Multiple Linear Regression model

- one response variable (Y)
- more than one explanatory variables (x's)

slope Associated to x_1

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{p-1} x_{i,p-1} + \varepsilon_i, \quad i = 1, \ldots, n$$

intercept (β_0)

Notation: we add $\underline{x_0}$ Associated to β_0

$$Y_i = \beta_0 \underbrace{x_{i0}}_{=1} + \beta_1 x_{i1} + \cdots + \beta_{p-1} x_{i,p-1} + \varepsilon_i, \quad i = 1, \ldots, n$$

Assumptions: $\varepsilon_i \Rightarrow$ Random variables (error)

$E(\varepsilon_i) = 0$; $Var(\varepsilon_i) = \sigma^2$; $Cov(\varepsilon_i, \varepsilon_j) = 0 \quad \forall i \neq j$

And usually we Assume that $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

1. $E[Y \mid x_0, \ldots, x_{p-1}] = \beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1}$
2. $Var(Y \mid x_0, \ldots, x_{p-1}) = Var(\varepsilon) = \sigma^2$

Matrix Notation will be more easy to work with this model

Matrix Notation

- Response random variable: $\underset{\sim}{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ vector ($n \times 1$)
- Regression parameters: $\underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}$ vector ($p \times 1$)

- Explanatory variables (not Random) : Matrix $(n \times p)$

$$\underset{\sim}{X} = \begin{array}{c c} \begin{matrix} x_0 & x_1 & \cdots & x_{p-1} \end{matrix} & \\ \begin{bmatrix} 1 & x_{11} & \cdots & x_{1,p-1} \\ 1 & x_{21} & \cdots & x_{2,p-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{n,p-1} \end{bmatrix} & \begin{matrix} obj_1 \\ \vdots \\ \\ obj_n \end{matrix} \end{array}$$

called Design Matrix

- Error random variable (unobserved) $(n \times 1)$: $\underset{\sim}{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

Model with Matrix Notation:

$$\boxed{\underset{\sim}{Y} = \underset{\sim}{X}\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}}$$

Assumptions with Matrix Notation:

$E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ vector $(n \times 1)$; $Var(\underset{\sim}{\varepsilon}) = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & \cdots & \sigma^2 \end{bmatrix}$ Matrix $(n \times n)$

$$= \sigma^2 \underset{\sim}{I}_{n \times n}$$

$E(\underset{\sim}{Y})$ and $var(\underset{\sim}{Y})$ with Matrix notation

1. $E(\underset{\sim}{Y}) = E(\underset{\sim}{X}\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}) \underset{\text{Linear operator}}{=} E(\underset{\sim}{X}\,\underset{\sim}{\beta}) + E(\underset{\sim}{\varepsilon}) =$

$$= \underset{\sim}{X}\,\underset{\sim}{\beta} + \underset{\sim}{0} = \underset{\sim}{X}\,\underset{\sim}{\beta}$$

2. $Var(\underset{\sim}{Y}) = Var(\underbrace{\underset{\sim}{X}\,\underset{\sim}{\beta}}_{\text{not Random variables}} + \underset{\sim}{\varepsilon}) = Var(\underset{\sim}{\varepsilon}) = \sigma^2 \underset{\sim}{I}$

$var(\underset{\sim}{X}\,\underset{\sim}{\beta}) = \underset{\sim}{0}$

Assuming that $\underset{\sim}{\varepsilon} \sim N_n(\underset{\sim}{0}, \sigma^2 \underset{\sim}{I}) \Rightarrow$

$$\underset{\sim}{Y} \sim N_n(\underset{\sim}{X}\underset{\sim}{\beta}; \sigma^2 \underset{\sim}{I})$$

Design Matrix : $\underset{\sim}{X} = \begin{bmatrix} \boxed{1 \quad x_{11} \quad \cdots \quad x_{1,p-1}} \\ \vdots \quad \vdots \quad \vdots \\ 1 \quad x_{n1} \quad x_{n,p-1} \end{bmatrix}$ ← $\underset{\sim}{x}$ for de 1st object

$$E[Y \mid \underset{\sim}{x}_1] = \beta_0 + \beta_1 x_{11} + \cdots + \beta_{p-1} x_{1,p-1}$$

