WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Chow Test: Detecting Structural Changes in Regression Models

***Siddamsetty Upendra* [a]*, Dr. R. Abbaiah* [b]*, Dr. P. Balasiddamani* [c]**

[a] *Researchsolor in Dept. of Statistics, S V University, Tirupathi*

[b] *Research Superviser, in Dept. of Statistics, S V University, Tirupathi*

[c] *Research Superviser, in Dept. of Statistics, S V University, Tirupathi*

A B S T R A C T

The Chow test, developed by economist Gregory Chow, is a test of statistical significance used to determine whether the coefficients included in two different regression models based on different data sets are comparable. The Chow test is commonly used in econometrics with time series data to determine if there is a structural break in the data. If a structural break is found in the test data, it indicates that the coefficients between the regression lines are not equal. In this study, we can fit one or more regression models and observe that the problem of how to test and confirm if there is a change arises.

Keywords: Chow test, structural change, statistical significance, coefficients, time series data, hypothesis testing, intercept terms, regression coefficients, stability analysis.

The basic idea of the regression model is that the data pattern is dependent and independent. The variables remained the same throughout the data collection period. Based on this assumption, fit a single linear regression model to the data set. Parameters remain the same across time periods and are used for the regression model's estimation, evaluation, and forecasting. When a change in the data pattern is assumed, it may not be appropriate to fit a single linear regression model and more than one regression model may need to be fitted. Before such a result can be adjusted to fit one or more regression models, the problem arises of how to test and determine if there is a change in the structure or pattern of the data. Such a change is characterized by a change in model parameters and is called a structural change.

Now we will study some examples to explore the problem of structural change in the data. Suppose that consumption pattern data is available for many years, and suppose that there is a war between the years for which consumption data is available. Obviously, before and after the consumption process, the war cannot continue because the country's economy is in disarray.

So if the model

$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + - - - + \beta_k x_{ik} + \mu_i, \quad i = 1,2, - - -, n$

If adjusted there, the regression coefficients change between before and after the war period. Such a change is called a structure change or data structure break. In this case, a better option is to fit two independent linear regression models: one for prewar data and one for postwar data. In a further example, suppose the study variable is an individual's income and the explanatory variable is the number of years of schooling. He understands that the objective is to find out if there is a slight discrimination between men's and women's salaries. To find out, you can fit two different regression models: one for male employees and one for female employees. By calculating and comparing the regression constants of the two models, we can verify the reality of gender discrimination in the salaries of male and female employees. The goal here and now is to test for structural change in the data and constant change of the regression coefficients in any way. In other words, we need to test the hypothesis that some or all of the regression coefficients change on different subsets of data.

For now, we study the situation where there is a structural change in the data. In this situation, all the data can be split into two separate data sets. Suppose a data set consisting of 'n' observations is split into two sets, one data set containing n_1 and the other data set containing n_2 observations, i.e. $n_1 + n_2 = n$.

Consider the model

$$y = \alpha\gamma + x\beta + \varepsilon$$

where is γa vector of (nx 1) with unity of all elements, $\alpha$is a scalar representing the intersection term, X is a (nxk) matrix of observations on k explanatory variables, $\beta$is a (kx 1) vector of coefficients of regression y, $\varepsilon$is a (n x 1) vector of disturbances.

Now γdivide x and εinto two subgroups based on n_1 and n_2 observations as follows

$\gamma = \begin{bmatrix}\gamma_1 \\ \gamma_2\end{bmatrix}, \quad x = \begin{bmatrix}x_1 \\ x_2\end{bmatrix}$and $\varepsilon = \begin{bmatrix}\varepsilon_1 \\ \varepsilon_2\end{bmatrix}$

Here $\gamma_1, \gamma_2, x_1, x_2$of orders ($n_1$ x 1), (n_2 x 1), (n_1 xk), (n_2 xk).

and the orders of the disturbances $\varepsilon_1$and ε_2 are (n_1 x 1) and (n_2 x 1), respectively.

Based on these divisions, the two samples correspond to two subgroups

$$y_1 = \alpha\gamma_1 + x_1\beta + \varepsilon_1$$

$$y_2 = \alpha\gamma_2 + x_2\beta + \varepsilon_2$$

In matrix notation, we can write

Model (1): $$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 & x_1 \\ \gamma_2 & x_2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

And name it Model (1).

