Date: 31.07.23

* Simple Linear Regression :-

y : dependent variable.

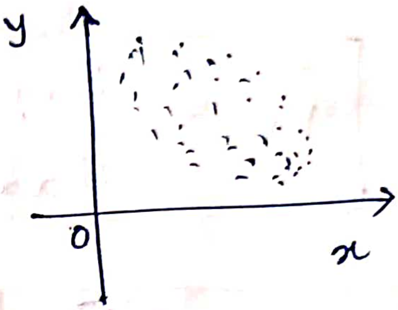
x : Regressor or independent variable.

Suppose,

$(y_1, x_1), (y_2, x_2), \ldots \ldots (y_n, x_n)$

y : shear strength

x : age of proppelant.



As x increases, y decreases.

$$f(y) = \beta_0 + \beta_1 f(x) + \varepsilon$$

$$\downarrow \qquad \downarrow$$

$$y^* = \beta_0 + \beta_1 x^* + \varepsilon$$

$$y_i^* = \beta_0 + \beta_1 x_i + \varepsilon_i \quad ; \quad i = 1(1)n.$$

$$\underset{\sim}{\beta} = (\beta_0, \beta_1)^T.$$

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = S(\underset{\sim}{\beta}).$$

Assumption: (Gauss Markov Assumption).

$$\left.\begin{array}{l} * \; E(\varepsilon_i) = 0 \\ * \; var(\varepsilon_i) = \sigma^2 \end{array}\right\} \forall i$$

$*$ $cov(\varepsilon_i, \varepsilon_j) = 0$ if $i \neq j$

$$\min_{\underset{\sim}{\beta} \in \mathbb{R}^2} S(\underset{\sim}{\beta})$$

$$\frac{\partial S(\underset{\sim}{\beta})}{\partial \beta_0} = 2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-1) = 0$$

$$\frac{\partial S(\underset{\sim}{\beta})}{\partial \beta_1} = 2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-x_i) = 0.$$

$$\boxed{\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}} \; ; \quad \text{where,}$$

$$S_{xy} = \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})$$

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 .$$

$$\boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}}$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i .$$

here,

$\hat{\beta}_1, \hat{\beta}_0$ are **least square estimators.**

Fitted Model $\longrightarrow$ $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$.

$E(\hat{\beta}_0) = ?$

$E(\hat{\beta}_1) = ?$

$$E(\hat{\beta}_1) = E\left[\frac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}\right]$$

$$\mathbb{E}(\hat{\beta}_1) = \sum_{i=1}^{n} c_i \, \mathbb{E}(y_i - \bar{y})$$

$$\bar{y} = \frac{1}{n} \Sigma y_i$$

$$= \frac{n\beta_0 + \beta_1 \Sigma x_i}{n}$$

$$= \beta_0 + \beta_1 \bar{x}.$$

$$E(y_i) = \beta_0 + \beta_1 x_i$$

$$E(\bar{y}) = \beta_0 + \beta_1 \bar{x}$$

$$\therefore \mathbb{E}(y_i - \bar{y}) = \beta_1 (x_i - \bar{x}).$$

$$\therefore \mathbb{E}(\hat{\beta}_1) = \sum c_i \, \mathbb{E}(y_i - \bar{y}) = \beta_1 \sum_{i=1}^{n} c_i (x_i - \bar{x}).$$

$$= \beta_1.$$

$$\frac{\beta_1 \Sigma \cancel{(x_i - \bar{x})^2}}{\Sigma \cancel{(x_i - \bar{x})^2}}$$

Hence, $\hat{\beta}_1$ is unbiased.

Now,

Show $\mathbb{E}(\hat{\beta}_0) = \beta_0$.

$$E(\hat{\beta}_0) = E(\bar{y} - \hat{\beta}_1 \bar{x}) = E(\beta_0 + \beta_1 \bar{x} - \hat{\beta}_1 \bar{x})$$

$$= \beta_0 + \beta_1 \bar{x} - E(\hat{\beta}_1) \bar{x}$$

$$= \beta_0 + \beta_1 \bar{x} - \beta_1 \bar{x} = \beta_0 \text{ . proved}$$

Hence, $\hat{\beta}_0$ is also unbiased.

Date 02.08.23

Consider,

Simple linear regression model:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \; ; \quad \varepsilon_i \sim E(\varepsilon_i) = 0, \quad Var(\varepsilon_i) = \sigma^2. \qquad \text{----} (*)$$

LSEs

$$\boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}}$$

$$\boxed{\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}}$$

where,

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

$$var(\hat{\beta}_1) = var\left(\sum c_i (y_i - \bar{y}) \right)$$

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

$$var(\hat{\beta}_1) = \sum c_i^2 \left[var(y_i) + var(\bar{y}) - 2\,cov(y_i, \bar{y}) \right]$$

$$= \sum c_i^2 \left[\sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n} \right]$$

$$= \sum c_i^2 \, \sigma^2 \left[1 - \frac{1}{n} \right].$$

Find $var(\hat{\beta}_0) = ?$ (Home work) suppose, $\varepsilon_i \sim N(0, \sigma^2)$ in ⊛.

we'll use MLE,

Likelihood, ~~then~~

$$L(\underset{\sim}{y}, \underset{\sim}{\beta}) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{\sum (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}}$$

Taking log both sides,

$$\log L = -\frac{n}{2}\log\sigma^2 - \frac{\sum (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2} + c.$$

$$\frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^4}$$

$$0 \overset{set}{=} -\frac{n}{\cancel{2}\cancel{\sigma^2}} + \frac{\sum (y_i - \beta_0 - \beta_1 x_i)^2}{\cancel{2}\cancel{\sigma^4}\,\sigma^2}$$

$$\boxed{\hat{\sigma}^2_{mle} = \frac{\sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{n}}$$

Multiple Linear Regression.

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} + \cdots + \beta_k x_{1k} + \varepsilon_1.$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \beta_3 x_{23} + \cdots + \beta_k x_{2k} + \varepsilon_2$$

$$\vdots$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots \beta_k x_{ik} + \varepsilon_i$$

$$\vdots$$

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3} + \cdots + \beta_k x_{nk} + \varepsilon_n..$$

where,

$$\underset{\sim}{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{bmatrix}_{n\times 1} ; \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & & & & \\ \vdots & & & & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} , \quad \underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}$$

$$\underset{n\times 1}{\underset{\sim}{y}} = \underset{n\times p}{\underset{\sim}{X}}\,\underset{p\times 1}{\underset{\sim}{\beta}} + \underset{n\times 1}{\underset{\sim}{\varepsilon}} , \quad p = k+1.$$

$$p < n.$$

<u>LSE</u>: $S(\beta) = \underset{\sim}{\varepsilon}^T \underset{\sim}{\varepsilon} = (y - X\beta)^T (y - X\beta) = S(\underset{\sim}{\beta})$.

$$\min_{\underset{\sim}{\beta} \in \mathbb{R}^p} \underset{\sim}{\varepsilon}^T \underset{\sim}{\varepsilon} = S(\underset{\sim}{\beta})$$

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

Results:

$$\boxed{\begin{array}{l} \text{(1) Let } \alpha = \underset{\sim}{y}^T A \underset{\sim}{x} \\[1em] \dfrac{\partial \alpha}{\partial \underset{\sim}{x}} = \underset{\sim}{y}^T A \quad ; \quad \dfrac{\partial \alpha}{\partial \underset{\sim}{y}} = \underset{\sim}{x}^T A^T \\[1em] \text{(2). } \alpha = \underset{\sim}{x}^T A \underset{\sim}{x} \\[1em] \dfrac{\partial \alpha}{\partial \underset{\sim}{x}} = \underset{\sim}{x}^T (A + A^T). \end{array}}$$

using above,

$$\frac{\partial s(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}} = -\underset{\sim}{y}^T X - \underset{\sim}{y}^T X + \beta^T (X^T X + X^T X) = 0$$

assuming X is of full rank.

$$\boxed{\hat{\beta}_{LSE} = \underset{p \times p}{(X^T X)^{-1}} \; \underset{p \times n}{X^T} \; \underset{n \times 1}{y}}$$

$1\ 2 .\ 3\ 4 \quad \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} = 1^2 + 2^2 + 3^2 + 4^2 \geq 0$

$$\frac{\partial s(\underset{\sim}{\beta})}{\partial \beta^T \partial \beta} = |(X^T X)|$$

Show $X^T X$ is positive definite.

$z^T X^T X z \geq 0 \ \forall \ z \in \mathbb{R}^p$

$(Xz)^T (Xz) = v^T v$ (let)

$= \sum v_i^2 \geq 0.$

$var(AY)$

$= A\,var(Y)\,A^T$

$\hat{\beta} = (X^TX)^{-1}X^TY$

$E(\hat{\beta}) = (X^TX)^{-1}X^T E(Y)$

$= (X^TX)^{-1}X^T(X\beta + \varepsilon)$

$= (X^TX)^{-1}X^TX\beta + 0$

$= \underset{\sim}{\beta}.$

$var(\underset{\sim}{\varepsilon}) = \sigma^2 I_{n\times n}.$

$Var(\hat{\beta}) = (X^TX)^{-1} X^T (\sigma^2 I) X (X^TX)^{-1}$

$= \sigma^2 (X^TX)^{-1}.$

use,

$cov(Y, X)$

$= E[(Y - E(Y))(X - E(X))^T]$

Date
04.08.23

$$\hat{\sigma}^2_{MLE} = \frac{(Y - X\hat{\beta})^T (Y - X\hat{\beta})}{n}$$

$$= \left[\frac{Y^T H Y}{n}\right]$$

where

① $H^T = H$

② $H^2 = H$

$$Y - X\hat{\beta} = Y - \underbrace{X(X^TX)^{-1}X^T}_{P_X \to \text{Projection matrix.}} Y$$

$$= \underbrace{\left[I - X(X^TX)^{-1}X^T\right]}_{H.} Y$$

$$H^2 = (I - X(X^TX)^{-1}X^T)(I - X(X^TX)^{-1}X^T)$$

$$= I - X(X^TX)^{-1}X^T - \cancel{X(X^TX)^{-1}X^T} + \cancel{[X(X^TX)^{-1}X^T][X(X^TX)^{-1}X^T]}$$

$$= I - X(X^TX)^{-1}X^T$$

$$= H.$$

Result: Z is a R.V with $E(Z) = \mu$.

