Computational Statistics and Data Analysis 74 (2014) 110–124

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Improved likelihood inference in generalized linear models

Tiago M. Vargas [a], Silvia L.P. Ferrari [a,*], Artur J. Lemonte [b]

[a] Departamento de Estatística, Universidade de São Paulo, São Paulo/SP, Brazil
[b] Departamento de Estatística, Universidade Federal de Pernambuco, Recife/PE, Brazil

ARTICLE INFO

Article history:
Received 28 May 2013
Received in revised form 18 November 2013
Accepted 20 December 2013
Available online 20 January 2014

Keywords:
Bartlett correction
Bootstrap
Generalized linear models
Gradient statistic
Likelihood ratio statistic
Score statistic
Bartlett-type correction
Wald statistic

ABSTRACT

We address the issue of performing testing inference in generalized linear models when the sample size is small. This class of models provides a straightforward way of modeling normal and non-normal data and has been widely used in several practical situations. The likelihood ratio, Wald and score statistics, and the recently proposed gradient statistic provide the basis for testing inference on the parameters in these models. We focus on the small-sample case, where the reference chi-squared distribution gives a poor approximation to the true null distribution of these test statistics. We derive a general Bartlett-type correction factor in matrix notation for the gradient test which reduces the size distortion of the test, and numerically compare the proposed test with the usual likelihood ratio, Wald, score and gradient tests, and with the Bartlett-corrected likelihood ratio and score tests, and bootstrap-corrected tests. Our simulation results suggest that the corrected test we propose can be an interesting alternative to the other tests since it leads to very accurate inference even for very small samples. We also present an empirical application for illustrative purposes.[1]

1. Introduction

The likelihood ratio (LR), Wald and Rao score tests are the large-sample tests usually employed for testing hypotheses in parametric models. Another criterion for testing hypotheses in parametric models, referred to as the gradient test, was proposed by Terrell (2002). An advantage of the gradient statistic over the Wald and the score statistics is that it does not involve knowledge of the information matrix, neither expected nor observed. Additionally, the gradient statistic is quite simple to be computed. Here, it is worthwhile to quote Rao (2005): "The suggestion by Terrell is attractive as it is simple to compute. It would be of interest to investigate the performance of the [gradient] statistic". Also, Terrell's statistic shares the same first order asymptotic properties with the LR, Wald and score statistics. That is, to the first order of approximation, the LR, Wald, score and gradient statistics have the same asymptotic distributional properties either under the null hypothesis or under a sequence of Pitman alternatives, i.e. a sequence of local alternative hypotheses that shrink to the null hypothesis at a convergence rate $n^{-1/2}$, n being the sample size; see Lemonte and Ferrari (2012a). Recently, the gradient test has been the subject of some research papers. In particular, Lemonte (2011, 2012, 2013) and Lemonte and Ferrari (2012b) provide comparison among the local power of the classic tests and the gradient test in some specific regression models. The authors showed that the gradient test can be an interesting alternative to the classic tests.

The LR, Wald, score and gradient statistics for testing composite or simple null hypothesis $\mathcal{H}_0$ against an alternative hypothesis $\mathcal{H}_a$, in regular problems, have a χ^2_k null distribution asymptotically, where k is the difference between the dimensions of the parameter spaces under the two hypotheses being tested. However, in small samples, the use

* Corresponding author. Tel.: +55 11 3091 6129; fax: +55 11 3091 6130.
E-mail addresses: silviaferrari@usp.br, silviaferrari.usp@gmail.com (S.L.P. Ferrari).

1 Supplementary Material presents derivation of Bartlett-type corrections to the gradient tests, and the computer code used in Section 6 (Appendix A).

0167-9473/$ – see front matter
http://dx.doi.org/10.1016/j.csda.2013.12.002

of these statistics coupled with their asymptotic properties become less justifiable. One way of improving the χ^2 approximation for the exact distribution of the LR statistic is by multiplying it by a correction factor known as the Bartlett correction (see Bartlett, 1937). This idea was later put into a general framework by Lawley (1956). The χ^2 approximation for the exact distribution of the score statistic can be improved by multiplying it by a correction factor known as the Bartlett-type correction. It was demonstrated in a general framework by Cordeiro and Ferrari (1991). Recently, Vargas et al. (2013) demonstrated how to improve the χ^2 approximation for the exact distribution of the gradient statistic in wide generality by multiplying it by a Bartlett-type correction factor. There is no Bartlett-type correction factor to improve the χ^2 approximation of the exact distribution of the Wald statistic in a general setting. The Bartlett and Bartlett-type corrections became widely used for improving the large-sample χ^2 approximation to the null distribution of the LR and score statistics in several special parametric models. In recent years there has been a renewed interest in Bartlett and Bartlett-type factors and several papers have been published giving expressions for computing these corrections for special models. Some references are Zuckeret al. (2000); Manor and Zucker (2004); Lagos and Morettin (2004); Tu et al. (2005); van Giersbergen (2009); Bai (2009); Lagos et al. (2010); Lemonte et al. (2010); Lemonte and Ferrari (2011); Noma (2011); Fujita et al. (2010); Bayer and Cribari-Neto (2013); Lemonte et al. (2012), among others. The reader is referred to Cordeiro and Cribari-Neto (1996) for a detailed survey on Bartlett and Bartlett-type corrections.

