SF 2930 Regression Analysis
Lecture 6
Multiple Linear Regression, Part 4 Hypothesis Testing for β in Ordinary Normal Multiple Regression, when the Null Hypothesis does not involve The Intercept

Timo Koski

KTH Royal Institute of Technology

2023

Learning outcomes

- The Centered Model
- Equivalence of the LSE in the Centered Model and the OLSE
- New Identities for SS_R
- Confidence Ellipsoid

FROM MVP P. 84

3.3.1 Test for Significance of Regression

The test for **significance of regression** is a test to determine if there is a **linear relationship** between the response y and any of the regressor variables $x_1, x_2, \dots, x_k$. This procedure is often thought of as an overall or global test of model adequacy. The appropriate hypotheses are

$$H_0: \beta_1 = \beta_1 = \cdots = \beta_k = 0$$

$$H_1: \beta_j \neq 0 \quad \text{for at least one } j$$

Rejection of this null hypothesis implies that at least one of the regressors $x_1, x_2, \dots, x_k$ contributes significantly to the model.

The test procedure is a generalization of the **analysis of variance** used in simple linear regression. The **total sum of squares** SS_T is partitioned into a **sum of squares**

FROM MVP P. 84

The authors of MVP do not explain, why β_0 is not involved in the null hypothesis.

From another source: Since β_0 is usually not zero, we would rarely be interested in including $\beta_0 = 0$ in the hypothesis. Rejection of H_o: $\boldsymbol{\beta} = \mathbf{0}_{k+1}$ might be due solely to β_0, and we would not learn whether the x variables predict y.

Hypotheses on the Regression coefficients

$$H_o\text{: } \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

$$H_1\text{: } \beta_j \neq 0 \textit{ for at least one } \beta_j$$

The null hypothesis says that there is no useful linear relationship between Y and any of the k predictors. If at least one of these β_j's is $\neq 0$, the model is deemed useful.
We could test each β_j separately (see the preceding lecture 5), but that would take time and be very conservative (if Bonferroni correction is used). A better test is a joint test, and is based on a statistic that has an F distribution when H_o is true.

Hypotheses on the Regression coefficients

$$H_o\text{: } \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

$$H_1\text{: } \beta_j \neq 0 \textit{ for at least one } \beta_j$$

We must find LSE of $\beta_1, \beta_2, \ldots, \beta_k$ without involving β_0. Then we must find a decomposition of variance to find the quadratic forms of an F-test without the presence of β_o in the F-statistic. The expressions from Lecture 5 must be modified. For this we need partitioned matrices and the centered model.

THE CENTERED MODEL

MVP pp. 84-86

Centered Multiple Linear Regression Model

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1, \ldots, n, \quad n > k + 1.$$

The centered model is

$$Y_i = \alpha + \beta_1 (x_{i1} - \bar{x}_1) + \cdots + \beta_k (x_{ik} - \bar{x}_k) + \varepsilon_i, \quad i = 1, \ldots, n \tag{1}$$

where $\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$, $j = 1, \ldots, k$, and

$$\alpha = \beta_0 + \beta_1 \bar{x}_1 + \cdots + \beta_k \bar{x}_k. \tag{2}$$

Centered Multiple Linear Regression Model

Introduce

$$X_R := \begin{pmatrix} x_{11} & \cdots & x_{1k} \\ x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{pmatrix}$$

so that the design matrix is partitioned as

$$X = \begin{pmatrix} \mathbf{I}_n & X_R \end{pmatrix}$$

The matrix of centered regressor/covariate values

We recall the centering matrix C_{ce} and compute

$$C_{ce}X_R = \left(X_R - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T X_R \right),$$

where we compute the $1 \times k$ matrix $\mathbf{1}_n^T X_R$ as

$$\mathbf{1}_n^T X_R = (1, 1, \ldots, 1) \begin{pmatrix} x_{11} & \cdots & x_{1k} \\ x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{pmatrix} = \left(\sum_{i=1}^{n} x_{i1}, \sum_{i=1}^{n} x_{i2}, \ldots, \sum_{i=1}^{n} x_{ik} \right)$$

The matrix of centered regressor/covariate values

Thus we get the $n \times k$ matrix

$$\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T X_R = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \left(\frac{1}{n}\sum_{i=1}^{n} x_{i1}, \frac{1}{n}\sum_{i=1}^{n} x_{i2}, \ldots, \frac{1}{n}\sum_{i=1}^{n} x_{ik} \right)$$

$$\& \tag{3}$$

$$\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T X_R = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k) = \begin{pmatrix} \bar{x}_1 & \cdots & \bar{x}_k \\ \bar{x}_1 & \cdots & \bar{x}_k \\ \vdots & \vdots & \vdots \\ \bar{x}_1 & \cdots & \bar{x}_k \end{pmatrix}$$

The matrix of centered regressor/covariate values matrix of centered regressor/covariate values

Hence

$$C_{ce}X_R = \left(X_R - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T X_R\right) = \begin{pmatrix} x_{11}-\bar{x}_1 & \cdots & x_{1k}-\bar{x}_k \\ x_{21}-\bar{x}_1 & \cdots & x_{2k}-\bar{x}_k \\ \vdots & \vdots & \vdots \\ x_{n1}-\bar{x}_1 & \cdots & x_{nk}-\bar{x}_k \end{pmatrix}$$

We set

$$X_c := \begin{pmatrix} x_{11}-\bar{x}_1 & \cdots & x_{1k}-\bar{x}_k \\ x_{21}-\bar{x}_1 & \cdots & x_{2k}-\bar{x}_k \\ \vdots & \vdots & \vdots \\ x_{n1}-\bar{x}_1 & \cdots & x_{nk}-\bar{x}_k \end{pmatrix} \tag{4}$$

This is the matrix of centered regressor/covariate values.

