The Annals of Statistics
1987, Vol. 15, No. 1, 314–324

STANDARDIZED LOG-LIKELIHOOD RATIO STATISTICS FOR MIXTURES OF DISCRETE AND CONTINUOUS OBSERVATIONS

BY J. L. JENSEN

University of Aarhus

When the log-likelihood statistic is divided by its mean, or an approximation to its mean, the limiting chi-squared distribution is often correct to order $n^{-3/2}$. Similarly, when the signed version of the likelihood ratio statistic is standardized with respect to its mean and variance the normal approximation is correct to order $n^{-3/2}$. Proofs for these statements have been given in great generality in the literature for the case of continuous observations. In this paper we consider cases where the minimal sufficient statistic is partly discrete and partly continuous. In particular, we consider testing problems associated with censored exponential life times.

1. Introduction. In this paper we shall be concerned with likelihood ratio statistics for situations where the minimal sufficient statistic is partly discrete and partly continuous. Discussing first the case of a one-dimensional parameter, we let w be the usual likelihood ratio statistic

$$w = 2\{l(\hat{\theta}) - l(\theta)\},$$

where l is the log-likelihood function. Letting R be the signed version,

$$R = \pm\sqrt{w},$$

with the sign determined by $\hat{\theta} - \theta$, we shall be interested in standardized statistics of the form

$$w' = \frac{w}{1 + B/n} \tag{1}$$

and

$$R' = \frac{R - \beta/\sqrt{n}}{\sqrt{1 + 1/n(B - \beta^2)}}, \tag{2}$$

where

$$Ew = 1 + \frac{B}{n} + O(n^{-3/2})$$

and

$$ER = \frac{\beta}{\sqrt{n}} + O\left(\frac{1}{n}\right).$$

In recent years there has been considerable interest in these standardized

Received July 1985; revised June 1986.
AMS 1980 *subject classifications*. Primary 62E20; secondary 62F05.
Key words and phrases. Censored life times, conditional expansion, transformed expansion.

statistics, and for a large class of problems with continuous variates it has been shown that the standardizations improve the chi-squared and normal approximations such that the error is of order $n^{-3/2}$; see Chandra and Ghosh (1979), Barndorff-Nielsen and Cox (1984), McCullagh (1984), Barndorff-Nielsen (1986) and DiCiccio (1983, 1984). These approximations may be summarized by stating that the distribution of R has the expansion

$$(3) \qquad P(R \le r) = \Phi(r) - \frac{\beta}{\sqrt{n}}\varphi(r) - \frac{B}{2n} r\varphi(r) + O(n^{-3/2}).$$

For a curved exponential family with continuous variates the proof of (3) is based on the following two parts. First an Edgeworth expansion is derived for the distribution of the minimal sufficient statistic, which is valid to order $n^{-3/2}$. Next R is replaced by its Taylor expansion in terms of the minimal sufficient statistic, and using this Taylor expansion the Edgeworth expansion is transformed to give an expansion for the distribution of R, see Chandra and Ghosh (1979) and McCullagh (1984). That the terms of order $n^{-1/2}$ and n^{-1} take the explicit form given in (3) is based on the fact that certain relations exist between the coefficients of the Edgeworth expansion for the minimal sufficient statistic and the coefficients of the Taylor expansion of R in terms of the minimal sufficient statistic. The case of an exponential family of order one is very illustrative and another instance of such relations may be seen in Jensen (1986).

When the minimal sufficient statistic is partly discrete and partly continuous, the formal Edgeworth expansion of its distribution is no longer valid to order $n^{-3/2}$, say. We can therefore not use the method outlined above to obtain the result (3). Instead we shall obtain an expansion of the distribution of R conditioned on the discrete part and then integrate out the conditioning variable. The use of a conditional argument to establish an Edgeworth expansion is not new. This approach has been extremely successful when dealing with linear rank statistics, see, e.g., Albers, Bickel and van Zwet (1976), Bickel and van Zwet (1978) and Does (1983). The main difference in the problems encountered here and in the above-mentioned papers is that in the conditional distribution, that we consider, the mean does not tend to zero in probability. Because of this it is not possible to write the conditional expansion as a main term, which does not depend on the conditioning variable, and a number of smaller terms. Conditional expansions have been considered in the literature recently, both for the case where the conditioning variable is continuous [Does, Helmers and Klaassen (1984)] and where it is discrete [Hipp (1984)].