$\mathrm{var}(Z) = \Sigma$

$$E(Z^T A Z) = \mathrm{trace}(A\Sigma) + \mu^T A \mu.$$

using this

$$E\left[\frac{Y^T H Y}{n}\right] = \frac{1}{n}\left[\mathrm{tr}(H\sigma^2 I)\right] + \beta^T X^T H X \beta$$

$$= \frac{1}{n}(\mathrm{tr}(H\sigma^2 I)) + 0 \quad \left\{ \because \begin{array}{l} HX = 0 \\ X - X = 0. \end{array} \right.$$

$$= \sigma^2 \mathrm{tr}(H)/n$$

$$\left\{\begin{array}{l}\text{using,}\\ tr(AB) = tr(BA)\end{array}\right.$$

$$= \sigma^2 \left[tr(I_{n\times n} - tr(X(X^TX)^{-1}X^T))\right]$$

$$= \sigma^2 \left[n - tr((X^TX)^{-1}X^TX)\right]$$

$$= \sigma^2 (n - tr(I_{p\times p})) = \sigma^2 (n-p).$$

$$E\left[\frac{Y^T H Y}{n}\right] = \frac{\sigma^2 (n-p)}{n}.$$

$$E\left[\frac{Y^T H Y}{(n-p)}\right] = \sigma^2 \quad \left\{ \hat{\sigma}^2 = \frac{Y^T H Y}{n-p} = \frac{SS_E}{n-p} \right.$$

$$= MS_E.$$

where,

$Y^T H Y$ is error term of squares (SS_E).

Gauss-Markov's Theorem :-
(G-M)

under G-M Assumptions,

$\hat{\beta}_{MLE/LSE} = (X^TX)^{-1}X^Ty$. is ~~BES~~ BLUE

(Best Linear Unbiased Estimator). of β.

This class should consists of all estimators of this type (Linear) : $C = \{Ay + b_0 : A, b_0 \in \mathbb{R}^p\}$

Best means, $Var(l^T \hat{\beta}_{LSE}) \le Var(l^T \tilde{\beta})$

$\forall \ \tilde{\beta} \in C.$
$\forall \ l \in \mathbb{R}^p$

Proof: $\tilde{\beta} \in C$.

$\tilde{\beta} \in C$, $(X^T X)^{-1} X^T +$

$\underbrace{A - (X^T X)^{-1} X^T}_{B}$

$$\tilde{\beta} = AY + b_0$$

$$= \left((X^T X)^{-1} X^T + B \right) Y + b_0 .$$

As, $\tilde{\beta}$ is unbiased,

$$\begin{cases} b_0 = 0 \\ BX\beta = 0 \quad \Rightarrow BX = 0 . \end{cases}$$

iff $\beta + BX\beta + b_0 = (I + BX)\beta + b_0$

$\Rightarrow I + BX = I$

$\Rightarrow BX = 0$

$b_0 = 0$.

Now,

$$Var(\tilde{\beta}) = var\left(\left[(X^T X)^{-1} X^T + B \right] Y \right)$$

$$= \sigma^2 \left[(X^T X)^{-1} X^T + B \right] \left[X (X^T X)^{-1} + B^T \right]$$

$$= \sigma^2 \left[(X^T X)^{-1} + BX (X^T X)^{-1} + (X^T X)^{-1} X^T B^T + BB^T \right]$$

$$= \sigma^2 \left[(X^T X)^{-1} + BB^T \right] = \sigma^2 (X^T X)^{-1} + \sigma^2 BB^T .$$

$$var(\tilde{\beta}) = var(\hat{\beta}_{LSE}) + \sigma^2 BB^T .$$

$$\text{var}(l^T\tilde{\beta}) = \text{Var}(l^T\hat{\beta}_{LSE}) + \sigma^2 \underbrace{l^T BB^T l}_{\geq 0}.$$

$$\Rightarrow l^T \text{var}(\tilde{\beta}) l$$

$$\text{var}(l^T\hat{\beta}_{LSE}) \leq \text{var}(l^T\tilde{\beta})$$

$$\text{var}(l^T\hat{\beta}_{LSE}) < \text{var}(l^T\tilde{\beta}) \quad ; \text{ for some } l.$$

Date 07. Aug, 23

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \quad S_{xy} = \Sigma(y_i - \bar{y})(x_i - \bar{x})$$

$$= \Sigma(x_i - \bar{x}) y_i - \Sigma \bar{y}(x_i - \bar{x})$$

$$\bar{y}\,\Sigma(x_i - \bar{x})$$

$$\bar{y}\,\Sigma(n\bar{x} - n\bar{x})$$

$$= 0.$$

$$S_{xx} = \Sigma(x_i - \bar{x})^2.$$

$$V(\hat{\beta}_1) = \Sigma c_i^2 \text{var}(y_i) + \sum_i \sum_{j \neq i} c_i c_j \text{cov}(y_i, y_j)$$

$$= \sigma^2 \Sigma c_i^2$$

$$= \sigma^2 \Sigma \frac{(x_i - \bar{x})^2}{(S_{xx})^2}.$$

$$= \sigma^2 \frac{S_{xx}}{S_{xx}^2} = \frac{\sigma^2}{S_{xx}}.$$

$\text{cov}(y_i, y_j) = 0$ as $(y_1, y_2, \dots, y_n)$ are independent.

$$\therefore \boxed{\text{var}(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}}$$

$$\text{var}(\hat{\beta}_0) = \text{var}(\bar{y}) + \bar{x}^2 \text{var}(\hat{\beta}_1) - 2\bar{x}\,\text{cov}(\bar{y}, \hat{\beta}_1).$$

Now, $\text{cov}(\bar{y}, \hat{\beta}_1) = \mathbb{E}[\{\bar{y} - E(\bar{y})\}\{\hat{\beta}_1 - E(\hat{\beta}_1)\}]$

$$= E[\bar{\epsilon}(\sum_{i=1}^n c_i y_i - \beta_1)] = \frac{1}{n} E[(\Sigma\epsilon_i)(\beta_0 \Sigma c_i + \beta_1 \Sigma c_i x_i + \Sigma c_i \epsilon_i) - \beta_1 \Sigma \epsilon_i]$$

$$= \frac{1}{n}[0 + 0 + 0 + 0] = 0.$$

$$\therefore \text{var}(\hat{\beta}_0) = \frac{\sigma^2}{n} + \bar{x}^2 \frac{\sigma^2}{S_{xx}} + 0 = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)$$

Multi-Variate Normal Distribution (MVN):

Let, $\underset{\sim}{u} = (u_1, u_2, \ldots, u_n)^T$

where, $u_i \sim N(0,1)$; $\forall\, i = 1(1)n$.

IID

Let $X_{r\times n}$, $\underset{\sim}{\mu}_{r\times 1}$.

$\underset{\sim}{Z} = X\underset{\sim}{u} + \underset{\sim}{\mu}$ → X is of full rank.

$var(\underset{\sim}{Z}) = XX^T = \Sigma_{r\times r}$

* Characteristic funⁿ: (CF)

$$\phi_Z(t) = \mathbb{E}\left[\exp\left(it^T\underset{\sim}{Z}\right)\right]$$

↳ for multivariate case

$$\phi_x(t) = \mathbb{E}\left[\exp(itx)\right]$$

↳ for univariate case.

First let's find CF of $\underset{\sim}{u} \sim N_{n\times n}(\underset{\sim}{0}, I_{n\times n})$.

$$\phi_{\underset{\sim}{u}}(t) = \mathbb{E}\left[\exp(it^T\underset{\sim}{u})\right]$$

$$e^{it_1u_1}e^{it_2u_2}\ldots e^{it_nu_n} = \mathbb{E}\left[e^{i\sum_{j=1}^{n} t_ju_j}\right]$$

$$= \prod_{j=1}^{n} E\left(e^{it_ju_j}\right)$$

$$\mathbb{E}\left(e^{it_j u_j}\right)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-u_j^2/2} e^{i t_j u_j} du_j$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[u_j^2 - 2i t_j u_j\right\}} du_j$$

$$= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left[u_j - it_j\right]^2 - t_j^2/2\right] du_j$$

$$= e^{-t_j^2/2} \cdot \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(u_j - it_j)^2} du_j}_{=1 \text{ where } \sigma^2 = 1}.$$

$$= e^{-t_j^2/2}$$

$$\phi_{\underset{\sim}{u}}(\underline{t}) = \prod_{j=1}^{n} e^{-t_j^2/2} = e^{-t^T t/2}$$

where, $\left\{ \underset{\sim}{u} \sim N(\underset{\sim}{0}, I_{n\times n}). \right.$

$$\underset{\sim}{z} = X\underset{\sim}{u} + \underset{\sim}{\mu}.$$

$$\phi_{\underset{\sim}{z}}(t) = E\left[\exp\{it^T z\}\right]$$

$$= E\left[\exp\{it^T(X\underset{\sim}{u} + \underset{\sim}{\mu})\}\right]$$

$$= e^{it^T \cdot \underset{\sim}{\mu}} \mathbb{E}\left[\exp\{it^T X \underset{\sim}{u}\}\right]$$

8.

$$= e^{it^T \underset{\sim}{\mu}} \; \phi_{\underset{\sim}{\mu}}(\cancel{X^T t})$$

$$= e^{it\underset{\sim}{\mu}} \; e^{-\frac{1}{2}(t^T x)(t^T)}$$

$$= e^{it^T \underset{\sim}{\mu}} \; \phi_{\underset{\sim}{t}}(X^T t)$$

$$= e^{it^T \underset{\sim}{\mu}} \; e^{-\frac{1}{2}(X^T t)^T (X^T t)}$$

$$= e^{it^T \mu} \; e^{-\frac{1}{2} t^T \underbrace{X X^T}_{\Sigma} t}$$

$$= e^{it^T \underset{\sim}{\mu} - \frac{1}{2} t^T \Sigma t}.$$

$$f(\underset{\sim}{z}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2} (z-\mu)^T \Sigma^{-1} (z-\mu) \right\}$$

$$\int_{-\infty}^{\infty} f(z)\, dz = 1.$$

Let's define, $\quad w = \Sigma^{-1/2}(z-\mu)$

$$(z-\mu)^T \Sigma^{-1} (z-\mu) = w^T w.$$

$$\frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \int e^{-\frac{1}{2} w^T w} |\Sigma|^{1/2} \, dw$$

$$= \frac{1}{(2\pi)^{n/2}} \int e^{-\frac{1}{2} w^T w} \, d\underline{w} \quad \left\{ \begin{array}{l} \left| \dfrac{\partial w_i}{\partial z_j} \right| = |\Sigma^{-1/2}| . \\ |\Sigma^{1/2}| \, d\underset{\sim}{w} = d\underset{\sim}{z} \\ (\text{H.W.}) \end{array} \right.$$

$$= \frac{1}{(2\pi)^{n/2}} \int \exp \frac{-\sum w_i^2}{2} \, dw_1 \, dw_2 \cdots dw_n .$$

$$= 1 .$$

$$\phi_{\underset{\sim}{z}}(t) = \mathbb{E}\left[e^{it^T z}\right]$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} f(\underset{\sim}{z}) \, dz$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(w^T \Sigma^{-1} w)\right\} \cdots \exp\left\{it^T(w + \underset{\sim}{\mu})\right\} dw_1 \, dw_2 \cdots dw_n$$

$\left\{ \text{put } w = z - \underset{\sim}{\mu} . \right.$

(Exercise).

$$= \exp\left(it^T \underset{\sim}{\mu} - \frac{1}{2} t^T \Sigma t\right) \quad (\text{H.W}) .$$

Date
09. Aug. 23

Linear estimation and Algebra)
(From R.B. Bapat

Theorem: Let $\underset{\sim}{y} \sim N(0, I_n)$ and let A be a symmetric $n \times n$ matrix. Then $\underset{\sim}{y}^T A \underset{\sim}{y}$ has a chi-square disn. with r-degrees of freedom if and only if A is idempotent and $\text{Rank}(A) = r$.