The centered model in a matrix form

Now we can write the equations in (1) in matrix form as

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = (\mathbf{1}_n, X_c) \begin{pmatrix} \alpha \\ \boldsymbol{\beta}_R \end{pmatrix} + \boldsymbol{\varepsilon}. \tag{5}$$

where

$$\boldsymbol{\beta}_R = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix},$$

and X_c is given in (4).

EXQ LSE FOR THE CENTERED MODEL

- Show that the normal equations (c.f. Lecture 3) for (5) are

$$\begin{pmatrix} n & \mathbf{0}_{\mathbf{k}}^{\mathbf{T}} \\ \mathbf{0}_{\mathbf{k}} & X_c^T X_c \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_R^c \end{pmatrix} = \begin{pmatrix} n\bar{y} \\ X_c^T \mathbf{y} \end{pmatrix} \tag{6}$$

Continued on the next slide →

EXQ LSE FOR THE CENTERED MODEL

- Why is $X_c^T X_c$ of full rank? Check formally that

$$\begin{pmatrix} n & \mathbf{0}_\mathbf{k}^\mathbf{T} \\ \mathbf{0}_\mathbf{k} & X_c^T X_c \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{n} & \mathbf{0}_\mathbf{k}^\mathbf{T} \\ \mathbf{0}_\mathbf{k} & \left(X_c^T X_c\right)^{-1} \end{pmatrix} \tag{7}$$

END of EXQ. □

Part 2: Equivalence

Part 2: Equivalence

We check that

$$\hat{\alpha} = \bar{y}$$

$$\hat{\boldsymbol{\beta}}_R^c = \left(X_c^T X_c\right)^{-1} X_c^T \mathbf{y} \tag{8}$$

and

$$\hat{\boldsymbol{\beta}} = \left(X^T X\right)^{-1} X^T \mathbf{y}$$

are the same, that is

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\boldsymbol{\beta}}_R^c \end{pmatrix}$$

$\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

The normal equations are

$$X^T X \hat{\boldsymbol{\beta}} = X^T \mathbf{y} \tag{9}$$

Partitioning of the Multiple Regression Eq's

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n & X_R \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \boldsymbol{\beta}_r \end{pmatrix}$$

where

$$\mathbf{1}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad X_R = \begin{pmatrix} x_{11} & \cdots & x_{1k} \\ x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{pmatrix}, \quad \boldsymbol{\beta}_R = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \tag{10}$$

Partitioning of the Multiple Regression Model

Then we can write the normal equations (9) as

$$\begin{pmatrix} \mathbf{1}_n & X_R \end{pmatrix}^T \begin{pmatrix} \mathbf{1}_n & X_R \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\boldsymbol{\beta}}_R \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n & X_R \end{pmatrix}^T \mathbf{y} \tag{11}$$

By multipication of partitioned matrices (see Appendix)

$$\begin{pmatrix} \mathbf{1}_n^T \mathbf{1}_n & \mathbf{1}_n^T X_R \\ X_R^T \mathbf{1}_n & X_R^T X_R \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\boldsymbol{\beta}}_R \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n \mathbf{y} \\ X_R^T \mathbf{y} \end{pmatrix}. \tag{12}$$

We verify that this in fact gives $\hat{\alpha}$ and that $\hat{\boldsymbol{\beta}}_R$ given here will satisfy (6), or, $\hat{\boldsymbol{\beta}}_R^C = \hat{\boldsymbol{\beta}}_R$. We obtain from (12)

$$n\widehat{\beta}_0 + \mathbf{1}_n^T X_R \widehat{\boldsymbol{\beta}}_R = n\bar{y} \tag{13}$$

and

$$X_R^T \mathbf{1}_n \widehat{\beta}_0 + X_R^T X_R \widehat{\boldsymbol{\beta}}_R = X_R^T \mathbf{y}. \tag{14}$$

Partitioning of the Multiple Regression Model

In (13) we get

$$\widehat{\beta}_0 + \frac{1}{n}\mathbf{1}_n^T X_R \widehat{\boldsymbol{\beta}}_R = \bar{y}$$

Here (see Lecture 2., Part 0)

$$\frac{1}{n}\mathbf{1}_n^T X_R = \frac{1}{n}(1 1 \ldots 1)\begin{pmatrix} x_{11} & \cdots & x_{1k} \\ x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{pmatrix} = \left(\frac{1}{n}\sum_{j=1}^{n} x_{j1} \ldots \frac{1}{n}\sum_{j=1}^{n} x_{jk}\right)$$

$$= (\bar{x}_1 \ldots \bar{x}_k) =: \bar{\mathbf{x}}^T.$$

Here $\bar{\mathbf{x}}$ has the means of the columns of X_R as components. Hence

$$\widehat{\beta}_0 + \bar{\mathbf{x}}^T \widehat{\boldsymbol{\beta}}_R = \bar{y} \tag{15}$$

Partitioning of the Multiple Regression Model

Recall $\alpha = \beta_0 + \beta_1 \bar{x}_1 + \cdots + \beta_k \bar{x}_k$ in (2). Hence

$$\widehat{\beta}_0 + \bar{\mathbf{x}}^T \widehat{\boldsymbol{\beta}}_R = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}_1 + \cdots + \hat{\beta}_k \bar{x}_k = \hat{\alpha}. \tag{16}$$

Hence we have obtained that $\hat{\alpha} = \bar{y}$.

