BARTLETT CORRECTIONS FOR ONE–PARAMETER EXPONENTIAL FAMILY MODELS

GAUSS M. CORDEIRO

Deptartamento de Estatística, Universidade Federal de Pernambuco
Cidade Universitária, Recife PE, 50740-540, Brazil

FRANCISCO CRIBARI–NETO

Department of Economics, Southern Illinois University, Carbondale IL, 62901-4515, USA

ELISETE C.Q. AUBIN

Deptartamento de Estatística, Universidade de São Paulo
Caixa Postal 66281, Ag. Cidade de São Paulo, São Paulo SP, 05389-970, Brazil

SILVIA L.P. FERRARI

Deptartamento de Estatística, Universidade de São Paulo
Caixa Postal 66281, Ag. Cidade de São Paulo, São Paulo SP, 05389-970, Brazil

In this paper we derive a general closed-form expression for the Bartlett correction for the test of $\mathrm{H}_0 : \theta = \theta^{(0)}$, where θ is a scalar parameter of a one-parameter exponential family model. Our results are general enough to cover many important and commonly used distributions. Several special cases and classes of variance functions of considerable importance are discussed, and some approximations based on asymptotic expansions are given. We also use a graphical analysis to examine how the correction varies with θ in some special cases. Simulation results are also given.

KEY WORDS: Bartlett correction; chi-squared distribution; exponential family; likelihood ratio statistic.

1. INTRODUCTION

The likelihood ratio (LR) test has been widely used to test hypotheses of interest in statistics and other sciences. As is well known, this test relies on a first order asymptotic chi-squared approximation, *i.e.*, under H_0 and mild regularity conditions, $LR \xrightarrow{d} \chi^2_q$ as $n \to \infty$, where '$\xrightarrow{d}$' denotes convergence in distribution, n is the number of observations and q is the number of restrictions imposed by H_0. Even though this approximation is known to hold for large sample sizes, it may deliver highly inaccurate approximations for samples of small to moderate size. It is thus important to consider corrections that can be applied to improve the finite-sample performance of the LR test.

A correction to the LR test that is commonly used is known as the Bartlett correction and dates back to 1937 (Bartlett, 1937). The idea here is to apply a scalar transformation to the LR statistic to obtain a corrected statistic, say LR^*, which is distributed as χ^2_q when terms of order $O(n^{-2})$ and smaller are neglected. That is, under the null hypothesis, $\Pr[LR \le x] = \Pr[\chi^2_q \le x] + O(n^{-1})$ whereas $\Pr[LR^* \le x] = \Pr[\chi^2_q \le x] + O(n^{-2})$. The corrected statistic is usually written as $LR^* = c^{-1}LR$, where $c = q^{-1}\mathrm{E}(LR)$ and $\mathrm{E}(LR)$

is evaluated up to order n^{-1}. When c depends on unknown parameters, these should be replaced by their restricted maximum likelihood estimates.

The purpose of our paper is to obtain a general closed-form expression for the Bartlett correction for tests involving the parameter of a distribution in the one-parameter exponential family. This family of distributions enjoys wide application and many useful mathematical properties; see, *e.g.*, Bickel and Doksum (1977). More precisely, consider the probability or density function

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

where θ is a scalar parameter and $\zeta(\cdot)$, $\alpha(\cdot)$, $d(\cdot)$ and $v(\cdot)$ are known functions, and it is assumed that the support of $\pi(y;\theta)$ does not depend upon θ. Also, α and ζ are assumed to have continuous first four derivatives with respect to θ, $\zeta(\cdot)$ is positive valued, and $\mathrm{d}\alpha(\theta)/\mathrm{d}\theta$ and $\mathrm{d}\beta(\theta)/\mathrm{d}\theta$ are different from zero for all θ in the parameter space, where $\beta(\theta)$ is defined in the next section as $\beta(\theta) = (\mathrm{d}\zeta(\theta)/\mathrm{d}\theta)(\zeta(\theta)\mathrm{d}\alpha(\theta)/\mathrm{d}\theta)^{-1}$. The family of distributions in (1) has many important and commonly used distributions as special cases (see Section 3). The null hypothesis under test is $\mathrm{H}_0 : \theta = \theta^{(0)}$ against a two-sided alternative, where $\theta^{(0)}$ is a given number. We write the Bartlett correction as

$$c = 1 + \frac{\rho(\theta)}{12n}, \tag{2}$$

and provide a closed-form expression for $\rho(\theta)$ in Section 2. Our goal here is to give a new simple expression for $\rho(\theta)$ which is algebraically more appealing than the general formulas developed by Lawley (1956) and McCullagh and Cox (1986). Our formula is readily applicable and involves only trivial operations on certain functions and their derivatives. The test of homogeneity of parameters of p independent populations is also considered as an application of our result. In Section 3 we show that our result can be easily used to obtain simple expressions in many special cases. The distributions we consider are very useful for modeling data in many applied sciences. Consider for example the logarithmic series distribution introduced by Fisher, Corbet and Williams (1943) when studying the sampling of butterflies. This distribution has many applications in ecology (see Williams, 1944, 1964) and business (see Chatfield, 1986, and Chatfield, Ehrenberg and Goodhardt, 1966). The other distributions we consider have equally important applications. In some cases, $\rho(\theta)$ requires the evaluation of polygamma, zeta or Bessel functions. We give simple approximations that can be used in such situations. We also plot $\rho(\theta)$ against θ for some distributions and thus examine how the correction varies with θ. This is done in Section 4. Section 5 considers Bartlett corrections for some classes of variance functions in natural exponential family models. Finally, Section 6 concludes the paper with some simulation results.

2. DERIVATION OF THE BARTLETT CORRECTION

Let $y_1, \ldots, y_n$ be a set of n independent and identically distributed random variables with density (or probability) function given by $\pi(y;\theta) = \exp\{t(y;\theta)\}$, where θ is a scalar parameter. The likelihood ratio statistic for the test of $\mathrm{H}_0 : \theta = \theta^{(0)}$ against $\mathrm{H}_1 : \theta \neq \theta^{(0)}$ is $LR = 2\sum_{i=1}^{n}\{t(y_i;\hat\theta) - t(y_i;\theta^{(0)})\}$, where $\hat\theta$ is the maximum likelihood estimator of θ. In order to derive an expression for c we need first to introduce some notation. In what follows, $v_r = v_r(\theta) = \mathrm{E}\{t^{(r)}(y;\theta)\}$, $r = 1, 2, 3, 4$, with $t^{(r)}(y;\theta) = \mathrm{d}^r t(y;\theta)/\mathrm{d}\theta^r$. Also, the following notation is used for the cumulants of log-likelihood derivatives (Lawley, 1956; Hayakawa, 1977): $\kappa_{\theta\theta} = \mathrm{E}(\mathrm{d}^2 l/\mathrm{d}\theta^2)$, $\kappa_{\theta\theta\theta} = \mathrm{E}(\mathrm{d}^3 l/\mathrm{d}\theta^3)$, $\kappa_{\theta\theta\theta\theta} = \mathrm{E}(\mathrm{d}^4 l/\mathrm{d}\theta^4)$, $\kappa_{\theta\theta}^{(\theta)} = \mathrm{d}\kappa_{\theta\theta}/\mathrm{d}\theta$, $\kappa_{\theta\theta}^{(\theta\theta)} = \mathrm{d}^2\kappa_{\theta\theta}/\mathrm{d}\theta^2$, $\kappa^{\theta\theta} = 1/\kappa_{\theta\theta}$, where l is the log-likelihood function for the full data.

