Biometrika (1991), **78**, 2, *pp.* 365–72
Printed in Great Britain

Bartlett-type adjustment for the conditional likelihood ratio statistic of Cox and Reid

BY RAHUL MUKERJEE
Indian Institute of Management, Post Box No. 16757, Calcutta 700 027, India
AND TAPAS K. CHANDRA
Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, Calcutta 700 035, India

SUMMARY

This paper explicitly derives a Bartlett-type adjustment for the conditional likelihood ratio statistic of Cox & Reid via that for the usual likelihood ratio statistic.

Some key words: Conditional likelihood ratio statistic; Likelihood ratio statistic; Parametric orthogonality.

1. INTRODUCTION

In recent pioneering work Cox & Reid (1987) introduced the notion of the conditional likelihood ratio statistic, derived results on it and posed several open problems; for a very illuminating further discussion, see Cox (1988, §§ 5.2, 5.4). One of these problems relates to the derivation of Bartlett-type adjustments (Bartlett, 1937; Barndorff-Nielsen & Cox, 1984) for the conditional likelihood ratio statistic via that for the usual likelihood ratio statistic. The present paper attempts to settle this problem to some extent.

We recall some definitions from Cox & Reid (1987). Let $\{X_i\}$ $(i \geq 1)$ be a sequence of independent and identically distributed random variables with common density $f(x;\, \theta, m)$ where θ is the one-dimensional parameter of interest and m is the nuisance parameter. Consider the null hypothesis H_0: $\theta = \theta_0$. For scalar m, the conditional likelihood ratio statistic is defined as

$$\lambda_n = 2\{h(\tilde{\theta}) - h(\theta_0)\},$$

where

$$h(\theta) = l_X(\theta, \hat{m}_\theta) - \tfrac{1}{2}\log\{nJ_{mm}(\theta, \hat{m}_\theta)\},$$

n is the sample size, $\hat{m}_\theta$ is the maximum likelihood estimator of m given θ,

$$J_{mm}(\theta, m) = -n^{-1}\sum \partial^2 \log f(X_i;\, \theta, m)/\partial m^2, \quad l_X(\theta, m) = \sum \log f(X_i;\, \theta, m)$$

$\sum$ denotes summation over i $(1 \leq i \leq n)$ and $h(\tilde{\theta}) = \sup_\theta h(\theta)$. Also, the usual likelihood ratio statistic is given by $\bar{\lambda}_n = 2\{l_X(\hat{\theta}, \hat{m}) - l_X(\theta_0, m_0^*)\}$, where $(\hat{\theta}, \hat{m})$ is the unrestricted maximum likelihood estimator of (θ, m) and $m_0^* = \hat{m}_{\theta_0}$.

In this paper, primarily for notational simplicity, we consider the situation where both θ and m are one-dimensional. Of course, the present discussion can be extended to the case of multi-dimensional m with additional algebra. However, the assumption that θ is one-dimensional is nontrivial. In particular, if θ and m are both multi-dimensional, then in general one cannot employ global parametric orthogonality as noted by Cox & Reid (1987).

2. Some preliminary results

The following notation will be helpful. For $i, j, i', j' = 0, 1, \ldots$, let

$$K_{ij}(\theta, m) = E_{\theta,m}\left\{\frac{\partial^{i+j}\log f(X;\theta,m)}{\partial\theta^i\partial m^j}\right\},$$

$$K_{ij.i'j'}(\theta, m) = E_{\theta,m}\left\{\frac{\partial^{i+j}\log f(X;\theta,m)}{\partial\theta^i\partial m^j}\frac{\partial^{i'+j'}\log f(X;\theta,m)}{\partial\theta^{i'}\partial m^{j'}}\right\},$$

$$H_{ij}(\theta, m) = n^{-1/2}\sum_{s=1}^{n}\left\{\frac{\partial^{i+j}\log f(X_s;\theta,m)}{\partial\theta^i\partial m^j} - K_{ij}(\theta, m)\right\},$$

$$K^*_{ij} = K_{ij}(\theta_0, m^*_0), \quad K^*_{ij.i'j'} = K_{ij.i'j'}(\theta_0, m^*_0), \quad H^*_{ij} = H_{ij}(\theta_0, m^*_0),$$

$$L_{ij} = K_{ij}(\theta_0, m), \quad L_{ij.i'j'} = K_{ij.i'j'}(\theta_0, m), \quad S_{ij} = H_{ij}(\theta_0, m),$$

$$a^*_{20} = -K^*_{20}, \quad a^*_{02} = -K^*_{02}, \quad a_{20} = -L_{20}, \quad a_{02} = -L_{02}.$$

Similarly we define $K_{ij.i'j'.i''j''}(\theta, m)$, $L_{ij.i'j'.i''j''}$, etc. Note that L_{ij}, $L_{ij.i'j'}$, etc. are functions of m. Since θ, m are both one-dimensional, we assume global parametric orthogonality (Cox & Reid, 1987), that is $K_{11}(\theta, m) \equiv 0$ for every θ, m. Then under standard regularity conditions, the per observation information matrix at θ_0 is given by $\mathrm{diag}\,(a_{20}, a_{02})$ and this is assumed to be positive definite for every m.

All formal expansions used in this paper are over a set A_n with $P_{\theta_0}(A_n) = 1 + o(n^{-1})$ (Chandra & Ghosh, 1979, p. 40). Most of the computational details have been omitted here to save space but may be obtained from the authors.