Proof: if $A^2 = A$ and $\text{Rank}(A) = r$.

$$A = P^T D P = P^T \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} P.$$

$$\underset{\sim}{Z} = P\underset{\sim}{y}, \quad \underset{\sim}{Z} \sim N(0, I$$

$$\underset{\sim}{y}^T A \underset{\sim}{y} = \underset{\sim}{y}^T P^T D P y = Z^T D Z = \sum_{i=1}^{r} Z_i^2$$

$$\sim \chi^2_r.$$

If we take
$A = X(X^T X)^{-1} X^T$
then $A^2 = A$

Conversely, let $\underset{\sim}{y} A \underset{\sim}{y} \sim \chi^2_r$.

Since, A is symmetric $\Rightarrow A = \underset{\text{orthogonal}}{P^T} \, \text{dia}(\underset{\text{eigen values of } A.}{\lambda_1, \ldots, \lambda_n}) P$

$$Z = P\underset{\sim}{y} \sim N(\underset{\sim}{0}, I)$$

Let's find characteristic function (CF) of $\underset{\sim}{y}^T A \underset{\sim}{y}$.

$$\phi(t) = \mathbb{E}\left[e^{it y^T A y}\right] = \mathbb{E}\left[e^{it z^T D z}\right]$$

$$= \mathbb{E}\left[e^{it \sum_{j=1}^{n} \lambda_j z_j^2}\right]$$

$$= \prod_{j=1}^{n} \mathbb{E}\left(e^{it\lambda_j z_j^2}\right)$$

$$= \prod_{j=1}^{n} (1 - 2it\lambda_j)^{-1/2} \quad \text{(Result)}$$

$$= (1-2it)^{-r/2}$$

$(1-2it)^r$

$\therefore$ roots are $\left(\frac{1}{2i}\right)$

$\left(\frac{1}{2i}\right)$ ... r-times.

$$(1-2it)^r = \prod_{j=1}^{n} (1-2i\lambda_j t)$$

$$\lambda_j = \begin{cases} 1 & \text{for } j = 1(1)r. \\ 0 & : \text{o/w} \end{cases}$$

<u>Theorem:</u> Let $\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}, \Sigma)$ and suppose that $\underset{\sim}{y}$, $\underset{\sim}{\mu}$, and Σ are conformally partitioned then $\underset{\sim}{y}_1$ and $\underset{\sim}{y}_2$ are independent iff $\Sigma_{12} = 0$.

$$\underset{\sim}{y} = \begin{pmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right]$$

Proof :

Independent $\Rightarrow$ $\text{cov}(\underset{\sim}{y}_1, \underset{\sim}{y}_2) = \Sigma_{12} = 0$.

For converse part :

characteristic funn of $\underset{\sim}{y}$.

$$\phi_{\underset{\sim}{y}}(t) = \exp\left[i t^T \underset{\sim}{\mu} - \frac{1}{2} t^T \Sigma t\right]$$

$$= \exp\left[i\,(t_1^T \; t_2^T)\begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix} - \frac{1}{2}(t_1^T \; t_2^T)\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}\right]$$

$$= \exp\left[i t_1^T \underset{\sim}{\mu}_1 + i t_2^T \underset{\sim}{\mu}_2 - \frac{1}{2} t_1^T \Sigma_{11} t_1 - \frac{1}{2} t_2^T \Sigma_{22} t_2\right]$$

$$= \exp\left[i t_1^T \mu_1 - \frac{1}{2} t_1^T \Sigma_{11} t_1\right] \exp\left[i t_2^T \mu_2\, t_2^T \Sigma_{22} t_2\right]$$

$$= \phi_{\underset{\sim}{y}_1}(t_1)\,\phi_{\underset{\sim}{y}_2}(t_2).$$

Theorem: $\underset{\sim}{y} \sim N(\underset{\sim}{\mu}, \sigma^2 I)$

Let A and B are matrices such that $AB^T = 0$ then Ay and By are idependent.

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\;\; {}^{\shortparallel}y_1 \quad\quad\quad\quad\quad {}^{\shortparallel}y_2$

Proof: $\begin{bmatrix} Ay \\ By \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} y$

$\text{Cov}(Ay, By) = \sigma^2 AB^T = 0$.

Date
11.08.23

Theorem: Let $\underset{\sim}{y} \sim N(\underset{\sim}{0}, I_n)$ and let A_1 and A_2 be symmetric idempotent matrices. Then $\underset{\sim}{y}^T A_1 \underset{\sim}{y}$ and $\underset{\sim}{y}^T A_2 \underset{\sim}{y}$ are independent if and only if $A_1 A_2 = 0$.

Proof: $A_1 A_2 = 0$, $A_1 \underset{\sim}{y}$ & $A_2 \underset{\sim}{y}$ are ~~idempotent~~ independent.

$y^T A_1 y = y^T A_1 A_1 y = (A_1 y)^T (A_1 y)$.

$y^T A_2 y = (A_2 y)^T (A_2 y)$.

Conversely: Let $y^T A_1 y$ and $y^T A_2 y$ are independent.

$y^T A_1 y + y^T A_2 y = y^T (A_1 + A_2) y \sim \chi^2_{r(A_1 + A_2)}$

Where,

$r(A_1+A_2)$ denotes the rank of (A_1+A_2).

$(A_1+A_2)(A_1+A_2) = A_1+A_2 \rightarrow A_1+A_2$ is idempotent.

$A_1 + A_2A_1 + A_1A_2 + A_2 = A_1+A_2$

$\Rightarrow A_2A_1 + A_1A_2 = 0$ --- (*)

(*) $\times A_2$

$A_1A_2 + A_2A_1 = 0 \times A_2$.

$A_2 \times A_1A_2 + A_2A_1A_2 = 0$

$A_2A_1A_2 + A_2A_1A_2 = 0$

$\Rightarrow A_2A_1A_2 = 0$. use this result to prove the theorem.

Theorem: Let $\underset{\sim}{y} \sim N(0; I_n)$. Let $A^T = A$, $A^2 = A$. Let $l \in \mathbb{R}^n$, $l \neq \underset{\sim}{0}$.

Then, y^TAy & l^Ty are independent iff $Al = 0$.

Proof: WLOG, $\|l\| = 1. = l^Tl$

$B = ll^T$

First let y^TAy & l^Ty are independent

$y^T A y$, $y^T l l^T y = y^T B y$.

we can say,

$AB = 0 \Rightarrow A l l^T = 0 \quad \times l$

$A l l^T l = 0$

$\Rightarrow Al = 0$;

conversely,

if $Al = 0 \Rightarrow Ay$ & $l^T y$ are independent.

$\Rightarrow y^T A y = (Ay)^T (Ay)$.

$\Rightarrow y^T A y$ and $l^T y$ are independent

(∵ Becoz, Ay & $l^T y$ are ind. and $(Ay)^T (Ay)$ is funⁿ. of Ay then $(Ay)^T (Ay)$ also independent ∴)

* Cochran's Theorem (Matrix version) : Let $A_1, A_2, \ldots, A_k$ be $n \times n$ matrices with $\sum_{i=1}^{k} A_i = I$, then following conditions are equivalent :-

(i) $\sum_{i=1}^{k} R(A_i) = n$.

(ii) $A_i^2 = A_i \quad \forall \ i = 1(1)k$.

(iii) $A_i A_j = 0 \quad \forall \ i \neq j$

Proof: (i) $\Rightarrow$ (iii).

Let $\underset{n\times r_i}{A_i} = B_i \quad C_i$

$$\underset{n\times r_i}{B_i} \quad \underset{r_i\times n}{C_i}$$

$$\sum A_i = I$$

$$\Rightarrow \sum B_i C_i = I$$

$$\Rightarrow \begin{bmatrix} \underset{n\times r_i}{B_1} & B_2 & \cdots & B_K \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_K \end{bmatrix} = I_{n\times n}$$

$$\sum_{i=1}^{K} R(A_i) = n \quad \Rightarrow \sum_{i=1}^{K} r_i = n.$$

If $AB = I$
$\Rightarrow BA = I$
in Prob Set (1).

$$\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_K \end{bmatrix} \begin{bmatrix} B_1 & B_2 & \cdots & B_K \end{bmatrix} = I_{n\times n}$$

$$\begin{bmatrix} C_1B_1 & C_1B_2 & \cdots & C_1B_K \\ C_2B_1 & C_2B_2 & \cdots & C_2B_K \\ \vdots & \vdots & \ddots & \vdots \\ C_KB_1 & C_KB_2 & \cdots & C_KB_K \end{bmatrix} = \begin{bmatrix} I_1 & & & 0 \\ & I_2 & & \\ & & I_3 & \\ 0 & & & \ddots \\ & & & & I_K \end{bmatrix}$$

$$C_i B_j' = 0 \quad \forall\, i \neq j$$

$$A_i A_j = B_i' (C_i' B_j) C_j$$

$$= 0. \quad \because 0$$

$(iii) \Rightarrow (ii)$ since $\sum_{i=1}^{k} A_i = I$

$$A_j \left(\sum_{i=1}^{k} A_i \right) = A_j$$

$$\Rightarrow \sum_{i=1}^{k} A_i A_j + A_j^2 = A_j$$

$$\Rightarrow A_j^2 = A_j$$

$(ii) \Rightarrow (i)$.

$\text{Rank}(A_i) = tr(A_i)$.

Now,

$$\sum \text{Rank}(A_i) = \sum tr(A_i)$$

$$= tr(\sum A_i)$$

$$= tr(I) = n.$$

* Cochran Theorem (For quadratic forms):

Let $u_1, u_2, \ldots u_N \overset{IID}{\sim} N(0, I)$

Let $B^{(1)}, B^{(2)}, \ldots, B^{(k)}$ be symmetric matrices,

Define, $r_i = \text{Rank}(B^{(i)}_{N\times N})$ & $Q_i = U^T B^{(i)}_{N\times N} U$.

where, $U^T = (u_1, u_2, \ldots u_N)$.

