Improved score tests for one–parameter exponential family models *

Silvia L.P. Ferrari
Departamento de Estatística
Universidade de São Paulo, Brazil

Gauss M. Cordeiro
Departamento de Estatística
Universidade Federal de Pernambuco, Brazil

Miguel A. Uribe–Opazo
Departamento de Matemática
Universidade Estadual do Oeste do Paraná, Brazil

Francisco Cribari–Neto
Department of Economics
Southern Illinois University, USA

Abstract: Under suitable regularity conditions, an improved score test was derived by Cordeiro and Ferrari (1991). The test is based on a corrected score statistic which has a chi-squared distribution to order n^{-1} under the null hypothesis, where n is the sample size. In this paper we follow their approach and obtain a Bartlett-corrected score statistic for testing $\mathrm{H}_0 : \theta = \theta^{(0)}$, where θ is the scalar parameter of a one-parameter exponential family model. We apply our main result to a number of special cases and derive approximations for corrections that involve unusual functions. We also obtain Bartlett-type corrections for natural exponential families.

Keywords: Bartlett-type correction; chi-squared distribution; exponential family; score statistic; variance function.

1 Introduction

Three commonly used large sample tests are the score (or Lagrange multiplier), likelihood ratio (LR) and Wald (W) tests. These three tests are asymptotically equivalent under the null hypothesis. It has been argued that after modifying their critical regions to force the three tests to have the same size to order n^{-1}, where n is the number of observations, the score test is usually more powerful than the LR and W tests to order n^{-1}; see Chandra and Joshi (1983), Chandra and Mukerjee (1984, 1985) and Mukerjee (1990a, 1990b). Another advantage of the score test is that it only requires estimation of the unknown parameters under the null model. All three tests rely on a first order asymptotic approximation which may be a poor one in samples of small to moderate size. It is possible to improve the chi-squared approximation to the LR test by multiplying the test statistic by a scalar correction factor known as the Bartlett correction; see Lawley (1956), Hayakawa (1977), Cordeiro (1987, 1993a) and the references therein. Recently, Cordeiro and Ferrari (1991) obtained a Bartlett-type correction to the score statistic. Their correction is defined as a second degree polynomial on the score statistic with coefficients that depend on cumulants of log-likelihood derivatives. Both unmodified statistics have a χ^2_q distribution to first order under the null hypothesis H_0, where q denotes the number of restrictions imposed by H_0. On the other hand, the modified statistics have a χ^2_q distribution

* *Correspondence to*: Silvia L.P. Ferrari, Departamento de Estatística, Universidade de São Paulo, Caixa Postal 66281, 05389-970 São Paulo, SP, Brazil.

to order n^{-1} under H_0. In other words, $\Pr[S \le x] = \Pr[\chi^2_q \le x] + o(1)$ and $\Pr[S^* \le x] = \Pr[\chi^2_q \le x] + o(n^{-1})$, where S and S^* represent the unmodified and modified statistics, respectively. The finite-sample accuracy of these corrections has been extensively evaluated through Monte Carlo simulation; see Cordeiro (1993b), Cordeiro and Cribari–Neto (1993), Cordeiro, Cribari–Neto, Aubin and Ferrari (1995), Cordeiro, Ferrari and Paula (1993), Cribari–Neto and Cordeiro (1995), Cribari-Neto and Ferrari (1995a, 1995b), Cribari–Neto and Zarkos (1995) and Ferrari and Cordeiro (1995). Such experiments have shown that Bartlett and Bartlett-type corrections usually lead to improvements in the size behaviour of the likelihood ratio and the score tests even in nonnormal and multivariate regression models.

The main purpose of the present paper is to obtain a general closed-form expression for the Bartlett-type correction to the score statistic in one-parameter exponential family models. This expression can be easily used to derive Bartlett-type corrections for many important distributions. Our aim here is to give a new formula for correcting the score test in one-parameter exponential family models which is algebraically more appealing for applications than the general formulas developed by Harris (1985). Unlike Harris' general formulas, our result can be readily used by applied researchers since it only requires trivial operations on suitably defined functions and their derivatives. It should be mentioned that the results in the present paper can be extended to more general problems, such as, two-parameter exponential family models.

Consider a set of n independent and identically distributed random variables with density (or probability) function

$$\pi(y;\theta) = \frac{1}{\zeta(\theta)} \exp\{-\alpha(\theta) d(y) + v(y)\}, \tag{1}$$

where θ is a scalar parameter, $\zeta(\cdot)$, $\alpha(\cdot)$, $d(\cdot)$ and $v(\cdot)$ are known functions and $\zeta(\cdot)$ is positive valued. We also assume that the support set of this distribution is independent of θ and that $\alpha(\cdot)$ and $\zeta(\cdot)$ have continuous first four derivatives, $\mathrm{d}\alpha(\theta)/\mathrm{d}\theta$ and $\mathrm{d}\beta(\theta)/\mathrm{d}\theta$ being different from zero for all θ in the parameter space. Here,

$$\beta(\theta) = \frac{\zeta'(\theta)}{\zeta(\theta)\alpha'(\theta)}, \tag{2}$$

primes denoting derivatives with respect to θ. This quantity will play an important role in the derivation of the correction.

In the next section, we derive a simple expression for the Bartlett-type correction to the score test in the exponential family (1). This new expression involves only the functions $\alpha(\cdot)$ and $\beta(\cdot)$ and their first three derivatives, and can be easily implemented in a computer algebra system such as MATHEMATICA (Wolfram, 1991) or MAPLE V (Abell and Baselton, 1994). In Section 3, we present a number of special cases thus showing that our main result has a wide range of important applications. Finally, Section 4 considers Bartlett-type corrections for the class of natural exponential families assuming different forms of variance functions.

2 Derivation of the correction

Consider n independent and identically distributed random variables $y_1, \ldots, y_n$ having any regular uniparametric distribution with density (or probability) function $\pi(y;\theta) = \exp\{t(y;\theta)\}$, where θ is a scalar parameter, and define $v_{(r)} = \mathrm{E}\{(t'(y;\theta))^r\}$ and $v_r = \mathrm{E}\{t^{(r)}(y;\theta)\}$, for $r = 1, \ldots, 4$, with $t'(y;\theta) = \mathrm{d}t(y;\theta)/\mathrm{d}\theta$

and $t^{(r)}(y;\theta) = \mathrm{d}^r t(y;\theta)/\mathrm{d}\theta^r$. The v's satisfy certain regularity relations such as $v_1 = 0$, $v_{(2)} = -v_2$, $v_{(3)} = 2v_3 - 3v_2'$ and $v_{(4)} = -3v_4 + 8v_3' - 6v_2'' + 3v_{2(2)}$, where $v_{2(2)} = \mathrm{E}\{(t''(y;\theta))^2\}$ (Lawley, 1956).