In this situation, the intercept terms and the regression coefficients are the same for both sub models. Therefore, there is no structural change in this state.

Intercepts and/or regression coefficients between subsamples can illustrate the problem of structural change.

If a structural change occurs due to a change in the interference terms, the situation is characterized by the following model:

$$y_1 = \alpha_1\gamma_1 + x_1\beta + \varepsilon_1$$

$$y_2 = \alpha_2\gamma_2 + x_2\beta + \varepsilon_2$$

In matrix notation, we can write

Model (2): $$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 & 0 & x_1 \\ 0 & \gamma_2 & x_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

If there is a structural change corresponding to different intercept terms and different regression coefficients, then the model:

$$y_1 = \alpha_1\gamma_1 + x_1\beta_1 + \varepsilon_1$$

$$y_2 = \alpha_2\gamma_2 + x_2\beta_2 + \varepsilon_2$$

In matrix notation, we can write

Model (3):

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 & 0 & x_1 & 0 \\ 0 & \gamma_2 & 0 & x_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

The hypothesis test for the structural change test is performed by testing the null value. Hypothesis depending on the situation.

(1) $H_0 : \alpha_1 = \alpha_2$

(2) $H_0 : \beta_1 = \beta_2$

(3) $H_0 : \alpha_1 = \alpha_2, \beta_1 = \beta_2$

To generate the test statistic, apply the ordinary least squares estimate to models (1), (2), and (3) and obtain the residual sum of squares, RSS_1, RSS_2, and RSS_3, respectively.

The degrees of freedom identified with respect to RSS for models (1), (2) and (3) are n-(k+1), n-(k+2) and n-2(k+1), respectively. .

The null hypothesis $H_0 : \alpha_1 = \alpha_2$, that is, different intercept terms, is tested statistically

$$F = \frac{(RSS_1 - RSS_2)/1}{RSS_2/n - k - 2}$$

It follows that $F(1, n - k - 2))$ under H_0. This statistic tests model (2) using model (1) $\alpha_1 = \alpha_2$, that is,

Model (1) contrasts with model (2).

Null hypothesis $H_0 : \beta_1 = \beta_2$, that is, it is tested using different regression coefficients

$$F = \frac{(RSS_2 - RSS_3)/k}{RSS_3/n - 2k - 2}$$

It follows that $F(k, n - 2k - 2)$under H_0. This statistic tests model (3) using model (2) $\beta_1 = \beta_2$, which means that model (2) differs from model (3).

A test of the null hypothesis $H_0 : \alpha_1 = \alpha_2, \beta_1 = \beta_2$, that is, different intercepts and different slope parameters

This can be tested together using test statistics

$$F = \frac{(RSS_1 - RSS_3)/(k+1)}{RSS_3/(n-2k-2)}$$

It follows that $F((k+1), n-2k-2)$ under H_0. This statistic tests model (3) using model (1) $\alpha_1 = \alpha_2$and $\beta_1 = \beta_2$is composite, which means that model (1) is compared to model (3).

This test is called the Chow test. This is necessary $n_1 > k$and $n_2 > k$for the stability of the regression coefficients in the two models. Analysis of the variance test setup is based on the development of this test.

References

Chow, G. C. (1960). "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." Econometrica, 28(3), 591-605.

Maddala, G. S., & Kim, I. M. (1998). "Unit Roots, Cointegration, and Structural Change." Cambridge University Press.

Quandt, R. E. (1960). "Tests of the Hypothesis That a Linear Regression Obeys Two Separate Regimes." Journal of the American Statistical Association, 55(291), 324-330.

Andrews, D. W. (1993). "Tests for Parameter Instability and Structural Change with Unknown Change Point." Econometrica, 61(4), 821-856.

Quandt, R. E. (1958). "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes." Journal of the American Statistical Association, 53(284), 873-880.

Bai, J., & Perron, P. (2003). "Critical Values for Multiple Structural Change Tests." Econometrics Journal, 6(1), 72-78.

Chow, G. C., & Lin, A. L. (1971). "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series." The Review of Economics and Statistics, 53(4), 372-375.

Perron, P. (1989). "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." Econometrica, 57(6), 1361-1401.

Andrews, D. W., & Ploberger, W. (1994). "Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative." Econometrica, 62(6), 1383-1414.

Franses, P. H., & McAleer, M. (1998). "Cointegration Analysis in the Presence of Structural Breaks." Econometric Reviews, 17(3), 251-275.