Notation:

$$\underset{\sim}{Z} \sim N(\underset{\sim}{\mu}, I_n)$$

A is an idempotent matrix $\operatorname{rank}(A) = r$.

$$\underset{\sim}{Z}^T A \underset{\sim}{Z} \sim \chi^2_{r, \underset{\sim}{\mu}^T A \underset{\sim}{\mu}}.$$

$$E(\underset{\sim}{Z}^T A \underset{\sim}{Z}) = r + \underset{\sim}{\mu}^T A \underset{\sim}{\mu}.$$

If $\underset{\sim}{X} \sim N(\underset{\sim}{\mu}, \sigma^2 I_n)$

$$\frac{\underset{\sim}{X}}{\sigma} \sim N\left(\frac{\underset{\sim}{\mu}}{\sigma}, I_n\right)$$

$$\underset{\sim}{X}^T A \underset{\sim}{X} \stackrel{d}{=} \sigma^2 \left(\frac{\underset{\sim}{X}^T}{\sigma} A \frac{\underset{\sim}{X}}{\sigma}\right) \stackrel{d}{=} \sigma^2 \chi^2_{r, \frac{\underset{\sim}{\mu}^T A \underset{\sim}{\mu}}{\sigma^2}}.$$

$$\boxed{E(\underset{\sim}{X}^T A \underset{\sim}{X}) = \sigma^2 r + \underset{\sim}{\mu}^T A \underset{\sim}{\mu}}$$

We "abuse" the notation and write.

$$\underset{\sim}{X}^T A \underset{\sim}{X} \sim \sigma^2 \chi^2_{r, \frac{\underset{\sim}{\mu}^T A \underset{\sim}{\mu}}{\sigma^2}}.$$

$$\Leftrightarrow \frac{\underset{\sim}{X}^T A \underset{\sim}{X}}{\sigma^2} \sim \chi^2_{r, \frac{\underset{\sim}{\mu}^T A \underset{\sim}{\mu}}{\sigma^2}}.$$

Multiple Linear Regression.

Data set : $D = \{(y_i, \underset{\sim}{x_i}) \mid i = 1, 2, \cdots n\}$. i.e. n observations.

y_is are independent but NOT identically distributed.

We assume that $y_i \in \mathbb{R}$ and $\underset{\sim}{x_i} \in \mathbb{R}^k$

$$\underset{\sim}{Y} = \underset{n \times 1}{\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}} \quad \underset{\sim}{\epsilon} = \underset{n \times 1}{\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}} \quad \underset{\sim}{\beta} = \underset{(k+1) \times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}} \quad X = \begin{bmatrix} 1 & \underset{\sim}{x_1}^T \\ 1 & \underset{\sim}{x_2}^T \\ \vdots & \vdots \\ 1 & \underset{\sim}{x}_{n-1}^T \\ 1 & \underset{\sim}{x}_n^T \end{bmatrix} = \underset{n \times (k+1)}{\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & & x_{2k} \\ \vdots & & & & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}}$$

As a consequence. the mode in vector notation can be written as

$$\underset{n \times 1}{\underset{\sim}{Y}} = \underset{n \times (k+1)}{X} \underset{(k+1) \times 1}{\underset{\sim}{\beta}} + \underset{n \times 1}{\underset{\sim}{\epsilon}}.$$

The least squared condition to estimate $\underset{\sim}{\beta}$ is to minimize.

$$S(\underset{\sim}{\beta}) = (\underset{\sim}{Y} - X\underset{\sim}{\beta})^T(\underset{\sim}{Y} - X\underset{\sim}{\beta}).$$

The least squared estimate of β is $\hat{\underset{\sim}{\beta}} = (X^T X)^{-1} X^T \underset{\sim}{Y}$.

assuming $|X^T X| \neq 0$.

- $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\epsilon}$; $\underset{\sim}{\epsilon} \sim N(\underset{\sim}{0}, \sigma^2 I_n)$.

  $\Rightarrow \underset{\sim}{Y} \sim N(X\underset{\sim}{\beta}, \sigma^2 I_n)$.

- $\hat{\underset{\sim}{\beta}} = (X^T X)^{-1} X^T \underset{\sim}{Y}$. where X is non-stochastic.