The score statistic for testing $\mathrm{H}_0 : \theta = \theta^{(0)}$ against a two-sided alternative can be written as $S_R = (\sum_{l=1}^{n} t'(y_l;\theta^{(0)}))^2/(nv_{(2)})$, where $v_{(2)}$ is evaluated at $\theta^{(0)}$. The evaluation of S_R does not require any estimation since it only depends on the value of θ specified under H_0. Cordeiro and Ferrari (1991) have shown that the modified score statistic

$$S_R^* = S_R\Big\{1 - \frac{1}{n}(c + bS_R + aS_R^2)\Big\}, \tag{3}$$

where

$$a = \frac{\gamma_1^2}{36}, \qquad b = \frac{3\gamma_2 - 10\gamma_1^2}{36}, \qquad c = \frac{5\gamma_1^2 - 3\gamma_2}{12}, \tag{4}$$

has a χ_1^2 distribution to order n^{-1}. Here, $\gamma_1 = v_{(3)}/v_{(2)}^{3/2}$ and $\gamma_2 = (v_{(4)} - 3v_{(2)}^2)/v_{(2)}^2$ are the usual measures of skewness and kurtosis of the score function for a single observation.

Next, let $\pi(y;\theta)$ be defined as in (1) so that $t(y;\theta) = -\log\zeta(\theta) - \alpha(\theta)d(y) + v(y)$. It then follows that $\mathrm{E}\{d(y)\} = -\beta$, $\mathrm{var}\{d(y)\} = \beta'/\alpha'$, $v_2 = -\alpha'\beta'$, $v_3 = -2\alpha''\beta' - \alpha'\beta''$, $v_4 = -3(\alpha'''\beta' + \alpha''\beta'') - \alpha'\beta'''$ and $v_{2(2)} = \alpha''^2\beta'/\alpha' + \alpha'^2\beta'^2$. Hence, the score statistic is given by $S_R = n\alpha'(\beta + \bar{d})^2/\beta'$, where $\bar{d} = n^{-1}\sum d(y_l)$ and β and β' are evaluated at $\theta^{(0)}$. Finally, using again the relations among the v's and plugging them into the expressions for a, b and c in (4), we get

$$a = \frac{(\beta'\alpha'' - \alpha'\beta'')^2}{36\alpha'^3\beta'^3}, \tag{5}$$

$$b = \frac{-\beta'^2\alpha''^2 + 11\alpha'\beta'\alpha''\beta'' - 10\alpha'^2\beta''^2 - 3\alpha'\beta'^2\alpha''' + 3\alpha'^2\beta'\beta'''}{36\alpha'^3\beta'^3}, \tag{6}$$

$$c = \frac{-4\beta'^2\alpha''^2 - \alpha'\beta'\alpha''\beta'' + 5\alpha'^2\beta''^2 + 3\alpha'\beta'^2\alpha''' - 3\alpha'^2\beta'\beta'''}{12\alpha'^3\beta'^3}. \tag{7}$$

Plugging (5)-(7) into (3) and evaluating α', α'', α''', β', β'' and β''' at $\theta^{(0)}$, we get Cordeiro and Ferrari's (1991) correction to the score statistic for testing H_0 in one-parameter exponential family models. Two features of equations (5)-(7) are noteworthy. First, c is equal to the coefficient ρ divided by 12 which determines the correction to the LR statistic (see eq. (4) of Cordeiro, Cribari–Neto, Aubin and Ferrari, 1995), and thus the present paper can be viewed as an extension of their paper. Second, a, b and c depend on the model (1) only through the functions α and β and their first three derivatives.

3 Special cases

In this section, we use equations (5)-(7) to derive Bartlett-type corrections to the score statistic for a number of distributions that belong to the one-parameter exponential family (1). The calculations were done using MATHEMATICA and MAPLE V. It should be remarked that it is not possible to guarantee that the modified statistic has a chi-squared distribution to order n^{-1} when the distribution is discrete. The special cases listed bellow have a wide range practical applications in various fields such as engineering, biology, medicine, economics, among others (Johnson and Kotz, 1970a, 1970b; Johnson, Kotz and Kemp, 1992).

The following distributions are considered:

(i) Binomial ($0<\theta<1$, $m\in I\!N$, m known, $y=0,1,2,\ldots,m$): $\alpha(\theta)=-\log\{\theta/(1-\theta)\}$, $\zeta(\theta)=(1-\theta)^{-m}$, $d(y)=y$, $v(y)=\log\binom{m}{y}$,

$$a=\frac{(2\theta-1)^2}{36m\theta(1-\theta)},\quad b=-\frac{22\theta(\theta-1)+7}{36m\theta(1-\theta)},\quad c=-\frac{\theta(1-\theta)-1}{6m\theta(1-\theta)}.$$

(ii) Negative binomial ($0<\theta<1$, $\gamma>0$, γ known, $y=0,1,2,\ldots$): $\alpha(\theta)=-\log\theta$, $\zeta(\theta)=(1-\theta)^{-\gamma}$, $d(y)=y$, $v(y)=\log\binom{\gamma+y-1}{y}$,

$$a=\frac{(\theta+1)^2}{36\gamma\theta},\quad b=-\frac{7\theta^2+8\theta+7}{36\gamma\theta},\quad c=\frac{1-\theta(1-\theta)}{6\gamma\theta}.$$

(iii) Poisson ($\theta>0$, $y=0,1,2,\ldots$): $\alpha(\theta)=-\log(\theta)$, $\zeta(\theta)=\exp\{\theta\}$, $d(y)=y$, $v(y)=-\log(y!)$, $a=1/(36\theta)$, $b=-7/(36\theta)$, $c=1/(6\theta)$.

