31 oct · lec 7

REMARK

the test about an individual coefficient β_j and about the overall significance are particular cases of this test.

- **TEST about a SINGLE PARAMETER β_j**

Special case with $p_0 = p-1$

Assume we are testing the significance of the last parameter β_p.

(or simply sort the columns of X so that the last covariate is the one corresponding to the parameter of interest)

$$\begin{cases} H_0: \beta_p = 0 \\ H_1: \beta_p \neq 0 \end{cases}$$

Testing β_p is equivalent to testing a model with $p_0 = p-1$ covariates

In this case we can partition $\underline{\beta}$ and X as

$$\underline{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{p-1} \\ \hline \beta_p \end{bmatrix} \begin{array}{l} p_0 = p-1 \\ \\ \\ 1 \end{array} \qquad X = [\underline{x}_1 \dots \underline{x}_{p-1} \mid \underline{x}_p]$$

the test becomes

$$F = \frac{\frac{\tilde{\Sigma}^2 - \hat{\Sigma}^2}{1}}{\frac{\hat{\Sigma}^2}{n-p}} \overset{H_0}{\sim} F_{1,n-p} \qquad F = (T_p)^2 \quad \text{with} \quad T_p = \frac{\hat{B}_p - 0}{\sqrt{\hat{V}(\hat{B}_p)}} \overset{H_0}{\sim} t_{n-p}$$

(recall: if $V \sim t_m$, then $V^2 \sim F_{1,m}$)

- **TEST ABOUT THE OVERALL SIGNIFICANCE**

if we consider $p_0 = 1$

$$\begin{cases} H_0: \beta_2 = \beta_3 = \dots = \beta_p = 0 \\ H_1: \overline{H_0} \end{cases}$$

then $$\underline{\beta} = \begin{bmatrix} \beta_1 \\ \hline \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \begin{array}{l} 1 = p_0 \\ \\ p-1 \\ \\ \end{array} \qquad X = [\underline{x}_1 \mid \underline{x}_2 \dots \underline{x}_p]$$

The restricted model corresponds to the NULL MODEL (model with only the intercept)

the test is $$F = \frac{\frac{\tilde{\Sigma}^2 - \hat{\Sigma}^2}{p-1}}{\frac{\hat{\Sigma}^2}{n-p}} \overset{H_0}{\sim} F_{p-1,n-p}$$

- **EQUIVALENCE WITH THE TEST ABOUT THE COEFFICIENT R^2**

Under H_0, all coefficients but β_1 (intercept) are zero: none of the covariates is useful to predict y.

The model assumed under H_0 is $Y_i = \beta_1 + \varepsilon_i$

We know that in the null model the estimate of β_1 is $\tilde{\beta}_1 = \bar{y}$.

The predicted values are $\tilde{y}_i = \bar{y}$ for all $i = 1, \dots, n$

The residuals are $\tilde{e}_i = y_i - \bar{y}$

The estimate of σ^2 is $\tilde{\sigma}^2 = \frac{1}{n} \tilde{\underline{e}}^T \tilde{\underline{e}}$

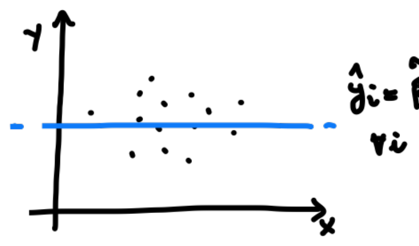
The distribution of the estimator is $\frac{n\tilde{\Sigma}^2}{\sigma^2} \sim \chi^2_{n-1}$

This model corresponds to the case of "no linear relationship between y and the covariates".

We have seen that the coefficient R^2 in this case is close to zero.

Similarly to what we have seen for the simple linear model, we can reformulate this hypothesis as a test on the value of the coefficient R^2 associated with the model:

$$\begin{cases} H_0: R^2 = 0 \\ H_1: R^2 \neq 0 \end{cases}$$



$\hat{y}_i = \tilde{\beta}_1 = \bar{y}$ and $R^2 = 0$

We used a transformation of R^2: $\frac{R^2}{1-R^2}$

Here, $$F = \frac{\frac{\tilde{\Sigma}^2 - \hat{\Sigma}^2}{p-1}}{\frac{\hat{\Sigma}^2}{n-p}} =$$

$$= \frac{\tilde{\Sigma}^2 - \hat{\Sigma}^2}{\hat{\Sigma}^2} \cdot \frac{n-p}{p-1} = \frac{\tilde{\underline{\varepsilon}}^T \tilde{\underline{\varepsilon}} - \underline{\varepsilon}^T \underline{\varepsilon}}{\underline{\varepsilon}^T \underline{\varepsilon}} \cdot \frac{n-p}{p-1}$$

$$= \frac{SSE_{H_0} - SSE_{H_1}}{SSE_{H_1}} \cdot \frac{n-p}{p-1} =$$

$$= \frac{SST - SSE}{SSE} \cdot \frac{n-p}{p-1} = \frac{SSR}{SSE} \cdot \frac{n-p}{p-1} = \frac{R^2}{1-R^2} \cdot \frac{n-p}{p-1} \overset{H_0}{\sim} F_{p-1,n-p}$$

$$\frac{R^2}{1-R^2} = \frac{SSR}{SST} \cdot \left(1 - \frac{SSR}{SST}\right)^{-1}$$

$$= \frac{SSR}{\cancel{SST}} \cdot \frac{\cancel{SST}}{SST - SSR} = \frac{SSR}{SSE}$$