RT-MAE 9521

LOCAL POWER OF THREE CLASSIC CRITERIA IN GENERALISED LINEAR MODELS WITH UNKNOWN DISPERSION

by

Silvia L.P. Ferrari, Denise A. Botter
and
Francisco Cribari-Neto

Key words: Asymptotic expansion; Chi-squared distribution; Composite hypothesis; Exponential family; Generalised linear models; Likelihood ratio test; Local power; Pitman alternatives; Score test; Wald test.

Classificação AMS: 62F05.
(AMS Classification)

- Agosto de 1995 -

Local power of three classic criteria in generalised linear models with unknown dispersion

BY SILVIA L.P. FERRARI

Departamento de Estatística, Universidade de São Paulo, Caixa Postal 66281, São Paulo/SP, 05389-970, Brazil

DENISE A. BOTTER

Departamento de Estatística, Universidade de São Paulo, Caixa Postal 66281, São Paulo/SP, 05389-970, Brazil

AND FRANCISCO CRIBARI-NETO

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

SUMMARY

Cordeiro, Ferrari & Paula (1994) obtained asymptotic expansions for the nonnull distribution functions of the likelihood ratio, score and Wald test statistics in generalised linear models under Pitman alternatives assuming known dispersion. This paper generalises some of their results by relaxing this assumption.

Some key words: Asymptotic expansion; Chi-squared distribution; Composite hypothesis; Exponential family; Generalised linear models; Likelihood ratio test; Local power; Pitman alternatives; Score test; Wald test.

1. INTRODUCTION

Three commonly used test criteria are the likelihood ratio, score and Wald tests. In a recent Biometrika paper, Cordeiro, Botter & Ferrari (1994) obtained asymptotic expansions to order $n^{-1/2}$, where n is the sample size, for the distribution functions of these statistics in generalised linear models under Pitman alternatives. They considered tests on a subset of linear parameters, on the dispersion parameter, and on all of the linear parameters. For the test on a subset of linear parameters, which is usually the most useful for practical applications, the authors have assumed that the dispersion parameter is known. This assumption is quite restrictive since for gamma, normal and inverse Gaussian models the dispersion is usually unkown. In this paper, we relax the assumption of known dispersion, and obtain nonnull asymptotic expansions for the likelihood ratio, score and Wald tests in generalised linear models for the test on a subset of parameters of the linear predictor. Our results show that these expansions are not sensitive to the uncertainty involved in the estimation of the dispersion parameter to order $n^{-1/2}$.

2. BACKGROUND

Let $y = (y_1, \ldots, y_n)'$ be a set of independent but not necessarily identically distributed observations, $l = l(\theta)$ be the log-likelihood function, where θ is an m-vector of unknown parameters, and U and K be the score function and Fisher's information matrix, respectively. The parameter vector θ is partitioned as $\theta = (\theta_1', \theta_2')'$, where the dimensions of θ_1 and θ_2 are q and $m - q$, respectively. The null hypothesis under test is $\mathrm{H}_0 : \theta_1 = \theta_1^{(0)}$ against a two-sided alternative, where $\theta_1^{(0)}$ is a known q-vector, and θ_2 is thus a vector of nuisance parameters. The following partitions are induced by the partition of the parameter vector:

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}, \quad K^{-1} = \begin{pmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ 0 & K_{22}^{-1} \end{pmatrix}.$$

We also define $M = K^{-1} - A$. The unrestricted and restricted maximum likelihood estimates of θ are denoted by $\hat{\theta} = (\hat{\theta}_1', \hat{\theta}_2')'$ and $\tilde{\theta} = (\theta_1^{(0)\prime}, \tilde{\theta}_2')'$, respectively. Functions evaluated at $\hat{\theta}$ are distinguished by the addition of a '^' whereas functions evaluated at $\tilde{\theta}$ are distinguished by the addition of a '~'. The likelihood ratio, Wald and score statistics are denoted by S_1, S_2 and S_3, respectively. Under mild regularity conditions, these three statistics have a limiting null χ_q^2 distribution, that is, $\mathrm{pr}(S_i \le c) = \mathrm{pr}(\chi_q^2 \le c) + o(1)$, $i = 1, 2, 3$.