It follows from Lawley (1956) that, in regular problems, $\mathrm{E}(LR) = 1+\epsilon_1+O(n^{-2})$, with $\epsilon_1 = l_{\theta\theta\theta\theta}-l_{\theta\theta\theta\theta\theta\theta}$, where

$$
\begin{aligned}
l_{\theta\theta\theta\theta} &= (\kappa^{\theta\theta})^2 \left(\frac{\kappa_{\theta\theta\theta\theta}}{4} - \kappa_{\theta\theta\theta}^{(\theta)} + \kappa_{\theta\theta}^{(\theta\theta)}\right), \\
l_{\theta\theta\theta\theta\theta\theta} &= (\kappa^{\theta\theta})^3 \left\{\kappa_{\theta\theta\theta}\left(\frac{\kappa_{\theta\theta\theta}}{6} - \kappa_{\theta\theta}^{(\theta)}\right) + \kappa_{\theta\theta\theta}\left(\frac{\kappa_{\theta\theta\theta}}{4} - \kappa_{\theta\theta}^{(\theta)}\right) + 2(\kappa_{\theta\theta}^{(\theta)})^2\right\};
\end{aligned}
$$

see also Cordeiro (1993).

When we apply Lawley's formula to a general uniparametric distribution, we get, after some algebra, that c is given by (2) with

$$
\rho(\theta) = \frac{3v_2(v_4 - 4v_3' + 4v_2'') - v_3(5v_3 - 24v_2') - 24v_2'^2}{v_2^3}, \tag{3}
$$

where primes denote derivatives with respect to θ.

The next step is to assume an exponential form for $\pi(\cdot;\cdot)$. Let $\pi(y;\theta)$ be as defined in (1) so that $t(y;\theta) = -\log\zeta(\theta) - \alpha(\theta)d(y) + v(y)$. Define $\beta = \beta(\theta) = \zeta'/(\zeta\alpha')$, where $\zeta' = \zeta'(\theta)$, $\alpha' = \alpha'(\theta)$. The maximum likelihood estimate $\hat{\theta}$ comes from $-n^{-1}\sum d(y_i) = \beta(\theta)$, and the likelihood ratio criterion for testing $\mathrm{H}_0 : \theta = \theta^{(0)}$ can be written as $LR = 2n\,\beta(\hat{\theta})\{\alpha(\hat{\theta}) - \alpha(\theta^{(0)})\} + 2n\,\log\{\zeta(\theta^{(0)})/\zeta(\hat{\theta})\}$. We then have that $t'(y;\theta) = -\alpha'\{\beta + d(y)\}$, and since $\mathrm{E}\{t'(y;\theta)\} = 0$ it follows that $\mathrm{E}\{d(y)\} = -\beta$, as expected. It can be shown that

$$
\begin{aligned}
v_2 &= -\alpha'\beta', \\
v_3 &= -2\alpha''\beta' - \alpha'\beta'', \\
v_4 &= -3(\alpha'''\beta' + \alpha''\beta'') - \alpha'\beta'''.
\end{aligned}
$$

Using the results above, it is possible to simplify the expression for $\rho(\theta)$ in (3). After some algebra, we obtain

$$
\rho(\theta) = -\frac{4\beta'^2\alpha''^2 + \alpha'\beta'\alpha''\beta'' - 5\alpha'^2\beta''^2 - 3\alpha'\beta'^2\alpha''' + 3\alpha'^2\beta'\beta'''}{(\alpha'\beta')^3}. \tag{4}
$$

It is noteworthy that the expression for $\rho(\theta)$ given in (4) only requires knowledge about α and β and their first three derivatives. $\rho(\theta)$ in (4) could also be expressed in terms of α and its first three derivatives together with ζ and its first four derivatives, although it is much simpler to leave $\rho(\theta)$ as in equation (4). When α is linear in θ, which corresponds to the natural exponential family, (4) reduces to $\rho(\theta) = (5\beta''^2 - 3\beta'\beta''')/\beta'^3$, which is in agreement with equation (13) in Cordeiro (1983); see Section 5. The other terms in (4) depend on the departure from this simple exponential family form. In particular, $\rho(\theta) = 0$ for distributions in the natural exponential family for which β satisfies the differential equation $5\beta''^2 = 3\beta'\beta'''$. In this case, LR is distributed as χ_1^2 to a second order (and not first order) of approximation. Thus, one can easily obtain the Bartlett correction for a given distribution in (1) by just plugging the corresponding α, β and their derivatives into (4). The derivation of (4) is omitted.

The goal of the analysis above is to obtain a simple formula for $\rho(\theta)$ rather than to explain its general structure. The main difficulty in interpreting (4) is that the individual terms are not invariant under reparameterization and therefore they have no geometric interpretation which is independent of the coordinate system chosen. In Section 5, a very simple interpretation of the Bartlett correction is given via a reparameterization.

Using equation (4) we can easily prove that $\rho(\theta) = 2$ if: (i) $\alpha(\theta)\zeta(\theta) = c_1$, or (ii) $\alpha(\theta)$ is linear, say $\alpha(\theta) = c_1\theta + c_2$, and $\zeta(\theta) = c_3/(\theta c_4^{\theta})$, where c_1, c_2, c_3, c_4 are arbitrary constants. These conditions are (individually) sufficient, but not necessary, to guarantee that $\rho(\theta) = 2$. In Section 3 we will consider several distributions for which conditions (i) and (ii) hold.

We now give an important application of equation (4) to the LR test of homogeneity of parameters of p independent populations taken from (1). Let y_{ij}, $i = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, n_i$, be independent random variables with probability or density function (1) with parameters θ_i, for $i = 1, 2, \ldots, p$. The likelihood ratio statistic for testing the null hypothesis $\mathrm{H}_0 : \theta_1 = \theta_2 = \cdots = \theta_p (= \theta)$, against the alternative hypothesis H_1 : the θ's are not all the same, is $LR = 2\sum_{i=1}^{p}\sum_{j=1}^{n_i}\{t(y_{ij};\hat{\theta}_i) - t(y_{ij};\hat{\theta})\}$, where $\hat{\theta}_i$ is the maximum likelihood estimator of θ_i and $\hat{\theta}$ is the maximum likelihood estimator of θ under H_0. We then have, under H_0 and ignoring terms of order n_i^{-2},

$$\begin{aligned}\mathrm{E}(LR) &= \mathrm{E}\left[2\sum_{i=1}^{p}\sum_{j=1}^{n_i}\{t(y_{ij};\hat{\theta}_i) - t(y_{ij};\theta)\}\right] - \mathrm{E}\left[2\sum_{i=1}^{p}\sum_{j=1}^{n_i}\{t(y_{ij};\hat{\theta}) - t(y_{ij};\theta)\}\right] \\ &= p - 1 + \frac{\rho(\theta)}{12}\left[\sum_{i=1}^{p}\frac{1}{n_i} - \frac{1}{n_+}\right],\end{aligned}$$

where ρ is given in (4) and $n_+ = \sum_{j=1}^{p} n_i$. Therefore, the Bartlett correction is given by

$$c = 1 + \frac{\rho(\theta)}{12(p-1)}\left[\sum_{i=1}^{p}\frac{1}{n_i} - \frac{1}{n_+}\right].$$

The expression above is a generalization of the well known Bartlett correction for the test of homogeneity of variances from p normal populations with known means. It follows from case (xx, a) in Section 3 that $\rho(\theta) = 4$. The application presented above can then be viewed as a generalization of this well known result (Bartlett, 1937).

