1'(u) = d\hat{p}_1(u)/du$, $u \in \mathbb{R}^1$. The bias-corrected (BC) and the bias-corrected-accelerated (BCa) bootstrap confidence intervals can then be obtained as

$$\hat{J}_{BC} = (\tilde{u}_{1-\beta} + n^{-1}\hat{\sigma}2\hat{p}_1(0),\ \tilde{u}_\beta + n^{-1}\hat{\sigma}2\hat{p}_1(0)), \quad (27)$$

$$\hat{J}_{BCa} = (\tilde{u}_{1-\beta} + n^{-1}\hat{\sigma}2\{\hat{p}_1(z_{1-\beta}) + \hat{q}_1(z_{1-\beta})\},\ \tilde{u}_\beta + n^{-1}\hat{\sigma}2\{\hat{p}_1(z_\beta) + \hat{q}_1(z_\beta)\}), \quad (28)$$

where, by Cornish-Fisher inversion, $\tilde{u}_\beta = z_\beta - n^{-1/2}\hat{p}_1(z_\beta) + n^{-1}\{\hat{p}_1(z_\beta)\hat{p}_1'(z_\beta) - z_\beta\hat{p}_1(z_\beta)^2/2 - \hat{p}_2(z_\beta)\}$, approximates the bootstrap quantile $\hat{u}_\beta$, so that $\tilde{u}_\beta = \hat{u}_\beta + O_p(n^{-3/2})$. We observe that $P(\theta \in \hat{J}_{BC}) - a = O(n^{-1})$ and $P(\theta \in \hat{J}_{BCa}) - a = O(n^{-1})$. See Appendix 7.4.

4.4. Short confidence intervals

Short bootstrap confidence intervals (Hall, 1988; and Hall, 1992b, section 3.7) can be defined as

$$\hat{J}_{SH} = (\hat{\theta} - n^{-1/2}\hat{\sigma}\hat{\gamma},\ \hat{\theta} + n^{-1/2}\hat{\sigma}\hat{\delta}), \quad (29)$$

where $\hat{\gamma}$ and $\hat{\delta}$ are the minimizers of the interval length $\hat{\gamma} + \hat{\delta}$ subject to $P(-\hat{\delta} \le T_n^* \le \hat{\gamma} | X) = a$. Using approximation W_n^*, it follows that $\hat{\gamma}_1(z_\beta) = \hat{q}_1'(z_\beta)/z_\beta - \hat{q}_1(z_\beta)$, $\hat{\gamma}_2(z_\beta) = -\hat{q}_2(z_\beta) - (z_\beta/2)\{\hat{q}_1'(z_\beta)/z_\beta - \hat{q}_1(z_\beta)\}^2$. Short two-sided bootstrap confidence intervals are given by the solutions $\hat{\gamma} = z_\beta + n^{-1/2}\hat{\gamma}_1(z_\beta) + n^{-1}\hat{\gamma}_2(z_\beta)$, $\hat{\delta} = z_\beta - n^{-1/2}\hat{\gamma}_1(z_\beta) + n^{-1}\hat{\gamma}_2(z_\beta)$. The coverage error is $P(\theta \in \hat{J}_{SH}) - a = O(n^{-1})$. See Appendix 7.4.

5. SIMIJLATION STUDY RESULTS

5.1. *The* bootstrap linear *approximation*

In order to see the effectiveness of the linear approximation W_n^* given by (19) to asymptotically pivotal bootstrap quantities T_n^*, we plot the difference $T_n^* - W_n^*$ from a set of m = 500 independent resamples, drawn with replacement from original samples, generated from different population distributions. For the univariate mean (example 1 above) and variance (example 2) we consider original samples from the Normal $N(0, 1)$ distribution of sizes n = 10, 30 and n = 25, 100, respectively. For the ratio of means (example 3) we generate original samples of sizes n = 10, 30 from the bivariate Folded-normal distribution, with correlated marginals $|N(0, 1)|$ and correlation coefficient p = 0.5. For the correlation coefficient (example 4) we generate original samples of sizes n = 15, 50 from the bivariate Lognormal distribution, with correlated marginals exp $\{N(0, 1)\}$ and correlation coefficient p = 0.3775. All these sample sizes represent the ranges chosen below to study the empirical coverage probabilities of the bootstrap confidence intervals described in section 4.