Partitioning of the Multiple Regression Model

We study now (14), i.e.,

$$X_R^T \mathbf{1}_n \widehat{\beta}_0 + X_R^T X_R \widehat{\boldsymbol{\beta}}_R = X_R^T \mathbf{y}$$

We have already observed that $\mathbf{1}_n^T X_R = n\bar{\mathbf{x}}^T$. Hence

$$n\bar{\mathbf{x}}\widehat{\beta}_0 + X_R^T X_R \widehat{\boldsymbol{\beta}}_R = X_R^T \mathbf{y} \tag{17}$$

We find (check details)

$$X_c^T X_c = X_R^T X_R - n\bar{\mathbf{x}}\bar{\mathbf{x}}^T$$

i.e.

$$X_R^T X_R = X_c^T X_c + n\bar{\mathbf{x}}\bar{\mathbf{x}}^T \tag{18}$$

Partitioning of the Multiple Regression Model

Furthermore

$$X_R^T \mathbf{y} = X_c^T \mathbf{y} + n\bar{\mathbf{x}}\bar{y} \tag{19}$$

By (18)

$$X_R^T X_R \widehat{\boldsymbol{\beta}}_R = X_c^T X_c \widehat{\boldsymbol{\beta}}_R + n\bar{\mathbf{x}}\bar{\mathbf{x}}^T \widehat{\boldsymbol{\beta}}_R$$

But by the above $\bar{\mathbf{x}}^T \widehat{\boldsymbol{\beta}}_R = \hat{\alpha} - \widehat{\beta}_0$, i.e.,

$$X_R^T X_R \widehat{\boldsymbol{\beta}}_R = X_c^T X_c \widehat{\boldsymbol{\beta}}_R + n\bar{\mathbf{x}} \left(\hat{\alpha} - \widehat{\beta}_0 \right). \tag{20}$$

We insert (20) in the the LHS (left hand side) of (17), which becomes

$$\text{LHS}: \quad n\bar{\mathbf{x}}\widehat{\beta}_0 + X_R^T X_R \widehat{\boldsymbol{\beta}}_R = n\bar{\mathbf{x}}\widehat{\beta}_0 + X_c^T X_c \widehat{\boldsymbol{\beta}}_R + n\bar{\mathbf{x}} \left(\hat{\alpha} - \widehat{\beta}_0 \right)$$

i.e.,

$$\text{LHS}: \quad X_c^T X_c \widehat{\boldsymbol{\beta}}_R + n\bar{\mathbf{x}}\hat{\alpha}$$

$$\text{LHS}: \quad X_c^T X_c \widehat{\boldsymbol{\beta}}_R + n\bar{\mathbf{x}}\hat{\alpha} \tag{21}$$

From (19) RHS (right hand side) of (17) is

$$\text{RHS}: \quad X_R^T \mathbf{y} = X_c^T \mathbf{y} + n\bar{\mathbf{x}}\bar{y} \tag{22}$$

Since we have found $\hat{\alpha} = \bar{y}$, as LHS= RHS in (17), (21) and (22) give

$$X_c^T X_c \widehat{\boldsymbol{\beta}}_R = X_c^T \mathbf{y} \tag{23}$$

But $\hat{\alpha} = \bar{y}$ and (23) can be written in matrix form as

$$\begin{pmatrix} n & \mathbf{0}_\mathbf{k}^\mathbf{T} \\ \mathbf{0}_\mathbf{k} & X_c^T X_c \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\boldsymbol{\beta}}_R \end{pmatrix} = \begin{pmatrix} n\bar{y} \\ X_c^T \mathbf{y} \end{pmatrix} \tag{24}$$

which is (6). By (7), $\hat{\boldsymbol{\beta}}_R = \hat{\boldsymbol{\beta}}_R^C$.

Fundamental Variance Identity Without β_0

Fundamental Variance Identity Without β_0

In Appendix D of Lecture 3 (see slides) we showed that

$$SS_{\mathrm{Res}} = Q\left(\widehat{\boldsymbol{\beta}}\right) = \mathbf{y}^T\mathbf{y} - \widehat{\boldsymbol{\beta}}^T X^T \mathbf{y} \tag{25}$$

We shall next use the partitioned equations to rewrite this. The first step is to write with the notations in (10)

$$\widehat{\boldsymbol{\beta}}^T X^T \mathbf{y} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\boldsymbol{\beta}}_R \end{pmatrix}^T (\mathbf{1}_n, X_R)^T \mathbf{y} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\boldsymbol{\beta}}_R \end{pmatrix}^T \begin{pmatrix} \mathbf{1}_n^T \\ X_R^T \end{pmatrix} \mathbf{y} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\boldsymbol{\beta}}_R \end{pmatrix}^T \begin{pmatrix} \mathbf{1}_n^T \mathbf{y} \\ X_R^T \mathbf{y} \end{pmatrix}$$

$$= n\bar{y}\widehat{\beta}_0 + \widehat{\boldsymbol{\beta}}_R^T X_R^T \mathbf{y} = n\bar{y}\left(\bar{y} - \bar{\mathbf{x}}^T \widehat{\boldsymbol{\beta}}_R\right) + \widehat{\boldsymbol{\beta}}_R^T X_R^T \mathbf{y}$$

where we used (15). (continued on the next slide $\rightarrow$)

