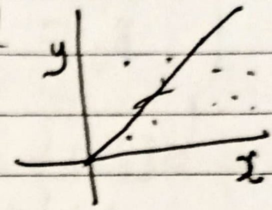
# Simple linear regression

$y \to$ dependent var.
$x \to$ Regressor / indep. var.

Ex> $y \to$ Shear strength
$x \to$ age of propellent
to predict $\hat{y}$.




$$f(y) = \beta_0 + \beta_2 f(x)$$
$$\parallel$$
$$y = \beta_0 + \beta_1 x + \varepsilon$$
$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad , \; i = 1, \cdots n$$

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

$$\sum \varepsilon_i^2 = \sum | \beta_0 + \beta_1 x_i - y_i |^2$$
$$= S(\underset{\sim}{\beta})$$

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

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

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

$$\frac{\partial S(\underset{\sim}{\beta})}{\partial \beta_1} = 2 \sum (y_i - \beta_0 - \beta_1 x_i)(-x_i) = 0$$

Date
Page

$$\hat{B}_1 = \frac{S_{xy}}{S_{xx}}, \text{ where}$$

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

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

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

Least square estimator.

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

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

$$y = B_0 + B_1 x$$

$H_0 : B_1 = 0$ vs
$H_1 : B_1 \neq 0$

$$E(\hat{B}_1) = E\left(\frac{\Sigma(y_i - y)(x_i - \bar{x})}{\Sigma(x_i - \bar{x})^2}\right)$$

$\rightarrow C_i$

{here $y_i$ is variable due to $\varepsilon_i$ which is changing}

$$E(\hat{B}_1) = \Sigma C_i \, E(y_i - \bar{y})$$

$$E(y_i) = B_0 + B_1 x_i$$

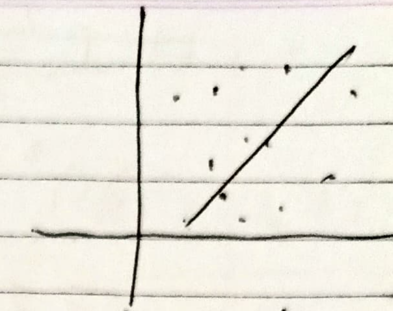
$$E(\bar{y}) = \cancel{\tfrac{1}{n}}(B_0 + B_1 \bar{x})$$

$$\Rightarrow E(\hat{B}_1) = \Sigma C_i (B_0 + B_1 x_i - B_0 - B_1 \bar{x})$$

$$= \Sigma C_i \, B_1 (x_i - \bar{x})$$

$E(\hat{B}_0) = B_0$

$\Rightarrow$ LSE are unbiased.



# We can hypothesis testing here to determine if the calculated slope is statistically significant.
By using statistical tests, such as t-test or P-Value calculations, we can determine whether to ~~determin~~ reject or accept null hypothesis.

$$y_i = B_0 + B_1 x_i + \varepsilon_i$$

LSEs -
$$\hat{B}_0 = \bar{y} - \hat{B}_1 \bar{x}$$

$$\hat{B}_1 = \frac{S_{xy}}{S_{xx}} \qquad S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y})$$
$$S_{xx} = \Sigma(x_i - \bar{x})^2$$

$$Var(\hat{B}_1) = Var\left(\Sigma c_i (y_i - \bar{y})\right)$$

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

$$Var\left(\Sigma \frac{(x_i - \bar{x})}{\Sigma(x_i - \bar{x})^2}(y_i - \bar{y})\right)$$

$$= \Sigma c_i^2 \, \cancel{Var}(y_i - \bar{y})$$

$$Cov(y_i \bar{y}) = Cov\left(y_i \frac{1}{n} \sum y_i\right)$$

$$= \frac{1}{n^2} Cov\ (y_i$$

$$\cancel{= c_i^2 \left(\sigma^2 - \frac{\sigma^2}{n}\right)}$$

$$= \sum c_i^2 \left( Var(y_i) + Var(\bar{y}) - 2Cov(y_i \bar{y})\right]$$

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

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

$$Var(\hat{\beta}_0) = Var(\bar{y} - \hat{\beta}_1 \bar{x})$$

# $\varepsilon_i \sim N(0, \sigma^2)$

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

MLES

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

$$\frac{\partial l}{\partial \beta_0} = \frac{-n}{2\sigma^2}$$

from here, we can conclude both LSE and MLE both are same.

Differentiating w.r.t $\sigma^2$,

$$-\frac{n}{2}\frac{\cancel{2\pi}}{\cancel{2\pi}\sigma^2} + \frac{\sum(\varepsilon_i)}{2\sigma^4} = 0$$

$$\frac{n}{\sigma^2} = \frac{\Sigma(\varepsilon_i)}{\sigma^4}$$

$$\Rightarrow \hat{\sigma}^2_{MLE} = \frac{\Sigma(\varepsilon_i)}{n} = \frac{\Sigma(y_i - \beta_0 - \beta_1 x_i)^2}{n}$$

## Multivariate linear regression

$$y_i = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{ik} + \varepsilon_i$$

$$\underset{\sim}{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & & \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}$$

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

$$y_{n\times1} = X_{n\times(k+1)} \underset{\sim}{\beta}_{(k+1)\times1} + \underset{\sim}{\varepsilon}_{n\times1}$$

$$\boxed{(k+1) < n}$$

**LSE :** $\underset{\sim}{\varepsilon}^T \underset{\sim}{\varepsilon} = \Sigma \varepsilon_i^2 = \Sigma(y_i - \beta_0 - \beta_1 x_1)^2$

