SC224:

Statistics:

$\boxed{\theta = 5}$

$\theta \to$ Unknown Parameter.

Metric

$e_i = \theta - \hat{\theta}_i$ —— (1)

$i = 1 \to \theta = 5,\quad \hat{\theta} = 5.2$

$i = 2 \to \quad \hat{\theta} = 4.8$

$i = 3 \to \quad \hat{\theta} = 5.1$

$i = 4 \to \quad \hat{\theta} = 4.9$

$\bar{e}_i = \frac{1}{N}\sum_{i=1}^{N}(\theta - \hat{\theta}_i)$

$\to$ "0"

$\hat{\theta} \to$ Estimated value of θ.

$\boxed{i = 1, 2, \ldots, N}$

$X_i \to \boxed{\text{Black Box}} \to y_i$

Mathematical Modeling

$\hat{\theta} \to [X, Y]$

$\hat{\theta} = (\quad)$

Boeing 737 Max → Crash.

MCAS

2 site.

10 March 2019.

Ethiopia to Kenya

Dec 2018

Indonesia.

Metric.

2

$$e_i = (\theta - \hat{\theta}_i)^2$$

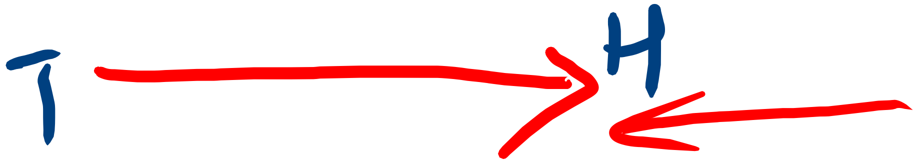
$$\bar{e}_i = \sum_{i=1}^{N} \frac{1}{N} (\theta - \hat{\theta}_i)^2$$