If $\sum_{i=1}^{k} B^{(i)} = I$

then following are equivalent :

(i) $r_1 + r_2 + \cdots + r_k = N$.

(ii) Q_i are indep.

(iii) Q_i are chi-square RV with r_i degrees of freedom.

(Testing of Hypothesis)

<u>Example</u>: consider the model,

$$\overset{Y}{\log F} = \overset{\beta_0}{\log c} - \overset{\beta_1}{\beta}\,\overset{x}{\log d}$$

represents the force of gravity between two bodies distance d apart.

$$y = \beta_0 + \beta_1 x \qquad (y_i, x_i) ; i = 1(1)n.$$

$H_0 : \beta_1 = 2$ Vs $H_1 : \beta_1 \neq 2$.

<u>Example</u>: (Full Model)

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{p-1} x_{i,p} + \epsilon_i$$

$H_0 : \beta_r = \cdots = \beta_{p-1} = 0$ Vs

H_1 : H_0 is NOT true.

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{r-1} x_{i,r-1} + \varepsilon_i$$

$$\underset{\sim}{y} = X_r \underset{\sim}{\beta}_r + \varepsilon .$$

$$H_0 : \begin{array}{c} \text{rth position} \\ \begin{pmatrix} 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \cdots & 0 \\ - & - & - & - & - & - & - \\ - & - & - & - & - & - & - \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \end{array} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \vdots \\ \beta_{p-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

$$\Rightarrow A\underset{\sim}{\beta} = \underset{\sim}{c} .$$

$H_0 : A\underset{\sim}{\beta} = c$ Vs $H_1 : A\underset{\sim}{\beta} \neq c$.

L.R.T (Likelihood Ratio Test):

$$LRT = \frac{\underset{\sigma, \beta \in H_0}{} L(\underset{\sim}{y}, \beta, \sigma)}{\underset{\sigma, \beta \in H_0 \cup H_1}{} L(\underset{\sim}{y}, \beta, \sigma)}$$

$H_0 : A\beta = c$

$H_1 : A\beta \neq c$.

$$L(\beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left[\frac{-1}{2\sigma^2} \| y - X\beta \|^2\right]$$

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

$$\hat{\sigma}^2 = \frac{\|y - X\hat{\beta}\|^2}{n}.$$

$$L(\hat{\beta}, \hat{\sigma}^2) = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

minimize:

$\|y - X\beta\|^2$ given $A\beta = c$.

$$S(\beta) = (y - X\beta)^T (y - X\beta) + \lambda(A\beta - c).$$

$$\frac{\partial S(\beta)}{\partial \beta} = -2X^T y + 2X^T X \beta + A^T\lambda = 0.$$

$$\hat{\beta}_R = (X^T X)^{-1} X^T y - \tfrac{1}{2} (X^T X)^{-1} A^T \lambda.$$

$$\frac{\partial S(\beta)}{\partial \lambda} = A\hat{\beta}_R = C$$

$$\Rightarrow \quad A\left[(X^T X)^{-1} X^T y - \tfrac{1}{2}(X^T X)^{-1} A^T \lambda\right] = 0.$$

$$\Rightarrow -\frac{1}{2}\hat{\lambda} = \left[A(X^T X)^{-1} A^T\right]^{-1} \left[c - A\hat{\beta}\right].$$

$$\hat{\beta}_R = \hat{\beta} + (X^T X)^{-1} A^T \left[A(X^T X)^{-1} A^T\right]^{-1} \left[c - A\hat{\beta}\right]$$

$$\hat{\sigma}^2_R = \frac{\|y - X\hat{\beta}_R\|^2}{n}.$$

$$L(\hat{\beta}_R, \hat{\sigma}_R) = (2\pi \hat{\sigma}^2_R)^{-n/2} \cdot e^{-n/2}.$$

$$LRT = \Lambda = \frac{L(\hat{\beta}_R, \sigma^2_R)}{L(\hat{\beta}, \hat{\sigma}^2)} = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}^2_R}\right)^{n/2}$$

$$\frac{n-p}{p}\left(\Lambda^{-2/n} - 1\right) \sim F_{q, n-p}.$$

$q = \text{Rank}(A)$.

Date 16 Aug '23

$H_0 : \beta_1 = \beta_2 = \cdots\cdots = \beta_p = 0.$

Vs H_1 : H_0 is not true.

$$F = \frac{n-p}{p}\left(\Lambda^{-2/n} - 1\right)$$

$$\Lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}^2_R}\right)^{n/2} = LRT.$$

Result: $Y \sim N(\underset{\sim}{\mu}, V)$ and $U = Y^T A Y$.

then $U \sim \chi^2_{r,\lambda}$ if (AV) is idempotent

when rank $(A) = r$.

& $\lambda = \mu^T A \underset{\sim}{\mu}$ non-centrality parameter.

Theorem:-

(i) $RSS_R - RSS = \| \hat{Y}_R - Y \|^2$

$$= (A\hat{\beta} - c)^T \left[A (X^T X)^{-1} A^T \right]^{-1} (A\hat{\beta} - c).$$

(ii) $E\left[RSS_R - RSS \right] = \sigma^2 q + (A\beta - c)^T \left[A (X^T X)^{-1} A^T \right]^{-1} (A\beta - c).$

(iii) when H_0 is true, then

$$F = \frac{(RSS_R - RSS)/q}{(RSS)/n-p} \sim F_{q,\,n-p}.$$

- RSS = Residual sum of squares,

$$= \| Y - \hat{Y} \|^2 = (Y - \hat{Y})^T (Y - \hat{Y})$$

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

where, $\hat{Y} = X\hat{\beta}$.

- $RSS_R = \| Y - X\hat{\beta}_R \|^2$

$A\beta = c$

Proof: (i) $RSS_R - RSS$

$$= (Y - X\hat{\beta}_R)^T (Y - X\hat{\beta}_R) - (Y - X\hat{\beta})^T (Y - X\hat{\beta})$$

since,

$$\hat{\beta}_R = \hat{\beta} + \underbrace{(X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}[C-A\hat{\beta}]}_{B}.$$

$$\hat{\beta}_R = \hat{\beta} + B.$$

$]^{-1}$

$$\therefore (Y - X\hat{\beta} - XB)^T (Y - X\hat{\beta} - XB) - (Y - X\hat{\beta})^T (Y - X\hat{\beta}).$$

$$= -B^T X^T (Y - X\hat{\beta}) - (Y - X\hat{\beta})^T XB + B^T X^T X B.$$

$$= B^T X^T X B.$$

$$(C - A\hat{\beta})^T [A(X^TX)^{-1}A^T]^{-1} A(X^TX)^{-1} X^T X (X^TX)^{-1} A^T [A(X^TX)^{-1}A^T]^{-1} (C - A\hat{\beta}).$$

$$= (C - A\hat{\beta})^T [A(X^TX)^{-1}A^T]^{-1} (C - A\hat{\beta}).$$

$$\underbrace{(A\hat{\beta} - C)^T}_{Y^{*T}} \underbrace{[A(X^TX)^{-1}A^T]^{-1}}_{B^*} \underbrace{(A\hat{\beta} - C)}_{Y^*}$$

$$= \| \hat{Y} - \hat{Y}_R \| \quad \text{(Home work)}.$$

(ii) $E\left[RSS_R - RSS\right] = E\left[y^{*T} B^{*} y^{*}\right]$.

$$y^{*} = A\hat{\beta} - c = A(X^TX)^{-1}X^Ty - c.$$

$$y^{*} \sim N\left(A\beta - c,\ \sigma^2 A(X^TX)^{-1}A^T\right)$$

$$\frac{y^{*}}{\sigma} \sim N\left(\frac{A\beta-c}{\sigma},\ \underset{r\times p}{A}\underset{p\times p}{(X^TX)^{-1}}\underset{p\times r}{A^T}\right)$$

$$E\left[\frac{RSS_R - RSS}{\sigma^2}\right] = E\left[\underbrace{\left(\frac{y^{*}}{\sigma}\right)^T\left[A(X^TX)^{-1}A^T\right]^{-1}\left(\frac{y^{*}}{\sigma}\right)}_{\sim \chi^2_{Rank(A),\ \lambda}}\right]$$

$\tilde{y}$

where,

$$\lambda = \left[\frac{A\beta - c}{\sigma}\right]^T\left[A(X^TX)^{-1}A^T\right]\left[\frac{A\beta - c}{\sigma}\right]$$

$$E\left[\frac{y^{*T}B^{*}y^{*}}{\sigma^2}\right] = \text{trace}\left\{\left[A(X^TX)^{-1}A^T\right]^{-1}\left[A(X^TX)^{-1}A^T\right]\right\}$$
$$+ \left(\frac{A\beta - c}{\sigma}\right)^T\left[A(X^TX)^{-1}A^T\right]\left(\frac{A\beta - c}{\sigma}\right)$$

(Complete it).

Under H_0 :

$$\frac{(RSS_R - RSS)/q}{RSS/(n-p)} \sim F_{q, n-p}$$ under H_0.

$$\frac{RSS_R - RSS}{\sigma^2} \sim \chi^2_{q, \lambda=0}$$

$$RSS/\sigma^2 \sim \chi^2_{n-p, \lambda=0}$$

$$(Y - X\hat{\beta}_R)^T (Y - X\hat{\beta}_R) - (Y - X\hat{\beta})^T (Y - X\hat{\beta})$$

$$RSS = \left(\frac{Y - X\hat{\beta}}{\sigma}\right)^T \left(\frac{Y - X\hat{\beta}}{\sigma}\right)$$

$$= \frac{Y^T}{\sigma} (I - X(X^TX)^{-1}X^T) \frac{Y}{\sigma}$$

$$Y \sim N(X\beta, \sigma^2 I)$$

Date: 18.08.23

$H_0 : A\beta = C$ vs

$H_1 : A\beta \neq C$

In particular.

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

vs H_1 : H_0 is not true.

$\hat{\beta}_R$

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & & & 0 \\ 0 & 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots \\ 0 & 0 & \cdots & & & & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\| \\ [\underset{\sim}{0} \quad I_{p-1}]$$

$F \sim F_{q=p-1,\, n-p}$. $\qquad y_i = \beta_0 + \epsilon_i$

$$\hat{\beta}_0 = \bar{y}.$$

$$\frac{RSS_R - RSS}{\sigma^2} \sim \chi^2_{q = \text{Rank}(A)}.$$

$$\frac{RSS}{\sigma^2} \sim \chi^2_{n-p}.$$

$$\frac{RSS}{\sigma^2} = \frac{Y^T}{\sigma}\left(I - X(X^TX)^{-1}X^T\right)\frac{Y}{\sigma}.$$

$$\frac{\dfrac{RSS_R - RSS}{\sigma^2} \Big/ q}{\dfrac{RSS}{\sigma^2} \Big/ n-p} \sim F_{q,\, n-p}.$$

$\Lambda = LRT$:

$$\frac{n-p}{q}\left(\Lambda^{-2/n} - 1\right) \sim F_{q,\, n-p}.$$

$$\Lambda = \left(\frac{\hat\sigma^2}{\hat\sigma_R^2}\right)^{n/2}$$

$$\frac{n-p}{q}\left(\frac{\hat\sigma_R^2}{\hat\sigma^2} - 1\right)$$

$$= \left(\frac{\hat\sigma_R^2 - \hat\sigma^2 / \sigma^2}{\hat\sigma^2/\sigma^2}\right).$$

Ex: $H_0 : \underset{\sim}{a}^T \beta = c$. vs $H_1 : \underset{\sim}{a}^T \beta \neq c$.