$(1, x_{11}, \ldots x_{1,p-1})$

but $\beta_0, \ldots, \beta_{p-1}$ are unknown ⇒ Estimation

$$\hat{E}[Y \mid \underset{\sim}{x}_1] = \hat{Y}_1 = \hat{\beta}_0 + \hat{\beta}_1 x_{11} + \cdots + \hat{\beta}_{p-1} x_{1,p-1}$$

Residual : $Y_1 - \hat{Y}_1 = e_1$ (or r_1)

Residuals : $Y_i - \hat{Y}_i = e_i$, $i = 1, \ldots, n$

Matrix Notation:

- residuals: vector $(n\times 1)$ $\underset{n\times 1}{\underset{\sim}{e}} = \begin{bmatrix} y_1-\hat{y}_1 \\ \vdots \\ y_n-\hat{y}_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \underset{\sim}{y} - \hat{\underset{\sim}{y}}$

where $\hat{\underset{\sim}{y}} = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} \underset{\sim}{x}_1^T \times \hat{\underset{\sim}{\beta}} \\ \vdots \\ \underset{\sim}{x}_n^T \hat{\underset{\sim}{\beta}} \end{bmatrix} = \underset{\sim}{X}\,\hat{\underset{\sim}{\beta}}$

$\underset{\sim}{x}_1 = \begin{bmatrix} x_{11} \\ \vdots \\ x_{n1} \end{bmatrix}$

$\underset{\sim}{e} = \underset{\sim}{y} - \hat{\underset{\sim}{y}} = \underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\beta}}$?

$\hat{y} = \underset{\sim}{X}\hat{\underset{\sim}{\beta}}$

Estimation of $\underset{\sim}{\beta}$ and σ^2 (Least square method)

Find the estimates that minimize the sum of square of residuals $= \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

Matrix Notation: $\underset{\sim}{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$; $\underset{\sim}{e}^T = [e_1, \dots, e_n]$

$$\underset{(1\times n)}{\underset{\sim}{e}^T}\ \underset{(n\times 1)}{\underset{\sim}{e}} = [e_1, \dots, e_n] \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} = e_1^2 + e_2^2 + \dots + e_n^2 = \sum_{i=1}^{n} e_i^2$$

SSE (Sum of square of residuals) =

$$= \sum_{i=1}^{n} e_i^2 = \underset{\sim}{e}^T \underset{\sim}{e} = (\underset{\sim}{y} - \hat{\underset{\sim}{y}})^T (\underset{\sim}{y} - \hat{\underset{\sim}{y}})$$

$$= (\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\beta}})^T (\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\beta}}) =$$

$$= (\underset{\sim}{y}^T - \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T)(\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\beta}}) =$$

$$= \underset{\sim}{y}^T \underset{\sim}{y} - \underset{\sim}{y}^T \underset{\sim}{X} \hat{\underset{\sim}{\beta}} - \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{y} + \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{X} \hat{\underset{\sim}{\beta}} =$$

$$= \underset{\sim}{y}^T \underset{\sim}{y} - 2 \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{y} + \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{X} \hat{\underset{\sim}{\beta}}$$

$$\frac{\partial\, SSE}{\partial \hat{\underset{\sim}{\beta}}} = 0 \iff -2 \underset{\sim}{X}^T \underset{\sim}{y} + 2 \underset{\sim}{X}^T \underset{\sim}{X} \hat{\underset{\sim}{\beta}} = 0$$

$$\iff \underset{\sim}{X}^T \underset{\sim}{y} = \underset{\sim}{X}^T \underset{\sim}{X} \hat{\underset{\sim}{\beta}} \iff$$

$$\iff \boxed{\hat{\underset{\sim}{\beta}} = (\underbrace{\underset{\sim}{X}^T \underset{\sim}{X}}_{})^{-1} \underset{\sim}{X}^T \underset{\sim}{y}} \quad \text{Solution}$$

