Lecture 9

September 19, 2018

Last class

- In the last class, we started the study of hypothesis testing in linear regression under the assumption of normality.

Last class

- In the last class, we started the study of hypothesis testing in linear regression under the assumption of normality. We assumed that $Y \sim N_n(X\beta, \sigma^2 I)$ (or equivalently that $e \sim N_n(0, \sigma^2 I)$.

Last class

- In the last class, we started the study of hypothesis testing in linear regression under the assumption of normality. We assumed that $Y \sim N_n(X\beta, \sigma^2 I)$ (or equivalently that $e \sim N_n(0, \sigma^2 I)$.
- Two consequences of the normality assumption are that $\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$ and that $RSS/\sigma^2 \sim \chi^2_{n-p-1}$.

Last class

- In the last class, we started the study of hypothesis testing in linear regression under the assumption of normality. We assumed that $Y \sim N_n(X\beta, \sigma^2 I)$ (or equivalently that $e \sim N_n(0, \sigma^2 I)$).
- Two consequences of the normality assumption are that $\hat{\beta} \sim N(\beta, \sigma^2(X^T X)^{-1})$ and that $RSS/\sigma^2 \sim \chi^2_{n-p-1}$.
- Moreover, $\hat{\beta}$ and RSS are statistically independent.

Testing $H_0 : \beta_j = 0$

- We started by testing the hypothesis $H_0 : \beta_j = 0$.

Testing $H_0 : \beta_j = 0$

- We started by testing the hypothesis $H_0 : \beta_j = 0$. We saw the t-test for testing this. The t-test is based on the statistic:

$$\frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)}$$

which is distributed according to the t-distribution with $n - p - 1$ degrees of freedom.

Testing $H_0 : \beta_j = 0$

- We started by testing the hypothesis $H_0 : \beta_j = 0$. We saw the t-test for testing this. The t-test is based on the statistic:

$$\frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)}$$

  which is distributed according to the t-distribution with $n - p - 1$ degrees of freedom.
- The p-value for testing $H_0 : \beta_j = 0$ can then be got by

$$\mathbb{P}\left(|t_{n-r}| > \left|\frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)}\right|\right).$$

Testing $H_0 : \beta_j = 0$

- We started by testing the hypothesis $H_0 : \beta_j = 0$. We saw the t-test for testing this. The t-test is based on the statistic:

$$\frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)}$$

which is distributed according to the t-distribution with $n - p - 1$ degrees of freedom.

- The p-value for testing $H_0 : \beta_j = 0$ can then be got by

$$\mathbb{P}\left(|t_{n-r}| > \left| \frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)} \right| \right).$$

R = p+1

- Note that when $n - r$ is large, the t-distribution is almost the same as a standard normal distribution.

Testing $H_0 : \beta_j = 0$

- The p-value above is for the two-sided alternative $H_1 : \beta_j \neq 0$.

Testing $H_0 : \beta_j = 0$

- The p-value above is for the two-sided alternative $H_1 : \beta_j \neq 0$. If you want to test against one-sided alternatives $H_1 : \beta_j > 0$ or $\beta_j < 0$, then just divide the above p-value by 2.

F-test for $H_0 : \beta_j = 0$

- There is another way of testing the null hypothesis that $H_0 : \beta_j = 0$.

F-test for $H_0 : \beta_j = 0$

- There is another way of testing the null hypothesis that $H_0 : \beta_j = 0$. This turns out to be identical to the t-test but can be generalized more easily to other tests.

F-test for $H_0 : \beta_j = 0$

- There is another way of testing the null hypothesis that $H_0 : \beta_j = 0$. This turns out to be identical to the t-test but can be generalized more easily to other tests.
- The F-test for $H_0 : \beta_j = 0$ is based on the following idea. The null hypothesis H_0 says that the explanatory variable x_j can be dropped from the linear model. Let us call this reduced model m.