3. SOME SPECIAL CASES

In this section we show that the expression for $\rho(\theta)$ in (4) can be used to obtain simple closed-form formulas for the Bartlett correction to the likelihood ratio statistic for many important distributions. Here, we consider 24 special cases and give closed-form expressions for $\rho(\theta)$ obtained using Mathematica (Wolfram, 1991). These cases cover more than 24 distributions since some of them are indeed families of distributions. For example, the Burr system of distributions considered here covers 10 distributions (all distributions in the Burr system with the exception of the Burr I and Burr IX). Several other distributions in (1) could also be analyzed. Most of the distributions considered here are well known and have a wide range of practical applications in fields such as engineering, biology, zoology, economics and medicine, among others. Cases (i) through (viii) involve discrete random variables whereas continuous random variables are considered in cases (ix) through (xxiv). For further details on the distributions considered here, see Johnson and Kotz (1970a, 1970b) and Johnson, Kotz and Kemp (1992).

It should be pointed out that although Bartlett corrections usually lead to an improvement in the rate of convergence to a chi-squared distribution in continuous models, there is no guarantee that this also holds for discrete models; see Frydenberg and Jensen (1989). However, the simulation results in Section 6 show that even in the discrete case the Bartlett correction leads to improvements in the finite-sample performance of the likelihood ratio test.

The following particular cases are considered:

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

$$\rho(\theta) = \frac{2(1-\theta+\theta^2)}{m\theta(1-\theta)}.$$

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

$$\rho(\theta) = \frac{2(1-\theta+\theta^2)}{\gamma\theta}.$$

(iii) Poisson ($\theta > 0$, $y = 0, 1, 2, \ldots$): $\alpha(\theta) = -\log(\theta)$, $\zeta(\theta) = \exp\{\theta\}$, $d(y) = y$, $v(y) = -\log(y!)$; $\rho(\theta) = 2\theta^{-1}$.

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

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

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

$$\begin{aligned}\rho(\theta) = &[\theta\{\theta + \log(1-\theta)\}^3]^{-1}[-2\theta^4 - 6\theta^3\log(1-\theta) - 8(\theta^2+\theta^3)\{\log(1-\theta)\}^2 \\ &- 3(2\theta + 2\theta^2 - \theta^3)\{\log(1-\theta)\}^3 - 2(1-\theta+\theta^2)\{\log(1-\theta)\}^4].\end{aligned}$$

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

$$\rho(\theta) = \frac{2g^2 + 6\theta gg' + 8\theta^2(3g'^2 - gg'') + 3\theta^3(4g'g'' - gg''') + \theta^4(5g''^2 - 3g'g''')}{\theta(g+\theta g')^3},$$

$g = g(\theta) = \mathrm{d}\log\zeta(\theta)/\mathrm{d}\theta$.

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

$$\rho(\theta) = \frac{5g''^2 - 3g'g'''}{g'^3},$$

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

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

$$\rho(\theta) = \frac{-2D_1^6 + 6D_0D_1^4D_2 - 8D_0^2D_1^3D_3 + 18D_0^3D_1D_2D_3 - 3D_0^3D_1^2D_4 - 9D_0^3D_2^3 + 3D_0^4D_2D_4 - 5D_0^4D_3^2}{(D_1^2 - D_0D_2)^3},$$

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

(ix) Maxwell ($\theta > 0$, $y > 0$): $\alpha(\theta) = (2\theta^2)^{-1}$, $\zeta(\theta) = \theta^3$, $d(y) = y^2$, $v(y) = \log(y^2\sqrt{2/\pi})$; $\rho(\theta) = 4/3$.

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

(a) k known: $\alpha(\theta) = \theta$, $\zeta(\theta) = \theta^{-k}$, $d(y) = y$, $v(y) = (k-1)\log(y) - \log\{\Gamma(k)\}$; $\rho(\theta) = 2k^{-1}$;

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

$$\rho(k) = \frac{5\psi''(k)^2 - 3\psi'(k)\psi'''(k)}{\psi'(k)^3},$$

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

(xi) Burr system of distributions ($\theta > 0$, $b > 0$, b known, $y > 0$): $\alpha(\theta) = \theta$, $\zeta(\theta) = c(\theta)/\theta$, $d(y) = -\log G(y)$, $v(y) = \log\{|\mathrm{d}\log G(y)/\mathrm{d}y|\}$; $\rho(\theta) = 2$, where the functions $c(\cdot)$ and $G(\cdot)$ are positive real-valued. Different choices for $c(\theta)$ and $G(y)$ lead to different distributions; see Burr (1942).

(xii) Rayleigh ($\theta > 0$, $y > 0$): $\alpha(\theta) = \theta^{-2}$, $\zeta(\theta) = \theta^2$, $d(y) = y^2$, $v(y) = \log(2y)$; $\rho(\theta) = 2$.

(xiii) Pareto ($\theta > 0$, $k > 0$, k known, $y > k$): $\alpha(\theta) = \theta + 1$, $\zeta(\theta) = (\theta k^{\theta})^{-1}$, $d(y) = \log(y)$, $v(y) = 0$; $\rho(\theta) = 2$.

(xiv) Weibull ($\theta > 0$, $\phi > 0$, ϕ known, $y > 0$): $\alpha(\theta) = \theta^{-\phi}$, $\zeta(\theta) = \theta^{\phi}$, $d(y) = y^{\phi}$, $v(y) = \log(\phi) + (\phi-1)\log(y)$; $\rho(\theta) = 2$.

(xv) Power ($\theta > 0$, $\phi > 0$, ϕ known, $0 < y < \phi$): $\alpha(\theta) = 1 - \theta$, $\zeta(\theta) = \theta^{-1}\phi^{\theta}$, $d(y) = \log(y)$, $v(y) = 0$; $\rho(\theta) = 2$.

(xvi) Laplace ($\theta > 0$, $-\infty < k < \infty$, k known, $y > 0$): $\alpha(\theta) = \theta^{-1}$, $\zeta(\theta) = 2\theta$, $d(y) = |y - k|$, $v(y) = 0$; $\rho(\theta) = 2$.

(xvii) Extreme value ($-\infty < \theta < \infty$, $\phi > 0$, ϕ known, $-\infty < y < \infty$): $\alpha(\theta) = \exp\{\theta/\phi\}$, $\zeta(\theta) = \phi\exp\{-\theta/\phi\}$, $d(y) = \exp\{-y/\phi\}$, $v(y) = -y/\phi$; $\rho(\theta) = 2$.

(xviii) Truncated extreme value ($\theta > 0$, $y > 0$): $\alpha(\theta) = \theta^{-1}$, $\zeta(\theta) = \theta$, $d(y) = \exp\{y\} - 1$, $v(y) = y$; $\rho(\theta) = 2$.

(xix) Lognormal ($\theta > 0$, $\mu > 0$, μ known, $y > 0$): $\alpha(\theta) = \theta^{-2}$, $\zeta(\theta) = \theta$, $d(y) = \{\log(y) - \mu\}^2/2$, $v(y) = -\log(y) + \{\log(2\pi)\}/2$; $\rho(\theta) = 4$.

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

(a) μ known: $\alpha(\theta) = (2\theta)^{-1}$, $\zeta(\theta) = \theta^{1/2}$, $d(y) = (y-\mu)^2$, $v(y) = -\{\log(2\pi)\}/2$; $\rho(\theta) = 4$.

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

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

(a) μ known: $\alpha(\theta) = \theta$, $\zeta(\theta) = \theta^{-1/2}$, $d(y) = (y-\mu)^2/(2\mu^2 y)$, $v(y) = -\{\log(2\pi y^3)\}/2$; $\rho(\theta) = 4$.