Considering first the likelihood ratio statistic, computations similar to those of Cox & Reid (1987, p. 13), see also Chandra & Ghosh (1979), with the use of parametric orthogonality yield

$$\hat\theta = \theta_0 + n^{-\frac{1}{2}}\left(\frac{H^*_{10}}{a^*_{20}}\right) + n^{-1}\{(a^*_{20})^{-2}H^*_{10}H^*_{20} + \tfrac{1}{2}(a^*_{20})^{-3}K^*_{30}(H^*_{10})^2\} + o(n^{-1}), \quad (2{\cdot}1\mathrm{a})$$

$$\hat m = m^*_0 + n^{-1}\{(a^*_{20}a^*_{02})^{-1}H^*_{10}H^*_{11} + \tfrac{1}{2}(a^*_{20})^{-2}(a^*_{02})^{-1}K^*_{21}(H^*_{10})^2\} + o(n^{-1}), \quad (2{\cdot}1\mathrm{b})$$

whence, using a Taylor's expansion about (θ_0, m^*_0),

$$\bar\lambda_n = \bar W_n^2 + o(n^{-1}), \quad (2{\cdot}2\mathrm{a})$$

where

$$\bar W_n = (a^*_{20})^{-\frac{1}{2}}H^*_{10} + n^{-\frac{1}{2}}\bar Q_1 + n^{-1}\bar Q_2, \quad (2{\cdot}2\mathrm{b})$$

$$\bar Q_1 = \tfrac{1}{2}(a^*_{20})^{-3/2}H^*_{10}H^*_{20} + \tfrac{1}{6}(a^*_{20})^{-5/2}K^*_{30}(H^*_{10})^2, \quad (2{\cdot}2\mathrm{c})$$

$$\begin{aligned}\bar Q_2 = {} & \tfrac{3}{8}(a^*_{20})^{-5/2}H^*_{10}(H^*_{20})^2 + \tfrac{5}{12}(a^*_{20})^{-7/2}K^*_{30}(H^*_{10})^2H^*_{20} \\ & + [\tfrac{1}{24}(a^*_{20})^{-7/2}\{K^*_{40} + 3(K^*_{21})^2/a^*_{02}\} + \tfrac{1}{9}(a^*_{20})^{-9/2}(K^*_{30})^2](H^*_{10})^3 \\ & + \tfrac{1}{6}(a^*_{20})^{-5/2}(H^*_{10})^2H^*_{30} + \tfrac{1}{2}(a^*_{20})^{-3/2}(a^*_{02})^{-1}H^*_{10}(H^*_{11})^2 \\ & + \tfrac{1}{2}(a^*_{20})^{-5/2}(a^*_{02})^{-1}K^*_{21}(H^*_{10})^2H^*_{11}. \end{aligned} \quad (2{\cdot}2\mathrm{d})$$

Next considering the conditional likelihood ratio statistic, similar computations show that

$$\tilde\theta = \hat\theta + n^{-1}K^*_{12}/(2a^*_{20}a^*_{02}) + o(n^{-1}), \quad \hat m_{\tilde\theta} = \hat m + o(n^{-1}), \quad (2{\cdot}3)$$

where $\hat\theta$, $\hat m$ are as in (2·1), and as before

$$\lambda_n = (W_n)^2 + o(n^{-1}), \quad (2{\cdot}4\mathrm{a})$$

where

$$W_n = (a^*_{20})^{-\frac{1}{2}} H^*_{10} + n^{-\frac{1}{2}} Q_1 + n^{-1} Q_2, \quad (2\cdot 4b)$$

$$Q_1 = \bar{Q}_1 + \tfrac{1}{2}(a^*_{20})^{-\frac{1}{2}}(a^*_{02})^{-1} K^*_{12}, \quad (2\cdot 4c)$$

$$\begin{aligned} Q_2 = \bar{Q}_2 &+ \tfrac{1}{2}(a^*_{20})^{-\frac{1}{2}}(a^*_{02})^{-1} H^*_{12} + \tfrac{1}{2}(a^*_{20})^{-\frac{1}{2}}(a^*_{02})^{-2}\{K^*_{12}H^*_{02} + K^*_{03}H^*_{11}\} \\ &+ [\tfrac{1}{4}(a^*_{02})^{-1}K^*_{22} + \tfrac{1}{4}(a^*_{02})^{-2}\{K^*_{21}K^*_{03} + (K^*_{12})^2\} + \tfrac{1}{6}(a^*_{20}a^*_{02})^{-1}K^*_{12}K^*_{30}](a^*_{20})^{-3/2}H^*_{10} \\ &+ \tfrac{1}{4}(a^*_{20})^{-3/2}(a^*_{02})^{-1}K^*_{12}H^*_{20}. \end{aligned} \quad (2\cdot 4d)$$

It should be noted that the derivation of (2·2a), (2·4a) also requires terms of order $O(n^{-3/2})$ in (2·1), (2·3). However, simple calculations show that the contribution of such terms eventually cancel out and hence these terms are not shown here. The expressions for $\bar{Q}_1$, Q_1, as in (2·2c), (2·4c), were also noted by Mukerjee (1989a). Observe that, as noted by Cox & Reid (1987), $\hat{m} - m^*_0 = O(n^{-1})$, $\tilde{\hat{\theta}} - \hat{\theta} = O(n^{-1})$; the latter observation, together with the fact, see (2·1a, b), that

$$\hat{m} = \hat{m}_{\hat{\theta}} = m^*_0 + n^{-1}(H^*_{11}d + \tfrac{1}{2}K^*_{21}d^2)/a^*_{02} + o(n^{-1}),$$

where $d = n^{\frac{1}{2}}(\hat{\theta} - \theta_0)$, provides a kind of intuitive justification for the second relation in (2·3).