Consider now a sequence of Pitman alternative hypotheses given by $\mathrm{H}_n : \theta_1 = \theta_1^{(0)} + \xi$, where $\xi = (\xi_1, \ldots, \xi_q)'$ with $\xi_i = O(n^{-1/2})$, $i = 1, \ldots, q$. In what follows, all suffices range from 1 to m. Let $U_i = \partial l / \partial \theta_i$, $U_{ij} = \partial^2 l / \partial \theta_i \partial \theta_j$, etc. We shall use the following notation for cumulants of log-likelihood derivatives (Lawley, 1956): $\kappa_{ij} = \mathrm{E}(U_{ij})$, $\kappa_{ijk} = \mathrm{E}(U_{ijk})$, $\kappa_{i,j} = \mathrm{E}(U_i U_j)$, $\kappa_{ij,k} = \mathrm{E}(U_{ij} U_k)$, $\kappa_{ij,kr} = \mathrm{E}(U_{ij} U_{kr}) - \kappa_{ij}\kappa_{kr}$, $\kappa_{i,j,kr} = \mathrm{E}(U_i U_j U_{kr}) - \kappa_{i,j}\kappa_{kr}$, $\kappa_{i,j,k,r} = \mathrm{E}(U_i U_j U_k U_r) - \kappa_{i,j}\kappa_{k,r} - \kappa_{i,k}\kappa_{j,r} - \kappa_{i,r}\kappa_{j,k}$. We have that the elements of K are $\kappa_{i,j} = -\kappa_{ij}$, and the corresponding elements of K^{-1} are $\kappa^{i,j} = -\kappa^{ij}$. Next, define δ as an m-vector given by

$$\delta = \begin{pmatrix} I_q \\ -K_{22}^{-1} K_{21} \end{pmatrix} \xi,$$

where I_q is the identity matrix of order q, and define the scalar parameter $\lambda = \delta' K \delta$.

Under a sequence of Pitman alternatives $\mathrm{H}_n : \theta_1 = \theta_1^{(0)} + \xi$ and regularity conditions which exclude lattice problems, the asymptotic expansions of the distribution functions of the three test statistics can be written as (Harris & Peers, 1980; Hayakawa, 1975)

$$\mathrm{pr}(S_i \le c) = G_{q,\lambda}(c) + \sum_{j=0}^{3} b_{ij} G_{q+2j,\lambda}(c) + o(n^{-1/2}),$$

where $G_{v,\lambda}(\cdot)$ is the distribution function of a noncentral chi-squared variate with v degrees of freedom and noncentrality parameter λ. The b's are written as follows: $b_{11} = (1/6)\sum_{i,j,k=1}^{p}\{(\kappa_{ijk} - 2\kappa_{i,j,k})\delta_i\delta_j\delta_k + 3(\kappa_{ijk} + 2\kappa_{i,jk})a_{ij}\delta_k\} - (1/2)\sum_{i=1}^{q}\sum_{j,k=1}^{p}(\kappa_{ijk} + \kappa_{i,jk})\xi_i\delta_j\delta_k$, $b_{12} = (1/6)\sum_{i,j,k=1}^{p}\kappa_{i,j,k}\delta_i\delta_j\delta_k$, $b_{13} = 0$, $b_{21} = (1/2)\sum_{i,j,k=1}^{p}\{(\kappa_{ijk} + 2\kappa_{i,jk})\delta_i\delta_j\delta_k - 2\kappa_{i,jk}m_{ij}\delta_k + (\kappa_{ijk} + 2\kappa_{i,jk})\kappa^{i,j}\delta_k\} - (1/2)\sum_{i=1}^{q}\sum_{j,k=1}^{p}(\kappa_{ijk} + \kappa_{i,jk})\xi_i\delta_j\delta_k$, $b_{22} = -(1/2)\sum_{i,j,k=1}^{p}(\kappa_{i,jk}\delta_i\delta_j\delta_k +$