F-test for $H_0 : \beta_j = 0$

- There is another way of testing the null hypothesis that $H_0 : \beta_j = 0$. This turns out to be identical to the t-test but can be generalized more easily to other tests.
- The F-test for $H_0 : \beta_j = 0$ is based on the following idea. The null hypothesis H_0 says that the explanatory variable x_j can be dropped from the linear model. Let us call this reduced model m.
- Also, let us call the original model M (this is the full model: $y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$).

F-test for $H_0 : \beta_j = 0$

- There is another way of testing the null hypothesis that $H_0 : \beta_j = 0$. This turns out to be identical to the t-test but can be generalized more easily to other tests.
- The F-test for $H_0 : \beta_j = 0$ is based on the following idea. The null hypothesis H_0 says that the explanatory variable x_j can be dropped from the linear model. Let us call this reduced model m.
- Also, let us call the original model M (this is the full model: $y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$).
- The model m is a *submodel* of M.

F-test for $H_0 : \beta_j = 0$

- The following presents another natural test for H_0. Let the Residual

F-test for $H_0 : \beta_j = 0$

- The following presents another natural test for H_0. Let the Residual Sum of Squares in the model m be denoted by $RSS(m)$ and let the RSS in the full model be $RSS(M)$.

F-test for $H_0 : \beta_j = 0$

- The following presents another natural test for H_0. Let the Residual Sum of Squares in the model m be denoted by $RSS(m)$ and let the RSS in the full model be $RSS(M)$.
- It is always true that $RSS(M) \leq RSS(m)$.

F-test for $H_0 : \beta_j = 0$

- The following presents another natural test for H_0. Let the Residual Sum of Squares in the model m be denoted by $RSS(m)$ and let the RSS in the full model be $RSS(M)$.
- It is always true that $RSS(M) \leq RSS(m)$.
- Now if $RSS(M)$ is *much smaller* than $RSS(m)$, it means that the explanatory variable x_j contributes a lot to the regression and hence cannot be dropped i.e., we reject the null hypothesis H_0.

F-test for $H_0 : \beta_j = 0$

- The following presents another natural test for H_0. Let the Residual Sum of Squares in the model m be denoted by $RSS(m)$ and let the RSS in the full model be $RSS(M)$.
- It is always true that $RSS(M) \leq RSS(m)$.
- Now if $RSS(M)$ is *much smaller* than $RSS(m)$, it means that the explanatory variable x_j contributes a lot to the regression and hence cannot be dropped i.e., we reject the null hypothesis H_0.
- On the other hand, if $RSS(M)$ is *only a little smaller* than $RSS(m)$, then x_j does not really contribute a lot in predicting y and hence can be dropped i.e., we do not reject H_0.

F-test for $H_0 : \beta_j = 0$

- Therefore one can test H_0 via the test statistic:

$$RSS(m) - RSS(M).$$

F-test for $H_0 : \beta_j = 0$

- Therefore one can test H_0 via the test statistic:

$$RSS(m) - RSS(M).$$

- We would reject the null hypothesis if this is large. How large?

F-test for $H_0 : \beta_j = 0$

- Therefore one can test H_0 via the test statistic:

$$RSS(m) - RSS(M).$$

- We would reject the null hypothesis if this is large. How large?
- To answer this, we need to look at the **null distribution** of $RSS(m) - RSS(M)$.

F-test for $H_0 : \beta_j = 0$

- Therefore one can test H_0 via the test statistic:

$$RSS(m) - RSS(M).$$

- We would reject the null hypothesis if this is large. How large?
- To answer this, we need to look at the **null distribution** of $RSS(m) - RSS(M)$.
- We will show later today that

$$\frac{RSS(m) - RSS(M)}{\sigma^2} \sim \chi_1^2$$

  under the null hypothesis.