(iv) Truncated Poisson ($\theta>0, y=1,2,\ldots$): $\alpha(\theta)=-\log(\theta)$, $\zeta(\theta)=e^{\theta}(1-e^{-\theta})$, $d(y)=y$, $v(y)=-\log(y!)$,

$$a=-\{\theta^2+3\theta+1+e^{\theta}(\theta^2-3\theta-2)+e^{2\theta}\}^2/\{36\theta e^{\theta}(1+\theta-e^{\theta})^3\},$$

$$\begin{aligned} b &= \{7+36\theta+71\theta^2+39\theta^3+7\theta^4+e^{\theta}(-28-108\theta-140\theta^2-9\theta^3+8\theta^4)\\ &+ e^{2\theta}(42+108\theta+67\theta^2-33\theta^3+7\theta^4)+e^{3\theta}(-28-36\theta+2\theta^2+3\theta^3)+7e^{4\theta}\}/\\ &\{36\theta e^{\theta}(1+\theta-e^{\theta})^3\},\end{aligned}$$

$$\begin{aligned} c &= -\{2+6\theta+16\theta^2+9\theta^3+2\theta^4+e^{\theta}(-8-18\theta-40\theta^2-9\theta^3-2\theta^4)\\ &+ e^{2\theta}(12+18\theta+32\theta^2-3\theta^3+2\theta^4)+e^{3\theta}(-8-6\theta-8\theta^2+3\theta^3)+2e^{4\theta}\}/\\ &\{12\theta e^{\theta}(1+\theta-e^{\theta})^3\}.\end{aligned}$$

(v) Logarithmic series ($0<\theta<1$, $y=1,2,\ldots$): $\alpha(\theta)=-\log(\theta)$, $\zeta(\theta)=-\log(1-\theta)$, $d(y)=y$, $v(y)=-\log(y)$,

$$a=-\{\log(1-\theta)(3\theta+\theta\log(1-\theta)+\log(1-\theta))+2\theta^2\}^2/\{36\theta(\theta+\log(1-\theta))^3\}.$$

$$\begin{aligned} b &= [22\theta^3(\theta+3\log(1-\theta))+\theta^2\{\log(1-\theta)\}^2(28\theta+73)+3\theta\{\log(1-\theta)\}^3(12\theta+12-\theta^2)\\ &+ \{\log(1-\theta)\}^4(7+8\theta+7\theta^2)]/\{36\theta(\theta+\log(1-\theta))^3\},\end{aligned}$$

$$\begin{aligned} c &= -[2\theta^3(\theta+3\log(1-\theta))+8\theta^2\{\log(1-\theta)\}^2(1+\theta)+3\theta\{\log(1-\theta)\}^3(2+2\theta-\theta^2)\\ &+ 2\{\log(1-\theta)\}^4(\theta^2-\theta+1)]/\{12\theta(\theta+\log(1-\theta))^3\}.\end{aligned}$$

(vi) Power Series ($\theta>0$, $a_y\geq 0$, $y=0,1,2,\ldots$): $\alpha=-\log(\theta)$, $\zeta(\theta)=\sum_{y=0}^{\infty}a_y\theta^y$, $d(y)=y$, $v(y)=\log(a_y)$,

$$\begin{aligned} a &= \frac{(g+3\theta g'+\theta^2g'')^2}{36\theta(g+\theta g')^3},\\ b &= -\frac{7g^2+36\theta gg'+\theta^2(69g'^2+2gg'')+3\theta^3(14g'g''-gg''')+\theta^4(10g''^2-3g'g''')}{36\theta(g+\theta g')^3},\\ c &= \frac{2g^2+6\theta gg'+8\theta^2(3g'^2-gg'')+3\theta^3(g'g''-gg''')+\theta^4(5g''^2-3gg''')}{12\theta(g+\theta g')^3},\end{aligned}$$

where $g = g(\theta) = \mathrm{d}\log\zeta(\theta)/\mathrm{d}\theta$. Note that cases (ii), (iii), (iv) and (v) can be obtained from this case by simple specification of the function $g(\cdot)$.

(vii) Zeta ($\theta > 0$, $y = 1, 2, 3, \ldots$): $\alpha(\theta) = \theta + 1$, $\zeta(\theta) = \mathrm{Zeta}(\theta + 1)$, $d(y) = \log(y)$, $v(y) = 0$,

$$a = \frac{g''^2}{36g'^3},\ b = -\frac{10g''^2 - 3g'g'''}{36g'^3},\ c = \frac{5g''^2 - 3g'g'''}{12g'^3},$$

where ζ is the Riemann zeta-function, *i.e.*, $\zeta(\theta) = \mathrm{Zeta}(\theta + 1) = \sum_{i=1}^{\infty} i^{-(\theta+1)}$ (see, *e.g.*, Patterson, 1988) and $g = g(\theta) = \mathrm{d}\log\mathrm{Zeta}(\theta + 1)/\mathrm{d}\theta$.

(viii) Non-central hypergeometric ($\theta > 0$, m_1, m_2, r are known positive integers, $k_1 = \max\{0, r - m_2\} \leq y \leq \min\{m_1, r\} = k_2$): $\alpha(\theta) = \theta$, $\zeta(\theta) = D_0(\theta)$, $d(y) = -y$, $v(y) = \log\{\binom{m_1}{y}\binom{m_2}{r-y}\}$,

$$a = (D_0^2 D_3 - 3D_0 D_1 D_2 + 2D_1^3)^2/\{36(D_1^2 - D_0 D_2)^3\},$$

$$\begin{aligned} b &= (10D_0^4 D_3^2 - 48D_0^3 D_1 D_2 D_3 + 28D_0^2 D_1^3 D_3 + 45D_0^2 D_1^2 D_2^2 - 66D_0 D_1^4 D_2 + 22D_1^6 - 3D_0^4 D_2 D_4 \\ &+ 9D_0^3 D_2^3 + 3D_0^3 D_1^2 D_4)/\{36(D_1^2 - D_0 D_2)^3\}, \end{aligned}$$

$$\begin{aligned} c &= (-5D_0^4 D_3^2 + 18D_0^3 D_1 D_2 D_3 - 8D_0^2 D_1^3 D_3 + 6D_0 D_1^4 D_2 - 2D_1^6 + 3D_0^4 D_2 D_4 - 9D_0^3 D_2^3 \\ &- 3D_0^3 D_1^2 D_4)/\{12(D_1^2 - D_0 D_2)^3\}, \end{aligned}$$

where $D_p = D_p(\theta) = \sum_{y=k_1}^{k_2} y^p \binom{m_1}{y}\binom{m_2}{r-y} \exp\{\theta y\}$, $p = 0, 1, 2, 3, 4$.

(ix) Maxwell ($\theta > 0$, $y > 0$): $\alpha(\theta) = (2\theta^2)^{-1}$, $\zeta(\theta) = \theta^3$, $d(y) = y^2$, $v(y) = \log(y^2\sqrt{2/\pi})$, $a = 2/27$, $b = -11/27$, $c = 1/9$.

