LINEAR REGRESSION ANALYSIS

MODULE – III

Lecture – 13

Multiple Linear Regression Analysis

Dr. Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Likelihood ratio test for $H_0 : R\beta = r$

The same logic and reasons used in the development of likelihood ratio test for $H_0 : \beta = \beta_0$ can be extended to develop the likelihood ratio test for $H_0 : R\beta = r$ as follows:

$$\Omega = \left\{(\beta, \sigma^2) : -\infty < \beta_i < \infty,\ \sigma^2 > 0,\ i = 1, 2, ..., k\right\}$$

$$\omega = \left\{(\beta, \sigma^2) : -\infty < \beta_i < \infty,\ R\beta = r,\ \sigma^2 > 0\right\}.$$

Let $\tilde{\beta} = (X'X)^{-1}X'y.$

Then

$$E(R\tilde{\beta}) = R\beta$$

$$V(R\tilde{\beta}) = E\left[R(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'R'\right]$$

$$= RV(\tilde{\beta})R'$$

$$= \sigma^2\ R(X'X)^{-1}R'.$$

Since $\tilde{\beta} \sim N\left[\beta, \sigma^2(X'X)^{-1}\right]$

so $R\tilde{\beta} \sim N\left[R\beta, \sigma^2 R(X'X)^{-1}R'\right]$

$$R\tilde{\beta} - r = R\tilde{\beta} - R\beta = R(\tilde{\beta} - \beta) \sim N\left[0, \sigma^2 R(X'X)^{-1}R'\right].$$

There exists a matrix Q such that

$$\left[R(X'X)^{-1}R'\right]^{-1} = QQ'$$

and then

$$\xi = QR(b-\beta) \sim N(0, \sigma^2 I_n).$$

Therefore under $H_0 : R\beta - r = 0,$

$$\begin{aligned}
\frac{\xi\xi'}{\sigma^2} &= \frac{(R\tilde{\beta}-r)'QQ'(R\tilde{\beta}-r)}{\sigma^2} \\
&= \frac{(R\tilde{\beta}-r)'\left[R(X'X)^{-1}R'\right]^{-1}(R\tilde{\beta}-r)}{\sigma^2} \\
&= \frac{(\tilde{\beta}-\beta)'R'\left[R(X'X)^{-1}R'\right]^{-1}R(\tilde{\beta}-\beta)}{\sigma^2} \\
&= \frac{\varepsilon'X(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R(X'X)^{-1}X'\varepsilon}{\sigma^2} \\
&\sim \chi^2(J)
\end{aligned}$$

which is obtained as $X(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R(X'X)^{-1}X'$ is an idempotent matrix and its trace is J which is the associated degrees of freedom with the respective quadratic form.

Also, irrespective of whether H_0 is true or not,

$$\frac{\tilde{e}'\tilde{e}}{\sigma^2} = \frac{(y-X\tilde{\beta})'(y-X\tilde{\beta})}{\sigma^2} = \frac{y'\bar{H}y}{\sigma^2} = \frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-k).$$

Moreover, the product of quadratic form matrices of $\tilde{e}'\tilde{e}$ and $(\tilde{\beta}-\beta)'R'\left[R(X'X)^{-1}R'\right]^{-1}R(\tilde{\beta}-\beta)$ is zero implying that both the quadratic forms are independent.

So in terms of likelihood ratio test statistic

$$\lambda_1 = \frac{\left(\dfrac{(R\tilde{\beta}-r)'\left[R(X'X)^{-1}R'\right]^{-1}(R\tilde{\beta}-r)}{\dfrac{\sigma^2}{J}}\right)}{\left(\dfrac{(n-k)\hat{\sigma}^2}{\sigma^2}\right)}$$

$$= \frac{\left(R\tilde{\beta}-r)'\left[R(X'X)^{-1}R'\right]^{-1}\left(R\tilde{\beta}-r\right)\right)}{J\hat{\sigma}^2}$$

$\sim F(J, n-k)$ under H_0.

So the decision rule is to reject H_0 whenever

$$\lambda_1 \geq F_\alpha(J, n-k)$$

where $F_\alpha(J, n-k)$ is the upper critical points on the central *F* distribution with *J* and (*n* – *k*) degrees of freedom.

Test of significance of regression (Analysis of variance)

If we set $R = [0 \quad I_{k-1}]$, $r = 0$, then the hypothesis $H_0 : R\beta = r$ reduces to the following null hypothesis:

$$H_0 : \beta_2 = \beta_3 = ... = \beta_k = 0$$

which is tested against the alternative hypothesis

$H_1 : \beta_j \neq 0$ for at least one *j = 2, 3,…, k*.

This hypothesis determines if there is a linear relationship between *y* and any set of the explanatory variables $X_2, X_3, ..., X_k$. Notice that X_1 corresponds to the intercept term in the model and hence $x_{i1} = 1$ for all $i = 1, 2, ..., n$.

This is an **overall** or **global test of model adequacy**. Rejection of the null hypothesis indicates that at least one of the explanatory variables among $X_2, X_3, ..., X_k$ contributes significantly to the model. This is called as **analysis of variance.**

Since $\varepsilon \sim N(0, \sigma^2 I)$,

so $y \sim N(X\beta, \sigma^2 I)$

$$b = (X'X)^{-1}X'y \sim N\left[\beta, \sigma^2 (X'X)^{-1}\right].$$

Also $\hat{\sigma}^2 = \dfrac{SS_{res}}{n-k}$

$$= \frac{(y-\hat{y})'(y-\hat{y})}{n-k}$$

$$= \frac{y'\left[I - X(X'X)^{-1}X'\right]y}{n-k} = \frac{y'\bar{H}y}{n-k} = \frac{y'y - b'X'y}{n-k}.$$

Since $(X'X)^{-1}X'\bar{H} = 0$,

so *b* and $\hat{\sigma}^2$ are independently distributed.

