Colloquium Biometricum 44
2014, 109−117

A COMPARISON OF SEVERAL TESTS FOR EQUALITY OF COEFFICIENTS IN QUADRATIC REGRESSION MODELS UNDER HETEROSCEDASTICITY

Małgorzata Szczepanik, Dorota Domagała

Department of Applied Mathematics and Computer Science
University of Life Sciences
Głęboka 28, 20-612 Lublin, Poland
e-mails: malgorzata.szczepanik@up.lublin.pl,
dorota.domagala@up.lublin.pl

Summary

This paper reviews and compares several tests of equality between the vectors of coefficients in the two polynomial (quadratic) regression models. The empirical significance levels of the examined tests are calculated by means of Monte Carlo simulations. The test ranks are presented based on the loss function.

Keywords and phrases: heteroscedasticity, regression models, Monte Carlo simulations

Classification AMS 2010: 62F03, 62J99, 65C05

1. Introduction

When a regression is used to represent a relationship between variables, the researcher could ask whether the same relationship holds for two groups of experimental units. The answer can be obtained by testing the equality of regression coefficients.

We consider the two regression models

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \; i = 1,2 \tag{1.1}$$

where $\mathbf{y}_i$ is $n_i \times 1$ vector, $\mathbf{X}_i$ is $n_i \times k$ matrix with rank k, $\boldsymbol{\beta}_i$ is $k \times 1$ vector of regression coefficients. We assume that error terms $\boldsymbol{\varepsilon}_i \sim N_{n_i}\left(\mathbf{0}, \sigma_i^2 \mathbf{I}_{n_i}\right)$ and $\boldsymbol{\varepsilon}_1$ is independent of $\boldsymbol{\varepsilon}_2$.

The null hypothesis to be tested is $H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ against the alternative $H_0 : \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$.

Chow (1960) proposed the following statistic to test the equality of regression coefficients

$$F = \frac{\mathbf{e}^{\mathrm{T}}\mathbf{e} - \mathbf{e}_1^{\mathrm{T}}\mathbf{e}_1 - \mathbf{e}_2^{\mathrm{T}}\mathbf{e}_2}{k} : \frac{\mathbf{e}_1^{\mathrm{T}}\mathbf{e}_1 + \mathbf{e}_2^{\mathrm{T}}\mathbf{e}_2}{n_1 + n_2 - 2k} , \tag{1.2}$$

where $n_1, n_2 > k$ and $\mathbf{e}^{\mathrm{T}}\mathbf{e} = \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)^{\mathrm{T}}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)$ is the residual sum of squares in the model containing observations of both models in (1.1) and $\mathbf{e}_i^{\mathrm{T}}\mathbf{e}_i = \left(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_i\right)^{\mathrm{T}}\left(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_i\right)$ are the residual sums of squares calculated separately for each of the models (1.1). The vectors $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}_i$ are the least squares estimators of $\boldsymbol{\beta}, \boldsymbol{\beta}_i$ respectively.

Chow showed that if $\sigma_1^2 = \sigma_2^2$ (when homoscedasticity holds) and under H_0 statistic (1.2) is distributed $F(k, n_1 + n_2 - 2k)$.

However, the assumption about homoscedasticity is not often fulfilled and there is heteroscedasticity between the models, which means that $\sigma_1^2 \neq \sigma_2^2$. A number of tests were devised to compare two heteroscedastic regression models, for example in the papers by Conerly and Mansfield (1988), Honda and Ohtani (1986), Thursby (1992), Weerahandi (1987). The comparisons of chosen tests with the assumption of linear regression were conducted by Thursby (1992), Tsurumi and Sheflin (1985) and Szczepanik and Wesołowska-Janczarek (2006).

2. Tests under heteroscedasticity

Kadiyala and Gupta (1978) proposed a test based on W/k statistics, where

$$W = \left(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2\right)^{\mathrm{T}} \left[S_1^2\left(\mathbf{X}_1^{\mathrm{T}}\mathbf{X}_1\right)^{-1} + S_2^2\left(\mathbf{X}_2^{\mathrm{T}}\mathbf{X}_2\right)^{-1}\right]^{-1} \left(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2\right) \tag{2.1}$$

and $S_i^2 = \mathbf{e}_i^{\mathrm{T}}\mathbf{e}_i / (n_i - k)$ and the critical value is $F(\alpha; k, n_1 + n_2 - 2k)$. This test will be referred to as the AChow test (follow by Thursby 1992).
Honda and Ohtani (1986) modified test statistic (2.1) as