$\kappa_{ijk} m_{ij} \delta_k)$, $b_{23} = -(1/6) \sum_{i,j,k=1}^{p} \kappa_{ijk} \delta_i \delta_j \delta_k$, $b_{31} = (1/6) \sum_{i,j,k=1}^{p} \{(\kappa_{ijk} - 2\kappa_{i,j,k}) \delta_i \delta_j \delta_k - 3\kappa_{i,j,k} m_{ij} \delta_k + 3(\kappa_{ijk} + 2\kappa_{i,jk}) a_{ij} \delta_k\} - (1/2) \sum_{i=1}^{q} \sum_{j,k=1}^{p} (\kappa_{ijk} + \kappa_{i,jk}) \xi_i \delta_j \delta_k$, $b_{32} = (1/2) \sum_{i,j,k=1}^{p} \kappa_{i,j,k} m_{ij} \delta_k$, $b_{33} = (1/6) \sum_{i,j,k=1}^{p} \kappa_{i,j,k} \delta_i \delta_j \delta_k$. Also, $b_{i0} = -(b_{i1} + b_{i2} + b_{i3})$, $i = 1, 2, 3$. All quantities, except for ξ, are evaluated under H_0.

3. Orthogonal parameters

Suppose $\theta = (\beta_1', \beta_2', \phi)'$ is our unknown parameter vector, where ϕ is a scalar parameter which is globally orthogonal to $\beta = (\beta_1', \beta_2')'$ in the sense of Cox & Reid (1987), with $\beta_1 = (\beta_1, \ldots, \beta_q)'$ and $\beta_2 = (\beta_{q+1}, \ldots, \beta_p)'$. Suppose also that the null hypothesis under test is $H_0 : \beta_1 = \beta_1^{(0)}$, where $\beta_1^{(0)}$ is a known q-vector. We can then write the information matrix as $K = \text{diag}\{K_\beta, \kappa_{\phi,\phi}\}$, where $K_\beta = E\{(\partial l/\partial \beta)(\partial l/\partial \beta)'\}$, $\kappa_{\phi,\phi} = -E(\partial^2 l/\partial \phi^2)$ and $l = l(\beta, \phi)$ is the log-likelihood function. K_β can be partitioned following the partition of the β vector as

$$K_\beta = \begin{pmatrix} K_{\beta 11} & K_{\beta 12} \\ K_{\beta 21} & K_{\beta 22} \end{pmatrix}.$$

We also have that

$$A = \begin{pmatrix} A_\beta & 0 \\ 0 & \kappa_{\phi,\phi}^{-1} \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} M_\beta & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{where} \quad A_\beta = \begin{pmatrix} 0 & 0 \\ 0 & K_{\beta 22}^{-1} \end{pmatrix}$$

and $M_\beta = K_\beta^{-1} - A_\beta$. It then follows that $m_{r\phi} = m_{\phi r} = m_{\phi\phi} = 0$, $a_{r\phi} = a_{\phi r} = 0$, for $r = 1, \ldots, p$, and $a_{\phi\phi} = \kappa_{\phi,\phi}^{-1}$, where $m_{\phi\phi}$ and $a_{\phi\phi}$ are the $(p+1, p+1)$th elements of the matrices M and A, respectively, and $m_{r\phi}$ and $a_{r\phi}$ are the $(r, p+1)$th elements of the matrices M and A, respectively.

Given the results above, it is possible to show that

$$\delta = \begin{pmatrix} \delta_\beta \\ 0 \end{pmatrix} \xi, \quad \text{where} \quad \delta_\beta = \begin{pmatrix} I_q \\ -K_{\beta 22}^{-1} K_{\beta 21} \end{pmatrix},$$

and $\lambda = \delta_\beta' K_\beta \delta_\beta$. That is, the noncentrality parameter λ is the same, regardless of whether or not ϕ is known. It is also possible to show using the general expressions for the b's in the previous section that $b_{ij} = b_{ij,\beta} + b_{ij,\beta\phi}$, where $b_{ij,\beta}$ equals the term b_{ij} for the case where ϕ is known. That is, the b's can be additively decomposed into their counterparts for the case where ϕ is known plus some extra terms that account for the uncertainty involved in the estimation of the parameter ϕ. It is also possible to show that $b_{ij,\beta\phi} = 0$ for $i = 1, 2, 3$ and $j = 2, 3$ and