(b) θ known: $\alpha(\mu) = \theta(2\mu^2)^{-1}$, $\zeta(\mu) = \exp\{-\theta/\mu\}$, $d(y) = y$, $v(y) = -\theta(2y)^{-1} + [\log\{\theta/(2\pi y^3)\}]/2$; $\rho(\mu) = 0$.

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

$$\rho(\theta) = \frac{-5\{\psi''(\theta+1) - \psi''(\theta+1/2)\}^2 + 3\{\psi'(\theta+1) - \psi'(\theta+1/2)\}\{\psi'''(\theta+1) - \psi'''(\theta+1/2)\}}{\{\psi'(\theta+1) - \psi'(\theta+1/2)\}^3},$$

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

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

$$\rho(\theta) = \frac{5r''(\theta)^2 - 3r'(\theta)r'''(\theta)}{r'(\theta)^3},$$

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

$$\rho(\theta) = \frac{2\{1 - 4(\cos\theta)^2\}}{r}.$$

The Bartlett correction for some other distributions can be obtained as special cases of the distributions listed above. For the Bernoulli distribution, one gets $\rho(\theta) = 2(1 - \theta + \theta^2)/\{\theta(1-\theta)\}$, which comes from the binomial distribution with $m = 1$. Similarly, the Bartllett correction for the geometric distribution is the same as the correction for the negative binomial distribution with $\gamma = 1$. The correction factor for the exponential distribution is obtained as a special case of the gamma distribution with $k = 1$, that is, $\rho(\theta) = 2$. The Bartlett correction for the Erlang distribution is obtained from the gamma distribution with $\theta = 1$ and k being an integer number. Also, the Bartlett correction for the log-gamma distribution is the same as for the gamma distribution. Finally, for the chi-squared distribution $\rho(\theta)$ can be obtained from the gamma distribution with $\theta = 1/2$ and $k = \theta/2$. It should also be mentioned that the following special cases were previously considered by Cordeiro (1983): binomial, Poisson, gamma (case a), inverse Gaussian (case b) and normal (case b).

It is interesting to note that for some distributions the correction does not depend on the value of the parameter specified under the null hypothesis, but this is not always the case. In some cases, the Bartlett correction and thus the first order χ^2 approximation to the LR test is affected by the value of θ specified in H_0; see Section 4 for more details. It is also noteworthy that our general expression for $\rho(\theta)$ in (4) is able to generate both simple formulas (for example, $\rho(\theta) = 2$) and complex expressions (for example, $\rho(\theta)$ for the truncated Poisson, logarithmic series, non-central hypergeometric and power series distributions) for different special cases.

In some cases the expression for $\rho(\theta)$ is quite complex, and it is thus important to obtain asymptotic expansions that yield simple approximations for $\rho(\theta)$ and do not require the evaluation of functions such as polygamma, Bessel or zeta functions. We start with the gamma distribution (θ known). For large values of k we have that

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

It is possible to use this result to simplify the expression for $\rho(k)$ for the gamma distribution. After some algebra, we obtain, for large k,

$$\rho(k) = -\frac{1}{k} - \frac{1}{2k^2} + \frac{1}{2k^3} + O(k^{-4}).$$

This expression is very convenient since it does not require the evaluation of polygamma functions. This expansion is also valid for the Erlang distribution and for the log-gamma distribution (with μ known). It should also be noted that a similar result for the chi-squared distribution can be obtained by just replacing k by $\theta/2$ in the expression above, *i.e.*, for the chi-squared distribution with θ degrees of freedom

$$\rho(\theta) = -\frac{2}{\theta} - \frac{2}{\theta^2} + \frac{4}{\theta^3} + O(\theta^{-4}),$$

which yields an approximation for large values of θ. By making similar developments and using the formula

$$\psi'(\theta+1) - \psi'(\theta+1/2) = 2\psi'(\theta) - 4\psi'(2\theta) - \frac{1}{\theta^2},$$

we obtain the following asymptotic expansion for the McCullagh distribution, for large θ:

$$\rho(\theta) = 4 + \frac{3}{2\theta^2} - \frac{3}{4\theta^3} + O(\theta^{-4}).$$

This should yield a good approximation for $\rho(\theta)$ when θ is large. Notice, for example, that for this expression $\lim_{\theta\to\infty} \rho(\theta) = 4$, which is in agreement with the graph displayed in Figure 4. Consider now the von Mises distribution where $\rho(\theta)$ is a function of Bessel functions of the first kind. For large values of θ, we can write $r(\theta)$ as (Abramowitz and Stegun, 1970, pp.416-421)

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

Using this result we get the following approximation for large values of θ

$$\rho(\theta) = 4 - \frac{9}{2\theta^2} - \frac{57}{2\theta^3} + O(\theta^{-4}).$$

For small θ we can use the fact that (Mardia, 1972, p.63)

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

to obtain the following approximation for $\rho(\theta)$

$$\rho(\theta) = \frac{9}{2} + \frac{3\theta^2}{2} + O(\theta^4).$$

These approximations, unlike the expression for $\rho(\theta)$ given in (xxiii), do not involve Bessel functions and can be easily evaluated. Next, we turn to the logarithmic series distribution. For small values of θ we have that

$$\log(1-\theta) = -\theta - \frac{\theta^2}{2} - \frac{\theta^3}{3} - \frac{\theta^4}{4} - \frac{\theta^5}{5} - \frac{\theta^6}{6} + O(\theta^7).$$

Using this result, we get, after some algebra, that for the logarithmic series distribution with small θ

$$\rho(\theta) = \frac{4}{\theta} - \frac{10}{3} + \frac{7\theta}{9} + \frac{74\theta^2}{135} + \frac{67\theta^3}{162} + O(\theta^4).$$

Consider next the truncated Poisson distribution. Here we make use of the expansion $e^\theta = 1 + \theta + \theta^2/2 + \theta^3/6 + \theta^4/24 + \cdots$ to obtain the following expansion for small θ:

$$\rho(\theta) = \frac{4}{\theta} - \frac{2}{3} + \frac{10\theta}{9} - \frac{89\theta^2}{135} + O(\theta^3).$$

Finally, consider the zeta distribution and let

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

$j = 0, 1, 2$, γ_0 being Euler's constant, *i.e.*, $\gamma_0 \approx 0.577$. After some algebra, we obtain that, for small values of θ,

$$\rho(\theta) = 2 + 24\theta^2(\gamma_0^2 + 2\gamma_1) - 44\theta^3(2\gamma_0^3 + 6\gamma_0\gamma_1 + 3\gamma_2).$$

It is possible to further simplify the expansion above. Using Maple V (Abell and Braselton, 1994) we obtain the following result:

$$\rho(\theta) = 2 + \frac{42594}{9463}\theta^2 - \frac{137395}{30206}\theta^3 + O(\theta^4).$$

Hence, $\rho(\theta) \approx 2$ when $\theta \approx 0$.

4. GRAPHICAL ANALYSIS

It is clear from the results in the previous section that for several distributions ρ is not a constant, but a function of θ, thus varying with this parameter. An interesting question is then: How do different values of θ specified in the null hypothesis affect the Bartlett correction and hence the first order χ_1^2 approximation? In order to shed some light on this issue, we plot $\rho(\theta)$ against θ for some distributions considered in the previous section. This shows how the correction changes with different values of θ. Here, we present plots of $\rho(\theta)$ versus θ for the following distributions: truncated Poisson, generalized hyperbolic secant (with $r = 1$), logarithmic series, McCullagh, Bernoulli, gamma (θ known), zeta and von Mises. These plots are given in Figures 1 through 8, respectively. The solid lines correspond to $\rho(\theta)$ and the dashed lines to the approximations given in the previous section, when available.