$$W - 2\left(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2\right)^{\mathrm{T}} \mathbf{A}^{-1}\left(d_1\mathbf{A}_1\mathbf{A}^{-1}\mathbf{A}_1 + d_2\mathbf{A}_2\mathbf{A}^{-1}\mathbf{A}_2\right)\mathbf{A}^{-1}\left(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2\right),$$

where $\mathbf{A}_i = S_i^2\left(\mathbf{X}_i^{\mathbf{T}}\mathbf{X}_i\right)^{-1}$, $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$, $d_i = 1/(n_i - k)$.
The critical value is $\chi^2(\alpha; k)$.
Weerahandi's (1987) test is based on *p*-value

$$p_v = 1 - \frac{1}{B\left((n_1 - k)/2, (n_2 - k)/2\right)} \times \int_0^1 f(R) R^{(n_1 - k - 2)/2} (1 - R)^{(n_2 - k - 2)/2}\, dR,$$

where $B\left((n_1 - k)/2, (n_2 - k)/2\right)$ is the beta function. The random variable *R* has the beta distribution with parameters $(n_1 - k)/2, (n_2 - k)/2$ and

$$f(R) = F_{k,r}\left(\frac{r}{k}\left(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2\right)^{\mathrm{T}} \left(\frac{\mathbf{e}_1^{\mathrm{T}}\mathbf{e}_1}{R}\left(\mathbf{X}_1^{\mathrm{T}}\mathbf{X}_1\right)^{-1} + \frac{\mathbf{e}_2^{\mathrm{T}}\mathbf{e}_2}{1 - R}\left(\mathbf{X}_2^{\mathrm{T}}\mathbf{X}_2\right)^{-1}\right)^{-1} \left(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2\right)\right)$$

where $r = n_1 + n_2 - 2k$ and $F_{k,r}$ is the cumulative distribution function of the *F* distribution with k, r degrees of freedom.

3. Monte Carlo design

The Monte Carlo experiments are based on the regression equations:

$$\begin{aligned} y_{1j} &= a_1 + b_1 x_{1j} + c_1 x_{1j}^2 + \varepsilon_{1j}, \quad j = 1, 2, \ldots, n_1, \\ y_{2j} &= a_2 + b_2 x_{2j} + c_2 x_{2j}^2 + \varepsilon_{2j}, \quad j = n_1 + 1, \ldots, n_1 + n_2, \end{aligned} \tag{3.1}$$

where we assume that $\varepsilon_{1j} \sim N\left(0, \sigma_1^2\right)$ and $\varepsilon_{2j} \sim N\left(0, \sigma_2^2\right)$.

To examine the empirical significance level we take in all the simulations: $a_i = b_i = c_i = \sigma_2^2 = 1,\ i = 1,2.$ For σ_1^2 the values 0.1, 0.2, 0.5, 0.8, 1, 2, 5, 10 are used to check tests under the extreme and moderate heteroscedasticity and also under homoscedasticity. The nine combinations of pairs n_1, n_2 were taken into consideration, namely (10,10), (20,20), (30,30), (50,50), (10,20), (10,50), (20,30), (20,50) and (30,50). The values of x_{1j} and x_{2j} are calculated as the equidistant and increasing numbers from [0,9] interval. When $n_1 = n_2$, we obtain $x_{1j} = x_{2j}$.

In every experiment, for appointed σ_1^2 and (n_1, n_2), the values of 1,000,000 y_{ij} according to (3.1) were generated. The Mersenne Twister pseudorandom number generator (Matsumoto and Nishimura 1998) was used in the simulations. The simulations and calculations were programmed in Pascal language. Also the software library LAPACK (Linear Algebra PACKage 2013) was very helpful to carry out the calculations. For each of 72 experiments the empirical significance level α^* was calculated as the quotient of the number of rejected null (true) hypothesis to the number of simulations (1,000,000). The nominal significance level is 0.05.
A part of the results is presented in Tables 1-4.