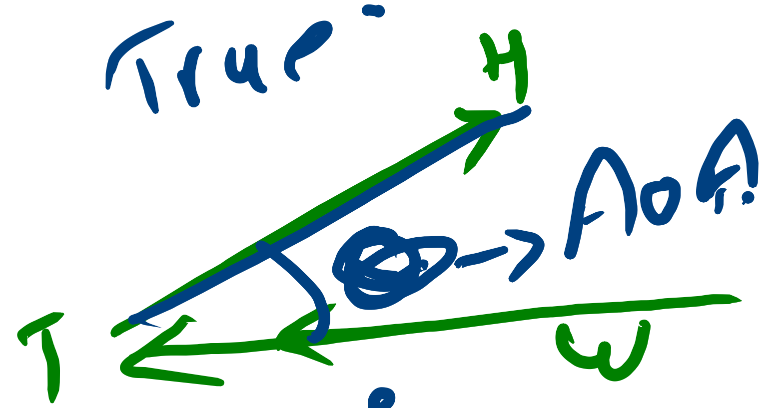
MSE

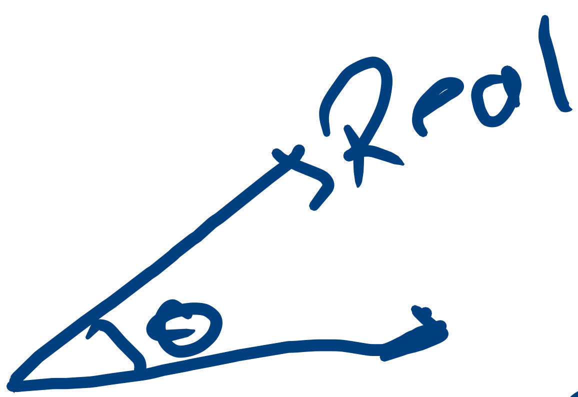
Erroneous MCAS

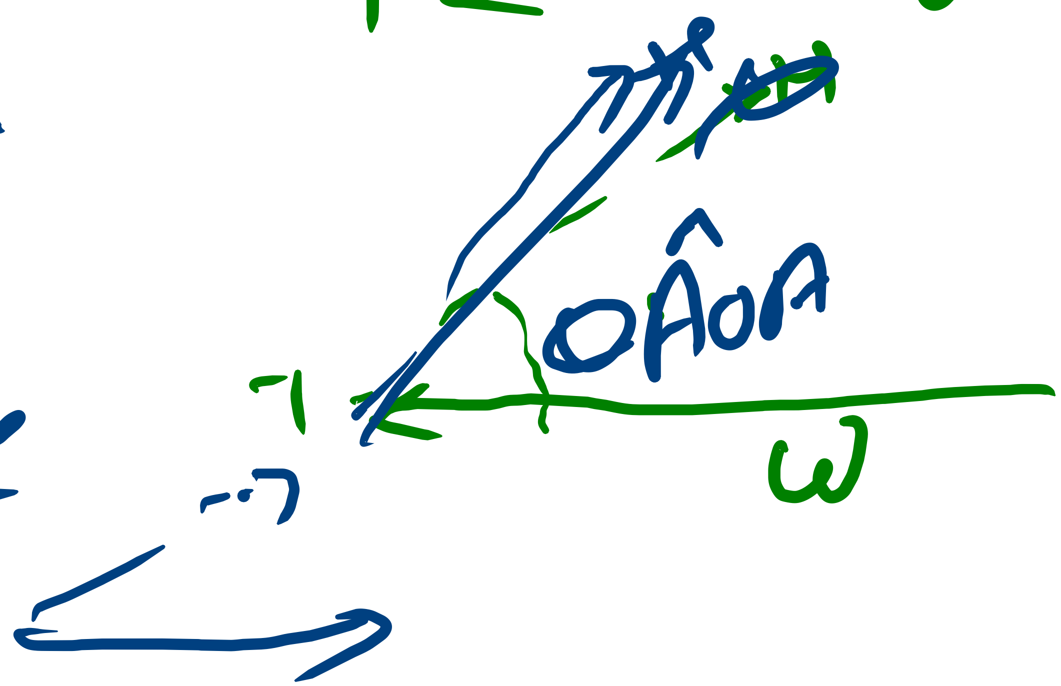
AoA -> Angle of Attack.

