The Annals of Statistics
1994, Vol. 22, No. 3, 1115–1132

SADDLEPOINT APPROXIMATIONS FOR MARGINAL AND CONDITIONAL PROBABILITIES OF TRANSFORMED VARIABLES

BY BINGYI JING[1] AND JOHN ROBINSON

University of Sydney

The Lugannani and Rice formula for tail areas in the univariate case has recently been extended to tail areas for marginal distributions and for conditional distributions in certain multivariate settings. However, the results on relative order of errors are given only formally or under the strong continuity assumptions necessary to obtain density approximations. This paper attempts to give a unified treatment of these results for smooth transformations of multivariate means under weaker conditions appropriate to indirect Edgeworth approximations for probabilities.

1. Introduction. Improvements to tail area approximations for sums of random vectors by the normal distribution may be made by using Edgeworth or saddlepoint approximations. The latter are frequently to be preferred because their errors are relative, so the approximations are accurate for both large and small tail probabilities. Generalisations of the Edgeworth expansions to smooth functions of means were derived rigorously by Bhattacharya and Ghosh (1978) and have been widely used, in particular, recently, in studies on the accuracy of bootstrap methods [e.g., in Hall (1988)]. It is the purpose of this paper to obtain saddlepoint approximations for marginal and conditional distributions of statistics which can be written as smooth functions of means and to prove that relative errors of order $n^{-3/2}$ exist under weak conditions. The numerical accuracy of tail areas using these methods would be expected to be considerably greater in the tails than that of Edgeworth methods.

There are essentially three methods of calculating saddlepoint tail area approximations in common use now: the indirect Edgeworth expansion; the numerically integrated saddlepoint density, usually renormalized for additional accuracy; and the formula of Lugannani and Rice (1980). Daniels (1987) compared the Lugannani and Rice formula and the indirect Edgeworth expansion with relative error of order $n^{-3/2}$ in the case of the mean, for exponential and inverse normal distributions, cases where the saddlepoint approximations to the density of the mean are "exact," and found that the indirect Edgeworth expansion performs slightly better than the Lugannani and Rice formula. However, the Lugannani and Rice formula is much simpler and easier to use than the indirect Edgeworth expansion for relative error of order $n^{-3/2}$.

Received January 1992; revised January 1994.

[1]Part of this work was completed when the author was at the Centre for Mathematics and its Applications, Australian National University.

AMS 1991 *subject classifications.* Primary 62E20; secondary 60F05.

Key words and phrases. Smooth functions of means, tail area approximations, Lugannani–Rice approximations.

The indirect Edgeworth method was studied in Robinson, Höglund, Holst and Quine (1990), where conditions were given for validity of order terms for relative errors. The proofs of the other methods are based on strong conditions involving identically distributed variables for which densities exist. It is not easy to see how the smoothing methods necessary to obtain the validity results in Robinson, Höglund, Holst and Quine (1990) could be used directly to prove the validity of the Lugannani and Rice result without the strong continuity assumptions necessary for the density approximations. So, in Section 2, we define the indirect Edgeworth approximation and obtain a general result for the error for rectangles and half-planes. We then discuss saddlepoint density approximations under strong conditions and give an explicit discussion of renormalisation; finally, we show that when an indirect Edgeworth approximation exists with relative error of order $n^{-3/2}$, then the corresponding integral of the formal saddlepoint density is approximately equal to it with a relative error of the same order and is thus a valid approximation with the same relative error. In Section 3, we discuss smooth transformations of multivariate means and obtain formal density approximations for these. Then, in Sections 4 and 5, we give a unified treatment of one-dimensional marginal and conditional results giving tail approximations in a Lugannani and Rice form, thus generalizing the results in Daniels and Young (1991) and Skovgaard (1987), respectively. Finally, in Section 6, we give an example of the results using marginal approximations in nonparametric settings. It must be pointed out that the methods used for this are based on those in Daniels and Young (1991).

It should be noted that all the saddlepoint tail area approximation results obtained for the above mentioned three cases take the following form:

$$\left(1 - \Phi(\widehat{w}\sqrt{n}) - \frac{1}{\sqrt{n}}\phi(\widehat{w}\sqrt{n})\left(\frac{1}{\widehat{w}} - \frac{\psi(\widehat{w})}{\widehat{w}}\right)\right)\left(1 + \frac{1}{n}p_n + O(n^{-3/2})\right), \tag{1.1}$$

where $\widehat{w}$ and $\psi(\widehat{w})$ are defined differently at each occurrence. If we consider the approximations in the normal range, then $p_n = O(n^{-1/2})$, giving the saddlepoint approximations a relative error of order $n^{-3/2}$. However, for large deviations, the error is of order n^{-1}. An approximation which has the same asymptotic properties as this but lies in the range $(0, 1)$ was suggested by Barndorff-Nielsen (1986, 1991) and proved to be equivalent to the Lugannani and Rice formula by Jensen (1992). From Jensen (1992), if

$$w^* = \widehat{w} - \frac{1}{n\widehat{w}} \log \psi(\widehat{w}),$$

then we can show that (1.1) can be replaced by

$$\left(1 - \Phi(w^*\sqrt{n})\right)\left(1 + \frac{1}{n}p_n + O(n^{-3/2})\right).$$

2. Indirect Edgeworth and saddlepoint methods. Let $X_1, \ldots, X_n$ be independent (but not necessarily identically distributed) random vectors in $\Re^d$, and let $S = \sum_{i=1}^{n} X_i$ and $\overline{X} = S/n$. Without loss of generality we may take $ES = 0$. Then let

$$\kappa(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log E e^{\theta \cdot X_i} = \frac{1}{n} \log E e^{\theta \cdot S}.$$

Let $\Theta = \{\theta \colon \kappa(\theta) < \infty\}$ and write the mean vector and the covariance matrix as

$$m(\theta) = \kappa'(\theta) = \frac{\partial \kappa(\theta)}{\partial \theta},$$
$$V(\theta) = \kappa''(\theta) = \frac{\partial^2 \kappa(\theta)}{\partial \theta^2}.$$

Let μ be the probability measure of $\overline{X}$. Then

$$\begin{aligned} P(\overline{X} \in B) &= \int_B \mu(dy) \\ &= \exp\Bigl(n\bigl[\kappa(\theta) - \theta \cdot m\bigr]\Bigr) \int_{B-m} \exp(-n\theta \cdot y)\nu_\theta(dy) \\ &\equiv \exp\Bigl(n\bigl[\kappa(\theta) - \theta \cdot m\bigr]\Bigr) I(\theta, B), \end{aligned}$$

where ν_θ is an exponentially shifted probability measure.