In Section 2 we prove a result that will enable us to establish (3) in the most simple i.i.d. cases. We shall prove an Edgeworth expansion for the one-dimensional statistic

$$\sqrt{n}\, g\left(\frac{1}{n}\sum_1^n X_i, \frac{1}{n}\sum_1^n K_i\right),$$

where $(X_1, K_1), \dots, (X_n, K_n)$ are i.i.d. with X and K both one-dimensional, X is a continuous variable and K is a lattice variable. Here $g(x, r)$ is a sufficiently smooth function of two variables. As an example we consider observations from an exponential distribution with fixed censoring time T. The variate X is then the censored life time and K is the indicator function for the event that the observation is censored. In Section 2 we also discuss the possibility of obtaining the result (3) in more general cases. The general cases cover both multi-dimensional parameters and testing problems with nuisance parameters. In Section 3 we find the coefficients β and B appearing in (3) for the case of varying censoring times and also for a two-sample problem involving a nuisance parameter. Finally, in Section 4 we give a numerical example.

2. Proof of the validity of the expansion. In this section we obtain an Edgeworth expansion for the one-dimensional statistic $\sqrt{n}\, g(\Sigma X_i/n, \Sigma K_i/n)$, where X and K are both one-dimensional and K is a lattice variable with minimal lattice equal to the integers. We make the following assumptions.

ASSUMPTION A.

(i) $EX = 0$, $\mathrm{Cov}(X, K) = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}$ with $\sigma^2 > 0$ and $|\rho| < 1$.

(ii) $E(e^{\tau K}) < \infty$ and $E(|x|^5 e^{\tau K}) < \infty$ for $-\tau_1 < \tau < \tau_2$ with $\tau_i > 0$.

(iii) For some $d > 0$ we may for all $c > 0$ find a $\delta < 1$ such that for $|t| > c$, $|\tau| \le d$ and all s

$$\left|E(e^{itX+(\tau+is)K})\right| \le \delta E(e^{\tau K}).$$

(iv) The function $g(x, r)$ is four times continuously differentiable for (x, r) in a region containing $(0, \mu)$, where $\mu = EK$.

(v) $\partial g(x, r)/\partial x|_{(0,\mu)} \ne 0$ and $g(0, \mu) = 0$.

Let $\psi(it, \tau + is) = E\{\exp[itX + (\tau + is)K]\}$ and define the cumulants $\gamma_{km} = \gamma_{km}(\tau)$ by

$$\gamma_{km}(\tau) = \frac{1}{i^{k+m}} \frac{\partial^{k+m} \ln \psi(it, \tau + is)}{\partial t^k \, \partial s^m}\bigg|_{(0,0)}.$$

In the following we let τ be the saddlepoint in the second variable determined by

$$\gamma_{01}(\tau) = r. \tag{4}$$

We first prove the following expansion which is of the mixed Edgeworth-saddlepoint type [Barndorff-Nielsen and Cox (1979)].

LEMMA 1. *There exist constants c_1 and c_2 such that*

$$\begin{aligned} P&\left(\frac{1}{\omega(\tau)}\left[\frac{1}{\sqrt{n}}\sum_1^n X_i - \sqrt{n}\,\gamma_{10}(\tau)\right] \le z \,|\, \bar{K} = r\right) \\ &= \Phi(z) + \frac{1}{\sqrt{n}}[\alpha_1(\tau)\varphi(z) + \alpha_2(\tau)\varphi''(z)] \\ &\quad + \frac{1}{n}[\beta_1(\tau)\varphi'(z) + \beta_2(\tau)\varphi^{(3)}(z) + \beta_3(\tau)\varphi^{(5)}(z)] + \theta\frac{c_2}{\sqrt{n}^{3}}, \end{aligned} \tag{5}$$

for $|\tau| \le c_1$ and with $|\theta| \le 1$. Here $\omega^2(\tau) = \gamma_{20} - \gamma_{11}^2/\gamma_{02}$ and the other coefficients may be obtained by following the guidelines of the proof.