$(\underset{\sim}{X}^T \underset{\sim}{X})^{-1}$ is called Matrix $\underset{\sim}{C}$

$$\underset{\sim}{C} = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \quad \therefore \quad \hat{\underset{\sim}{\beta}} = \underset{\sim}{C} \underset{\sim}{X}^T \underset{\sim}{y}$$

Matrix $\boxed{\underset{\sim}{C} = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1}}$

Matrix $\underset{\sim}{C}^{-1} = \underset{\sim}{X}^T \underset{\sim}{X}$, Let's have a Look

taking $\underline{\underline{p = 3}}$

$$\underset{\substack{\sim \\ (n \times 3)}}{X} = \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \end{array}$$

$$\underset{\sim}{C}^{-1} = \underset{\sim}{X}^T \underset{\sim}{X} = \underset{(3 \times n)}{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \end{bmatrix}} \cdot \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \iff$$

$$\underset{\sim}{C}^{-1} = \underset{\sim}{X}^T \underset{\sim}{X} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i2} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1} x_{i2} \\ \sum_{i=1}^{n} x_{i2} & \sum_{i=1}^{n} x_{i1} x_{i2} & \sum_{i=1}^{n} x_{i2}^2 \end{bmatrix}$$

$(n \times 3)$

3×3

$\underset{\sim}{C}^{-1}$ is symmetric $\Rightarrow \underset{\sim}{C}$ is symmetric

$\underset{\sim}{C}$ is a symmetric matrix $\quad \underset{\sim}{C}^T = \underset{\sim}{C}$

Fitted values: $\hat{E}[\underset{\sim}{Y}] = \hat{\underset{\sim}{Y}} = \hat{\mu}_{\underset{\sim}{Y}}$

$$\hat{\underset{\sim}{Y}} = \underset{\sim}{X}\hat{\underset{\sim}{\beta}} = \underset{\sim}{X}\underset{\sim}{C}\underset{\sim}{X}^T\underset{\sim}{Y} = \underbrace{\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T}_{H}\underset{\sim}{Y}$$

$$= \underset{\sim}{H}\underset{\sim}{Y}$$

$\underset{\sim}{H}$ - Hat Matrix

$\underset{\sim}{H}$ is a symmetric and Idempotent Matrix

$\underset{\sim}{H}^T = \underset{\sim}{H}$

$\underset{\sim}{H}\,\underset{\sim}{H} = \underset{\sim}{H}$

SSE using Matrix $\underset{\sim}{H}$

$\underset{\sim}{X}\hat{\underset{\sim}{\beta}} = \underset{\sim}{H}\underset{\sim}{Y}$

$$SSE = \underset{\sim}{Y}^T\underset{\sim}{Y} - \underset{\sim}{Y}^T\underset{\sim}{X}\hat{\underset{\sim}{\beta}} - \hat{\underset{\sim}{\beta}}^T\underset{\sim}{X}^T\underset{\sim}{Y} + \hat{\underset{\sim}{\beta}}^T\underset{\sim}{X}^T\underset{\sim}{X}\hat{\underset{\sim}{\beta}} =$$

$$= \underset{\sim}{Y}^T\underset{\sim}{Y} - \underset{\sim}{Y}^T\underset{\sim}{H}\underset{\sim}{Y} - \cancel{\underset{\sim}{Y}^T\underset{\sim}{H}^T\underset{\sim}{Y}} + \underset{\sim}{Y}^T\underbrace{\cancel{\underset{\sim}{H}^T}\underset{\sim}{H}}_{H}\underset{\sim}{Y}$$

$$= \underset{\sim}{Y}^T\underset{\sim}{Y} - \underset{\sim}{Y}^T\underset{\sim}{H}\underset{\sim}{Y} = \underset{\sim}{Y}^T\underset{\sim}{Y} - \underset{\sim}{Y}^T\underset{\sim}{X}\hat{\underset{\sim}{\beta}}$$

$$SSE = \underset{\sim}{Y}^T\underset{\sim}{Y} - \underset{\sim}{Y}^T\underset{\sim}{X}\hat{\underset{\sim}{\beta}}$$

$$SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n} \hat{y}_i^2 - n\bar{y}^2 = \hat{\underset{\sim}{y}}^T \hat{\underset{\sim}{y}} - n\bar{y}^2$$