In order to derive the Bartlett-type adjustment for λ_n through that for $\bar{\lambda}_n$, it will be convenient to work with their 'square root' versions W_n and $\bar{W}_n$. In fact, as recent studies (Chandra & Ghosh, 1979; Bickel & Ghosh, 1990; DiCiccio, Field & Fraser, 1990; Levin & Kong, 1990; Mukerjee, 1990) show, such an approach is useful in other contexts as well. In order to obtain the approximate cumulants of W_n, $\bar{W}_n$, we note that

$$m^*_0 = m + n^{-\frac{1}{2}}\left(\frac{S_{01}}{a_{02}}\right) + n^{-1}(a_{02}^{-2}S_{01}S_{02} + \tfrac{1}{2}a_{02}^{-3}L_{03}S_{01}^2) + o(n^{-1}), \quad (2\cdot 5)$$

and make a further Taylor's expansion about (θ_0, m) to get

$$\bar{W}_n = a_{20}^{-\frac{1}{2}}S_{10} + n^{-\frac{1}{2}}\bar{Q}'_1 + n^{-1}\bar{Q}'_2 + o(n^{-1}), \quad (2\cdot 6a)$$

where

$$\bar{Q}'_1 = v_1 S_{10}S_{20} + v_2 S_{10}^2 + v_3 S_{10}S_{01} + v_4 S_{01}^2 + v_5 S_{01}S_{11}, \quad (2\cdot 6b)$$

$$\begin{aligned} \bar{Q}'_2 = y_1 S_{10}S_{20}^2 &+ y_2 S_{10}^2 S_{20} + y_3 S_{10}^3 + y_4 S_{10}^2 S_{30} + y_5 S_{10}S_{11}^2 + y_6 S_{10}S_{01}S_{11} + y_7 S_{10}^2 S_{11} \\ &+ y_8 S_{10}^2 S_{01} + y_9 S_{10}S_{01}S_{21} + y_{10}S_{20}S_{01}^2 + y_{11}S_{20}S_{01}S_{11} + y_{12}S_{20}S_{10}S_{01} + y_{13}S_{01}S_{10}S_{02} \\ &+ y_{14}S_{10}S_{01}^2 + y_{15}S_{01}^2 S_{11} + y_{16}S_{01}^3 + y_{17}S_{01}S_{02}S_{11} + y_{18}S_{01}^2 S_{02} + y_{19}S_{01}^2 S_{12}, \end{aligned} \quad (2\cdot 6c)$$

$$v_1 = \tfrac{1}{2}a_{20}^{-3/2}, \quad v_2 = \tfrac{1}{6}L_{30}a_{20}^{-5/2}, \quad v_3 = \tfrac{1}{2}(a_{20}^{3/2}a_{02})^{-1}L_{21}, \quad v_4 = \tfrac{1}{2}(a_{20}^{\frac{1}{2}}a_{02}^2)^{-1}L_{12},$$

$$v_5 = (a_{20}^{\frac{1}{2}}a_{02})^{-1}, \quad y_1 = \tfrac{3}{8}a_{20}^{-5/2}, \quad y_2 = \tfrac{5}{12}a_{20}^{-7/2}L_{30}, \quad y_3 = \tfrac{1}{24}a_{20}^{-7/2}(L_{40} + 3L_{21}^2/a_{02}) + \tfrac{1}{9}a_{20}^{-9/2}L_{30}^2,$$

$$y_4 = \tfrac{1}{6}a_{20}^{-5/2}, \quad y_5 = \tfrac{1}{2}(a_{20}^{3/2}a_{02})^{-1}, \quad y_6 = (a_{20}^{3/2}a_{02}^2)^{-1}L_{12} + \tfrac{1}{3}(a_{20}^{5/2}a_{02})^{-1}L_{30}, \quad (2\cdot 7a)$$

$$y_7 = \tfrac{1}{2}(a_{20}^{5/2}a_{02})^{-1}L_{21}, \quad y_8 = \tfrac{1}{2}(a_{20}^{5/2}a_{02}^2)^{-1}L_{21}L_{12} + \tfrac{1}{6}(a_{20}^{5/2}a_{02})^{-1}L_{31} + \tfrac{5}{12}(a_{20}^{7/2}a_{02})^{-1}L_{30}L_{21},$$

$$y_9 = \tfrac{1}{2}(a_{20}^{3/2}a_{02})^{-1}, \quad y_{10} = \tfrac{1}{4}(a_{20}^{3/2}a_{02}^2)^{-1}L_{12}, \quad y_{11} = \tfrac{1}{2}(a_{20}^{3/2}a_{02})^{-1},$$

$$y_{12} = \tfrac{3}{4}(a_{20}^{5/2}a_{02})^{-1}L_{21}, \quad y_{13} = \tfrac{1}{2}(a_{20}^{3/2}a_{02}^2)^{-1}L_{21},$$