PROOF. Define

$$H(t, s) = \frac{\psi(it, \tau + is)}{\psi(0, \tau)} e^{-isr - it\gamma_{10} + \omega^2 t^2/2}.$$

From Bartlett's (1938) representation of a conditional characteristic function we find, after shifting to the conjugate distribution determined by τ,

$$E\left(e^{it(1/\sqrt{n})\Sigma X_i} \middle| \overline{K} = r\right)$$

$$= e^{it\sqrt{n}\gamma_{10} - \omega^2 t^2/2} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} H\left(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}\right)^n ds \Big/ \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} H\left(0, \frac{s}{\sqrt{n}}\right)^n ds$$

$$(6) \quad = e^{it\sqrt{n}\gamma_{10} - \omega^2 t^2/2}\left\{1 + \int_0^t \left[\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} n \frac{\partial H(u/\sqrt{n}, s/\sqrt{n})}{\partial u} H\left(\frac{u}{\sqrt{n}}, \frac{s}{\sqrt{n}}\right)^{n-1} ds\right]\right.$$

$$\left. \Big/ \left[\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} H\left(0, \frac{s}{\sqrt{n}}\right)^n ds\right] du\right\}.$$

The expansion (5) is now obtained by a standard argument from (6) via a Taylor expansion of $\ln H(t, s)$ and the use of Esseen's smoothing theorem [Esseen (1945)]. The Taylor expansion is used for $|t| < c_3$ and $|s| < c_4$, say. For $|t| > c_3$ the characteristic function is bounded by the use of assumption A(iii) and a similar bound applies for $|s| > c_4$ due to the lattice character of K. □

From Lemma 1 we may obtain an expansion of the conditional distribution of $\sqrt{n}\, g(\bar{x}, \overline{K})$ given that $\overline{K} = r$. Without loss of generality we assume that $\partial g(x, \mu)/\partial x|_0 > 0$. Let in the following

$$g_k(r) = \frac{\partial^k g(x, r)}{\partial x^k}\bigg|_{\gamma_{10}(\tau)},$$

where τ is determined by (4).

LEMMA 2. *There exist constants c_3, c_4 and c_5 such that the probability $P(\sqrt{n}\, g(\bar{x}, \overline{K}) \le w | \overline{K} = r)$ has the expansion (5) with z given by*

$$(7) \quad z = u - \frac{\omega}{2\sqrt{n}} \frac{g_2(r)}{g_1(r)} u^2 + \frac{1}{n}\left[\frac{1}{2}\left(\omega \frac{g_2(r)}{g_1(r)}\right)^2 - \frac{\omega^2}{6} \frac{g_3(r)}{g_1(r)}\right] u^3 + \theta c_3 \frac{|u|^4}{\sqrt{n}^3}.$$

Here $|\theta| \le 1$,

$$(8) \quad u = \frac{w - \sqrt{n}\, g(\gamma_{10}(\tau), r)}{\omega(\tau) g_1(r)}$$

and the expansion holds for $|u/\sqrt{n}| \le c_4$ and $|r - \mu| \le c_5$.

PROOF. We want the probability of the set

$$g\left(\gamma_{10} + \frac{1}{\sqrt{n}} y, r\right) - g(\gamma_{10}, r) \le \frac{1}{\sqrt{n}}\left(w - \sqrt{n} g(\gamma_{10}, r)\right), \tag{9}$$

where $y = \Sigma x_i / \sqrt{n} - \sqrt{n}\gamma_{10}$. From the conditions A(iv) and (v) we find that the left-hand side of (9) is strictly increasing in $y/\sqrt{n}$ for $|y/\sqrt{n}| < c_6$ and $|r - \mu| < c_7$, say. Taylor expanding we obtain for $|u/\sqrt{n}| < c_8$, with u given in (8) and c_8 suitably small, that the set (9) includes

$$c_9\sqrt{n} \le \frac{1}{\omega} y \le z, \tag{10}$$

with z given in (7). Furthermore, the remaining part of the set (9) is contained in the set $|y/\omega| > c_9\sqrt{n}$ which, according to (5), has probability of order $O(n^{-3/2})$. The result of the lemma now follows from (10). □