Figures 1 and 3 show the good performance of W_n^* in approximating to T_n^*,

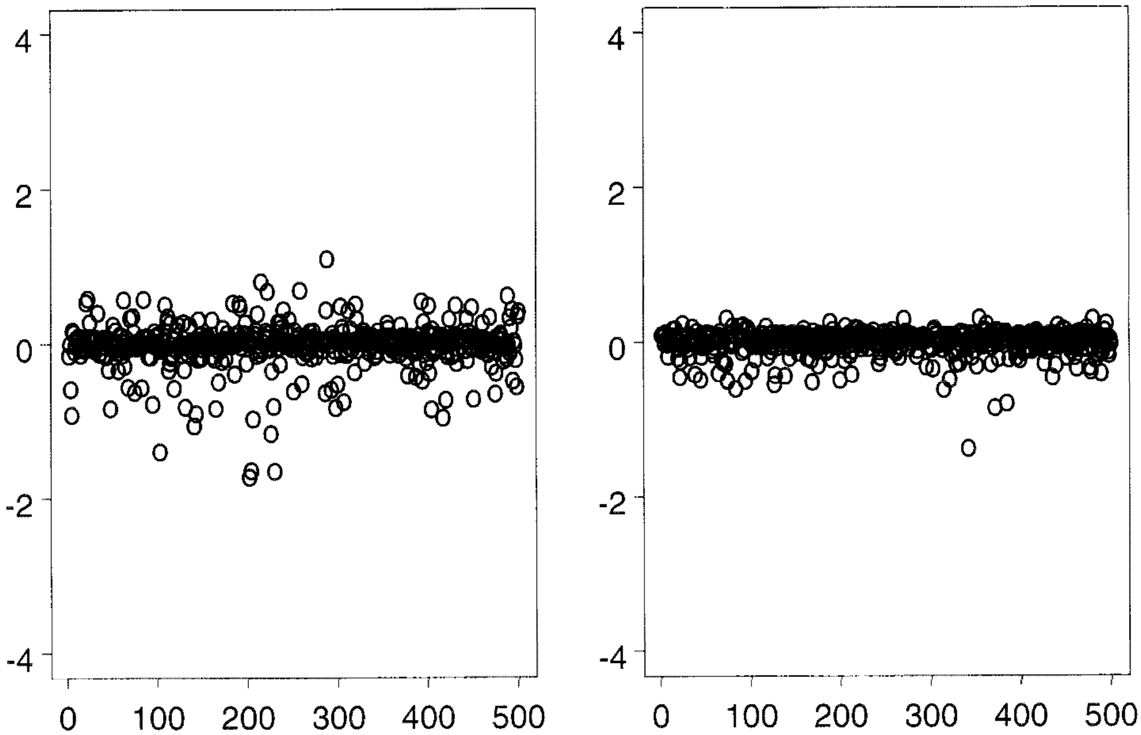

Figure 1 – Asymptotically pivotal univariate mean: behavior of $T_n^* - W_n^*$ From m = 500 resamples of sizes $n = 10$ (left panel) and $n = 30$ (right panel). Original samples of size n drawn from the Normal $N(0, 1)$ distribution

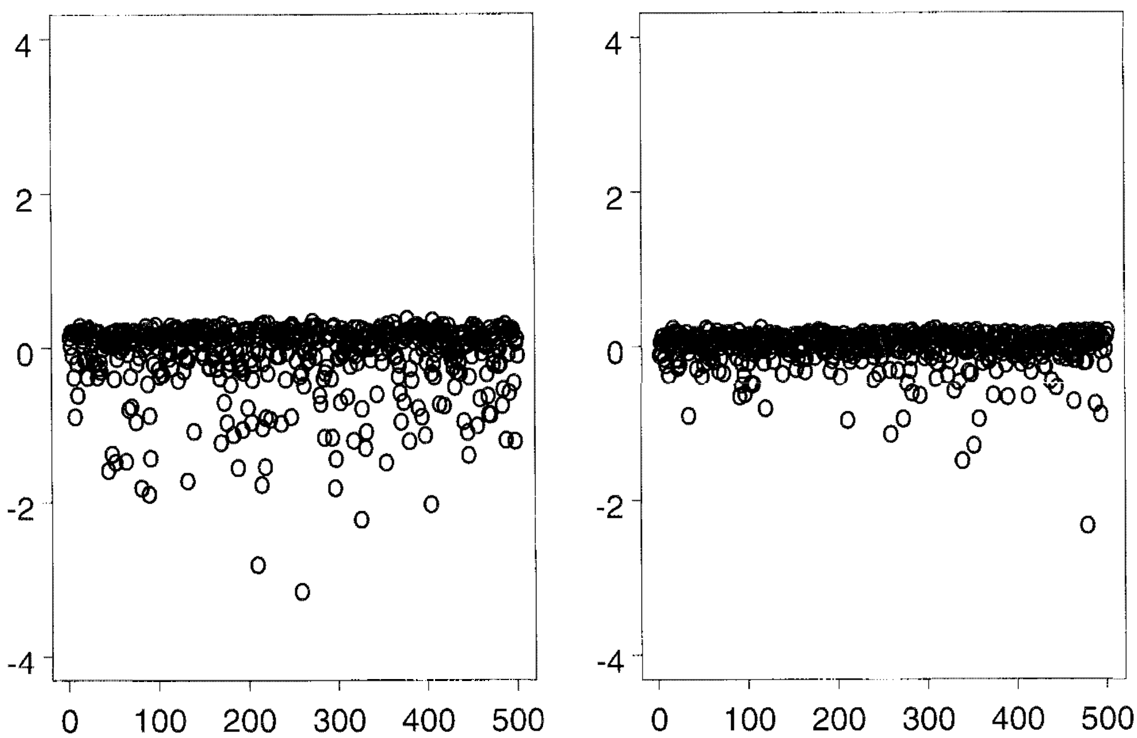

Figure 2 – Asymptotically pivotal univariate variance: behavior of $T_n^* - W_n^*$ from m = 500 resamples of sizes $n = 25$ (left panel) and $n = 100$ (right panel). Original samples of size n drawn from the Normal $N(0, 1)$ distribution.