2.1. *Indirect Edgeworth approximations.* Using an Edgeworth approximation for ν_θ gives the formal indirect Edgeworth approximation [Barndorff-Nielsen and Cox (1979), and Robinson, Höglund, Holst and Quine (1990)],

$$P(\overline{X} \in B) = \exp\Bigl(n\bigl[\kappa(\theta) - \theta \cdot m\bigr]\Bigr)\bigl(e_{s-3}(\theta, B) + \delta(\theta, B)\bigr) \tag{2.1}$$

where

$$\begin{aligned} e_{s-3}(\theta, B) &= \int_{B-m} \frac{\exp\bigl(-n\theta \cdot y - n\|y^*\|^2/2\bigr)}{\bigl(2\pi/n\bigr)^{d/2}\Delta^{1/2}} \left[1 + \sum_{i=1}^{s-3} n^{-i/2} Q_{i\theta}(\sqrt{n}y^*)\right] dy, \\ \delta(\theta, B) &= I(\theta, B) - e_{s-3}(\theta, B), \end{aligned} \tag{2.2}$$

where $y^* = V^{-1/2}y$, $\|y^*\|^2 = y^* \cdot y^*$, $\Delta = \det(V)$; $Q_{i\theta}$ are polynomials with coefficients based on cumulants calculated at θ; and $\delta(\theta, B)$ is the error term. The result (2.1) is exactly that of Barndorff-Nielsen and Cox (1979), who restrict attention to the two-term expansions for the one- and two-dimensional cases.

The order of the error $\delta(\theta, B)$ determines the worth of the approximation. It is difficult to give a simple expression for the error bound in general. By using the method of Skovgaard (1986), we can get $\delta(\theta, B) = O(n^{-(s-2)/2} \log n)$ for any Borel set B, but this can be improved. Notice that since $e_{s-3}(\theta, B)$ is $O(1)$, the order of the error $\delta(\theta, B)$ is a relative error.

In the case $\theta = 0$, von Bahr (1967) considered the restrictions that $X_1, \dots, X_n$ were i.i.d. random vectors, that $E\|X_i\|^s < \infty$ and that Cramér's condition was satisfied. He used the Edgeworth expansion up to $s - 2$ terms and showed that, for rectangles and bounded convex sets, the error $\delta(\theta, B) = o(n^{-(s-2)/2})$ [see von Bahr (1967), Theorem 2]. For unbounded convex sets, he imposed the additional condition that a number of moments of higher order exist, and he was able to show that $\delta(\theta, B) = O(n^{-(s-2)/2})$ [see von Bahr (1967), Theorem 2]. Bhattacharya and Ranga Rao [(1976), Section 20] extended von Bahr's results. They showed that if $X_1, \dots, X_n$ were independent (not necessarily identically distributed) random vectors, $E\|X_i\|^s < \infty$ and Cramér's condition (20.55) of Bhattacharya and Ranga Rao (1976) was satisfied, then for a subset of

$$\mathcal{A}_1(c\colon \Phi_{0,V}) = \left\{A \in \Re^d\colon \Phi_{0,V}\big((\partial A)^{\varepsilon}\big) \le c\varepsilon \quad \text{for all } \varepsilon > 0\right\}$$

which includes the class of all convex sets, the error $\delta(\theta, B) = o(n^{-(s-2)/2})$ [see (20.48) and (20.49) of Bhattacharya and Ranga Rao (1976)]. Note that $\Phi_{0,V}$ above stands for the multivariate normal distribution function with vector mean 0 and covariance matrix V, and ∂A means the boundary of A. However, Bhattacharya and Ranga Rao (1976) and von Bahr (1967) assumed Cramér's condition, which is rather restrictive. For example, it is not satisfied by the Wilcoxon rank sum statistic. Robinson, Höglund, Holst and Quine (1990) relaxed Cramér's condition to (S.4) of the conditions (S.1)–(S.4) below; however, their results are only useful for bounded sets. In this paper, we will extend those results of Robinson, Höglund, Holst and Quine (1990) to half-planes and rectangles, in which we are mainly interested here. The techniques involved are those of von Bahr (1967) and Robinson, Höglund, Holst and Quine (1990).

Consider the following conditions:

(S.1) There exists a convex, compact $K \subset \operatorname{int}(\Theta)$, for all n, such that $\operatorname{int}(K)$ is not empty.

(S.2) There exists a positive-definite matrix $V_0(\theta)$ such that, as $n \to \infty$,

$$V(\theta) \to V_0(\theta) \quad \text{uniformly for } \theta \in K.$$

(S.3) $\eta_s(\theta) \equiv E_\theta\|X\|^s < C$, for all $\theta \in K$.

(S.4) $q_\theta(n^{(s-2)/2}) < Cn^{-\lambda}$, for $\lambda = (d+1)(s-2)/2$, where

$$q_\theta(T) = \sup_\eta \left\{|\widehat{\nu}_\theta(\eta)|\colon cn^{1/2} \le n^{1/2}\|V(\theta)^{1/2}\eta\| < n^{(s-2)/2}\right\}.$$

THEOREM 1. *Assume conditions* (S.1)–(S.4) *hold. If B is a rectangle* (*or a half-plane*) *and $0 \notin B$, then we can choose θ such that, for every $u \in B - m$, $\theta \cdot u \ge 0$; and then the error in* (2.1) *satisfies*

$$|\delta(\theta, B)| \le Cn^{-(s-2)/2}.$$

The proof of Theorem 1 is given in the Appendix.

REMARK 2.1. If, in particular, for some $\widehat{x} \in \partial B$, the boundary of B, we can choose $\widehat{\theta} = \theta(\widehat{x})$, the solution of $m(\theta) = \widehat{x}$, such that, for every $u \in B - \widehat{x}$, $\widehat{\theta} \cdot u \geq 0$, then we get the most useful indirect Edgeworth approximation.

REMARK 2.2. The difficulty with approximation (2.1) lies in the complexity of the terms of first and second order and the difficulty of obtaining reasonable formulae for the integral in any but the simplest cases. Robinson, Höglund, Holst and Quine (1990) used this method to obtain results for a number of examples in one dimension and to obtain some approximations for conditional results.

REMARK 2.3. Here we consider only nonlattice variables because our interest is in transformed variables and our techniques would not give relative errors of $O(n^{-3/2})$. In special cases when the lattice variables are not transformed, results have been obtained by Daniels (1987), Skovgaard (1987), Robinson, Höglund, Holst and Quine (1990) and Jensen (1992).

REMARK 2.4. Durbin (1980) obtained a result clearly related to Theorem 1 but under rather different conditions. He considers statistics which are not necessarily sums of i.i.d. random variables but imposes conditions on the characteristic function and the cumulants.