Date
Page

$\boxed{p = k+1}$

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

$$\min_{\underset{\sim}{\beta} \in R^p} S(\beta) = \underset{\sim}{y}^T y - \underset{\sim}{\beta}^T X \underset{\sim}{y} - \underset{\sim}{y} X \beta$$
$$+ \underset{\sim}{\beta}^T X^T X \underset{\sim}{\beta}$$

<u>Results</u> -

1) Let $\alpha = \underset{\sim}{y}^T A \underset{\sim}{x}$

$$\frac{\partial \alpha}{\partial \underset{\sim}{x}} = \underset{\sim}{y}^T A \, , \quad \frac{\partial \alpha}{\partial \underset{\sim}{y}} = \underset{\sim}{x}^T A^T$$

2) $\alpha = \underset{\sim}{x}^T A \underset{\sim}{x}$ , $\dfrac{\partial \alpha}{\partial \underset{\sim}{x}} = \underset{\sim}{x}^T (A + A^T)$

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

X is of full rank.

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

$$\frac{\partial S(\beta)}{\partial \beta^T \partial \beta} = |(X^T X)|$$

$$\left\{ X^T X \rightarrow Z^T X^T X Z \rightarrow (XZ)^T (XZ) = V^T V \atop \qquad\qquad\qquad\qquad\qquad \Sigma V_i^2 \geq 0 \right.$$

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

Date
Page

$X = [x_0, x_1, x_2, \ldots, x_k] \begin{bmatrix} z_0 \\ z_1 \\ \vdots \\ z_k \end{bmatrix}$

$\sum_{j=0}^{k} z_j x_j = 0$

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

# Calculating Mean & var.

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

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

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

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

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

$= Var(Ay)$

$= A\, Var(y)\, A^T$

$= (X^T X)^{-1} X^T (\sigma^2 I) X (X^T X)^{-1}$

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

$Cov(Y, X)$

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

4/08/23

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

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

$= \underbrace{(I - X(X^T X)^{-1} X^T)}_{H} y$

(1) $X^T = X$

(2) $H^2 = H$

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

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

$+$

<u>Result</u>- $Z$ is a RV with $E(Z) = \mu$

$Var(Z) = \Sigma$

$E(Z^TAZ) = \text{trace}(A\Sigma) + \mu^TA\mu$

$\Rightarrow E\left(\frac{Y^THY}{n}\right) = \frac{1}{n}\left[\left[\text{trace}(H\sigma^2 I)\right]\right.$

$\left. + \beta^TX^THX\beta\right]$

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

$= X - X = 0$

$\sigma^2 tr(H) = \sigma^2 tr(I - X(X^TX)^{-1}X)$

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

$= \sigma^2 [n-p]$

$E\left(\frac{Y^THY}{n}\right) = \frac{\sigma^2(n-p)}{n}$

$E\left(\frac{Y^THY}{n-p}\right) = \sigma^2$

$\hat{\sigma}^2 = \frac{Y^THY}{n-p} = \cancel{SSE}$ $\rightarrow$ Error ~~Sum of Squa~~ SSE

$$= \frac{SS_E}{\underbrace{n-p}_{\text{degree of freedom}}} = MS_E$$

degree of freedom

**Theorem (Gauss - Markov thm):**

Under GM assumptions, $\hat{\beta}_{MLE/LSE} = (X^TX)^{-1}X^TY$ is " BEST LINEAR UNBIASED ESTIMATOR" (BLUE) of $\beta$.

$$C = \{AY + b_0 : A\}_{b_0 \in R^p}$$

$$Var(l^T \tilde{\beta}_{LSE}) \le Var(l^T \tilde{\beta}) \quad \forall \tilde{\beta} \in C$$
$$\forall l \in R^p$$

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

$$\tilde{\beta} = AY + b_0$$
$$= [(X^TX)^{-1}X^T + B]Y + b_0$$
$$\{(X^TX)^{-1}X^T + \underbrace{A - (X^TX)^{-1}X^T}_{B}\}$$

$$E(\tilde{\beta}) = ((X^TX)^{-1}X^T + B)X\beta + b_0$$
$$= \beta + BX\beta + b_0$$

because $\tilde{\beta}$ is unbiased

$b_0 = 0$

$BX\beta = 0 \Rightarrow BX = 0$

$BX\beta + b_0 = 0$

Date
Page

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

for it to be unbiased estimator,
$I + BX = I$ and $b_0 = 0$
$\Rightarrow BX = 0$

Calculating var. of $\tilde{\beta}$ -

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

{ using formula, $Var(AY) = A\,Var(Y)\,A^T$ }

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

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

$$~~Var~~\quad Var(l^T\tilde{\beta}) = Var(l^T\hat{\beta}_{LSE}) + \underbrace{\sigma^2 l^TBB^Tl}_{\geq 0}$$