It is clear from Figure 1 that the correction becomes very large for small values of θ and very small for large values of θ for the truncated Poisson distribution. The approximation given in the previous section works well for $\theta < 1$. For the generalized hyperbolic secant distribution (Figure 2), we observe that the correction displays an oscillatory behavior, ranging from -6 to 2. In the case of the logarithmic series distribution (Figure 3), the correction becomes large for values of θ close to 0 or 1, and our approximation works well for values of θ up to 0.7. The plot of $\rho(\theta)$ versus θ for the McCullagh distribution (Figure 4) shows

Figure 1: Truncated Poisson

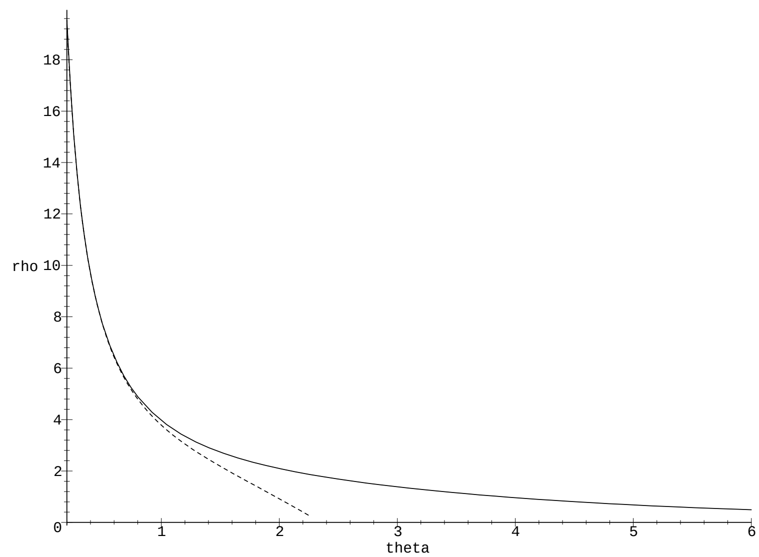


Figure 2: Generalized Hyperbolic Secant

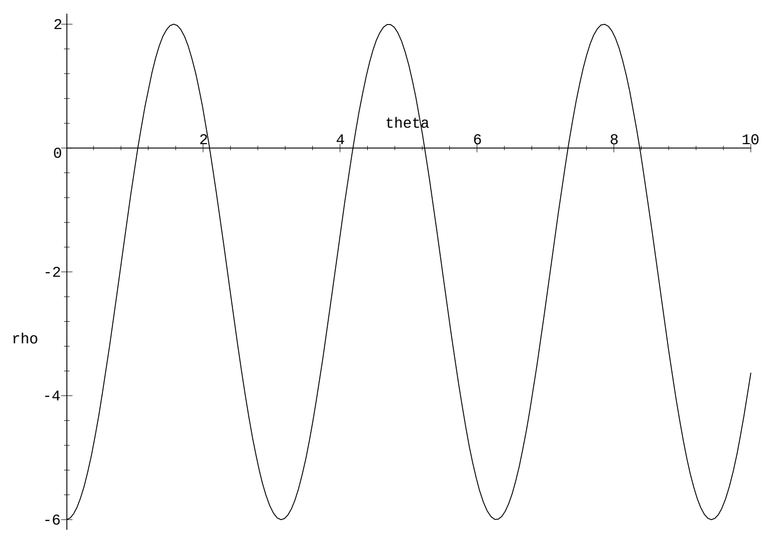


Figure 3: Logarithmic Series

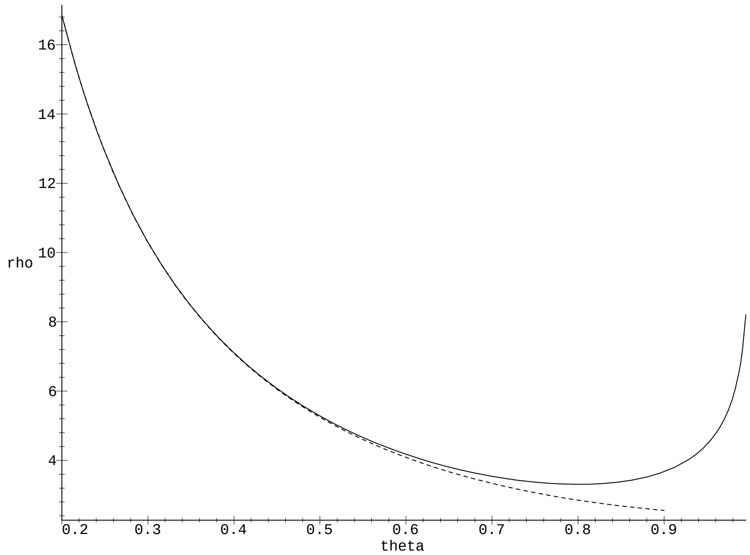


Figure 4: McCullagh

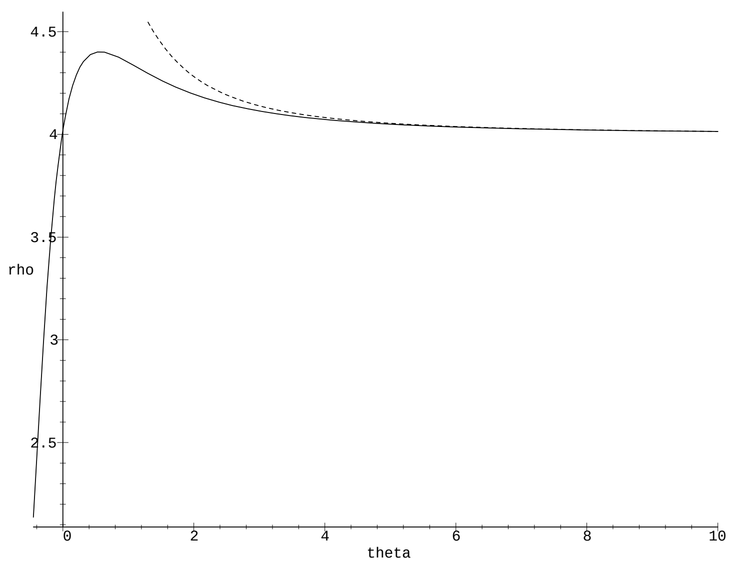


Figure 5: Bernoulli

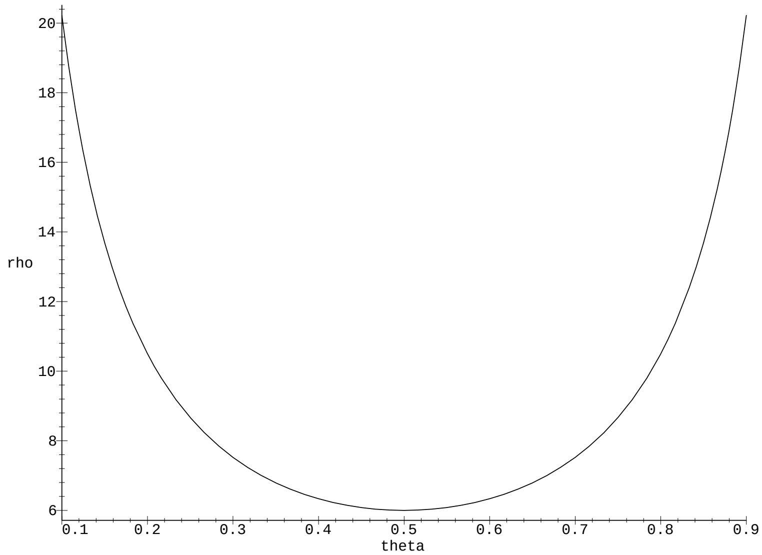


Figure 6: Gamma

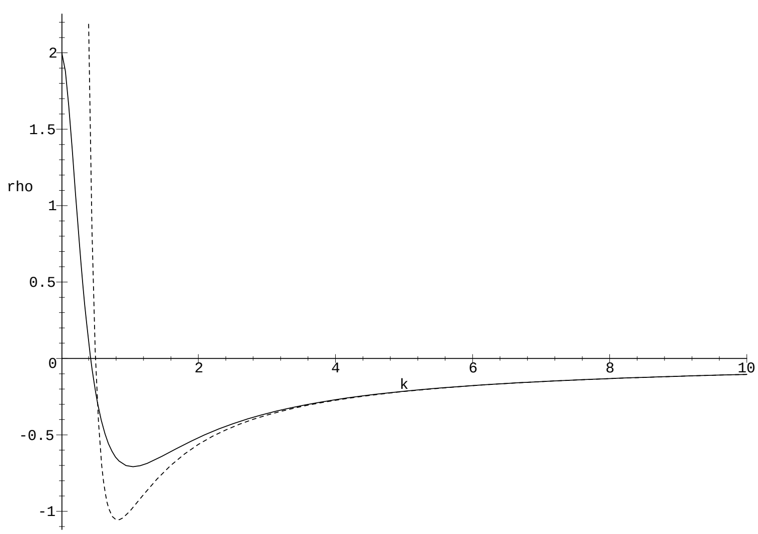


Figure 7: Zeta

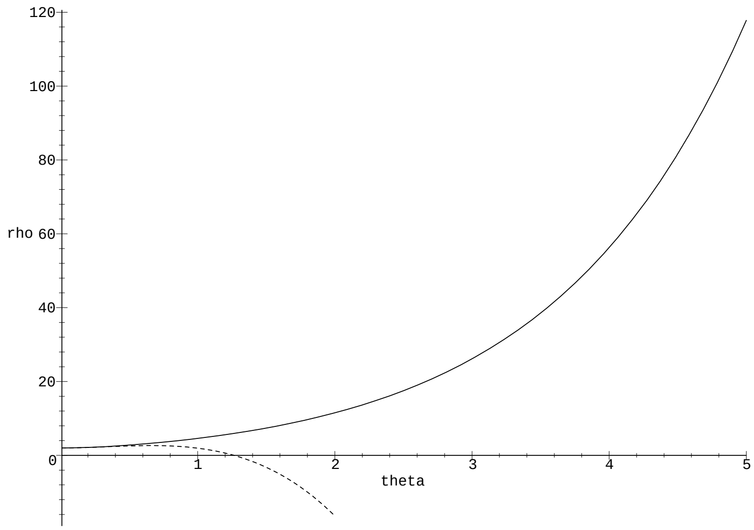
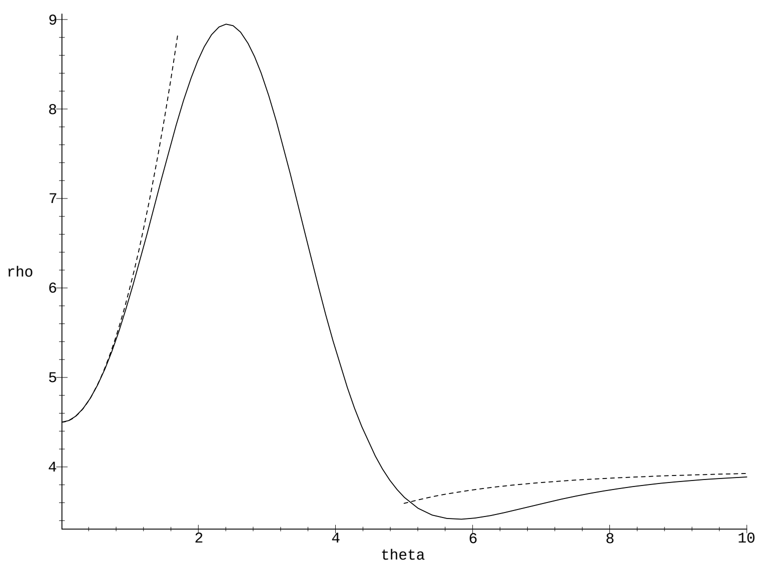


Figure 8: Von Mises



a rather interesting behavior. $\rho(\theta)$ increases quite fast for small θ, and after reaching a peak at approximately 4.4 decreases continuously approaching an asymptotic level as θ increases. Here, $\lim_{\theta\to\infty}\rho(\theta) = 4$. Also, our approximation seems to work well for $\theta > 2$. For the Bernoulli distribution (Figure 5) the correction is U-shaped, reaching a minimum at $\theta = 1/2$. Next, consider the gamma distribution with θ known (Figure 6). The correction is close to 2 when k is close to zero. It then drops rapidly and becomes negative, and then approaches zero from below as θ increases. Again, our approximation delivers accurate results for $k > 2$. The graph of $\rho(\theta)$ vs. θ for the zeta distribution is given in Figure 7. It is possible to see that $\rho(\theta)$ is