2.2. *Saddlepoint density approximations.* Under sufficient continuity conditions such that smoothing is unnecessary, such as if $X_1, \dots, X_n$ were independent and identically distributed random vectors with integrable characteristic functions, it is possible to obtain an approximation which is appropriate for arbitrarily small sets B. In this case we can show, from Robinson, Höglund, Holst and Quine [(1990), Theorem 1], that

$$P(\overline{X} \in x + B) = \frac{\exp\left[n\left(\kappa(\theta) - \theta \cdot x\right)\right]}{\left(2\pi/n\right)^{d/2}\Delta^{1/2}}\left(1 + \frac{1}{n}Q_{2\theta}(0)\right)\mathrm{vol}(B)\left(1 + O\left(n^{-3/2}\right)\right).$$

Then there is an indirect Edgeworth expansion for the density,

$$f_n(x) = g_n(x)\left(1 + n^{-1}Q_{2\theta}(0) + O\left(n^{-3/2}\right)\right),$$

where

$$g_n(x) = \frac{\exp\left[-n\Lambda(x)\right]}{\left(2\pi/n\right)^{d/2}\Delta^{1/2}},$$

$$\Lambda(x) = \sup_t \left[t \cdot x - \kappa(t)\right] = \theta \cdot x - \kappa(\theta).$$

This is the usual saddlepoint estimate of the density. See Robinson, Höglund, Holst and Quine [(1990), Remark 6] for further details.

2.3. *Integrated saddlepoint approximations.* To obtain $P(\overline{X} \in B)$, for an arbitrary set B, we can integrate the density if it exists. In particular, if we integrate over $\Re^d$, we can renormalize the density by dividing by this integral. We will show that this improves the order of approximation. Integrating the density over $\Re^d$ gives

$$1 = \int_{\Re^d} g_n(x)\left(1 + \frac{1}{n}Q_{2\theta(x)}(0) + O\left(n^{-3/2}\right)\right)dx.$$

So

$$\begin{aligned}\int_{\Re^d} g_n(x)\,dx &= 1 - \frac{1}{n}\int_{\Re^d} g_n(x)Q_{2\theta(x)}(0)\,dx + O\left(n^{-3/2}\right)\\ &= 1 - \frac{1}{n}Q_{20}(0) + O\left(n^{-3/2}\right),\end{aligned}$$

from the Laplace approximation. So the renormalized saddlepoint density is

$$\frac{g_n(x)}{\int_{\Re^d} g_n(x)\,dx} = f_n(x) - \frac{1}{n}g_n(x)\left(\left[Q_{2\theta(x)}(0) - Q_{20}(0)\right] + O\left(n^{-1/2}\right)\right).$$

In the region $\|x\| = O(n^{-1/2})$, the second term above is $O(n^{-3/2})$, while for large deviations the error is $O(n^{-1})$.

We need a method of justifying the approximation obtained by integration of the formal density estimate as above when the conditions implying validity for the approximation to the density are not satisfied, but those for the indirect Edgeworth expansion are. This would be the case, for example, if we approximate the tail area of a rank statistic. In the case of a Wilcoxon statistic, the jumps in the distribution function are of $O(n^{-3/2})$, so an indirect Edgeworth expansion with this error could be obtained, but a density estimate is not available. If we integrate the formal density, then we obtain a result equal to the indirect Edgeworth expansion to $O(n^{-3/2})$. This equality follows in cases where the continuity assumptions are satisfied, and so in other cases, since this is not dependent on the continuity assumptions, but purely on the form of the expansions. So if the conditions of Theorem 1 are satisfied, we can use the integral of the formal density to obtain the following theorem.

THEOREM 2. *Assume that conditions* (S.1)–(S.4) *hold with* $s = 5$, *and that* B *is a rectangle* (*or a half-plane*) *and* $0 \notin B$. *Then*

$$P(\overline{X} \in B) = \int_B g_n(x)\left(1 + \frac{1}{n}Q_{2\theta(x)}(0)\right)dx\left(1 + O\left(n^{-3/2}\right)\right).$$

These approximations can be obtained by numerical integration, but it is convenient to have a simpler formula for the integrated result. Such a result is provided by the indirect Edgeworth method considered above, but in a number of cases a simpler form can be obtained by applying the method of Temme (1982) to the integral of the density. This technique was suggested by Daniels and Young (1991) and gives results of the same form as were obtained by Lugannani and Rice (1980). This third approach appears to be the most useful, and the remainder of this note is an attempt to put these problems in a single framework for smooth transforms of sample means.

3. Transformations. Consider a C^∞-transformation $g\colon \Re^d \to \Re^d$. Write $a = g(x)$. Let

$$J = \left| \left(\det\left(\frac{\partial g}{\partial x} \right) \right)^{-1} \right|,$$

and assume that $J > 0$ in $M_\delta = \{x\colon \|x\| < \delta\}$. Then g is a C^∞-diffeomorphism on M_δ onto its image M_δ^*. Let

$$L(a) = \Lambda\big(x(a)\big),$$

where $x(a) = g^{-1}(a)$ for $a \in M_\delta^*$. Let

$$L(\alpha) = \inf_a L(a) = 0,$$

so $x(\alpha) = 0$ if we take $ES = 0$. Choose δ small enough so that $L(a)$ is convex in M_δ^*; this is possible since

$$\left(\frac{\partial^2 L(a)}{\partial a^2} \right)_{x=0} = \left(\frac{\partial x}{\partial a} \right)_{x=0}^T \cdot \left(\frac{\partial^2 \Lambda(x)}{\partial x^2} \right)_{x=0} \cdot \left(\frac{\partial x}{\partial a} \right)_{x=0},$$

$(\partial\Lambda(x)/\partial x)_{x=0} = 0$ and $(\partial^2\Lambda(x)/\partial x^2)_{x=0}$ is positive definite. Assume also that conditions ensuring that the formal density estimate exists are satisfied. Then the joint density of $A = g(\overline{X})$ is $f(a)(1 + O(n^{-3/2}))$, where

$$f(a) = \frac{\exp\big[-nL(a)\big]J}{\big(2\pi/n\big)^{d/2}\Delta^{1/2}}\big(1 + n^{-1}Q_{2\theta}(0)\big). \tag{3.1}$$

Note that here $\theta = \theta(x(a))$.

Let B be a convex set which intersects M_δ^* but such that B does not contain α. Then we can minimize $L(a)$ in $B \cap M_\delta^*$ and the minimum occurs on the boundary of B.

4. Marginal densities and distributions. Let $A = (A_1, A_2) = (g_1(\overline{X}), g_2(\overline{X}))$ with $A_1 \in \Re^k$, $0 < k < d$. In this section, we are interested in getting saddlepoint density approximations for A_1, for general k, and saddlepoint approximations for the tail area probability $P(A_1 \geq \widehat{a}_1)$, for the case $k = 1$.