$F \sim F_{1,\, n-p}$.

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

$$Z = \frac{a^T\hat\beta - a^T\beta}{\sqrt{\sigma^2 \underset{\sim}{a}^T (X^T X)^{-1} \underset{\sim}{a}}} \sim N(0, 1)$$

$$\boxed{\hat\sigma^2 = MSE.} \qquad \frac{RSS}{\sigma^2} \sim \chi^2_{n-p}.$$

$$\frac{RSS}{\sigma^2} = X \sim \chi^2_{n-p}$$

$$\boxed{t_n^2 \sim F_{1,n}}$$

$$\frac{\dfrac{a^T\hat{\beta} - c}{\sqrt{\sigma^2 a^T (X^T X)^{-1} a}}}{\sqrt{\dfrac{RSS}{\sigma^2} / n-p}} \sim t_{n-p}.$$

$H_0 : \beta_j = 0$

$c = 0$

$\underset{\sim}{a}^T = (0, 0, \ldots, 1, \ldots 0)$

* Multiple correlation coefficient (R^2) :-

$y = \beta_0 + \beta_1 x + \varepsilon.$

Sample correlation x & y.

$$r^2 = \frac{\{\Sigma (y_i - \bar{y})(x_i - \bar{x})\}^2}{\Sigma (y_i - \bar{y})^2 \, \Sigma (x_i - \bar{x})^2}$$

$$= \frac{S_{xy}^2}{S_{xx} S_{yy}}$$

$$RSS = \Sigma (y_i - \hat{y_i})^2$$

$$\hat{y_i} = \hat{\beta}_0 + \hat{\beta}_1 x_i , \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}.$$

$$RSS = \sum (y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i)^2$$

$$= \Sigma \{ (y_i - \bar{y} - \hat{\beta}_1 (x_i - \bar{x}) \}^2$$

$$= \Sigma (y_i - \bar{y})^2 + \hat{\beta}_1^2 \Sigma (x_i - \bar{x})^2$$

$$- 2 \hat{\beta}_1 \Sigma (y_i - \bar{y})(x_i - \bar{x}).$$

$$= \Sigma (y_i - \bar{y})^2 + \frac{S_{xy}^2}{S_{xx}^2} \times S_{xx} - 2 \frac{S_{xy}}{S_{xx}} \times S_{xy}$$

$$= \Sigma (y_i - \bar{y})^2 - \frac{S_{xy}^2}{S_{xx}} \times \frac{S_{yy}}{S_{yy}}$$

$$= (1 - r^2) \Sigma (y_i - \bar{y})^2 = RSS$$

$$\left\{ \begin{array}{l} \text{Now, as } |r| \longrightarrow 1 \\ RSS \longrightarrow 0. \end{array} \right.$$

$$r^2 = 1 - \frac{RSS}{\Sigma (y_i - \bar{y})^2} = 1 - \frac{RSS}{SST}$$

$$= \frac{SSR}{SST} = R^2.$$

Lack of fit test:

Date: 21 Aug, 23

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

$x_{\cdot 1} \equiv f(x_1)$ like $x_1^2, x_1^3, \log x_1$.

In this test y_{ij} denotes the j'th observations on the response at x_i, $i = 1, \dots m$,

$j = 1, 2, \dots n_i$.

$$n = \sum_{i=1}^{m} n_i$$

x_{ij} = grade of jth person

x_i = grade

$$SS_{Res} \ (RSS) = SS_{PE} + SS_{LOF}.$$

$$(y_{ij} - \hat{y}_i) = (y_{ij} - \bar{y}_i) + (\bar{y}_i - \hat{y}_i).$$

$$(y_{ij} - \hat{y}_i)^2 = (y_{ij} - \bar{y}_i)^2 + (\bar{y}_i - \hat{y}_i)^2 + 2(y_{ij} - \bar{y}_i)(\bar{y}_i - \hat{y}_i).$$

$\{ \hat{y}_{ij} = \hat{y}_i \ \forall j$

$$\sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_i)^2 = \underbrace{\sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}_{SSPE} + \sum_{i=1}^{m} n_i (\bar{y}_i - \hat{y}_i)^2$$

Aug, 23

x_1 .

y_{ij} = grade of j th person

grade

$= \hat{y}_i \ \forall j$

1 . 2

$$\underset{\sim}{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n_1} \\ \vdots \\ y_{21} \\ y_{22} \\ y_{2n_2} \\ \vdots \\ y_{m1} \\ y_{m2} \\ \vdots \\ y_{mn_m} \end{bmatrix}$$

$$\bar{y}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} y_{ij}$$

$$= \frac{1}{n_1} \mathbf{1}_{n_1}^T \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n_1} \end{bmatrix}$$

$$\frac{SS_{PE}}{\sigma^2} \sim \chi^2_{n-m}$$

$$\sum_{j=1}^{n_1} (y_{ij} - \bar{y}_1)^2 \sim \chi^2_{n_1-1}.$$

$$SS_{LOF} \sim \chi^2_{m-2}$$

$$\underset{n-m}{SS_{PE}} + \underset{m-2}{SS_{LOF}} = RSS$$

$$F_0 = \frac{SS_{LOF} / m-2}{SS_{PE} / n-m} \sim F_{m-2,\, n-m}.$$

Confidence intervals and Regions :

(CI)

CI on the Regression coefficients :

$H_0 : \beta_j = 0$ Vs $H_1 : \beta_j \neq 0$.

$(1-\alpha)\%$ CI for β_j

$$t = \frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{var}(\hat{\beta}_j)}}$$

$P\left[|t| > t^{\alpha/2}_{n-p}\right] = \alpha$

$P\left(|t| \leq t^{\alpha/2}_{n-p}\right) = 1-\alpha.$

$$\hat{\beta}_j - t^{\alpha/2}_{n-p}\sqrt{\hat{var}(\hat{\beta}_j)} \leq \beta_j$$

$$\leq \hat{\beta}_j + t^{\alpha/2}_{n-p}\sqrt{\hat{var}(\hat{\beta}_j)}.$$

$$\hat{Var}(\hat{\beta}_j) = MS_{Res}.$$

$$= \hat{\sigma}^2 (X^T X)^{-1}_{jj}$$

Date
23.08.23

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \varepsilon.$$

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$. $A\underset{\sim}{\beta} = c = 0$.

$$\hat{\beta}_R = \begin{pmatrix} \bar{y} \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$F_R = \frac{(RSS_R - RSS)/d.f}{RSS / d.f}$$

$\alpha/2$ $n-p$ $] = \alpha$

$1-\alpha$.

<u>C.I</u> <u>on regression</u> <u>parameters</u>:

$H_0: \beta_j = 0$

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 c_{jj}}} \quad j = 0, 1, \ldots K.$$

c_{jj} is (j,j)th element of $(X^T X)^{-1}$

$$F_j = \hat{\beta}_j \mp t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 c_{jj}}$$

Suppose, E_j $(j = 1, 2, \ldots, K)$ is the event that the jth statement is correct and let

$$P[E_j] = 1 - \alpha_j$$

$$P\left(\bigcap_{j=1}^{k} E_j\right) = 1 - P\left(\cup E_j^c\right)$$

$$\geq 1 - \sum_{j=1}^{k} P(E_j^c)$$

$$= 1 - \sum_{i=1}^{k} \alpha_j$$

If $\alpha_j = \alpha \ \forall j$

$$= 1 - k\alpha.$$

* Bonferroni's t-intervals:

$t_{\alpha/2, n-p}$ if $t_{\alpha/2k, n-p}$

$P(E_j) = 1-\alpha$ $P(E_j) = 1 - \alpha/k$

$$P(\cap E_j) \geq 1 - \frac{k\alpha}{k}$$

$$= 1 - \alpha$$

* Max^m Modulus Intervals:

Define $$T_j = \frac{\underset{\sim}{a}_j^T \hat{\beta} - \underset{\sim}{a}_j \beta_{H_0}}{\sqrt{\widehat{var}(\underset{\sim}{a}_j^T \hat{\beta})}}$$

$$\hat{var}(a_j^T \hat{\beta}) = \hat{\sigma}^2 \, \hat{a}_j (X^T X)^{-1} a_j$$

Set,

$$1-\alpha = P\left[\max_j |T_j| \leq u^{\alpha}_{k,n-p}\right]$$

$\left\{ \text{from Book by Seber \& Lee.} \right.$

Interval is:

$$\hat{\beta}_j \pm u^{\alpha}_{k,n-p} \sqrt{\hat{var}(\hat{\beta}_j)}$$

K is rank of A

$A\beta = C.$

Scheffe's Method:

Without loss of any generality, Assume that first d vectors of the set $\{a_1, a_2, \ldots a_K\}$ are linearly independent and remaining (if any) are linearly dependent on the first d vectors thus. $d \leq \min(K, p)$.