Before integrating the conditional expansion given in Lemma 2 with respect to the distribution of $\overline{K}$ we put $v = \sqrt{n}(r - \mu)$ and expand (7), (8) and the coefficients in (5) in powers of $n^{-1/2}$. Let in the following $g_{km} = \partial^{k+m} g(x, r)/(\partial x^k \partial r^m)|_{(0,\mu)}$. A careful analysis of the error term gives the following expansion,

$$\begin{aligned} &P\left(\sqrt{n} g(\bar{x}, \overline{K}) \le w | \overline{K} = r\right) \\ &\quad = \Phi(\xi) + \frac{1}{\sqrt{n}} Q_1(w, v)\varphi(\xi) + \frac{1}{n} Q_2(w, v)\varphi(\xi) \\ &\qquad + \frac{\theta}{\sqrt{n}^3} c_6 \sum_{k+l \le 6} |w|^k |v|^l, \end{aligned} \tag{11}$$

for $|v/\sqrt{n}| \le c_7$ with c_6 and c_7 suitable constants. Here

$$\xi = \left[w - (g_{10}\rho/\sigma + g_{01})v\right] / \left(g_{10}\sqrt{1 - \rho^2}\right)$$

and Q_1, Q_2 are polynomials of degree 3 and 5, respectively. From Lemma 2 the validity of (11) is established for $|\zeta/\sqrt{n}| \le c_8$, say. From the magnitude of the terms in (11) this, however, implies that (11) is valid for all ξ.

We may now show our final theorem.

THEOREM 1. *The formal Edgeworth expansion of the distribution of* $\sqrt{n} g(\bar{x}, \overline{K})$, *as defined in Bhattacharya and Ghosh* (1978), *with* $O(n^{-1})$ *terms included is uniformly valid to order* $o(n^{-1})$.

PROOF. The expansion (11) is only valid for $|v/\sqrt{n}| \le c_7$. However, when integrating with respect to the distribution of v we use (11) for all v. This is because the probability $P(|v| > c_7\sqrt{n}) = P(|\overline{K} - \mu| > c_7)$ is exponentially small in n, see Petrov (1975), and so by Hölder's inequality $E|v|^k 1\{|v| > c_7\sqrt{n}\}$ is also exponentially small.

Since the right-hand side of (11) is a smooth function in v with bounded derivatives we obtain from the results of Götze and Hipp (1978) that we may replace the distribution of v by its formal Edgeworth expansion when integrating (11). Since w^k appears in the error term of (11) we may use the resulting expansion for $|w| < d \log n$ thus giving an error of order $O(n^{-3/2}(\log n)^6)$. Furthermore, the same expansion shows that $P(|w| > d \log n) = O(n^{-3/2}(\log n)^6)$, for d sufficiently large, so that the expansion may be used for all w.

It is a trivial matter to check that the expansion obtained here is identical to the formal Edgeworth expansion. However, it also follows directly from the way the formal Edgeworth expansion is obtained, see Bhattacharya and Ghosh (1978), and the fact that we could replace the distribution of v by its formal expansion above. □

Let us now return to expansions on the form (3) for the signed version R of the log-likelihood ratio statistic. As explained in the introduction we only have to establish the validity of the formal Edgeworth expansion for R in order to obtain (3). This we may do from Theorem 1. As an example we consider observations $x_1, \dots, x_n$ from an exponential distribution possibly censored at the fixed time T. If the mean of the exponential distribution is θ^{-1} the log-likelihood function $l(\theta)$ is

$$l(\theta) = (n - K.)\ln\theta - \theta\sum_1^n x_i,$$

where $K. = \Sigma K_i$ is the number of censored individuals. The signed version R of the log-likelihood ratio statistic becomes $R = \sqrt{n}\, g(\bar{x} - \psi/\theta, \bar{K})$ where

$$g(y, r) = \sqrt{2}\left\{\theta y + \psi - (1 - r)\left[1 + \ln\left(\frac{\theta y + \psi}{1 - r}\right)\right]\right\}^{1/2} \operatorname{sign}(\theta y + (r - e^{-\theta T})),$$

with $\psi = 1 - e^{-\theta T}$. It is easy to check the conditions in Assumption A, in particular we have that $\partial g(y, r)/\partial y|_{(0, e^{-\theta T})} = \theta/\sqrt{\psi} \neq 0$. The coefficients β and B in the expansion (3) are given in Section 3 below.