4.1. *Saddlepoint approximations for the marginal density.* Under the conditions of Section 2.2, which ensure the existence of the density of A, we can consider a saddlepoint approximation to the marginal density of A_1, $f_{A_1}(a_1)$. Let

$$H(a_1) = \inf_{a_2} L(a) = L(\widetilde{\alpha}),$$

where $\tilde{\alpha} = (a_1, \tilde{\alpha}_2(a_1))$. So from the Laplace approximation, the marginal density of A_1 is

$$
\begin{aligned}
f_{A_1}(a_1) &= \frac{\exp[-nH(a_1)]}{(2\pi/n)^{k/2}} \int_{\Re^{d-k}} \frac{\exp\left[-n\left(L(a) - L(\tilde{\alpha})\right)\right]}{(2\pi/n)^{(1/2)(d-k)}\Delta^{1/2}} \\
&\quad \times J\left(1 + n^{-1}Q_{2\theta}(0) + \left(n^{-3/2}\right)\right) da_2 \qquad (4.1)\\
&= (n/2\pi)^{k/2} \exp[-nH(a_1)]G(a_1)\left(1 + n^{-1}Q_2(a_1) + O\left(n^{-3/2}\right)\right),
\end{aligned}
$$

where

$$
G(a_1) = \frac{\tilde{J}}{\tilde{\Delta}^{1/2}\det(\tilde{L}_{22})^{1/2}}, \tag{4.2}
$$

$L_{22}(a) = \partial^2 L(a)/\partial a_2^2$, and $\tilde{J}$, $\tilde{\Delta}$ and $\tilde{L}_{22}$ are J, Δ and L_{22} evaluated at $\tilde{\alpha}$ and $Q_2(a_1)$ is a function of a_1 obtained from the integration.

As in Daniels and Young (1991), we can write

$$
H''(a_1) = L_{11} - L_{12}L_{22}^{-1}L_{21},
$$

so another expression for $G(a_1)$ is

$$
G(a_1) = \frac{\tilde{J}\det\left[\Lambda''(x(\tilde{\alpha}))\right]^{1/2} H''(a_1)^{1/2}}{\det(\tilde{L}'')^{1/2}}. \tag{4.3}
$$

4.2. *An alternative form for the marginal density.* Using (4.2), and two matrix identities, Tierney, Kass and Kadane (1989, 1991) derived an expression for $G(a_1)$ which does not involve $g_2(x)$. For a fixed $a_1 \in \Re^k$, let $\lambda \in \Re^k$ be the Lagrange multiplier for the problem of minimizing the function $\Lambda(x)$, subject to the constraint $g_1(x) = a_1$. Let $x(a_1)$ be the solution to the problem, and $\overline{\Lambda}(x) = \Lambda(x) - \lambda \cdot g_1(x)$. From Tierney, Kass and Kadane (1989, 1991),

$$
G(a_1) = \left(\frac{\det[\Lambda''(x)]}{\det\left[\overline{\Lambda}''(x)\right] \cdot \det\left\{ (g_1'(x))^T \left[\overline{\Lambda}''(x)\right]^{-1} (g_1'(x)) \right\}} \right)^{1/2}_{x = x(a_1)}, \tag{4.4}
$$

where by $\overline{\Lambda}''(x)$ we mean $\Lambda''(x) - \lambda \cdot g_1''(x)$. Also, $H(a_1)$ from Section 4.1 is equal to $\Lambda(x(a_1))$. Therefore

$$
f_{A_1}(a_1) = (n/2\pi)^{k/2}\exp\left[-n\Lambda(x(a_1))\right]G(a_1)\left(1 + n^{-1}Q_2(a_1) + O\left(n^{-3/2}\right)\right) \tag{4.5}
$$

with $G(a_1)$ defined as in (4.4).

4.3. *Integrated saddlepoint approximations for the marginal distribution.* In this section, we restrict attention to $k = 1$. Note we can write $P(A_1 \geq \widehat{a}_1) = P(A \in B)$ with $B = \{a: a_1 \geq \widehat{a}_1\}$, a half-plane. If the assumptions ensuring that the formal density estimate exists are satisfied, we can integrate (4.1) over B to get a saddlepoint approximation for $P(g(\overline{X}) \in B)$ with relative error $O(n^{-3/2})$. However, even if these assumptions are not satisfied, we can use the methods of Section 2.3 to show that the integral of (4.1) over B gives a saddlepoint approximation with relative error $O(n^{-3/2})$. The following theorem is related to results obtained under stronger conditions involving existence of densities by, for example, Tierney, Kass and Kadane (1989), DiCiccio, Field and Fraser (1990), DiCiccio and Martin (1991) and Daniels and Young (1991).

THEOREM 3. *If conditions* (S.1)–(S.4) *hold with* $s = 5$, *and if* $B = \{a: a_1 \geq \widehat{a}_1\}$, *then, for* $\sqrt{n}(\widehat{a}_1 - \alpha_1) \geq c > 0$,

$$P(A_1 \geq \widehat{a}_1) = \int_B f(a)\,da\Big(1 + O\big(n^{-3/2}\big)\Big); \tag{4.6}$$

this can be written

$$\begin{aligned} P(A_1 \geq \widehat{a}_1) = {} & \left(1 - \Phi(\widehat{w}\sqrt{n}) - \frac{1}{\sqrt{n}}\varphi(\widehat{w}\sqrt{n})\left(\frac{1}{\widehat{w}} - \frac{\psi(\widehat{w})}{\widehat{w}}\right)\right) \\ & \times\left(1 + \frac{1}{n}\big(Q_2(\widehat{a}_1) - Q_2(\alpha_1)\big) + O\big(n^{-3/2}\big)\right), \end{aligned} \tag{4.7}$$

where $\widetilde{\alpha} = (\widehat{a}_1, \widetilde{\alpha}_2(\widehat{a}_1))$,

$$\widehat{w} = \sqrt{2H(\widehat{a}_1)}\ \operatorname{sgn}(\widehat{a}_1 - \alpha_1)$$

and

$$\frac{\psi(\widehat{w})}{\widehat{w}} = \frac{\widetilde{J}}{\widetilde{\Delta}^{1/2}\det[\widetilde{L}_{22}]^{1/2}H'(\widehat{a}_1)}.$$

PROOF. The proof of (4.6) concerns only technical smoothing arguments and so it is given in the appendix. However, the derivation of (4.7) from (4.6) is given below.