Fundamental Variance Identity Without β_0

We have found up to this point that

$$\widehat{\boldsymbol{\beta}}^T X^T \mathbf{y} = n\bar{y}\left(\bar{y} - \bar{\mathbf{x}}^T \widehat{\boldsymbol{\beta}}_R\right) + \widehat{\boldsymbol{\beta}}_R^T X_R^T \mathbf{y}$$

We continue to trim the right hand side as

$$= n\bar{y}^2 - n\bar{y}\widehat{\boldsymbol{\beta}}_R^T \bar{\mathbf{x}} + \widehat{\boldsymbol{\beta}}_R^T X_R^T \mathbf{y} = n\bar{y}^2 + \widehat{\boldsymbol{\beta}}_R^T \left(X_R^T \mathbf{y} - n\bar{y}\bar{\mathbf{x}}\right)$$

$$= n\bar{y}^2 + \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y},$$

where used (19) above, that is, $X_R^T \mathbf{y} = X_c^T \mathbf{y} + n\bar{\mathbf{x}}\bar{y}$. Now we return to (25)

Fundamental Variance Identity Without β_0

The final result above is

$$\widehat{\boldsymbol{\beta}}^T X^T \mathbf{y} = n\bar{y}^2 + \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y}.$$

Hence we get in (25)

$$SS_{\mathrm{Res}} = \mathbf{y}^T\mathbf{y} - \widehat{\boldsymbol{\beta}}^T X^T \mathbf{y} = \mathbf{y}^T\mathbf{y} - n\bar{y}^2 - \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y}.$$

We note by scalar product and an Appendix to Lecture 1, slides, that

$$\mathbf{y}^T\mathbf{y} - n\bar{y}^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2 = \sum_{i=1}^n (y_i - \bar{y})^2 .$$

Hence we have found

$$\boxed{SS_{\mathrm{Res}} = \sum_{i=1}^n (y_i - \bar{y})^2 - \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y} \qquad (26)}$$

Fundamental Variance Identity Without β_0

$$SS_{\mathrm{Res}} = \sum_{i=1}^{n} (y_i - \bar{y})^2 - \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y}$$

A phantastic step (if you mind)

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \widehat{\boldsymbol{\beta}}_R^T X_c^T + \left(\sum_{i=1}^{n} (y_i - \bar{y})^2 - \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y} \right) = \widehat{\boldsymbol{\beta}}_R^T X_c^T + SS_{\mathrm{Res}}.$$

Or,

$$SS_{\mathrm{T}} = \widehat{\boldsymbol{\beta}}_R^T X_c^T + SS_{\mathrm{Res}} \tag{27}$$

This means that the regression or model sum of squares SS_{R} is identified as

$$SS_{\mathrm{R}} = \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y} \tag{28}$$

Fundamental Variance Identity Without β_0

In view of (8), we have $\hat{\boldsymbol{\beta}}_R = \left(X_c^T X_c\right)^{-1} X_c^T \mathbf{y}$. Hence

$$\widehat{\boldsymbol{\beta}}_R^T X_c^T X_c \widehat{\boldsymbol{\beta}}_R = \widehat{\boldsymbol{\beta}}_R^T \underbrace{X_c^T X_c \left(X_c^T X_c\right)^{-1}}_{=\mathbb{I}_k} X_c^T \mathbf{y} = \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y}$$

Hende the regression or model sum of squares from (28) is

$$\boxed{SS_{\mathrm{R}} = \widehat{\boldsymbol{\beta}}_R^T X_c^T X_c \widehat{\boldsymbol{\beta}}_R \qquad (29)}$$

The advantage of (29) is that SS_{R} is a quadratic form in the normal r.v. $\widehat{\boldsymbol{\beta}}_R$.

Another Round of Phantastic Identities

$$SS_{\mathrm{R}} = \widehat{\boldsymbol{\beta}}_R^T X_c^T X_c \widehat{\boldsymbol{\beta}}_R = \widehat{\boldsymbol{\beta}}_R^T X_c^T X_c \left(X_c^T X_c \right)^{-1} X_c^T \mathbf{y}$$

$$= \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y} = \mathbf{y}^T X_c \left(X_c^T X_c \right)^{-1} X_c^T \mathbf{y}$$

as $\left(X_c^T X_c\right)^{-1}$ is symmetric, as shown in Lecture 3 in an Appendix. We introduce a centered hat matrix by

$$H_c := X_c \left(X_c^T X_c \right)^{-1} X_c^T \tag{30}$$

and

$$SS_{\mathrm{R}} = \mathbf{y}^T H_c \mathbf{y}$$

Another Round of Phantastic Identities

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \mathbf{y}^T C_{ce}\mathbf{y} = \mathbf{y}^T H_c\mathbf{y} + \left(\mathbf{y}^T C_{ce}\mathbf{y} - \mathbf{y}^T H_c\mathbf{y}\right)$$

$$= \mathbf{y}^T H_c\mathbf{y} + \mathbf{y}^T\left(C_{ce} - H_c\right)\mathbf{y}$$

This is

$$SS_T = SS_{\mathrm{R}} + SS_{\mathrm{Re}},$$

hence

$$SS_{\mathrm{Re}} = \mathbf{y}^T\left(C_{ce} - H_c\right)\mathbf{y}.$$

Proposition

1. I) $H_c C_{ce} = H_c$
2. II) H_c *is idempotent and* $\mathrm{rank}\, H_c = k$.
3. III) $C_{ce} - H_c$ *is idempotent and* $\mathrm{rank}\,(C_{ce} - H_c) = n - k - 1$
4. IV) $H_c\,(C_{ce} - H_c) = \mathbf{O}_{k \times k}$.

