Chapter 12

Tests for Structural Change and Stability

A fundamental assumption in regression modeling is that the pattern of data on dependent and independent variables remains the same throughout the period over which the data is collected. Under such an assumption, a single linear regression model is fitted over the entire data set. The regression model is estimated and used for prediction assuming that the parameters remain same over the entire time period of estimation and prediction. When it is suspected that there exists a change in the pattern of data, then the fitting of single linear regression model may not be appropriate, and more than one regression models may be required to be fitted. Before taking such a decision to fit a single or more than one regression models, a question arises how to test and decide if there is a change in the structure or pattern of data. Such changes can be characterized by the change in the parameters of the model and are termed as structural change.

Now we consider some examples to understand the problem of structural change in the data. Suppose the data on the consumption pattern is available for several years and suppose there was a war in between the years over which the consumption data is available. Obviously, the consumption pattern before and after the war does not remain the same as the economy of the country gets disturbed. So if a model

$$y_i = \beta_0 + \beta_1 X_{i1} + ... + \beta_k X_{ik} + \varepsilon_i,\ i = 1, 2, ..., n$$

is fitted then the regression coefficients before and after the war period will change. Such a change is referred to as a structural break or structural change in the data. A better option, in this case, would be to fit two different linear regression models- one for the data before the war and another for the data after the war.

In another example, suppose the study variable is the salary of a person, and the explanatory variable is the number of years of schooling. Suppose the objective is to find if there is any discrimination in the salaries of males and females. To know this, two different regression models can be fitted-one for male employees and another for females employees. By calculating and comparing the regression coefficients of both the models, one can check the presence of sex discrimination in the salaries of male and female employees.

Consider another example of structural change. Suppose an experiment is conducted to study certain objectives and data is collected in the USA and India. Then a question arises whether the data sets from both the countries can be pooled together or not. The data sets can be pooled if they originate from the same model in the sense that there is no structural change present in the data. In such case, the presence of structural change in the data can be tested and if there is no change, then both the data sets can be merged and single regression model can be fitted. If structural change is present, then two models are needed to be fitted.

The objective is now how to test for the presence of a structural change in the data and stability of regression coefficients. In other words, we want to test the hypothesis that some of or all the regression coefficients differ in different subsets of data.

Analysis

We consider here a situation where only one structural change is present in the data. The data, in this case, be divided into two parts. Suppose we have a data set of n observations which is divided into two parts consisting of n_1 and n_2 observations such that

$$n_1 + n_2 = n.$$

Consider the model

$$y = \alpha\ell + X\beta + \varepsilon$$

where ℓ is a $(n\times 1)$ vector with all elements unity, α is a scalar denoting the intercept term, X is a $(n\times k)$ matrix of observations on k explanatory variables, β is a $(k\times 1)$ vector of regression coefficients and ε is a $(n\times 1)$ vector of disturbances.

Now partition ℓ, X and ε into two subgroups based on n_1 and n_2 observation as

$$\ell = \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix},\ X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},\ \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

where the orders of ℓ_1 is $(n_1\times 1)$, ℓ_2 is $(n_2\times 1)$, X_1 is $(n_1\times k)$, X_2 is $(n_2\times k)$, ε_1 is $(n_1\times 1)$ and ε_2 is $(n_2\times 1)$.

Based on this partitions, the two models corresponding to two subgroups are

$$y_1 = \alpha\ell_1 + X_1\beta + \varepsilon_1$$
$$y_2 = \alpha\ell_2 + X_2\beta + \varepsilon_2.$$

In matrix notations, we can write

$$\text{Model (1)}: \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \ell_1 & X_1 \\ \ell_2 & X_2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

and term it as Model (1).

In this case, the intercept terms and regression coefficients remain the same for both the submodels. So there is no structural change in this situation.

The problem of structural change can be characterized if intercept terms and/or regression coefficients in the submodels are different.