Let us now briefly discuss the case of exponential life times with varying censoring time T. We then have to extend the proof of Lemma 1 and the result of Götze and Hipp (1978) to the case of independent but not identically distributed observations. For the case in hand this may be done if we assume that there exist $T_1 < T_2$ and $\alpha > 0$ such that the fraction of T_i values between T_1 and T_2 is greater than α for $n \geq n_0$, say.

Let us conclude this section by discussing general testing problems with a partly discrete and a partly continuous minimal sufficient statistic. As mentioned in the introduction the reason that an expansion of the form (3) exists in the continuous case is that there are certain relations between the coefficients of the Edgeworth expansion for the distribution of the minimal sufficient statistic and the coefficients of a Taylor expansion of the likelihood ratio statistic. These relations of course hold also in the partly discrete case using the formal

Edgeworth expansion. What we have seen in the proof above is that by going via the conditional argument the results obtained from the formal expansion actually hold. This is based on the result given for example in Götze and Hipp (1978) that the mean of a sufficiently smooth function $f(x)$, where x is a standardized sum, may be approximated to the required order by replacing the distribution of x by its formal Edgeworth expansion. If therefore it is possible to obtain a continuous expansion of the distribution of the likelihood ratio statistic in the conditional distribution given the discrete part, as in (11), the final expansion (3) should be generally true. Furthermore, in order that the conditional expansion exists in exponential models, it seems that the only requirement is that the problem is not degenerate. By this we mean that the main term of the Taylor expansion of the likelihood ratio statistic should not be entirely in terms of the discrete part of the minimal sufficient statistic but should also include the continuous part, see assumption A(v) above.

The above discussion covers both multi-dimensional parameters and problems with nuisance parameters. An example with a nuisance parameter is given in the next section where we consider a two-sample problem. Another quite special case is given in Jensen (1984) where the pure birth process is considered. Here the discrete part, i.e., the number of births, acts as an approximate ancillary statistic, and an expansion similar to (3) exists both in the conditional distribution given the number of births and in the marginal distribution.

3. Approximations for the one-sample and two-sample cases. In Section 2 we showed that an expansion of the form (3) holds for censored exponential life times with identical censoring times and we also discussed more general situations. However, the detailed discussion of Section 2 is not needed in order to derive the bias correction β and the Bartlett adjustment B. We now derive these coefficients for the one-sample and two-sample situation, respectively, and with varying censoring times T_i.

The Bartlett adjustment may be obtained by using the method given in Lawley (1956). For this we need no more than the means of the first four derivatives for the log-likelihood function. For the one-sample case we find that

$$B = \lambda_2^{-2}\left\{\frac{1}{4}\lambda_4 - (\lambda_3)_1 + (\lambda_2)_2\right\} - \lambda_2^{-3}\left\{\frac{5}{12}\lambda_3^2 - 2\lambda_3(\lambda_2)_1 + 2(\lambda_2)_1^2\right\}$$
$$= \frac{1}{\psi_0}\left\{\frac{1}{6} + \frac{\psi_2}{\psi_0} - 2\frac{\psi_1}{\psi_0} + 2\frac{\psi_1^2}{\psi_0^2}\right\}, \tag{12}$$

where

$$\lambda_i = E\left(\frac{1}{n}\frac{d^i l(\theta)}{d\theta^i}\right) = (-1)^{i-1}(i-1)!\frac{1}{\theta^i}E\left(\frac{n-K}{n}\right)$$
$$= (-1)^{i-1}(i-1)!\frac{\psi_0}{\theta^i},$$
$$(\lambda_i)_j = \frac{d^j\lambda_i(\theta)}{d\theta^j}$$

and

$$(13)\quad \psi_0 = 1 - \frac{1}{n}\sum_1^n e^{-\theta T_i}, \qquad \psi_1 = \frac{1}{n}\sum_1^n \theta T_i e^{-\theta T_i}, \qquad \psi_2 = \frac{1}{n}\sum_1^n (\theta T_i)^2 e^{-\theta T_i}.$$