Consider the change of variable

$$w = \sqrt{2H(a_1)}\ \operatorname{sgn}(a_1 - \alpha_1),$$

where $L(\alpha) = \inf_{a_1} H(a_1) = 0$. Then integrating (4.1) gives

$$\begin{aligned} P(A_1 \geq \widehat{a}_1) = {} & \frac{1}{\sqrt{2\pi/n}}\int_{\widehat{w}}^{\infty} \exp(-nw^2/2)\psi(w) \\ & \times\Big(1 + n^{-1}Q_2\big(a_1(w)\big) + O\big(n^{-3/2}\big)\Big)dw, \end{aligned} \tag{4.8}$$

where

$$\psi(w) = G(a_1)\left|\frac{da_1}{dw}\right|.$$

Now

$$\frac{dw}{da_1} = \frac{H'(a_1)}{w}.$$

So

$$\left(\frac{dw}{da_1}\right)_{a_1=\alpha_1} = \lim_{a_1\to\alpha_1}\frac{H'(a_1)}{w} = \frac{(H''(\alpha_1))}{(dw/da_1)_{a_1=\alpha_1}}.$$

Thus

$$\left|\frac{dw}{da_1}\right|_{a_1=\alpha_1} = \big(H''(\alpha_1)\big)^{1/2}.$$

So from (4.3),

$$\psi(0) = \frac{G(\alpha_1)}{\big(H''(\alpha_1)\big)^{1/2}} = 1.$$

Using the expression $\psi(w) = 1 + (\psi(w) - 1)$ in (4.8), we have

$$P(A_1 \geq \widehat{a}_1) = \left(1 - \Phi(\widehat{w}\sqrt{n}) - \frac{1}{\sqrt{n}}\varphi(\widehat{w}\sqrt{n})\left(\frac{1}{\widehat{w}} - \frac{\psi(\widehat{w})}{\widehat{w}}\right)\right)$$
$$\times\left(1 + \frac{1}{n}\psi_1(0) + \frac{1}{n}Q_2(\widehat{a}_1) + O\big(n^{-3/2}\big)\right),$$

where $\psi_1(w) = [1 - \psi(w)]/w^2 + \psi'(w)/w$. Taking $\widehat{w} = -\infty$ in (4.8), we obtain

$$1 = 1 + n^{-1}\psi_1(0) + n^{-1}Q_2(\alpha_1) + O\big(n^{-3/2}\big).$$

So $\psi_1(0) = -Q_2(\alpha_1) + O(n^{-1/2})$, and (4.7) follows immediately. □

4.4. *An alternative form for the marginal distribution.* Corresponding to the alternative form of the saddlepoint density given by Tierney, Kass and Kadane (1989) and discussed in Section 4.2, we can derive an integrated saddlepoint formula which involves only the transformation $a_1 = g_1(x)$.

From Section 4.2, $H(a_1) = \Lambda[x(a_1)]$, so we get

$$H'(a_1) = \Lambda'\big[x(a_1)\big]\cdot x'(a_1),$$

and thus

$$P(A_1 \geq \widehat{a}_1) = \left(1 - \Phi(\widehat{w}\sqrt{n}) - \frac{1}{\sqrt{n}}\varphi(\widehat{w}\sqrt{n})\left(\frac{1}{\widehat{w}} - \frac{\psi(\widehat{w})}{\widehat{w}}\right)\right)$$
$$\times\left(1 + \frac{1}{n}\big(Q_2(\widehat{a}_1) - Q_2(\alpha_1)\big) + O\big(n^{-3/2}\big)\right),$$

where

$$\widehat{w} = \sqrt{2\Lambda\left[x(\widehat{a}_1)\right]}\,\operatorname{sgn}\left(\widehat{a}_1 - \alpha_1\right),$$

$$\frac{\psi(\widehat{w})}{\widehat{w}} = \left(\frac{\det\left[\Lambda''(x)\right]}{\det\left[\overline{\Lambda}''(x)\right]\cdot\det\left[\left(g_1'(x)\right)^T\left[\overline{\Lambda}''(x)\right]^{-1}\left(g_1'(x)\right)\right]}\right)^{1/2}_{x=x(\hat{a}_1)}\frac{1}{\Lambda'\left[x(\widehat{a}_1)\right]\cdot x'(\widehat{a}_1)}.$$

5. Conditional densities and distributions.

5.1. *Conditional densities.* Under the conditions of Section 2.2, which ensure the existence of the density of A, we can consider a saddlepoint approximation to the conditional density $f_{A_1\mid A_2=a_2}(a_1)$, the density of A_1 at a_1 given $A_2 = a_2$, where $a = (a_1, a_2)$ with $a_1 \in \Re^k$ and $a_2 \in \Re^{d-k}$. Let

$$L\left((\widetilde{\alpha}_1, a_2)\right) = \inf_{a_1} L\left((a_1, a_2)\right).$$

Now from (3.1),

$$f_A(a) = \frac{\exp\left[-nL(a)\right]J}{\left(2\pi/n\right)^{d/2}\Delta^{1/2}}\left(1 + n^{-1}Q_2(a) + O\left(n^{-3/2}\right)\right),$$

and also from the marginal density result,

$$f_{A_2}(a_2) = \frac{\exp\left[-nL(\widetilde{\alpha}_1, a_2)\right]\widetilde{J}_2}{\left(2\pi/n\right)^{(d-k)/2}\widetilde{\Delta}_2^{1/2}\det(\widetilde{L}_{11})^{1/2}}\left(1 + n^{-1}Q_2^*(a_2) + O\left(n^{-3/2}\right)\right),$$

where $L_{11}(a) = \partial^2 L(a)/\partial a_1^2$, and $\widetilde{J}_2$, $\widetilde{\Delta}_2$ and $\widetilde{L}_{11}$ are J, Δ and L_{11} evaluated at $(\widetilde{\alpha}_1(a_2), a_2)$. So

$$f_{A_1\mid A_2=a_2}(a_1) = \frac{f_A(a)}{f_{A_2}(a_2)}$$

$$= \left(\frac{n}{2\pi}\right)^{k/2}\frac{J}{\widetilde{J}_2}\left(\frac{\det(\widetilde{L}_{11})\widetilde{\Delta}_2}{\Delta}\right)^{1/2}\exp\left\{-n\left[L(a) - L\left((\widetilde{\alpha}_1, a_2)\right)\right]\right\} \qquad (5.1)$$

$$\times\left(1 + n^{-1}q_2(a_1) + O\left(n^{-3/2}\right)\right).$$

5.2. *Conditional distributions.* Again we will restrict our attention to $k = 1$. We will find a saddlepoint approximation to

$$P\left(A_1 \geq \widehat{a}_1 \mid A_2 = a_2\right).$$

Unlike Section 4.3, here we will assume a strong Cramér condition because we use the density in the conditioning. Again in the following, we will take $s = 5$.