  $\hat{\underset{\sim}{\beta}} \sim N\left((X^T X)^{-1} X^T (X\underset{\sim}{\beta}), (X^T X)^{-1} X^T (\sigma^2 I_n) X (X^T X)^{-1}\right)$

  $= N\left(\underset{\sim}{\beta}, \sigma^2 (X^T X)^{-1} (X^T X) (X^T X)^{-1}\right)$

  $= N\left(\underset{\sim}{\beta}; \sigma^2 (X^T X)^{-1}\right)$

  Note 1. $E(\hat{\beta}) = \beta$. Hence it is an unbiased estimator.

- $\hat{y}_0 = (1 \; \underset{\sim}{x}_0^T)\, \hat{\underset{\sim}{\beta}} = \begin{pmatrix} 1 \\ \underset{\sim}{x}_0 \end{pmatrix}^T \hat{\underset{\sim}{\beta}}$

- $\hat{\underset{\sim}{Y}} = X\hat{\underset{\sim}{\beta}} = \underline{X (X^T X)^{-1} X^T}\, \underset{\sim}{Y}$.

  prediction vector.

  Note 2 $X(X^T X)^{-1} X^T$ is the orthogonal projection matrix of $\mathcal{C}(X) = \mathcal{C}(XX^T)$.

- $\underset{\sim}{e} = (\underset{\sim}{Y} - \hat{\underset{\sim}{Y}}) = (I_n - P_X)\underset{\sim}{Y}$.

  estimated error vector.

  where $P_X = X(X^T X)^{-1} X^T$.

  $P_X = P_X^T$, $P_X^2 = P_X$.

  Symmetric. Idempotent.

- $Cov(\hat{\underset{\sim}{Y}}, \underset{\sim}{e}) = 0_{\text{matrix}}$.

- $\hat{Y} = P_X \underset{\sim}{Y} \sim N(X\beta, P_X \sigma^2)$
- $\underset{\sim}{e} = (I_n - P_X)\underset{\sim}{Y} \sim N(\underset{\sim}{0}, (I_n - P_X)\sigma^2)$.
- $\text{cov}(\hat{\underset{\sim}{Y}}, \underset{\sim}{e}) = 0$.

$\underset{\sim}{Y} \sim N(X\beta, \sigma^2 I_n)$

$P_X X = X$

$P_X^2 = P_X, \quad P_X^T = P_X$.

Analysis of Variance. (ANOVA)

⊛ Whether the model building is at all necessary or not??

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$

vs H_1: H_0 is not true. (atleast one β_i $i \geq 1$ is non zero)

If H_0 is true then. $Y_i \underset{\sim}{\text{iid}} (\beta_0, \sigma^2)$

Total sum of square (SST) $= \sum_{i=1}^{n} (y_i - \bar{y})^2$

$= \underset{\sim}{y}^T \left(I_n - \frac{1}{n} \underset{\sim}{1} \underset{\sim}{1}^T\right) \underset{\sim}{Y}$.

$SST \sim \sigma^2 \chi^2_{n-1,\ ncp = \frac{(X\beta)^T}{\sigma}\left(I_n - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\frac{(X\beta)}{\sigma}}$

$= \sigma^2 \chi^2_{n-1,\ \frac{\underset{\sim}{\beta}^T}{\sigma}\left(X^T\left[I_n - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right]X\right)\frac{\underset{\sim}{\beta}}{\sigma}}$.

$\boxed{\text{Under } H_0\ ncp = 0}$

$\left(I_n - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)$. is

→ symmetric.

→ idempotent.

→ with rank $(n-1)$

$$SST = \underset{\sim}{y}^T\left(I_n - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\underset{\sim}{y}$$

$$= \underset{\sim}{y}^T\left(I_n - P_x + P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\underset{\sim}{y}$$

$$= \underset{\sim}{y}^T(I_n - P_x)\underset{\sim}{y} + \underset{\sim}{y}^T\left(P_x - \frac{1}{n}11^T\right)\underset{\sim}{y}$$

$$= \underset{\sim}{y}^T(I_n - P_x)^2\underset{\sim}{y} + \underset{\sim}{y}^T\left(P_x - \frac{1}{n}11^T\right)\underset{\sim}{y}$$

$$= \underset{\sim}{y}^T(I_n - P_x)^T(I_n - P_x)\underset{\sim}{y} + \underset{\sim}{y}^T\left(P_x - \frac{1}{n}11^T\right)\underset{\sim}{y}$$

