Title
A Conditional Saddlepoint Approximation for Testing Problems

Journal
Journal of the American Statistical Association, 94(446)

ISSN
0162-1459

Authors
Gatto, R
Jammalamadaka, SR

Publication Date
1999-06-01

DOI
10.1080/01621459.1999.10474148

Peer reviewed

A Conditional Saddlepoint Approximation for Testing Problems

Riccardo GATTO and S. Rao JAMMALAMADAKA

A saddlepoint approximation is provided for the distribution function of one M statistic conditional on another M statistic. Many interesting statistics based on dependent quantities (e.g., spacings, multinomial frequencies, rank differences) can be expressed in terms of independent identically distributed random variables conditioned on their sum, so that this conditional saddlepoint approximation yields accurate approximations for the distribution of such statistics. This saddlepoint approximation can also be used in conditional testing, where nuisance parameters are eliminated by conditioning on sufficient statistics.

KEY WORDS: Conditional test; Exponential distribution; Geometric distribution; Goodness-of-fit test; M statistics; Multinomial distribution; Nonparametric test; Poisson distribution; Rank test; Spacings; Two-sample test.

1. INTRODUCTION

In this article we propose a saddlepoint approximation for the distribution function of an M statistic conditional on another M statistic. Many important test statistics can be rewritten as conditional M statistics in this sense, and thus our conditional saddlepoint approximation can be exploited to obtain accurate approximations to p values or critical values. Such test statistics include the class of spacing statistics, which are based on the gaps between the successive values of the ordered sample. They have proven useful in various statistical problems, chief among them the goodness-of-fit tests. (For a general review on spacings, see Pyke 1965.) Except in a few special cases, the exact distribution of such statistics based on uniform spacings is unknown. For most cases, the asymptotic distribution is known, but it can be potentially misleading, especially when the sample size is moderate to small. We exploit the fact that the uniform spacings have the same distribution as exponential random variables conditioned on their sum. Hence any function of uniform spacings has the same distribution as the corresponding function in terms of independent and identically distributed (iid) exponentials random variables conditioned on their sum. In testing multinomial/grouped data, the likelihood ratio test or the Pearson chi-squared test can be written similarly in terms of iid Poisson random variables conditioned on their total, because multinomial frequencies have such a conditional Poisson representation. We thus can provide more accurate approximations for the distribution of the likelihood ratio or the chi-squared statistics, useful for small to moderate sample sizes. In testing whether two samples are from the same population, Holst and Rao (1980) showed that a whole range of classical rank tests have a simpler equivalent representation in terms of what are called the "spacing frequencies"; that is, the frequencies of one sample that fall in between the successive order statistics of the other sample. We can use our saddlepoint approximation to improve the accuracy of the asymptotic normal approximation in this context, by using the fact that these spacing frequencies have again the same distribution as iid geometric random variables conditioned on their sum. Tests in this class include the van der Waerden/normal score test and the Wilcoxon/Mann–Whitney test. A further important use of our saddlepoint approximation is in conditional testing, where nuisance parameters are eliminated by conditioning on their sufficient statistics.

The article is organized as follows. In Section 2 we give the essential computational steps of our saddlepoint approximation for conditional M statistics. In Section 3 we show how our saddlepoint approximation can be used with the specific one-sample tests. Several numerical computations illustrate the high level of accuracy of our methods. We devote Section 4 to the two-sample test statistics, giving many examples in which our saddlepoint approximation can be used advantageously.

2. CONDITIONAL SADDLEPOINT APPROXIMATION

The saddlepoint technique of asymptotic analysis was introduced into statistics by Daniels (1954) for deriving a very accurate approximation to the density of the mean of a sample of iid observations. In contrast with normal limits or Edgeworth approximations, the numerical accuracy of the saddlepoint approximation is surprisingly good, particularly in the tails of the distribution, and even for very small sample sizes. The Edgeworth expansion is known to inherit undesirable oscillations from its Hermite polynomials, sometimes leading to negative tail probabilities. Several extensions of Daniels's first formula have been proposed (see, e.g., Booth and Butler 1990 for randomization distributions; Field 1982 for M estimators; Gatto and Ronchetti 1996 for marginal densities of general statistics; Lugannani and Rice 1980 for tail probabilities). Some general texts or reviews include works by Barndorff–Nielsen and Cox (1989), Davison and Hinkley (1997), Field and Ronchetti (1990), Field and Tingley (1997), Jensen (1995), and Reid (1988). Sad-

Riccardo Gatto is Assistant Professor, Department of Statistics, ESSEC, Cergy-Pontoise, 95021 France, (E-mail: *gatto@edu.essec.fr*). S. Rao Jammalamadaka is Professor, Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 91306 (E-mail *rao@pstat.ucsb.edu*). Gatto's research is supported in part by the Swiss Federal Office of Science and Education. Rao's research is supported in part by National Science Foundation grant DMS-9803600. The authors thank the editor, an associate editor, and two anonymous referees for thoughtful comments that improved the quality of the article.

Journal of the American Statistical Association
June 1999, Vol. 94, No. 446, Theory and Methods

dlepoint approximations for conditional distributions have been proposed by Skovgaard (1987) for the distribution of a sample mean given another mean, by Wang (1993) for a mean given a nonlinear function of another mean, and by Jing and Robinson (1994) for a nonlinear function of a mean given another nonlinear function of another mean. DiCiccio, Martin, and Young (1993) proposed a different type of conditional approximation, which, however, requires that the sample have a distribution within the exponential class. The saddlepoint approximation that we propose for testing problems exploits some conditioning properties shared by many classes of statistics, which are reexpressed as conditional M statistics (see Secs. 3 and 4). By generalizing Skovgaard's (1987) conditional approximation for sample means, we first derive a saddlepoint approximation for conditional M statistics.

The main steps for the derivation of our conditional saddlepoint approximation to the distribution function are as follows. Consider n independent random variables $X_1 \ldots, X_n$ (scalar or vectors) and a M statistic $(T_{1n}, \mathbf{T}_{2n}^T), T_{1n} = T_{1n}(X_1, \ldots, X_n) \in \mathbb{R}$, and $\mathbf{T}_{2n} = \mathbf{T}_{2n}(X_1, \ldots, X_n) \in \mathbb{R}^p$, defined by

$$\sum_{i=1}^{n} \begin{pmatrix} \psi_{1i}(X_i, T_{1n}, \mathbf{T}_{2n}) \\ \boldsymbol{\psi}_{2i}(X_i, \mathbf{T}_{2n}) \end{pmatrix} = \mathbf{0}.$$

The joint cumulant generating function of the sum of score functions ψ_{1i} and ψ_{2i} is given by

$$K_n(\boldsymbol{\lambda}, \mathbf{t}) = \sum_{i=1}^{n} \log[E \exp\{\lambda_1 \psi_{1i}(X_i, t_1, \mathbf{t}_2) + \boldsymbol{\lambda}_2^T \boldsymbol{\psi}_{2i}(X_i, \mathbf{t}_2)\}], \quad (1)$$

where $\boldsymbol{\lambda} = (\lambda_1, \boldsymbol{\lambda}_2^T)$ and $\mathbf{t} = (t_1, \mathbf{t}_2^T)$, with t_1 the point at which we evaluate the conditional distribution and $\mathbf{t}_2$ the point of conditioning.