approximately equal to 2 for small values of θ, and that $\rho(\theta)$ diverges to ∞ when $\theta \to \infty$. It is also clear that the approximation given in previous section only works well when $\theta < 1$. Finally, consider the von Mises distribution (Figure 8). For small values of θ, $\rho(\theta) \approx 4.5$. Indeed, $\lim_{\theta\to 0}\rho(\theta) = 4.5$. It is interesting to note that $\rho(\theta)$ increases substantially for θ between 1 and 3. The approximation for small θ works well for, say, $\theta < 1.75$, and the approximation for large θ delivers accurate results for $\theta > 8$.

5. NATURAL EXPONENTIAL FAMILY

In this section we study the natural exponential family which is indexed by the natural parameter α in the form

$$\pi(y;\alpha) = \frac{1}{\delta(\alpha)}\exp\{-\alpha d(y) + v(y)\}, \tag{5}$$

where $-d(y)$ is the canonical statistic. Similar conditions to the ones stated for the probability or density function in (1) are assumed to hold. Here $\delta(\alpha)$ is the cumulant generating function of $-d(y)$, *i.e.*, the rth cumulant of $-d(y)$ is $\kappa_r = \mathrm{d}^r \log\delta(\alpha)/\mathrm{d}\alpha^r$, for $r = 1, 2, \ldots$, where $\beta(\alpha) = \delta'(\alpha)/\delta(\alpha)$. Also, $\kappa_{r+1} = \mathrm{d}^r\beta/\mathrm{d}\alpha^r$, for $r = 1, 2, \ldots$ In particular, the mean and variance of $-d(y)$ are $\kappa_1 = \beta$ and $\kappa_2 = \mathrm{d}\beta/\mathrm{d}\alpha = \beta'$. From now on primes will denote derivatives with respect to α. The variance function β' does not depend on the particular parameterization used, and it is an important tool for handling Bartlett corrections in exponential family models for two reasons. First, it defines the distribution in (5). Second, it is much simpler to express the variance function than the generator function $\delta(\alpha)$. The first fact was first suggested by Jørgensen (1987) who showed that the variance function uniquely defines, up to a linear transformation, a distribution in (5).