$v_i$ are independent ⇒ exponential $v_i$ also independent

Date
Page

## Multivariate normal distr$^n$ (MVN)

Let $\underset{\sim}{v} = (v_1, \dots, v_n)^T$ where

$v_i \sim N(0,1)$, IID

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

$$\underset{\sim}{Z} = \underset{\substack{\downarrow \\ \text{full rank.}}}{X}\underset{\sim}{v} + \underset{\sim}{\mu}$$

$$E(\underset{\sim}{Z}) = \underset{\sim}{\mu}$$

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

### Characteristics function

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

First let us find CF of $\underset{\sim}{v} \sim N(0, I_{n \times n})$

$$\phi_{\underset{\sim}{v}}(t) = E\left[e^{i t^T \underset{\sim}{v}}\right]$$

$$= E\left[e^{i \sum_{j=1}^{n} t_j^T v_j}\right]$$

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

Now

$$E\left(e^{i t_j v_j}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-v_j^2/2} e^{i t_j v_j} \, dv_j$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[v_j^2 - 2 i t_j v_j\right]} \, dv_j$$

Date
Page

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

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

$$= e^{-t_j^2/2} \qquad \begin{array}{l} w = u_j - it_j \\ dw = du_j \end{array}$$

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

$$= e^{-\underset{\sim}{t}^T t/2} \qquad \underset{\sim}{u} \sim N(0, I)$$

$$\underset{\sim}{Z} = X\underset{\sim}{u} + \underset{\sim}{\mu}$$

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

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

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

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

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

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

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

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

let us define $\quad w = \Sigma^{-1/2} (z-\mu)$

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

$$\Rightarrow f(\underset{\sim}{z}) = \frac{1}{(2\pi)^{\frac{n}{2}} \Sigma^{\frac{1}{2}}} \exp\left[-\frac{1}{2} w^T w\right]$$

$$|\Sigma^{\frac{1}{2}}| \, d\underset{\sim}{w} = d\underset{\sim}{z}$$

$$\left|\frac{dw_i}{\partial z_j}\right| = |\Sigma^{-\frac{1}{2}}|$$

$$\int_{-\infty}^{\infty} f(\underset{\sim}{z}) \, d\underset{\sim}{z} = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} w^T w} |\Sigma|^{\frac{1}{2}} dw$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} w^T w} \, dw$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{-\infty}^{\infty} e^{-\frac{\Sigma w_i^2}{2}} \, dw_1 \, dw_2 - d$$

$$= 1$$

Date
Page

$$\phi_{\underset{\sim}{Z}}(t) = E\left(e^{it^T Z}\right)$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \cdot f(Z)\, dz$$

$$w = Z - \underset{\sim}{\mu}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} w^T \Sigma^{-1} w\right)$$

$$\exp\left\{ it^T (w + \underset{\sim}{\mu}) \right\} dw_1 - dw_n$$

— exercise.

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

**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-squared distribution with $r$ degree of freedom if and only if $A$ is idempotent and $Rank(A) = r$.

**Proof:** If $A^2 = A$ and $Rank(A) = r$

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

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

$$\underset{\sim}{y}^T A \underset{\sim}{y} = \underset{\sim}{y}^T P^T D P y = Z^T D Z$$

$$= \sum_{i=1}^{r} Z_i^2 \sim \chi_r^2$$

Date
Page

Conversly, let $\underset{\sim}{y} A \underset{\sim}{y} \sim \chi^2_\gamma$

Since A is symmetric $\Rightarrow A = \underset{\text{orthogonal}}{P^T} \operatorname{diag}(\lambda_1, \ldots, \lambda_n) P$ (eigenvalues of A)

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

Characteristics functions of $\underset{\sim}{y}^T A \underset{\sim}{y}$

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

$$= E\left[e^{it\, z^T D z}\right]$$

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

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

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

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

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

Root $\Rightarrow$ $t = \frac{1}{2i}$ — $\gamma$ times $\Rightarrow \lambda_j = 1$ for $j = 1, \ldots, \gamma$

$= 0$ o.w

<u>Theorem</u>: Let $y \sim N(\underset{\sim}{\mu}, \Sigma)$ and suppose that $y$, $\underset{\sim}{\mu}$ and $\Sigma$ are conformally partitioned. Then, $\underset{\sim}{y_1}$ & $y_2$ are independent iff $\Sigma_{12} = 0$

Date
Page

$$\underset{\sim}{y} = \begin{pmatrix} \underset{\sim}{y_1} \\ \underset{\sim}{y_2} \end{pmatrix} \sim N\left( \begin{pmatrix} \underset{\sim}{\mu_1} \\ \underset{\sim}{\mu_2} \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right)$$

**Proof -**

Independent $\Rightarrow Cov(\underset{\sim}{y_1}, \underset{\sim}{y_2}) = \Sigma_{12} = 0$

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

$$t = \begin{bmatrix} \underset{\sim}{t_1} \\ t_2 \end{bmatrix}$$

$= \cancel{exp\left[ i t^T \left[ \mu \right.\right.}$

$$exp\left[ i \left[ t_1^T \; t_2^T \right] \begin{bmatrix} \underset{\sim}{\mu_1} \\ \underset{\sim}{\mu_2} \end{bmatrix} - \frac{1}{2} \left[ t_1^T \; t_2^T \right] \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \right\}$$

$$= exp\left[ it_1^T \mu_1 + it_2^T \underset{\sim}{\mu_2} - \frac{1}{2} t_1^T \Sigma_{11} t_1 - \frac{1}{2} t_2^T \Sigma_{22} t_2 \right]$$

$$= exp\left[ it_1^T \mu_1 - \frac{1}{2} t_1^T \Sigma_{11} t_1 \right] exp\left[ it_2^T \mu_2 - \frac{t_2^T \Sigma_{22} t_2}{2} \right]$$

$$= \phi_{\underset{\sim}{y_1}}(t_1)\, \phi_{\underset{\sim}{y_2}}(t_2)$$

**Theorum -** $\underset{\sim}{y} \sim N(\underset{\sim}{\mu}, \sigma^2 I)$

Let $A$ & $B$ are matrices such that $AB^T = 0$. then, $Ay$ & $By$ are idenpodent.