Step 1. Find $\boldsymbol{\alpha} \in \mathbb{R}^{p+1}$ and $\boldsymbol{\beta} \in \mathbb{R}^p$, the saddlepoints associated with $(T_{1n}, \mathbf{T}_{2n}^T)$ and to $\mathbf{T}_{2n}$, solutions of the joint and the conditioning saddlepoint equations, given by

$$\frac{\partial}{\partial \boldsymbol{\lambda}} K_n(\boldsymbol{\lambda}, \mathbf{t}) = \mathbf{0}$$

and

$$\frac{\partial}{\partial \boldsymbol{\lambda}_2} K_n((0, \boldsymbol{\lambda}_2), \mathbf{t}) = \mathbf{0}.$$

Step 2. Define

$$K_n''(\boldsymbol{\lambda}, \mathbf{t}) = \frac{\partial^2}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T} K_n(\boldsymbol{\lambda}, \mathbf{t}),$$

$$K_{2n}''(\boldsymbol{\lambda}_2, \mathbf{t}) = \frac{\partial^2}{\partial \boldsymbol{\lambda}_2 \partial \boldsymbol{\lambda}_2^T} K_n((0, \boldsymbol{\lambda}_2), \mathbf{t}),$$

$$s = \alpha_1 \left| \frac{\det(K_n''(\boldsymbol{\alpha}, \mathbf{t}))}{\det(K_{2n}''(\boldsymbol{\beta}, \mathbf{t}))} \right|^{1/2},$$

$$r = \operatorname{sgn}(\alpha_1)\{2[K_n((0, \boldsymbol{\beta}), \mathbf{t}) - K_n(\boldsymbol{\alpha}, \mathbf{t})]\}^{1/2},$$

and

$$P_n(t_1 | \mathbf{t}_2) = 1 - \Phi(r) + \phi(r) \left\{ \frac{1}{s} - \frac{1}{r} \right\}, \quad (2)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution functions and α_1 is the first element of $\boldsymbol{\alpha}$. Then, uniformly for $(t_1, \mathbf{t}_2^T)$ in sets where $(t_k - ET_{kn}) = O(1)$, $k = 1, 2$ (i.e., in large deviation regions),

$$P[T_{1n} \geq t_1 | \mathbf{T}_{2n} = \mathbf{t}_2] = P_n(t_1 | \mathbf{t}_2)\{1 + O(n^{-1})\}. \quad (3)$$

The error is also $O(n^{-3/2})$ over regions where $(t_k - ET_{kn}) = O(n^{-1/2}), k = 1, 2$; that is, in normal deviation regions.

Alternatively, the Barndorff–Nielsen tail area formula is given by

$$P_n^*(t_1 | \mathbf{t}_2) = 1 - \Phi\left(r + \frac{1}{r} \log\left\{ \frac{s}{r} \right\} \right), \quad (4)$$

and, by lemma 2.1 of Jensen (1992), $P_n(t_1|\mathbf{t}_2) = P_n^*(t_1|\mathbf{t}_2)\{1 + O(n^{-1})\}$, uniformly in large deviation regions. The asymptotic error becomes $O(n^{-3/2})$ in normal deviation regions.

In some cases we are interested in obtaining an approximation to the conditional density function instead of the cumulative distribution function. A straightforward approach is to compute the ratio of the saddlepoint approximations to the joint density of $(T_{1n}, \mathbf{T}_{2n}^T)$, with the marginal density of $\mathbf{T}_{2n}$. This is often referred to as the "double saddlepoint approximation" and maintains the same properties as an individual saddlepoint approximation in terms of relative error. It leads to the following approximation to the conditional density:

$$g_n(t_1 | \mathbf{t}_2) \propto |\det J_n(\boldsymbol{\alpha}, \mathbf{t})| \cdot |\det K_n''(\boldsymbol{\alpha}, \mathbf{t})|^{-1/2} \times \exp\{K_n(\boldsymbol{\alpha}, \mathbf{t})\}, \quad (5)$$

where

$$J_n(\boldsymbol{\lambda}, \mathbf{t}) = \sum_{i=1}^{n} E\left[\begin{pmatrix} \frac{\partial}{\partial t_1} \psi_{1i}(X_i, t_1, \mathbf{t}_2) & \frac{\partial}{\partial \mathbf{t}_2^T} \psi_{1i}(X_i, t_1, \mathbf{t}_2) \\ \mathbf{0} & \frac{\partial}{\partial \mathbf{t}_2^T} \boldsymbol{\psi}_{2i}(X_i, \mathbf{t}_2) \end{pmatrix} \times \exp\{\lambda_1 \psi_{1i}(X_i, t_1, \mathbf{t}_2) + \boldsymbol{\lambda}_2^T \boldsymbol{\psi}_{2i}(X_i, \mathbf{t}_2)\} \right],$$

where α solves the joint saddlepoint equation $\partial K_n(\boldsymbol{\lambda}, \mathbf{t})/\partial \boldsymbol{\lambda} = \mathbf{0}$. Normalizing (5) so that $\int g_n(t_1|\mathbf{t}_2)\, dt_1 = 1$ leads to a uniform relative error $O(n^{-1})$ in large deviation regions and to a relative $O(n^{-3/2})$ in normal deviation regions. This density approximation could allow for $\dim(T_{1n}) \geq 1$, although the computational effort would increase drastically with the dimensionality. Also, when computing a double saddlepoint approximation, global uniformity can be of crucial importance. For the density of a sample mean conditioned on another, Jensen (1991) showed that a sufficient condition for global uniformity is the log-concavity of the joint underlying density (see Jensen 1991, thm. 4). These results should in principle generalize to M estimators.

The justification of the foregoing formulas is based on the equivalence

$$[T_{1n} \geq t_1 | \mathbf{T}_{2n} = \mathbf{t}_2]$$
$$\text{iff } \left[\sum_{i=1}^{n} \psi_{1i}(X_i, t_1, \mathbf{t}_2) \geq 0 \middle| \sum_{i=1}^{n} \boldsymbol{\psi}_{2i}(X_i, \mathbf{t}_2) = \mathbf{0} \right],$$

which holds under the usual condition that ψ_{1i} and $\boldsymbol{\psi}_{2i}$ are decreasing functions in t_1 and t_2. Moreover, regularity conditions for (3) are similar to the conditions (I), (II), and (III) of Skovgaard (1987, sec. 3), with $\psi_{1i}(X_i, t_1, t_2)$ and $\psi_{2i}(X_i, t_2)$ replacing "X_i" and "Y_i" in the original reference.

Remark 1. From (1), we see that $X_1, \ldots, X_n$ need not be identically distributed, as it is often assumed. (See also Strawderman, Casella, and Wells 1996 for saddlepoint approximations for the non-iid case for [unconditional] M estimators.)

Remark 2. In most of our applications in the next two sections, we have, for $i = 1, \ldots, n, \psi_{1i}(\cdot) = a_i \psi_1(\cdot)$ for some fixed constants a_i that are identified in each case, and $\boldsymbol{\psi}_{2i}(\cdot) = \boldsymbol{\psi}_2(\cdot)$.