The likelihood ratio criterion for testing $\mathrm{H}_0 : \theta = \theta^{(0)}$ is invariant to reparameterizations, and since (5) defines a one-to-one correspondence between α and θ, equation (4) reduces to

$$\rho(\alpha) = 5\gamma_1^2 - 3\gamma_2, \tag{6}$$

where $\gamma_1^2 = \beta''^2/\beta'^3$ and $\gamma_2 = \beta'''/\beta'^2$ are the third and fourth standardized cumulants of $-d(y)$. Equation (6) was first given by Cordeiro (1983) and McCullagh and Cox (1986). This formula has a more elegant form than (4) although the latter is more convenient since it gives the Bartlett correction in terms of the original parameter θ. Figure 9 plots $\rho(\alpha)$ against γ_1 and γ_2. Note that the range of values for $\rho(\alpha)$ is limited by $\gamma_1^2 + \gamma_2 + 2 \geq 0$ and $\gamma_2 - \gamma_1^2 \geq 1$ which imply that $-3\gamma_2 \leq \rho(\alpha) \leq 2\gamma_1^2 - 3$. It is clear from this figure that $\rho(\alpha)$ decreases as γ_2 increases, thus assuming large negative values for large γ_2 when γ_1 is small. Also, $\rho(\alpha)$ displays a quadratic behavior as a function of γ_1 for given γ_2.

The Bartlett correction (6) reduces to a constant $k \neq 0$ when the generator function is given by $\delta(\alpha) = c_1 \exp\{c_2\alpha\}(k\alpha + c_3)^{-2/k}$, where the c_i's are arbitrary constants. The $\delta(\alpha)$ functions of all special cases in Section 3 having non-zero constant corrections verify this condition. On the other hand, the Bartlett correction vanishes when $\delta(\alpha) = \exp\{c_1 + c_2\alpha + c_3\alpha^2\}$ or $\delta(\alpha) = \exp\{c_1 + c_2\alpha^{1/2} + c_3\alpha\}$, where again the c_i's are arbitrary constants with $c_3 \neq 0$. From the 24 special cases considered in Section 3, only two yield $\rho(\alpha) = 0$: the normal case with known mean for which the first equation holds, and the inverse Gaussian case with known index whose $\delta(\alpha)$ satisfies the latter equation.

The aim of the section is to give Bartlett corrections in closed-form for special families of variance functions. We begin with the power variance function defined by $\beta' = c^{-1}\beta^p$ where the domain of β is $\mathbb{R}$ for $p = 0$ and $\mathbb{R}_+$ otherwise, and $c > 0$ is a constant. The Bartlett correction for this variance function reduces to

$$\rho(\beta) = \frac{p(3-p)\beta^{p-2}}{c}. \tag{7}$$

Equation (7) covers many important cases including the classical distributions: normal ($p = 0$), Poisson ($p = 1$), gamma ($p = 2$), and inverse Gaussian ($p = 3$). Other values of p define a number of distributions which have been classified by Jørgensen (1987). Power variance functions with $p \leq 0$ correspond to distributions generated by extreme stable distributions whose support is $\mathbb{R}$. For $1 < p < 2$ and for $p > 2$, (5) corresponds to certain compound Poisson distributions and to distributions generated by positive stable distributions, respectively. In both cases, the distributions are continuous and have support on $\mathbb{R}_+$. The remaining cases with power $0 < p < 1$ do not correspond to distributions in (5). The plot of $\rho(\beta)$ against β and p is given in Figure 10 (for $c = 1$). The correction only vanishes for $p = 0, 3$. It is also clear that $\rho(\beta)$ assumes large negative values when β is small and p is negative.

Consider now the family of variance functions defined as polynomials of order less than or equal to 3, say $\beta' = c_0 + c_1\beta + c_2\beta^2 + c_3\beta^3$. For the case where $c_3 = 0$, six types of natural exponential families were defined by Morris (1982). Letac and Mora (1990) extended Morris' classification and showed that there are only six types of distributions in (5) whose variance functions are polynomials of degree 3 in β. The general formula for $\rho(\beta)$ for testing $\mathrm{H}_0 : \theta = \theta^{(0)}$ in the case of cubic variance functions is

$$\rho(\beta) = 2c_2 + \frac{2}{\beta'}\{c_1 - 4c_o c_2 - c_3\beta(9c_0 + 3c_1\beta + c_2\beta^2)\}. \tag{8}$$

The corrections for all 12 cases discussed in Morris (1982) and Letac and Mora (1990) can be obtained from equation (8); see Table 1. For quadratic variance functions, the correction is obtained setting $c_3 = 0$. The corrections in Table 1 for the normal, Poisson, binomial, gamma and generalized hyperbolic secant distributions agree with the results in Section 3. Two of the six distributions in Table 1 with cubic variance function are continuous: the inverse Gaussian and Ressel distributions. The remaining four are concentrated on the set $\mathbb{N}$ of nonnegative integers: Abel, Takács and the two arcsine (strict and extended) distributions.

We have also attempted to obtain Bartlett corrections for wider classes of variance functions such as, for example, the variance function $\beta' = P + Q\sqrt{R}$, where P, Q and R are polynomials in β of degrees less than or equal to 3, 2 and 2, respectively. This variance function includes the Babel class of variance functions for which $\beta' = b\Delta + (a\beta + c)\sqrt{\Delta}$, where Δ is a polynomial of degree ≤ 2 in β which is not a perfect square