Proof:

I) $H_c C_{ce} = H_c$

$$H_c C_{ce} = H_c \left(\mathbb{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) = H_c - \frac{1}{n} H_c \mathbf{1}_n \mathbf{1}_n^T.$$

By construction of the centered hat matrix

$$H_c \mathbf{1}_n \mathbf{1}_n^T = X_c \left(X_c^T X_c \right)^{-1} X_c^T \mathbf{1}_n \mathbf{1}_n^T.$$

Here

$$X_c^T \mathbf{1}_n = \left(\mathbf{1}_n^T X_c \right)^T = \left((1 1 \ldots 1) \begin{pmatrix} x_{11} - \bar{x}_1 & \cdots & x_{1k} - \bar{x}_k \\ x_{21} - \bar{x}_1 & \cdots & x_{2k} - \bar{x}_k \\ \vdots & \vdots & \vdots \\ x_{n1} - \bar{x}_1 & \cdots & x_{nk} - \bar{x}_k \end{pmatrix} \right)^T$$

Here

$$X_c^T \mathbf{1}_n = \left((11\ldots 1) \begin{pmatrix} x_{11}-\bar{x}_1 & \cdots & x_{1k}-\bar{x}_k \\ x_{21}-\bar{x}_1 & \cdots & x_{2k}-\bar{x}_k \\ \vdots & \vdots & \vdots \\ x_{n1}-\bar{x}_1 & \cdots & x_{nk}-\bar{x}_k \end{pmatrix} \right)^T$$

E.g., if we take by rules of matrix multiplication the scalar product of the first column with $\mathbf{1}_n^T$ we get

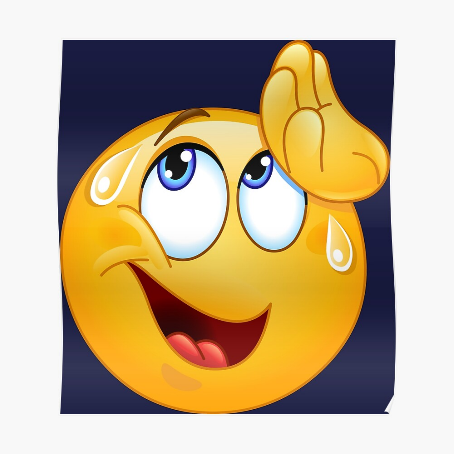
$$\sum_{j=1}^{n} x_{j1} - n\bar{x}_1 = \sum_{j=1}^{n} x_{j1} - \sum_{j=1}^{n} x_{j1}$$

by definition of $\bar{x}_1$ in the environment of (3). The same holds for every other column of X_c. Therefore $X_c^T \mathbf{1}_n = \left(\mathbf{0}_k^T\right)^T = \mathbf{0}_k$ and

$$H_c C_{ce} = H_c + \frac{1}{n} X_c \left(X_c^T X_c\right)^{-1} \mathbf{0}_k \mathbf{1}_n^T$$

$$= H_c + \frac{1}{n} X_c \left(X_c^T X_c\right)^{-1} \mathbf{O}_{k,n} = H_c$$

The part II) is checked by direct multiplication. Parts III) and IV) follow from I) and II). $\square$



Proposition

If $\mathbf{Y} \in N_n(X\boldsymbol{\beta})$*, then*

$$SS_{\mathrm{R}}/\sigma^2 = \widehat{\boldsymbol{\beta}}_R^T X_C^T X_C \widehat{\boldsymbol{\beta}}_R/\sigma^2 \sim \chi^2(k,\lambda), \textit{where} \quad \lambda = \frac{1}{2}\boldsymbol{\beta}_R^T X_C^T X_C \boldsymbol{\beta}_R$$

and

$$SS_{\mathrm{Res}}/\sigma^2 = \left(\sum_{i=1}^{n} (y_i - \bar{y})^2 - \widehat{\boldsymbol{\beta}}_R^T X_C^T \mathbf{y} \right) /\sigma^2 \sim \chi^2(n-k-1)$$

These follow due to Proposition 1 in the same way as the corresponding statements in Lecture 6. □

$$SS_{\mathrm{R}}/\sigma^2 = \widehat{\boldsymbol{\beta}}_R^T X_c^T X_c \widehat{\boldsymbol{\beta}}_R/\sigma^2 \sim \chi^2(k, \lambda), \text{where} \quad \lambda = \frac{1}{2}\boldsymbol{\beta}_R^T X_c^T X_c \boldsymbol{\beta}_R$$

and

$$SS_{\mathrm{Res}}/\sigma^2 = \left(\sum_{i=1}^{n} (y_i - \bar{y})^2 - \widehat{\boldsymbol{\beta}}_R^T X_c^T \mathbf{y} \right) /\sigma^2 \sim \chi^2(n - k - 1)$$

These statistics do not contain β_0!

Proposition

If $\mathbf{Y} \in N_n(X\beta)$, then SS_{R} and SS_{Res} are independent.

Proof as in Lecture 5. $\square$

Now we can state the F-test for

$$H_o\colon \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

$$H_1\colon \beta_j \neq 0 \textit{ for at least one } \beta_j$$

Set

$$F = \frac{SS_{\mathrm{R}}/(k\sigma^2)}{SS_{\mathrm{Res}}/(n-k-1)\sigma^2} = F = \frac{SS_{\mathrm{R}}/k}{SS_{\mathrm{Res}}/(n-k-1)}$$

I) If H_o is true, then $\lambda = 0$ and $F \sim F(k, n-k-1)$.

II) If H_o not true, then $F \sim F(k, n-k-1, \lambda)$.