(sum squares of the regression)

$$= \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \boxed{\underset{\sim}{X} \hat{\underset{\sim}{\beta}}} - n\bar{y}^2 =$$

$$= \hat{\underset{\sim}{\beta}}^T \underbrace{\underset{\sim}{X}^T \underset{\sim}{X} (\underset{\sim}{X}^T \underset{\sim}{X})^{-1}}_{\underset{\sim}{I}} \underset{\sim}{X}^T \underset{\sim}{y} - n\bar{y}^2$$

$$= \boxed{\hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{y} - n\bar{y}^2}$$

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \underset{\sim}{y}^T \underset{\sim}{y} - n\bar{y}^2$$

(sum squares of y's)

SST = SSR + SSE

- SST is easy to obtain
- SSR is also not difficult since sometimes we have $\underset{\sim}{X}^T \underset{\sim}{y}$ $\Rightarrow$ SSE = SST − SSR

$$\underset{\sim}{X}^T \underset{\sim}{y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{21} & x_{22} & \cdots & x_{n2} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_i x_{i1} y_i \\ \sum_i x_{i2} y_i \end{bmatrix}$$

$\boxed{P = 3}$

So, we have the LS estimator of $\underset{\sim}{\beta}$

$$\hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{y} = \underset{\sim}{C}\, \underset{\sim}{X}^T \underset{\sim}{y}$$

and as in the simple L.R. the estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{SSE}{n-p} = \frac{SST - SSR}{n-p} = MSE$$ (Mean Square Error)

Important quantities in Regression

- **SST** (total) $= \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2$

  $= \underset{\sim}{y}^T \underset{\sim}{y} - n\bar{y}^2$

  matrix notation

- **SSR** (regression) $= \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n} \hat{y}_i^2 - n\bar{y}^2$

  $\hat{y} = \underset{\sim}{X}\hat{\beta}$

  $\hat{\beta} = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{y}$

  $= \hat{\underset{\sim}{y}}^T \hat{\underset{\sim}{y}} - n\bar{y}^2 = \hat{\beta}^T \underset{\sim}{X}^T \underset{\sim}{X} \hat{\beta} - n\bar{y}^2$

  $= \hat{\underset{\sim}{\beta}}^T \underbrace{\underset{\sim}{X}^T \underset{\sim}{X} (\underset{\sim}{X}^T \underset{\sim}{X})^{-1}}_{\underset{\sim}{I}} \underset{\sim}{X}^T \underset{\sim}{y} - n\bar{y}^2 =$

  $= \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{y} - n\bar{y}^2$

$$\boxed{SST = SSR + SSE}$$

- **SSE** (residuals) $= SST - SSR = \underset{\sim}{y}^T \underset{\sim}{y} - \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{y}$

Properties of the estimators

$E(\hat{\underset{\sim}{\beta}}) = ?$

$$E(\hat{\underset{\sim}{\beta}}) = E\left((\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{y}\right) = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T E(\underset{\sim}{y}) =$$

$$= \underbrace{(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{X}}_{\underset{\sim}{I}} \underset{\sim}{\beta} = \underset{\sim}{\beta}$$

$\therefore\ E(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{\beta}$ (unbiased estimator)

$$Var(\hat{\underset{\sim}{\beta}}) = Var\left(\underbrace{(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T}_{\underset{\sim}{A}} \underset{\sim}{y}\right) =$$