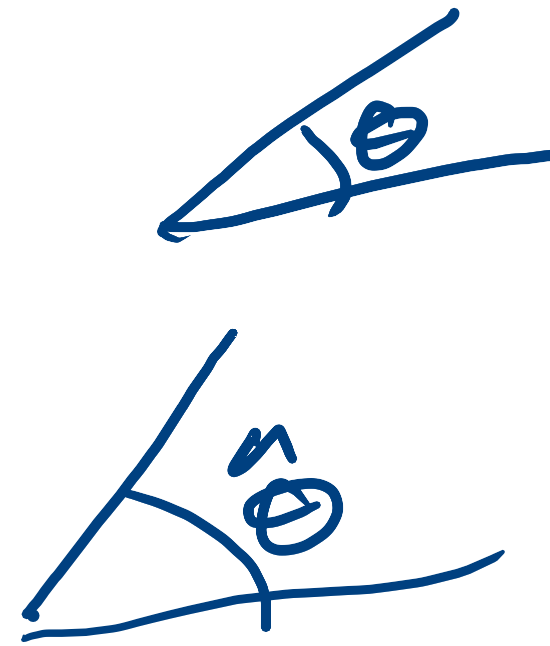
$\hat{AoA}$



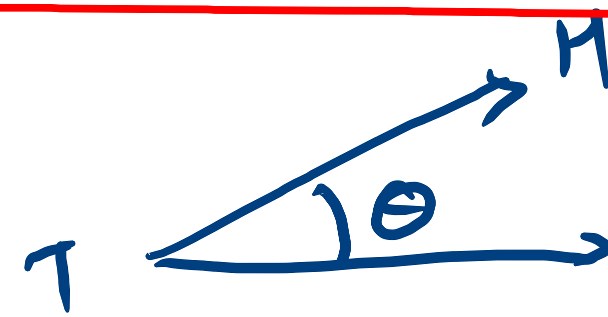






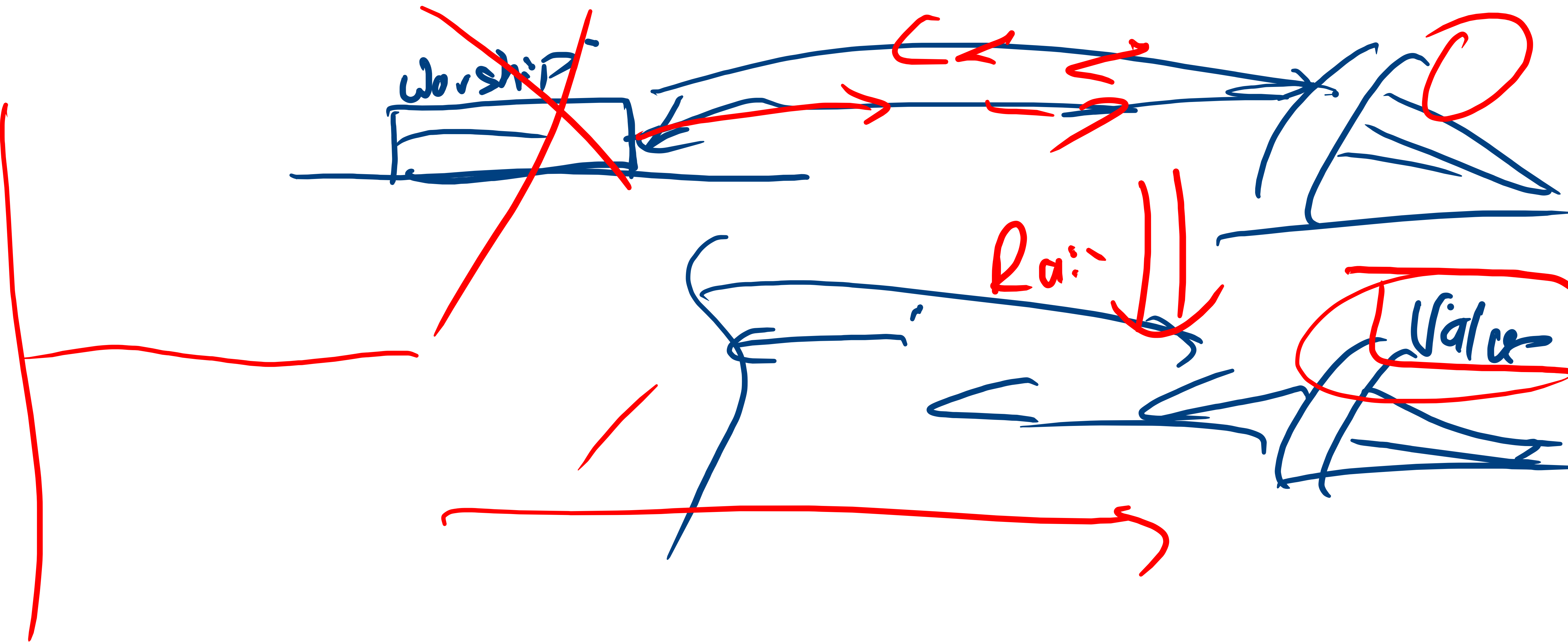


$\boxed{\text{Downfall: The case against Boeing.}}$



cont'd.

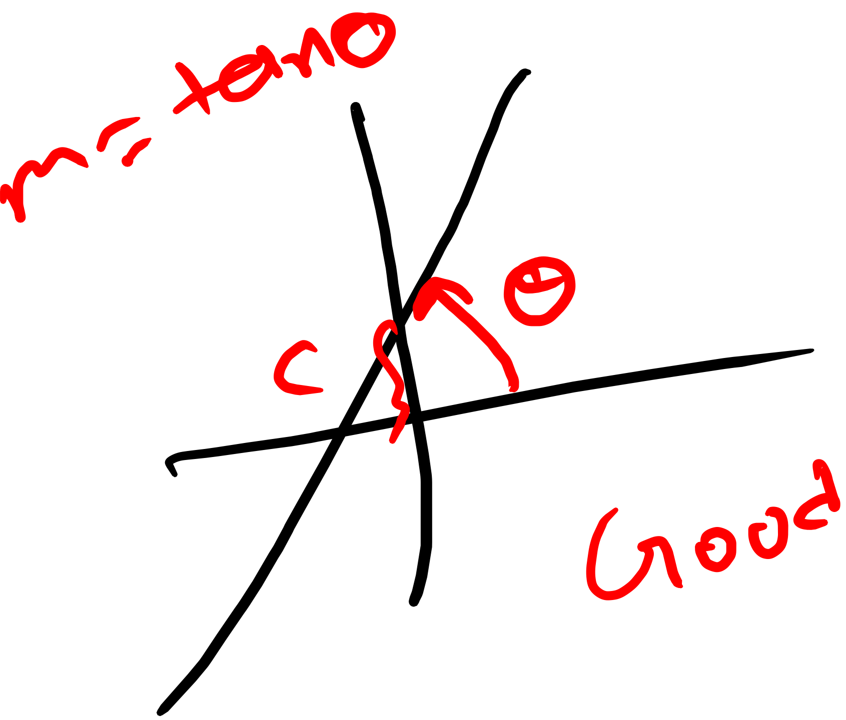
Est.	True	
P	P	
N	P	Misdetection.
P	N	False Alarm
N	N	