Date
Page

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

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

11/8/23 **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 iff $A_1 A_2 = 0$.

**Proof**: $A_1 A_2 = 0$ $\quad$ $A_1 \underset{\sim}{y}$ & $A_2 \underset{\sim}{y}$ are independent.

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

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

Let $y^T A_1 y$ and $y^T A_2 y$ ind.

from prev. thms $\Rightarrow$

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

$$(A_1 + A_2)(A_1 + A_2) = (A_1 + A_2)$$

$$A_1 + A_2 A_1 + A_1 A_2 + A_2 = A_1 + A_2$$

$$\Rightarrow A_1 A_2 + A_2 A_1 = 0$$

By multiplica

Multiply By $A_2 \rightarrow$ $A_2 A_1 A_2 + A_2 A_1 = 0$ $\Rightarrow A_2 A_1 A_2 = 0$

Date

Page

Theorem - Let $y \sim N(0, I_n)$. Let $A^T = A$, $A^2 = A$. Let $l \in \mathbb{R}^n$, $l \neq 0$. Then $y^T A y$ & $l^T y$ are ind. iff $Al = 0$

WLOG, $\|l\| = 1$

$$B = ll^T$$

First let $y^T A y$ & $l^T y$ are ind.

$$y^T A y, \quad y^T l l^T y = y^T B y$$

$AB = 0 \Rightarrow All^T = 0$ (Multiply both s by $l$)

$$\Rightarrow All^T l = 0$$

$$\Rightarrow Al = 0$$

Conv if $Al = 0 \Rightarrow Ay$ & $l^T y$ ind.

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

$\Rightarrow y^T Ay$ & $l^T y$ are indp.

(Matrix version) Cochran's thm $\Rightarrow$ Let $A_1, A_2, \ldots, A_k$ be $n \times n$ matrices with $\sum_{i=1}^{k} A_i = I$. Then, following condn are equivalent:

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

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

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

Date
Page

**Proof:** Let $A_i = \underset{n \times r_i}{B_i} \; \underset{r_i \times n}{C_i}$

(i) ⇒ (iii)

Rank factorisation ⇒ full-column rank, full-row rank

$$\sum A_i = I$$

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

$$\Rightarrow [B_1 \; B_2 \; \cdots \; B_k] \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_k \end{bmatrix} = I$$

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

$AB = I$
$\Rightarrow BA = I$
by Prob1 ...

$$\begin{bmatrix} C_1 \\ \vdots \\ C_k \end{bmatrix} [B_1 \cdots B_k] = I$$

$$\begin{bmatrix} C_1 B_1 & C_1 B_2 & \cdots & C_1 B_k \\ & C_2 B_2 & & \\ & & \ddots & \\ & & & C_k B_k \end{bmatrix} = \begin{bmatrix} I_1 & 0 & \cdots & 0 \\ 0 & I_2 & & \\ \vdots & & \ddots & \\ & & & I_k \end{bmatrix}$$

(iii) ⇒ (ii) Since $\sum A_i = I$

$$A_j \left( \sum A_i \right) = A_j$$

$$\Rightarrow \sum_{i \neq j} A_i A_j + A_j^2 = A_j$$

$$\Rightarrow A_j^2 = A_j$$

(ii) ⇒ (i) ~~Rank~~ $A_i$ is idempotent

$$\text{Rank}(A_i) = \text{trace}(A_i)$$

Date
Page

## Cochran thm for quadratic forms.

Let $U_1, U_2, \ldots, U_N \sim N(0,1)$ I.I.D

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

Define $r_i = \text{Rank}(B^{(i)})$ &

$$Q_i = U^T \underset{N \times N}{B^{(i)}} U \quad \text{where } U^T = (U_1, \ldots, U_N)$$

If $\sum_{i=1}^{k} B^{(i)} = I$, then following are equiv.

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

(ii) $Q_i$ are independent

(iii) $Q_i$ are chi-squared RV with $r_i$ df.

Ex ⇒ Consider the model:
$\log F = \log C - B \log d$ represents the force of gravity between two bodies distance 'd' apart.

$$y = B_0 + B_1 X$$

$B_1 = ?$ $\quad (y_i, x_i), i = 1, \ldots n$

$H_0: B_1 = 2$ vs $H_1: B_1 \neq 2$

Ex ⇒ Full model, $\quad y = XB + \varepsilon$

$$Y_i = B_0 + B_1 X_{i1} + \cdots + B_{p-1} X_{i(p-1)} + \varepsilon_i$$

$H_0: B_2 = \cdots = B_{p-1} = 0$ vs $H_1$: $H_0$ is not true.