$$y_{14} = (a_{20}^{3/2}a_{02}^2)^{-1}\left\{\frac{1}{4}\left(\frac{L_{21}L_{03}}{a_{02}}\right) + \frac{3}{8}\left(\frac{L_{21}^2}{a_{20}}\right) + \frac{1}{4}L_{22} + \frac{1}{2}\left(\frac{L_{12}^2}{a_{02}}\right) + \frac{1}{6}\left(\frac{L_{30}L_{12}}{a_{20}}\right)\right\}, \quad (2\cdot 7b)$$

$$y_{15}=\tfrac{1}{2}(a_{20}^{\frac{1}{2}}a_{02}^{2})^{-1}\left(\frac{L_{21}}{a_{20}}+\frac{L_{03}}{a_{02}}\right),\quad y_{16}=(a_{20}^{\frac{1}{2}}a_{02}^{3})^{-1}\left\{\frac{1}{6}L_{13}+\frac{1}{2}\left(\frac{L_{12}L_{03}}{a_{02}}\right)+\frac{1}{4}\left(\frac{L_{12}L_{21}}{a_{20}}\right)\right\},$$
$$y_{17}=(a_{20}^{\frac{1}{2}}a_{02}^{2})^{-1},\quad y_{18}=(a_{20}^{\frac{1}{2}}a_{02}^{3})^{-1}L_{12},\quad y_{19}=\tfrac{1}{2}(a_{20}^{\frac{1}{2}}a_{02}^{2})^{-1}.$$

Similarly,

$$W_n=a_{20}^{-\frac{1}{2}}S_{10}+n^{-\frac{1}{2}}Q_1'+n^{-1}Q_2'+o(n^{-1}), \tag{2·8a}$$

where

$$Q_1'=\bar{Q}_1'+v,\quad Q_2'=\bar{Q}_2'+Y, \tag{2·8b}$$
$$v=\tfrac{1}{2}(a_{20}^{\frac{1}{2}}a_{02})^{-1}L_{12}, \tag{2·9a}$$
$$Y=y_1'S_{12}+y_2'S_{02}+y_3'S_{11}+y_4'S_{20}+y_5'S_{10}+y_6'S_{01}, \tag{2·9b}$$

$$\begin{gathered}
y_1'=\tfrac{1}{2}(a_{20}^{\frac{1}{2}}a_{02})^{-1},\quad y_2'=\tfrac{1}{2}(a_{20}^{\frac{1}{2}}a_{02}^{2})^{-1}L_{12},\\
y_3'=\tfrac{1}{2}(a_{20}^{1/2}a_{02}^{2})^{-1}L_{03},\quad y_4'=\tfrac{1}{4}(a_{20}^{3/2}a_{02})^{-1}L_{12},\\
y_5'=(a_{20}^{3/2}a_{02})^{-1}\left\{\frac{1}{4}L_{22}+\frac{1}{4}a_{02}^{-1}(L_{21}L_{03}+L_{12}^{2})+\frac{1}{6}\left(\frac{L_{12}L_{30}}{a_{20}}\right)\right\},\\
y_6'=(a_{20}^{\frac{1}{2}}a_{02}^{2})^{-1}\left\{\frac{1}{2}L_{13}+\frac{1}{4}\left(\frac{L_{12}L_{21}}{a_{20}}\right)+\left(\frac{L_{12}L_{03}}{a_{02}}\right)\right\}.
\end{gathered} \tag{2·10}$$

From (2·8b), (2·9), observe that Q_1' differs from $\bar{Q}_1'$ only by a constant while Q_2' differs from $\bar{Q}_2'$ only by a linear term, a fact which plays a crucial role in the subsequent derivations. Since the expressions in (2·6), (2·7), (2·9), (2·10) are, indeed, involved, the computations have been verified by alternative methods, for example, by considering expansions about (θ_0, m) directly instead of working in two stages as shown above. An advantage of the present two-stage derivation is that at the intermediate stage one obtains expansions, up to $o(n^{-1})$ and free from the nuisance parameter, for $\hat{\theta}$, $\hat{m}$, $\tilde{\theta}$, $\hat{m}_{\tilde{\theta}}$, $\bar{\lambda}_n$ and λ_n, see (2·1)–(2·4), and these may be useful in other contexts as well. Moreover, as our computational experience suggests, the algebra in the two-stage derivation appears to be a little simpler than that in a direct one-stage expansion about (θ_0, m).

3. The Bartlett-type adjustment

As indicated in the Appendix, the approximate cumulants of $\bar{W}_n$, under θ_0, are given by, say,

$$\begin{gathered}
k_{1n}(\bar{W}_n)=n^{-\frac{1}{2}}\bar{R}_1+o(n^{-1}),\quad k_{2n}(\bar{W}_n)=1+n^{-1}\bar{R}_2+o(n^{-1}),\\
k_{rn}(\bar{W}_n)=o(n^{-1})\quad (r\geqslant 3),
\end{gathered} \tag{3·1}$$

where

$$\bar{R}_1=-\frac{1}{2}a_{20}^{-\frac{1}{2}}\left\{\left(\frac{L_{12}}{a_{02}}\right)+\frac{1}{3}\left(\frac{L_{10.10.10}}{a_{20}}\right)\right\}, \tag{3·2a}$$