$$b_{i1,\beta\phi} = \frac{1}{2} \sum_{k=1}^{p} (\kappa_{\phi\phi k} + 2\kappa_{\phi,\phi k}) \kappa_{\phi,\phi}^{-1} \delta_k \qquad (1)$$

for $i = 1, 2, 3$, where $\kappa_{\phi\phi k} = E(\partial^3 l/\partial^2 \phi \partial \beta_k)$ and $\kappa_{\phi,\phi k} = E\{(\partial l/\partial \phi)(\partial^2 l/\partial \phi \partial \beta_k)\}$.

4. NONNULL ASYMPTOTIC EXPANSIONS IN GENERALISED LINEAR MODELS

Let each observation y_l, $l = 1,\ldots,n$, be the realization of a continuous random variable Y_l with density function

$$\pi(y;\theta,\phi) = \exp[\phi\{y\theta_l - b(\theta_l)\} + a(y,\phi)], \tag{2}$$

where $a(\cdot,\cdot)$, $b(\cdot)$ and $c(\cdot)$ are known functions and θ_l and ϕ are possibly unknown parameters. Here, $\mathrm{E}(Y_l) = \mu_l = b'(\theta_l)$ and $\mathrm{var}(Y_l) = V_l/\phi$, where ϕ^{-1} is the dispersion parameter, $V = V(\mu) = \mathrm{d}\mu/\mathrm{d}\theta$ is the variance function, and $\theta = \int(1/V)\,\mathrm{d}\mu = q(\mu)$ is a strictly monotonic function of the mean. The linear predictor is given by $\eta = \sum_{j=1}^{p}\beta_j x_j = X\beta$, where X is an $n \times p$ $(p \leq n)$ matrix with rank p and β is a p-vector of unknown parameters. A generalised linear model is defined by a distribution in (2) and by a strictly monotonic, twice-differentiable link function $d(\mu) = \eta$ relating the mean to the linear predictor.

Next, we partition the p-vector β as in the previous section, that is, $\beta = (\beta_1', \beta_2')'$, where $\beta_1 = (\beta_1,\ldots,\beta_q)'$ and $\beta_2 = (\beta_{q+1},\ldots,\beta_p)'$ for $q \leq p$, which induces the partition of the model matrix as $X = (X_1 \;\; X_2)$. The null hypothesis is $\mathrm{H}_0 : \beta_1 = \beta_1^{(0)}$ and the alternative $\mathrm{H}_n : \beta_1 = \beta_1^{(0)} + \xi$, with $\xi_i = O(n^{-1/2})$, as in Section 3. The likelihood ratio, Wald and score statistics can be written as $S_1 = 2\{l(\hat\beta,\hat\phi) - l(\tilde\beta,\tilde\phi)\}$, $S_2 = (\hat\beta_1 - \beta_1^{(0)})'(\hat K^{11})^{-1}(\hat\beta_1 - \beta_1^{(0)})$ and $S_3 = \tilde U_1'\tilde K^{11}\tilde U_1$, respectively, where $\tilde\phi$ is the restricted maximum likelihood estimate of the unknown precision parameter ϕ.