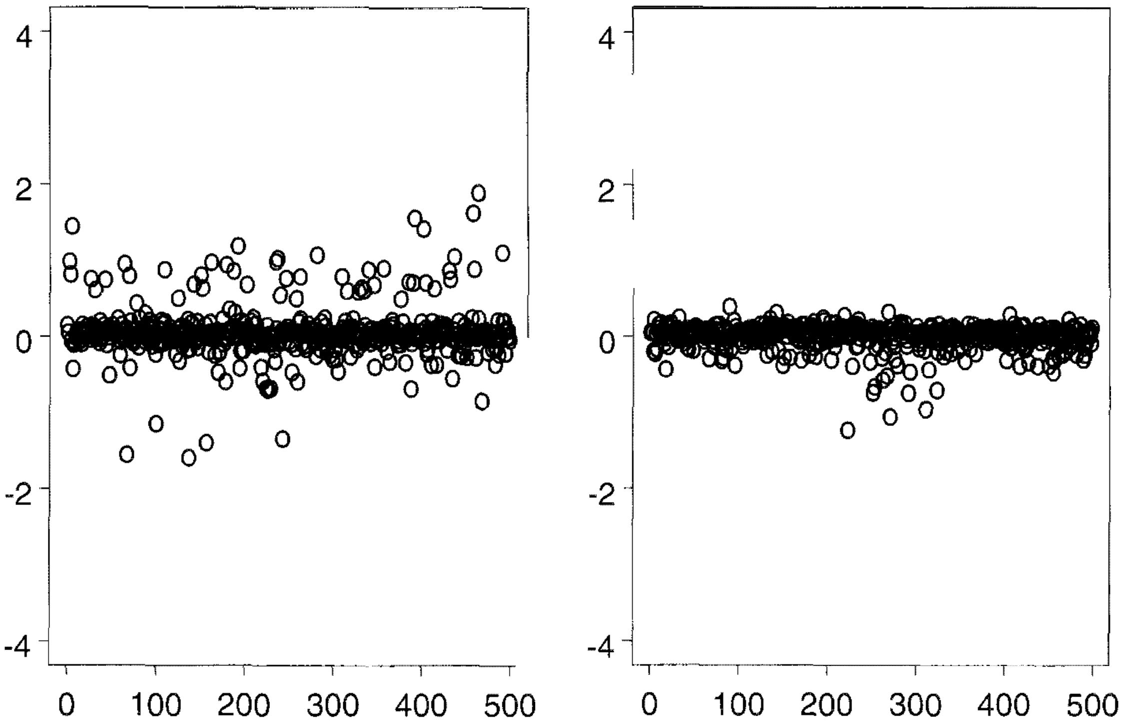

Figure 3 - Asymptotically pivotal ratio of means: behavior of $T_n^* - W_n^*$ from m = 500 resamples of sizes n = 10 (left panel) and n = 30 (right panel). Original samples of size n from the bivariate Folded-normal distribution with correlation coefficient p = 0.5 between marginals.

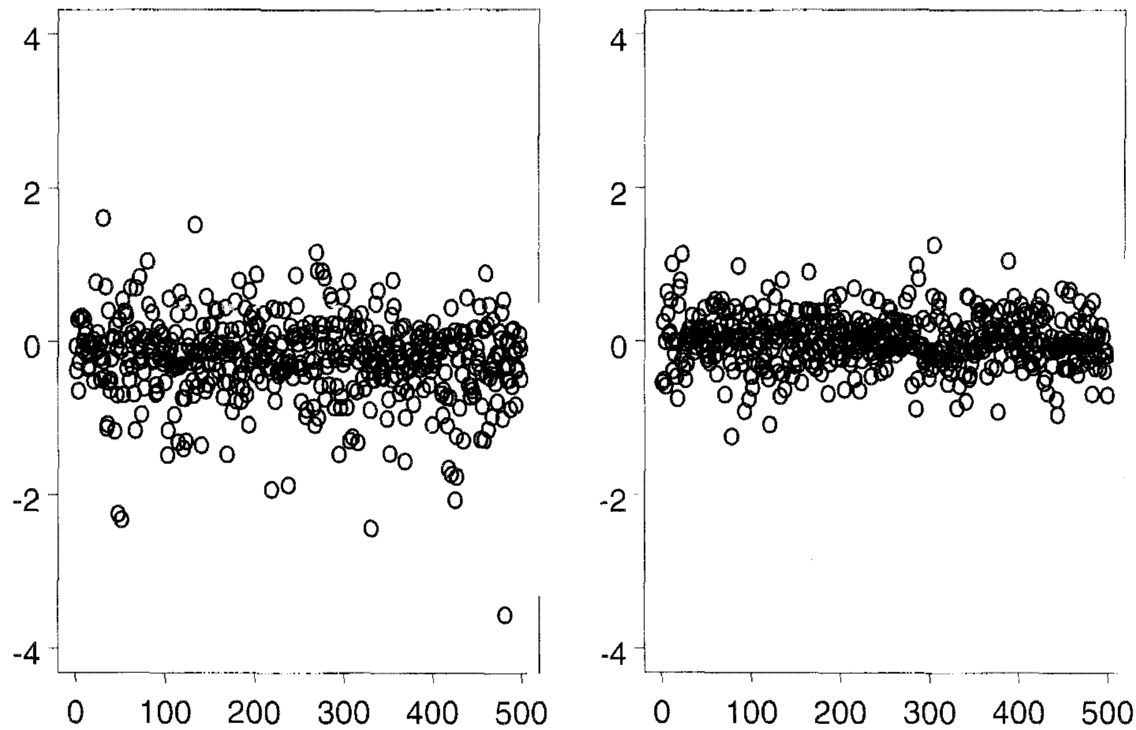

Figure 3 - Asymptotically pivotal correlation coefficient: behavior of $T_n^* - W_n^*$ from $m = 500$ resamples of sizes n = 15 (left panel) and $n = 50$ (right panel). Original samples of size n from the bivariate Lognormal distribution with correlation coefficient p = 0.5775 between marginals.

whereas T_n^* is the asymptotically pivotal version of the univariate mean and ratio of means, respectively. Variability around zero of the difference $T_n^* - W_n^*$ is very small. Figures 2 and 4 show a slower convergence of W_n^* toward T_n^*, as n increases, whereas T_n^* is the asymptotically pivotal version of the univariate variance and correlation coefficient, respectively.