In the two-sample case we consider a sample of size n_1 from a censored exponential distribution with mean θ^{-1} and a sample of size n_2 with mean $(\chi\theta)^{-1}$. We then want to test the hypothesis $\chi = \chi_0$ with θ a nuisance parameter. The new Bartlett adjustment takes the form

$$(14)\qquad B = \frac{n}{n_1}B_1 + \frac{n}{n_2}B_2 - B_0,$$

where $n = n_1 + n_2$ and B_1 and B_2 are the Bartlett adjustments from the two samples, respectively, calculated from (12) [with $\theta = \chi_0\theta$ in (13) for the second sample]. The term B_0 is also on the form (12) but with the ψ_i coefficient given as a weighted mean of the coefficients from the two samples, i.e.,

$$\psi_i = \frac{n_1}{n}\psi_{1i} + \frac{n_2}{n}\psi_{2i},$$

with obvious notation. When using the Bartlett adjustment (14) we replace the nuisance parameter θ with its estimate under the null hypothesis.

The bias term may also be calculated from the formulas given in Lawley (1956). For the one-sample case we find

$$(15)\qquad \beta = -\frac{1}{6\psi_0^{3/2}}(2\psi_0 - 3\psi_1),$$

with the sign of the likelihood ratio statistic determined by $\theta - \hat{\theta}$. And for the two-sample case we find after some calculation and with $\alpha_i = n_i/n$,

$$\beta = \gamma_1\beta_1 - \gamma_2\beta_2$$

and β_i is the coefficient (15) for sample i and

$$\gamma_i = \left(\frac{\alpha_1\alpha_2\psi_{10}\psi_{20}}{\psi_0\psi_{i0}}\right)^{1/2}\left(1 + \frac{\alpha_i\psi_{i0}}{\psi}\right)\frac{1}{\alpha_i}.$$

Here the sign of the likelihood ratio statistic is given by the sign of $\hat{\chi} - \chi_0$ with $\hat{\chi} = \hat{\theta}_2/\hat{\theta}_1$.

4. Numerical examples. In this section we discuss two numerical examples for the one-sample case, one to show that the approximations work satisfactorily for moderate sample sizes, when the probability of censoring is small, and one to give a warning when the probability of censoring is high.

In Bartholomew (1963) an example is considered with 20 survival times subject to censoring at $T = 150$ hours. During that period 15 items fail with the following life times, measured in hours: 3, 19, 23, 26, 27, 37, 38, 41, 45, 58, 84, 90, 99, 109, 138. A test for the hypothesis $\theta = 1/65$ is wanted and under this hypothesis the bias term β from (15) and the Bartlett adjustment B from (12) become

$$\beta = -0.217, \qquad B = 0.417.$$

To check the approximations to the mean of the likelihood ratio statistic w and the signed version r and to check the adequacy of the distributional approximations 100,000 samples of the size 20 were generated. For the mean of R we find

$$ER \approx \frac{\beta}{\sqrt{n}} = -0.049, \qquad \widehat{ER} = -0.051 \pm 0.003,$$

where here and in the following $\pm\sigma$ gives the standard deviation of the estimate $\widehat{ER}$ based on 100,000 samples. For the mean of w we get

$$Ew \approx 1 + \frac{B}{n} = 1.021, \qquad \widehat{Ew} = 1.024 \pm 0.005.$$

In both cases we see that the theoretical approximations to order $1/n$ are good. Next, we compare the approximation to the distributions of r, r', w and w' with the estimates of the exact distributions as obtained from the simulation. For a few selected values of u_α, where $\Phi(u_\alpha) = \alpha$ with Φ the normal distribution function, we compare α, $P(R \le u_\alpha)$ and $P(R' \le u_\alpha)$ in the first table below. In the second table we compare 2α, $P(w \le u_\alpha^2)$ and $P(w' \le u_\alpha^2)$. All the values in the tables are given in percentages.

Approximation		1.0	2.5	97.5	99.0
Exact	r	1.17 ± 0.03	2.96 ± 0.05	97.72 ± 0.05	99.10 ± 0.03
	r'	0.97 ± 0.03	2.54 ± 0.05	97.55 ± 0.05	99.02 ± 0.03

Approximation		95.0	97.5	99.0
Exact	w	94.76 ± 0.07	97.36 ± 0.05	98.92 ± 0.03
	w'	94.99 ± 0.07	97.52 ± 0.05	99.00 ± 0.03

We see that an improvement is obtained by using the standardized statistics and that the approximations are satisfactory.