F-test for $H_0 : \beta_j = 0$

- Therefore one can test H_0 via the test statistic:

$$RSS(m) - RSS(M).$$

- We would reject the null hypothesis if this is large. How large?
- To answer this, we need to look at the **null distribution** of $RSS(m) - RSS(M)$.
- We will show later today that

whether use M or m to estimate sigma square?

$$\frac{RSS(m) - RSS(M)}{\sigma^2} \sim \chi^2_1$$

under the null hypothesis.
- Since we do not know σ^2, we estimate it by the square of the residual standard error:

$$\hat{\sigma}^2 = \frac{RSS(M)}{n-p-1},$$

This one is independent of RSS(m)-RSS(M)

F-test for $H_0 : \beta_j = 0$

- Since we do not know σ^2, we estimate it by the square of the residual standard error:

  $$\hat{\sigma}^2 = \frac{RSS(M)}{n-p-1},$$

  to obtain the test statistic:

  $$\frac{RSS(m) - RSS(M)}{RSS(M)/(n-p-1)}$$

F-test for $H_0 : \beta_j = 0$

- Since we do not know σ^2, we estimate it by the square of the residual standard error:
$$\hat{\sigma}^2 = \frac{RSS(M)}{n-p-1},$$
to obtain the test statistic:
$$\frac{RSS(m) - RSS(M)}{RSS(M)/(n-p-1)}$$
- The numerator and the denominator are independent (to be shown later). This independence will not hold if the denominator were $RSS(m)/(n-p)$.

F-test for $H_0 : \beta_j = 0$

- Since we do not know σ^2, we estimate it by the square of the residual standard error:
$$\hat{\sigma}^2 = \frac{RSS(M)}{n-p-1},$$
to obtain the test statistic:
$$\frac{RSS(m) - RSS(M)}{RSS(M)/(n-p-1)}$$
- The numerator and the denominator are independent (to be shown later). This independence will not hold if the denominator were $RSS(m)/(n-p)$.
- Thus under the null hypothesis
$$\frac{RSS(m) - RSS(M)}{RSS(M)/(n-m)} \sim F_{1,n-p-1}.$$

F-test for $H_0 : \beta_j = 0$

- Since we do not know σ^2, we estimate it by the square of the residual standard error:
$$\hat{\sigma}^2 = \frac{RSS(M)}{n-p-1},$$
to obtain the test statistic:
$$\frac{RSS(m) - RSS(M)}{RSS(M)/(n-p-1)}$$
- The numerator and the denominator are independent (to be shown later). This independence will not hold if the denominator were $RSS(m)/(n-p)$.
- Thus under the null hypothesis
$$\frac{RSS(m) - RSS(M)}{RSS(M)/(n-m)} \sim F_{1,n-p-1}.$$
p-value can therefore be got by
$$\mathbb{P}\left(F_{1,n-p-1} > \frac{RSS(m) - RSS(M)}{RSS(M)/(n-p-1)}\right).$$

F-test for $H_0 : \beta_j = 0$

- This is the p-value for the two-sided alternative. For the one-sided alternative, divide by two.

Equivalence of These Two Tests

- It turns out that these two tests for testing $H_0 : \beta_j = 0$ are equivalent in the sense that they give the same p-value.

Equivalence of These Two Tests

- It turns out that these two tests for testing $H_0 : \beta_j = 0$ are equivalent in the sense that they give the same p-value.
- This is because

$$\left(\frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)}\right)^2 = \frac{RSS(m) - RSS(M)}{RSS(M)/(n-p-1)}$$

This is not very difficult to prove but we shall skip its proof.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

which has p explanatory variables.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.
- Let m denote a sub-model of M that is obtained by a linear constraint on the parameter $\beta = (\beta_0, \ldots, \beta_p)$ of M.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.
- Let m denote a sub-model of M that is obtained by a linear constraint on the parameter $\beta = (\beta_0, \ldots, \beta_p)$ of M. Examples include:

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

  $$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.
- Let m denote a sub-model of M that is obtained by a linear constraint on the parameter $\beta = (\beta_0, \ldots, \beta_p)$ of M. Examples include:
  1. For the constraint $\beta_1 = 0$, the model m becomes: $y_i = \beta_0 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.