$$\begin{aligned}
\bar{R}_2={}&a_{20}^{-2}(\tfrac{1}{4}L_{40}+L_{20.20}+L_{10.30}+L_{10.10.20})+a_{20}^{-3}(\tfrac{7}{4}L_{10.20}^{2}+\tfrac{11}{6}L_{30}L_{10.20}+\tfrac{7}{18}L_{30}^{2})\\
&+(a_{20}a_{02})^{-1}(2L_{11.11}+L_{01.21}+2L_{10.01.11}+L_{10.12}+\tfrac{1}{2}L_{22})\\
&-(a_{20}a_{02}^{2})^{-1}(\tfrac{1}{2}L_{12}^{2}+L_{21}L_{01.02}+\tfrac{1}{2}L_{21}L_{03}+L_{12}L_{10.02})\\
&-(a_{20}^{2}a_{02})^{-1}(\tfrac{1}{4}L_{21}^{2}+\tfrac{1}{3}L_{30}L_{12}+\tfrac{1}{2}L_{12}L_{10.20}+L_{21}L_{20.01}).
\end{aligned} \tag{3·2b}$$

The expressions in (3·1) and (3·2) may be compared with equation (9) of DiCiccio et al..(1990).

Observe that by (2·6), (2·8), $W_n = \bar{W}_n + n^{-\frac{1}{2}}v + n^{-1}Y + o(n^{-1})$. Hence using the facts that (i) v is a constant, (ii) Y consists only of linear terms, (iii) $\bar{W}_n = a_{20}^{-\frac{1}{2}} S_{10} + O(n^{-\frac{1}{2}})$, and (iv) the limiting joint distribution of S_{10} and Y, under θ_0, is bivariate normal, up to the first order of approximation, with

$$E_{\theta_0}(S_{10}) = E_{\theta_0}(Y) = 0, \quad \mathrm{var}_{\theta_0}(S_{10}) = a_{10}, \quad \mathrm{cov}_{\theta_0}(S_{10}, Y) = c,$$

for each m, where, by (2·9b), (2·10),

$$c = (a_{20}^{\frac{1}{2}} a_{02})^{-1}\{\tfrac{1}{2}L_{10.12} + \tfrac{1}{4}L_{22} + a_{20}^{-1}L_{12}(\tfrac{1}{4}L_{10.20} + \tfrac{1}{6}L_{30}) + a_{02}^{-1}(\tfrac{1}{2}L_{12}L_{10.02} - \tfrac{1}{4}L_{03}L_{21} + \tfrac{1}{4}L_{12}^2)\}, \tag{3·3}$$

it follows by a little algebra from (3·1) that the approximate cumulants of W_n, under θ_0, are given by

$$k_{1n}(W_n) = n^{-\frac{1}{2}}R_1 + o(n^{-1}), \quad k_{2n}(W_n) = 1 + n^{-1}R_2 + o(n^{-1}),$$
$$k_{3n}(W_n) = o(n^{-1}), \quad k_{4n}(W_n) = o(n^{-1}), \tag{3·4}$$

where

$$R_1 = \bar{R}_1 + v, \quad R_2 = \bar{R}_2 + 2a_{20}^{-\frac{1}{2}}c. \tag{3·5}$$

Furthermore, calculations similar to those in the Appendix show that

$$k_{rn}(W_n) = o(n^{-1}) \quad (r \geqslant 5). \tag{3·6}$$

By (2·4a), (3·4),

$$E_{\theta_0,m}(\lambda_n) = E_{\theta_0,m}(W_n^2) + o(n^{-1}) = 1 + n^{-1}a(m) + o(n^{-1}), \tag{3·7a}$$

where

$$a(m) = R_2 + R_1^2. \tag{3·7b}$$

Consider now the Bartlett adjusted statistic

$$\lambda_{Bn} = \lambda_n / \{1 + n^{-1}a(m_0^*)\}. \tag{3·8}$$

By (2·4a), (2·5),

$$\lambda_{Bn} = W_{Bn}^2 + o(n^{-1}), \tag{3·9}$$

where

$$W_{Bn} = W_n\{1 - \tfrac{1}{2}n^{-1}a(m)\}.$$

Hence by (3·4), (3·6), (3·7b), the approximate cumulants of W_{Bn}, under θ_0, are given by

$$k_{1n}(W_{Bn}) = n^{-\frac{1}{2}}R_1 + o(n^{-1}), \quad k_{2n}(W_{Bn}) = 1 - n^{-1}R_1^2 + o(n^{-1}),$$
$$k_{rn}(W_{Bn}) = o(n^{-1}) \quad (r \geqslant 3), \tag{3·10}$$

so that using an Edgeworth expansion for W_{Bn} (Bhattacharya & Ghosh, 1978) and recalling the symmetry of the normal distribution, it is clear from (3·9) that

$$\mathrm{pr}_{\theta_0}(\lambda_{Bn} \leqslant x) = \mathrm{pr}_{\theta_0}(|W_{Bn}| \leqslant x^{\frac{1}{2}}) + o(n^{-1}) = \int_0^x g(z)\,dz + o(n^{-1}), \tag{3·11}$$

for all $x \geqslant 0$ and all m, where $g(z)$ is the density of the chi-squared distribution with 1 degree of freedom. Hence by (3·7), (3·8), a Bartlett-type adjustment is available for λ_n.