(x) Gamma ($k > 0, \theta > 0$, $y > 0$):

(a) k known: $\alpha(\theta) = \theta$, $\zeta(\theta) = \theta^{-k}$, $d(y) = y$, $v(y) = (k-1)\log(y) - \log\{\Gamma(k)\}$, $a = 1/(9k)$, $b = -11/(18k)$, $c = 1/(6k)$.

(b) θ known: $\alpha(k) = 1 - k$, $\zeta(k) = \theta^{-k}\Gamma(k)$, $d(y) = \log(y)$, $v(y) = -\theta y$,

$$a = \frac{\psi''(k)^2}{36\psi'(k)^3},\quad b = \frac{-10\psi''(k)^2 + 3\psi'(k)\psi'''(k)}{36\psi'(k)^3},\quad c = \frac{5\psi''(k)^2 - 3\psi'(k)\psi'''(k)}{12\psi'(k)^3},$$

where $\Gamma(\cdot)$ and $\psi(\cdot)$ are the gamma and digamma functions, respectively.

(xi) Burr system of distributions ($\theta > 0$, $b > 0$, b known, $y > 0$): $\alpha(\theta) = \theta$, $\zeta(\theta) = g(\theta)/\theta$, $d(y) = -\log G(y)$, $v(y) = \log\{|\mathrm{d}\log G(y)/\mathrm{d}y|\}$, $a = 1/9$, $b = -11/18$, $c = 1/6$, where the functions $g(.)$ and $G(.)$ must be positive. Different choices for $g(\theta)$ and $G(y)$ lead to different distributions; see Burr (1942). Burr I and Burr IX distributions do not belong to the exponential family.

(xii) Rayleigh ($\theta > 0$, $y > 0$): $\alpha(\theta) = \theta^{-2}$, $\zeta(\theta) = \theta^2$, $d(y) = y^2$, $v(y) = \log(2y)$, $a = 1/9$, $b = -11/18$, $c = 1/6$.

(xiii) Pareto ($\theta > 0$, $k > 0$, k known, $y > k$): $\alpha(\theta) = \theta + 1$, $\zeta(\theta) = (\theta k^{\theta})^{-1}$, $d(y) = \log(y)$, $v(y) = 0$, $a = 1/9$, $b = -11/18$, $c = 1/6$.

(xiv) Weibull ($\theta > 0,\ \ \phi > 0,\ \ \phi$ known, $\ y > 0$): $\alpha(\theta) = \theta^{-\phi}$, $\zeta(\theta) = \theta^{\phi}$, $d(y) = y^{\phi}$, $v(y) = \log(\phi) + (\phi - 1)\log(y)$, $a = 1/9,\ b = -11/18,\ c = 1/6$.

(xv) Power ($\theta > 0,\ \phi > 0,\ \phi$ known, $y > \phi$): $\alpha(\theta) = 1 - \theta$, $\zeta(\theta) = \theta^{-1}\phi^{\theta}$, $d(y) = \log(y)$, $v(y) = 0$, $a = 1/9,\ b = -11/18,\ c = 1/6$.

(xvi) Laplace ($\theta > 0,\ -\infty < k < \infty$, k known, $y > 0$): $\alpha(\theta) = \theta^{-1}$, $\zeta(\theta) = 2\theta$, $d(y) = |y - k|$, $v(y) = 0$, $a = 1/9,\ b = -11/18,\ c = 1/6$.

(xvii) Extreme value ($-\infty < \theta < \infty,\ \phi > 0,\ \phi$ known, $-\infty < y < \infty$): $\alpha(\theta) = \exp\{\theta/\phi\}$, $\zeta(\theta) = \phi\exp\{-\theta/\phi\}$, $d(y) = \exp\{-y/\phi\}$, $v(y) = -y/\phi$, $a = 1/9,\ b = -11/18,\ c = 1/6$.

(xviii) Truncated extreme value ($\theta > 0,\ y > 0$): $\alpha(\theta) = \theta^{-1}$, $\zeta(\theta) = \theta$, $d(y) = \exp\{y\} - 1$, $v(y) = y$, $a = 1/9,\ b = -11/18,\ c = 1/6$.

(xix) Lognormal ($\theta > 0,\ -\infty < \mu < \infty$, μ known, $y > 0$): $\alpha(\theta) = \theta^{-2}$, $\zeta(\theta) = \theta$, $d(y) = (\log y - \mu)^2/2$, $v(y) = -\log y - \{\log(2\pi)\}/2$, $a = 2/9,\ b = -11/9,\ c = 1/3$.

(xx) Normal ($\theta > 0,\ -\infty < \mu < \infty,\ -\infty < y < \infty$):

(a) μ known: $\alpha(\theta) = (2\theta)^{-1}$, $\zeta(\theta) = \theta^{1/2}$, $d(y) = (y - \mu)^2$, $v(y) = -\{\log(2\pi)\}/2$, $a = 2/9,\ b = -11/9,\ c = 1/3$.

(b) θ known: $\alpha(\mu) = -\mu/\theta$, $\zeta(\mu) = \exp\{\mu^2/(2\theta)\}$, $d(y) = y$, $v(y) = -\{y^2 + \log(2\pi\theta)\}/2$, $a = 0$, $b = 0,\ c = 0$.

(xxi) Inverse Gaussian ($\theta > 0,\ \mu > 0,\ y > 0$):

(a) μ known: $\alpha(\theta) = \theta$, $\zeta(\theta) = \theta^{-1/2}$, $d(y) = (y - \mu)^2/(2\mu^2 y)$, $v(y) = -\{\log(2\pi y^3)\}/2$, $a = 2/9$, $b = -11/9,\ c = 1/3$.

(b) θ known: $\alpha(\mu) = \theta/(2\mu^2)$, $\zeta(\mu) = \exp\{-\theta/\mu\}$, $d(y) = y$, $v(y) = -\theta/(2y) + [\log\{\theta/(2\pi y^3)\}]/2$, $a = \mu/(4\theta),\ b = -5\mu/(4\theta),\ c = 0$.

