Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computational Statistics & Data Analysis 50 (2006) 3179–3198

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

Approximate power of score test for variance heterogeneity under local alternatives in nonlinear models☆

Jin-Guan Lin[a, b, *], Bo-cheng Wei[a]

[a] *Department of Mathematics, Southeast University, Nanjing 210096, China*
[b] *Department of Mathematics, Jiangsu Institute of Education, Nanjing 210013, China*

Received 1 June 2005; accepted 22 June 2005
Available online 20 July 2005

Abstract

Lin and Wei [2003. Testing for heteroscedasticity in nonlinear regression models. Comm. Statist. Theory Methods 32, 171–192] developed the score test for heteroscedasticity in nonlinear regression models and investigated the power of this test through Monte Carlo simulations. The main purpose of this paper is to present an approach for estimating the local power for the score test, based on a noncentral χ^2 approximation to such power under contiguous alternatives. The approach is also extended to nonlinear models with AR(1) errors. The methods are applied to the problem of local power calculations for the score tests of heteroscedasticity in European rabbit data [Ratkowsky, 1983. Nonlinear Regression Modelling. Marcel Dekker, New York, pp. 108–110]. Simulation studies are presented which indicate that the asymptotic approximation to the finite-sample situation is good over a wide range of parameter configurations.

Keywords: AR(1) errors; Heteroscedasticity; Local power; Noncentral χ^2 approximation; Nonlinear models; Residual plots; Score test; Simulation

---

☆ The project supported by NSFC 10371016, SSFC 04BTJ002, NSSPFC LXZ0415 and the Post-Doctorial Grant in Southeast University.

* Corresponding author. Tel.: +86 2586 271015; fax: +86 2586 271015.
E-mail address: ljgzxc@sina.com.cn (J.-G. Lin).

0167-9473/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2005.06.013

1. Introduction

It is a standard assumption for nonlinear regression models that the error terms all have equal variances (Bates and Watts, 1988; Seber and Wild, 1989). The violation of this assumption can have adverse consequences for the efficiency of estimators, so it is important to detect the variance heterogeneity in regression. The general form of nonlinear models with heteroscedasticity is

$$Y_i = f(\mathbf{x}_i, \boldsymbol{\beta}) + \varepsilon_i, \quad \varepsilon_i \sim N\left(0, \sigma_i^2\right), \quad i = 1, 2, \ldots, n, \tag{1}$$

where the ε_i's are assumed to be independent and $\boldsymbol{\beta}$ is a $p \times 1$ unknown vector of regression coefficients; f is a known and twice differentiable function. To test for heteroscedasticity, the variance model considered is of the following form:

$$\sigma_i^2 = \sigma^2 m(\mathbf{z}_i, \boldsymbol{\lambda}) = \sigma^2 m_i,$$

where $\sigma^2 > 0$ is a scale parameter, $m_i = m(\mathbf{z}_i, \boldsymbol{\lambda})$ is the variance weight function of Y_i, $\mathbf{z}_i$'s are covariates, which are related to the covariates $\mathbf{x}_i$'s or to other covariates (Cook and Weisberg, 1983), $\boldsymbol{\lambda}$ is a $q \times 1$ unknown vector of variance parameters, and m is a known differentiable function of variance in $\boldsymbol{\lambda}$. We assume that there exists a unique value $\boldsymbol{\lambda}_0$ of $\boldsymbol{\lambda}$ such that $m(\mathbf{z}_i, \boldsymbol{\lambda}_0) = 1$ for all i. Obviously, if $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0$, then $\sigma_i^2 = \sigma^2$ and the Y_i's have constant variance. Model (1) is a direct extension of the model presented in Cook and Weisberg (1983). They suggested using log-linear and power product models as the variance function of m. Under the above assumptions, the test for heterogeneity of variance is equivalent to a test of hypotheses

$$\mathrm{H}_0:\ \boldsymbol{\lambda} = \boldsymbol{\lambda}_0, \quad \mathrm{H}_1:\ \boldsymbol{\lambda} \neq \boldsymbol{\lambda}_0. \tag{2}$$

Let $\boldsymbol{\theta} = \left(\boldsymbol{\lambda}^{\mathrm{T}}, \boldsymbol{\beta}^{\mathrm{T}}, \sigma^2\right)^{\mathrm{T}}$, then for hypothesis (2), $\boldsymbol{\lambda}$ is the parameter of interest and $\boldsymbol{\gamma} = \left(\boldsymbol{\beta}^{\mathrm{T}}, \sigma^2\right)^{\mathrm{T}}$ is the nuisance parameter. Thus, the log-likelihood of $\boldsymbol{\theta}$ for $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^{\mathrm{T}}$ can be written as

$$l(\boldsymbol{\theta}) = C - \frac{n}{2}\log\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\log m_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n} m_i^{-1}(Y_i - f(\mathbf{x}_i, \boldsymbol{\beta}))^2,$$

where C is a constant. Furthermore, the Fisher information matrix of $\mathbf{Y}$ for $\boldsymbol{\theta}$ under the null hypothesis H_0 is

$$\mathrm{I}_{\mathbf{Y}}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{1}{2}\dot{M}^{\mathrm{T}} M^{-2}\dot{M} & 0 & \frac{1}{2\sigma^2}\dot{M}^{\mathrm{T}} M^{-1}\mathbf{1}_n \\ 0 & \frac{1}{\sigma^2}\dot{F}^{\mathrm{T}} M^{-1}\dot{F} & 0 \\ \frac{1}{2\sigma^2}\mathbf{1}_n^{\mathrm{T}} M^{-1}\dot{M} & 0 & \frac{n}{2\sigma^4} \end{bmatrix}, \tag{3}$$

where $M = \mathrm{diag}(m_1, m_2, \ldots, m_n)$, $\dot{M} = \left(\partial m(\mathbf{z}_i, \boldsymbol{\lambda})/\partial\lambda_j\right)_{n\times q}$, $\mathbf{1}_n = \left(\underbrace{1, 1, \ldots, 1}_{n}\right)^{\mathrm{T}}$, $\dot{F} = \left(\partial f(\mathbf{x}_i, \boldsymbol{\beta})/\partial\beta_j\right)_{n\times p}$. From Lin and Wei (2003), the score test of hypothesis (2) is based on