The above line of argument works primarily because, as a consequence of the rather special nature of the correction term $J_{mm}(\theta, \hat{m}_\theta)$ involved in the definition of λ_n, the third and higher order cumulants of W_n, like those of $\bar{W}_n$, are of order $o(n^{-1})$ under θ_0. It may also be emphasized that, in consideration of (3·7)–(3·10), by correcting the mean of the conditional likelihood ratio statistic λ_n, all cumulants of λ_n, under θ_0, are simultaneously corrected up to $o(n^{-1})$. This may be contrasted with the findings of Harris (1987) who worked directly with the moment generating function of λ_n and perhaps overlooked the simplification that the consideration of the square root version W_n entails. As a final remark, in the spirit of equations (10)–(12) of DiCiccio et al. (1990) and in view of (3·4), one could as well have corrected W_n by $W_n^* = (W_n - n^{-\frac{1}{2}} R_1)/(1 + n^{-1} R_2)^{\frac{1}{2}}$ and then λ_n by $\lambda_n^* = W_n^{*2}$. Clearly, under θ_0 all the cumulants of W_n^*, unlike those of W_{Bn}, agree, up to $o(n^{-1})$, with those of a standard normal variate and hence if interest lies in the one-sided tail probabilities for the square root version of the conditional likelihood ratio statistic then use of W_n^* rather than W_{Bn} gives a closer approximation to normality. However, in the present context interest lies more often in tail probabilities concerning the conditional likelihood ratio statistic itself and it is easily seen that to this effect, use of λ_n^* will give no better approximation to the chi-squared probability integral than the Bartlett-type adjusted statistic λ_{Bn} which satisfies (3·11).

We now indicate how the adjustment factor for λ_n can be calculated from that for the usual likelihood ratio statistic $\bar{\lambda}_n$. By (2·2a), (2·5), (3·1), it can be seen that, analogously to (3·8), the Bartlett-adjusted statistic corresponding to $\bar{\lambda}_n$ is

$$\bar{\lambda}_{Bn} = \bar{\lambda}_n / \{1 + n^{-1} \bar{a}(m_0^*)\},$$

where

$$\bar{a}(m) = \bar{R}_2 + \bar{R}_1^2. \tag{3·12}$$

By (3·5), (3·7b), (3·12),

$$a(m) = \bar{a}(m) + 2a_{20}^{-\frac{1}{2}} c + 2\bar{R}_1 v + v^2, \tag{3·13}$$

where v, $\bar{R}_1$, c are as in (2·9a), (3·2a) and (3·3) respectively. Equation (3·13) gives an explicit formula for deriving the Bartlett adjustment for λ_n via that for $\bar{\lambda}_n$.

Example 3·1. Let

$$f(x; \theta, m) = (2\pi\theta)^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}(x-m)^2/\theta\}.$$

Then it can be seen that

$$a_{20} = \tfrac{1}{2}\theta_0^{-2}, \quad a_{02} = \theta_0^{-1}, \quad L_{12} = \theta_0^{-2}, \quad L_{10.10.10} = L_{11.11} = \theta_0^{-3},$$
$$L_{01.21} = L_{30} = -L_{22} = 2\theta_0^{-3}, \quad L_{10.20} = L_{10.01.11} = -\theta_0^{-3}, \quad L_{40} = -9\theta_0^{-4},$$
$$L_{10.30} = 3\theta_0^{-4}, \quad L_{20.20} = -L_{10.10.20} = \tfrac{9}{4}\theta_0^{-4}, \quad L_{21} = L_{10.12} = L_{10.02} = 0.$$

Hence, by (2·9a), (3·2), (3·3), (3·10), (3·13),

$$v = 2^{-\frac{1}{2}}, \quad \bar{R}_1 = -\tfrac{5}{3}(2^{-\frac{1}{2}}), \quad \bar{R}_2 = \tfrac{4}{9}, \quad c = -\frac{2^{\frac{1}{2}}}{12}\theta_0^{-1}, \quad \bar{a}(m) = \tfrac{33}{18}, \quad a(m) = \tfrac{1}{3}.$$

In this example, both $\bar{a}(m)$ and $a(m)$ do not depend on m.

ACKNOWLEDGEMENTS

The work of the first author, now on leave from the Indian Statistical Institute, was supported by a grant from the Centre for Management and Development Studies, Indian

Institute of Management, Calcutta. The authors are grateful to the referees for their very constructive suggestions.

Appendix

Approximate cumulants of $\bar{W}_n$

A technique discussed by Mukerjee (1989b) will be helpful in the derivation. From elementary considerations, as by, for example, Chandra & Joshi (1983), first note that the approximate cumulants of S_{10}, under θ_0, are given by

$$k_{1n}(S_{10}) = 0, \quad k_{2n}(S_{10}) = a_{20}, \quad k_{3n}(S_{10}) = n^{-\frac{1}{2}} L_{10.10.10},$$

$$k_{4n}(S_{10}) = n^{-1}(L_{10.10.10.10} - 3a_{20}^2), \quad k_{rn}(S_{10}) = o(n^{-1}) \quad (r \geq 5).$$

Hence the approximate characteristic function of S_{10}, under θ_0, is

$$\{1 + n^{-\frac{1}{2}} F_1(\xi, m) + n^{-1} F_2(\xi, m)\} \exp(\tfrac{1}{2} a_{20} \xi^2) + o(n^{-1}),$$

where

$$\xi = (-1)^{\frac{1}{2}} t, \quad F_1(\xi, m) = \tfrac{1}{6} \xi^3 L_{10.10.10}, \tag{A·1a}$$

$$F_2(\xi, m) = \tfrac{1}{24} \xi^4 (L_{10.10.10.10} - 3a_{20}^2) + \tfrac{1}{72} \xi^6 L_{10.10.10}^2. \tag{A·1b}$$