The generalized linear models (GLMs), first defined by Nelder and Wedderburn (1972), are a large class of statistical models for relating responses to linear combinations of predictor variables, including many commonly encountered types of dependent variables and error structures as special cases. It generalizes the classical normal linear model, by relaxing some of its restrictive assumptions, and provides methods for the analysis of non-normal data. Additionally, the GLMs have applications in disciplines as widely varied as agriculture, demography, ecology, economics, education, engineering, environmental studies, geography, geology, history, medicine, political science, psychology, and sociology. We refer the read to Lindsey (1997) for applications of GLMs in these areas. In summary, the GLM approach is attractive because it (1) provides a general theoretical framework for many commonly encountered statistical models; (2) simplifies the implementation of these different models in statistical software, since essentially the same algorithm can be used for estimation, inference and assessing model adequacy for all GLMs. Introductions to the area are given by Firth (1991) and Dobson and Barnett (2008), whereas McCullagh and Nelder (1989) and Hardin and Hilbe (2007) give more comprehensive treatments.

The asymptotic χ^2 distribution of the LR, Wald, score and gradient statistics is used to test hypotheses on the model parameters in GLMs, since their exact distributions are difficult to obtain in finite samples. However, for small sample sizes, the χ^2 distribution may not be a trustworthy approximation to the exact null distributions of the LR, Wald, score and gradient statistics. Higher order asymptotic methods, such as the Bartlett and Bartlett-type corrections, can be used to improve the LR, Wald, score and gradient tests. Several papers have focused on deriving matrix formulas for the Bartlett and Bartlett-type correction factors in GLMs. For example, some efforts can be found in the works by Cordeiro (1983, 1987), who derived an improved LR statistic. An improved score statistic was derived by Cordeiro et al. (1993) and Cribari-Neto and Ferrari (1995). These results will be revised in this paper. Although the algebraic forms of the Bartlett and Bartlett-type correction factors are somewhat complicated, they can be easily incorporated into a computer program. This might be a worthwhile practice, since the Bartlett and Bartlett-type corrections act always in the right direction and, in general, give a substantial improvement.

This paper is concerned with small sample likelihood inference in GLMs. First, we derive a general Bartlett-type correction factor in matrix notation to improve the inference based on the gradient statistic in the class of GLMs when the number of observations available to the practitioner is small. Further, in order to evaluate and compare the finite-sample performance of the improved gradient test in GLMs with the usual LR, Wald, score and gradient tests, and with the improved LR and score tests, we also perform Monte Carlo simulation experiments by considering the gamma regression model and the inverse Gaussian regression model. Bootstrap-based tests and monotonic versions of the Bartlett-corrected score and gradient statistics are also included in the Monte Carlo experiments. The simulation study on the size properties of these tests evidences that the improved gradient test proposed in this paper can be an appealing alternative to the classic asymptotic tests in this class of models when the number of observations is small. We shall emphasize that we have not found any comprehensive simulation study in the statistical literature comparing the classical uncorrected and corrected large-sample tests in GLMs. This paper fills this gap, and includes the gradient test and its Bartlett-type corrected version derived here in the simulation study.

The article is organized in the following form. In Section 2, we define the class of GLMs and discuss estimation and hypothesis testing inference on the regression parameters. Improved likelihood-based inference is presented in Section 3. We present the Bartlett-corrected LR and score statistics, and derive a Bartlett-type correction factor for the gradient statistic. Tests on the precision parameter are provided in Section 4. Monte Carlo simulation results are presented and discussed in Section 5. An application to real data is considered in Section 6. The paper closes up with a brief discussion in Section 7.

2. The model, estimation and testing

Suppose the univariate random variables $Y_1, \ldots, Y_n$ are independent and each Y_l has a probability density function in the following family of distributions:

$$\pi(y; \theta_l, \phi) = \exp\big\{\phi[y\theta_l - b(\theta_l) + c(y)] + a(y, \phi)\big\}, \quad l = 1, \ldots, n, \tag{1}$$