14.1. Notes on the F Test

1. If $\text{rank}(\mathbf{A}) = q$, then $\hat{\mathbf{Y}}_H = \mathbf{X}\hat{\boldsymbol{\beta}}_H$, with

$$\hat{\boldsymbol{\beta}}_H = \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}\mathbf{A}\hat{\boldsymbol{\beta}},$$

   where $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$.

   Proof: We use the Lemma from Lecture 13 that gives an expression for $\mathbf{P}_\omega - \mathbf{P}_\Omega$.

$$\begin{aligned}
\hat{\mathbf{Y}}_H &= \mathbf{P}_\omega\mathbf{Y} = \mathbf{P}_\Omega\mathbf{Y} + (\mathbf{P}_\omega - \mathbf{P}_\Omega)\mathbf{Y} \\
&= \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}\mathbf{A}\underbrace{(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}} \\
&= \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}\mathbf{A}\hat{\boldsymbol{\beta}} \\
&= \mathbf{X}(\hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}\mathbf{A}\hat{\boldsymbol{\beta}})
\end{aligned}$$

   Since $\mathbf{X}$ times $\hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}\mathbf{A}\hat{\boldsymbol{\beta}}$ gives $\hat{\mathbf{Y}}_H$, $\hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}\mathbf{A}\hat{\boldsymbol{\beta}}$ must be a least-squares estimate of $\boldsymbol{\beta}_H$.

2. The F-test extends to $H : \mathbf{A}\boldsymbol{\beta} = \mathbf{c}$, for a constant $\mathbf{c}$. In this case, our previous results become (if $\text{rank}(\mathbf{A}) = q$)

$$\hat{\boldsymbol{\beta}}_H = \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c}),$$

$$RSS_H - RSS = (\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c}).$$

   and F has the same distribution as before.

14.2. Examples of the F Test

Example 1 (t-Test):

Let $U_1, \ldots, U_{n_1}$ be i.i.d. $N(\mu_1, \sigma^2)$ and $V_1, \ldots, V_{n_2}$ be i.i.d. $N(\mu_2, \sigma^2)$, independently of the U_i. As a linear model

$$\begin{pmatrix} U_1 \\ \vdots \\ U_{n_1} \\ V_1 \\ \vdots \\ V_{n_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{n_1} \\ \varepsilon_{n_1+1} \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

The hypothesis $H : \mu_1 = \mu_2$ can be written

$$H : \mathbf{A}\boldsymbol{\beta} = (1, -1) \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \mu_1 - \mu_2 = 0.$$

It is easily verified that $\hat{\boldsymbol{\beta}} = \begin{pmatrix} \bar{U} \\ \bar{V} \end{pmatrix}$, so that $\mathbf{A}\hat{\boldsymbol{\beta}} = \bar{U} - \bar{V}$.

$$\begin{aligned} (RSS_H - RSS)/\sigma^2 &= (\mathbf{A}\hat{\boldsymbol{\beta}})^2 [\text{var}(\mathbf{A}\hat{\boldsymbol{\beta}})]^{-1} \\ &= (\bar{U} - \bar{V})^2 \left[\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]^{-1}. \end{aligned}$$

Also, $RSS/(n-2)$ is the pooled sample variance ($n = n_1 + n_2$):

$$RSS/(n-2) = [\sum_i (U_i - \bar{U})^2 + \sum_j (V_j - \bar{V})^2]/(n-2) \equiv S^2.$$

Therefore ($p = 2, q = 1$),

$$F = \frac{RSS_H - RSS}{RSS/(n-2)} = (\bar{U} - \bar{V})^2 \left[S^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]^{-1} = T^2,$$

where $T = (\bar{U} - \bar{V})/(S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}})$ is the two-sample t statistic. The t-test is equivalent to the F test because the square of a t_{n-2} random variable has the $F_{1,n-2}$ distribution.