(xxii) McCullagh ($\theta > -1/2,\ -1 \le \mu \le 1$, μ known, $0 < y < 1$): $\alpha(\theta) = -\theta$, $\zeta(\theta) = 4^{-\theta}\,\mathrm{B}(\theta + 1/2, 1/2)$, $d(y) = \log[y(1-y)/\{(1+\mu)^2 - 4\mu y\}]$, $v(y) = -[\log\{y(1-y)\}]/2$,

$$a = \frac{(\psi''(\theta+1) - \psi''(\theta+0.5))^2}{36(\psi'(\theta+0.5) - \psi'(\theta+1))^3},$$

$$b = \frac{-10(\psi''(\theta+1) - \psi''(\theta+0.5))^2 + 3(\psi'(\theta+1) - \psi'(\theta+0.5))(\psi'''(\theta+1) - \psi'''(\theta+0.5))}{36(\psi'(\theta+0.5) - \psi'(\theta+1))^3},$$

$$c = \frac{5(\psi''(\theta+1) - \psi''(\theta+0.5))^2 - 3(\psi'(\theta+1) - \psi'(\theta+0.5))(\psi'''(\theta+1) - \psi'''(\theta+0.5))}{12(\psi'(\theta+0.5) - \psi'(\theta+1))^3},$$

where $\mathrm{B}(\cdot,\cdot)$ is the beta function (see McCullagh, 1989).

(xxiii) von Mises ($\theta > 0,\ 0 < \mu < 2\pi$, μ known, $0 < y < 2\pi$): $\alpha(\theta) = -\theta$, $\zeta(\theta) = 2\pi I_0(\theta)$, $d(y) = \cos(y - \mu)$, $v(y) = 0$,

$$a = \frac{r''(\theta)^2}{36r'(\theta)^3}, \quad b = \frac{-10r''(\theta)^2 + 3r'(\theta)r'''(\theta)}{36r'(\theta)^3}, \quad c = \frac{5r''(\theta)^2 - 3r'(\theta)r'''(\theta)}{12r'(\theta)^3}.$$

where $I_\nu(\cdot)$ is the modified Bessel function of first kind and νth order, and $r(\theta) = I_0'(\theta)/I_0(\theta)$.

(xxiv) Generalized hyperbolic secant ($-\pi/2 \le \theta \le \pi/2$, $r > 0$, r known, $0 < y < 1$): $\alpha(\theta) = \theta$, $\zeta(\theta) = \pi(\sec\theta)^r$, $d(y) = -\log\{y/(1-y)\}/\pi$, $v(y) = -\log\{y(1-y)\}/2$,

$$a = \frac{1-(\cos\theta)^2}{9r}, \quad b = -\frac{11-14(\cos\theta)^2}{18r}, \quad c = \frac{1-4(\cos\theta)^2}{6r}.$$

Among the cases studied here, the following distributions were previously considered by Cordeiro, Ferrari and Paula (1993): binomial, Poisson, gamma (case a), inverse Gaussian (case b) and normal (case b). The latter case corresponds to testing the mean of a normal distribution with known variance and is the only one for which $a = b = c = 0$ which is in agreement with the fact that in this case S_R has an exact χ_1^2 null distribution. An interesting feature of the results obtained here is that several cases correspond to the following constants $a = 1/9$, $b = -11/18$ and $c = 1/6$, thus implying that $S_R^* = S_R\{1-(3-11S_R+2S_R^2)/(18n)\}$. Since a, b and c determine the expansion to the distribution function of the score statistic to order $O(n^{-1})$ under the null hypothesis (see Cordeiro and Ferrari, 1991), an interesting conclusion is that the score statistics for such examples have the same distribution to order $O(n^{-1})$ and not only to order $O(n^0)$. It is easy to verify that a, b and c equal the constants given above when one of the following conditions holds: (a) $\alpha(\theta)\zeta(\theta) = c_1$ or (b) $\alpha(\theta) = c_1\theta + c_2$, ($c_1$ and c_2 are usually equal to 1 or -1) and $\zeta(\theta) = c_3/(\theta c_4^\theta)$, where $c_1, \ldots, c_4$ are known scalars. These conditions are individually sufficient, but not necessary.

For some of the special cases considered here, the correction has a very simple form and for some of them the correction does not even depend on θ. In some cases, however, the correction is a very complicated function of θ (*e.g.*, McCullagh, von Mises and zeta distributions) requiring the evaluation of Bessel, polygamma and zeta functions. In order to simplify the evaluation of the Bartlett-type correction in such cases, we shall derive simple approximations for a, b and c.

We start by considering the corrections that involve polygamma functions. For large values of k,

$$\psi'(k) = \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} - \frac{1}{30k^5} + \frac{1}{42k^7} + O(k^{-9}).$$

Then, for the gamma distribution (θ known)

$$a = \frac{1}{36k} + \frac{1}{72k^2} + O(k^{-4}), \quad b = -\frac{1}{9k} - \frac{1}{18k^2} - \frac{1}{72k^3} + O(k^{-4}), \quad c = -\frac{1}{12k} - \frac{1}{24k^2} + \frac{1}{24k^3} + O(k^{-4})$$

for large k.

For the McCullagh distribution, we use the equation $\psi'(z+1) - \psi'(z+1/2) = 2\psi'(z) - 4\psi'(2z) - z^{-2}$, to obtain

$$a = \frac{2}{9} - \frac{1}{24\theta^2} + O(\theta^{-3}), \quad b = -\frac{11}{9} + \frac{1}{6\theta^2} + O(\theta^{-3}), \quad c = \frac{1}{3} + \frac{3}{8\theta^2} + O(\theta^{-3}),$$

for large values of θ. For the von Mises case, the correction involves the function $r(\cdot)$ and its first three derivatives. For large values of θ (Abramowitz and Stegun, 1970, pp.416-421)

$$r(\theta) = 1 - \frac{1}{2\theta} - \frac{1}{8\theta^2} - \frac{1}{8\theta^3} - \frac{25}{128\theta^4} - \frac{13}{32\theta^5} + \cdots$$

By making use of this expansion, we get

$$a = \frac{2}{9} + \frac{1}{8\theta^2} + \frac{23}{48\theta^3} + O(\theta^{-4}), \quad b = -\frac{11}{9} - \frac{1}{2\theta^2} - \frac{77}{48\theta^3} + O(\theta^{-4}), \quad c = \frac{1}{3} - \frac{3}{8\theta^2} - \frac{19}{8\theta^3} + O(\theta^{-4})$$

for large θ. For small θ, we have (Mardia, 1972, p.63)

$$r(\theta) = \frac{\theta}{2}\left\{1 - \frac{\theta^2}{8} + \frac{\theta^4}{48} - \cdots\right\}$$

and it then follows that

$$a = \frac{\theta^2}{32} + O(\theta^4), \quad b = -\frac{1}{8} - \frac{19\theta^2}{96} + O(\theta^4), \quad c = \frac{3}{8} + \frac{\theta^2}{8} + O(\theta^4).$$