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

$$= SS\,Error + SS\,Model$$

$$= \sum_{i=1}^n(\hat{y}_i - y_i)^2 + \sum_{i=1}^n(\hat{y}_i - \bar{y})^2$$

$$Q = A_1 - A_2.$$

$$(I_n - P_x)^T = (I_n - P_x)$$

$$(I_n - P_x)^2 = (I_n - P_x)$$

$$\sum_{i=1}^n(\hat{y}_i - y_i)^2$$

$$= \underset{\sim}{y}^T(I_n - P_x)\underset{\sim}{y}$$

= Error Sum of squares

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

$$= \underset{\sim}{y}\left(P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\underset{\sim}{y}$$

= Model Sum of Squares

or Regression sum of squares..

5

- $\underset{\sim}{Y}^T \left(I_n - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\underset{\sim}{Y} = SST \sim \sigma^2 \chi^2_{n-1}$, $ncp = \underset{\sim}{\beta}^T X^T \left(I_n - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right) X \underset{\sim}{\beta}$.

$\underline{SST}$ = SS Error + SS Model.

$= \underset{\sim}{Y}^T (I_n - P_x)\underset{\sim}{Y} + \underset{\sim}{Y}^T \left(P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\underset{\sim}{Y}$.

if. $|X^T X| \neq 0$.

- $\underline{\text{SS Error}} = \underset{\sim}{Y}^T (I_n - P_x)\underset{\sim}{Y} \sim \sigma^2 \chi^2_{n-k-1,\ ncp = 0}$

$ncp\left(\sigma^2 \left(\frac{Y}{\sigma}\right)^T (I_n - P_x)\left(\frac{Y}{\sigma}\right)\right) = \sigma^2 \left(\frac{X\beta}{\sigma}\right)^T (I_n - P_x)\left(\frac{X\beta}{\sigma}\right) = \underset{\sim}{\beta}^T X^T (I_n - P_x) X \underset{\sim}{\beta} = 0$ $\boxed{P_x X = X}$

$\Rightarrow E(\text{SS Error}) = \sigma^2 (n-k-1)$.

$\Rightarrow E\left(\frac{\text{SS Error}}{n-k-1}\right) = \sigma^2$

The unbiased estimator of σ^2 is $\frac{\text{SS Error}}{n-k-1}$ = Mean squared error = MSError.

- $\underline{\text{SS Model}}$: $\underset{\sim}{Y}^T \left(P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\underset{\sim}{Y} \sim \cancel{\sigma^2} \chi^2_{k}$, $ncp = \underset{\sim}{\beta}^T X^T \left(P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right) X \underset{\sim}{\beta} / \sigma^2$

$ncp\left(\cancel{\sigma^2} \left(\frac{Y}{\sigma}\right)^T \left(P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\left(\frac{Y}{\sigma}\right)\right) = \cancel{\sigma^2}\left(\frac{X\beta}{\sigma}\right)^T \left(P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right)\left(\frac{X\beta}{\sigma}\right) = \frac{\beta^T X^T \left(P_x - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right) X \underset{\sim}{\beta}}{\sigma^2}$.

$$\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \hline \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \cdots \\ \underset{\sim}{\beta}_R \end{pmatrix} \qquad X = \left(\underset{\sim}{1} \;\vdots\; X_R\right)$$

(6)

$$ncp\left(\frac{SSModel}{\sigma^2}\right) = \underset{\sim}{\beta}^T X^T \left(P_X - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\right) X \underset{\sim}{\beta}.$$

$\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \underset{\sim}{\beta}_R \end{pmatrix}$ $X = [1 \,\vdots\, X_R]$

$P_X X = X.$

$$= \frac{1}{\sigma^2} \underset{\sim}{\beta}^T \left(X^T P_X X - \frac{1}{n} X^T \underset{\sim}{1}\underset{\sim}{1}^T X\right)\underset{\sim}{\beta}.$$

$$= \frac{1}{\sigma^2} \underset{\sim}{\beta}^T \left((X^T X) - \frac{1}{n} X^T \underset{\sim}{1}\underset{\sim}{1}^T X\right)\underset{\sim}{\beta}.$$

Note: $ncp(SST) = ncp(SSM)$.