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

$$\underset{\sim}{Y} = X_{r} \underset{\sim}{\beta}_r + \varepsilon$$

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

$$\Rightarrow A\underset{\sim}{\beta} = \underset{\sim}{C}$$

$$H_0: A\underset{\sim}{B} = C \quad \text{vs} \quad H_1: A\underset{\sim}{B} \neq C$$

<u>LRT</u>

$H_0: AB = C$

$H_1: AB \neq C$

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

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

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

$$\hat{\sigma}^2 = \frac{\|Y - X\hat{B}\|^2}{n}$$

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

<u>Minimize</u> $\|Y - XB\|^2$ given $AB = C$

$$S(B) = (Y - XB)^T (Y - XB) + \lambda(AB - C)$$

Date
Page

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

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

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

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

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

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

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

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

$$LRT = \Lambda = \frac{L(\hat{\beta}_R, \sigma_R^2)}{L(\hat{\beta}, \hat{\sigma}^2)}$$

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

$$\sim F_{q, n-p}$$

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

$H_0 : B_1 = \cdots = B_p = 0$
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 $rank(A) = r$
& $\lambda = \underset{\sim}{\mu}^T A \underset{\sim}{\mu}$ non-certainity paramete

**Theorem** :

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

$$= (A\hat{B} - C)^T \left[A(X^TX)^{-1}A^T\right]^{-1} (A\hat{B} - C)$$

(ii) $E[RSS_R - RSS]$

$$= \sigma^2 q + (AB - C)^T \left(A(X^TX)^{-1}A^T\right)^{-1} (AB - C)$$

(iii) when $H_0$ is true, then

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

Date
Page

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

Residual Sum of Squares

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

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

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

Now,

$$RSS_R - RSS = (y - X\hat{\beta}_R)^T (y - X\hat{\beta}_R) - (y - X\hat{\beta})^T (y - X\hat{\beta})$$

Since,

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

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

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

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

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

Putting value of B,

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

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

In Idempotent matrix, Rank of matrix = trace of matrix.

$$= (A\hat{B} - C)^T \left[ A(X^TX)^{-1} A^T \right]^{-1} (A\hat{B} - C)$$

$$\quad\; Y^{*T} \qquad\qquad B^* \qquad\qquad Y^*$$

$$= \| \hat{Y} - \hat{Y}_R \| \qquad (H.W)$$

(ii) $E(RSS_R - RSS) = E(Y^{*T} B^* Y^*)$

$$Y^* = A\hat{B} - C = A(X^TX)^{-1} X^T Y - C$$

$$Y^* \sim N(AB - C, \sigma^2 A(X^TX)^{-1} A^T)$$

(Multivariate N)

$$\frac{Y^*}{\sigma} \sim N\left(\frac{AB-C}{\sigma}, A(X^TX)^{-1} A^T\right)$$

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

$$E\left(\frac{Y^{*T} B^* Y^*}{\sigma^2}\right) = trace\left\{ (A(X^TX)^{-1} A^T)(\;)^{-1} \right\}$$

$$+ \left(\frac{AB - C}{\sigma}\right)^T \left[A(X^TX)^{-1}A^T\right]\left(\frac{AB-C}{\sigma}\right)$$

$$= q + -$$

$q$ = no. of restr

Under $H_0$,

$$\frac{(RSS_R - RSS)/q}{RSS/n-p}$$

$$\frac{X/n_1 \sim \chi^2_{df=n_1}}{Y/n_2 \sim \chi^2_{df=n_2}}$$ ← Separately ↙ $\sim F_{n_1, n_2}$

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

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

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

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

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

18/8/23

$H_0 : A\beta = C$

vs $H_1 : A\beta \neq C$

In particular,

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

vs $H_1$ : $H_0$ is not true.

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

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

$$= [\underline{0} \;\; I_{p-1}]$$

$$y_i = \beta_0 + \varepsilon_i$$

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

$RSS_R - RSS/... \; \chi^2$

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

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

$q^T \hat{\beta} \sim \chi^2$

① ÷ ② follows $F_{q, n-q}$.

$\Lambda = 2RT$

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

$$\Lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}^2_R}\right)^{n/2} = \frac{n-p}{q}\left(\frac{\hat{\sigma}^2_R}{\hat{\sigma}^2} - 1\right)$$

$$= \frac{n-p}{q}\left[\frac{(\hat{\sigma}_R^2 - \hat{\sigma}^2)/\sigma^2}{\hat{\sigma}^2/\sigma^2}\right]$$

Ex- $H_0: \underset{\sim}{q}^T \beta = C$ vs $H_1: \underset{\sim}{q}^T \beta \neq C$

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

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

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

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

Result: $X \sim \chi^2_n$, $Y \sim N(0,1)$

$$T = \frac{Y}{\sqrt{X/n}} \sim t_n$$ where $X$ & $Y$ are indepen.

$$\frac{\chi^{2}}{y \sim N} \sim F$$

$$\frac{a^T \hat{\beta} - c}{\sqrt{MSE\, \underset{\sim}{a}^T (X^T X)^{-1} \underset{\sim}{a}}} \sim t_{n-p}$$

$$\downarrow$$

$$RSS / n-p$$

## Multiple Correlation Coefficient

$R^2 \rightarrow$ the coefficient of determination.
is a measure that provides information about the goodness of fit of a model.

$|R| \rightarrow$ Best Estimate

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

Sample correlation $X$ & $Y$,

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

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

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

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

(Refer book)

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

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

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

$\Rightarrow RSS \to 0 \Rightarrow$ Best line fit.