Table 1. Chow test. Empirical significance level x 100

	σ_1^2/σ_2^2							
$(\boldsymbol{n_1}, \boldsymbol{n_2})$	**0.1**	**0.2**	**0.5**	**0.8**	**1**	**2**	**5**	**10**
(10,10)	7.12	6.32	5.28	5.03	5.01	5.30	6.31	7.13
(20,20)	6.03	5.68	5.18	5.05	5.00	5.16	5.70	6.03
(50,50)	5.41	5.3	5.09	5.03	5.01	5.07	5.27	5.44
(10,20)	0.83	1.16	2.41	3.96	5.02	9.97	18.86	24.53
(10,50)	0.01	0.05	0.82	2.96	5.01	17.23	40.86	54.94
(20,30)	1.73	2.11	3.21	4.34	4.97	7.71	11.89	14.36
(30,50)	1.05	1.43	2.75	4.15	4.97	8.57	13.8	16.77

Table 2. AChow test. Empirical significance level x 100

	σ_1^2/σ_2^2							
(n_1, n_2)	**0.1**	**0.2**	**0.5**	**0.8**	**1**	**2**	**5**	**10**
(10,10)	7.12	6.32	5.28	5.03	5.01	5.30	6.31	7.13
(20,20)	6.03	5.68	5.18	5.05	5.00	5.16	5.70	6.03
(50,50)	5.41	5.3	5.09	5.03	5.01	5.07	5.27	5.44
(10,20)	5.16	4.99	5.29	5.83	6.15	7.34	8.79	9.47
(10,50)	5.24	5.88	7.50	8.51	8.97	10.15	11.18	11.6
(20,30)	5.34	5.14	5.00	5.13	5.17	5.61	6.23	6.6
(30,50)	5.14	5.03	5.02	5.15	5.18	5.54	5.93	6.11

The following conclusions can be drawn in relation to Chow test:

– experiments with $n_1 = n_2$: empirical significance levels are always equal or larger than 0.05. The results are close to the nominal (0.05) significance level when moderate heteroscedasticity ($\sigma_1^2 = 0.5, 0.8, 2$) and homoscedasticity occur. The empirical levels for extreme heteroscedasticity are markedly different from 0.05,

– experiments with $n_1 < n_2$: the values α^* are different from the nominal significance level when $\sigma_1^2 \neq 1$. Moreover, the empirical levels of significance are always less than the nominal when $\sigma_1^2 < \sigma_2^2$. On the other hand, α^* is larger than 0.05 when $\sigma_1^2 > \sigma_2^2$.

The conclusions about AChow test are as follows:

– the results are always larger than or equal to 0.05 and are also less dispersed compared with Chow test when $n_1 < n_2$,

– experiments with $n_1 = n_2$: the same α^* values are obtained as in Chow test because $x_{1j} = x_{2j}$,

– experiments with $n_1 < n_2$: the test performs well in situations where $\sigma_1^2 = 0.2, 0.5$. The larger heteroscedasticity the larger α^* (except for the case $n_1 = 10, n_2$=50).

Table 3. Honda -Ohtani test. Empirical significance level x 100

	σ_1^2/σ_2^2							
(n_1, n_2)	**0.1**	**0.2**	**0.5**	**0.8**	**1**	**2**	**5**	**10**
(10,10)	7.24	6.94	6.31	6.11	6.09	6.32	6.93	7.25
(20,20)	6.08	5.97	5.71	5.67	5.65	5.70	5.98	6.07
(50,50)	5.42	5.39	5.3	5.27	5.25	5.28	5.37	5.45
(10,20)	5.93	5.8	5.93	6.25	6.42	7	7.38	7.43
(10,50)	5.49	5.83	6.63	7.02	7.16	7.38	7.40	7.37
(20,30)	5.66	5.57	5.48	5.56	5.56	5.78	6.03	6.14
(30,50)	5.36	5.31	5.32	5.38	5.38	5.57	5.71	5.76

Table 4. Weerahandi test. Empirical significance level x 100

	σ_1^2/σ_2^2							
(n_1, n_2)	**0.1**	**0.2**	**0.5**	**0.8**	**1**	**2**	**5**	**10**
(10,10)	4.18	3.62	2.93	2.76	2.74	2.93	3.62	4.19
(20,20)	4.66	4.45	4.12	4.02	4.01	4.11	4.46	4.68
(50,50)	4.88	4.83	4.69	4.66	4.64	4.67	4.8	4.90
(10,20)	3.34	2.82	2.07	2.56	2.54	3.34	3.51	4.07
(10,50)	0.80	0.47	0.53	1.04	1.44	2.72	2.46	3.75
(20,30)	4.07	3.73	3.21	3.45	3.29	3.82	4.53	4.82
(30,50)	4.37	4.13	3.86	3.71	3.83	4.27	4.59	4.89

The conclusions about Honda-Ohtani test:

- the results are always larger than or equal to 0.05,
- experiments with $n_1 = n_2$: α^* values are slightly larger compared to Chow and AChow tests,

– experiments with $n_1 < n_2$: the empirical significance levels are slightly larger compared to AChow test except for the situations when $\sigma_1^2 = 5, 10$ and the case $n_1 = 10, n_2$=50.