- Let m denote a sub-model of M that is obtained by a linear constraint on the parameter $\beta = (\beta_0, \ldots, \beta_p)$ of M. Examples include:
  1. For the constraint $\beta_1 = 0$, the model m becomes: $y_i = \beta_0 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$.
  2. For $\beta_1 = \beta_2 = 0$, the model m becomes $y_i = \beta_0 + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + e_i$.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.

- Let m denote a sub-model of M that is obtained by a linear constraint on the parameter $\beta = (\beta_0, \ldots, \beta_p)$ of M. Examples include:
  1. For the constraint $\beta_1 = 0$, the model m becomes: $y_i = \beta_0 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$.
  2. For $\beta_1 = \beta_2 = 0$, the model m becomes $y_i = \beta_0 + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + e_i$.
  3. For $\beta_1 = \cdots = \beta_p = 0$, the model m becomes $y_i = \beta_0 + e_i$.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.

- Let m denote a sub-model of M that is obtained by a linear constraint on the parameter $\beta = (\beta_0, \ldots, \beta_p)$ of M. Examples include:

  1. For the constraint $\beta_1 = 0$, the model m becomes: $y_i = \beta_0 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$.
  2. For $\beta_1 = \beta_2 = 0$, the model m becomes $y_i = \beta_0 + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + e_i$.
  3. For $\beta_1 = \cdots = \beta_p = 0$, the model m becomes $y_i = \beta_0 + e_i$.
  4. For $\beta_1 = \beta_2$, the model m becomes $y_i = \beta_0 + \beta_1(x_{i1} + x_{i2}) + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + e_i$.

General Linear Hypothesis Testing via the F-test

- Let M denote the full regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i$$

  which has p explanatory variables. Let r denote the rank of X in this full model.

- Let m denote a sub-model of M that is obtained by a linear constraint on the parameter $\beta = (\beta_0, \ldots, \beta_p)$ of M. Examples include:
  1. For the constraint $\beta_1 = 0$, the model m becomes: $y_i = \beta_0 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$.
  2. For $\beta_1 = \beta_2 = 0$, the model m becomes $y_i = \beta_0 + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + e_i$.
  3. For $\beta_1 = \cdots = \beta_p = 0$, the model m becomes $y_i = \beta_0 + e_i$.
  4. For $\beta_1 = \beta_2$, the model m becomes $y_i = \beta_0 + \beta_1 (x_{i1} + x_{i2}) + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + e_i$.
  5. For $\beta_1 = 3$, the model m becomes $y_i = \beta_0 + 3x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$.

General Linear Hypothesis Testing via the F-test

- How do we test the hypothesis $H_0 : m$ against $H_1 : M$?

General Linear Hypothesis Testing via the F-test

- How do we test the hypothesis $H_0 : m$ against $H_1 : M$?
- This test can be carried out by noting that if $RSS(m) - RSS(M)$ is large, then m is not a good model for the data and we therefore reject H_0.

General Linear Hypothesis Testing via the F-test

- How do we test the hypothesis $H_0 : m$ against $H_1 : M$?
- This test can be carried out by noting that if $RSS(m) - RSS(M)$ is large, then m is not a good model for the data and we therefore reject H_0. On the other hand, if $RSS(m) - RSS(M)$ is small, we do not reject H_0.

General Linear Hypothesis Testing via the F-test

- How do we test the hypothesis $H_0 : m$ against $H_1 : M$?
- This test can be carried out by noting that if $RSS(m) - RSS(M)$ is large, then m is not a good model for the data and we therefore reject H_0. On the other hand, if $RSS(m) - RSS(M)$ is small, we do not reject H_0.
- The test is therefore based on $RSS(m) - RSS(M)$.