$$= \frac{1}{\sigma^2} \underset{\sim}{\beta}^T \left([\underset{\sim}{1}\ X_R]^T [\underset{\sim}{1}\ X_R] - \frac{1}{n} \underline{[\underset{\sim}{1}\ X_R]^T \underset{\sim}{1}}\ \underline{\underset{\sim}{1}^T [\underset{\sim}{1}\ X_R]} \right)\underset{\sim}{\beta}.$$

$$= \frac{1}{\sigma^2} \underset{\sim}{\beta}^T \left(\left[\begin{array}{c|c} n & 1^T X_R \\ \hline X_R^T 1 & X_R^T X_R \end{array}\right] - \frac{1}{n} [n \,\vdots\, 1^T X_R]^T [n \,\vdots\, 1^T X_R] \right)\underset{\sim}{\beta}$$

$$= \frac{1}{\sigma^2} \underset{\sim}{\beta}^T \left(\left[\begin{array}{c|c} n & 1^T X_R \\ \hline X_R^T 1 & X_R^T X_R \end{array}\right] - \left[\begin{array}{c|c} n & 1^T X_R \\ \hline X_R^T 1 & \frac{1}{n} X_R^T 1 1^T X_R \end{array}\right] \right)\underset{\sim}{\beta}.$$

$$= \frac{1}{\sigma^2} \underset{\sim}{\beta} \left(\begin{array}{c|c} 0 & \underset{\sim}{0}^T \\ \hline \underset{\sim}{0} & X_R^T X_R - \frac{1}{n}(1^T X_R)^T (1^T X_R) \end{array} \right)\underset{\sim}{\beta}.$$

(7)

$$\text{ncp}\left(\frac{SS\,Model}{\sigma^2}\right) = \frac{\beta^T X^T}{\sigma}\left(P_X - \frac{1}{n}\underset{\sim}{1}\,\underset{\sim}{1}^T\right)\frac{X\beta}{\sigma}$$

$$= \frac{1}{\sigma^2}\beta^T \left(\begin{array}{c|c} 0 & \underset{\sim}{0}^T \\ \hline \underset{\sim}{0} & X_R^T X_R - \frac{1}{n}(1^T X_R)^T(1^T X_R) \end{array}\right)\beta \qquad \beta = \begin{pmatrix}\beta_0 \\ \hline \underset{\sim}{\beta}_R\end{pmatrix}$$

$$= \frac{1}{\sigma^2}\underset{\sim}{\beta}_R^T\left(X_R^T X_R - \frac{1}{n}(1^T X_R)^T(1^T X_R)\right)\underset{\sim}{\beta}_R = \frac{\lambda}{\sigma^2} \quad |X_R^T X_R| \neq 0$$

$\text{ncp}\,(SS\,Model) = 0$ if and only if $\underset{\sim}{\beta}_R = \underset{\sim}{0}$.

H.W.. Show that $\left(X_R^T X_R - \frac{1}{n}(1^T X_R)^T(1^T X_R)\right)$ can be ~~writ~~ written as a matrix $Z^T Z$ ~~where~~ with centered vector of X_R.

$$\sum (x_i - \bar{x})^2$$

$$\sum (x_{ri} - \bar{x}_r)^2$$

$$SST \sim \sigma^2 \chi^2_{n-1},\ \text{ncp} = \lambda/\sigma^2$$

$$SSM \sim \sigma^2 \chi^2_{k},\ \text{ncp} = \lambda/\sigma^2$$

$$SSE \sim \sigma^2 \chi^2_{n-k-1},\ \text{ncp} = 0.$$

independent.

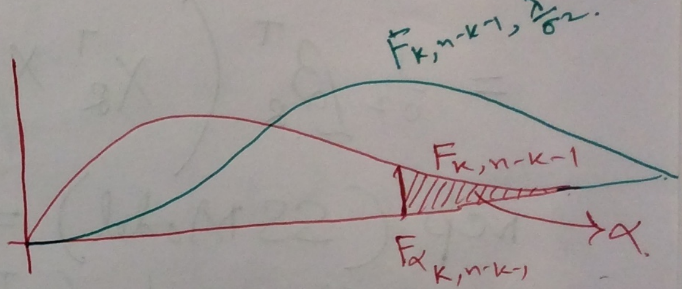
$$\frac{SSM}{\sigma^2} \sim \chi^2_{k},\ \lambda/\sigma^2.$$

$$\frac{SSE}{\sigma^2} \sim \chi^2_{n-k-1},\ 0.$$

$$F = \frac{SS\ Model / k}{SS\ Error / (n-k-1)} \sim F_{k,\, n-k-1,\, \left(\frac{\lambda}{\sigma^2}\right),\, 0}$$