Remark 3. The conditioning statistic T_{2n} can be continuous or discrete. However, when T_{1n} is discrete with sum of score functions defined on the lattice $\{\omega + j\delta | j = 0, \pm 1, \ldots\}$, a similar formula can be derived with slight modifications by replacing s in (2) or (4) by

$$s_D = (1 - \exp\{-\delta \alpha_1\}) \left| \frac{\det(K_n''(\boldsymbol{\alpha}, \mathbf{t}))}{\det(K_{2n}''(\boldsymbol{\beta}, \mathbf{t}))} \right|^{1/2}. \qquad (6)$$

A continuity correction could also be derived, extending the results of Daniels (1987) and Skovgaard (1987). In all of our examples, even though T_{1n} is discrete, the domains are not lattices (i.e., not equally spaced), and these modifications for discreteness would not be helpful. Nevertheless, these domains are sufficiently fine so that the continuous saddlepoint approximations are numerically accurate.

3. ONE-SAMPLE TESTS

3.1 Tests Based on Spacings

Spacing statistics are based on the gaps between the successive values of the ordered sample. If $U_1, \ldots, U_{n-1}$ are iid random variables uniformly distributed on [0, 1], and $0 \leq U_{(1)} \leq \cdots \leq U_{(n-1)} \leq 1$, are their order statistics, we define the "uniform spacings" as

$$D_i = U_{(i)} - U_{(i-1)}, \qquad i = 1, \ldots, n,$$

where $U_{(0)} \stackrel{\text{def}}{=} 0$ and $U_{(n)} \stackrel{\text{def}}{=} 1$. Spacing statistics are used in various statistical problems, such as in testing goodness of fit, hazard rates, and so on. Spacings form a maximal invariant statistic with respect to changes of origin and sense of rotation (clockwise or counter-clockwise) in connection with circular data, so that any origin and sense invariant statistical procedure is actually a function of these spacings (see Rao 1969). For some appropriate choices of real-valued functions $h(\cdot)$ or $h_i(\cdot), i = 1, \ldots, n$, many of these statistics can be expressed as

$$T_n^* = \sum_{i=1}^{n} h_i(nD_i) \qquad (7)$$

or

$$T_n = \frac{1}{n} \sum_{i=1}^{n} h(nD_i). \qquad (8)$$

In all but a few special cases, the exact distribution of spacing statistics is unknown and, although the limiting distribution may be known for most cases, can be quite an inaccurate approximation. The following conditional representation of the spacings allows us to apply the saddlepoint approximation described in Section 2. If $E_1, \ldots, E_n$ are iid exponential random variables with distribution function $P[E_1 \leq x] = 1 - \exp\{-x\}, x \geq 0$, then it is known that

$$\{nD_1, \ldots, nD_n\} \sim \{E_1, \ldots, E_n\} \left| \sum_{i=1}^{n} E_i = n. \right.$$

Example 1: Rao's Spacing Test (Batschelet 1981; Rao 1969, 1976). Directions in two dimensions can be represented as points on the circumference of a unit circle or as angles. (See, e.g., Batchelet 1981 for an introduction.) One of the first steps before further modeling or inference is to check whether the data are isotropic (i.e., uniformly distributed without any preferred direction). The more general problem of goodness-of-fit testing on the circle can also be reduced to this via the probability integral transform. Formally, suppose that $X_1, \ldots, X_n$ are n iid random variables with circular continuous distribution F over the unit circle, and that we want to test the null hypothesis H_0: $F(x) = x/(2\pi), 0 \leq x < 2\pi$, versus the general alternative that it is not. If we place n arcs of equal length $(2\pi/n)$ starting at each sample point, then it can be seen that $\tilde{T}_n$, the "total uncovered part of the circumference," is given by

$$\tilde{T}_n = \sum_{i=1}^{n} \max\left\{\tilde{D}_i - \frac{2\pi}{n}, 0\right\} = \frac{1}{2} \sum_{i=1}^{n} \left| \tilde{D}_i - \frac{2\pi}{n} \right|,$$

where $\tilde{D}_i = X_{(i)} - X_{(i-1)}, i = 2, \ldots, n$, and $\tilde{D}_1 = 2\pi - X_{(n)} + X_{(1)}$. This defines Rao's statistic for testing uniformity over the unit circle. When all of the observations lie at the same point, corresponding to extreme clustering, then $\tilde{T}_n = 2\pi(1 - 1/n)$; when the observations are exactly equally spaced, then $\tilde{T}_n = 0$. In the related problem of testing uniformity on [0, 1], the statistic T_n defined in (8) is distributed as $\tilde{T}_n/(2\pi)$, and the corresponding score function is given by $h(x) = 1/2|x - 1|$. It has been shown this test statistic has a normal limit distribution (Rao 1969), namely

$$\sqrt{n}\left(T_n - \frac{1}{e}\right) \stackrel{\mathcal{D}}{\to} \mathcal{N}\left(0, \frac{2}{e} - \frac{5}{e^2}\right).$$

To apply steps 1 and 2 of the saddlepoint approximation, we must find the joint cumulant generating function (1) associated with the score functions

$$\psi_1(x,t_1) = \begin{cases} \frac{1}{2}\,(1-x-t_1), & \text{if } x \in [0,1] \\ \frac{1}{2}\,(x-1-t_1), & \text{if } x \in [1,\infty) \end{cases}$$

and

$$\psi_2(x,t_2) = x - t_2,$$

with $\psi_{ji} = \psi_j, i = 1,\ldots,n, j = 1,2$. For $\lambda_1/2 + \lambda_2 \le 1$, this cumulant generating function can be shown to be

$$\begin{aligned} K_n(\boldsymbol{\lambda},\mathbf{t}) &= -n\left[\frac{\lambda_1 t_1}{2} + \lambda_2(1-t_2) \right. \\ &\quad + \log\left\{\lambda_1 - \exp\left\{\frac{\lambda_1}{2} - \lambda_2 + 1\right\}\left(\frac{\lambda_1}{2} + \lambda_2 - 1\right)\right\} \\ &\quad \left. - \log\left(\frac{\lambda_1}{2} + \lambda_2 - 1\right) - \log\left(-\frac{\lambda_1}{2} + \lambda_2 - 1\right)\right]. \end{aligned}$$

Although the two first derivatives of $K_n(\lambda,t)$ are complicated, they can be obtained by a software for symbolic computation (e.g., Maple). The following results illustrate the very good accuracy of the saddlepoint approximation. Figure 1 shows how both the Lugannani and Rice (2) and Barndorff–Nielsen (4) approximations err for the case of four observations (or $n = 5$ spacings). We notice that all of the errors displayed are very small, even though the Lugannani and Rice formula generally outperforms the Barndorff–Nielsen formula. The exact probabilities are taken from Russell and Levitin (1995). Figure 2 emphasizes the tail behavior of the saddlepoint approximation and confirms numerically that the saddlepoint approximation has a bounded relative error. Table 1 also presents the approximation to the cumulative distribution function obtained by numerical integration of the saddlepoint density function. All three possibilities given by Lugannani and Rice, Barndorff–Nielsen, and integrated density lead to highly accurate results.