$Var(\underset{\sim}{A}\underset{\sim}{X}) = \underset{\sim}{A}\, Var(\underset{\sim}{X})\, \underset{\sim}{A}^T$

generalization of $Var(aX) = a^2 Var(X)$

$$= \underset{\sim}{A}\, Var(\underset{\sim}{y})\, \underset{\sim}{A}^T = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \sigma^2 \underset{\sim}{I}\, \underset{\sim}{X} (\underset{\sim}{X}^T \underset{\sim}{X})^{-1}$$

$$= \sigma^2 \underbrace{(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{X}}_{\underset{\sim}{I}} (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} = \sigma^2 \underbrace{(\underset{\sim}{X}^T \underset{\sim}{X})^{-1}}_{\underset{\sim}{C}}$$

$$= \sigma^2 \underset{\sim}{C}$$

taking the k-th component of the vector $\hat{\underset{\sim}{\beta}}$ we have

$$\hat{\beta}_k \quad \text{and} \quad E(\hat{\beta}_k) = \beta_k$$

$$Var(\hat{\beta}_k) = \sigma^2 \underbrace{c_{k+1,k+1}}_{\substack{(k+1)\text{ element} \\ \text{of the main diagonal} \\ \text{of } \underset{\sim}{C}}}$$

But as σ^2 is unknow we need to estimate the variance of $\hat{\beta}_K$, using the estimator:

$$\hat{Var}(\hat{\beta}_K) = \hat{\sigma}^2 c_{K+1,K+1} \quad \text{where } \hat{\sigma}^2 = MSE$$

tests and confidence Intervals for the parameters of regressions (β's)

supposing that $\underset{\sim}{\varepsilon} \sim N_n(\underset{\sim}{0}, \sigma^2 \underset{\sim}{I}) \Rightarrow \underset{\sim}{Y} \sim N_n(\underset{\sim}{X}\underset{\sim}{\beta}, \sigma^2 \underset{\sim}{I})$

and $\hat{\underset{\sim}{\beta}} \sim N_p(\underset{\sim}{\beta}, \sigma^2 \underset{\sim}{C})$; $\hat{\beta}_K \sim N(\beta_K, \sigma^2 c_{K+1,K+1})$

Confidence Intervals and tests for Individuals slope Coefficients (β_K, $K = 0, \dots, p-1$)

Pivotal Quantity: $T = \dfrac{\hat{\beta}_K - \beta_K}{\sqrt{\hat{\sigma}^2 c_{K+1,K+1}}} \sim t_{(n-p)}$ $K = 0, \dots, p-1$

- $CI_{(1-\alpha)\times 100\%}(\beta_K) = \left[\hat{\beta}_K \pm t_{1-\alpha/2}(n-p)\sqrt{\hat{\sigma}^2 c_{K+1,K+1}}\right]$
- test of Individual coefficients (β's)

$H_0: \beta_K = 0$ vs $H_1: \begin{cases} \beta_K \neq 0 \text{ (two sided) } 1 \\ \beta_K < 0 \text{ (one sided Left tail) } 2 \\ \beta_K > 0 \text{ (one sided Right tail) } 3 \end{cases}$

(x_K Is not significant to explain $E[Y]$)

(under H_0)

test statistics : $T_0 = \dfrac{\hat{\beta}_k - 0}{\sqrt{\hat{\sigma}^2 c_{k+1,k+1}}} \sim t(n-p)$

rej H_0 if $\begin{cases} ① & |T_0| > a \\ ② & T_0 < b \\ ③ & T_0 > c \end{cases}$

Confidence Interval for the mean response

$E[Y \mid \underset{\sim}{x}_0] = \mu_{Y|\underset{\sim}{x}_0} = \beta_0 + \beta_1 x_{0,1} + \cdots + \beta_{p-1} x_{0,p-1}$

one obs. $\underset{\sim}{x}$

$= \underset{\sim}{x}_0^T \underset{\sim}{\beta}$

Estimator of $\mu_{Y|\underset{\sim}{x}_0}$ is $\hat{\mu}_{Y|\underset{\sim}{x}_0}$

$\hat{\mu}_{Y|\underset{\sim}{x}_0} = \hat{\beta}_0 + \cdots + \hat{\beta}_{p-1} x_{0,p-1} = \underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}$

$E[\hat{\mu}_{Y|\underset{\sim}{x}_0}] = \beta_0 + \cdots + \beta_{p-1} x_{0,p-1} = \mu_{Y|\underset{\sim}{x}_0}$ (unbiased)