For the second example we let again the sample size be $n = 20$ and choose the censoring time T such that under the null hypothesis the probability of censoring is $p = 0.9$. For the mean of R and w we find

$$ER \approx \frac{\beta}{\sqrt{n}} = 0.100, \qquad \widehat{ER} = 0.125 \pm 0.003,$$

$$Ew \approx 1 + \frac{B}{n} = 1.084, \qquad \widehat{Ew} = 1.136 \pm 0.005,$$

which show some discordance. The table of the distribution functions becomes

Approximation		1.0	2.5	97.5	99.0
Exact	R	1.11 ± 0.03	1.90 ± 0.04	87.95 ± 0.10	100.00 ± 0.00
	R'	1.12 ± 0.03	2.40 ± 0.05	100.00 ± 0.00	100.00 ± 0.00

Approximation		95.0	97.5	99.0
Exact	w	86.05 ± 0.11	98.88 ± 0.03	99.70 ± 0.02
	w'	86.75 ± 0.11	98.89 ± 0.03	99.76 ± 0.02

We see now that serious problems occur in the upper tail of the distribution of the signed likelihood ratio statistic R. The reason is, however, quite simple. The exact distribution of R is bounded to the right and has a point probability—in the simulations at $\sqrt{2nT}$—corresponding to all the observations being censored. This point probability is in the above example 12.2 percent and so the upper 10–20 percent of the distribution cannot be well approximated by a continuous distribution.

For a more thorough simulation study with small samples from an exponential distribution with censoring see Schou and Væth (1980).

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. and COX, D. R. (1979). Edgeworth and saddlepoint approximations with statistical applications (with discussion). *J. Roy. Statist. Soc. Ser. B* **41** 279–312.

BARNDORFF-NIELSEN, O. E. and COX, D. R. (1984). Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator. *J. Roy. Statist. Soc. Ser. B* **46** 483–495.

BARTHOLOMEW, D. J. (1963). The sampling distribution of an estimate arising in life testing. *Technometrics* **5** 361–374.

BARTLETT, M. S. (1938). The characteristic function of a conditional statistic. *J. London Math. Soc.* **13** 62–67.

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

BICKEL, P. J. and VAN ZWET, W. R. (1978). Asymptotic expansions for the power of distribution free tests in the two-sample problem. *Ann. Statist.* **6** 937–1004.

CHANDRA, T. K. and GHOSH, J. K. (1979). Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables. *Sankhyā Ser. A* **41** 22–47.

DICICCIO, T. J. (1983). Ph.D. thesis, Univ. Waterloo.

DICICCIO, T. J. (1984). On parameter transformations and interval estimation. *Biometrika* **71** 477–485.

DOES, R. J. M. M. (1983). An Edgeworth expansion for simple linear rank statistics under the null-hypothesis. *Ann. Statist.* **11** 607–624.

DOES, R. J. M. M., HELMERS, R. and KLAASSEN, C. A. J. (1984). On the Edgeworth expansion for the sum of a function of uniform spacings. To appear in *J. Statist. Plann. Inference.*

ESSEEN, C. G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta. Math.* **77** 1–125.

GÖTZE, F. and HIPP, C. (1978). Asymptotic expansions in the central limit theorem under moment conditions. *Z. Wahrsch. verw. Gebiete* **42** 67–87.

HIPP, C. (1984). Asymptotic expansions for conditional distributions: the lattice case. *Probab. Math. Statist.* **4** 207–219.

JENSEN, J. L. (1984). Asymptotic expansions in the pure birth process. Research Report 123, Dept. Theoretical Statistics, Aarhus Univ.

JENSEN, J. L. (1986). Similar tests and the standardized log likelihood ratio statistic. *Biometrika* **73** 567–572.

LAWLEY, D. N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika* **43** 295–303.

MCCULLAGH, P. (1984). Local sufficiency. *Biometrika* **71** 233–244.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York.

SCHOU, G. and VÆTH, M. (1980). A small sample study of occurrence/exposure rates for rare events. *Scand. Actuar. J.* **63** 209–225.

DEPARTMENT OF THEORETICAL STATISTICS
INSTITUTE OF MATHEMATICS
UNIVERSITY OF AARHUS
NY MUNKEGADE
DK-8000 AARHUS C
DENMARK