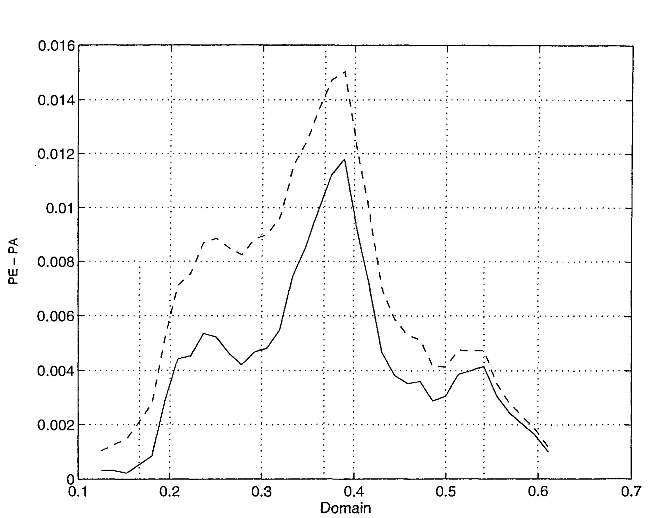

Figure 1. Rao Test Statistic; $n = 5$. $P_E - P_A$ is the error in the cumulative distribution; P_E, the exact cumulative distribution; P_A, the approximated cumulative distribution. Lugannani and Rice approximation (———) and Barndorff–Nielsen approximation (- - -). The dotted vertical lines indicate the .025 quantile (left), the median (middle), and the .975 quantile (right).

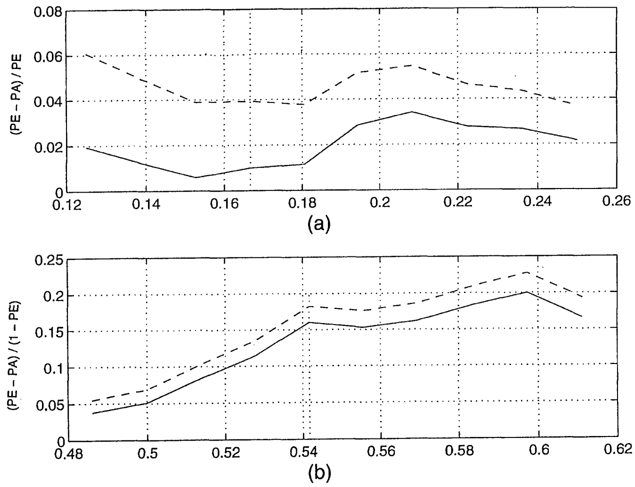

Figure 2. Rao Test Statistic, $n = 5$: Relative Errors in the Cumulative Distribution. Lugannani and Rice approximation (———) and Barndorff–Nielsen approximation (- - -). (a) Left tail, $(P_E - P_A)/P_E$; (b) Right tail, $(P_E - P_A)/(1 - P_E)$. P_E is the exact cumulative distribution; P_A, the approximated cumulative distribution. The dotted vertical lines indicate the .025 quantile (a) and the .975 quantile (b).

Example 2: Log Spacing Test. The choice of the score function $h(x) = \log(x)$ in (8) defines an alternative test statistic discussed by Darling (1953). This statistic has been shown to be the best with respect to Bahadur efficiency (see Zhao and Jammalamadaka 1989). It is also asymptotically normally distributed,

$$\sqrt{n}(T_n + \gamma) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\pi^2}{6} - 1\right),$$

where $\gamma = \int_0^\infty \log\{x\}\exp\{-x\}\,dx$ is Euler's constant. For $\lambda_2 < 1$, the joint cumulant generating function (1) for the

Table 1. Rao Test Statistic; Exact Cumulative Distribution (P_E), Lugannani and Rice (P_{LR}), Barndorff–Nielsen (P_{BN}), and Integrated Saddlepoint Density (P_{ID}) Approximations; $n = 10$

t_1	$P_E[T_n < t_1]$	$P_{LR}[T_n < t_1]$	$P_{BN}[T_n < t_1]$	$P_{ID}[T_n < t_1]$
.139	.001	.001	.001	0
.167	.004	.004	.004	.004
.194	.015	.015	.015	.015
.222	.042	.041	.041	.043
.250	.093	.093	.092	.096
.278	.178	.176	.175	.182
.306	.294	.292	.291	.300
.333	.433	.430	.429	.439
.361	.577	.580	.578	.583
.389	.708	.706	.705	.713
.417	.815	.813	.812	.819
.444	.892	.891	.890	.895
.472	.943	.942	.941	.945
.500	.972	.972	.971	.973
.528	.988	.987	.987	.989
.556	.995	.995	.995	.996
.583	.998	.998	.998	.999
.611	.999	.999	.999	1.000

log statistic is given by

$$K_n(\boldsymbol{\lambda}, \mathbf{t}) = n[-\lambda_1 t_1 - \lambda_2 t_2 - (1+\lambda_1)\log\{1-\lambda_2\} + \log\{\Gamma(1+\lambda_1)\}].$$

The derivatives of $K_n(\lambda, t)$ are functions involving mainly polygamma functions. Figure 3 compares our saddlepoint approximation (based on the Lugannani and Rice formula) and the normal approximation to the exact distribution obtained by 10^5 Monte Carlo simulations. The curves are transformed into the logit scale to emphasize the tail behavior. Although the saddlepoint clearly matches the exact curve overall, it can also be seen that the normal approximation is generally misleading. Figure 4 gives an idea of how fast the density of the log statistic converges toward normality. The plotted curves are saddlepoint approximations, and at $n = 15$ we are still some distance from normality.

Example 3: Greenwood Test Statistic. The choice of the score function $h(x) = x^2$ into (8) defines the so-called Greenwood test statistic. It is also asymptotically normally distributed,

$$\sqrt{n}(T_n - 2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4).$$

To compute our saddlepoint approximation, we need the joint cumulant generating function (1), which is given by

$$K_n(\boldsymbol{\lambda}, \mathbf{t}) = \begin{cases} n\left[\lambda_1 t_1 - \lambda_2 t_2 + \frac{(\lambda_2-1)^2}{4\lambda_1} + \frac{1}{2}\log\left\{\frac{\pi}{\lambda_1}\right\} + \log\left\{\Phi\left(\frac{\lambda_2-1}{\sqrt{2\lambda_1}}\right)\right\}\right], & \text{if } \lambda_1 < 0, \\ n[-\lambda_2 t_2 - \log\{1-\lambda_2\}], & \text{if } \lambda_1 = 0. \\ & \text{and } \lambda_2 < 1. \end{cases}$$

Table 2 shows the good accuracy of the saddlepoint approximation compared with the exact values taken from Burrows (1979). As was the case with the Rao test statistic,

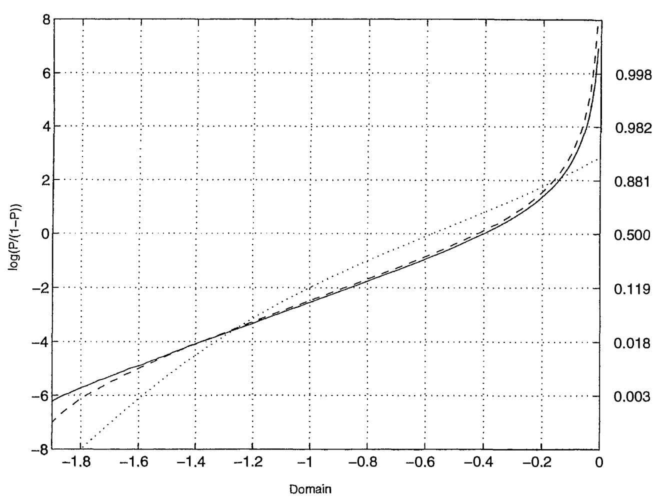

Figure 3. Log Test Statistic, n = 5: Cumulative Distribution. The scale on the right axis is the cumulative distribution P; that on the left axis is log{P/(1 − P)}. Exact from 10⁵ Monte Carlo simulations (———), Lugannani and Rice (– – –), and asymptotic normal (· · ·).