THEOREM 4. *If conditions* (S.1)–(S.3) *hold and also*

(S.4′) $$\lim_{n\to\infty}\sup_{\|\eta\|>c}\left|\hat{\nu}_\theta(\eta)\right| < 1,$$

then

(5.2) $$P(A_1 \ge \hat{a}_1 \mid A_2 = a_2) = \left(1 - \Phi(\hat{w}\sqrt{n}) - \frac{1}{\sqrt{n}}\varphi(\hat{w}\sqrt{n})\left(\frac{1}{\hat{w}} - \frac{\psi(\hat{w})}{\hat{w}}\right)\right) \times\left(1 + \frac{1}{n}(q_2(\hat{a}_1) - q_2(\tilde{\alpha}_1)) + O(n^{-3/2})\right)$$

where

$$\hat{w} = \sqrt{2\left[L(\hat{a}_1, a_2) - L((\tilde{\alpha}_1, a_2))\right]}\ \operatorname{sgn}(\hat{a}_1 - \tilde{\alpha}_1),$$

$$\frac{\psi(\hat{w})}{\hat{w}} = \frac{J}{\tilde{J}_2}\left(\frac{\det(\tilde{L}_{11})\tilde{\Delta}_2}{\Delta}\right)^{1/2}\left(\frac{\partial L}{\partial a_1}\right)^{-1}.$$

PROOF. Consider the change of variable a to (w, a_2), where

(5.3) $$w = \sqrt{2(L(a) - L(\tilde{\alpha}_1, a_2))}\ \operatorname{sgn}(a_1 - \tilde{\alpha}_1).$$

Then integrating (5.1),

(5.4) $$P(A_1 \ge \hat{a}_1 \mid A_2 = a_2) = \sqrt{\frac{n}{2\pi}}\int_{\hat{w}}^{\infty}\exp\left(\frac{-nw^2}{2}\right)\psi(w)\left(1 + n^{-1}q_2(a_1) + O(n^{-3/2})\right)dw,$$

where

$$\psi(w) = \frac{J}{\tilde{J}_2}\left(\frac{\det(\tilde{L}_{11})\tilde{\Delta}_2}{\Delta}\right)^{1/2}\frac{1}{dw/da_1} = \frac{J}{\tilde{J}_2}\left(\frac{\det(\tilde{L}_{11})\tilde{\Delta}_2}{\Delta}\right)^{1/2}\frac{w}{\partial L/\partial a_1}.$$

Similarly to Theorem 3, we can easily show that $\psi(0) = 1$. Then writing $\psi(w) = 1 + (\psi(w) - 1)$ in (5.4), we get Theorem 4 in the same way as Theorem 3. □

REMARK 5.1. Theorem 4 extends Skovgaard (1987), which gave equivalent results for a linear transformation. It also generalizes the results by Wang (1993).

6. Applications. One important application of saddlepoint approximations is in bootstrap analysis. Davison and Hinkley (1988) applied saddlepoint approximations to replace intensive Monte Carlo simulation for unstudentized bootstrap means with great accuracy. Daniels and Young (1991) extended the technique to the studentized bootstrap means. Jing, Feuerverger and Robinson (1994) used the saddlepoint approximations to look at the accuracy and coverage of bootstrap confidence intervals by studying relative errors in much the way Hall (1988) used Edgeworth methods.

In the following, we will give an example to illustrate the use of Theorem 3. Consider the simple linear regression model

$$y_i = \alpha + \beta z_i + \varepsilon_i, \qquad i = 1, \dots, n,$$

where $\varepsilon_1, \dots, \varepsilon_n$ are an independent sample from the distribution function F. Here we assume $z_1, \dots, z_n$ are fixed and $\sum_{i=1}^n z_i = 0$. Then the least-squares estimate of β is $\widehat{\beta} = \sum_{i=1}^n z_i y_i / \sum_{i=1}^n z_i^2$. Robinson (1987) considers the bootstrap (percentile and percentile-t) and randomization methods for constructing confidence intervals for β and indicates that the percentile-t method gives the best approximations. See also Hall [(1992), Section 4.3] for comparisons of different bootstrap methods in the regression context.

A confidence interval for β is

$$\left(\widehat{\beta} - \frac{u_L(F) s}{\left(\sum_{i=1}^n z_i^2\right)^{1/2}}, \widehat{\beta} - \frac{u_U(F) s}{\left(\sum_{i=1}^n z_i^2\right)^{1/2}} \right),$$

where $s^2 = (n-2)^{-1} \sum_{i=1}^n \widehat{\varepsilon}_i^2$, $u_U(F)$ is the solution of

$$P\left(\frac{\widehat{\beta} - \beta}{s / \left(\sum_{i=1}^n z_i^2\right)^{1/2}} \le u_U(F) \right) = \frac{1}{2}\alpha,$$

with $u_L(F)$ being defined similarly. The percentile-t confidence interval for β is

$$\left(\widehat{\beta} - \frac{u_L^* s}{\left(\sum_{i=1}^n z_i^2\right)^{1/2}}, \widehat{\beta} - \frac{u_U^* s}{\left(\sum_{i=1}^n z_i^2\right)^{1/2}} \right),$$

where u_U^* is the bootstrap approximation for $u_U(F)$,

$$P^*\left(\frac{\beta^* - \widehat{\beta}}{s^* / \left(\sum_{i=1}^n z_i^2\right)^{1/2}} \le u_U^* \right) = \frac{1}{2}\alpha, \tag{6.1}$$

where $\beta^* - \widehat{\beta} = \sum_{i=1}^n \varepsilon_i^* z_i / \sum_{i=1}^n z^2$,

$$s^{*2} = (n-2)^{-1} \left[\sum_{i=1}^n \varepsilon_i^{*2} - n(\overline{\varepsilon}^*)^2 - (\beta^* - \widehat{\beta})^2 \sum_{i=1}^n z_i^2 \right]$$

and the ε_i^*'s are generated randomly with replacement from the residuals $\widehat{\varepsilon}_i$'s, and where P^* denotes probability conditional on the $\widehat{\varepsilon}_i$'s, and u_L^* is defined similarly. It can easily be shown that (6.1) is equivalent to

$$P^*\left(\frac{n^{1/2}\left[\sum_{i=1}^n p_i\varepsilon_i^*/n\right]}{\left[\sum_{i=1}^n \varepsilon_i^{*2}/n-(\bar{\varepsilon}^*)^2\right]^{1/2}}\le t_U^*\right)=\frac{1}{2}\alpha,$$

where $p_i = n^{1/2}z_i/(\sum_{i=1}^n z_i^2)^{1/2}$ and $t_U^* = [n/(n-2)]^{1/2}u_U^*[1+u_U^{*2}/(n-2)]^{-1/2}$.