Example 2 (Multiple Regression):

$$Y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_{p-1} x_{i,p-1} + \varepsilon_i.$$

Test $H : \beta_j = 0$ $(j \neq 0)$, i.e.,

$$H : \mathbf{A}\boldsymbol{\beta} = (0, 0, \ldots, 0, 1, 0, \ldots, 0)\boldsymbol{\beta} = 0$$

($\mathbf{A}$ has 1 in the jth position). Assume $\mathbf{X}$ has rank p so $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. Then

$$\text{var}(\hat{\beta}_j) = \sigma^2[(\mathbf{X}'\mathbf{X})^{-1}]_{j,j}$$

Applying Theorem 13.4.1,

$$\begin{aligned}(RSS_H - RSS)/\sigma^2 &= \frac{1}{\sigma^2}(A\hat{\boldsymbol{\beta}})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(A\hat{\boldsymbol{\beta}}) \\ &= \hat{\beta}_j[\text{var}(\hat{\beta}_j)]^{-1}\hat{\beta}_j \\ &= (\hat{\beta}_j)^2/[\text{var}(\hat{\beta}_j)] \\ &= (\hat{\beta}_j)^2/(\sigma^2[(\mathbf{X}'\mathbf{X})^{-1}]_{j,j})\end{aligned}$$

and

$$\begin{aligned}F = \frac{RSS_H - RSS}{RSS/(n-p)} &= (\hat{\beta}_j)^2/(S^2[(\mathbf{X}'\mathbf{X})^{-1}]_{j,j}) \\ &= (\hat{\beta}_j)^2/[\widehat{SE}(\hat{\boldsymbol{\beta}}_j)]^2\end{aligned}$$

where $S^2 = RSS/(n-p)$.

Let $T = \hat{\boldsymbol{\beta}}_j/\widehat{SE}(\hat{\boldsymbol{\beta}}_j)$ be the usual t statistic for testing the significance of coefficients in a multiple regression model. See that $F = T^2$.

Example 3 (Simple Linear Regression):

$$Y_i = \beta_0 + \beta_1(x_i - \bar{x}) + \varepsilon_i.$$

Then

$$\hat{\beta}_1 = \frac{\sum_i x_i Y_i - \sum_i x_i \sum_i Y_i/n}{\sum_i (x_i - \bar{x})^2} = \frac{\sum_i (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_i (x_i - \bar{x})^2}$$

and

$$\mathrm{var}(\hat{\beta}_1) = \sigma^2 / \sum_i (x_i - \bar{x})^2.$$

From the previous example, the F statistic for testing $H : \beta_1 = 0$ is

$$F = \frac{\hat{\beta}_1^2}{S^2 / \sum_i (x_i - \bar{x})^2}.$$

It can be shown that

$$RSS = (1 - r^2) \sum_i (Y_i - \bar{Y})^2 = (1 - r^2) RSS_H,$$

where r is the sample correlation coefficient between the Y's and the x's.

This leads to the interpretation that $r^2 = (RSS_H - RSS)/RSS_H$ is the proportion of variance explained by the regression relationship. We will later generalize this to the sample multiple correlation coefficient (R^2).

14.3. Power of the F-Test

Consider the model with the usual assumptions and also the normality assumption for the residuals:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I}),$$

with $\text{rank}(\mathbf{X}_{n\times p}) = r$. The F statistic for testing $H : \mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ is

$$F = \frac{(RSS_H - RSS)/q}{RSS/(n-r)},$$

where $\text{rank}(\mathbf{A}_{q\times p}) = q$. Our goal is to calculate

$$\text{Power } = P(F > F^{\alpha}_{q,n-r} | H \text{ not true}).$$

We have seen that

$$\frac{RSS}{\sigma^2} \sim \chi^2_{n-r}, \quad \underline{\text{ whether } H \text{ is true or not}}.$$

By the results in lecture notes #6,

$$\frac{RSS_H - RSS}{\sigma^2} = \frac{\mathbf{Y}'(\mathbf{P}_\Omega - \mathbf{P}_\omega)\mathbf{Y}}{\sigma^2}$$

has a non-central chi-squared distribution with non-centrality parameter

$$\lambda = \boldsymbol{\mu}'(\mathbf{P}_\Omega - \mathbf{P}_\omega)\boldsymbol{\mu}/2\sigma^2.$$

The df is given by the rank of $\mathbf{P}_\Omega - \mathbf{P}_\omega$, which must be q because $(RSS_H - RSS)/\sigma^2 \sim \chi^2_q$ if H is true.