Figure 9: Bartlett Surface

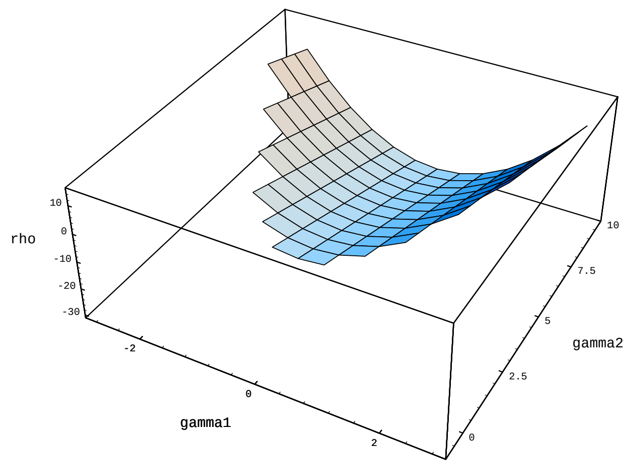


Figure 10: Bartlett Surface for Power Variance Function

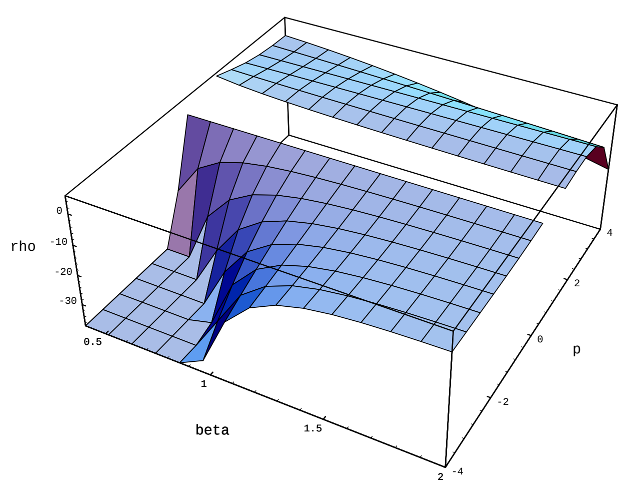


Table 1 Bartlett corrections for cubic variance functions

Distribution	c_0	c_1	c_2	c_3	Range of β	$\rho(\beta)$
normal (xxi, b)	θ	0	0	0	$\mathbb{R}$	0
Poisson (iii)	0	1	0	0	$\mathbb{R}_+$	$\frac{2}{\beta}$
binomial (i)	0	1	$-\frac{1}{m}$	0	$(0,\, m)$	$\frac{2(m^2-m\beta+\beta^2)}{m\beta(m-\beta)}$
negative binomial (ii)	0	1	$\frac{1}{\gamma}$	0	$\mathbb{R}_+$	$\frac{\beta^2+\beta\gamma+\gamma^2}{\beta\gamma(\beta+\gamma)}$
gamma (x, a)	0	0	$\frac{1}{k}$	0	$\mathbb{R}_+$	$\frac{2}{k}$
generalized hyperbolic secant (xxiv)	r	0	$\frac{1}{r}$	0	$\mathbb{R}$	$\frac{2(\beta^2-3r^2)}{r(\beta^2+r^2)}$
Abel (p known)	0	1	$\frac{2}{p}$	$\frac{1}{p^2}$	$\mathbb{R}_+$	$\frac{2}{\beta}$
Takács (p, $a>0$ known)	0	1	$\frac{2a+1}{ap}$	$\frac{a+1}{ap^2}$	$\mathbb{R}_+$	$(*)$
strict arcsine (p known)	0	1	0	$\frac{1}{p^2}$	$\mathbb{R}_+$	$-\frac{2(3\beta^2-p^2)}{\beta(\beta^2+p^2)}$
large arcsine (p, $a>0$ known)	0	1	$\frac{2}{ap}$	$\frac{1+a^2}{a^2p^2}$	$\mathbb{R}_+$	$(**)$
Ressel ($p>0$)	0	0	$\frac{1}{p}$	$\frac{1}{p^2}$	$\mathbb{R}_+$	$\frac{2}{\beta+p}$
inverse Gaussian (xxi, b)	0	0	0	$-\frac{1}{\theta^2}$	$\mathbb{R}_+$	0

Note:

$$(*) = \frac{2\{a^2p^2 + ap(1+2a)\beta + (1+a+a^2)\beta^2\}}{a\beta(p+\beta)(ap+\beta+a\beta)}$$

$$(**) = \frac{2\{a^2p^2 + 2ap\beta + (1-3a^2)\beta^2\}}{\beta\{a^2p^2 + 2ap\beta + (1+a^2)\beta^2\}}$$

and a, b and c are real numbers. The closed-form expressions for the Bartlett corrections for these variance functions were obtained using Mathematica but they are too cumbersome to be reported here.

6. SIMULATION RESULTS

We present simulation results from two experiments, both based on 10,000 replications. The first simulation involves the logarithmic series distribution and the following values for θ: 0.1, 0.2, 0.4, 0.6 and 0.8. We used the following approximation for the maximum likelihood estimate of θ (Birch, 1963):

$$\hat{\theta} \approx 1 - \left[1 + \left\{\left(\frac{5}{3} - \frac{1}{16}\log(\bar{y})\right)(\bar{y}-1) + 2\right\}\log(\bar{y})\right]^{-1}.$$

The second experiment focuses on the binomial distribution, where we have that $y_1, \ldots, y_n$ are independently distributed as $\mathrm{B}(m, \theta)$. We set $m = 5$, and take the following values for θ: 0.1, 0.2, 0.3, 0.4 and 0.5. The results for the nominal levels of $\nu = 0.10, 0.05$ are given in Tables 2 and 3 (entries are percentages).

It is clear that the Bartlett-corrected statistic delivers small-sample improvements, in the sense that the estimated sizes of the Bartlett-corrected tests are closer to the nominal levels than the sizes of their uncorrected counterparts. Moreover, it is possible to see that when the correction is larger, the first order

Table 2 Estimated rejection probabilities—logarithmic series

		$\theta = 0.1$		$\theta = 0.2$		$\theta = 0.4$		$\theta = 0.6$		$\theta = 0.8$	
n	ν	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
5	10	19.9	14.7	14.5	11.7	12.0	10.1	11.1	10.2	11.0	10.3
	5	11.8	9.0	8.3	6.4	6.4	5.4	5.8	5.2	5.6	5.1
10	10	15.2	11.9	12.4	11.3	11.1	10.2	10.9	10.1	10.5	10.1
	5	8.6	7.2	6.5	5.7	5.5	5.2	5.4	5.2	5.3	5.1
20	10	12.4	11.0	11.3	10.4	10.6	10.1	10.3	10.0	10.3	10.0
	5	6.8	6.0	5.8	5.3	5.3	5.1	5.2	5.1	5.1	5.0

Note: (1) denotes $\Pr[LR \geq c]$ and (2) denotes $\Pr[LR^* \geq c]$.

Table 3 Estimated rejection probabilities—binomial

		$\theta = 0.1$		$\theta = 0.2$		$\theta = 0.3$		$\theta = 0.4$		$\theta = 0.5$	
n	ν	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
1	10	13.9	12.2	13.0	11.0	12.1	10.4	11.8	10.1	11.7	10.1
	5	8.8	6.8	7.0	5.3	6.5	5.2	6.3	5.2	6.1	5.2
2	10	12.6	11.4	11.5	10.2	11.1	10.2	10.9	10.1	10.4	9.9
	5	6.9	5.9	6.5	5.1	5.9	5.1	5.6	5.1	5.5	5.0
4	10	11.3	10.5	10.7	9.9	10.4	9.8	10.4	9.9	10.2	9.9
	5	6.6	5.7	6.1	5.1	5.9	5.1	5.3	5.0	5.1	5.0

Note: (1) denotes $\Pr[LR \geq c]$ and (2) denotes $\Pr[LR^* \geq c]$.

approximation to the likelihood ratio statistic is worse. This is clear when one contrasts the results in Table 2 (logarithmic series distribution) with the plot of $\rho(\theta)$ against θ given in Figure 3. That is, the figures in Table 2 show that the likelihood ratio test tends to be more oversized when θ is small, and Figure 3 reveals that $\rho(\theta)$ increases when θ approaches 0. Similar conclusions for the binomial distribution are obtained by comparing the figures in Table 3 with the plot of $\rho(\theta)$ against θ in Figure 5.

ACKNOWLEDGMENTS

The authors wish to thank Miguel Angel Uribe–Opazo and an anonymous referee for constructive comments and suggestions. The partial financial support of CNPq/Brazil is also gratefully acknowledged.

REFERENCES

Abell, M.L. and Braselton, J.P. (1994). *The Maple V Handbook*. AP Professional, New York.

Abramowitz, M. and Stegun, I.R. (1970). *Handbook of Mathematical Functions*. National Bureau of Standards, Washington, D.C.

Bartlett, M.S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society A*, **160**, 268-282.

Bickel, P.J. and Doksum, K.A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day, Oakland.

Birch, M.W. (1963). An algorithm for the logarithmic series distribution. *Biometrics*, **19**, 651-652.

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

Chatfield, C., Ehrenberg, A.S.C. and Goodhardt, G.J. (1966). Progress on a simplified model of stationary purchasing behavior. *Journal of the Royal Statistical Society A*, **129**, 317-367.

Chatfield, C. (1986). Discrete distributions and purchasing models. *Communications in Statistics–Theory and Methods*, **15**, 697-708.

Cordeiro, G.M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B*, **45**, 404-413.

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

Fisher, R.A., Corbet, A.S. and Williams, C.B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*, **12**, 42-58.

Frydenberg, M. and Jensen, J.L. (1989). Is the "improved likelihood statistic" really improved in the discrete case? *Biometrika*, **76**, 655-661.

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

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

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

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

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

Lawley, D.N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, **71**, 233-244.

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

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

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

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

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

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

Williams, C.B. (1944). The logarithmic series and its applications to biological problems. *Journal of Ecology*, **34**, 253-272.

Williams, C.B. (1964). *Patterns in Balance of Nature*. Academic Press, London.

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