Finally, consider the zeta distribution and let

$$\gamma_j = \lim_{m\to\infty}\left\{\sum_{k=1}^{m}\frac{(\log k)^j}{k} - \frac{(\log m)^{j+1}}{j+1}\right\},$$

$j = 0, 1, 2, 3$, γ_0 being Euler's constant, *i.e.*, $\gamma_0 \approx 0.577$. It is possible to obtain

$$a = \frac{1}{9} + \frac{1}{3}(\gamma_0^2 + 2\gamma_1)\theta^2 - \frac{4}{9}(2\gamma_0^3 + 6\gamma_0\gamma_1 + 3\gamma_2)\theta^3 + \frac{1}{9}(21\gamma_0^4 + 84\gamma_0^2\gamma_1 + 54\gamma_1^2 + 30\gamma_0\gamma 2 + 10\gamma_3)\theta^4 + O(\theta^5),$$

$$\begin{aligned} b &= -\frac{11}{18} - \frac{7}{3}(\gamma_0^2 + 2\gamma_1)\theta^2 + \frac{31}{9}(2\gamma_0^3 + 6\gamma_0\gamma_1 + 3\gamma_2)\theta^3 - \frac{1}{9}(174\gamma_0^4 + 696\gamma_0^2\gamma_1 + 441\gamma_1^2 + 255\gamma_0\gamma_2 \\ &+ 85\gamma_3)\theta^4 + O(\theta^5), \end{aligned}$$

$$c = \frac{1}{6} + 2(\gamma_0^2 + 2\gamma_1)\theta^2 - \frac{11}{3}(2\gamma_0^3 + 6\gamma_0\gamma_1 + 3\gamma_2)\theta^3 + \frac{1}{3}(69\gamma_0^4 + 276\gamma_0^2\gamma_1 + 171\gamma_1^2 + 105\gamma_0\gamma_2 + 35\gamma_3)\theta^4 + O(\theta^5)$$

for small values of θ. It is possible to use MAPLE V (Abell and Baselton, 1994) to numerically evaluate the γ's and simplify these expansions as

$$\begin{aligned} a &= \frac{1}{9} + \frac{7099}{113556}\theta^2 - \frac{15043}{327410}\theta^3 + \frac{24134}{42607}\theta^4 + O(\theta^5), \\ b &= -\frac{11}{18} - \frac{18068}{29741}\theta^2 + \frac{29835}{83788}\theta^3 - \frac{24407}{62468}\theta^4 + O(\theta^5), \\ c &= \frac{1}{6} + \frac{7099}{18926}\theta^2 - \frac{137395}{362472}\theta^3 + \frac{17629}{39036}\theta^4 + O(\theta^5). \end{aligned}$$

The approximations given above are not expected to work well in practice for all values of θ in the parameter space. Some of them are good approximations for "large" θ and the others, for "small" θ. Strictly speaking, an approximated corrected statistic obtained through these approximations does not have a χ_1^2 distribution to order n^{-1} under H_0. However, if the approximations are good, the approximated corrected statistic and the exact corrected statistic are approximately equal, any difference being negligible. An interesting question is then: How large or how small θ should be for obtaining good approximations to a, b and c? To shed some light on this issue, we present plots of a, b and c against θ in Figure 1 for the gamma and McCullagh distributions. For the gamma distribution, our approximation works well when $k \geq 2$, whereas for the McCullagh distribution the approximations become reliable when, say, $\theta \geq 4$. A similar graphical analysis can be performed for the other approximations obtained above. This is not done to save space. For the von Mises distribution, the approximations work well for $\theta \leq 1.5$ and $\theta \geq 5$ for small and large values of θ, respectively. The zeta distribution requires $\theta \leq 0.3$ for the approximations to work well.

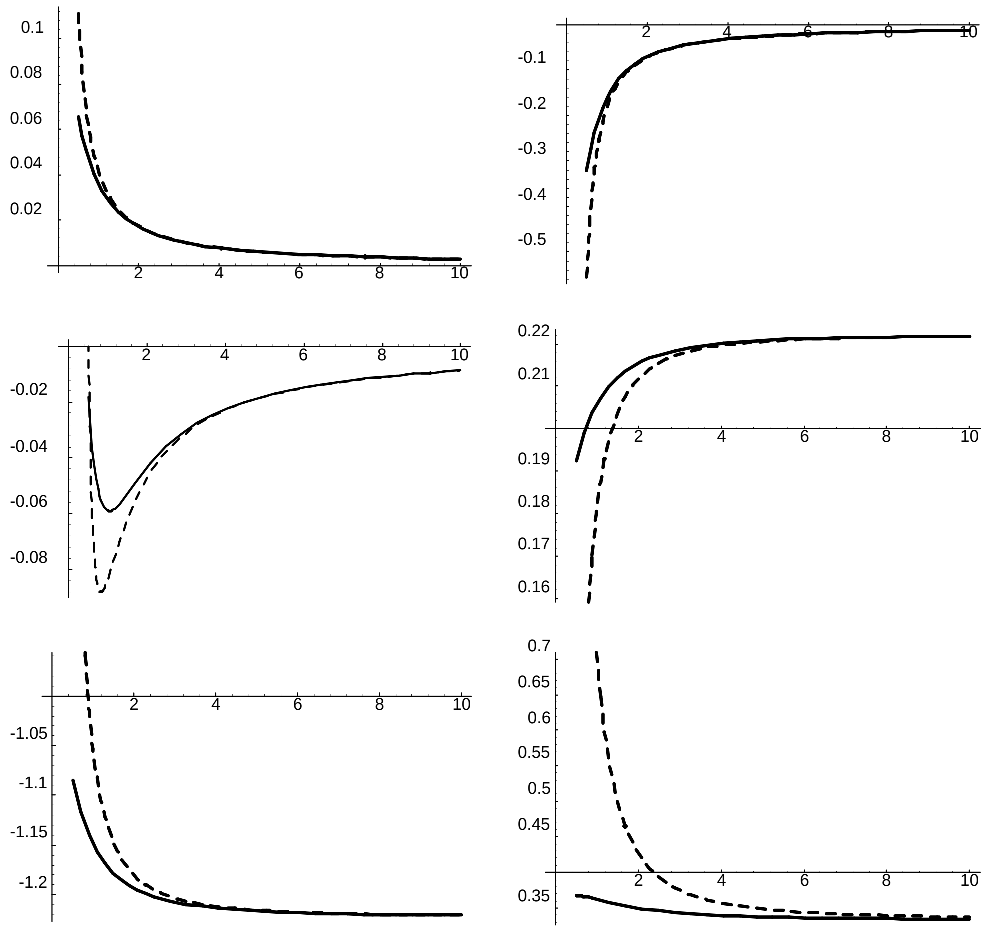

Figure 1. Aproximations for the gamma (θ known) and McCullagh distributions. The top two panels show the approximations for a and b for the gamma distribution (small k), the two panels on the second row show the approximations for c for the gamma distribution (small k) and for a for the McCullagh distribution (large θ), whereas approximations for b and c for the McCullagh distribution (large θ) are given in the bottom two panels, respectively. Solid lines indicate exact values and dashed lines indicate approximations.