5.2. *Empirical coverage probabilities*

Here, we report on a simulation experiment conducted to investigate the coverage properties of the newly-approximated asymptotic bootstrap confidence intervals $\hat{J}_{EC}$ (Edgeworth-corrected), $\hat{J}_{TB}$!transformation-based), $\hat{J}_{BC}$ (bias-corrected), $\hat{J}_{BCa}$ (bias-corrected accelerated) and J_{SH} (short) studied in section 4. Empirical coverage probabilities are estimated from 2000 independent original samples with different sample sizes n, drawn from various population distributions. It should be remembered that the new linear approximation method does not need Monte Carlo simulation. In this sense the bootstrap is exact (without simulation error), being automatically based on an ideal (infinite) number of resamples. Empirical coverage probability of percentile-t bootstrap confidence intervals $\hat{J}$ is obtained by performing 2000 repetitions of a single simulation round, which consists of $m = 1300$ independent resamples. We always consider a nominal coverage $a = 0.9$.

Table 1 reports on the coverage probability of the bootstrap confidence interval for the univariate mean (example 1). We simulate original samples from population distributions with various degrees of skewness and kurtosis. We consider the Normal $N(0, 1)$ (symmetric distribution with no kurtosis) the Chi-square χ_4^2 distribution (asymmetric distribution with no kurtosis), the Lognormal $\exp\{N(0, 1)\}$ distribution (asymmetric distribution with kurtosis). Note that asymptotic bootstrap confidence intervals $\hat{J}_{EC}$, $\hat{J}_{TB}$, $\hat{J}_{BC}$, $\hat{J}_{BCa}$ and $\hat{J}_{SH}$ produce similar and satisfactory results. The performance confirms the good behavior of W_n^* as linear approximation to T_n^*, shown in figure 1. Table 1 also shows the coverage probability of bootstrap confidence intervals for the univariate variance (example 2). Both classical asymptotic confidence intervals and bootstrap confidence intervals typically have a small coverage probability in this case, whereas the population distribution is not Normal. Observe in this sense the clear variability of $T_n^* - W_n^*$ around 0 shown in figure 2 (left panel). In particular, notice how confidence interval J_{TB} gives the worst performance of all.

Table 2 illustrates examples 3 and 4 introduced above. We consider different population bivariate distributions with different values for the correlation coefficient p between their marginals. We generate the bivariate Folded-normal $N(0, 1)$ distribution (with correlation coefficient $p = 0.5$ between the marginals), the bivariate Lognormal distribution (with correlations $p = 0.3775$ and $p = 0$) and the bivariate Normal distribution (with correlations $p = 0.5$ and $p = 0$). Table 3 shows the coverage probability of bootstrap confidence intervals for the ratio of means (example 3). Confidence interval $\hat{J}_{BCa}$ may apparently be preferred to intervals $\hat{J}_{EC}$, $\hat{J}_{TB}$, $\hat{J}_{BC}$ and $\hat{J}_{SH}$. In any case, confidence intervals $\hat{J}_{EC}$, $\hat{J}_{TB}$, $\hat{J}_{BC}$ and $\hat{J}_{SH}$

TABLE 1

Empirical coverage probabilities of two-sided bootstrap confidence intervals (with nominal coverage $\alpha = 0.9$) for the univariate mean and variance

			$\hat{J}$	$\hat{J}_{EC}$	$\hat{J}_{TB}$	$\hat{J}_{BC}$	$\hat{J}_{BCa}$	$\hat{J}_{SH}$
mean	(I)	$n = 10$	0.8150	0.8185	0.8175	0.8320	0.8360	0.8180
		$n = 30$	0.8820	0.8790	0.8765	0.8855	0.8845	0.8770
	(II)	$n = 10$	0.7850	0.7955	0.7940	0.8165	0.8135	0.7935
		$n = 30$	0.8500	0.8515	0.8545	0.8710	0.8665	0.8555
	(III)	$n = 10$	0.7400	0.7510	0.7395	0.7700	0.7690	0.7420
		$n = 30$	0.8050	0.8080	0.7895	0.8240	0.8280	0.7970
variance	(I)	$n = 25$	0.7250	0.7780	0.6520	0.8065	0.8050	0.7705
		$n = 100$	0.8450	0.8840	0.7265	0.8915	0.8885	0.8740
	(II)	$n = 25$	0.6050	0.6915	0.5835	0.7140	0.7240	0.6800
		$n = 100$	0.7550	0.7900	0.6575	0.8265	0.8240	0.7845
	(III)	$n = 25$	0.4010	0.4625	0.3910	0.5135	0.5015	0.4605
		$n = 100$	0.5600	0.5895	0.4715	0.6510	0.6390	0.5765

TABLE 2

Empirical coverage probabilities of two-sided bootstrap confidence intervals (with nominal coverage $\alpha = 0.9$) for the ratio of means and correlation coefficient

			$\hat{J}$	$\hat{J}_{EC}$	$\hat{J}_{TB}$	$\hat{J}_{BC}$	$\hat{J}_{BCa}$	$\hat{J}_{SH}$
ratio of means	(I)	$n = 10$	0.8200	0.8445	0.8390	0.8565	0.8620	0.8445
		$n = 30$	0.8650	0.8795	0.8800	0.8905	0.8900	0.8790
	(II)	$n = 10$	0.7300	0.7550	0.7835	0.7420	0.7870	0.7545
		$n = 30$	0.8200	0.8380	0.8350	0.8310	0.8655	0.8260
	(III)	$n = 10$	0.7700	0.7970	0.7790	0.7820	0.8105	0.7790
		$n = 30$	0.8250	0.8265	0.8130	0.8165	0.8505	0.8120
correlation coefficient	(IV)	$n = 15$	0.8200	0.8330	0.7500	0.9130	0.8360	0.8035
		$n = 50$	0.8550	0.8950	0.8555	0.9405	0.8615	0.8805
	(V)	$n = 15$	0.8350	0.7895	0.8000	0.8545	0.8325	0.7940
		$n = 50$	0.8640	0.8625	0.8780	0.8835	0.8940	0.8725
	(II)	$n = 15$	0.7550	0.7595	0.7570	0.8755	0.8815	0.7580
		$n = 50$	0.8050	0.8090	0.8095	0.8965	0.9070	0.8100