If the structural change is caused due to change in the intercept terms only then the situation is characterized by the following model:

$$\begin{aligned} y_1 &= \alpha_1 \ell_1 + X_1 \beta + \varepsilon_1 \\ y_2 &= \alpha_2 \ell_2 + X_2 \beta + \varepsilon_2 \end{aligned}$$

or

$$\text{Model (2)}: \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \ell_1 & 0 & X_1 \\ 0 & \ell_2 & X_2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \beta \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$

If the structural change is due to different intercept terms as well as different regression coefficients, then the model is

$$\begin{aligned} y_1 &= \alpha_1 \ell_1 + X_1 \beta_1 + \varepsilon_1 \\ y_2 &= \alpha_2 \ell_2 + X_2 \beta_2 + \varepsilon_2 \end{aligned}$$

or

$$\text{Model (3)}: \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \ell_1 & 0 & X_1 & 0 \\ 0 & \ell_2 & 0 & X_2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$

The test of hypothesis related to the test of structural change is conducted by testing anyone of the null hypothesis depending upon the situation

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

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

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

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

Note that the degrees of freedom associated with

- RSS_1 from the model (1) is $n-(k+1)$.
- RSS_2 from the model (2) is $n-(k+1+1)=n-(k+2)$.
- RSS_3 from the model (3) is $n-(k+1+k+1)=n-2(k+1)$.

The null hypothesis $H_0 : \alpha_1 = \alpha_2$ i.e., different intercept terms is tested by the statistic

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

which follows $F(1, n-k-2)$ under H_0. This statistic tests $\alpha_1 = \alpha_2$ for model (2) using model (1), i.e., model (1) contrasted with model (2).

The null hypothesis $H_0 : \beta_1 = \beta_2$, i.e., different regression coefficients is tested by

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

which follows $F(k, n-2k-2)$ under H_0. This statistic tests $\beta_1 = \beta_2$ from the model (3) using the model (2), i.e., model (2) contrasted with the model (3).

The test of the null hypothesis $H_0 : \alpha_1 = \alpha_2, \beta_1 = \beta_2$, i.e., different intercepts and different slope parameters can be jointly tested by the test statistic

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

which follows $F(k+1, n-2k-2)$ under H_0. This statistic tests jointly $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$ for model (3) using model (1), i.e., model (1) contrasted with model (3). This test is known as **Chow test.** It requires $n_1 > k$ and $n_2 > k$ for the stability of regression coefficients in the two models. The development of this test is as follows which is based on the set up of analysis of variance test.

Development of Chow test:

Consider the models

$$\underset{n_1\times 1}{y_1} = \underset{n_1\times p}{X_1}\,\underset{p\times 1}{\beta_1} + \underset{n_1\times 1}{\varepsilon_1} \qquad (i)$$

$$\underset{n_2\times 1}{y_2} = \underset{n_2\times p}{X_2}\,\underset{p\times 1}{\beta_2} + \underset{n_2\times 1}{\varepsilon_2} \qquad (ii)$$

$$\underset{n\times 1}{y} = \underset{n\times p}{X}\,\underset{p\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad (iii)$$

$$n = n_1 + n_2$$

where $p = k+1$ which includes k explanatory variables and an intercept term.

Define

$$\bar{H}_1 = I_1 - X_1\left(X_1'X_1\right)^{-1}X_1'$$

$$\bar{H}_2 = I_2 - X_2\left(X_2'X_2\right)^{-1}X_2'$$

$$\bar{H} = I - X\left(X'X\right)^{-1}X'$$

where I_1 and I_2 are the identity matrices of the orders $\left(n_1 \times n_1\right)$ and $\left(n_2 \times n_2\right)$. The residual sums of squares based on models (i), (ii) and (iii) are obtained as

$$RSS_1 = \varepsilon_1'\bar{H}_1\varepsilon_1$$

$$RSS_2 = \varepsilon_2'\bar{H}_2\varepsilon_2$$

$$RSS = \varepsilon'\bar{H}\varepsilon.$$

Then define

$$\bar{H}_1^* = \begin{bmatrix} \bar{H}_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \bar{H}_2^* = \begin{bmatrix} 0 & 0 \\ 0 & \bar{H}_2 \end{bmatrix}.$$

Both $\bar{H}_1^*$ and $\bar{H}_2^*$ are $\left(n \times n\right)$ matrices. Now the RSS_1 and RSS_2 can be re-expressed as

$$RSS_1 = \varepsilon'\bar{H}_1^*\varepsilon$$

$$RSS_2 = \varepsilon'\bar{H}_2^*\varepsilon$$

where ε is the disturbance term related to model (iii) based on $\left(n_1 + n_2\right)$ observations.

Note that $\bar{H}_1^*\bar{H}_2^* = 0$ which implies that RSS_1 and RSS_2 are independently distributed.