Let $\theta_n = \theta_0 + n^{-\frac{1}{2}} \delta$, where δ is free from n, and

$$E_{\theta_n, m}(\bar{Q}_1') = \bar{C}_1(\delta, m) + n^{-\frac{1}{2}} \bar{M}_1(\delta, m) + O(n^{-1}),$$

$$E_{\theta_n, m}(\bar{Q}_2') = \bar{C}_2(\delta, m) + O(n^{-\frac{1}{2}}), \quad E_{\theta_n, m}(S_{20} \bar{Q}_1') = \bar{C}_3(\delta, m) + O(n^{-\frac{1}{2}}),$$

$$E_{\theta_n, m}\{(\bar{Q}_1')^2\} = \bar{C}_4(\delta, m) + O(n^{-\frac{1}{2}}),$$

where $\bar{C}_i(\delta, m)$ $(1 \leq i \leq 4)$, $\bar{M}_1(\delta, m)$ are free from n. Then following Lemma 3 of Mukerjee (1989b), the approximate characteristic function of $\bar{W}_n$, under θ_0, is given by

$$E_{\theta_0, m}\{\exp(\xi \bar{W}_n)\} = \chi(\xi, m) \exp(\tfrac{1}{2} \xi^2) + o(n^{-1}), \tag{A·2a}$$

where

$$\begin{aligned} \chi(\xi, m) = 1 &+ n^{-\frac{1}{2}} \{F_1(a_{20}^{-\frac{1}{2}} \xi, m) + \xi \bar{C}_1(a_{20}^{-\frac{1}{2}} \xi, m)\} \\ &+ n^{-1} \{F_2(a_{20}^{-\frac{1}{2}} \xi, m) + \xi \bar{M}_1(a_{20}^{-\frac{1}{2}} \xi, m) + \xi \bar{C}_2(a_{20}^{-\frac{1}{2}} \xi, m) - \tfrac{1}{2} a_{20}^{-1} \xi^3 \bar{C}_3(a_{20}^{-\frac{1}{2}} \xi, m) \\ &\quad + \tfrac{1}{2} \xi^2 \bar{C}_4(a_{20}^{-\frac{1}{2}} \xi, m) - \tfrac{1}{6} a_{20}^{-3/2} \xi^4 L_{30} \bar{C}_1(a_{20}^{-\frac{1}{2}} \xi, m)\}. \end{aligned} \tag{A·2b}$$

Using (2·6), (2·7), calculations similar to those of Mukerjee (1989a) show that

$$\bar{C}_1(\delta, m) = -\tfrac{1}{2}(a_{20}^{\frac{1}{2}} a_{02})^{-1} L_{12} - \tfrac{1}{6} a_{20}^{-3/2} L_{10.10.10}(1 + \delta^2 a_{20}), \tag{A·3a}$$

$$\begin{aligned} \bar{M}_1(\delta, m) = \delta a_{20}^{-3/2} \{ & \tfrac{1}{2}(L_{10.10.20} + a_{20}^2) \\ & + \tfrac{1}{6}(L_{30} L_{10.10.10} / a_{20}) + \tfrac{1}{2}(a_{20} L_{12} L_{10.01.01} / a_{02}^2) \\ & + \tfrac{1}{2}(L_{21} L_{10.10.01} / a_{02}) + (a_{20} L_{10.01.11} / a_{02})\} \\ + \tfrac{1}{2} \delta^3 \{ & a_{20}^{-3/2} (\tfrac{1}{2} L_{10.20}^2 + \tfrac{1}{2} L_{10.10.10} L_{10.20} + \tfrac{1}{2} a_{20} L_{20.20} + \tfrac{1}{2} a_{20} L_{10.10.20} \\ & + \tfrac{1}{3} L_{30} L_{10.20} + \tfrac{1}{3} L_{30} L_{10.10.10}) - \tfrac{1}{2}(a_{20}^{\frac{1}{2}} a_{02})^{-1} L_{21}(L_{20.01} + L_{10.10.01})\}, \end{aligned} \tag{A·3b}$$

$$\begin{aligned} \bar{C}_2(\delta, m) = \delta [& a_{20}^{-5/2} \{\tfrac{3}{8} a_{20}(L_{20.20} - a_{20}^2) + \tfrac{3}{4} L_{10.20}^2 + \tfrac{5}{4} L_{30} L_{10.20} + \tfrac{1}{3} L_{30}^2\} \\ & + a_{20}^{-3/2} (\tfrac{1}{8} L_{40} + \tfrac{1}{2} L_{10.30}) + (a_{20}^{\frac{1}{2}} a_{02})^{-1} (\tfrac{1}{2} L_{11.11} + \tfrac{1}{2} L_{21.01} + \tfrac{1}{4} L_{22} + \tfrac{1}{2} L_{10.12}) \\ & + (a_{20}^{3/2} a_{02})^{-1} (\tfrac{1}{4} L_{21} L_{20.01} - \tfrac{1}{4} L_{21}^2 - \tfrac{1}{6} L_{30} L_{12} - \tfrac{1}{4} L_{12} L_{10.20}) \\ & - (a_{20}^{\frac{1}{2}} a_{02}^2)^{-1} (\tfrac{1}{2} L_{21} L_{01.02} + \tfrac{1}{4} L_{21} L_{03} + \tfrac{1}{2} L_{12}^2)] \\ + \delta^3 \{ & a_{20}^{-3/2} (\tfrac{3}{8} L_{10.20}^2 + \tfrac{5}{12} L_{30} L_{10.20} + \tfrac{1}{9} L_{30}^2) + a_{20}^{-\frac{1}{2}} (\tfrac{1}{24} L_{40} + \tfrac{1}{6} L_{10.30}) + \tfrac{1}{8}(a_{20}^{\frac{1}{2}} a_{02})^{-1} L_{21}^2\}, \end{aligned} \tag{A·3c}$$