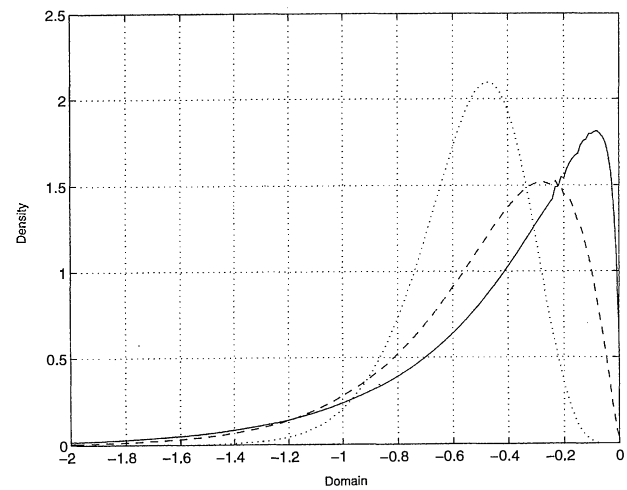

Figure 4. Log Test Statistic: Saddlepoint Approximations to Density Functions (——— n = 3; – – –, n = 5; · · ·, n = 15).

the Barndorff–Nielsen approximation is slightly less accurate than the Lugannani and Rice approximation.

Example 4: A Locally Most Powerful (LMP) Spacings Test. Holst and Rao (1981) considered LMP tests based on spacings. In particular, the nonsymmetric statistic defined by the score function

$$h_i(nD_i) = \Phi^{(-1)}\left(\frac{i}{n+1}\right) nD_i$$

in (7) provides the LMP test based on spacings for testing whether the data are from a $\mathcal{N}(0,1)$ distribution against the alternative that they are from a $\mathcal{N}(\mu, 1)$ distribution, $\mu \in \mathbb{R}$. This test has power at the usual $n^{-1/2}$ alternatives. The joint cumulant generating function is given by

$$K_n(\boldsymbol{\lambda}, \mathbf{t}) = -\lambda_1 t_1 - n\lambda_2 t_2 - \sum_{i=1}^{n} \log\left\{\lambda_1 \Phi^{(-1)}\left(\frac{i}{n+1}\right) - \lambda_2 + 1\right\},$$

which exists for

$$\lambda_2 < \begin{cases} 1 + \lambda_1 \Phi^{(-1)}\left(\frac{1}{n+1}\right), & \text{if } \lambda_1 \le 0 \\ 1 + \lambda_1 \Phi^{(-1)}\left(\frac{n}{n+1}\right), & \text{if } \lambda_1 > 0. \end{cases}$$

Table 3 shows the left tail of this LMP statistic for the case of five observations (i.e., $n = 6$ spacings). The conditional saddlepoint approximation based on steps 1 and 2 is ex-

Table 2. Greenwood Test Statistic; Exact Left-Tail Probabilities (P_E), Lugannani and Rice (P_{LR}), and Barndorff–Nielsen (P_{BN}) Approximations; n = 4

t_1	$P_E[T_n < t_1]$	$P_{LR}[T_n < t_1]$	$P_{BN}[T_n < t_1]$
1.032	.010	.010	.008
1.100	.050	.057	.051
1.156	.100	.107	.098
1.252	.200	.204	.191
1.328	.300	.302	.285
1.408	.400	.398	.377
1.496	.500	.500	.475

Table 3. LMP Spacings Test; Exact Cumulative Probabilities (P_E), Lugannani and Rice Approximations (P_{LR}); n = 6

t_1	$P_E[T_n < t_1]$	$P_{LR}[T_n < t_1]$
−4.500	.002	.002
−4.000	.005	.006
−3.500	.013	.014
−3.000	.030	.031
−2.500	.060	.061
−2.000	.107	.108
−1.500	.175	.177
−1.000	.266	.268
−.500	.377	.379

tremely close to the exact distribution as obtained through 10^5 simulations.

3.2 Tests Based on Grouped Data

Suppose that we have a sample $X_1, \dots, X_n$ of iid observations with underlying continuous distribution F, and that the support of F has been divided into m nonoverlapping intervals or classes. Let p_j denote the probability that a sample value belongs to the jth interval, $j = 1, \dots, m$. Let S_j denote the number of sample values belonging to the jth class interval, $j = 1, \dots, m$. Under H_0: $\mathbf{p} = \mathbf{p}_0$, where $\mathbf{p} = (p_1, \dots, p_m)$ and $\mathbf{p}_0 = (p_{01}, \dots, p_{0m})$ is a specified vector, $S_1, \dots, S_m$ are distributed according to the multinomial probability distribution with parameters $(n; \mathbf{p}_0)$. Our conditional saddlepoint can be used by means of the following conditional representation of multinomial random variables. If $W_1, \dots, W_m$ are independent Poisson random variables with probability distribution function $P[W_j = w] = e^{-\xi p_{0j}}(\xi p_{0j})^w / w!, j = 1, \dots, m$, then, for all $\xi \in \mathbb{R}_+^*$, it is easy to check that

$$\{S_1, \dots, S_m\} \sim \{W_1, \dots, W_m\} \left| \sum_{i=1}^{m} W_i = n. \right. \qquad (9)$$

Example 2: Likelihood Ratio Test. The generalized likelihood ratio test statistic for this situation is derived from

$$\Lambda = \frac{\sup_{\mathbf{p} \in \Pi_0} \left\{ \frac{n!}{S_1! \dots S_m!} p_1^{S_1}, \dots, p_m^{S_m} \right\}}{\sup_{\mathbf{p} \in \Pi} \left\{ \frac{n!}{S_1! \dots S_m!} p_1^{S_1}, \dots, p_m^{S_m} \right\}},$$

where Π is the entire parametric space and Π_0 is the parametric space restricted by H_0. Without loss of generality (by the probability integral transform), when all of the $\{p_{0j}\}$ are equal, $-2\log\{\Lambda\}$ is equal to $2\sum_{j=1}^{m} S_j \log\{S_j\}$ plus a constant term, and a large value of this latter statistic provides evidence against H_0 with all $\{p_{0j}\}$ equal.