Consider, a, $d \times p$ matrix A of Rank d

Set $\hat{\phi} = A\hat{\beta}$.

$$\frac{(\hat{\phi} - \phi)^T [A (X^T X)^{-1} A^T]^{-1} (\hat{\phi} - \phi)}{\sigma^2} \sim \chi^2_{(d)}$$

B-440

using, $\underset{\sim}{Y} \sim N(\underset{\sim}{0}, \underset{\sim}{V})$

$$(AV)^2 = AV.$$

$$y^T A y \sim \chi^2_{Rank(AV)}.$$

$$\frac{\hat{\phi} - \phi}{\sigma} \sim N\left[0,\ A(X^TX)^{-1}A^T\right]$$

$$\frac{MS_{Res}}{\sigma^2} \sim \chi^2_{n-p}.$$

$$\frac{(\hat{\phi} - \phi)^T \left[A(X^TX)^{-1}A^T\right]^{-1}\left[\hat{\phi} - \phi\right]}{d\ MS_{Res}}$$

$$\sim F_{d,n-p}.$$

* **Maximum modulus t- intervals :-**

* **Scheffe's Method :** $\phi = A\beta$.

$$F = \frac{(\hat{\phi}-\phi)^T [A(X^TX)^{-1}A^T]^{-1}(\hat{\phi}-\phi)/d}{MSRes = \frac{RSS}{n-p} = \hat{\sigma}^2} \sim F_{d,n-p}.$$

$$1-\alpha = P[F \le F^{\alpha}_{d,n-p}]$$

$$= P[(\hat{\phi}-\phi)^T L^{-1}(\hat{\phi}-\phi) \le m] \text{ (say)}.$$

where, $L = A(X^TX)^{-1}A^T$

$m = MSRes\, F_{d,n-p}$.

$$= P[b^T L^{-1} b \le m] \quad \text{where } b = (\hat{\phi}-\phi).$$

$$\boxed{\textbf{Result;}\ \text{If } L \text{ is p.d. then for any } b \quad \sup_{h\neq 0}, \max \frac{(h^T b)^2}{(h^T L h)} = b^T L^{-1} b.}$$

$$= P\left[\sup_{h\neq 0} \frac{(h^T b)^2}{h^T L h} \le m\right]$$

B-440

$$= P\left[\frac{(h^T b)^2}{h^T L h} \le m \ \forall\ h \right]$$

$$= P\left[\frac{|h^T\hat{\phi} - h^T\phi|}{\sqrt{MSRes}\,(h^T L h)^{1/2}} \le \left(d\, F^{\alpha}_{d,\,n-p}\right)^{1/2} \right.$$

$$\forall\ h \neq 0.$$

$\therefore$ CI for $h^T\phi$.

$$\left[h^T\hat{\phi} \pm \left(d\, F^{\alpha}_{d,\,n-p}\right)^{1/2} \sqrt{MSRes\,(h^T L h)} \right]$$

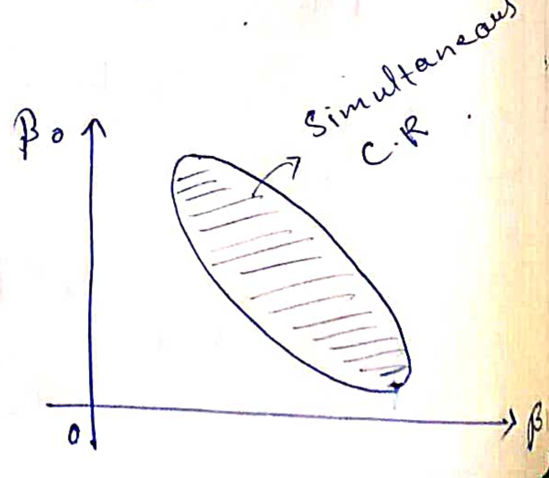
* <u>Confidence Region</u> :- (R)

We know that $F = \dfrac{(\hat{\beta}-\beta)^T (X^TX)(\hat{\beta}-\beta)}{p \times MS_{Res}} \sim F_{p,\,n-p}$

$$MS_{Res}\ (= \hat{\sigma}^2)$$

$$R : \left\{ \beta : F \le F^{\alpha}_{p,\,n-p} \right\}$$

$\beta = (\beta_0, \beta_1)'$



$$P\left[F \leq F^{\alpha}_{p,\, n-p}\right] = 1-\alpha .$$

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

$$(X^T X) = \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}$$

$$(\hat{\beta} - \beta)^T = \left[(\hat{\beta}_0 - \beta), \quad (\hat{\beta}_1 - \beta_1)\right]$$

$$\left[(\hat{\beta}_0 - \beta_0) \quad (\hat{\beta}_1 - \beta_1)\right] \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix} \begin{bmatrix} (\hat{\beta}_0 - \beta_0) \\ (\hat{\beta}_1 - \beta_1) \end{bmatrix}$$

p.

Date
26.08.23

* Residual Analysis *

$$e_i = y_i - \hat{y}_i . \qquad \hat{Y} = X\hat{\beta} .$$

(1) Sample mean:-

$$\bar{e} = \frac{\sum e_i}{n} = 0 .$$

where

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

Proof:

$$\underset{\sim}{1}^T \underset{\sim}{e} = \underset{\sim}{1}^T (Y - \hat{Y})$$

$$= \underset{\sim}{1}^T (Y - X(X^TX)^{-1}X^T Y)$$

$$= \underset{\sim}{1}^T [I - X(X^TX)^{-1}X^T] Y .$$

We have,

$$(X^TX)^{-1}X^TX = I .$$

$$X((X^TX)^{-1}X^TX) = X$$

where

$$X = [\underset{\sim}{1} \;\; X^*]$$

$$\cancel{[\underset{\sim}{1} \;\; X^*][X^TX]^{-1}X^TX = [\underset{\sim}{1} \;\; X^*]}$$

$$X(X^TX)^{-1}X^T[\underset{\sim}{1} \;\; X^*] = [\underset{\sim}{1} \;\; X^*]$$

$$\Rightarrow \quad X(X^TX)^{-1}X^T \underset{\sim}{1} = 1 .$$

$$\underset{\sim}{1}^T X (X^TX)^{-1} X^T = \underset{\sim}{1}^T$$

$$\underset{\sim}{1}^T \underset{\sim}{e} = \underset{\sim}{1}^T (Y - \hat{Y})$$

$$= \underset{\sim}{1}^T [Y - X(X^TX)^{-1}X^TY]$$

$$= \underset{\sim}{1}^T [I - X(X^TX)^{-1}X^T] Y = \underset{\sim}{1}^T - \underset{\sim}{1}^T = 0$$

$Var(\bar{e})$

Sample variance : $\underset{\sim}{\sim} \dfrac{\Sigma (e_i - \bar{e})^2}{n-p}$ when n is large.

$$= \frac{\Sigma e_i^2}{n-p} = MSRes.$$

Methods for Scaling Residuals :-

(1) Standardized Residuals :

$$d_i = \frac{e_i - 0}{\sqrt{MSRes}} \quad i = 1(1)n.$$

- $d_i > 3$

$\Rightarrow$ i^{th} observation is outlier.

(2). Studentized Residuals ;

$$\underset{\sim}{e} = y - \underset{\sim}{\hat{y}} = [I - H] \underset{\sim}{y}$$

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

$$Var(\underset{\sim}{e}) = \sigma^2 [I - H].$$

where, $H = (h_{ij})$.

$$var(e_i) = \sigma^2 (1 - h_{ii})$$

$$r_i = \frac{e_i}{\sqrt{MSRes(1 - h_{ii})}}$$

$\hookrightarrow$ Studentized Residuals.

(3). PRESS Residuals ;

$$e_{(i)} = y_i - \hat{y}_{(i)} = \frac{e_i}{1 - h_{ii}}$$

$$e_{(1)} = y_1 - \hat{y}_{(1)}$$

$$\hat{y}_{(1)} = \underset{\sim}{x}_i \, (x_{(i)}^T x)^{-1} \ldots\ldots$$

Proof: TST : $e_{(i)} = y_i - \hat{y}_{(i)} = \dfrac{e_i}{1-h_{ii}}$

Result: $\left((X^T X) - \underset{\sim}{x}\,\underset{\sim}{x}^T\right)^{-1}$

$$= (X^T X)^{-1} + \frac{(X^T X)^{-1} \underset{\sim}{x}\,\underset{\sim}{x}^T (X^T X)^{-1}}{1 - x^T (X^T X)^{-1} x}$$

$$X^T_{(i)} X_{(i)} = X^T X - X_{(i)} X_{(i)}^T$$

$$y = \beta_0 + \beta_1 x_{i1} \, \beta_1 x_{i2}$$

$$X^T_{(1)} X_{(1)} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{21} & x_{31} & \cdots & x_{n1} \\ \vdots & & & \\ x_{22} & & \cdots & x_{n2} \end{bmatrix} \begin{bmatrix} 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix}$$

$$X^T X = \begin{matrix} n & \sum x_{i1} & \sum x_{i2} \\ \sum x_{i1} & \sum x_{i1}^2 & \end{matrix}$$

$$e_{(i)} = y_i - \hat{y}_{(i)}$$

$$= y_i - x_i^T (X_{(i)}^T X_{(i)})^{-1} X_{(i)}^T y_{(i)}$$

$$= y_i - \left[x_i^T \frac{\left[(X^TX)^{-1} + (X^TX)^{-1} x_i x_i^T (X^TX)^{-1} \right]}{1 - h_{ii}} \right] X_{(i)}^T y_{(i)}$$

$$= y_i - x_i^T (X^TX)^{-1} X_{(i)}^T y_{(i)} - \frac{x_i^T (X^TX)^{-1} x_i x_i^T (X^TX)^{-1} X_{(i)}^T y_{(i)}}{1 - h_{ii}}$$

$$= (1 - h_{ii}) y_i - (1 - h_{ii}) x_i^T (X^TX)^{-1} X_{(i)}^T y_{(i)}$$

$$- h_{ii} \, x_i^T (X^TX)^{-1} X_{(i)}^T y_{(i)} \Big] \Big/ (1 - h_{ii}).$$

$$= \frac{(1 - h_{ii}) y_i - x_i^T (X^TX)^{-1} X_{(i)}^T y_{(i)}}{1 - h_{ii}}$$

Since, $X^T Y = X_{(i)}^T Y_{(i)} + x_i y_i$

$$= \frac{(1 - h_{ii}) y_i - x_i^T (X^TX)^{-1} (X^Ty - x_i y_i)}{(1 - h_{ii})}$$

$$= \frac{y_i - \hat{y}_i}{1-h_{ii}} = \frac{e_i}{1-h_{ii}}$$

using

$\hat{Y} = X\hat{\beta}$

$\hat{y}_i = x_i^T\hat{\beta}$

$$e_{(i)} = \frac{e_i}{1-h_{ii}}$$

$$\sum_{i=1}^{n} e_{(i)}$$

$$\text{var}(e_{(i)}) = \frac{1}{(1-h_{ii})^2} \text{var}(e_i)$$

$$= \frac{\sigma^2(1-h_{ii})}{(1-h_{ii})^2} = \frac{\sigma^2}{(1-h_{ii})}$$

Standard PRESS Residuals :

$$= \frac{e_i/(1-h_{ii})}{\sqrt{\hat{\sigma}^2/(1-h_{ii})}}$$

$$= \frac{e_i}{\sqrt{\hat{\sigma}^2(1-h_{ii})}}$$

P.S. - 1

$X_i \sim N(0,1)$.