4 Natural exponential family

In Section 3, we presented the Bartlett-type correction for the score test in the one-parameter exponential family parameterized in terms of the parameter θ (see eq. (1)). This parameterization is quite convenient for studying how the correction varies with $\theta^{(0)}$. However, it is also informative to write the correction for the exponencial family parameterized in the natural form. By doing this, it is possible to write the coefficients a, b and c in a very ellegant and simple way, and also give an interpretation for these coefficients.

The one-parameter natural exponential family is written as

$$\pi(y;\alpha) = \frac{1}{\delta(\alpha)}\exp\{-\alpha d(y) + v(y)\}, \tag{8}$$

where α is the natural parameter and $-d(y)$ is the canonical statistic. Since (1) and (8) define a one-to-one correspondence between θ and α, the coefficients a, b and c given in (5)-(7) reduce to the corresponding expressions in (4) with $\gamma_1^2 = \beta''^2/\beta'^3$ and $\gamma_2 = \beta'''/\beta'^2$, where $\beta = \mathrm{d}\log\{\delta(\alpha)\}/\mathrm{d}\alpha$, with primes denoting derivatives with respect to α. In order to give an interpretation for the coefficients a, b and c, it should be noticed that for distributions in (8) one has that $\delta(\alpha)$ is the cumulant generator of $-d(y)$ with the $(r+1)$th cumulant given by $\mathrm{d}^r\beta/\mathrm{d}\alpha^r$. Hence, γ_1^2 and γ_2 are the third and fourth standardized cumulants of $-d(y)$. Then, a is proportional to γ_1^2 which is the usual measure of skewness of $-d(y)$ and c is proportional to $5\gamma_1^2 - 3\gamma_2$ and may be viewed as a measure of nonnormality or noninverse normality of $-d(y)$ (see McCullagh and Cox, 1986). The coefficient b is a linear combination of a and c.

When the cumulant generating function of $-d(y)$ can be written as $\delta(\alpha) = \exp\{c_0 + c_1\alpha + c_2\alpha^2\}$, where c_0, c_1 and c_2 are arbitrary constants, we have that $a = b = c = 0$, that is, the Bartlett-type correction vanishes. This is the case for the normal distribution with known variance. On the other hand, if

$$\delta(\alpha) = \exp\{c_0\alpha\}(k\alpha + c_1)^{-1/(9k)}c_2, \tag{9}$$

where c_0, c_1 and c_2 are any real numbers and $k > 0$, we have that $a = k$, $b = -11k/2$ and $c = 3k/2$. For 13 of the cases considered in the previous section, a, b and c are constant with $a > 0$, and it is possible to show that their cumulant generating functions satisfy (9) and that the above relations hold.

Next, we shall derive simple expressions for the coefficients a, b and c in terms of the mean β of the canonical statistic $-d(y)$ for some subclasses of distributions in the one-parameter exponential family. First, it should be noticed that there is, up to a linear tranformation, only one distribution in (8) with a specified variance function (Jørgensen, 1987) and hence β' uniquely characterizes any distribution in (8). The following families of variance functions are considered ($c_0, \ldots, c_3 \in I\!R$) :

(1) Power variance function: $\beta' = \beta^p/c_0$, for $p \le 0$ and $p \ge 1$, $c_0 > 0$:

$$a = \frac{p^2\beta^{p-2}}{36c_0}, \quad b = -\frac{p(4p+3)\beta^{p-2}}{36c_0}, \quad c = \frac{p(3-p)\beta^{p-2}}{12c_0}.$$

Here, $p = 0, 1, 2, 3$ for the normal, Poisson, gamma and inverse Gaussian distributions, respectively. For other distributions in this class, see Jørgensen (1987).

(2) Quadratic variance function: $\beta' = c_0 + c_1\beta + c_2\beta^2$, where a, b and c are given by the corresponding expressions for the cubic variance function (see below) with $c_3 = 0$. Morris (1982) showed that there

are only six distributions in this class, namely: normal ($c_0 = \theta$, $c_1 = c_2 = 0$, $\theta > 0$, θ known), Poisson ($c_0 = c_2 = 0$, $c_1 = 1$), binomial ($c_0 = 0$, $c_1 = 1$, $c_2 = -1/m$, $m \in I\!N$, m known), negative binomial ($c_0 = 0$, $c_1 = 1$, $c_2 = 1/\gamma$, $\gamma > 0$, known), gamma ($c_0 = c_1 = 0$, $c_2 = 1/k$, $k > 0$ known) and generalized hyperbolic secant ($c_0 = r$, $c_1 = 0$, $c_2 = 1/r$, $r > 0$ known).

(3) Cubic variance function: $\beta' = c_0 + c_1\beta + c_2\beta^2 + c_3\beta^3$:

$$a = (36\beta')^{-1}\{c_1^2 - 4c_0c_2 + c_3\beta^2(6c_1 + 8c_2\beta + 9c_3\beta^2)\} + 4c_2,$$

$$b = (36\beta')^{-1}\{-7(c_1^2 - 4c_0c_2) + c_3\beta(18c_0 - 24c_1\beta - 38c_2\beta^2 - 45c_3\beta^3)\} - 22c_2,$$

$$c = (6\beta')^{-1}\{c_1^2 - 4c_0c_2 - c_3\beta(9c_0 + 3c_1\beta + c_2\beta^2)\} + 2c_2.$$

Letac and Mora (1990) showed that there are only six distributions in this class with $c_3 \neq 0$, namely: Abel ($c_0 = 0$, $c_1 = 1$, $c_2 = 2/p$, $c_3 = 1/p^2$, $p > 0$ known), Takács ($c_0 = 0$, $c_1 = 1$, $c_2 = (2m+1)/(mp)$, $c_3 = (m+1)/(mp^2)$, $p > 0$ and $m > 0$ are known), strict arcsine ($c_0 = c_2 = 0$, $c_1 = 1$, $c_3 = 1/p^2$, p known), large arcsine ($c_0 = 0$, $c_1 = 1$, $c_2 = 2/(mp)$, $c_3 = (1+m^2)/(mp)^2$, $p > 0$ and $m > 0$ are known), Ressel ($c_0 = c_1 = 0$, $c_2 = 1/p$, $c_3 = 1/p^2$, $p > 0$ p known) and inverse Gaussian ($c_0 = c_1 = c_2 = 0$, $c_3 = -1/\theta$, $\theta > 0$, θ known).