$\mathrm{var}(\hat{\mu}_{Y|\underset{\sim}{x}_0}) = \mathrm{var}(\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}) = \underset{\sim}{x}_0^T \mathrm{var}(\hat{\underset{\sim}{\beta}}) \underset{\sim}{x}_0 =$

$= \underset{\sim}{x}_0^T \hat{\sigma}^2 \underset{\sim}{C} \underset{\sim}{x}_0 = \sigma^2 \underset{\sim}{x}_0^T \underset{\sim}{C} \underset{\sim}{x}_0$

Pivotal Quantity: $T = \dfrac{\hat{\mu}_{Y|\underset{\sim}{x}_0} - \mu_{Y|\underset{\sim}{x}_0}}{\sqrt{\hat{\sigma}^2 \underset{\sim}{x}_0^T \underset{\sim}{C} \underset{\sim}{x}_0}} \sim t(n-p)$

- $CI_{(1-\alpha)\times 100\%}(\mu_{Y|\underset{\sim}{x}_0}) = \left[\hat{\mu}_{Y|\underset{\sim}{x}_0} \pm t_{1-\alpha/2}(n-p)\sqrt{\hat{\sigma}^2 \underset{\sim}{x}_0^T \underset{\sim}{C} \underset{\sim}{x}_0} \right]$

Prediction Interval for a new observation of Y: Y_0

① $Y_0 = Y|\underset{\sim}{x}_0 = \underbrace{\underset{\sim}{x}_0^T \underset{\sim}{\beta}}_{\mu_{Y|\underset{\sim}{x}_0}} + \varepsilon = \mu_{Y|\underset{\sim}{x}_0} + \varepsilon$

② $\hat{Y}_0 = \hat{\mu}_{Y|\underset{\sim}{x}_0} = \underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}$

$$(\hat{Y}_0 - Y_0) = \underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}} - \underset{\sim}{x}_0^T \underset{\sim}{\beta} - \varepsilon$$

$$E(\hat{Y}_0 - Y_0) = E(\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}} - \underset{\sim}{x}_0^T \underset{\sim}{\beta} - \varepsilon) = \underset{\sim}{x}_0^T \beta - \underset{\sim}{x}_0^T \underset{\sim}{\beta} - 0 = 0$$

$E(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{\beta}$ and $E(\varepsilon) = 0$

$$Var(\hat{Y}_0 - Y_0) = Var(\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}} - \underbrace{\underset{\sim}{x}_0^T \underset{\sim}{\beta}}_{\text{constant}} - \varepsilon) = Var(\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}) + Var(\varepsilon)$$

$$= \underset{\sim}{x}_0^T Var(\hat{\underset{\sim}{\beta}}) \underset{\sim}{x}_0 + \sigma^2 = \sigma^2 \underset{\sim}{x}_0^T \underset{\sim}{C} \underset{\sim}{x}_0 + \sigma^2$$

$$= \sigma^2(1 + \underset{\sim}{x}_0^T \underset{\sim}{C} \underset{\sim}{x}_0)$$

Pivotal Quantity: $T = \dfrac{\hat{Y}_0 - Y_0}{\sqrt{\hat{\sigma}^2(1 + \underset{\sim}{x}_0^T \underset{\sim}{C} \underset{\sim}{x}_0)}} \sim t_{(n-p)}$

- $PI_{(1-\alpha)\times 100\%}(Y_0) = \left[\hat{Y}_0 \pm t_{1-\alpha/2}(n-p) \sqrt{\hat{\sigma}^2(1 + \underset{\sim}{x}_0^T \underset{\sim}{C} x_0)} \right]$

One test that is very Important is the overall test - test for significance of the regression model to the data set

Does the explanatory variables $(x_1, \ldots, x_{p-1})$ explain significantly the expected value of Y ?