(4) $\bar{x} = \frac{1}{n}\sum x_i = \frac{1}{n} \underset{\sim}{1}^T x$

$$= \frac{1}{n}[1, 1, \cdots 1]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

$$\sum (x_i - \bar{x})^2 = (x - \underset{\sim}{1}\bar{x})^T (x - \underset{\sim}{1}\bar{x})$$

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

$$= (x - Jx)^T (x - Jx)$$

where $\frac{\underset{\sim}{1}\underset{\sim}{1}^T}{n} = J$

$$= x^T (I-J)' (I-J) x.$$

$$= x^T (I-J) x$$

As $(I-J)(I-J)$
$= I - J + J^2 - J = I - J$
J is idempotent

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

$$\underset{\sim}{1} - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}^T\underset{\sim}{1} = \underset{\sim}{1} - \underset{\sim}{1} = 0$$

using $l^T y$ and $y^T A y$ are ind. if Al

⑤.

$$\Sigma (x_i - x_{i+1})^2 = x^{*T} x^* = x^T \boxed{A^T A} x$$

$A\underset{\sim}{1} = 0, \quad A^T A \underset{\sim}{1} = 0$

$$[x_1 - x_2, \quad x_2 - x_3 \quad \cdots \quad x_{n-1} - x_n] \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \\ \vdots \\ x_{n-1} \; x_n \end{bmatrix}$$

$$x^* = \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \\ \vdots \\ x_{n-1} \; x_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & -1 & & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & \cdots & & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$Ax \qquad\qquad A \qquad\qquad x$$

⑥ $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim N(\underset{\sim}{0}, \Sigma) \; ; \; \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$

$$\frac{y_1^2}{\sigma_{11}} = [y_1 \quad y_2] \underbrace{\begin{bmatrix} 1/\sigma_{11} & 0 \\ 0 & 0 \end{bmatrix}}_{= A} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$y^T \underbrace{[\Sigma^{-1} - A]}_{B} y \qquad\qquad (B\Sigma)^2 = B\Sigma$$

$= X^T \boxed{A^T A} X$

$$\begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \\ \vdots \\ x_{n-1} \quad x_n \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

X

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

$(B\Sigma)^2 = B\Sigma$

$$A\Sigma = \begin{bmatrix} 1/\sigma_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \sigma_{12}/\sigma_{11} \\ 0 & 0 \end{bmatrix}$$

$(I - A\Sigma)(I - A\Sigma) = (I - A\Sigma)$ idempotent

As

$$\begin{bmatrix} 1 & \sigma_{12}/\sigma_{11} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & \sigma_{12}/\sigma_{11} \\ 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \sigma_{12}/\sigma_{11} \\ 0 & 0 \end{bmatrix} = A\Sigma$$

(8) hint: $Y = X\beta + \varepsilon$.

$$X = \begin{bmatrix} 1 & x_{11} - \bar{x}_1 & \cdots & x_{1,p-1} - \bar{x}_{p-1} \\ 1 & x_{21} - \bar{x}_1 & & x_{2,p-1} - \bar{x}_{p-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} - \bar{x}_1 & & x_{n,p-1} - \bar{x}_{p-1} \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 1/n & 0 \\ 0 & (X^{*T} X^*)^{-1} \end{bmatrix}$$

$\hat{\beta} = (X^T X)^{-1} X^T Y$

$$X^T Y = \begin{bmatrix} \underset{\sim}{1}^T \\ X^{*T} \end{bmatrix} Y = \begin{bmatrix} 1^T Y \\ X^{*T} Y \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & (X^{*T} X^{*})^{-1} \end{bmatrix} \begin{bmatrix} 1^T y \\ X^{*T} y \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \bar{y} \\ (X^{*T} X^{*})^{-1} X^{*T} y \end{bmatrix}$$

Prob. (9) is similar to (8).

(10).

$$F = \frac{\dfrac{(A\hat{\beta} - c)^T [A (X^T X)^{-1} A^T]^{-1} (A\beta - c)}{q = \text{Rank}(A)}}{M.S Res}$$

$H_0 : \alpha_1 = \alpha_2$ $\quad A = [1 \ \ -1]$

$A\underset{\sim}{\alpha} = 0$ where

$$X = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 6 & 0 \\ 0 & 5 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 1/6 & 0 \\ 0 & 1/5 \end{bmatrix}, \quad \hat{\beta} = \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix}$$

$$(\hat{\alpha}_1 - \hat{\alpha}_2)^T \left(11/30\right)^{-1} (\hat{\alpha}_1 - \hat{\alpha}_2) = \frac{(\hat{\alpha}_1 - \hat{\alpha}_2)^2}{(11/30)\, MSRes}.$$

Date
28.08.23

R- Student / Externally studentized Residuals

$$S_{(i)}^2 = \frac{(n-p)\,MSRes - e_i^2/(1-h_{ii})}{n-p-1}.$$

$$H_{(i)} = X_{(i)} (X^T_{(i)} X_{(i)})^{-1} X^T_{(i)}.$$

Normal Probability plot :-

Let $t_{[1]} < t_{[2]} < \cdots < t_{[n]}$

be the R - Student in increasing order.

If we plot $t_{[i]}$ against the cumulative probability

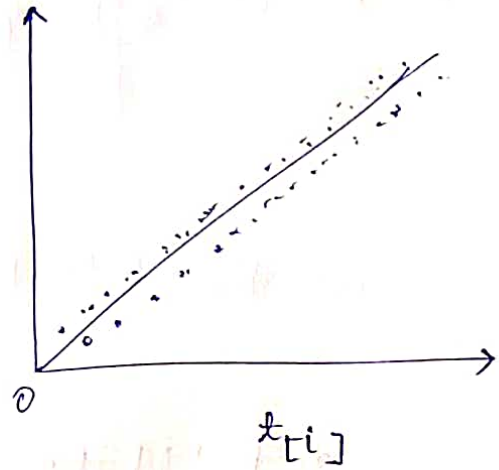
$P_i = (i - 1/2)/n$ for $i = 1(1)n$.

on the normal prob plot, the resulting points should lie approximately on a straight line.

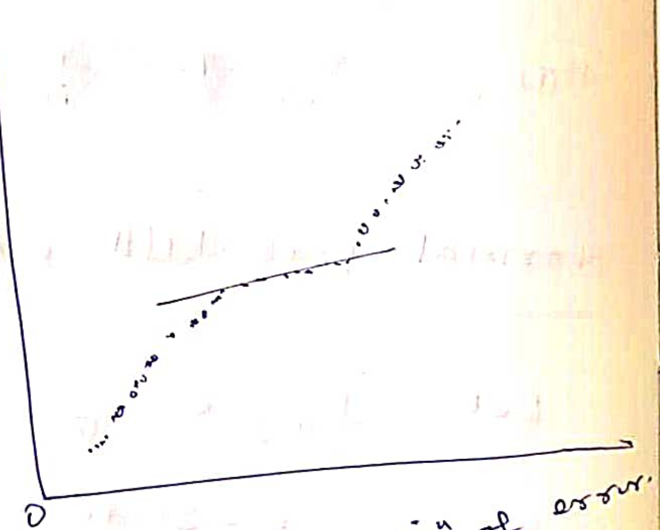
$t_{[i]}$ against $E[t_{[i]}] \simeq \Phi^{-1}[(i-1/2)/n]$

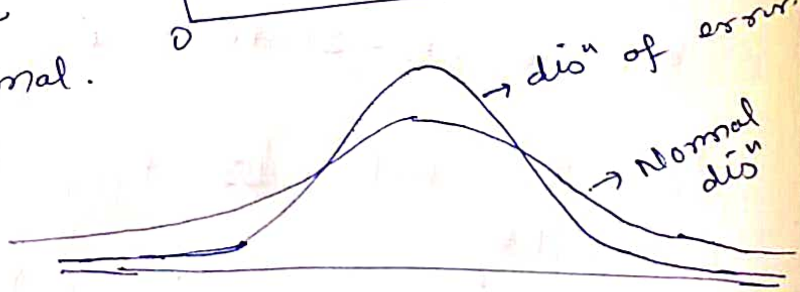
(a)

Ideal plot, error is normally distributed.

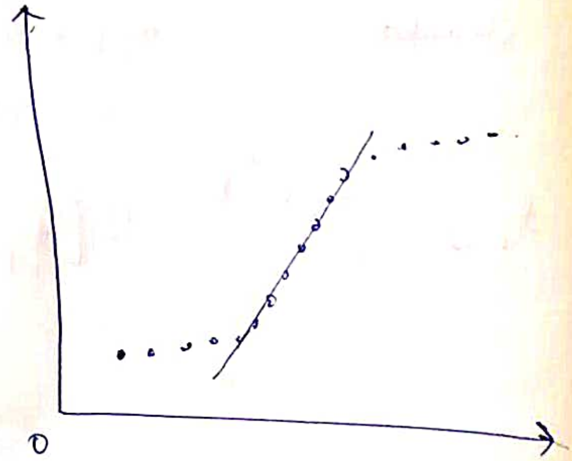


(b) Sharp upward & downward curve at both extremes indicating that the tails of this distribution are too light to be considered normal.





(c)



(d)

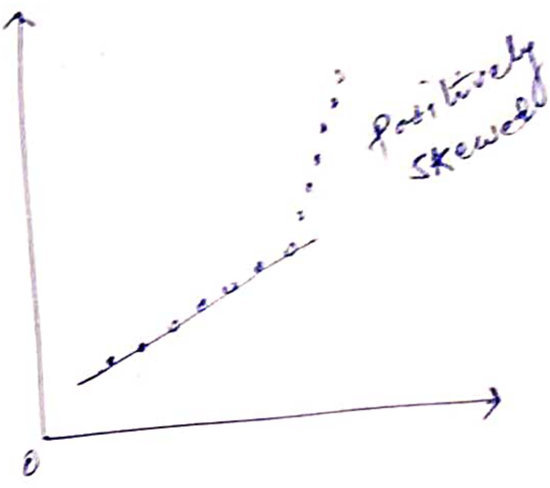
(e)

⇒ Pl

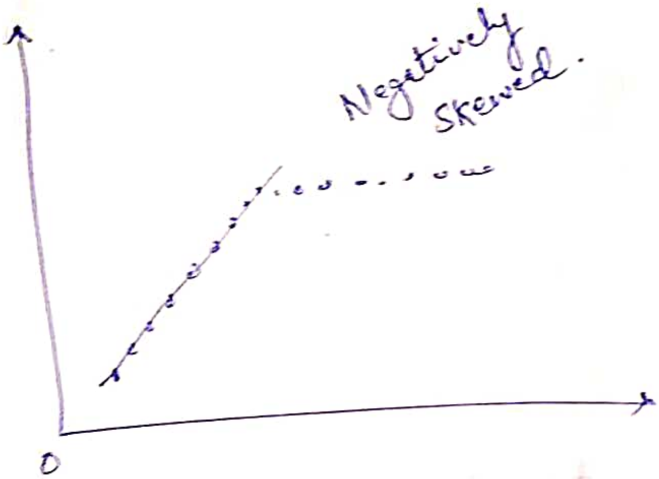
(a)

t_i

(d)



(e)



⇒ Plot of Residuals against the fitted Values :-

(a)