Since $y'\bar{H}y = \varepsilon'\bar{H}\varepsilon$ and $\bar{H}$ is an idempotent matrix, so

$$\frac{SS_{res}}{\sigma^2} \sim \chi^2_{(n-k)},$$

i.e., central χ^2 distribution with (n - k) degrees of freedom.

Partition $X = [X_1, \quad X_2^*]$ where the submatrix X_2^* contains the explanatory variables $X_2, X_3, ..., X_k$ and partition $\beta = [\beta_1, \quad \beta_2^*]$ where the subvector β_2^* contains the regression coefficients $\beta_2, \beta_3, ..., \beta_k$.

Now partition the total sum of squares due to *y's* as

$$SS_T = y'Ay$$
$$= SS_{reg} + SS_{res}$$

where

$$SS_{reg} = b_2^{*\prime} X_2^{*\prime} A X_2^* b_2^*$$

is the **sum of squares due to regression** and b_2^* is the OLSE of β_2^*.

The **sum of squares due to residuals** is given by

$$SS_{res} = (y - Xb)'(y - Xb)$$
$$= y'\bar{H}y$$
$$= SS_T - SS_{reg}.$$

Further

$\frac{SS_{reg}}{\sigma^2} \sim \chi^2_{k-1}\left(\frac{\beta_2^{*\prime} X_2^{*\prime} A X_2^* \beta_2^*}{2\sigma^2}\right)$, i.e., non-central χ^2 distribution with non-centrality parameter $\frac{\beta_2^{*\prime} X_2^{*\prime} A X_2^* \beta_2^*}{2\sigma^2}$,

$\frac{SS_T}{\sigma^2} \sim \chi^2_{n-1}\left(\frac{\beta_2^{*\prime} X_2^{*\prime} A X_2^* \beta_2^*}{2\sigma^2}\right)$, i.e., non-central χ^2 distribution with non-centrality parameter $\frac{\beta_2^{*\prime} X_2^{*\prime} A X_2^* \beta_2^*}{2\sigma^2}$.

Since $X_2 \bar{H} = 0$, so SS_{reg} and SS_{res} are independently distributed. The mean squares due to regression is

$$MS_{reg} = \frac{SS_{reg}}{k-1}$$

and the mean square due to error is

$$MS_{res} = \frac{SS_{res}}{n-k}.$$

Then

$$\frac{MS_{reg}}{MS_{res}} \sim F_{k-1,n-k}\left(\frac{\beta_2^{*\prime} X_2^{*\prime} A X_2^* \beta_2^*}{2\sigma^2}\right)$$

which is a non-central *F* -distribution with $(k-1)(n-k)$ degrees of freedom and non-centrality parameter $\frac{\beta_2^{*\prime} X_2^{*\prime} A X_2^* \beta_2^*}{2\sigma^2}$.

Under $H_0 : \beta_2 = \beta_3 = \ldots = \beta_k = 0,$

$$F = \frac{MS_{reg}}{MS_{res}} \sim F_{k-1,n-k}.$$

The decision rule is to reject at α level of significance whenever

$$F \geq F_\alpha (k-1, n-k).$$

The calculation of *F*-statistic can be summarized in the form of an analysis of variance (ANOVA) table given as follows:

Source of variation	**Sum of squares**	**Degrees of freedom**	**Mean squares**	F
Regression	SS_{reg}	$k-1$	$MS_{reg} = SS_{reg} / k-1$	F
Error	SS_{res}	$n-k$	$MS_{res} = SS_{res} /(n-k)$	
Total	SS_T	$n-1$		

Rejection of H_0 indicates that it is likely that atleast one $\beta_i \neq 0 \quad (i = 1, 2, ..., k)$.

Test of hypothesis on individual regression coefficients

In case the test in analysis of variance is rejected, then another question arises is that which of the regression coefficients is/are responsible for the rejection of null hypothesis. The explanatory variables corresponding to such regression coefficients are important for the model.

Adding such explanatory variables also increases the variance of fitted values $\hat{y}$, so one needs to be cautious that only those regressors are added which are really important in explaining the response. Adding unimportant explanatory variables may increase the residual mean square which may decrease the usefulness of the model.

To test the null hypothesis $H_0 : \beta_j = 0$

versus the alternative hypothesis $H_1 : \beta_j \neq 0$

has already been discussed is the case of simple linear regression model. In present case, if H_0 is accepted, it implies that the explanatory variable X_j can be deleted from the model. The corresponding test statistic is

$$t = \frac{b_j}{se(b_j)} \sim t(n-k-1) \text{ under } H_0$$

where the standard error of OLSE b_j of β_j is

$se(b_j) = \sqrt{\hat{\sigma}^2 C_{jj}}$ where C_{jj} denotes the j^{th} diagonal element of $(X'X)^{-1}$ corresponding to b_j.

The decision rule is to reject H_0 at α level of significance if $|t| > t_{\frac{\alpha}{2}, n-k-1}$.

Note that this is only a **partial or marginal test** because b_j depends on all the other explanatory variables $X_i (i \neq j)$ that are in the model. This is a test of the contribution of X_j given the other explanatory variables in the model.