We can approximate t_U^*, hence u_U^* by Monte Carlo simulation. However, we now apply Theorem 4 in an attempt to replace Monte Carlo simulation, in a way similar to Davison and Hinkley (1988) for unstudentized means and to Daniels and Young (1991) for studentized means. To approximate P^*, the conditional probability given the sample Cramér's condition is not satisfied, but if it holds for the ε's, then (S4) holds for the ε^*'s [see, e.g., Albers, Bickel and van Zwet (1976)]. Then, to apply Theorem 3, we consider $X_i = (\varepsilon_i^*, \varepsilon_i^{*2}, p_i\varepsilon_i^*)$, $i = 1,\dots,n$. So $\overline{X} = (\bar{\varepsilon}^*, \sum\varepsilon_i^{*2}/n, \sum p_i\varepsilon_i^*/n)$, and

$$\begin{aligned}\kappa(\theta_1,\theta_2,\theta_3) &= \frac{1}{n}\sum_{i=1}^n \log E\exp\Big(\theta_1\varepsilon_i^*+\theta_2\varepsilon_i^{*2}+\theta_3\big(p_i\varepsilon_i^*\big)\Big)\\ &= \frac{1}{n}\sum_{i=1}^n \log\left[\frac{1}{n}\sum_{j=1}^n \exp\big((\theta_1+p_i\theta_3)\widehat{\varepsilon}_i+\theta_2\widehat{\varepsilon}_i^2\big)\right].\end{aligned}$$

We will choose the following transformation:

$$\begin{aligned}a_1 &= n^{1/2}x_3/\big(x_2-x_1^2\big)^{1/2},\\ a_2 &= x_1,\\ a_3 &= x_2.\end{aligned}$$

For the law school data in Efron (1982), we obtained $u_L^* = 1.777$ and $u_U^* = -1.728$ for $\alpha = 0.10$. So the 90% percentile-t confidence interval for β by using the saddlepoint approximations is (0.271, 0.628). By using Monte Carlo simulation (with 500,000 resamples drawn), we get a 90% percentile-t confidence interval for β to be (0.2692, 0.6284). In this case the saddlepoint method gives a very good approximation.

APPENDIX

A.1. Proof of Theorem 1. We will first get a bound for $\delta(\theta,B,T)$, which is the convolution of $\delta(\theta,B)$ with a smoothing density over the d variables, and then through a smoothing lemma similar to Lemma 2 of Robinson, Höglund, Holst and Quine (1990), we will get a bound for $\delta(\theta,B)$.

From (2.2),

$$\delta(\theta,B)=\int_{\Re^d}\chi_{m-B}(-y)\exp(-n\theta\cdot y)H(dy),$$

where

$$H(dy) = \nu_{n\theta}(dy) - e_{s-3}(y, \nu_\theta)\,dy,$$

where $e_k(y, \nu_\theta)$ is the formal Edgeworth expansion for ν_θ of order k, so

$$\delta(\theta, B, T) \equiv \int_{\Re^d} \delta(\theta, B - u) K_L(du)$$
$$= (2\pi)^{-d} \int_{\Re^d} \widehat{K}_L(\xi) \widehat{\chi}_{\theta, m - B}(\xi) \widehat{h}(\xi)\,d\xi,$$

where

$$\widehat{\chi}_{\theta, m - B}(\xi) = \int_{\Re^d} \chi_{m - B}(u) \exp(n\theta \cdot u + i\xi \cdot u)\,du,$$

$\widehat{h}$ is the Fourier transformations of H and K_L is a smoothing function to be chosen later.

We first consider the case that B is a rectangle. There exists a matrix Q such that

$$B' \equiv \{w = Qu\colon u \in B\} = \{w \mid a_j \le w_j \le b_j,\ j = 1, \dots, d\}.$$

Let $\eta = Q\xi$. Then

$$\delta(\theta, B, T) = (2\pi)^{-d} \int_{\Re^d} \widehat{K}_L(Q'\eta) \widehat{\chi}_{\theta, m - B}(Q'\eta) \widehat{h}(Q'\eta)\,d\eta,$$

where K_L is chosen such that

$$\widehat{K}_L(Q'\eta) = \prod_{j=1}^{d} \left(1 - \frac{|\eta_j|}{T}\right) \quad \text{for } |\eta_j| \le T.$$

We should note that all the C's used in the following are positive constants and may be different at each occurrence.

LEMMA 1. *Under the conditions of Theorem* 1,

$$|\delta(\theta, B, T)| \le Cn^{-(s-2)/2} + Cn^{-\lambda} T^d$$

PROOF. Using the method and notation of von Bahr [(1967), pages 76–77], we can show that

$$|\delta(\theta, B, T)| \le C {\sum_{(\Gamma, \Lambda)}}' \int_{n^{1/2}\|\eta_\Gamma\| \le T,\, \gamma \in \Gamma} \frac{\left| \prod_{\gamma \in \Gamma} \left(1 - \exp(-n\alpha\eta_\gamma^2) P_\gamma\right) \widehat{h}(Q'\eta_\Gamma)\right|}{\prod_{\gamma \in \Gamma} |\eta_\gamma|}\,d\eta_\Gamma$$
$$\le I_1 + I_2 + I_3,$$

where I_i is integral over S_i, $i = 1, 2, 3$, with

$$S_1 = \left\{\eta_\Gamma: n^{1/2}\|W(\theta)^{1/2}\eta_\Gamma\| < Cn^{(s-2)/(2s)}\right\},$$
$$S_2 = \left\{\eta_\Gamma: Cn^{(s-2)/(2s)} < n^{1/2}\|W(\theta)^{1/2}\eta_\Gamma\| < Cn^{1/2}\right\},$$
$$S_3 = \left\{\eta_\Gamma: Cn^{1/2} < n^{1/2}\|W(\theta)^{1/2}\eta_\Gamma\| < T\right\},$$

for $W = QVQ'$. It is easy to check, using von Bahr's method, that $I_1 \le Cn^{-(s-2)/2}$, and we can follow Robinson, Höglund, Holst and Quine (1990) to get $I_2 \le Cn^{-(s-2)/2}$ and $I_3 \le Cn^{-(s-2)/2}$, for $T = n^{(s-2)/2}$, from which the lemma follows. □

The following lemma can be obtained from Lemma 2 of Robinson, Höglund, Holst and Quine (1990) with slight changes, but applies only to half-planes or rectangles.