From classical theory, the asymptotic distribution of $-2\log\Lambda$, as $n \to \infty, m$ fixed, is a χ^2_{m-1}; on the other hand, if both m and $n \to \infty$ and $n/m \to c^*, 1 < c^* < \infty$, then we have a normal limiting distribution. In this more practical case, where both m and n are large, by defining the quantities $c = n/m$,

$$T_{1m} = \frac{1}{m} \sum_{j=1}^{m} W_j \log\{W_j\}$$

and

$$T_{2m} = \frac{1}{m} \sum_{j=1}^{m} c^{-1} W_j,$$

the saddlepoint approximation to the distribution of $(T_{1m} | T_{2m} = 1)$ can be obtained by steps 1 and 2 of Section 2, which provides an approximation for the original statistic $T_m = m^{-1} \sum_{j=1}^{m} S_j \log\{S_j\}$. The score functions are now $\psi_1(w, t_1) = w\log\{w\} - t_1$ and $\psi_2(w, t_2) = w - ct_2$, and $\psi_{kj} = \psi_k, j = 1, \dots, m, k = 1, 2$. For $-\infty < \lambda_1 < 1$, the joint cumulant generating function (1) is given by

$$K_m(\boldsymbol{\lambda}, \mathbf{t}) = m \left[-\lambda_1 t_1 - \xi\lambda_2 t_2 - c + \log\left\{ 1 + \sum_{k=1}^{\infty} \frac{k^{k\lambda_1} e^{k\lambda_2} \xi^k}{\Gamma(k+1)} \right\} \right].$$

Example 3: Pearson Chi-Squared Test. The chi-squared test provides a different statistic for testing H_0: $\mathbf{p} = \mathbf{p}_0$. It is defined by

$$T_m = \frac{1}{m} \sum_{j=1}^{m} \frac{(S_j - np_{0j})^2}{np_{0j}}.$$

A large value of T_m gives evidence for rejecting H_0: $\mathbf{p} = \mathbf{p}_0$. The asymptotic distribution of T_m, as $n \to \infty, m$ fixed, is a χ^2_{m-1}, whereas if both m and $n \to \infty, n/m \to c^*, 1 < c^* < \infty$, we have a normal limiting distribution. Again by the conditional Poisson representation (9), we can reexpress T_m as $(T_{1m} | T_{2m} = 1)$, where

$$T_{1m} = \frac{1}{m} \sum_{j=1}^{m} \frac{(W_j - c)^2}{c}$$

and

$$T_{2m} = \frac{1}{m} \sum_{j=1}^{m} c^{-1} W_j.$$

We consider this case where both m and n are large and obtain the saddlepoint approximation by steps 1 and 2. The cumulant generating function for $c = n/m$ and $\lambda_1 \leq 1$ is given by

$$K_m(\boldsymbol{\lambda}, \mathbf{t}) = m \left[-\lambda_1 t_1 - c\lambda_2 t_2 - \xi + \log\left\{ \exp\{\lambda_1 c\} + \sum_{k=1}^{\infty} \frac{\exp\left\{\lambda_1 \frac{(k-c)^2}{c} + \lambda_2 k\right\} \xi^k}{\Gamma(k+1)} \right\} \right].$$

Note that the conditional Poisson representation (9) could be used for specific alternatives, so that we could compute the distribution of T_m under alternatives; that is, obtain the power function of T_m.

3.3 Conditional Tests

A method for eliminating nuisance parameters in a test statistic is by conditioning on statistics that are sufficient (under the null hypothesis) for these nuisance parameters. If the distribution of the test statistic T_n depends on a nuisance parameter whose sufficient statistic is S_n, then a p value can be obtained by $p = P[T_n < t_n | S_n = s_n]$, assuming that small T_n indicate evidence against H_0, where t_n and s_n are T_n and S_n of an observed sample. This approach, referred to as conditional testing, can advantageously be used with our saddlepoint approximation.

Example 4: Skewness-to-Variance Test. Suppose that we wish to test whether $W_1, \ldots, W_n$ are counts of events of a homogeneous Poisson process with unknown nuisance parameter $\xi \in \mathbb{R}_+^*$. It is known that the Poisson distribution has all cumulants equal to the parameter and that it is the only distribution with this property, so that ratios of empirical cumulants are used to test various types of departures from the hypothesized distribution. When a departure of variability is of concern, the empirical variance over the mean is a possible test statistic; when a departure in skewness is of concern, the empirical third cumulant over the second,

$$T_n = \frac{\sum_{i=1}^n (W_i - \bar{W})^3}{\sum_{i=1}^n (W_i - \bar{W})^2},$$

where $\bar{W} = n^{-1} \sum_{i=1}^n W_i$ is an adequate test statistic. The nuisance parameter ξ can be eliminated by conditioning on $\bar{W}$, its minimal sufficient statistic, and it can be shown that $\{W_1, \ldots, W_n\} | \sum_{i=1}^n W_i$ is multinomial with equal cell probabilities. For $\lambda_1 \leq 0$, the joint cumulant generating function is given by

$$K_n(\mathbf{t}, \boldsymbol{\lambda}) = n \left[-\lambda_1(t_1 t_2^2 + t_2^3) - \lambda_2 t_2 - \xi + \log \sum_{k=0}^{\infty} \frac{\exp\{\lambda_1 k^3 - \lambda_1(t_1 + 3t_2)k^2 + [\lambda_1(2t_1 t_2 + 3t_2^2) + \lambda_2]k\}\xi^k}{\Gamma(k+1)} \right].$$

The data selected are yearly fatal accidents suffered by scheduled American domestic-operated passenger aircraft from 1948 until 1961 (see Pyke 1965). The sample value of T_n is -1.277, and the alternative hypothesis is a left overskewness. The conditional p value obtained by the saddlepoint approximation is .092, which is close to the exact p value .075 obtained by 10^5 Monte Carlo simulations. The approximation to the whole left tail is shown in Table 4, which reflects a quite good accuracy of our saddlepoint approximation based on steps 1 and 2.

4. TWO-SAMPLE TESTS

Consider a first sample of $(m-1)$ iid random variables $X_1 \ldots, X_{m-1}$, with underlying continuous distribution F defined on $A \subset \mathbb{R}$, and a second sample of n iid random variables $Y_1 \ldots, Y_n$, with underlying distribution G, also defined on $A \subset \mathbb{R}$. The general two-sample problem is to test the null hypothesis H_0: $F = G$. Define the random variables

$$S_j = \sum_{i=1}^n I\{Y_i \in [X_{(j-1)}, X_{(j)})\}, \qquad j = 1, \ldots, m,$$

where, for convenience, we take $X_{(0)} \stackrel{\text{def}}{=} \inf\{A\}$ and $X_{(m)} \stackrel{\text{def}}{=} \sup\{A\}$. The numbers $S_1, \ldots, S_m$ may be called the "spacing frequencies," because they correspond to the frequencies or counts of the $\{Y_i\}$ that fall in between successive $X_{(j)}$. In fact, if $R(X_{(k)})$ denotes the rank of the kth largest $\{X_j\}$ in the combined sample, $k = 1, \ldots, m$, then it is easily seen that $R(X_{(k)}) = \sum_{j=1}^k (S_j + 1)$ or $S_k = R(X_{(k)}) - R(X_{(k-1)}) - 1, k = 1, \ldots, m$, so that the $\{S_j\}$ are also the "rank differences." Let $h(\cdot)$ and $h_j(\cdot), j = 1, \ldots, m$, be real-valued functions satisfying certain regularity conditions (see Holst and Rao 1980, sec. 2, cond. A), and define the general classes of test statistics

$$T_\nu^* = \sum_{j=1}^m h_j(S_j) \tag{10}$$

and

$$T_\nu = \frac{1}{m} \sum_{j=1}^m h(S_j), \tag{11}$$

which represent the nonsymmetric and the symmetric test statistics based on the rank spacings. We consider the asymptotic properties of these statistics based on two samples when both m and n tend to infinity; formally through nondecreasing sequences of positive integers $\{m_\nu\}$ and $\{n_\nu\}$ such that, as $\nu \to \infty$,

$$m_\nu \to \infty, n_\nu \to \infty$$

and

$$\frac{m_\nu}{n_\nu} = \rho_\nu \to \rho, \qquad 0 < \rho < \infty.$$

First-order asymptotics under the null hypothesis H_0 for the more general nonsymmetric test (10) is stated in corollary 3.1 of Holst and Rao (1980). If $V_1, \ldots, V_m$ are iid geo-