The conclusions about Weerahandi test:

– the results are always less than 0.05,
– experiments with $n_1 = n_2$: α^* values are close to the nominal when the assumed heteroscedasticity is more extremal: $\sigma_1^2 = 0.1, 0.2, 5, 10$,
– experiments with $n_1 < n_2$: the best results are obtained for $\sigma_1^2 = 5, 10$.

4. The comparison of the tests

To sum up and compare results of the examined test, the loss function proposed in Thursby (1992) is used:

$$LS_1 = \left(\sum_{i=1}^{n_e} \left(100 \times \alpha_i^* - 5\right)^2 / n_e \right)^{1/2},$$

where n_e denotes number of experiments.
The values of LS_1 are presented in Tables 5 and 6.

Table 5. The loss values and ranks of tests

	all combinations of sample sizes (all experiments) $n_e = 72$		$n_1 = n_2$, $n_e = 36$	
Test	LS_1	rank	LS_1	rank
Chow	9.26	4	0.66	1
AChow	1.75	3	0.66	1
Honda-Ohtani	1.14	1	0.92	2
Weerahandi	1.71	2	0.95	3

Table 6. The loss values and ranks of tests with respect to the degree of heteroscedasticity

	moderate heteroscedasticity and homoscedasticity $\sigma_1^2 = 0.5, 0.8, 1, 2$ $n_e = 36$		$n_1 < n_2$ $\sigma_1^2 = 5, 10$ $n_e = 10$		$n_1 < n_2$ $\sigma_1^2 = 0.1, 0.2$ $n_e = 10$	
Test	LS_1	rank	LS_1	rank	LS_1	rank
Chow	2.85	4	23.85	4	4.16	4
AChow	1.42	2	3.62	3	0.32	1
Honda-Ohtani	1.04	1	1.70	2	0.61	2
Weerahandi	1.91	3	1.10	1	2.44	3

5. Conclusions

We have considered the empirical significance levels of four tests. As shown in Table 5 the results of all 72 experiments imply that Honda-Ohtani test is the best. However, when $n_1 = n_2$ Chow and AChow tests seem to be good. In the experiments with moderate heteroscedasticity and homoscedasticity (Table 6) again Honda-Ohtani test seems to be the best. Weerahandi test is the best in the case of heteroscedasticity of 5 and 10 when $n_1 < n_2$. For the experiments where $n_1 < n_2$ and the variance of the first model is 0.1 or 0.2 the AChow test and the Honda-Ohtani test appear to be good choices.

The obtained ranks of tests are similar to results in Szczepanik and Wesołowska-Janczarek (2006). The differences concern extreme heteroscedasticity. However it should be noted that the assumptions about the models were not the same and the x_{1j} and x_{2j} in (3.1) were chosen in different way at the present experiments.

Examining the empirical powers of these tests will be the next step of our studies.

References

Chow G. C. (1960). Tests of Equality between Sets of Coefficients in Two Linear Regressions. *Econometrica* 28, 3, 591─605.

Conerly M. D., Mansfield E. R. (1988). An Approximate Test for Comparing Heteroscedastic Regression Models. *Journal of the American Statistical Association* 83, No. 403, 811─817.

Honda Y., Ohtani K. (1986). Modified Wald tests in tests of equality between sets of coefficients in two linear regressions under heteroscedasticity. *The Manchester School of Economic and Social Studies* 54, 208─218.

Kadiyala K.R., Gupta S. (1978). Tests for pooling cross-sectional data in the presence of heteroskedasticity, *Bulletin of the Institute of Mathematical Statistics* 7, 275─276.

Matsumoto M., Nishimura T. (1998). Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans. on Modeling and Computer Simulation* 8, No. 1, 3─30.

Szczepanik M., Wesołowska-Janczarek M. (2006). Porównanie testów do sprawdzania identyczności dwóch modeli regresji w przypadku heteroscedastyczności. *Colloquium Biometryczne* 36, 257─266.

Thursby J. G. (1992). A comparison of several exact and approximate tests for structural shift under heteroscedasticity. *Journal of Econometrics* 53, 363─386.

Tsurumi H., Sheflin N. (1985). Some tests for the constancy of regression under heteroscedasticity. *Journal of Econometrics* 27, 221─234.

Weerahandi S. (1987). Testing regression equality with unequal variances. *Econometrica* 55, No. 5, 1211─1215.

LAPACK–Linear Algebra PACKage. (2013). www.netlib.org/lapack.