1984

Estimation and Testing in a Regression Model with Spherically Symmetric Errors

Aman Ullah

Victoria Zinde-Walsh

ISSN: 0318-725X
ISBN: 0-7714-0604-5

RESEARCH REPORT 8422

ESTIMATION AND TESTING IN A REGRESSION MODEL WITH SPHERICALLY SYMMETRIC ERRORS

by

Aman Ullah* and Victoria Zinde-Walsh

November, 1984

ABSTRACT

In this paper we show that both the restricted and unrestricted likelihood estimators of the error variance, obtained under normality, are not in general robust against the spherical distribution of errors. Thus, care should be taken in interpreting the residual variances in the applied work. The robustness of Rao's score, Wald's and the likelihood ratio tests, are also analyzed.

*The research support from the NSERC to A. Ullah is gratefully acknowledged. The authors are grateful to B. K. Sinha and R. S. Singh for useful discussions and suggestions.

ESTIMATION AND TESTING IN A REGRESSION MODEL WITH SPHERICALLY SYMMETRIC ERRORS

by

Aman Ullah and Victoria Zinde-Walsh

1. INTRODUCTION

In this paper we consider the problem of estimation and testing in the context of a regression model in which errors follow a spherically symmetric distribution. It has been shown that the maximum likelihood (ML) estimator of the regression coefficients under normal errors is the same as derived under spherically symmetric errors. However, the ML estimator of the error variance under normal distribution is not the same as obtained under the spherical distribution of errors. Finally, it has been shown that the exact likelihood ratio (LR) test, Rao's Score (RS) test and Wald's (W) test for testing restrictions are robust against the spherical distribution of errors. In the special case of multivariate-t errors, the results compare with those in Ullah and Zinde-Walsh (1983).

2. THE MAIN RESULTS

Let us consider the regression model

(2.1) $$y = X\beta + u$$

where y is an $n \times 1$ vector of the dependent variable, X is an $n \times p$ matrix of non-stochastic exogenous variables, β is a $p \times 1$ vector of unknown parameters and u is an $n \times 1$ error vector.

It is known that if u is distributed as multivariate normal with mean vector zero and variance-covariance matrix $\sigma^2 I$, then the maximum likelihood (ML) estimators for β and σ^2, respectively, are given by

$$(2.2) \qquad b = (X'X)^{-1}X'y \qquad \text{and} \qquad s^2 = \frac{1}{n}(y - Xb)'(y - Xb).$$

Also, for testing H_o: $R\beta = r$ against H_1: $R\beta \neq r$, where R is an $m \times p$ constant matrix of rank m, and r is an $m \times 1$ vector of constants, the LR test criterion turns out to be

$$(2.3) \qquad LR = -2 \log \ell, \text{ with } \ell = \max_{R\beta = r, \sigma^2} L(\beta,\sigma^2) \Big/ \max_{\beta,\sigma^2} L(\beta,\sigma^2) = (s^2/\hat{\sigma}^2)^{n/2}.$$

in which we have used the restricted ML estimators of σ^2 and β given, respectively, as

$$(2.4) \qquad \hat{\sigma}^2 = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta}) \quad \text{and} \quad \hat{\beta} = b - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(Rb - r).$$

Likewise, the RS and W test statistics are, respectively, given by

$$(2.5) \qquad RS = (Rb - r)'[R(X'X)^{-1}R']^{-1}(Rb - r)/\hat{\sigma}^2 \quad \text{and} \quad W = \frac{\hat{\sigma}^2}{s^2} RS.$$

We now consider the class of spherically symmetric distributions

$$(2.6) \qquad f(u) = \sigma^{-n}\phi\left(\frac{u'u}{\sigma^2}\right) = \sigma^{-n}\phi(w), \qquad w = \frac{u'u}{\sigma^2}$$

where $\phi(w)$ is a decreasing function on $[0,\infty)$. Note that both the multivariate normal and the student-t are members of the class of distributions in (2.6).

We introduce b_ϕ, s^2_ϕ; and $\hat{\beta}_\phi$, $\hat{\sigma}^2_\phi$ to represent the unrestricted and restricted ML estimators, respectively, of β and σ^2 under the spherically symmetric distributions in (2.6). Similarly, LR_ϕ, W_ϕ and RS_ϕ will represent LR, W and RS test statistics under (2.6).