General Linear Hypothesis Testing via the F-test

- How do we test the hypothesis $H_0 : m$ against $H_1 : M$?
- This test can be carried out by noting that if $RSS(m) - RSS(M)$ is large, then m is not a good model for the data and we therefore reject H_0. On the other hand, if $RSS(m) - RSS(M)$ is small, we do not reject H_0.
- The test is therefore based on $RSS(m) - RSS(M)$.
- What is the distribution of this quantity under the null hypothesis?

General Linear Hypothesis Testing via the F-test

- How do we test the hypothesis $H_0 : m$ against $H_1 : M$?
- This test can be carried out by noting that if $RSS(m) - RSS(M)$ is large, then m is not a good model for the data and we therefore reject H_0. On the other hand, if $RSS(m) - RSS(M)$ is small, we do not reject H_0.
- The test is therefore based on $RSS(m) - RSS(M)$.
- What is the distribution of this quantity under the null hypothesis?
- We will show that

  $$\frac{RSS(m) - RSS(M)}{\sigma^2} \sim \chi^2_{p-q}$$

  where q is the number of explanatory variables in m and p is the number of explanatory variables in M. Let us now prove this fact.

General Linear Hypothesis Testing via the F-test

- Let the hat matrices in the two models be denoted by $H(m)$ and $H(M)$ respectively.

General Linear Hypothesis Testing via the F-test

- Let the hat matrices in the two models be denoted by $H(m)$ and $H(M)$ respectively. Write

$$RSS(m) = Y^T(I - H(m))Y \quad \text{and} \quad RSS(M) = Y^T(I - H(M))Y$$

General Linear Hypothesis Testing via the F-test

- Let the hat matrices in the two models be denoted by $H(m)$ and $H(M)$ respectively. Write

$$RSS(m) = Y^T(I - H(m))Y \quad \text{and} \quad RSS(M) = Y^T(I - H(M))Y$$

so that

$$RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y.$$

General Linear Hypothesis Testing via the F-test

- Let the hat matrices in the two models be denoted by $H(m)$ and $H(M)$ respectively. Write

$$RSS(m) = Y^T(I - H(m))Y \quad \text{and} \quad RSS(M) = Y^T(I - H(M))Y$$

so that

$$RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y.$$

- We need the null distribution of $RSS(m) - RSS(M)$.

General Linear Hypothesis Testing via the F-test

- Let the hat matrices in the two models be denoted by $H(m)$ and $H(M)$ respectively. Write

$$RSS(m) = Y^T(I - H(m))Y \quad \text{and} \quad RSS(M) = Y^T(I - H(M))Y$$

so that

$$RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y.$$

- We need the null distribution of $RSS(m) - RSS(M)$. So we shall assume that $Y = X(m)\beta(m) + e$ (where $X(m)$ is the X-matrix in the model m). Why?

General Linear Hypothesis Testing via the F-test

- Let the hat matrices in the two models be denoted by $H(m)$ and $H(M)$ respectively. Write

  $$RSS(m) = Y^T(I - H(m))Y \quad \text{and} \quad RSS(M) = Y^T(I - H(M))Y$$

  so that

  $$RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y.$$

- We need the null distribution of $RSS(m) - RSS(M)$. So we shall assume that $Y = X(m)\beta(m) + e$ (where $X(m)$ is the X-matrix in the model m). Why? It is important to realize that $H(m)X(m) = X(m)$ and also $H(M)X(m) = X(m)$.

General Linear Hypothesis Testing via the F-test

- Let the hat matrices in the two models be denoted by $H(m)$ and $H(M)$ respectively. Write

  $RSS(m) = Y^T(I - H(m))Y$ and $RSS(M) = Y^T(I - H(M))Y$

  so that

$$RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y.$$

- We need the null distribution of $RSS(m) - RSS(M)$. So we shall assume that $Y = X(m)\beta(m) + e$ (where $X(m)$ is the X-matrix in the model m). Why? It is important to realize that $H(m)X(m) = X(m)$ and also $H(M)X(m) = X(m)$. So

  Plug in Y = X(m)ß(m)+e

$$RSS(m) - RSS(M) = e^T(H(M) - H(m))e.$$

General Linear Hypothesis Testing via the F-test

- $H(M) - H(m)$ is a symmetric and idempotent $n \times n$ matrix of rank $p - q$.