LEMMA 2. *If* $|\theta| < C$ *and* $\varepsilon = c/T$, *where* $K_T(S(\varepsilon)) = 1 - \alpha > 2/3$, *for* $S(\varepsilon) = \{y: y \in \Re^d, \|y\| < \varepsilon\}$, *and if* B *is a half-plane or a rectangle*, *then writing* $(B)_\varepsilon = \{x + y: x \in B, y \in S(\varepsilon)\}$ *and* $(B)_{-\varepsilon} = ((B^c)_\varepsilon)^c$ *we have*

$$|\delta(\theta, B)| \le C\left\{\max\left(|\delta(\theta, B_\varepsilon, T)|, |\delta(\theta, B_{-\varepsilon}, T)|\right) + C\varepsilon\right\}.$$

Choosing $T = n^{(s-2)/2}$ in Lemmas 1 and 2, we get Theorem 1. Notice that there is no restriction on the size of B, so the theorem is true for half-planes as well.

A.2. Proof of Theorem 3. Choose ε small enough such that $N_\varepsilon^* \equiv \{x: |x_i| < \varepsilon, i = 1, \dots, d\} \subset M_\delta^*$ and $N_\varepsilon = g^{-1}(N_\varepsilon^*)$. So we have

$$P\left(g(\overline{X}) \in B\right) = P\left(\overline{X} \in g^{-1}\left(B \cap N_\varepsilon^*\right)\right) + P\left(g(\overline{X}) \in B \cap \left(N_\varepsilon^*\right)^c\right),$$

where $P(g(\overline{X}) \in B \cap (N_\varepsilon^*)^c) = n^{-c(\log n)^2}$, for $\varepsilon = \log n/\sqrt{n}$.

Let $A = B \cap N_\varepsilon^*$, $g^{-1} = h$ and $A' = h(A)$. If v_1 is a vertex of A, then, for any $v \in A$,

$$h(v) = h(v_1) + h'(v_1)(v - v_1) + O\left(\delta^2\right) = h_1(v) + O\left(\delta^2\right).$$

Consider the parallelogram P given by $h_1(A)$. From $A' \subset P \cup (A' - P)$, we get

$$|\delta(\theta, A', T)| \le |\delta(\theta, P, T)| + \left|\delta\left(\theta, (A' - P), T\right)\right|,$$

where

$$\delta(\theta, A, T) = I(\theta, A, T) - e_{s-3}(\theta, A, T).$$

By the method of von Bahr (1967), we can show that, as in the proof of Theorem 1, for $T = n^{(s-2)/2}$,

$$|\delta(\theta, P, T)| = O\left(n^{-(s-2)/2}\right).$$

From Robinson, Höglund, Holst and Quine (1990, Lemma 1),

$$|\delta(\theta,(A'-P),T)| \le C\widehat{\chi}_{\theta,(A'-P)}(0)n^{d/2} \times\Big(\eta_s(\theta)\big(\det V(\theta)\big)^{-1/2}n^{-(s-2)/2}+T^d q_{n\theta}(T)\Big).$$

We can show, for $\sqrt{n}(\widehat{a}_1-\alpha_1)\ge c>0$,

$$\widehat{\chi}_{\theta,(A'-P)}(0)=O\big(\delta^{(d+1)}\big)=O\big(n^{-(d+1)/2}(\log n)^{d+1}\big).$$

Now taking $T=n^{(s-2)/2}$ in condition (S.4), we get $|\delta(\theta,(A'-P),T)|=O(n^{-(s-2)/2})$. Hence

$$|\delta(\theta,A',T)|=O\big(n^{-(s-2)/2}\big). \tag{A.1}$$

It can also be shown that, similarly to Lemma 1,

$$|\delta(\theta,A')|\le C\Big\{\max\big(|\delta(\theta,A'_\varepsilon,T)|,|\delta(\theta,A'_{-\varepsilon},T)|\big)+C\varepsilon\Big\}. \tag{A.2}$$

Then choosing $T=n^{(s-2)/2}$ in (A.2), and using (A.1), we get the theorem.

REFERENCES

ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1976). Asymptotic expansions for the power of distribution free tests in the one-sample problem. *Ann. Statist.* **4** 108–156.

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

BARNDORFF-NIELSEN, O. E. (1991). Modified signed log likelihood ratio. *Biometrika* **78** 557–563.

BARNDORFF-NIELSEN, O. E. and COX, D. R. (1979). Edgeworth and saddle-point approximations with statistical applications (with discussion). *J. Roy. Statist. Soc. Ser. B* **41** 279–312.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

DANIELS, H. E. (1987). Tail probability approximations. *Internat. Statist. Rev.* **55** 37–48.

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., FIELD, C. A. and FRASER, D. A. S. (1990). Approximations of marginal tail probabilities and inference for scalar parameters. *Biometrika* **77** 77–95.

DICICCIO, T. J. and MARTIN, M. A. (1991). Approximations of marginal tail probabilities for a class of smooth functions with applications to Bayesian and conditional inference. *Biometrika* **78** 891–902.

DURBIN, J. (1980). Approximations for densities of sufficient statistics. *Biometrika* **67** 311–333.

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

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.

JENSEN, J. L. (1992). The modified signed likelihood statistic and saddlepoint approximations. *Biometrika* **79** 693–703.

JING, B. Y., FEUERVERGER, A. and ROBINSON, J. (1994). On the bootstrap saddlepoint approximations. *Biometrika* **81** 211–215.

LUGANNANI, R. and RICE, S. O. (1980). Saddlepoint approximation for the distribution function of the sum of independent random variables. *Adv. in Appl. Probab.* **12** 475–490.

ROBINSON, J. (1987). Nonparametric confidence intervals in regression: the bootstrap and randomization methods. In *New Perspectives in Theoretical and Applied Statistics* (M. L. Puri, U. P. Vilaplana and W. Werts, eds.) 243–256. Wiley, New York.

ROBINSON, J., HÖGLUND, T., HOLST, L. and QUINE, M. P. (1990). On approximating probabilities for small and large deviations in $\Re^d$. *Ann. Probab.* **18** 727–753.

SKOVGAARD, IB. M. (1986). On multivariate Edgeworth expansions. *Internat. Statist. Rev.* **54** 169–186.

SKOVGAARD, IB. M. (1987). Saddlepoint expansions for conditional distributions. *J. Appl. Probab.* **24** 875–887.

TEMME, N. M. (1982). The uniform asymptotic expansion of a class of integrals related to cumulative distribution functions. *SIAM J. Math.* **13** 239–253.

TIERNEY, L., KASS, R. E. and KADANE, J. B. (1989). Approximate marginal densities of nonlinear functions. *Biometrika* **76** 425–433. [Correction: **78** (1991) 233–234.]

VON BAHR, B. (1967). Multi-dimensional integral limit theorems. *Ark. Mat.* **7** 71–88.

WANG, S. (1993). Saddlepoint approximations in conditional inference. *J. Appl. Probab.* **30** 397–404.

SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF SYDNEY
NSW 2006
AUSTRALIA