Next, we consider w_ϕ as the unique solution for w in the equation

$$(2.7) \qquad \frac{n}{2} + w\,\frac{\phi'(w)}{\phi(w)} = 0$$

where $\phi'(w)$ is the first derivative of $\phi(w)$. A sufficient condition for the existence of a unique solution w_ϕ would be $\phi''(w)\phi(w) - \phi'^2(w) \le 0$, that is the concavity of $\log \phi(w)$.

The main results of the paper can then be stated in the following Theorems:

THEOREM 1: Under (2.6), the unrestricted and restricted ML estimators respectively, of the parameters β and σ^2 are given by

$$(2.8) \qquad b_\phi = b \quad , \qquad s^2_\phi = \frac{1}{w_\phi}(y - Xb)'(y - Xb)$$

and

$$(2.9) \qquad \hat{\beta}_\phi = \hat{\beta} \quad , \qquad \hat{\sigma}^2_\phi = \frac{1}{w_\phi}(y - X\hat{\beta})'(y - X\hat{\beta})$$

where b and $\hat{\beta}$ are as given in (2.2) and (2.4), respectively, and w_ϕ is the unique solution of w in (2.7).

PROOF: Using (2.6), write the log of the likelihood function for the parameters of the model $y = X\beta + u$ as

$$(2.10) \qquad \log L(\beta, \sigma^2) = -\frac{n}{2}\log \sigma^2 + \log \phi(w) = f(y)$$

where $w = (y - X\beta)'(y - X\beta)/\sigma^2 = u'u/\sigma^2$. The first-order conditions for the maximization of $\log L(\beta,\sigma^2)$ are

$$(2.11) \qquad \frac{\partial}{\partial\beta}\log L(\beta,\sigma^2) = -\frac{2\phi'(w)}{\phi(w)}\frac{X'(y-X\beta)}{\sigma^2} = 0$$

$$(2.12) \qquad \frac{\partial}{\partial\sigma^2}\log L(\beta,\sigma^2) = -\frac{n}{2\sigma^2} - \frac{w}{\sigma^2}\frac{\phi'(w)}{\phi(w)} = 0.$$

It is clear from (2.11) that the unrestricted ML estimator of β is $b_\phi = b$ as given in (2.2) and (2.8). Further, from (2.12) we get

$$\phi(w) + \frac{2}{n}w\phi'(w) = 0$$

as given in (2.7). Let $w_\phi = (y-X\beta)'(y-X\beta)/\sigma^2$ be the only solution of this equation for which the $\log L(\beta,\sigma^2)$ is at maximum. Then the ML estimator of σ^2 is

$$(2.13) \qquad s_\phi^2 = \frac{1}{w_\phi}(y-Xb_\phi)'(y-Xb_\phi) = \frac{1}{w_\phi}(y-Xb)'(y-Xb)$$

as given in (2.8).

For the restricted ML estimators of β and σ^2 we maximize $H = \log L(\beta,\sigma^2) + \lambda(r-R\beta)$ where λ is the vector of Lagrangian coefficients. The first-order conditions are

$$(2.14) \qquad \frac{\partial H}{\partial\beta} = -\frac{2\phi'(w)}{\phi'(w)}\frac{X'(y-X\beta)}{\sigma^2} - \lambda R = 0$$

$$(2.15) \qquad \frac{\partial H}{\partial\sigma^2} = -\frac{n}{2\sigma^2} - \frac{w}{\sigma^2}\frac{\phi'(w)}{\phi(w)} = 0, \qquad \frac{\partial H}{\partial\lambda} = r - R\beta = 0.$$

The solutions for β and σ^2 from (2.14) and (2.15) can then be verified to be as given in (2.9).

THEOREM 2: Under (2.6), the LR, RS and W test statistics for testing $H_o: R\beta = r$ against $H_1: R\beta \neq r$ are given by

$$\text{(2.16)} \qquad LR_\phi = LR, \quad RS_\phi = c_\phi^{-1}\delta_\phi RS \quad \text{and} \quad W_\phi = c_\phi \delta_\phi W$$

where LR, RS and W are as given in (2.3) and (2.5), respectively, c_ϕ is a constant whose value depends on ϕ, and $\delta_\phi = w_\phi/n$.

PROOF: Note that the likelihood function for β and σ^2 is

$$\text{(2.17)} \qquad L(\beta,\sigma^2) = \sigma^{-n}\phi(w) = \sigma^{-n}\phi\left[\frac{(y - X\beta)'(y - X\beta)}{\sigma^2}\right].$$

Thus, using the results in (2.8) and (2.9)

$$\text{(2.18)} \qquad \max_{\beta,\sigma^2} L(\beta,\sigma^2) = s_\phi^{-n}\,\phi(w_\phi)$$

$$\text{(2.19)} \qquad \max_{R\beta = r,\sigma^2} L(\beta,\sigma^2) = \hat{\sigma}_\phi^{-n}\,\phi(w_\phi).$$

It is then obvious that $LR_\phi = LR$ as given in (2.16).