We can write

$$\begin{aligned}\bar{H} &= I - X\left(X'X\right)^{-1}X' \\ &= \begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix} - \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}\left(X'X\right)^{-1}\left(X_1' \quad X_2'\right) \\ &= \begin{pmatrix} \bar{H}_{11} & \bar{H}_{12} \\ \bar{H}_{21} & \bar{H}_{22} \end{pmatrix}\end{aligned}$$

where $\bar{H}_{11} = I_1 - X_1\left(X'X\right)^{-1}X_1'$

$$\begin{aligned}\bar{H}_{12} &= -X_1\left(X'X\right)^{-1}X_2' \\ \bar{H}_{21} &= -X_2\left(X'X\right)^{-1}X_1' \\ \bar{H}_{22} &= I_2 - X_2\left(X'X\right)^{-1}X_2'.\end{aligned}$$

Define

$$\bar{H}^* = \bar{H}_1^* + \bar{H}_2^*$$

so that

$$\begin{aligned}RSS_3 &= \varepsilon'\bar{H}^*\varepsilon \\ RSS_1 &= \varepsilon'\bar{H}\varepsilon \quad .\end{aligned}$$

Note that $\left(\bar{H} - \bar{H}^*\right)$ and $\bar{H}*$ are idempotent matrices. Also $\left(\bar{H} - \bar{H}^*\right)\bar{H}^* = 0.$ First, we see how this result holds.

Consider

$$\begin{aligned}\left(\bar{H} - \bar{H}_1^*\right)\bar{H}_1^* &= \begin{pmatrix} \bar{H}_{11} - \bar{H}_1 & \bar{H}_{12} \\ \bar{H}_{21} & \bar{H}_{22} \end{pmatrix}\begin{pmatrix} \bar{H}_1 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} (\bar{H}_{11} - \bar{H}_1)\bar{H}_1 & 0 \\ \bar{H}_{21}\bar{H}_1 & 0 \end{pmatrix}.\end{aligned}$$

Since $X_1'\bar{H}_1 = 0,$ so we have

$$\begin{aligned}\bar{H}_{21}\bar{H}_1 &= 0 \\ \bar{H}_{11}\bar{H}_1 &= 0.\end{aligned}$$

Also, since $\bar{H}_1$ is idempotent, it follows that

$$\left(\bar{H}_{11}-\bar{H}_1\right)\bar{H}_1=0.$$

Thus $\left(\bar{H}-\bar{H}_1^*\right)\bar{H}_1^*=0$

or $\quad \bar{H}\bar{H}_1^*=\bar{H}_1^*.$

Similarly, it can be shown that

$$\left(\bar{H}-\bar{H}_2^*\right)\bar{H}_2^*=0$$

or $\quad \bar{H}\bar{H}_2^*=\bar{H}_2^*.$

Thus this implies that

$$\left(\bar{H}-\bar{H}_1^*-\bar{H}_2^*\right)\left(\bar{H}_1^*+\bar{H}_2^*\right)=0$$

or $\quad \left(\bar{H}-\bar{H}^*\right)\bar{H}^*=0.$

Also, we have

$$\begin{aligned}
tr\ \bar{H} &= n-p \\
tr\ \bar{H}^* &= tr\,\bar{H}_1^*+tr\,\bar{H}_2^* \\
&= n_1-p+n_2-p \\
&= n-2p \\
&= n-2k-2.
\end{aligned}$$

Hence RSS_1 and RSS_3 are independently distributed. Further

$$\frac{RSS_1-RSS_3}{\sigma^2}\sim\chi_p^2,$$

$$\frac{RSS_3}{\sigma^2}\sim\chi_{(n-2p)}^2.$$

Also, $\left(\dfrac{RSS_1-RSS_3}{\sigma^2}\right)$ and $\dfrac{RSS_3}{\sigma^2}$ are independently distributed. Hence under the null hypothesis

$$\frac{\left(\dfrac{RSS_1-RSS_3}{\sigma^2}\right)\Big/p}{\left(\dfrac{RSS_3}{\sigma^2}\right)\Big/(n-2p)}\sim F\left(p,n-2p\right)$$

or $\quad \dfrac{\left(RSS_1-RSS_3\right)/\left(k+1\right)}{RSS_3/\left(n-2k-2\right)}\sim F\left(k+1,n-2k-2\right).$

Limitations of these tests

1. All tests are based under the assumption that σ^2 remains the same. So first the stability of σ^2 should be checked and then these tests can be used.
2. It is assumed in these tests that the point of change is exactly known. In practice, it is difficult to find such a point at which the change occurs. It is more difficult to know such point when the change occurs slowly. These tests are not applicable when the point of change is unknown. An ad-hoc technique when the point of change is unknown is to delete the data of transition period.
3. When there are more than one points of structural change, then the analysis becomes difficult.