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

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

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

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

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

$$= 1 - \frac{RSS}{SST} = \frac{SS_R}{SS_T} = R^2$$

## Lack of fit test -

$y_{ij}$ denotes the j-th obs. on the response at $x_i$, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$

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

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

$$(y_{ij} - \hat{y}_i) = (y_{ij} - \bar{y}_i) + (\bar{y}_i - \hat{y}_i) \qquad \bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

Square on both sides,

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

Distribution of $\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_i)^2 \sim \chi^2_{n-m}$

$$\underset{\sim}{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ \vdots \\ y_{2n_2} \\ \vdots \\ y_{m1} \\ \vdots \\ y_{mn_m} \end{bmatrix} \qquad A = \begin{bmatrix} 0_{n_1 \times n_1} & 0 \\ 0 & 0 \end{bmatrix} y$$

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

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

$$\left(A\underset{\sim}{y} - \frac{1}{n_1} 1_{n_1}^T Ay\right)^T (\quad)$$

# $SS_{LOF} \sim \chi^2$

$SS_{PE} + SS_{LOF} \rightarrow m-2$

RSS $\quad n-2$

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

Confidence intervals & regions (CI)

CI on the regression coefficient :

$H_0 : B_j = 0$ vs $H_1 : B_j \neq 0$

$(1-\alpha)\%$ CI for $B_j$

$t \geq t_{n-p}^{\alpha/2}$

$P(|t| > t_{n-p}^{\alpha/2}) = \alpha$

$P(|t| \leq t_{n-p}^{\alpha/2}) = 1-\alpha$

$$t = \frac{\hat{B}_j - B_j}{\sqrt{var(\hat{B}_j)}}$$

$$\cancel{\hat{B}_j}\ \hat{B}_j - t_{n-p}^{\alpha/2}\sqrt{\hat{var}(\hat{B}_j)} \le B_j$$
$$\le \hat{B}_j + t_{n-p}^{\alpha/2}\sqrt{\hat{var}(\hat{B}_j)}$$

$$\hat{var}(\hat{B}_j) = \sigma^2 (X^T X)_{jj} = MS_{e_{jj}}$$

*

## CI on regression parameters

$H_0: B_j = 0 \qquad \dfrac{\hat{B}_j - B_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \quad ; j = 0, \ldots$

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

$$E_j = [\hat{B}_j - t_{\alpha/2, n-p}\sqrt{\hat{\sigma}^2 C_{jj}}$$

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

$$P(E_j) = 1 - \alpha_j$$

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

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

$j$ = no. of statement

Date
Page

## Bonferroni's $t$ - intervals

$t_{\alpha/2k, n-p}$ $\qquad P(\cap E_j) \geq 1 - k\frac{\alpha}{k}$

$$P(E_j) = 1 - \frac{\alpha}{2k} \qquad = 1-\alpha$$

## Maximum modulus intervals-

Define $T_j = \dfrac{\alpha_j^T \hat{\beta} - \alpha_j \beta_{H_0}}{\sqrt{\hat{var}(\alpha_j^T \hat{\beta})}}$

$$\hat{var}(\alpha_j^T \hat{\beta}) = \hat{\sigma}^2 \alpha_j^T (x^T x)^{-1} \underset{\sim}{\alpha_j}$$

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

$$\left[\hat{\beta}_j \pm U^{\alpha}_{k, n-p} \sqrt{\hat{var}(\hat{\beta}_j)}\right.$$

## Scheffe's method

Without loss of generality assume that first $d$ vector of the set $\{a_1, \ldots, a_k\}$ 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 $\hat{\phi} = A\hat{\beta}$

Date
Page

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

## Calculating degree of freedom of quadratic

$$y \sim N(0, V)$$

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

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

## Maximum modulus t-intervals:

$$P\left[\max_{j=1,\dots,k} |T_j| \le U_{k, n-p}\right] = 1 - \alpha$$

$H_0$: $A\beta = 0$

$$T = \max_{j=1 \dots k} |T_j|$$

$$\begin{cases} \alpha_1^T \beta = 0 \\ \alpha_2^T \beta = 0 \Rightarrow \text{Indep} \\ \alpha_3^T \beta = 0 \\ \vdots \\ \alpha_k^T \beta = 0 \end{cases}$$

## Bonferroni's interval -

$$\alpha_j^T \hat{\beta} \pm t_{n-p}^{\frac{\alpha}{2k}} \sqrt{\hat{var}(\alpha_j^T \beta)}$$

## Scheff's method

$$\phi = A\beta$$

$$F = \frac{(\hat{\phi} - \phi)^T \left[A (X^T X)^{-1} A^T\right]^{-1} [\hat{\phi} - \phi]}{MS_{Res} = \frac{RSS}{n-p} = \hat{\sigma}^2} \sim F_{d, n-p}$$

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

$$= \left[\cancel{F}(\hat{\phi}-\phi)^{T} L^{-1} (\hat{\phi}-\phi) \le m\right] \text{ say}$$

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

$$m = MS_{Res}\, F^{\alpha}_{d,n-p}$$

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

Result: If $L$ is p.d. then, for any $b$

$$\left\{ \underset{h \neq 0}{\overset{\text{Sup}}{\max}} \frac{(h^{T}b)^{2}}{h^{T}Lh} = b^{T}L^{-1}b \right\}$$

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

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

$$= P\left[\frac{|h^{T}\hat{\phi} - h^{T}\phi|}{\sqrt{MS_{Res}}\,(h^{T}Lh)^{\frac{1}{2}}} \le \left(dF^{\alpha}_{d,n-b}\right)^{\frac{1}{2}}\right]$$

C.I for $h^{T}\phi$ $\forall h \neq 0$

$$\left[h^{T}\phi \pm \left(dF^{\alpha}_{d,n-b}\right)^{\frac{1}{2}} \sqrt{MS_{Res}(h^{T}Lh)}\right]$$