This theorem also implies that RSS_H/σ^2 has a non-central χ^2 distribution. Therefore we have a decomposition of non-central chi-squared variables and Theorem 1.10 (Seber) implies that RSS and $RSS - RSS_H$ are independent. We have proved that F has a non-central F distribution as stated in Theorem 14.3.1.

Definition: Let X_1 and X_2 be independent random variables with $X_1 \sim \chi^2_{d_1}(\lambda)$ and $X_2 \sim \chi^2_{d_2}$. Then the distribution of the ratio

$$F = \frac{X_1/d_1}{X_2/d_2}$$

is defined as the *non-central F distribution* with d_1 numerator degrees of freedom (df), d_2 denominator df, and non-centrality parameter λ, and is denoted $F_{d_1,d_2}(\lambda)$.

Theorem 14.3.1: The F statistic for testing $H : \mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ has the non-central F distribution $F \sim F_{q,n-r}(\lambda)$, where $\lambda = \boldsymbol{\mu}'(\mathbf{P}_\Omega - \mathbf{P}_\omega)\boldsymbol{\mu}/2\sigma^2$.

14.4. Calculating the Non-Centrality Parameter

We have the following representations:

$$\begin{aligned}
\sigma^2 2\lambda &= \boldsymbol{\mu}'(\mathbf{P}_\Omega - \mathbf{P}_\omega)\boldsymbol{\mu} \\
&= \mathbf{Y}'(\mathbf{P}_\Omega - \mathbf{P}_\omega)\mathbf{Y}\,\Big|_{\mathbf{Y}=\boldsymbol{\mu}} \\
&= \mathbf{Y}'((\mathbf{I} - \mathbf{P}_\omega) - (\mathbf{I} - \mathbf{P}_\Omega))\mathbf{Y}\,\Big|_{\mathbf{Y}=\boldsymbol{\mu}} \\
&= (RSS_H - RSS)\,\Big|_{\mathbf{Y}=\boldsymbol{\mu}} \\
&= (\mathbf{A}\hat{\boldsymbol{\beta}})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}})\,\Big|_{\mathbf{Y}=\boldsymbol{\mu}} \\
&= (\mathbf{A}\boldsymbol{\beta})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}(\mathbf{A}\boldsymbol{\beta}).
\end{aligned}$$

To calculate the non-centrality parameter, we can substitute the true mean $\boldsymbol{\mu}$ under the alternative hypothesis or the true parameter $\mathbf{A}\boldsymbol{\beta}$ into the appropriate formula.

We can use this result to calculate the expected value of the F statistic. We have

$$\begin{aligned}
E[RSS_H - RSS] &= \sigma^2 q + (\mathbf{A}\boldsymbol{\beta})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}']^{-1}(\mathbf{A}\boldsymbol{\beta}) \\
&= \sigma^2(q + 2\lambda).
\end{aligned}$$

Therefore, using $E[1/(RSS/\sigma^2)] = E[1/\chi^2_{n-r}] = (n-r-2)^{-1}$:

$$\begin{aligned}
E[F] &= E[(RSS_H - RSS)/q]E[1/\{RSS/(n-r)\}] \\
&= [\sigma^2(q + 2\lambda)/q]\{(n-r)/\sigma^2\}E[1/(RSS/\sigma^2)] \\
&= [\sigma^2(q + 2\lambda)/q]\{(n-r)/\sigma^2\}(n-r-2)^{-1} \\
&= (1 + 2\lambda/q)(\frac{n-r}{n-r-2})
\end{aligned}$$