always have a good coverage probability with this example (cf. also figure 3). Table 2 also shows the coverage probability of bootstrap confidence interval: for the correlation coefficient (example 4). Bias-corrected confidence intervals $\hat{J}_{BC}$ and $\hat{J}_{BCa}$ outperform intervals $\hat{J}_{EC}$, $\hat{J}_{TB}$ and $\hat{J}_{SH}$. However, interval $\hat{J}_{BC}$ tends to have an empirical coverage larger than the desired nominal level α, with the bivariate Normal distribution and $\rho = 0.5$. The correlation coefficient example proves to be fairly difficult also for the linear approximation W_n^*. The good performance in terms of coverage error is possibly due to the small range of values (both negative and positive) of the basic smooth function of means, which makes W_n^* more efficient in estimating cumulants of T_n^*. Compare figure 4, for instance, with figure 2, which explains the behavior of W_n^* in approximating the univariate variance. In that case, we have a similar variability, but we have poorer coverage of nominal level α in most situations.

As can be observed from the present simulation study, at least in terms of coverage error, the linear approximation method can be regarded as a good way of simplifying the use of bootstrap confidence intervals $\hat{J}_{EC}$, $\hat{J}_{TB}$, $\hat{J}_{BC}$, $\hat{J}_{BCa}$ and $\hat{J}_{SH}$. Such complicated bootstrap confidence intervals are equivalent to asymptotic confidence intervals, simply determined by the bootstrap cumulants of a mean of smooth functions.

Software written in S-PLUS (cf. Becker *et* al., 1988) is available from the Author on request.

6. CONCLUSION

Empirical exponential families as formulated by DiCiccio and Efron (1992) serve as an elegant setting for studying application of the linear approximation W, given by (12), under the mean-value parametrization. DiCiccio and Efron (1992) introduce these families of distributions to obtain approximations of bootstrap cumulants of smooth functions of means by numerical differentiation of the cumulant generating function. See also DiCiccio and Efron (1996). Application of the bootstrap version W_n^* does not require this specific procedure to be applied in empirical exponential families and admits a more direct expression (W_n^* being a mean) for its cumulant generating function.

Asymptotic iterated bootstrap confidence intervals in Lee and Young (1995) which are based on the symbolic computation of bootstrap cumulants instead of bootstrap resampling through Monte Carlo simulation, can alternatively be approximated by linear approximations W, or Q, given by (12) and (26) as well. However, a slight increase in coverage error may result with the univariate mean, variance and the correlation coefficient examples discussed in sections 3 and 5.

Linear approximations W, and Q, both require a preliminary standardization for location $\mu = E(\mathbf{X}_1)$ of original sample observations X,, i = 1, ..., n. Recall definitions (12) and (26). A standardization for location and scale can be of interest, if one accepts more difficult definitions for smooth functions T_n and Q, and a sensible increase in computational burden. Notice that every component $\mathbf{X}_1^{(s_1)}$ in X, must be transformed into $(\mathbf{X}_1^{(s_1)} - \mu^{(s_1)})/(\xi^{(s_1)})^{1/2}$, where $\xi^{(s_1)} = E((\mathbf{X}_1^{(s_1)} - \mu^{(s_1)})^2)$, $s_1 = 1, ..., e$. On the other hand, the order of error in approximations W, and Q, does not vary, as may be deduced from *(33)*in Appendix 7.1.

The starting point of linear approximations W, and Q, given by (12) and (26) is analytical, but W, and Q, do not represent an attempt to completely remove Monte Carlo simulation from the bootstrap. They behave as means and so they do not require simulation in order to estimate their cumulants. A comparison with more classical bootstrap approaches (based exclusively on Monte Carlo simulation), such as in Oldford (1985), ought to confirm that W, and Q, are computationally efficient and easy to run.

Keeping in mind that W, makes important bootstrap confidence intervals feasible (beyond the univariate mean case), W, may be further improved (in terms of order of error) by Monte Carlo simulation. Let us write $T_n = W_n + R_n$, where T_n is given by *(11)* and $R_n = T_n - W_n = O_p(n^{-3/2})$. For a fixed n, we may estimate the distribution of R_n by Monte Carlo simulation of m independent bootstrap resamples, where $m = l(n)$, for some optimal function I, and then recover variability of R_n in the approximation of bootstrap cumulants. The same conclusions hold for D_n and Z_n given by (24), with $R_n = O_p(n^{-2})$, and Q_n. This idea may be related to recent work by DiCiccio *et al.* (1997), Lee and Young *(1997)* and Hall *et al. (1999)*.

Alternative simulation-based bootstrap confidence intervals may be constructed by substituting analytical approximations $\tilde{v}_\beta$ and $\tilde{u}_\beta$ in asymptotic intervals (23), *(27)* and (28), with bootstrap quantiles $\hat{v}_\beta$ and $\hat{u}_\beta$ estimated from a sufficiently large number m of bootstrap resamples.

Transformations of the non-pivotal quantity D_n able to produce some desired effect, can easily be incorporated into the linear approximation Z_n given by *(24)*. For instance, the variance-stabilizing bootstrap transformations of Tibshirani *(1988)* $\tau(D_n)$ produce the linear approximation $Z_n = \sum_{i=1}^{n} \tau(g(n^{-1}\mathbf{Y}_i + \nu))$. The use of such transformations in bootstrap confidence intervals is discussed, among others, in Canty, Davison and Hinkley (1996).