Next, we derive the RS test statistic. For this, we first obtain the second-order derivatives of $\log L(\beta,\sigma^2)$ in (2.10). These are

$$\text{(2.20)} \qquad \frac{\partial^2}{\partial\beta\partial\beta'}\log L(\beta,\sigma^2) = \frac{2\phi'(w)}{\phi'(w)}\frac{X'X}{\sigma^2} + 4\left[\frac{\phi''(w)}{\phi(w)} - \left(\frac{\phi'(w)}{\phi(w)}\right)^2\right]\frac{X'uu'X}{\sigma^4}$$

$$\text{(2.21)} \qquad \frac{\partial^2}{\partial\beta\partial\sigma^2}\log L(\beta,\sigma^2) = \frac{2}{\sigma^4}\left[\frac{\phi'(w)}{\phi(w)} + w\left(\frac{\phi''(w)}{\phi(w)} - \left(\frac{\phi'(w)}{\phi(w)}\right)^2\right)\right]X'u$$

where $\phi''(w)$ is the second derivative of $\phi(w)$.

Now notice that the right-hand side of (2.21) is an odd function of $u = y - X\beta$. Therefore

$$E \frac{\partial^2}{\partial\beta\partial\sigma^2} \log L(\beta,\sigma^2) = 0. \tag{2.22}$$

Further, we observe that

$$E \frac{\phi'(w)}{\phi(w)} = \int_u \frac{\phi'(w)}{\phi(w)} \sigma^{-n}\phi(w)\,du = \sigma^{-n}\int_u \phi'(w)\,du = c_{1\phi}$$

$$E\left(\frac{\phi'(w)}{\phi(w)}\right)^2 \frac{uu'}{\sigma^2} = \sigma^{-n}\int_u \left(\frac{\phi'(w)}{\phi(w)}\right)^2 \frac{uu'}{\sigma^2}\phi(w)\,du = c_{2\phi}I \tag{2.23}$$

$$E \frac{\phi''(w)}{\phi(w)} \frac{uu'}{\sigma^2} = \sigma^{-n}\int_u \phi''(w) \frac{uu'}{\sigma^2}du = c_{3\phi}I$$

where $c_{1\phi}$, $c_{2\phi}$ and $c_{3\phi}$ are constants whose values depend on ϕ.

Thus, using (2.23)

$$E[-\frac{\partial^2}{\partial\beta\partial\beta'} \log L(\beta,\sigma^2)] = c_\phi \frac{X'X}{\sigma^2} = I(\beta) \tag{2.24}$$

where

$$c_\phi = 4(c_{2\phi} - c_{3\phi}) - 2c_{1\phi} \tag{2.25}$$

whose values depend on ϕ (see (2.23)) and $I(\beta)$ represents the information matrix of β. Further, because of (2.22), the information matrix of β and σ^2, $I(\beta,\sigma^2)$ will be a block diagonal matrix. Therefore, the RS test statistic can be written as

$$RS_\phi = \hat{d}' I^{-1}(\hat{\beta})\hat{d} \tag{2.26}$$

where $I(\hat{\beta})$ and $\hat{d}$ are $I(\beta)$ and $\partial \log L(\beta,\sigma^2)/\partial\beta$, respectively, evaluated at the restricted estimators $\beta = \hat{\beta}_\phi = \hat{\beta}$ and $\sigma^2 = \hat{\sigma}^2_\phi$. These are $\hat{d} = X'(y - X\hat{\beta})/\hat{\sigma}^2_\phi$ and $I(\hat{\beta}) = c_\phi\hat{\sigma}_\phi^{-2}X'X$. The result in (2.16) for RS_ϕ is then obvious, by noting that $(y - X\hat{\beta})'X(X'X)^{-1}X'(y - X\hat{\beta}) = (Rb - r)'[R(X'X)^{-1}R']^{-1}(Rb - r)$.

Finally, consider the W statistic

$$W_\phi = (Rb - r)' [RI^{-1}(b)R']^{-1}(Rb-r); \tag{2.27}$$

where $I(b) = I(b_\phi) = c_\phi X'X/s_\phi^2$. Substituting this in (2.27) the result in (2.16) for W_ϕ follows.