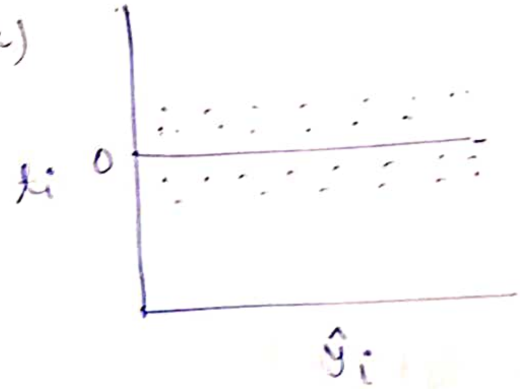


fig: Ideal situation

(b)

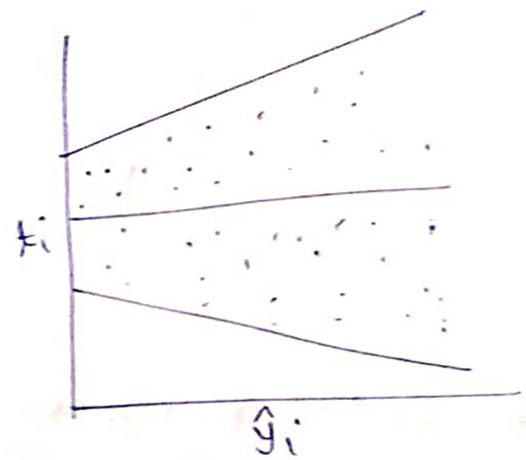
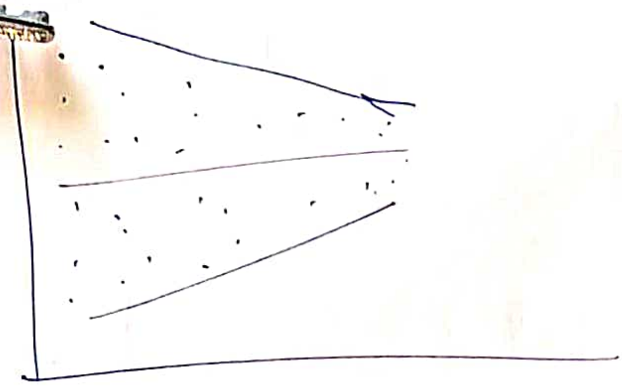
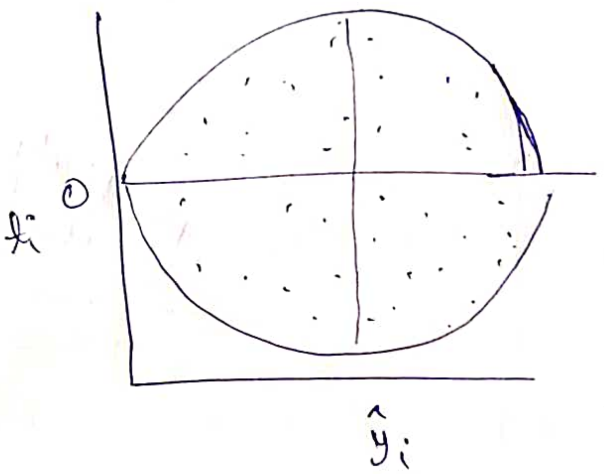


fig: outward funnel pattern implies that the Variance is an incrasing funⁿ of y.

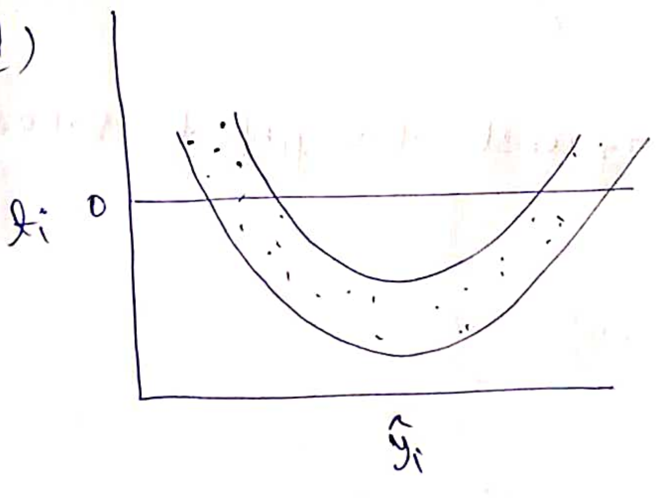


(c)



The double bow pattern often occurs when y is a proportion between 0 & 1.. For example, Binomial proportion have variance.

(d)



A curved plot such as this indicates non-linearity.

* Partial Regression plots (PReg.P.)

consider the marginal rate of the regressor x_j given other regressors that are already in the model.

To illustrate :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon .$$

We are concern about the nature of the marginal relationship for regressor x_1.

First regress y on x_2 and obtain the fitted values and residuals.

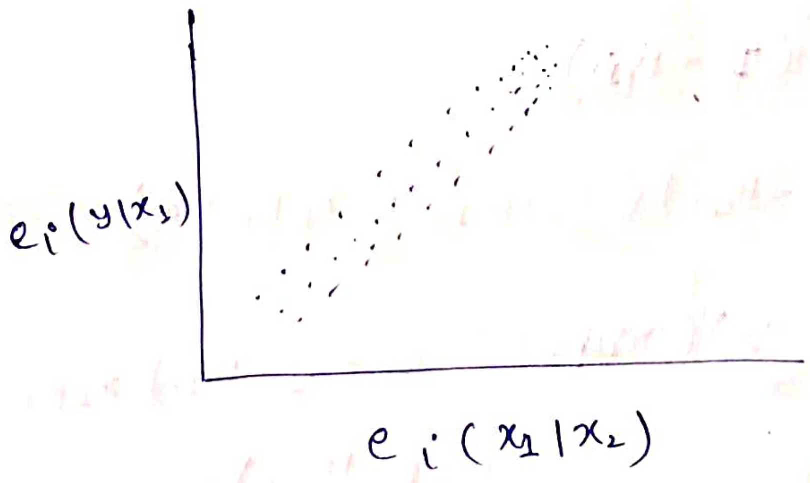
$$\hat{y}_i (x_1) = \hat{\theta}_0 + \hat{\theta}_1 x_{i2}$$

$$e_i (y | x_1) = y_i - \hat{y}_i (x_2) .$$

Now, regress x_1 on x_2 and calculate the residuals

$$\hat{x}_{i1} = \hat{\alpha}_0 + \hat{\alpha}_1 x_{i2} \quad \text{for } i = 1(1)n .$$

$$e_i (x_1 | x_2) = x_{i1} - \hat{x}_{i1}$$



$$Y \longleftrightarrow x_j$$

$e(y \mid X_{(j)}) \rightarrow$ residuals when x_j is removed to see the relationship b/w y & x_j

$$Y = X\beta + \varepsilon = X_{(j)}\beta^* + \beta_j x_j + \varepsilon.$$

$e(x_j \mid X_{(j)}) \rightarrow$ residuals when you regress x_j on remaining x_j's.

Date 30.08.23

PRegP : Partial Regression Plots.

$x_j \rightarrow$ removed.

$$e(\underline{y} \mid X_{(j)})$$

$$e(\underline{x}_j \mid X_{(j)})$$

$$Y = X\beta + \varepsilon \quad , \quad X = [X_{(j)}\ x_j]$$

$$Y = X_{(j)}\beta_{(j)} + x_j \cdot \beta_j + \varepsilon.$$

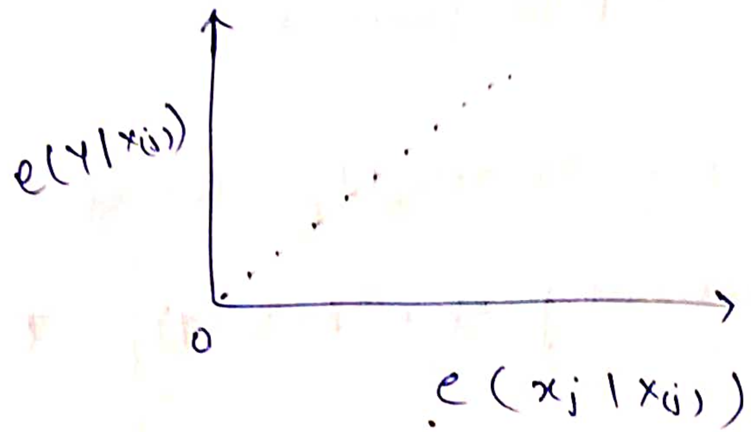
Pre-multiply $(I - H_{(j)})$

$$(I - H_{(j)})Y = (I - H_{(j)})(X_{(j)}\beta_{(j)} + x_j\beta_j + \varepsilon)$$

$$= \cancel{(I-H_{(j)})(X_{(j)}\beta_{(j)}} \quad 0 + (I - H_{(j)})x_j\beta_j$$

$$+ (I - H_{(j)})\varepsilon$$

$$e(Y \mid X_{(j)}) = \beta_j \cdot e(x_j \mid X_{(j)}) + \varepsilon^*.$$



$$\hat{Y} = X_{(j)}\hat{\beta}_{(j)} + x_j\hat{\beta}_j.$$

$$e(Y \mid X_{(j)}) = Y - X_{(j)}\hat{\beta}_{(j)}$$
$$= Y - \hat{Y} + x_j\hat{\beta}_j = e + x_j\hat{\beta}_j$$

$$e(Y \mid X_{(j)}) = e + x_j\hat{\beta}_j$$

$\downarrow$

partial residual.

$$e_i(Y \mid X_{(j)}) = e_i + x_{ij}\hat{\beta}_j.$$

PRESS statistics:

$$PRESS = \sum_{i=1}^{n}(y_i - \hat{y}_{(i)})^2 = \sum_{i=1}^{n}\left(\frac{e_i}{1-h_{ii}}\right)^2$$

Smaller PRESS value is preferred..

R^2 for prediction Based on PRESS:

$$R^2_{prediction} = 1 - \frac{PRESS}{SS_T}$$

- Transformation and weighting to correct model inadequacies:

Variance stabilizing transformation on y:

Relationship of σ^2 to $E(Y)$	Transformation
$\sigma^2 \propto$ ~~cost~~ const.	$y' = y$
$\sigma^2 \propto E(y)$	$y' = \sqrt{y}$ (Poisson / count type data)
$\sigma^2 \propto E(Y)(1 - E(y))$	$y' = \sin^{-1}(\sqrt{y})$ (Binomial proportion)

Transformation and weighting to correct model inequalities:

$\sigma^2 \propto (E(y))^2$ $\quad y' = \ln(y)$

$\sigma^2 \propto (E(y))^3$ $\quad y' = y^{-1/2}$

$\sigma^2 \propto (E(y))^4$ $\quad y' = y^{-1}$.