It is sometimes useful to enlarge the domain of a smooth function of means or consider irregular situations for its application. The linear approximations W_n and Q_n given by *(12)* and *(26)* is naturally flexible in most of these cases. Recent results in Kano (1999), for instance, could appropriately be combined with W_n in the present bootstrap context.

7. APPENDIX

7.1. Proofs of (12) and (13)

Using the Taylor expansion of $T_n = n^{1/2}A(\bar{\mathbf{y}} + \mu)$ and $A_i = A(n^{-1}\mathbf{Y}_i + \mu)$ around $\mu = E(\mathbf{X}_1)$, where mean $\bar{\mathbf{y}} = n^{-1}\sum_{i=1}^{n}\mathbf{Y}_i$ and $\mathbf{Y}_i = \mathbf{X}_i - \mu$, we may deduce

$$T_n = n^{1/2}\sum_{j=1}^{3}\frac{1}{j!}\sum_{s_1=1}^{e}\cdots\sum_{s_j=1}^{e} a_{s_1\ldots s_j}(\mu)\sum_{i=1}^{n}(n^{-1}\mathbf{Y}_i)^{(s_1)}\cdots\sum_{i=1}^{n}(n^{-1}\mathbf{Y}_i)^{(s_j)} + O_p(n^{-3/2}), \tag{30}$$

$$A_i = \sum_{j=1}^{3}\frac{1}{j!}\sum_{s_1=1}^{e}\cdots\sum_{s_j=1}^{e} a_{s_1\ldots s_j}(\mu)(n^{-1}\mathbf{Y}_i)^{(s_1)}\cdots(n^{-1}\mathbf{Y}_i)^{(s_j)} + O_p(n^{-3}). \tag{31}$$

Keep in mind that $\mu_{s_1\ldots s_j} = E\{(\mathbf{X}_1 - \mu)^{(s_1)}\ldots(\mathbf{X}_1 - \mu)^{(s_j)}\}$ and $\mu = 0$. For $j = 1, 2, 3$,

$$\mu_{s_1\ldots s_j} = E\left\{\sum_{i=1}^{n}(n^{-1}\mathbf{Y}_i)^{(s_1)}\cdots\sum_{i=1}^{n}(n^{-1}\mathbf{Y}_i)^{(s_j)}\right\} = n\cdot E\left\{(n^{-1}\mathbf{Y}_1)^{(s_1)}\cdots(n^{-1}\mathbf{Y}_1)^{(s_j)}\right\}. \tag{32}$$

Cross products $(n^{-1}\mathbf{Y}_{i_1})^{(s_{j1})}(n^{-1}\mathbf{Y}_{i_2})^{(s_{j2})}$ and $(n^{-1}\mathbf{Y}_{i_1})^{(s_{j1})}((n^{-1}\mathbf{Y}_{i_2})^{(s_{j2})})^2$, in (32), with $i_1 \neq i_2$, $i_1, i_2 = 1, \ldots, n$, are zero under expectation. If $j \geq 4$, the j-th term in Taylor expansion of $n^{1/2}A(\bar{\mathbf{y}} + \text{p})$ around p is not equal (under expectation) to the j-th term in Taylor expansion of $A(n^{-1}\mathbf{Y}_i + \text{p})$ around p. In particular, it can be observed that

$$\mathrm{E}(T_n - W_n) = n^{-3/2}\frac{1}{4!}\sum_{s_1=1}^{e}\sum_{s_2=1}^{e}\sum_{s_3=1}^{e}\sum_{s_4=1}^{e} a_{s_1s_2s_3s_4}(\mu)\,\mu_{s_1s_2}\mu_{s_3s_4} + O(n^{-2}), \tag{33}$$

and E{(T, - $W_n)^2\} = O(n^{-3})$. Thus, approximation W, differs from T_n by an error of order $n^{-3/2}$, T_n = W, + $O_p(n^{-3/2})$.

7.2. Approximation W_n satisfies (7) with the same constant5 K,,, κ_{22}, κ_{31} and κ_{41} *as* T_n

Recall that K,, = 0 and $\kappa_{21} = 1$. Focussing on κ_{12} and T_n, we may conclude that

$$\begin{aligned}\mathrm{E}(T_n) &= n^{1/2}\sum_{s_1=1}^{e}\sum_{s_2=1}^{e} a_{s_1}(\mu)\,\mu_{s_1} + n^{-1/2}\frac{1}{2!}\sum_{s_1=1}^{e}\sum_{s_2=1}^{e} a_{s_1s_2}(\mu)\,\mu_{s_1s_2} + O(n^{-1})\\ &= \kappa_{11} + n^{-1/2}\kappa_{12} + O(n^{-1}),\end{aligned}$$

where $\kappa_{11} = 0$, because $\mu_{s_1} = 0$. Considering W_n, we observe that

$$\begin{aligned}\mathrm{E}(W_n) &= \mathrm{E}\left\{n^{1/2}\sum_{i=1}^{n}\left(\sum_{s_1=1}^{e} a_{s_1}(\mu)(n^{-1}\mathbf{Y}_i)^{(s_1)}\right.\right.\\ &\quad\left.\left.+\frac{1}{2!}\sum_{s_1=1}^{e}\sum_{s_2=1}^{e} a_{s_1s_2}(\mu)(n^{-1}\mathbf{Y}_i)^{(s_1)}(n^{-1}\mathbf{Y}_i)^{(s_2)}\right)\right\} + O(n^{-1})\\ &= n^{1/2}\{n\cdot(n^{-1}\kappa_{11} + n^{-2}\kappa_{12})\} + O(n^{-1}) = n^{-1/2}\kappa_{12} + O(n^{-1})\end{aligned}$$