Q.E.D.

2.1 Remarks on the Results

The following remarks are based on the results in the Theorems.

1. The results in Theorm 1 show that $b_\phi = b$ and $\hat{\beta}_\phi = \hat{\beta}$. Thus both the unrestricted and restricted ML estimators of β are numerically robust against the spherically symmetric distribution of errors. However, $s_\phi^2 \neq s^2$ and $\hat{\sigma}_\phi^2 \neq \hat{\sigma}^2$ which imply that the restricted as well as unrestricted estimators of σ^2 are not in general robust. For example, if u follows a multivariate student-t distribution with mean zero and variance $\sigma^2 I$, then it can easily be verified that $s_\phi^2 = \gamma(\gamma-2)^{-1}s^2 \neq s^2$ and $\hat{\sigma}_\phi^2 = \gamma(\gamma-2)^{-1}\hat{\sigma}^2 \neq \hat{\sigma}^2$. Further, if

$$f(u) = c(\sigma^2)^{\frac{-n}{2}} \exp[\frac{1}{2}\frac{u'u}{\sigma^2}] \qquad \text{for } \frac{u'u}{\sigma^2} \leq \omega \tag{2.27}$$

$$= 0 \qquad \text{for } \frac{u'u}{\sigma^2} > \omega ,$$

where c is a constant and $\omega < n$, then the ML estimator of σ^2 is $s_\phi^2 = \frac{1}{\omega}\hat{u}'\hat{u} \neq s^2$. Similarly, if we consider a distribution

$$f(u) = c(\sigma^2)^{\frac{-n}{2}} (1 + \frac{u'u}{\sigma^2})^{-k} , \quad 2k > n, \tag{2.28}$$

then $s_\phi^2 = (\frac{2k}{n} - 1)\hat{u}'\hat{u}$, which will coincide with s^2 only in the case where $k = n+1/2$.

2. The result in Theorem 2 shows that although the LR test is numerically robust ($LR_\phi = LR$) the RS and W tests are not numerically robust ($RS_\phi \neq RS$ and $W_\phi \neq W$) against the spherically symmetric distribution of errors.

This result is similar to one in Ullah and Zinde-Walsh (1983) in the multivariate-t case.

Next, we observe that, under H_o, all the LR_ϕ, RS_ϕ and W_ϕ tests are monotonic functions of central $F = q(RRSS - URSS)/mURSS \sim f(m,q)$ where $RRSS = (y - x\hat{\beta})'(y - x\hat{\beta})$, $URSS = (y - xb)'(y - xb)$ and $q = n-p$. That the F-statistic is distributed as a central F under H_o whatever $\phi(\omega)$ may be follows from the result of David (1977). The implication of this result is that the test criterion for the normal distribution of errors remains the same under the spherically symmetric distribution of errors. That is we reject H_o if

$$LR_\phi > LR_\phi^\alpha \quad \text{or} \quad LR > LR^\alpha$$

$$RS_\phi = c_\phi^{-1}\delta_\phi RS > c_\phi^{-1}\delta_\phi RS^\alpha \quad \text{or} \quad RS > RS^\alpha$$

$$W_\phi = c_\phi \delta_\phi W > c_\phi \delta_\phi W^\alpha \quad \text{or } W > W^\alpha$$

where LR^α, RS^α and W^α are the α% critical points of LR, RS and W, respectively.

From the above findings it is clear that although RS_ϕ and W_ϕ are not numerically robust they are robust for testing purposes. This latter robustness we call inference robustness. Note that the LR test achieves both inference and numerical robustness. These distinct ideas of robustness were not explicitly stated in Ullah and Zinde-Walsh (1983).

The inequality results for LR_ϕ, RS_ϕ and W_ϕ can be developed for a given ϕ as in Ullah and Zinde-Walsh. Also, as in the multivariate-t case of Ullah and Zinde-Walsh if the large sample critical value based on chi-square is used, then the tests LR_ϕ, RS_ϕ and W_ϕ will differ with respect to their sizes and powers in small samples and there may be conflict between their conclusions.

References

David, A. P., "Spherical Matrix Distributions and a Multivariate Model," J. Roy. Statistic. Soc., Sec. B., 1977, 39, 254-261.

Ullah, A. and V. Zinde-Walsh, "On the Robustness of LM, LR and W Tests in Regression Models," Research Report 8316, University of Western Ontario, 1983. Forthcoming in Econometrica, July 1984.