These statements follow as in Lecture 5.

The F-statistic from Lecture 5 (see the slides)

$$F = \frac{SS_{\mathrm{R}}/k}{SS_{\mathrm{Res}}/(n-k-1)}$$

We shall first analyze (we simplify reading and writing by $\boldsymbol{\beta}_* \mapsto \boldsymbol{\beta}$)

$$\frac{1}{\sigma^2} SS_{\mathrm{R}} \sim \chi^2(k, \lambda), \quad \lambda = \frac{1}{2} (X\boldsymbol{\beta})^T \left(H - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) X\boldsymbol{\beta}. \qquad (31)$$

If $\mathbf{v} \in \mathbb{R}^n$ and A is a real, symmetric, positive-definite $n \times n$ matrix, then the set

$$\mathbb{D}(\mathbf{v}, h) = \{\mathbf{x} \in \mathbb{R}^n | (\mathbf{x} - \mathbf{v})^T A(\mathbf{x} - \mathbf{v}) \leq h\}$$

is an ellipsoid with radius h centered at $\mathbf{v}$. The eigenvectors of A are the principal axes of $\mathbb{D}$.

Cook, R Dennis: Detection of influential observation in linear regression, Technometrics, 19, 1, 15–18, 1977.

$$\left(\frac{\left(\boldsymbol{\beta}_R - \hat{\boldsymbol{\beta}}_R\right)^T X_c^T X_c \left(\boldsymbol{\beta}_R - \hat{\boldsymbol{\beta}}_R\right)}{k} / \mathrm{SS}_{\mathrm{Res}} / (n-k-1) \right) \sim F(k, n-k-1)$$

Confidence Ellipsoid

$100 \cdot (1-\alpha)\%$ confidence ellipsoid for $\boldsymbol{\beta}_{*,R}$ is

$$\left\{ \boldsymbol{\beta}_R \middle| \left(\boldsymbol{\beta}_R - \hat{\boldsymbol{\beta}}_R \right)^T X_c^T X_c \left(\boldsymbol{\beta}_R - \hat{\boldsymbol{\beta}}_R \right) \leq \frac{k \mathrm{SS}_{\mathrm{Res}}}{(n-k-1)} F_{1-\alpha}(k, n-k-1) \right\}.$$

where $F_{1-\alpha}(k, n-k-1)$ is the $1-\alpha$ probability point of the central F-distribution.

APPENDIX A

The F-statistic from Lecture 5 (see the slides)

$$F = \frac{SS_{\mathrm{R}}/k}{SS_{\mathrm{Res}}/(n-k-1)}$$

We shall first analyze (we simplify reading and writing by $\boldsymbol{\beta}_* \mapsto \boldsymbol{\beta}$)

$$\frac{1}{\sigma^2} SS_{\mathrm{R}} \sim \chi^2(k, \lambda), \quad \lambda = \frac{1}{2} (X\boldsymbol{\beta})^T \left(H - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) X\boldsymbol{\beta}. \tag{32}$$

Partitioned Matrices

We write

$$(X\boldsymbol{\beta})^T\left(H-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)X\boldsymbol{\beta}=\boldsymbol{\beta}^TX^T\left(H-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)X\boldsymbol{\beta}$$

$$=\left(\beta_0,\boldsymbol{\beta}_R^T\right)\begin{pmatrix}\mathbf{1}_n^T\\ X_R^T\end{pmatrix}\left(H-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)\begin{pmatrix}\mathbf{1}_n & X_R\end{pmatrix}\begin{pmatrix}\beta_0\\ \boldsymbol{\beta}_R\end{pmatrix} \tag{33}$$

(Convince yourself that the following is a conformable matrix multiplication)

$$\begin{pmatrix}\mathbf{1}_n^T\\ X_R^T\end{pmatrix}\left(H-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)=\begin{pmatrix}\mathbf{1}_n^T\left(H-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)\\ X_R^T\left(H-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)\end{pmatrix}$$

Partitioned Matrices

$$\begin{pmatrix} \mathbf{1}_n^T\left(H - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right) \\ X_R^T\left(H - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right) \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n^T H - \frac{1}{n}\mathbf{1}_n^T\mathbf{1}_n\mathbf{1}_n^T \\ X_R^T H - \frac{1}{n}X_R^T\mathbf{1}_n\mathbf{1}_n^T \end{pmatrix} = \begin{pmatrix} \mathbf{0}_n^T \\ X_R^T - \frac{1}{n}X_R^T\mathbf{1}_n\mathbf{1}_n^T \end{pmatrix}$$

Here we used, as often before, $\mathbf{1}_n^T H = \left(H^T\mathbf{1}_n\right)^T = \left(H\mathbf{1}_n\right)^T = \mathbf{1}_n^T$ and $\mathbf{1}_n^T\mathbf{1}_n = n$ and the rule B, i.e., $X_R^T H = X_R^T$ in the technical appendix XXX. When we insert this in (33) we have

$$= \left(\beta_0, \boldsymbol{\beta}_R^T\right) \begin{pmatrix} \mathbf{0}_n^T \\ X_R^T - \frac{1}{n}X_R^T\mathbf{1}_n\mathbf{1}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{1}_n & X_R \end{pmatrix} \begin{pmatrix} \beta_0 \\ \boldsymbol{\beta}_R \end{pmatrix}$$

Here $X_R^T\mathbf{1}_n - \frac{1}{n}X_R^T\mathbf{1}_n\mathbf{1}_n^T\mathbf{1}_n = X_R^T\mathbf{1}_n - X_R^T\mathbf{1}_n = \mathbf{0}_k$ and $\mathbf{0}_n^T\mathbf{1}_n = 0$ and $\mathbf{0}_n^T X_R = \mathbf{0}_k^T$, as $\mathbf{0}_n^T$ is $1 \times n$ and X_R is $n \times k$.

