▩ 연구논문

Comparison of Score Test and Other Tests in Testing Equality between Two Linear Regressions

Jeong-il Kim
Dept. of Statistics, Taejon University

Abstract

Testing problem of equality between two linear regressions is considered. Chow test is discussed and likelihood ratio test(LR test), F-test and Score test are derived. Score test are compared with other tests and show they are asymptotically equivalent.

1. Introduction

The problem of testing the equality of sets of regression coefficients in two regressions arises with considerable frequency in econometrics and biostatistics literature. Since Chow(1960) proposed the test for the equality of two linear regression models, various methods have been proposed(see Fisher(1970), Ali and Silver(1985), Tsurumi and Shelfin(1985) and Conerly and Mansfield(1988)). Chow presented systematically the test involved, related the prediction interval and the analysis of covariance within the framework of general linear hypothesis, and extended the results to testing the equality between subsets of coefficients.

We can also develop the LR test and F-test along Chow's idea. On the other hand, Rao(1948) introduced Score test and it has an advantage over LR test because Score test statistic only requires estimation under the null. Futhermore Score test statistics are asymptotically equivalent to Wald and LR test statistics under both null and Pitman alternative hypothesis(Serfling 1980).

In this paper Score test will be presented for testing the equality of two linear regressions. Two general linear regression models with p coefficients can be written in the following form:

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \quad \boldsymbol{\varepsilon}_i \sim N(\boldsymbol{0}, \theta \boldsymbol{I}_{n_i}), \quad i=1, 2 \tag{1}$$

or in matrix notations

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix} \sim N(\mathbf{0}, \theta \mathbf{I}_N), \qquad (2)$$

where the $\mathbf{y}_i$ are $n_i \times 1$ vectors of observable responses, the $\mathbf{X}_i$ are $n_i \times p$ matrices of known constants, $\boldsymbol{\beta}_i$ are $p \times 1$ vectors of unknown parameters, $\boldsymbol{\varepsilon}_i$ are independent $n_i \times 1$ vectors of unobservable random errors, $N = n_1 + n_2$, and $\mathbf{I}_N$ is the identity matrix. Assume that $n_1, n_2 > p$ and the $\mathbf{X}_i$ are of full-column rank and it is desired to test $H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$.

In Section 2 we will present Chow-test, LR test and F-test and show they are equivalent. In Section 3 we will derive Score-test and finally we will show Chow test and LR test are asymptotically equivalent to Score test in Section 4.

2. Chow test, LR test and F-test

To test the equality between sets of coefficients in two linear regressions, Chow compared the sum of squares of the residuals(SSE) under the null hypothesis to those under the alternative hypothesis. He derived the test statistic from the reason that the sum of squares of the residuals under H_0 will be shown to equal to the sum of squares of residuals under the alternative hypothesis plus the sum of squares of the deviations between the two sets of estimates of $\mathbf{y}$ under two hypothesis. He proposed the following test statistic

$$F_0 = \frac{(SSE(H_0) - SSE(H_1))/p}{SSE(H_1) - /(N-2p)}, \qquad (3)$$

where $SSE(H_0)$ is SSE of the reduced model under null and $SSE(H_1)$ SSE of the full model under alternative. This test statistic turns out to be the standard analysis of covariance test.

For an alternative test one may employ the method of likelihood ratios. From the likelihood function, we can obtain the following MLEs of $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and θ by differentiating the log-likelihood with respect to corresponding parameters:

$$\hat{\boldsymbol{\beta}}_{1\Omega} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}_1,$$

$$\hat{\boldsymbol{\beta}}_{2\Omega} = (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{y}_2,$$

$$\hat{\theta}_\Omega = \sum_{i=1}^{2} \| \boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}}_i \|^2 / N.$$

Under the null hypothesis, $H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 (= \boldsymbol{\beta})$, the MLEs of $\boldsymbol{\beta}$ and θ are given as:

$$\hat{\boldsymbol{\beta}}_0 = \{(\boldsymbol{X}_1' \boldsymbol{X}_1) + (\boldsymbol{X}_2' \boldsymbol{X}_2)\}^{-1} (\boldsymbol{X}_1' \boldsymbol{y}_1 + \boldsymbol{X}_2' \boldsymbol{y}_2),$$
$$\hat{\theta}_0 = \sum_{i=1}^{2} \| \boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}}_i \|^2 / N.$$

Then the critical region of LR test is determined by

$$\frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{\hat{\theta}_\Omega}{\hat{\theta}_0} < \lambda_0 . \tag{4}$$

and this test statistic will be shown to be equal to the Chow test statistic.