$$\bar{C}_3(\delta, m) = \delta\{\tfrac{1}{2}a_{20}^{-\frac{1}{2}}(L_{20.20} - a_{20}^2) + a_{20}^{-3/2}(L_{10.20}^2 + \tfrac{1}{2}L_{30}L_{10.20}) - \tfrac{1}{2}(a_{20}^{\frac{1}{2}}a_{02})^{-1}(L_{21}L_{20.01} + L_{12}L_{10.20})\}$$
$$+ \tfrac{1}{2}\delta^3 a_{20}^{-\frac{1}{2}}L_{10.20}(L_{10.20} + \tfrac{1}{3}L_{30}), \qquad \text{(A·3d)}$$

$$\bar{C}_4(\delta, m) = \{\bar{C}_1(\delta, m)\}^2 + \tfrac{1}{4}a_{20}^{-2}(L_{20.20} - a_{20}^2) + (a_{20}a_{02})^{-1}L_{11.11} + a_{20}^{-3}(\tfrac{1}{4}L_{10.20}^2 + \tfrac{1}{3}L_{10.20}L_{30} + \tfrac{1}{18}L_{30}^2)$$
$$- \tfrac{1}{2}(a_{20}a_{02}^2)^{-1}L_{12}^2 - \tfrac{1}{2}(a_{20}^2a_{02})^{-1}(L_{21}L_{20.01} + \tfrac{3}{2}L_{21}^2)$$
$$+ \delta^2\{\tfrac{1}{4}a_{20}^{-1}(L_{20.20} - a_{20}^2) + a_{20}^{-2}(\tfrac{3}{4}L_{10.20}^2 + \tfrac{2}{3}L_{30}L_{10.20} + \tfrac{1}{9}L_{30}^2)$$
$$- (a_{20}a_{02})^{-1}L_{21}(\tfrac{1}{2}L_{20.01} - \tfrac{1}{4}L_{21})\}. \qquad \text{(A·3e)}$$

From (A·1), (A·2), (A·3), after considerable algebra, which is omitted here, the relations (3·1), (3·2) follow. Simple regularity conditions like the following, some of these being consequences of parametric orthogonality, are useful in the derivation:

$$L_{10.11} = -L_{21}, \quad L_{01.11} = -L_{12}, \quad L_{10.10.10} + 3L_{10.20} + L_{30} = 0,$$
$$L_{40} + 4L_{10.30} + 3L_{20.20} + 6L_{10.10.20} + L_{10.10.10.10} = 0,$$
$$L_{20.01} + L_{10.11} + L_{10.10.01} = 0, \quad L_{11.01} + L_{10.02} + L_{10.01.01} = 0.$$

REFERENCES

BARNDORFF-NIELSEN, O. E. & COX, D. R. (1984). Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator. *J. R. Statist. Soc.* B **46**, 483–95.

BARTLETT, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. R. Soc.* A **160**, 268–82.

BHATTACHARYA, R. N. & GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansions. *Ann. Statist.* **6**, 434–51.

BICKEL, P. J. & GHOSH, J. K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction—a Bayesian argument. *Ann. Statist.* **18**, 1070–90.

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

CHANDRA, T. K. & JOSHI, S. N. (1983). Comparison of the likelihood ratio, Rao's and Wald's tests and a conjecture of C. R. Rao. *Sankhyā* A **45**, 226–46.

COX, D. R. (1988). Some aspects of conditional and asymptotic inference: A review. *Sankhyā* A **50**, 314–37.

COX, D. R. & REID, N. (1987). Parameter orthogonality and approximate conditional inference (with discussion). *J. R. Statist. Soc.* B **49**, 1–39.

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

HARRIS, P. (1987). Discussion of paper by D. R. Cox & N. Reid. *J. R. Statist. Soc.* B **49**, 30.

LEVIN, B. & KONG, F. (1990). Bartlett's bias correction to the profile score is a saddlepoint correction. *Biometrika* **77**, 219–21.

MUKERJEE, R. (1989a). Comparison of tests in the presence of a nuisance parameter. In *Statistical Data Analysis and Inference*, Ed. Y. Dodge, pp. 131–9. Amsterdam: North-Holland.

MUKERJEE, R. (1989b). Third-order comparison of unbiased tests: A simple formula for the power difference in the one-parameter case. *Sankhyā* A **51**, 212–32.

MUKERJEE, R. (1990). Comparison of tests in the multiparameter case II. A third-order optimality property of Rao's test. *J. Mult. Anal.* **33**, 31–48.

[*Received November* 1989. *Revised July* 1990]