Partitioned Matrices

Hence we get (please control the required conformabilities in all of this)

$$
= \left(\beta_0, \boldsymbol{\beta}_R^T\right) \begin{pmatrix} 0 & \mathbf{0}_k^T \\ \mathbf{0}_k & X_R^T X_R - \frac{1}{n} X_R^T \mathbf{1}_n \mathbf{1}_n^T X_R \end{pmatrix} \begin{pmatrix} \beta_0 \\ \boldsymbol{\beta}_R \end{pmatrix}
$$

$$
= \boldsymbol{\beta}_R^T \left(X_R^T X_R - \frac{1}{n} X_R^T \mathbf{1}_n \mathbf{1}_n^T X_R \right) \boldsymbol{\beta}_R .
$$

Here the centering matrix re-appears by

$$
X_R^T X_R - \frac{1}{n} X_R^T \mathbf{1}_n \mathbf{1}_n^T X_R = X_R^T \left(\mathbb{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) X_R = X_R^T C_{ce} X_R
$$

Partitioned Matrices

By idempotence and symmetry of C_{ce} we have

$$X_R^T C_{ce} X_R = X_R^T C_{ce} C_{ce} X_R = X_R^T C_{ce}^T C_{ce} X_R = (C_{ce} X_R)^T C_{ce} X_R.$$

But

$$C_{ce} X_R = \left(X_R - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T X_R \right) = X_c.$$

Hence we have written the non-centrality parameter as a quadratic form in the reduced vector of regression coefficients

$$\boxed{\lambda = \frac{1}{2} \boldsymbol{\beta}_R^T X_c^T X_c \boldsymbol{\beta}_R. \qquad (34)}$$

Technical Appendix: Partitioned Matrices

Technical Appendix: An Important Identity for a Partitioned Matrix

Let X be an $n \times p$ matrix, $X = \begin{pmatrix} X_1 & X_1 \end{pmatrix}$

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{X}$$

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[\mathbf{X}_1\mathbf{X}_2] = \mathbf{X}$$

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[\mathbf{X}_1\mathbf{X}_2] = [\mathbf{X}_1\mathbf{X}_2]$$

Consequently,

A $$\boxed{\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}_1 = \mathbf{X}_1 \quad \text{and} \quad \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}_2 = \mathbf{X}_2}$$

Similarly,

B $$\boxed{\mathbf{X}_1'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{X}_1' \quad \text{and} \quad \mathbf{X}_2'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{X}_2'}$$

Multiplication of Partitioined Matrices

If two matrices **A** and **B** are conformal for multiplication, and if **A** and **B** are partitioned so that the submatrices are appropriately conformal, then the product **AB** can be found using the usual pattern of row by column multiplication with the submatrices as if they were single elements; for example

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{pmatrix}. \tag{2.35}$$

The Inverse

We now give the inverses of some special matrices. If $\mathbf{A}$ is symmetric and nonsingular and is partitioned as

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

and if $\mathbf{B} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$, then, provided $\mathbf{A}_{11}^{-1}$ and $\mathbf{B}^{-1}$ exist, the inverse of $\mathbf{A}$ is given by

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}^{-1} \\ -\mathbf{B}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{B}^{-1} \end{pmatrix}. \tag{2.50}$$

Selected Topics from Preceding Lecture Required in this Lecture.

Ordinary Normal (Gaussian) Multiple Regression

$\varepsilon \sim N_n\left(\mathbf{0}, \sigma^2 \mathbb{I}_n\right)$ and $\boldsymbol{\beta}_*$ such that

$$\mathbf{Y} = X\boldsymbol{\beta}_* + \varepsilon \quad \text{True model} \tag{35}$$

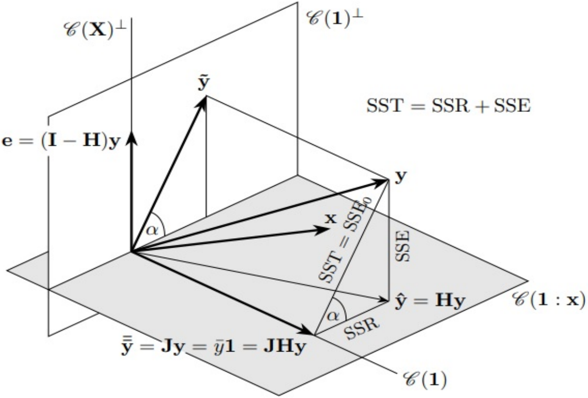
$$\widehat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \mathbf{Y}$$

$$\widehat{\boldsymbol{\beta}} \sim N_{k+1}\left(\boldsymbol{\beta}_*, \sigma^2 (X^T X)^{-1}\right) \tag{36}$$

and

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}_* + (X^T X)^{-1} X^T \varepsilon \tag{37}$$

SS_R as a Quadratic Form


$\mathscr{C}(\mathbf{X})^\perp$
$\mathscr{C}(\mathbf{1})^\perp$
$\tilde{\mathbf{y}}$
$\mathrm{SST} = \mathrm{SSR} + \mathrm{SSE}$
$\mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{y}$
$\mathbf{y}$
$\mathbf{x}$
α
$\mathrm{SST} = \mathrm{SSE}_0$
SSE
$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$
$\mathscr{C}(\mathbf{1} : \mathbf{x})$
α
SSR
$\bar{\bar{\mathbf{y}}} = \mathbf{J}\mathbf{y} = \bar{y}\mathbf{1} = \mathbf{J}\mathbf{H}\mathbf{y}$
$\mathscr{C}(\mathbf{1})$