If H_0 is true: $\underset{\sim}{\beta}_R = \underset{\sim}{0} \iff \beta_1 = \beta_2 = \cdots = \beta_k = 0$

$$F \sim F_{k,\, n-k-1}$$



If H_0 is not true then

$$F \sim F_{k,\, n-k-1,\, \frac{\lambda}{\sigma^2}} \quad \left\{ \frac{\lambda}{\sigma^2} > 0 \right.$$

We will reject the H_0 in favour of $H_1 \left(\underset{\sim}{\beta}_R \neq \underset{\sim}{0}\right)$ if

$$F_{obsd.} > F_{\alpha,\, k,\, n-k-1}$$

at level-α

This analysis is known as ANOVA.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \cdots\cdots + \beta_K x_{Ki} + \epsilon_i$$

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\epsilon} \qquad \underset{\sim}{\epsilon} \sim N(\underset{\sim}{0}, \sigma I_n).$$

- The impact of x_1 and x_2 are same on Y or NOT.
- There is no impact of x_5 on Y or NOT.

We want to test H_0: X_5 has no impact on Y.
vs H_1: X_5 has some impact on Y.

We can redefine the hypothesis as.

$$H_0: \beta_5 = 0$$
$$H_1: \beta_5 \neq 0.$$

We know $\underset{\sim}{\hat{\beta}} \sim N\left(\underset{\sim}{\beta}, \sigma^2 (X^T X)^{-1}\right)$

$$\equiv N\left(\underset{\sim}{\beta}, \sigma^2 C\right).$$

$$\hat{\beta}_5 \sim N\left(\beta_5, \sigma^2 c_{55}\right)$$

$$\underset{\sim}{l}^T = (0\,0\,0\,0\,0\,1\,0 \cdots 0).$$

$$\hat{\beta}_5 = \underset{\sim}{l}^T \hat{\beta} \sim N\left(\underset{\sim}{l}^T \underset{\sim}{\beta}, \underset{\sim}{l}^T (\sigma^2 C) \underset{\sim}{l}\right).$$

Let us denote.

$$C = (X^T X)^{-1}_{(k+1)\times(k+1)}$$

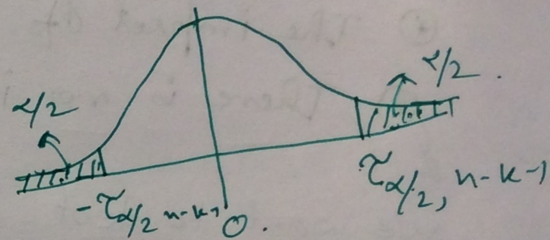
index are from. $(0, 1, 2 \cdots K.)$

$$\begin{array}{c|c} & \hat{\beta}_0 \quad \hat{\beta}_1 \quad \cdots \quad \hat{\beta}_k \\ \hline \hat{\beta}_0 & c_{00} \quad c_{11} \quad \cdots \quad c_{0k} \\ \hat{\beta}_1 & \\ \vdots & \\ \hat{\beta}_k & c_{k0} \quad \cdots \quad c_{kk} \end{array}$$

$$c_{ij}\sigma^2 = \mathrm{cov}(\hat{\beta}_i, \hat{\beta}_j)$$

σ^2 is unknown

Hence $\hat{\sigma}^2 = \dfrac{SS\ Error}{n-k-1}$. $\qquad \dfrac{SS\ Error}{\sigma^2} \sim \chi^2_{n-k-1,\,0}$.



$$\hat{\beta}_5 \sim N(\beta_5, \sigma^2 C_{55}).$$

$$\Rightarrow \frac{\hat{\beta}_5 - \beta_5}{\sqrt{\sigma^2 C_{55}}} \sim N(0,1).$$

$$\Rightarrow \frac{\hat{\beta}_5 - 0}{\sqrt{\sigma^2 C_{55}}} \sim N(0,1) \text{ under } H_0.$$

$$\Rightarrow T = \frac{\hat{\beta}_5}{\sqrt{\hat{\sigma}^2 C_{55}}} \sim t_{n-k-1,\,0} \text{ under } H_0.$$

We reject H_0 in favour of H_1 if

$$|T_{observed}| > \tau_{\alpha/2,\,n-k-1}$$

at level α.