Prove it by taking square and see if we can get the same matrix

General Linear Hypothesis Testing via the F-test

- $H(M) - H(m)$ is a symmetric and idempotent $n \times n$ matrix of rank $p - q$. To see that $H(M) - H(m)$ is idempotent notice that

$$(H(M) - H(m))^2 = H(M) + H(m) - 2H(M)H(m). \quad (1)$$

General Linear Hypothesis Testing via the F-test

- $H(M) - H(m)$ is a symmetric and idempotent $n \times n$ matrix of rank $p - q$. To see that $H(M) - H(m)$ is idempotent notice that

$$(H(M) - H(m))^2 = H(M) + H(m) - 2H(M)H(m). \quad (1)$$

- Now $H(M)H(m) = H(m)$.

General Linear Hypothesis Testing via the F-test

- $H(M) - H(m)$ is a symmetric and idempotent $n \times n$ matrix of rank $p - q$. To see that $H(M) - H(m)$ is idempotent notice that

$$(H(M) - H(m))^2 = H(M) + H(m) - 2H(M)H(m). \qquad (1)$$

- Now $H(M)H(m) = H(m)$. To see this, it is enough to show that $H(M)H(m)v = H(m)v$ for every vector v.

General Linear Hypothesis Testing via the F-test

- $H(M) - H(m)$ is a symmetric and idempotent $n \times n$ matrix of rank $p - q$. To see that $H(M) - H(m)$ is idempotent notice that

Take transpose

$$(H(M) - H(m))^2 = H(M) + H(m) - 2H(M)H(m). \quad (1)$$

- Now $H(M)H(m) = H(m)$. To see this, it is enough to show that $H(M)H(m)v = H(m)v$ for every vector v.
- Now recall that $H(m)v$ denotes the projection of v onto the column space of $X(m)$. And $H(M)H(m)v$ projects $H(m)v$ projects onto the column space of $X(M)$ (which equals the original X matrix).

General Linear Hypothesis Testing via the F-test

- $H(M) - H(m)$ is a symmetric and idempotent $n \times n$ matrix of rank $p - q$. To see that $H(M) - H(m)$ is idempotent notice that

$$(H(M) - H(m))^2 = H(M) + H(m) - 2H(M)H(m). \qquad (1)$$

Why H(m)H(M) = H(m)^T?

- Now $H(M)H(m) = H(m)$. To see this, it is enough to show that $H(M)H(m)v = H(m)v$ for every vector v.
- Now recall that $H(m)v$ denotes the projection of v onto the column space of $X(m)$. And $H(M)H(m)v$ projects $H(m)v$ projects onto the column space of $X(M)$ (which equals the original X matrix).
- But because the column space of $X(m)$ is contained in the column space of $X(M)$, it follows that $H(m)v$ is already contained in the column space of $X(M)$. Thus its projection onto the column space of $X(M)$ equals itself. So $H(M)H(m)v = H(m)v$.

General Linear Hypothesis Testing via the F-test

- Because $H(M)H(m) = H(m)$, it follows from (1) that

$$(H(M) - H(m))^2 = H(M) - H(m).$$

General Linear Hypothesis Testing via the F-test

- Because $H(M)H(m) = H(m)$, it follows from (1) that

$$(H(M) - H(m))^2 = H(M) - H(m).$$

- Therefore, under the null hypothesis

$$\frac{RSS(m) - RSS(M)}{\sigma^2} = \left(\frac{e}{\sigma}\right)^T (H(M) - H(m))\frac{e}{\sigma} \sim \chi^2_{p-q}.$$