For two-parameter full exponential family distributions with canonical parameters ϕ and $\phi\theta$, the term $a(y,\phi)$ in (2) can be written as $a(y,\phi) = d_1(\phi) + d_2(y)$. It then follows that in generalised linear models $\kappa_{\phi\phi k} = \kappa_{\phi,\phi k} = 0$, and hence equation (1) reduces to $b_{i1,\beta\phi} = 0$, $i = 1,2,3$. This implies that the b's are given by $b_{11} = (1/2)\sum[\phi\{(f-g)_l e_l t_l^2 - f_l t_l^3\} - f_l t_l z_{2ll}]$, $b_{12} = (\phi/6)\sum(f-g)_l t_l^3$, $b_{13} = 0$, $b_{21} = (1/2)\sum[\phi\{(f+g)_l e_l t_l^2 - f_l t_l^3\} - 2g_l t_l(z_{ll} - z_{2ll}) - f_l t_l z_{ll}]$, $b_{22} = (1/2)\sum\{-\phi g_l t_l^3 + (f+2g)_l t_l(z_{ll} - z_{2ll})\}$, $b_{23} = (\phi/6)\sum(f+2g)_l t_l^3$, $b_{31} = (1/2)\sum[\phi\{(f+g)_l e_l t_l^2 - f_l t_l^3\} - (f-g)_l t_l(z_{ll} - z_{2ll}) - f_l t_l z_{2ll}]$, $b_{32} = (1/2)\sum(f-g)_l t_l(z_{ll} - z_{2ll})$ and $b_{33} = (\phi/6)\sum(f-g)_l t_l^3$, where summations are over the n observations and $Z = \{z_{lm}\} = X(X'WX)^{-1}X'$, $Z_2 = \{z_{2lm}\} = X_2(X_2'WX_2)^{-1}X_2'$, $W = \mathrm{diag}\{w_1,\ldots,w_n\}$ with $w_l = V_l^{-1}(\mathrm{d}\mu_l/\mathrm{d}\eta_l)^2$, $t = (t_1,\ldots,t_n)' = X\delta$, $e = (e_1,\ldots,e_n) = X_1\delta$,

$$f = \frac{1}{V}\frac{\mathrm{d}\mu}{\mathrm{d}\eta}\frac{\mathrm{d}^2\mu}{\mathrm{d}\eta^2} \quad \text{and} \quad g = \frac{1}{V}\frac{\mathrm{d}\mu}{\mathrm{d}\eta}\frac{\mathrm{d}^2\mu}{\mathrm{d}\eta^2} - \frac{1}{V^2}\frac{\mathrm{d}V}{\mathrm{d}\mu}\left(\frac{\mathrm{d}\mu}{\mathrm{d}\eta}\right)^3.$$

Also, it can be shown that the noncentrality parameter is $\lambda = \phi\sum t_l^2 w_l$. That is, we obtain the same b's and λ as did Cordeiro, Botter and Ferrari (1994) for the case of known dispersion.

Some previous results have shown that the uncertainty involved in the estimation of the dispersion parameter does change the null distributions of the likelihood ratio and score statistics in generalised linear models to order n^{-1}; see Cordeiro (1987) and Cribari–Neto & Ferrari (1995), respectively. However, the results above show that the nonnull distributions of the likelihood ratio, Wald and score statistics to order $n^{-1/2}$ in generalised linear models do not change when one introduces unknown dispersion.

Acknowledgement

The partial financial support of Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brazil) is gratefully acknowledged.

References

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

Cordeiro, G.M., Botter, D.A. & Ferrari, S.L.P. (1994). Nonnull asymptotic distributions of three classic criteria in generalised linear models. *Biometrika* **81**, 709-20.

Cox, D.R. & Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B* **49**, 1-39.

Cribari-Neto, F. & Ferrari, S.L.P. (1995). Second order asymptotics for score tests in generalised linear models. *Biometrika* **82**, 426-32.

Hayakawa, T. (1975). The likelihood ratio criterion for a composite hypothesis under a local alternative. *Biometrika* **62**, 451-60.

Harris, P. & Peers, H.W. (1980). The local power of the efficient scores test statistic. *Biometrika* **67**, 525-29.