(4) Babel class: $\beta' = c_0\Delta + (c_1\beta + c_2)\Delta^{1/2}$, where Δ is a polynomial of degree smaller than 3 which is not a perfect square. For this class a, b and c are easily obtained using MATHEMATICA but the resulting expressions are very cumbersome and consequently they are not presented here.

Acknowledgement

We thank an anonymous referee for suggestions that led to improvements in our paper. The partial financial support of CNPq and CAPES/Brazil is also gratefully acknowledged.

References

[1] Abell, M.L. and J.P. Baselton (1994), *The MAPLE V Handbook* (AP Professional, New York).

[2] Abramowitz, M. and I.R. Stegun (1970), *Handbook of Mathematical Functions* (National Bureau of Standards, New York).

[3] Burr, I.W. (1942), Cumulative frequency functions. *Annals of Mathematical Statistics* **13**, 215-232.

[4] Chandra, T.K. and S.N. Joshi (1983), Comparison of the likelihood ratio, Rao's and Wald's tests and a conjecture of C.R. Rao. *Sankhyā A* **45**, 226-246.

[5] Chandra, T.K. and R. Mukerjee (1984), On the optimality of Rao's statistic. *Communications in Statistics—Theory and Methods* **13**, 1507-1515.

[6] Chandra, T.K. and R. Mukerjee (1985) Comparison of the likelihood ratio, Wald's and Rao's tests. *Sankhyā A* **47**, 270-284.

[7] Cordeiro, G.M. (1987), On the corrections to the likelihood ratio statistics. *Biometrika* **74,** 265-274.

[8] Cordeiro, G.M. (1993a), General matrix formulae for computing Bartlett corrections. *Statistics and Probability Letters* **16**, 11-18.

[9] Cordeiro, G.M. (1993b), Bartlett corrections and bias reduction for two heteroscedastic regression models. *Communications in Statistics—Theory and Methods* **22**, 169-182.

[10] Cordeiro, G.M. and F. Cribari–Neto (1993), On bias reduction, Bartlett corrections and a new class of transformations. Brazilian Journal of Probability and Statistics **7**, 179-200.

[11] Cordeiro, G.M., F. Cribari–Neto, E.C.Q. Aubin and S.L.P. Ferrari (1995), Bartlett corrections for one-parameter exponential family models. *Journal of Statistical Computation and Simulation*, forthcoming.

[12] Cordeiro, G.M. and S.L.P. Ferrari (1991), A modified score test statistic having chi-squared distribution to order n^{-1}. *Biometrika* **78**, 573-582.

[13] Cordeiro, G.M., S.L.P. Ferrari and G.A. Paula (1993), Improved score tests for generalized linear models. *Journal of the Royal Statistical Society B* **55**, 661-674.

[14] Cribari–Neto, F. and G.M. Cordeiro (1995), On Bartlett and Bartlett-type corrections. Working Paper 95-17, Department of Economics, Southern Illinois University.

[15] Cribari–Neto, F. and S.L.P. Ferrari (1995a), An improved Lagrange multiplier test of heteroskedasticity. *Communications in Statistics—Simulation and Computation* **24**, 31-44.

[16] Cribari–Neto, F. and S.L.P. Ferrari (1995b), Second order asymptotics for score tests in generalized linear models. *Biometrika* **82**, 426-432.

[17] Cribari–Neto, F. aand S. Zarkos (1995), Improved test statistics for multivariate regressions. *Economics Letters* **49**, forthcoming.

[18] Ferrari, S.L.P. and G.M. Cordeiro (1995), Corrected score tests for exponential family nonlinear models. *Statistics and Probability Letters*, forthcoming.

[19] Harris, P. (1985), An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika* **72**, 653-659.

[20] Hayakawa, T. (1977), The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics A* **29**, 359-378.

[21] Johnson, N.L. and S. Kotz (1970a), *Continuous Univariate Distributions*, vol.1 (Houghton Mifflin, Boston).

[22] Johnson, N.L. and S. Kotz (1970b), *Continuous Univariate Distributions*, vol.2 (Houghton Mifflin, Boston).

[23] Johnson, N.L., S. Kotz and A.W. Kemp (1992), *Univariate Discrete Distributions* (Wiley, New York, 2nd ed.).

[24] Jørgensen, B. (1987), Exponential dispersion models. *Journal of the Royal Statistical Society B* **49**, 127-162.

[25] Lawley, D.N. (1956), A generalized method for approximating to the distribution of the likelihood ratio criteria. *Biometrika* **43**, 295-303.

[26] Letac, G. and M. Mora (1990), Natural real exponential families with cubic variance functions. *Annals of Statistics* **18**, 1-37.

[27] Mardia, K.V. (1972), *Statistics of Directional Data* (Academic Press, London).

[28] McCullagh, P. (1989), Some statistical properties of a family of continuous univariate distributions. *Journal of the American Statistical Association* **84**, 125-141.

[29] McCullagh, P. and D.R. Cox (1986), Invariants and likelihood ratio statistics. *Annals of Statistics* **14**, 1419-1430.

[30] Morris, C.N. (1982), Natural exponential families with quadratic variance fucntions. *Annals of Statistics* **10**, 65-80.

[31] Mukerjee, R. (1990a), Comparison of tests in the multiparameter case I: second-order power. *Journal of Multivariate Analysis* **33**, 17-30.

[32] Mukerjee, R. (1990b), Comparison of tests in the multiparameter case II: a third-order optimality property of Rao's test. *Journal of Multivariate Analysis* **33**, 31-48.

[33] Patterson, S.J. (1988), *An Introduction to the Theory of the Riemann Zeta-Function* (Cambridge University Press, New York).

[34] Wolfram, S. (1991), *Mathematica: A System for Doing Mathematics by Computer* (Addison-Wesley, New York, 2nd ed.).