(*) $H_0: \beta_3 = 2\beta_4$ vs. $H_1: \beta_3 > 2\beta_4$ $\Big\} \iff$ $H_0: \beta_3 - 2\beta_4 = 0$, $H_1: \beta_3 - 2\beta_4 > 0$

$$\hat{\beta} \sim N(\beta, \sigma^2 C) \quad \Rightarrow \quad \ell^T \hat{\beta} \sim N(\ell^T \beta, \sigma^2 \ell^T C \ell)$$

define $\ell^T = (0,0,0,1,-2,0\cdots 0)$

$$\ell^T \hat{\beta} \sim N\left(\beta_3 - 2\beta_4,\ \sigma^2 (c_{33} + 4c_{44} - 4c_{34})\right)$$

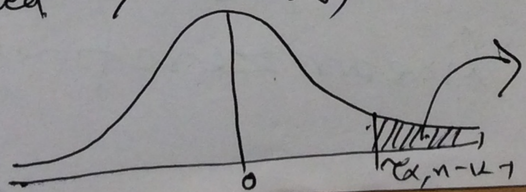
$$\Rightarrow \frac{\ell^T\hat{\beta} - \ell^T\beta}{\sqrt{\sigma^2 \ell^T C \ell}} \sim N(0,1)$$

$$\Rightarrow T = \frac{\ell^T\hat{\beta} - 0}{\sqrt{\hat{\sigma}^2 \ell^T C \ell}} \sim t_{n-k-1,0}$$

under H_0.

We reject H_0 in favor of H_1 at level α

if $T_{observed} > \tau_{\alpha, n-k-1}$.



$$(1, -2, 0)\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i x_j$$

$$= a_{11}\, 1\cdot 1 + a_{22}(-2)(-2) + a_{12}\, 1(-2) + a_{12}\, 1(-2)$$

$$= a_{11} + 4a_{22} - 4a_{12}.$$

$\alpha = P_{H_0}(\text{Rejecting } H_0)$

= Prob of rejecting H_0 when H_0 is true

= P(Type-I error)

- **Linear parametric function: (LPF).**

A linear parametric function of $\underset{\sim}{\beta}$ is a linear combination of it. i.e. $\underset{\sim}{p}^T \underset{\sim}{\beta} = \sum_{i=0}^{k+1} p_i \beta_i \quad \forall \underset{\sim}{\beta}, \underset{\sim}{p} \in \mathbb{R}^{k+1}$.

(*) If $\sum_{i=0}^{k} p_i = 0$ then the LPF is known a contrast.

- **Linear unbiased estimator: (LUE)**

If $E(\underset{\sim}{l}^T \underset{\sim}{y}) = \underset{\sim}{p}^T \underset{\sim}{\beta} \quad \forall \underset{\sim}{\beta} \in \mathbb{R}^{k+1}$.

then $\underset{\sim}{l}^T \underset{\sim}{y}$ is considered as ~~an~~ a linear unbiased estimator of $\underset{\sim}{p}^T \underset{\sim}{\beta}$. If there exists a LUE of $\underset{\sim}{p}^T \beta$ then we call $\underset{\sim}{p}^T \underset{\sim}{\beta}$ is estimable.

- **Linear Zero function. (LZF)**

If $E(\underset{\sim}{l}^T \underset{\sim}{y}) = 0. \quad \forall \underset{\sim}{\beta} \in \mathbb{R}^{k+1}$. then.

$\underset{\sim}{l}^T \underset{\sim}{y}$ is known as. a linear zero function.

(*) We can express any linear zero function LZF in the form of.

$$\underset{\sim}{m}^T (I - P_x) \underset{\sim}{Y}$$

for some $\underset{\sim}{m}$ and orthogonal projection matrix P_x of $\mathcal{C}(X)$.

By dfn. $E(\underset{\sim}{l}^T \underset{\sim}{Y}) = 0 \quad \forall \underset{\sim}{\beta} \in \mathbb{R}^{k+1}$. for LZF.

$\Leftrightarrow \quad \underset{\sim}{l}^T X \underset{\sim}{\beta} = 0 \quad \forall \underset{\sim}{\beta} \in \mathbb{R}^{k+1}$.

$\Leftrightarrow \quad \underset{\sim}{l}^T X = \underset{\sim}{0}^T$.