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

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

9501 - FAVA, L. On Interpenetrating Order Statistics. São Paulo, IME-USP. 1995. 32p. (RT-MAE-9501)

9502 - FERRARI,P.A.,KESTEN,H., MARTÍNEZ,S. R-Positivity, Quasi-Stationary Distributions and Ratio Limit Theorems for a Class of Probabilistic Automata. São Paulo, IME-USP. 1995. 39p. (RT-MAE-9502)

9503 - FERRARI, S.L.P., CORDEIRO, G.M., URIBE-OPAZO, M.A., CRIBARI-NETO, F. Improved Score Tests for One-Parameter Exponential Family Models. São Paulo, IME-USP. 1995. 21p. (RT-MAE-9503)

9504 - ARIZONO, H., BUENO, V.C. Bounds for System Reliability of Used Components under MIFR/$\Im_t$ Assumptions. São Paulo, IME-USP. 1995. 10p. (RT-MAE-9504)

9505 - SINGER, J.M., ANDRADE, D.F. Regression Models for the Analysis of Pretest/Posttest Data. São Paulo, IME-USP. 1995. 8p. (RT-MAE-9505)

9506 - SILVA, A.J., ZUAZOLA, P.L.I., WECHSLER, S. Partial exchangeability and extendibility. São Paulo, IME-USP. 1995. 21p. (RT-MAE-9506)

9507 - MADRUGA, M.G., OCHI-LOHMANN, T.H., OKAZAKI, K., PEREIRA, C.A.B., RABELLO-GAY, M.Z. Bayesian Dosimetry II: Credibility Intervals for Radiation Dose. São Paulo, IME-USP. 1995. 12p. (RT-MAE-9507)

9508 - FERRARI, S.L.P.; URIBE-OPAZO, M.A.; CRIBARI-NETO, F. Second order asymptotics for score tests in exponential family nonlinear models. São Paulo, IME-USP. 1995. 07p. (RT-MAE-9508)

9509 - WECHSLER, S.; BALLAS, D.; CARDOSO, R.H.A. Bayesian Optimization of Quality-of-life. São Paulo, IME-USP. 1995. 06p. (RT-MAE-9509)

9510 - BOTTER, D.A; CORDEIRO, G.M. Bartlett Corrections for Generalized Linear Models with Dispersion Covariates. São Paulo, IME-USP. 1995. 31p. (RT-MAE-9510)

9511 - PAULA, G.A.; ROJAS, O.V. On Restricted Hypotheses in Extreme Value Regression Models. São Paulo, IME-USP. 1995. 14p. (RT-MAE-9511)

9512 - ARIZON, H.; BUENO, V.C. Bounds for System Reliability of Used Components Under NBU$|\Im_t$ Assumptions. São Paulo, IME-USP. 1995. 10p. (RT-MAE-9512)

9513 - FERRARI, S.L.P.; BOTTER, D.A.; CORDEIRO, G.M.; CRIBARI-NETO, F. Second and Third Order Bias Reduction for One-Parameter Family Models. São Paulo, IME-USP. 1995. 07p. (RT-MAE-9513)

9514 - BELITSKY, V.; KOHAYAKAWA, Y. Inference on the Flip Rates from Density Dynamics for a Class of Spin-Flip Systems. São Paulo, IME-USP. 1995. 21p. (RT-MAE-9514)

9515 - BELITSKY, V. The change of sign of derivatives of the density function of a spin flip system with the order of derivation. São Paulo, IME-USP. 1995. 27p. (RT-MAE-9515)

9516 - PAULA, G.A. On Approximation of the Level Probabilities for Testing Ordered Parallel Regression Lines. São Paulo, IME-USP. 1995. 15p. (RT-MAE-9516)

9517 - PAULA, G.A. The Concept of Leverage in Restricted Regression. São Paulo, IME-USP. 1995. 09p. (RT-MAE-9517)

9518 - YOSHIDA, A.S.; LEITE, J.G.; BOLFARINE, H. Bayes Estimation of the Number of Component Processes of a Superimposed Process. São Paulo, IME-USP. 1995. 19p. (RT-MAE-9518)

9519 - SIMONIS, A. Metastability of the d-Dimensional Contact Process. São Paulo, IME-USP. 1995. 17p. (RT-MAE-9519)

9520 - FONTES, L.R.G.; ISOPI, M.; SIDORAVICIUS, V. Analyticity of the Density and Exponential Decay of Correlations in 2-d Bootstrap Percolation. São Paulo, IME-USP. 1995. 17p. (RT-MAE-9520)

The complete list of Relatórios do Departamento de Estatística, IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05389-970 - São Paulo, Brasil