Also this critical region may be derived by the F-test. That is, first note that

$$\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2 \sim N(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2 , \{(\boldsymbol{X}_1' \boldsymbol{X}_1)^{-1} + (\boldsymbol{X}_2' \boldsymbol{X}_2)^{-1}\} \theta)$$

Thus under the null, $\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2 = \boldsymbol{0}$,

$$(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2)' \{(\boldsymbol{X}_1' \boldsymbol{X}_1)^{-1} + (\boldsymbol{X}_2' \boldsymbol{X}_2)^{-1}\}^{-1} (\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2)/\theta \sim \chi^2(p),$$

and

$$\frac{(N-2p)\hat{\theta}_\Omega}{\theta} \sim \chi^2(N-2p). \tag{5}$$

The ratio of the above two independent quantities ends up with the same as the statistics of LR test and Chow-test.

3. Score test

In relationship of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, we can express $\boldsymbol{\beta}_2 = \boldsymbol{\beta}_1 + \boldsymbol{\delta}$ and the equivalent hypothesis $H_0 : \boldsymbol{\delta} = \boldsymbol{0}$. But the reparametrization like this has some difficulties because of information matrix. Thus we reparametrize

$$\boldsymbol{\beta}_2 = \boldsymbol{\beta}_1 + \{(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} + (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\}\Psi. \tag{6}$$

Therefore the test for $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ is equivalent to $\Psi = \mathbf{0}$, i.e., the null hypothesis now becomes $H_0: \Psi = \mathbf{0}$ with $(\boldsymbol{\beta}_1, \theta)$ as nuisance parameter. Now we shall derive the Score test statistic

From the joint density of $(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)$,

$$f(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2) \propto \theta^{-(n_1+n_2)/2} \exp\{-\frac{\boldsymbol{V}}{2\theta}\},$$

where

$$\boldsymbol{V} = \|\boldsymbol{y}_1 - \boldsymbol{X}_1\boldsymbol{\beta}_1\|^2 + \|\boldsymbol{y}_2 - \boldsymbol{X}_2(\boldsymbol{\beta}_1 + \boldsymbol{\Delta}\Psi)\|^2,$$

$$\boldsymbol{\Delta} = \{(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} + (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\},$$

one can derive the following marginal Score function and Information matrix for Ψ.

$$\boldsymbol{U}_\Psi = \{\boldsymbol{\Delta}'(\boldsymbol{X}_2\boldsymbol{y}_2 - (\boldsymbol{X}_2'\boldsymbol{X}_2)\boldsymbol{\beta}_1)\boldsymbol{\Delta}'(\boldsymbol{X}_2'\boldsymbol{X}_2)\boldsymbol{\Delta}\Psi\}/\theta,$$
$$\boldsymbol{I}_{\Psi\Psi} = \boldsymbol{\Delta}'(\boldsymbol{X}_2'\boldsymbol{X}_2)\boldsymbol{\Delta}/\theta. \tag{7}$$

And we can derive the following Information matrix of $(\Psi, \boldsymbol{\beta}_1, \theta)$

$$\boldsymbol{I} = \frac{1}{\theta}\begin{pmatrix} \boldsymbol{I}_{\Psi\Psi} & \boldsymbol{I}_{\Psi\beta_1} & \boldsymbol{I}_{\Psi\theta} \\ \boldsymbol{I}_{\beta_1\Psi} & \boldsymbol{I}_{\beta_1\beta_1} & \boldsymbol{I}_{\beta_1\theta} \\ \boldsymbol{I}_{\theta\Psi} & \boldsymbol{I}_{\theta\beta_1} & \boldsymbol{I}_{\theta\theta} \end{pmatrix}$$

$$= \frac{1}{\theta}\begin{pmatrix} \boldsymbol{\Delta}(\boldsymbol{X}_2'\boldsymbol{X}_2)\boldsymbol{\Delta} & \boldsymbol{\Delta}(\boldsymbol{X}_2'\boldsymbol{X}_2) & \mathbf{0} \\ (\boldsymbol{X}_2'\boldsymbol{X}_2)\boldsymbol{\Delta} & (\boldsymbol{X}_1'\boldsymbol{X}_1)+(\boldsymbol{X}_2'\boldsymbol{X}_2) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{\theta\theta} \end{pmatrix}$$