From (7), it is known that $\mathrm{E}(T_n) - \{\mathrm{E}(T_n)\}^2 = \kappa_{21} + n^{-1}\kappa_{22} + O(n^{-2})$, where $\kappa_{21} = 1$, because (6) holds. Considering W_n, we observe that

$$\begin{aligned}&\mathrm{E}(W_n^2) - \{\mathrm{E}(W_n)\}^2\\ &= \mathrm{E}\left\{n\sum_{i=1}^{n}\sum_{s_1=1}^{e}\sum_{s_2=1}^{e} a_{s_1}(\mu)\,a_{s_2}(\mu)(n^{-1}\mathbf{Y}_i)^{(s_1)}(n^{-1}\mathbf{Y}_i)^{(s_2)}\right\} + n^{-1}\kappa_{22} + O(n^{-2})\\ &= n\cdot(n\cdot n^{-2}\kappa_{21}) + n^{-1}\kappa_{22} + O(n^{-2}) = 1 + n^{-1}\kappa_{22} + O(n^{-2}).\end{aligned}$$

Following lengthy algebraic calculations, it is similarly seen that W_n has cumulants with the same constants κ_{22}, κ_{31} and K,, as T_n.

7.3. Proofs *of (20)* and *(21)*

Expansion of j-th bootstrap cumulant λ_j of B, = nA_1^* can be deduced from expansion (7), by taking the bootstrap version $A(n^{-1}\mathbf{Y}_1^* + \%)$ of $A(n^{-1}\mathbf{Y}_1 + \text{p})$ in place of the function $n^{1/2}A(\bar{\mathbf{y}} + \mu) = n^{1/2}A(\bar{\mathbf{x}})$.

Recall that $\hat{\kappa}_j$ and $\hat{\lambda}_j$ are the j-th cumulants of W_n^* and B , $= nA_1^*$, respectively. Since B_1^* is embedded in mean

$$W_n^* = n^{1/2}\bar{b}^* ,$$

we observe that $\hat{\kappa}_j = n^{-(j-2)/2}\hat{\lambda}_j$, where $\hat{\lambda}_j$ confirms (20), with $\hat{\lambda}_{11} = 0$ and $\hat{\lambda}_{21} = 1$. It follows that $\hat{\kappa}_1 = n^{1/2}\hat{\lambda}_1 = n^{-1/2}\hat{\lambda}_{12} + O_p(n^{-3/2})$, $\hat{\kappa}_2 - 1 = \hat{\lambda}_2 - 1 = n^{-1}\hat{\lambda}_{22} + O_p(n^{-2})$, $\hat{\kappa}_3 = n^{-1/2}\hat{\lambda}_3 = n^{-1/2}\hat{\lambda}_{31} + O_p(n^{-3/2})$, $\hat{\kappa}_4 = n^{-1}\hat{\lambda}_4 = n^{-1}\hat{\lambda}_{41} + O_p(n^{-2})$. We define by $\hat{\kappa}_{11}$, $\hat{\kappa}_{12}$, $\hat{\kappa}_{21}$, $\hat{\kappa}_{22}$, $\hat{\kappa}_{31}$ and $\hat{\kappa}_{41}$ the bootstrap counterparts of constants κ_{11}, κ_{12}, κ_{21}, κ_{22}, κ_{31} and κ_{41} in (7), respectively. Observe that $\hat{\kappa}_{11} = \hat{\lambda}_{11} = 0$, $\hat{\kappa}_{12} = \hat{\lambda}_{12}$, $\hat{\kappa}_{21} = \hat{\lambda}_{21} = 1$, $\hat{\kappa}_{22} = \hat{\lambda}_{22}$, $\hat{\kappa}_{31} = \hat{\lambda}_{31}$ and $\hat{\kappa}_{41} = \hat{\lambda}_{41}$. Constants κ_{12}, κ_{22}, κ_{31} and κ_{41} can thus be approximated by $n\hat{\lambda}_1$, $n(\hat{\lambda}_2 - 1)$, $\hat{\lambda}_3$ and $\hat{\lambda}_4$, respectively, with an error of order $O_p(n^{-1})$, obtaining *(21)*.

7.4. Coverage errors of confidence intervals (22), (23), (27), (28) and (29)

Recalling T_n given by (11), we let $S_n = T_n + n^{-1}\Delta_n$, where $\Delta_n = n^{1/2}\{\hat{q}_1(z_\beta) - q_1(z_\beta)\}$ and $\beta = (1+\alpha)/2$. We let $a_\beta = \mathrm{E}(T_n \Delta_n) + O(n^{-1})$. Notice that polynomials $q_3(u)$ and $\hat{q}_3(u)$ in Edgeworth expansions (8) and *(21)* are even functions of $u \in R'$. Proposition 3.1 in Hall (1992a) shows that two-sided bootstrap percentile-t confidence intervals J have a coverage error with expansion $P(\theta \in J) - \alpha = -2n^{-1}a_\beta z_\beta \phi(z_\beta) + O(n^{-2})$. A parallel expansion holds for $S_n = Q_n + n^{-1}\Delta_n$, where Q_n is given by *(26)* and $\Delta_n = n^{1/2}\{\hat{p}_1(z_\beta) - p_1(z_\beta)\}$. Considering basic definitions of two-sided bootstrap confidence intervals described in section 4 (Hall, 1992a, chapter 3; and Hall, 1992b), coverage errors of asymptotic confidence intervals $\hat{J}_{EC}$, $\hat{J}_{TB}$, $\hat{J}_{BC}$, $\hat{J}_{BCa}$ and $\hat{J}_{SH}$ may similarly be obtained. In particular, observe that polynomials $q_1(u)$ and $q_2(u)$, and $p_1(u)$ and $p_2(u)$, $u \in \mathbb{R}^1$, can be approximated through bootstrap linear approximations W_n^* and Q_n^* with an error of order $O(n^{-1})$ or smaller.