Table 4. Skewness-to-Variance Test Statistic; Exact Cumulative Probabilities (P_E), Lugannani and Rice Approximations (P_{LR}): American Scheduled Aircraft Accidents, 1948–1961 (n = 14).

t_1	$P_E[T_n < t_1]$	$P_{LR}[T_n < t_1]$
−2.600	.008	.011
−2.400	.011	.016
−2.200	.016	.023
−2.000	.023	.032
−1.800	.033	.044
−1.600	.046	.059
−1.400	.063	.078
−1.277	.075	.092
−1.200	.085	.102
−1.000	.111	.132
−.800	.144	.168
−.600	.186	.210

metric random variables with probability distribution function

$$P[V_1 = k] = (1-p)^k p, \qquad k = 0, 1, 2, \ldots, \tag{12}$$

then, for $p = 1/(\rho + 1)$ and under H_0,

$$\sum_{j=1}^{m} h_j(S_j) \xrightarrow{\mathcal{D}} \mathcal{N}(\mu, \sigma^2), \tag{13}$$

where $\mu = E(\sum_{j=1}^m h_j(V_j))$ and $\sigma^2 = \text{var}(\sum_{j=1}^m h_j(V_j) - \beta \sum_{i=j}^m V_j)$, in which β is the regression coefficient given by $\beta = \text{cov}(\sum_{j=1}^m h_j(V_j), \sum_{j=1}^m V_j)/\text{var}(\sum_{j=1}^m V_j)$. The same conditioning idea used for obtaining the first-order approximation (13) can be exploited for the construction of our saddlepoint approximation. Under H_0, the m spacing frequencies have the same distribution as m iid geometric random variables conditioned to sum up to n. Namely, if $V_1, \ldots, V_m$ are iid geometric random variables with probability distribution function given by (12), then, for all $p \in (0, 1)$, it is easily verified that

$$\{S_1, \ldots, S_m\} \sim \{V_1, \ldots, V_m\} \,\Bigg|\, \sum_{j=1}^m V_j = n.$$

By defining

$$T_{1\nu}^* = \sum_{j=1}^m h_j(V_j)$$

and

$$T_{2\nu} = \frac{1}{m} \sum_{j=1}^m \rho_\nu V_j,$$

the conditional distribution of $(T_{1\nu}^* | T_{2\nu} = 1)$ can be obtained again by steps 1 and 2 of Section 2. The next examples are derived from the following optimality result. Consider G_m, a smooth sequence of distribution functions converging toward F, as $m \to \infty$. It turns out that the asymptotically most powerful test for the null hypothesis H_0 against the sequence of simple alternatives,

$$A_m\colon\ G = G_m,$$

is to reject H_0 when

$$\sum_{j=1}^m l\left(\frac{j}{m+1}\right) S_j > c, \tag{14}$$

where $l(\cdot)$ is the derivative of $L(u) = \lim_{m\to\infty} m^{1/2} [G_m(F^{(-1)}(u)) - u], 0 \le u \le 1$ (see Holst and Rao 1980, thm. 3.2, for further details).

Example 5: van der Waerden/Normal Score Test. Suppose that we wish to test $H_0\colon F = G$ against the sequence of translation alternatives $G(x) = G_m(x) \stackrel{\text{def}}{=} F(x - \theta m^{-1/2})$ for all $x \in A$. If $f'(x) = F''(x)$ exists and is almost everywhere continuous, then, for $0 \le u \le 1$ and, at the continuity points of f', we have $l(u) = -\theta f'[F^{(-1)}(u)]/f[F^{(-1)}(u)]$. From (14), when $F = \Phi$, the asymptotically most powerful test is the van der Waerden or normal score test statistic, given by

$$T_\nu^* = \sum_{j=1}^m \Phi^{(-1)}\left(\frac{j}{m+1}\right) S_j.$$

The joint cumulant generating function is given by

$$K_\nu(\boldsymbol{\lambda}, \mathbf{t}) = m(\log\{p\} - \lambda_2 t_2) - \lambda_1 t_1 + \sum_{j=1}^m \log\left\{1 + \sum_{k=1}^\infty \exp\left\{k\left[\lambda_1 \Phi^{(-1)}\left(\frac{j}{m+1}\right) + \lambda_2 \rho_\nu\right]\right\}(1-p)^k\right\},$$

where the infinite sum converges for

$$\lambda_2 < \begin{cases} -\frac{n}{m}\left[\Phi^{(-1)}\left(\frac{1}{m+1}\right)\lambda_1 + \log\left\{\frac{1}{1-p}\right\}\right], & \text{if } \lambda_1 \le 0 \\ -\frac{n}{m}\left[\Phi^{(-1)}\left(\frac{m}{m+1}\right)\lambda_1 + \log\left\{\frac{1}{1-p}\right\}\right], & \text{if } \lambda_1 > 0. \end{cases}$$

Example 6: Wilcoxon/Mann–Whitney Test. This test is known to be the locally most powerful rank test for detecting change in location in a logistic distribution. If the $\{R(X_j)\}$ are the ranks of the first sample in the combined sample, then this test statistic is simply $\sum_{j=1}^m R(X_j)$. But because $R(X_{(k)}) = \sum_{j=1}^k S_j + k, k = 1, \ldots, m$, it can be rewritten as $\sum_{j=1}^m R(X_{(j)}) = m + \sum_{j=1}^m (m + 1 - j) S_j$, so that the test statistic can be reexpressed by

$$T_\nu^* = \sum_{j=1}^m \left(1 - \frac{j}{m+1}\right) S_j.$$

The joint cumulant generating function is given by

$$K_\nu(\boldsymbol{\lambda}, \mathbf{t}) = m(\log\{p\} - \lambda_2 t_2) - \lambda_1 t_1 + \sum_{j=1}^m \log\left\{1 + \sum_{k=1}^\infty \exp\left\{k\left[\lambda_1\left(1 - \frac{j}{m+1}\right) + \lambda_2 \rho_\nu\right]\right\}(1-p)^k\right\},$$

and convergence of the sum is for

$$\lambda_2 < \begin{cases} -\frac{n}{m}\left[\frac{1}{m+1}\lambda_1 + \log\left\{\frac{1}{1-p}\right\}\right], & \text{if } \lambda_1 \le 0 \\ -\frac{n}{m}\left[\frac{m}{m+1}\lambda_1 + \log\left\{\frac{1}{1-p}\right\}\right], & \text{if } \lambda_1 > 0. \end{cases}$$