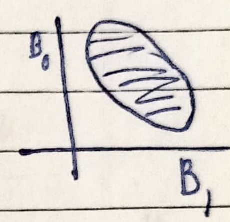
## Confidence region

we know that $\dfrac{(\hat{\beta}-\beta)^T (X^TX)(\hat{\beta}-\beta)}{p\, MS_{Res}} \sim F^{\alpha}_{p,n-p}$

$(\hat{\sigma}^2)$

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

$$\mathbb{R} = \{ B : F \le F^{\alpha}_{p,n-p} \}$$



---

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

$$X^TX = \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}$$

$$(\hat{\beta}-\beta)^T = \left[ (\hat{\beta}_0-\beta_0), (\hat{\beta}_1-\beta_1) \right]$$

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

## Residual analysis

$e_i = y_i - \hat{y}_i$ , $\hat{Y} = X\hat{B}$

(1) Sample mean

$$\bar{e} = \frac{\Sigma e_i}{n} = 0 \qquad e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

Date ______
Page ______

**Proof** $\underset{\sim}{1}^T \underset{\sim}{e} = \underset{\sim}{1}^T (y - \hat{y})$

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

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

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

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

$$X(X^TX)^{-1} X^T [1 \quad x^*] = [1 \quad x^*]$$

$$[X(X^TX)^{-1}X^T 1 \quad X(X^TX)^{-1}X^Tx^*] = [1 \quad x^*]$$

expanding (1) - $\underset{\sim}{1}^T - \underset{\sim}{1}^T = 0$

# $var(\bar{e}) =$

Sample variance $\simeq \dfrac{\sum(e_i - \bar{e})^2}{n-p}$ when $n$ is large.

$$= \frac{\sum e_i^2}{n-p} = MS_R$$

**Methods for scaling residuals**

$$d_i = \frac{e_i}{\sqrt{MS_{Res}}}, \quad i = 1, 2, \ldots n$$

$d_i > 3$

$\Rightarrow$ $i$-th obs is outlier.

Date
Page

## Studentized Residuals-

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

$$Var(\underset{\sim}{e}) = \sigma^2[I - H] \quad \text{where } H = X(X^TX)^{-1}X^T$$

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

$$r_i = \frac{e_i}{\sqrt{MS_{Res}(1 - h_{ii})}} \qquad H = (h_{ij})$$

## PRESS residuals.

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

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

$$\hat{y}_{(1)} = \underset{\sim}{x}_i (X_1^T X_1)^{-1} X_{(1)}^T y_{(1)}$$

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

**Result:** $(X^TX - \underset{\sim}{x}\underset{\sim}{x}^T)^{-1}$

$\underset{\sim}{x} \to$ some vector

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

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

In book (Appendix)

Date
Page

$$y = B_0 + B_1 x_{i1} + B_2 x_{i2}$$

deleted 1st row

$$X_{(1)}^T X_{(1)} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{21} & x_{31} & \cdots & x_{n1} \\ \vdots & & & \\ x_{h1} & & \cdots & x_{nn} \end{bmatrix} \begin{bmatrix} 1 & x_{21} & x_{22} \\ 1 & \vdots & \vdots \\ \vdots & & \\ 1 & x_{n1} & n2 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} n & \Sigma x_{i1} & - \\ & & \\ & & \end{bmatrix}$$

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

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

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

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

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

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

$$= \big[(1 - h_{ii}) y_i - (1 - h_{ii}) x_i^T (X^T X)^{-1} X_{(i)}^T y_{(i)}$$
$$- h_{ii} x_i^T (X^T X)^{-1} X_{(i)}^T y_{(i)} \big] / (1 - h_{ii})$$

Date
Page

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

Since $\quad X^T Y = X_{(i)}^T Y_i + x_i y_i$

$$= (1-h_{ii}) y_i - x_i^T (X^T X)^{-1} (X^T y - x_i y_i)$$

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

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

$$\mathrm{Var}(e_{(i)}) = \frac{\sigma^2 (1-h_{ii})}{(1-h_{ii})^2} = \frac{\sigma^2}{1-h_{ii}}$$

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

R - Student / Externally studentized residuals

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

$$\frac{e_i}{1-h_{ii}}$$

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

$$H_{(i)} = X_{(i)} (X_{(i)}^T X_{(i)})^{-1} X_{(i)}^T \qquad H = X (X^T X)^{-1} X^T$$

$$S_{(i)}^2 = \frac{SS_{Res(i)}}{n-p-1}$$

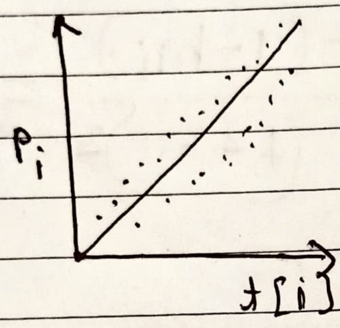
## Normal ~~normal~~ probability ~~plot~~ plot:

Let $t_{[1]} < t_{[2]} < \cdots < t_{[n]}$ be the R-student in increasing order. If we plot $t_{[i]}$ against the cumulative probability $P_i = (i - \frac{1}{2})/n$ for $i = 1, 2, \ldots n$

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

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

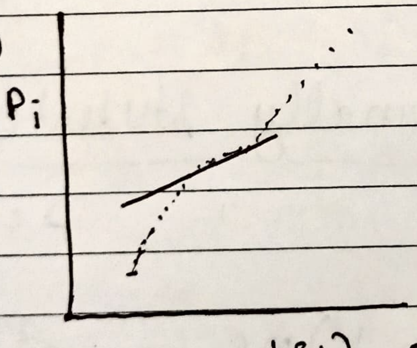
(a)



this plot should come in case of normal variables.