General Linear Hypothesis Testing via the F-test

- Because $H(M)H(m) = H(m)$, it follows from (1) that

$$(H(M) - H(m))^2 = H(M) - H(m).$$

trace(H(M)-H(m)) = rank of it because symmetric and idempotent

- Therefore, under the null hypothesis

$$\frac{RSS(m) - RSS(M)}{\sigma^2} = \left(\frac{e}{\sigma}\right)^T (H(M) - H(m))\frac{e}{\sigma} \sim \chi^2_{p-q}.$$

- Since we do not know σ^2, we estimate it by

$$\hat{\sigma}^2 = \frac{RSS(M)}{n - p - 1},$$

to obtain the test statistic:

$$\frac{RSS(m) - RSS(M)}{RSS(M)/(n - p - 1)}$$

General Linear Hypothesis Testing via the F-test

- The numerator and the denominator are independent because $RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y$ and $RSS(M) = Y^T(I - H(M))Y$ and the product of the matrices

$$(H(M) - H(m))(I - H(M)) = H(M) - H(M)^2 - H(m) + H(m)H(M)$$

General Linear Hypothesis Testing via the F-test

- The numerator and the denominator are independent because $RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y$ and $RSS(M) = Y^T(I - H(M))Y$ and the product of the matrices

$$(H(M) - H(m))(I - H(M)) = H(M) - H(M)^2 - H(m) + H(m)H(M)$$

- Thus under the null hypothesis

$$\frac{(RSS(m) - RSS(M))/(p-q)}{RSS(M)/(n-p-1)} \sim F_{p-q,n-p-1}.$$

General Linear Hypothesis Testing via the F-test

- The numerator and the denominator are independent because $RSS(m) - RSS(M) = Y^T(H(M) - H(m))Y$ and $RSS(M) = Y^T(I - H(M))Y$ and the product of the matrices

  $$(H(M) - H(m))(I - H(M)) = H(M) - H(M)^2 - H(m) + H(m)H(M)$$

- Thus under the null hypothesis

  $$\frac{(RSS(m) - RSS(M))/(p-q)}{RSS(M)/(n-p-1)} \sim F_{p-q,n-p-1}.$$

- p-value can therefore be got by

  $$\mathbb{P}\left(F_{p-q,n-p-1} > \frac{(RSS(m) - RSS(M))/(p-q)}{RSS(M)/(n-p-1)}\right).$$

General Linear Hypothesis Testing via the F-test

- If the null hypothesis can be written in terms of a single linear function of β, such as $H_0 : \beta_1 + 5\beta_3 = 5$. Then it can also be tested via the t-test; using the statistic:

$$\frac{\hat{\beta}_1 + 5\hat{\beta}_3 - 5}{s.e(\hat{\beta}_1 + 5\hat{\beta}_3)}$$

  which has the t-distribution with $n - p - 1$ degrees of freedom under H_0. This test and the corresponding F-test will have the same p-value.

Testing for all explanatory variables

- How do we test $H_0 : \beta_1 = \cdots = \beta_p = 0$ against its complement?

Testing for all explanatory variables

- How do we test $H_0 : \beta_1 = \cdots = \beta_p = 0$ against its complement?
- Just take m to be the model $y_i = \beta_0 + e_i$.

Testing for all explanatory variables

- How do we test $H_0 : \beta_1 = \cdots = \beta_p = 0$ against its complement?
- Just take m to be the model $y_i = \beta_0 + e_i$. In this case, $RSS(m) = TSS$ and $q = 0$ and $RSS(M) = RSS$.

Testing for all explanatory variables

- How do we test $H_0 : \beta_1 = \cdots = \beta_p = 0$ against its complement?
- Just take m to be the model $y_i = \beta_0 + e_i$. In this case, $RSS(m) = TSS$ and $q = 0$ and $RSS(M) = RSS$.
- Thus the p-value is

$$\mathbb{P}\left\{ F_{p,n-p-1} > \frac{(TSS - RSS)/p}{RSS/(n-p-1)} \right\}.$$