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_{p-1} = 0$$

vs

$$H_1: \exists^1 \beta_k \neq 0, \quad k = 1, \ldots, p-1$$

Pivotal Quantity ??

to build the Pivotal Quantity we use the sum of the squares identity:

$SST = SSR + SSE$, where

$$SST = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = \underset{\sim}{Y}^T \underset{\sim}{Y} - n\bar{Y}^2$$

(Total)

$$SSR = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 = \underset{\sim}{\hat{Y}}^T \underset{\sim}{\hat{Y}} - n\bar{Y}^2 =$$

(Regression)

$$= \underset{\sim}{\hat{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y} - n\bar{Y}^2$$

$$SSE = \sum_{i=1} (y_i - \hat{y}_i)^2 = \underset{\sim}{e}^T \underset{\sim}{e} = \underset{\sim}{y}^T \underset{\sim}{y} - \underset{\sim}{y}^T \underset{\sim}{X} \underset{\sim}{\hat{\beta}}$$

(residuals)

And take the Mean squares as:

$$MST = \frac{SST}{n-1} \; ; \; MSR = \frac{SSR}{p-1} \; ; \; MSE = \frac{SSE}{n-p}$$

to buid the following table : ANOVA table

Sourse	SS (Sum squares)	d.f. (degrees of freedom)	MS (Mean squares)	F_ratio
Regression	SSR	$(p-1)$	MSR	$F_0 = \frac{MSR}{MSE}$
Residuals	SSE	$(n-p)$	MSE	
totae	SST	$(n-1)$	MST	

under Ho, $F_0 = \frac{MSR}{MSE}$ is small (if $\beta_1 = \beta_2 = \dots = \beta_{p-1} = 0$)

why?? taking the R^2 (coefficient of Determination)

$$R^2 = \frac{SSR}{SST} \quad \text{and} \quad \frac{SST}{SST} = \frac{(SSR + SSE)}{SST} = 1$$

$$\Leftrightarrow \frac{SSR}{SST} + \frac{SSE}{SST} = R^2 + \frac{SSE}{SST} \Rightarrow R^2 + \frac{SSE}{SST} = 1$$

$$\iff \frac{SSE}{SST} = 1 - R^2$$

And
$$F_0 = \frac{MSR}{MSE} = \frac{\frac{SSR}{p-1}}{\frac{SSE}{n-p}} = \frac{SSR \times (n-p)}{SSE(p-1)} = \frac{\frac{SSR}{SST}(n-p)}{\frac{SSE}{SST}(p-1)} = \frac{R^2(n-p)}{(1-R^2)(p-1)}$$

$R^2 \to 0$ $(\beta_1 = \beta_2 = \dots = \beta_{p-1} = 0) \Rightarrow F_0$ small

$R^2 \to 1$ $(\exists\, \hat{\beta}_k \neq 0) \Rightarrow F_0$ increases

Critical Region: Reject H_0 if $F_0 > c$

$F_0 \sim$?? Under H_0 is possible to prove that:

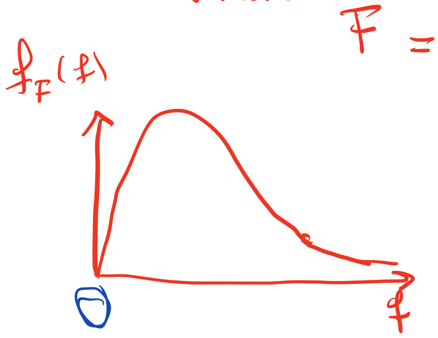
$$SSR \sim \chi^2_{(p-1)} \text{ and } SSE \sim \chi^2_{(n-p)}$$
$$SSR \perp\!\!\!\perp SSE$$

Result: $X \sim \chi^2_{(n)}$ and $Y \sim \chi^2_{(m)}$ and $X \perp\!\!\!\perp Y$

then $F = \left(\dfrac{X/n}{Y/m} \right) \sim F_{(n,m)}$

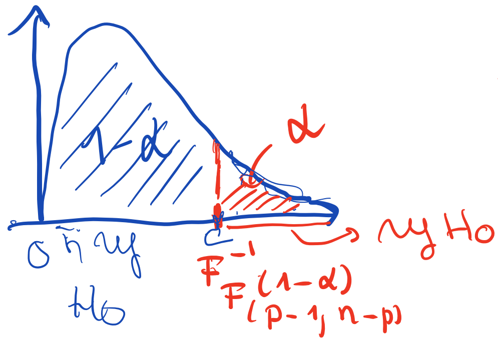
$\downarrow$

F- distribution



So, under H_0 $\quad F_0 = \dfrac{MSR}{MSE} \sim F_{\underbrace{(p-1,\, n-p)}_{d.f.}}$

For a fix significance level α, we rej H_0 if



$F_0 > c = F^{-1}_{F_{(p-1,\,n-p)}}(1-\alpha)$

(quantile $(1-\alpha)$ of the $F_{(p-1,\,n-p)}$ dist.)