And the inverse matrix of $\boldsymbol{I}$, $\boldsymbol{I}^{-1}$, is

$$\boldsymbol{I}^{-1} = \theta\begin{pmatrix} \boldsymbol{I}_{\Psi\Psi}^{-} & \boldsymbol{I}_{\Psi\beta_1}^{-} & \boldsymbol{I}_{\Psi\theta}^{-} \\ \mathrm{I}_{\Psi\beta_1}^{-} & \mathrm{I}_{\beta_1\beta_1}^{-} & \mathrm{I}_{\beta_1\theta}^{-} \\ \mathrm{I}_{\Psi\theta}^{-} & \mathrm{I}_{\beta_1\theta}^{-} & \mathrm{I}_{\theta\theta}^{-} \end{pmatrix}.$$

and

$$I_{\bar{\Psi\Psi}} = \{(X_1' X_1)^{-1} + (X_2' X_2)^{-1}\}^{-1} \theta.$$

And under the null hypothesis, the MLEs of $\beta_1 = \beta$ and θ are

$$\hat{\beta} = \{(X_1' X_1) + (X_2' X_2)\}^{-1} (X_1' y_1 + X_2' y_2),$$
$$\hat{\theta} = \sum_{i=1}^{2} \| y_i - X_i \hat{\beta} \|^2 / N.$$

Denoting $\eta = (\Psi, \beta, \theta)$, Score function and Fisher information matrix estimated under the null are

$$U_\Psi(\hat{\eta}_0) = \{(X_1' X_1)^{-1} X_1' y_1 - (X_2' X_2)^{-1} X_2' y_2\} / \hat{\theta}$$
$$I_{\bar{\Psi}}(\hat{\eta}_0) = \{\sum_{i=1}^{2} (X_i' X_i)^{-1}\}^{-1} \hat{\theta}.$$

The statitic of Score test becomes

$$S = U'_\Psi(\hat{\eta}_0) I_{\bar{\Psi}}(\hat{\eta}_0) U_\Psi(\hat{\eta}_0)$$
$$= (\hat{\beta}_1 - \hat{\beta}_2)' [(X_1' X_1)^{-1} + (X_2' X_2)^{-1}]^{-1} (\hat{\beta}_1 - \hat{\beta}_2) / \hat{\theta}. \quad (8)$$

This statistic S is asymptotically distributed as χ^2 with p degrees of freedom when both n_1 and n_2 tend to infinity, a necessary condition for unbounded growth of the information of Ψ (Cox and Hinkley, 1974).

4. Relationships among the statistics

We shall show that the LR-test or Chow-test is asymptotically equivalent to the Score test. Suppose that any random variable, T_n, is distributed as central $\chi^2(n)$, then $T_n / n \rightarrow 1$ in probability by Chebyshev's inequality. Hence Using this fact the denominator of equation (3) converges to 1 in probability as N goes to infinity. And since $\theta_0 \rightarrow \theta$ a.s.. by Slutsky theorem, the $p \times F_0$ and S have the same asymptotic distribution, that is, the central χ^2 distribution with degrees of freedom p.

REFERENCES

[1] Ali, M. M. and Silver, J. L. (1985), "Tests for equality between sets of coefficients in two linear regressions under heteroscedasticity," *Journal of American Statistical Association*, Vol. 80, pp. 730-735.

[2] Chow, G. C. (1960), "Tests of equality between sets of coefficients in two linear regressions," *Econometrica*, Vol. 28, pp. 591-605.

[3] Conerly, M. D. and Mansfield, E. R. (1988), "An approximate test for comparing heteroscedastic regression models," *Journal of American Statistical Association*, Vol. 83, pp. 811-817.

[4] Cook, R. D. and Weisberg S. (1983), "Diagnostics for heteroscedasticity in regression," *Biometrika*, Vol. 70, No. 1, pp. 1-10.

[5] Cox, D. R. and Hinkley, D. V. (1974), *Theoretical Statistics*, London Chapman & Hall.

[6] Fisher, F. M. (1970), "Tests of equality between sets of coefficients in two linear regressions: An expository note," *Econometrica*, Vol. 38, pp. 361-366.

[7] Rao, C. R. (1948), "Large sample tests of statistical hypothesis concerning several parameters with applications to problems of estamation," *Proceedings of the Cambridge Philosophical Society*, 44 in Regression, *Biometrika*, Vol. 70, No. 1, pp. 1-10.

[8] Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

[9] Tsurumi, H. and Sheflin, N. (1985), "Some tests for the constancy of regressions under heteroscedasticity," *Econometrica*, Vol. 27, pp. 221-234.