Example 7: Savage/Exponential Score Test. We wish to test $H_0\colon F = G$ against the sequence of scale alternatives $G(x) = G_m(x) \stackrel{\text{def}}{=} F(x[1 - \theta m^{-1/2}])$, for all $x \in A$. If the density $f(x) = F'(x)$ is continuous, then $l(u) = -\theta[1 + f'(F^{(-1)}(u)) F^{(-1)}(u)/f(F^{(-1)}(u))]$. As an example, when $F(x) = 1 - \exp\{-x\}$, we have $l(u) = -\theta(1 + \log\{1 - u\})$, and from (14) the asymptotically most powerful test statistic is given by

$$T_\nu^* = \sum_{j=1}^{m} \log\left\{1 - \frac{j}{m+1}\right\} S_j.$$

The joint cumulant generating function is given by

$$\begin{aligned} K_\nu(\boldsymbol{\lambda}, \mathbf{t}) &= m(\log\{p\} - \lambda_2 t_2) - \lambda_1 t_1 \\ &\quad + \sum_{j=1}^{m} \log\left\{1 + \sum_{k=1}^{\infty}\left(1 - \frac{j}{m+1}\right)^{\lambda_1 k} \right. \\ &\qquad \left. \times \exp\{\lambda_2 \rho_\nu k\}(1-p)^k\right\}, \end{aligned}$$

where the infinite sum converges for

$$\lambda_2 < \begin{cases} \frac{n}{m}\left[\log\{m+1\}\lambda_1 + \log\left\{\frac{1}{1-p}\right\}\right], & \text{if } \lambda_1 \le 0 \\ \frac{n}{m}\left[\log\left\{\frac{m+1}{m}\right\}\lambda_1 + \log\left\{\frac{1}{1-p}\right\}\right], & \text{if } \lambda_1 > 0. \end{cases}$$

5. DISCUSSION

This article illustrates that the conditional saddlepoint approximations discussed here can be advantageously used in various testing problems. It leads to accurate inference and often requires less computing time than Monte Carlo simulation. An important advantage of saddlepoint approximations, with respect to Monte Carlo simulation, is that they lead to accurate numerical results through analytical formulas, which can be useful for deriving other related properties. For example, the saddlepoint approximation can be used for computing the "tail area influence function," which describes the normalized influence on a tail area of a small amount of contamination at a fixed point. That is, for a statistic T_n depending on observations with underlying distribution F, it is defined by

$$\begin{aligned} &\mathrm{TAIF}(x, t, T_n, F) \\ &\quad = \lim_{\varepsilon \to 0, \varepsilon > 0} \varepsilon^{-1}[P_{F_{\varepsilon,x}}(T_n \ge t) - P_F(T_n \ge t)], \end{aligned}$$

for all x in the sample space, where $F_{\varepsilon,x} = (1-\varepsilon)F + \varepsilon\Delta_x, \Delta_x$ being Dirac's distribution with mass 1 at x (see Field and Ronchetti 1985).

The complexity of our formulas shows the importance of automatic symbolic computation, which allows one to compute the derivatives of the cumulant generating functions in a reasonable amount of time. An important illustration of the connection between statistical theory and symbolic computation is the treatment of asymptotic expansions by Andrews and Stafford (1993).

[Received January 1997. Revised September 1998.]

REFERENCES

Andrews, D. F., and Stafford, J. E. (1993), "Tools for the Symbolic Computation of Asymptotic Expansions," *Journal of the Royal Statistical Society*, Ser. B, 55, 613–627.

Barndorff–Nielsen, O. E., and Cox, D. R. (1989), *Asymptotic Techniques for Use in Statistics*, London: Chapman and Hall.

Batschelet, E. (1981), *Circular Statistics in Biology*, New York: Academic Press.

Booth, J. G., and Butler, R. W. (1990), "Randomization Distributions and Saddlepoint Approximations in Generalized Linear Models," *Biometrika*, 77, 787–796.

Burrows, P. M. (1979), "Selected Percentage Points of Greenwood's Statistic," *Journal of the Royal Statistical Society*, Ser. A, 142, 256–258.

Daniels, H. E. (1954), "Saddlepoint Approximations in Statistics," *The Annals of Mathematical Statistics*, 25, 631–650.

——— (1987), "Tail Probability Approximations," *International Statistical Review*, 55, 37–48.

Darling, D. A. (1953), "On a Class of Problems Related to the Random Division of an Interval," *The Annals of Mathematical Statistics*, 24, 239–253.

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

DiCiccio, T. J., Martin, M. A., and Young, G. A. (1993), "Analytical Approximations to Conditional Distribution Functions," *Biometrika*, 80, 781–790.

Field, C. (1982), "Small Sample Asymptotic Expansions for Multivariate M-Estimates," *The Annals of Statistics*, 10, 672–689.

Field, C., and Ronchetti, E. (1985), "A Tail Area Influence Function and Its Application to Testing," *Communications in Statistics, Part C*, 4, 19–41.

——— (1990), *Small Sample Asymptotics*, Lecture Notes Monograph Series, Vol. 13, Hayward, CA: Institute of Mathematical Statistics.

Field, C. A., and Tingley, M. A. (1997), "Small Sample Asymptotics: Applications in Robustness," in *Handbook of Statistics*, Vol. 15, eds. Maddala and Rao, Amsterdam: North-Holland, pp. 513–536.

Gatto, R., and Ronchetti, E. (1996), "General Saddlepoint Approximations of Marginal Densities and Tail Probabilities," *Journal of the American Statistical Association*, 91, 666–673.

Holst, L., and Rao, J. S. (1980), "Asymptotic Theory for Some Families of Two-Sample Nonparametric Statistics," *Sankhyā*, Ser. A, 42, 19–52.

——— (1981), "Asymptotic Spacings Theory With Applications to the Two-Sample Problem," *The Canadian Journal of Statistics*, 9, 79–89.

Jensen, J. L. (1991), "Uniform Saddlepoint Approximations and Log-Concave Densities," *Journal of the Royal Statistical Society*, Ser. B, 53, 157–172.

——— (1992), "The Modified Signed Likelihood Statistic and Saddlepoint Approximations," *Biometrika*, 79, 693–703.

——— (1995), *Saddlepoint Approximations*, New York: Oxford University Press.

Jing, B., and Robinson, J. (1994), "Saddlepoint Approximations for Marginal and Conditional Probabilities of Transformed Variables," *The Annals of Statistics*, 22, 1115–1132.

Lugannani, R., and Rice, S. (1980), "Saddlepoint Approximation for the Distribution of the Sum of Independent Random Variables," *Advances in Applied Probability*, 12, 475–490.

Pyke, R. (1965), "Spacings," *Journal of the Royal Statistical Society*, Ser. B, 27, 395–449.

Rao, J. S. (1969), *Some Contributions to the Analysis of Circular Data*, doctoral thesis, Indian Statistical Institute, Calcutta.

——— (1976), "Some Tests Based on Arc Lengths for the Circle," *Sankhyā*, Ser. B, 4, 329–338.

Reid, N. (1988), "Saddlepoint Methods and Statistical Inference," *Statistical Science*, 3, 213–238.

Russell, G. S., and Levitin, D. J. (1995), "An Expanded Table of Probability Values for Rao's Spacing Test," *Communications in Statistics, Simulation and Computation*, 24, 879–888.

Skovgaard, I. M. (1987), "Saddlepoint Expansions for Conditional Distributions," *Journal of Applied Probability*, 24, 875–887.

Strawderman, R. L., Casella, G., and Wells, M. T. (1996), "Practical Small-Sample Asymptotics for Regression Problems," *Journal of the American Statistical Association*, 91, 643–654.

Wang, S. (1993), "Saddlepoint Approximations in Conditional Inference," *Journal of Applied Probability*, 30, 397–404.

Zhao, X., and Jammalamadaka, S. R. (1989), "Bahadur Efficiencies of Spacing Tests for Goodness of Fit," *Annals of the Institute of Mathematical Statistics*, 41, 541–553.