$\Rightarrow \quad \underset{\sim}{l} \in (\mathcal{C}(X))^{\perp}$

$\quad \underset{\sim}{l} \in \mathcal{C}(I_n - P_x)$.

$\Rightarrow \quad \underset{\sim}{l} = (I_n - P_x) \underset{\sim}{m} \qquad m \in \mathbb{R}^n$.

$\Rightarrow \exists \ \underset{\sim}{m}$ such that. $\underset{\sim}{l}^T X = \underset{\sim}{m}^T (I_n - P_x) X = \underset{\sim}{0}^T$.

$\Rightarrow$

Every LZF can be written as. $\boxed{\underset{\sim}{m}^T (I_n - P_x) \underset{\sim}{Y}}$.

$E\left[\underset{\sim}{m}^T (I_n - P_x) \underset{\sim}{Y}\right] = \underset{\sim}{0}^T \underset{\sim}{\beta} = \underline{0 \text{ scalar}}.$

1. Show that $\underset{\sim}{l}^T \underset{\sim}{y}$ is a LUE of $\underset{\sim}{p}^T \underset{\sim}{\beta}$ iff $X^T \underset{\sim}{l} = \underset{\sim}{p}$.

<u>where</u> $\underset{\sim}{Y} \sim N(X\underset{\sim}{\beta}, \sigma^2 I_n)$

$$E(\underset{\sim}{l}^T \underset{\sim}{y}) = \underset{\sim}{p}^T \underset{\sim}{\beta} \quad \forall \underset{\sim}{\beta} \in \mathbb{R}^{k+1}.$$

$$\Leftrightarrow \quad \underset{\sim}{l}^T X \underset{\sim}{\beta} = \underset{\sim}{p}^T \underset{\sim}{\beta} \quad \forall \underset{\sim}{\beta} \in \mathbb{R}^{k+1}$$

$$\Leftrightarrow \quad \underset{\sim}{l}^T X = \underset{\sim}{p}^T. \qquad \text{even if } |X^T X| = 0.$$

HW.

2. $\underset{\sim}{l}^T \underset{\sim}{y}$ is a LZF iff $\underset{\sim}{X}^T \underset{\sim}{l} = \underset{\sim}{0}$.

HW

3. An 'iff' condition for $\underset{\sim}{p}^T \underset{\sim}{\beta}$ to be estimable is $\underset{\sim}{p} \in \mathcal{C}(X^T)$.

Best Linear Unbiased Estimator (BLUE)

Dfn. Let us consider the set of unbiased estimators of a LPF $\underset{\sim}{p}^T \underset{\sim}{\beta}$ as

$$U = \left\{ \underset{\sim}{l}^T \underset{\sim}{y} \;\middle|\; E(\underset{\sim}{l}^T \underset{\sim}{y}) = \underset{\sim}{p}^T \underset{\sim}{\beta} \;\; \forall \underset{\sim}{\beta} \in \mathbb{R}^{k+1} \right\}$$

We call $\underset{\sim}{l_0}^T \underset{\sim}{y}$ the ~~best~~ BLUE of $\underset{\sim}{p}^T \underset{\sim}{\beta}$ if.

$$V(\underset{\sim}{l_0}^T \underset{\sim}{y}) < V(\underset{\sim}{l}^T \underset{\sim}{y}) \quad \forall \; \underset{\sim}{l_0} \neq l \text{ and } \underset{\sim}{l_0}^T y \in U$$

$l_0^T y \in U$.

Theorem: A linear function of $\underset{\sim}{y}$ is the BLUE of its expectation iff it is uncorrelated with all linear zero functions (LZF).

$$E(\underset{\sim}{l}^T \underset{\sim}{y}) = \underset{\sim}{p}^T \underset{\sim}{\beta}$$

$$\text{iff } \operatorname{cov}(\underset{\sim}{l}^T \underset{\sim}{y}, \underset{\sim}{z}^T y) = 0.$$

$$\left\{ \begin{array}{l} E(\underset{\sim}{z}^T \underset{\sim}{y}) = 0 \\ \forall \; \underset{\sim}{\beta} \in \mathbb{R}^{k+1}. \end{array} \right.$$

Consider $\underset{\sim}{l}^T \underset{\sim}{Y} = \underset{\sim}{l}^T I_n \underset{\sim}{Y}$

$= \underset{\sim}{l}^T P_X \underset{\sim}{Y} + \underset{\sim}{l}^T (I_n - P_X) \underset{\sim}{Y}$.