Dipartimento di Scienze statistiche "Paolo Fortunati" ANDREA PALLINI
Università di Bologna

REFERENCES

R.A. BECKER, J.M. CHAMBERS, A.R. WILKS (1988), *The new S language,* Wadsworth and Brooks/ Cole, Pacific Grove, California.

R.N. BHATTACHARYA, J.K. GHOSH (1978), *On the validity of the formal Edgeworth expansion,* "Annals of Statistics", 6, pp. 434-451.

A.J. CANTY, A.C. DAVISON, D.V. HINKLEY (1996), *Discussion of the paper by T.J. DiCiccio and B. Efron,* "Statistical Science", *3,* pp. 214-219.

P.J. DAVIS, P. RABINOWITZ (1984), *Methods of numerical integration,* 2nd *Ed.,* Wiley and Sons, New York.

A.C. DAVISON, D.V. HINKLEY (1997), *Bootstvap methods and their application,* Cambridge University Press, Cambridge.

T.J. DICICCIO, R. EFRON (1992), *More accurate confidence intervals in exponential families*, "Biometrika", 79, pp. 231-245.

T.J. DICICCIO, B. EFRON (1996), *Bootstrap confidence intervals (with discussion)*, "Statistical Science", 3, pp. 189-228.

T.J. DICICCIO, R.E. KASS, A. RAFTERY, L. WASSERMAN (1997), *Computing Bayes factors by combining simulation and asymptotic approximations*, "Journal of the American Statistical Association", 92, pp. 903-915.

T.J. DICICCIO, J.P. ROMANO (1988), *A review of bootstrap confidence intervals (with discussion)*, "Journal of the Royal Statistical Society R", 50, pp. 338-370.

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

B. EFRON (1987), *Better bootstrap confidence intervals (with discussion)*, "Journal of the American Statistical Association", 82, pp. 171-200.

B. EFRON, R.J. TIBSHIRANI (1993), *An introduction to the bootstrap*, Chapman and Hall, London.

W. FLEMING (1977), *Functions of several variables*, Springer-Verlag, New York.

P. HALL (1986), *On the number of bootstrap simulations required to construct a confidence interval*, "Annals of Statistics", 14, pp. 1453-1462.

P. HALL (1988), *Theoretical comparisons of bootstrap confidence intervals (with discussion)*, "Annals of Statistics", 16, pp. 927-985.

P. HALL (1992a), *The bootstrap and Edgeworth expansion*, Springer-Verlag, New York.

P. HALL (1992b), *On the removal of skewness by transformation*, "Journal of the Royal Statistical Society B", 54, pp. 221-228.

P, HALL, M.A. MARTIN, S. SUN (1999), *Monte Carlo approximation to Edgeworth expansions*, "Canadian Journal of Statistics", 27, pp. 579-584.

Y. KANO (1999), *Delta method approach in a certain irregular condition*, "Communications in Statistics Theory and Methods", 28, pp. 789-807.

S. KONISHI (1991), *Normalizing transformations and bootstrap confidence intervals*, "Annals of Statistics", 19, pp. 2209-2225.

S.M.S. LEE, G.A. YOUNG (1995), *Asymptotic iterated bootstrap confidence intervals*, "Annals of Statistics", 23, pp. 1301-1330.

S.M.S. LEE, G.A. YOUNG (1997), *Estimation of the distribution function of a standardized statistic*, "Journal of the Royal Statistical Society B", 59, pp. 383-400.

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

P. MCCULLAGH (1987), *Tensor methods in statistics*, Chapman and Hall, London.

R.W. OLDFORD (1985), *Bootstrapping by Monte Carlo versus approximating the estimator and bootstrapping exactly: cost and performance*, "Communications in Statistics - Simulation and Computation", 14, pp. 395-424.

V.V. PETROV (1995), *Limit theorems of probability theory*, Clarendon Press, Oxford.

C.R. RAO (1973), *Linear statistical inference and its application*, 2nd Ed., Wiley and Sons, New York.

R.J. SERFLING (1980), *Approximations theorems in mathematical statistics*, Wiley and Sons, New York.

J. SHAO, D. TU (1995), *The Jackknife and the bootstrap*, Springer-Verlag, New York.

R. TIBSHIRANI (1988), *Variance stabilization and the bootstrap*, "Biometrika", 75, pp. 433-444.

RIASSUNTO

Su un metodo lineare negli intervalli di confidenza boostrap

Viene proposto un metodo di approssimazione lineare per la costruzione di intervalli di confidenza bootstrap asintotici. Vengono approssimate quantità asintoticamente pivotali e non pivotali, che sono funzioni regolari di medie di n variabili casuali indipendenti ed identicamente distribuite, utilizzando una somma di n funzioni regolari indipendenti della stessa forma analitica. Gli errori sono rispettivamente di ordine $O_p(n^{-3/2})$ e $O_p(n^{-2})$. Il metodo lineare consente un'approssimazione diretta dei cumulanti bootstrap, considerando l'insieme di n funzioni regolari indipendenti come un campione originale da ricampionare con ripetizione.

SUMMARY

On a linear method in bootstrap confidence intervals

A linear method for the construction of asymptotic bootstrap confidence intervals is proposed. We approximate asymptotically pivotal and non-pivotal quantities, which are smooth functions of means of n independent and identically distributed random variables, by using a sum of n independent smooth functions of the same analytical form. Errors are of order $O_p(n^{-3/2})$ and $O_p(n^{-2})$, respectively. The linear method allows a straightforward approximation of bootstrap cumulants, by considering the set of n independent smooth functions as an original random sample to be resampled with replacement.