Non-central chi-square

Definition

$\mathbf{X} \sim N_n(\mu, \mathbb{I}_n)$ *(i.e. $X_1, \ldots, X_n$ are independent, $X_i \sim N(\mu_i, 1)$). Set*

$$W := \mathbf{X}^T\mathbf{X} = \sum_{i=1}^{n} X_i^2, \quad \lambda := \sum_{i=1}^{n} \mu_i^2.$$

W has the **non-central chi-square distribution** *with n degrees of freedom and non-centrality parameter λ, coded as $W \sim \chi^2(n, \lambda)$*

Note that $\chi^2(n, 0) = \chi^2(n)$

Non-central chi-square & Quadratic Forms

Proposition

Let $\mathbf{X} \sim N_n(\mu, \Sigma)$, let A be a symmetric $n \times n$ matrix of constants of rank r, and let $\lambda := \frac{1}{2}\mu^T A \mu$. Then $\mathbf{X}^T A \mathbf{X} \sim \chi^2(r, \lambda)$ if and only if $A\Sigma$ is idempotent.

Theorem 5.5 on p. 117 in Rencher, Alvin C and Schaalje, G Bruce: *Linear Models in Statistics*, 2008. Proof by momentgenerating functions.
We shall now apply this to the quadratic forms SS_{T} and SS_{R}.

Probability Distributions Related to the Normal Distribution: F-distribution

Proposition

If $X_1 \sim \chi^2(n_1)$ and $X_2 \sim \chi^2(n_2)$ are independent. Let

$$V := \frac{X_1/n_1}{X_2/n_2} \tag{38}$$

V has the F-distribution with (n_1, n_2) degrees of freedom, coded as $V \sim F(n_1, n_2)$

This proposition is Problem 10. of Chapter 1 Section 3, in Gut, Allan: An Intermediate Course in Probability. Second Edition, Springer, 2009.

If $X \sim \chi^2(r, \lambda)$, $Y \sim \chi^2(s)$ and X and Y are independent, then

$$Q = \frac{X/r}{Y/s} \sim F(r, s, \lambda) \tag{39}$$

Here $F(r, q, \lambda)$ is the **non-central F-distribution** with non-centrality parameter λ. The pdf is is a noncentral F-distributed random variable.

Hence, by the preceding

$$\frac{SS_{\mathrm{R}}/k}{SS_{\mathrm{Res}}/(n-k-1)} \sim F(k, n-k-1, \lambda) \tag{40}$$

The ratio

$$F = \frac{SS_{\mathrm{R}}/k}{SS_{\mathrm{Res}}/(n-k-1)}$$

is thus called the F-statistic.

F-distribution, critical values

E.g., $F_{0.05}(3,8)4.07$

s/r	1	2	3	4	5	6
1	161	200	216	225	230	234
2	18.5	19	19.2	19.2	19.3	19.3
3	10.1	9.55	9.28	9.12	9.01	8.94
4	7.71	6.94	6.59	6.39	6.26	6.16
5	6.61	5.79	5.41	5.19	5.05	4.95
6	5.99	5.14	4.76	4.53	4.39	4.28
7	5.59	4.74	4.35	4.12	3.97	3.87
8	5.32	4.46	4.07	3.84	3.69	3.58
9	5.12	4.26	3.86	3.63	3.48	3.37
10	4.96	4.1	3.71	3.48	3.33	3.22
11	4.84	3.98	3.59	3.36	3.2	3.09

F-Test

Source	df	Sum of Squares	MSS
Regression	k	SS_{R}	SS_{R}/k
Residual	$n-k-1$	SS_{Res}	$\widehat{\sigma}^2$=SS_{Res}/(n-k-1)
Total	$n-1$	SS_{T}	

Source = source of variation, df= degrees of freedom, SS= sum of squares, MSS= mean sum of squares.

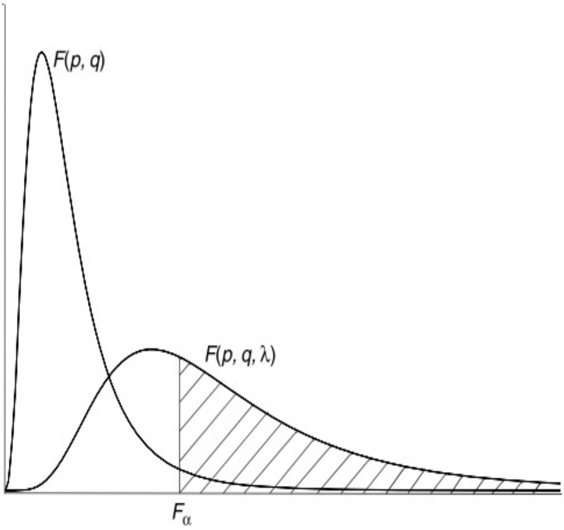


Figure 5.2 Central F, noncentral F, and power of the F test (shaded area).

When an F statistic is used to test a hypothesis H_o, the distribution will typically be central if the (null) hypothesis is true and noncentral if the hypothesis is false.
Thus the noncentral F distribution can often be used to evaluate the power of an F- test. The power of a test is the probability of rejecting H_o for a given value of l. If Fa is the upper a percentage point of the central F distribution, then the power, P(p, q, a, l), can be defined as