Observe

LUE ① $E(\underset{\sim}{l}^T P_X \underset{\sim}{Y}) = l^T P_X (X\beta) = \underset{\sim}{l}^T X \underset{\sim}{\beta} = \underset{\sim}{b}^T \underset{\sim}{\beta}$

$\boxed{\text{if } E(\underset{\sim}{l}^T \underset{\sim}{Y}) = \underset{\sim}{b}^T \underset{\sim}{\beta}}$

LZF ② $E(l^T (I_n - P_X) \underset{\sim}{Y}) = 0 = l^T [(I_n - P_X) X] \underset{\sim}{\beta} = \underset{\sim}{l}^T 0 \underset{\sim}{\beta}$

③ $\mathrm{cov}(\underset{\sim}{l}^T P_X \underset{\sim}{Y}, \underset{\sim}{l}^T (I_n - P_X) \underset{\sim}{Y}) = 0$

$\Rightarrow$ $l^T P_X Y$ is the BLUE of $\underset{\sim}{b}^T \beta = E(\underset{\sim}{l}^T \underset{\sim}{Y})$

(*) If we have $(\underset{\sim}{l}^T \underset{\sim}{Y})$ a LUE of $\underset{\sim}{b}^T \underset{\sim}{\beta}$ we can convert it to the BLUE of $\underset{\sim}{b}^T \underset{\sim}{\beta}$ as

$\underset{\sim}{l}^T P_X \underset{\sim}{Y}$

$\boxed{\text{even if } |X^T X| = 0}$

Ex 4. Every estimable LPF has unique BLUE.

Let us assume there exists more than one ~~BLUE~~ BLUE of LPF $\underset{\sim}{p}^T\underset{\sim}{\beta}$ as $\underset{\sim}{l}_1^T\underset{\sim}{y}$ and $\underset{\sim}{l}_2^T\underset{\sim}{y}$.

$$\underset{\sim}{l}_1^T\underset{\sim}{y} = \underset{\sim}{l}_2^T\underset{\sim}{y} + \underset{\sim}{l}_1^T\underset{\sim}{y} - \underset{\sim}{l}_2^T\underset{\sim}{y}.$$

$$= \underset{\sim}{l}_2^T\underset{\sim}{y} + (\underset{\sim}{l}_1^T - \underset{\sim}{l}_2^T)\underset{\sim}{y}.$$

Note:
$(\underset{\sim}{l}_1 - \underset{\sim}{l}_2)^T\underset{\sim}{y}$ is LZF.
$\underset{\sim}{l}_1^T\underset{\sim}{y}$ is BLUE.

$$Var(\underset{\sim}{l}_1^T\underset{\sim}{y}) = V(\underset{\sim}{l}_2^T\underset{\sim}{y}) + V((\underset{\sim}{l}_1^T - \underset{\sim}{l}_2^T)\underset{\sim}{y}) + cov(\underset{\sim}{l}_2^T\underset{\sim}{y}, (\underset{\sim}{l}_1 - \underset{\sim}{l}_2)^T\underset{\sim}{y})$$

$$= V(\underset{\sim}{l}_2^T\underset{\sim}{y}) + V[(\underset{\sim}{l}_1^T - \underset{\sim}{l}_2^T)\underset{\sim}{y}]$$

$E(\underset{\sim}{l}_1^T\underset{\sim}{y}) = \underset{\sim}{p}^T\underset{\sim}{\beta} = E(\underset{\sim}{l}_2^T\underset{\sim}{y})$
as LUE.
$V(\underset{\sim}{l}_1^T\underset{\sim}{y}) = V(\underset{\sim}{l}_2^T\underset{\sim}{y})$
as BLUE.

$$\Rightarrow V[(\underset{\sim}{l}_1^T - \underset{\sim}{l}_2^T)\underset{\sim}{y}] = 0$$

also $$E[(\underset{\sim}{l}_1^T - \underset{\sim}{l}_2^T)\underset{\sim}{y}] = 0$$

$(\underset{\sim}{l}_1 - \underset{\sim}{l}_2)^T\underset{\sim}{y} = 0$ with prob 1.

$\underset{\sim}{l}_1^T\underset{\sim}{y} = \underset{\sim}{l}_2^T\underset{\sim}{y}$ with prob 1.