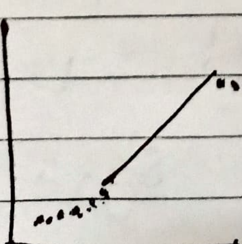
Ideal plot, error is normally distrib.

(b)

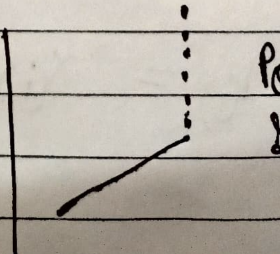


Sharp upward and downward curve at both extremes indicating that the tails of this distribution are too light to be considered normal.

(c)

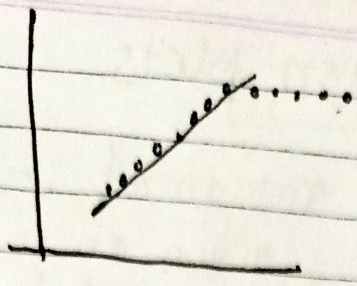


(d)



Positive skewed.

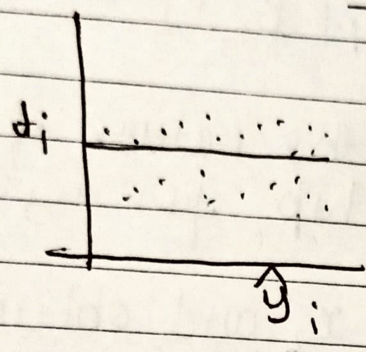
(e)



Negative Skewed

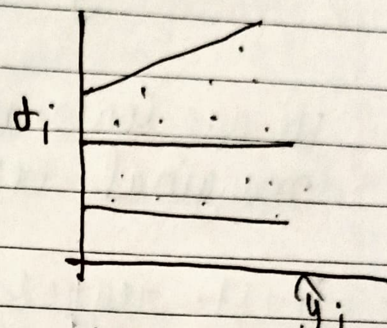
## Plot of residuals against the fitted values

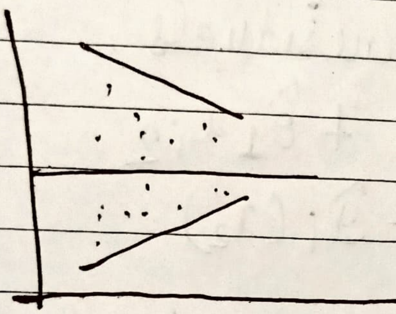
a)




Ideal

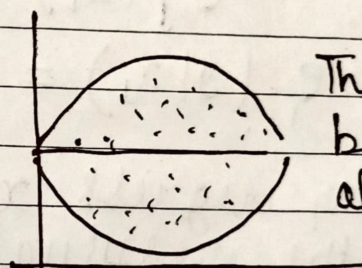
b)



Outward funnel pattern implies that variance is an increa func.

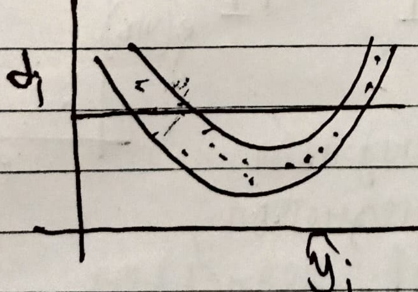
c)



The double bow pattern often occurs when $y$ is proportion b/w 0 & 1.

Ex→ Binomial proportion varn.

(d)




A curved plot such ~~Heal~~ as this indicates ~~normality~~ non-linear.

# Partial regression plots

PRsPs consider the marginal role of the regressor $x_j$ given other regressor that are already in model.

To illustrate:

$$y = B_0 + B_1 x_1 + B_2 x_2 + \varepsilon$$

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

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

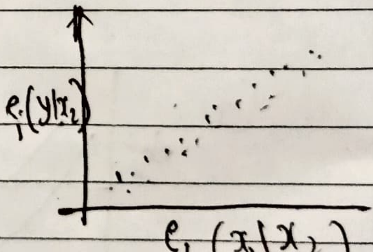
$$\hat{y}_j(x_1) = \hat{\theta}_0 + \hat{\theta}_1 x_{i2}$$

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

Now, regress $x_1$ on $x_2$ and calculate the residuals

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

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




\# first apply on one regressor, then w.r.t that regressor, apply on rest of regressor.

$$e(Y|X_{(j)})$$

$$e(x_j|X_{(j)})$$

Type I error → when $H_0$ is true, but we reject $H_0$

Type II error → when $H_0$ is false, but we do not reject $H_0$.

t-test → It is used for comparing means of two proportions. the data used used in this test should follow normal distrn.

F-test → It is used for comparing variance of two proportions. It is not necessary to values used in this test not to be follow from N-distrn.

Deterministic model → a deterministic function always returns the same result if given the same input values. Actually, it allows you to calculate a future event exactly, without the involvement of randomness.

$$E(Y) = \beta_0 + \beta_1 x \rightarrow \text{Linear model}$$

$$S_{xy} = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$S_{xx} = \frac{\Sigma (x_i - \bar{x})^2}{n-1}$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$S_x = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$$

Correla