Available online at www.sciencedirect.com
ScienceDirect
Economics Letters 95 (2007) 374–379
economics letters
www.elsevier.com/locate/econbase

Score tests of normality in bivariate probit models

Anthony Murphy *

Hertford College, Catte Street, Oxford OX1 3BW, UK

Received 17 October 2005; received in revised form 27 October 2006; accepted 21 November 2006
Available online 26 March 2007

Abstract

A relatively simple and convenient score test of normality in the bivariate probit model is derived. Monte Carlo simulations show that the small sample performance of the bootstrapped test is quite good. The test may be readily extended to testing normality in related models.

Keywords: Score test; Bivariate probit; Normality; Gram–Charlier series

JEL classification: C25

1. Introduction

In this paper I show how to construct a simple score (LM) test of the normality assumption in bivariate probit and related models. To date, the normality assumption in these models has seldom been tested, even though the parameter estimates are inconsistent when the distribution of the random error terms is mis-specified.

A score test of normality has an obvious advantage over likelihood ratio or Wald tests. I follow convention and focus on skewness and excess kurtosis when deriving the test statistic. The alternative hypothesis used is based on a truncated or type AA bivariate Gram Charlier series used by Lee (1984) and Smith (1985) for example. The score test involves conditional/truncated expectations of terms such as $u_1^j u_2^k$, where $u' = (u_1, u_2)'$ is a bivariate normal random vector. Unfortunately, the cited papers do not contain explicit expressions for these expectations. These may be simulated but explicit expressions are more accurate and convenient.

* Tel.: +44 1865 279455; fax: +44 1865 279466.
E-mail address: anthony.murphy@nuffield.ox.ac.uk.

0165-1765/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2006.11.010

2. The bivariate probit model

The bivariate probit model may be derived from a pair of regression or reduced form equations $y_1^* = x_1' \beta_1 + u_1$ and $y_2^* = x_2' \beta_2 + u_2$ with latent dependent variables y_1^* and y_2^*. The random errors u_1 and u_2 are distributed as standard bivariate normal variables with correlation coefficient D. Only the signs of the latent variables are observed. It is useful to define the indicator and sign variable y_j and s_j, with $y = 1(y_1^* > 0)$ and $s_j = 2y_j - 1$ for $j = 1,2$.

In order to derive the log likelihood for this model, expressions for the four probabilities P_{11}, P_{10}, P_{01} and P_{00} are required, where $P_{10} = \text{prob}(y_1 = 1, y_2 = 0) = \text{prob}(u_1 > -x_1'\beta_1, u_2 \leq -x_2'\beta_2)$ for example. In the bivariate probit model, the probabilities are given by:

$$P_{y_1 y_2} = \phi(s_1 x_1' \beta_1, s_2 x_2' \beta_2, s_1 s_2 \rho) = \phi_{y_1 y_2} \tag{1}$$

where ϕ is the standard normal bivariate c.d.f., and $P_{y_1 y2}$ and $\phi_{y_1 y2}$ are used for short. Thus, for example, $\phi_{10} = \phi(x_1'\beta_1, -x_2'\beta_2, -\rho)$.

3. The type AA Gram Charlier alternative

The results in Lee (1984) and Smith (1985), inter alia, suggest that a truncated or type AA bivariate Gram Charlier series may be a suitable alternative to the standard bivariate normal density. The Gram Charlier expansion for a regular standardized bivariate density $f(u_1, u_2)$ with correlation coefficient ρ is:

$$\begin{aligned} f(u_1, u_2) &= \phi(u_1, u_2, \rho) + \sum_{j+k} \sum_{\geq 3} (-1)^{j+k} \frac{c_{jk}}{j!k!} D_1^j D_2^k \phi(u_1, u_2, \rho) \\ &= \phi(u_1, u_2, \rho) \left\{ 1 + \sum_{j+k} \sum_{\geq 3} \frac{K_{jk}}{j!k!} H_{jk}(u_1, u_2, \rho) \right\} \end{aligned} \tag{2}$$

where the $H_{jk}(u_1, u_2, \rho) = ((-1)^{j+k} D_1^j\ D_1^k\ \phi(u_1, u_2, \rho)) / \phi(u_1, u_2, \rho)$ are bivariate Hermite polynomials, the Ds are differentiation operators, the K_{jk}s are cumulants and $\phi(u_1, u_2, \rho)$ is the bivariate standard normal density.

Truncating Eq. (2) by omitting all terms with $j+k>4$ yields the type AA Gram Charlier series, which may not be a proper p.d.f. However, it is only being used to generate a test statistic with, it is hoped, some power against local departures from normality. Pagan and Vella (1989) suggest using the density in Gallant and Nychka (1987) to test for normality.

However, this approach is no simpler than the one used in this paper.

Expressions for the four probabilities P_{00}, P_{10}, P_{01} and P_{11} are required. Under the type AA Gram Charlier alternative, the probabilities are:

$$\begin{aligned} P_{y_1 y_2} &= \int \int_R f(u_1, u_2, \rho) \mathrm{d}u_1 \mathrm{d}u_2 = \phi_{y_1 y_2} + \sum_{j+k=3,4} \sum \frac{K_{jk}}{j!k!} \int \int_R H_{jk}(u_1, u_2, \rho) \phi(u_1, u_2, \rho) \mathrm{d}u_1 \mathrm{d}u_2 \\ &= \phi_{y_1 y_2} \left(1 + \sum_{j+k=3,4} \sum \frac{K_{jk}}{j!k!} E_{y_1 y_2} H_{jk}(u_1, u_2, \rho) \right) \end{aligned} \tag{3}$$