How to consult the F_ quantiles table?

n_1 = numerator degrees of freedom
n_2 = denominator degrees of freedom

F Distribution Table (Percentiles)

n_1 = graus liberdade Numerador
n_2 = graus liberdade Denominador

	$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10
0.95	1	161.5	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9
0.975		647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.7
0.99		4052.2	4999.5	5403.4	5624.6	5763.7	5859.0	5928.4	5981.07	6022.4	6055.9
0.95	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
0.975		38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40
0.99		98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40
0.95	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
0.975		17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42
0.99		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23
0.95	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
0.975		12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84
0.99		21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55
0.95	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
0.975		10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62
0.99		16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05

Example: $F^{-1}_{F_{(3,4)}}(0.95) = 6.59$

$F^{-1}_{(3,4)}(0.05) = ??$

$F \sim F_{(n,m)}$

$X \sim \chi^2_{(n)}$

$Y \sim \chi^2_{(m)}$ $\quad X \perp\!\!\!\perp Y$

$$F = \frac{\frac{X}{n}}{\frac{Y}{m}} \; ; \; \frac{1}{F} = \frac{\frac{Y}{m}}{\frac{X}{n}} \sim F_{(m,n)}$$

number $x = F^{-1}_{F(n,m)}(\alpha)$; x is the α-quantile of F

$$P(F \le x) = \alpha \iff P\left(\frac{1}{F} \geq \frac{1}{x}\right) = \alpha \iff$$

$$1 - P\left(\frac{1}{F} \le \frac{1}{x}\right) = \alpha \iff P\left(\frac{1}{F} \le \frac{1}{x}\right) = 1 - \alpha$$

so $\frac{1}{x}$ is the $(1-\alpha)$-Quantile of an $F_{(m,n)}$

$$F^{-1}_{F(3,4)}(0.05) = \frac{1}{F^{-1}_{(4,3)}(0.95)} = \frac{1}{6.39}$$

If we reject $H_0: \beta_1 = \beta_2 = \cdots = \beta_{p-1} = 0$ We just know that at Least one of the explanatory variables is useful.

We can check if a subset of explanatory variables are or not useful

test F-Partial (Some x's are not need to explain the $E[Y]$??)

$H_0: \beta_1 = \beta_2 = \cdots = \beta_r = 0$ vs $H_1: \exists^1 \beta_k \neq 0, \; k = 1, \ldots, r$

$x_1, x_2, \ldots, x_r$ $(r < p-1)$ are not Important $\qquad (r < p-1)$

Pivotal Quantity ??

considering $\underset{\sim}{\beta}_1 = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_r \end{bmatrix}$ and $\underset{\sim}{\beta}_2 = \begin{bmatrix} \beta_{r+1} \\ \vdots \\ \beta_{p-1} \end{bmatrix}$

and defining $SSR(\underset{\sim}{\beta}_1 | \underset{\sim}{\beta}_2) = \underbrace{SSR(\underset{\sim}{\beta}_1, \underset{\sim}{\beta}_2)}_{\text{sum squares regression with all variables}} - \underbrace{SSR(\underset{\sim}{\beta}_2)}_{\text{sum squares reg. with variables } x_{k+1}, \ldots, x_{p-1}}$

Under H_0 $F_0 = \dfrac{SSR(\underset{\sim}{\beta}_1 | \underset{\sim}{\beta}_2)}{r\, MSE} \sim F_{(r, n-p)}$

and we will reject H_0 if $F_0 > F^{-1}_{F_{(r,n-p)}}(1-\alpha)$
for a fix significant Level α