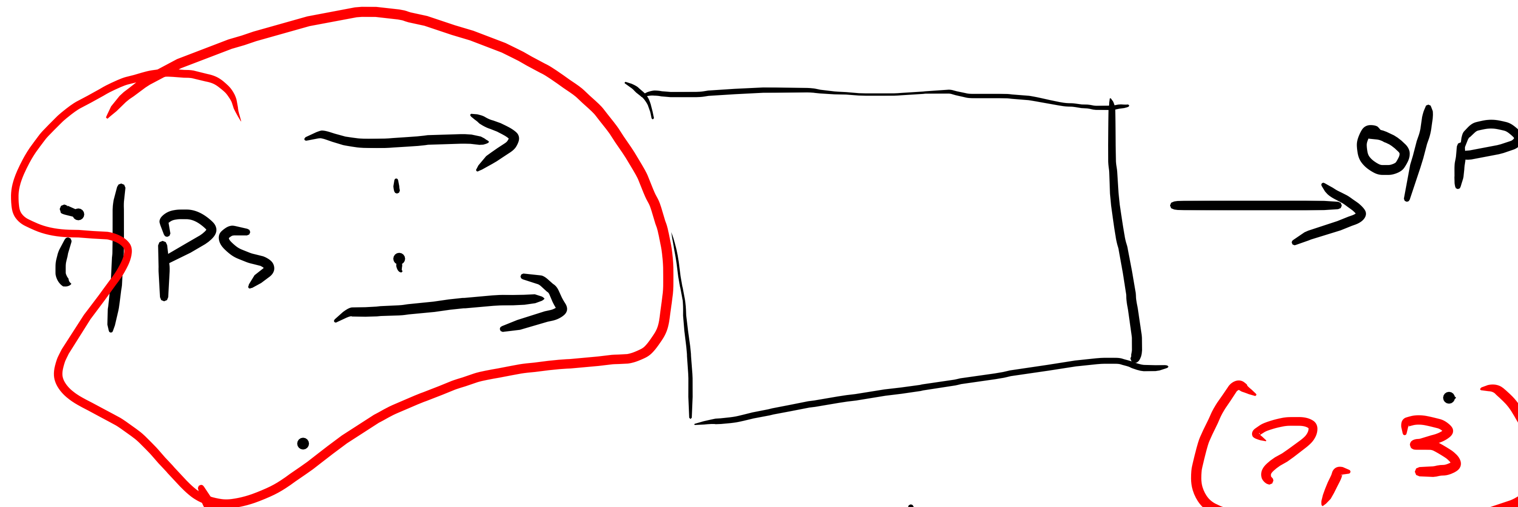


Worship
Ra
Value

Model: → i/Ps , O/P: [x, y]

Model: → Parameter [m, c]





Good Model

$y = f(i/Ps)$

(1) => $\boxed{y = 2x + 4}$

(2, 3)

$y = mx + c$

$\boxed{2 = m(3) + c}$

(2) ↑ Bad Model.

$$\boxed{(x_i,\ y_i)}$$

$$y = f(x_1, x_2 \, \ldots \, x_N)$$

x_i's		y_i's
1		1
2		2
3		3
⋮		⋮
N	X	N

$$y = mx + c$$

$$\boxed{y_i = \beta_0 x_i + \beta_1}$$

LRM.

$x_i \longrightarrow \boxed{\beta_0 \quad \beta_1} \longrightarrow y_i$

$\boxed{y_i = \beta_0 + \beta_1 x_i}$ ⇒ Exact Mathematical Model (Theoretical Model)

$\boxed{y_i = \beta_0 + \beta_1 x_i + \varepsilon_i}$

⇓

Statistical Model (Practical Model)

$$y = f(x_1, x_2, x_3, \dots)$$

y → Yield of crop

x_1 → fertilizers

x_2 → Rain

x_3 → Temp

Model ①

$$\boxed{y = \beta_0 x_1 + \beta_1^2 x_2 + \beta_2 x_3}$$

Model ②

$$\boxed{y = \beta_0 x_1 + \beta_1^2 x_2 + \beta_2 x_3 + \epsilon}$$

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

$$\boxed{y_i = \beta_0 + \beta_1 x_i + \epsilon}$$

LRM

$N = 10$

$i = 1, 2, \ldots N$

$x_i : \{ - - - \}$

$y_i : \{ - - - \}$

$\boxed{(x_i, y_i)\text{'s}}$

Characterize the Parameters

$(\beta_0, \beta_1 : \text{Unknown})$

$\hat{\beta}_0 \quad \hat{\beta}_1$: Estimator for unknown Parameter

* *

eqn 1
$$\hat{\beta}_1 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

Eqn 2
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

RHS

Check for Unbiasedness of Estimate.

$E\{\hat{\beta}_0\} = \beta_0$

$\checkmark$ $\Downarrow$

$y = \beta_0 + \beta_1 x \rightarrow$ linear Model

$y = \beta_0 + \beta_1^2 x$

$\Downarrow$

$y = \beta_0 + \beta_1^* x$ $\quad \beta_1^* = \beta_1^2$

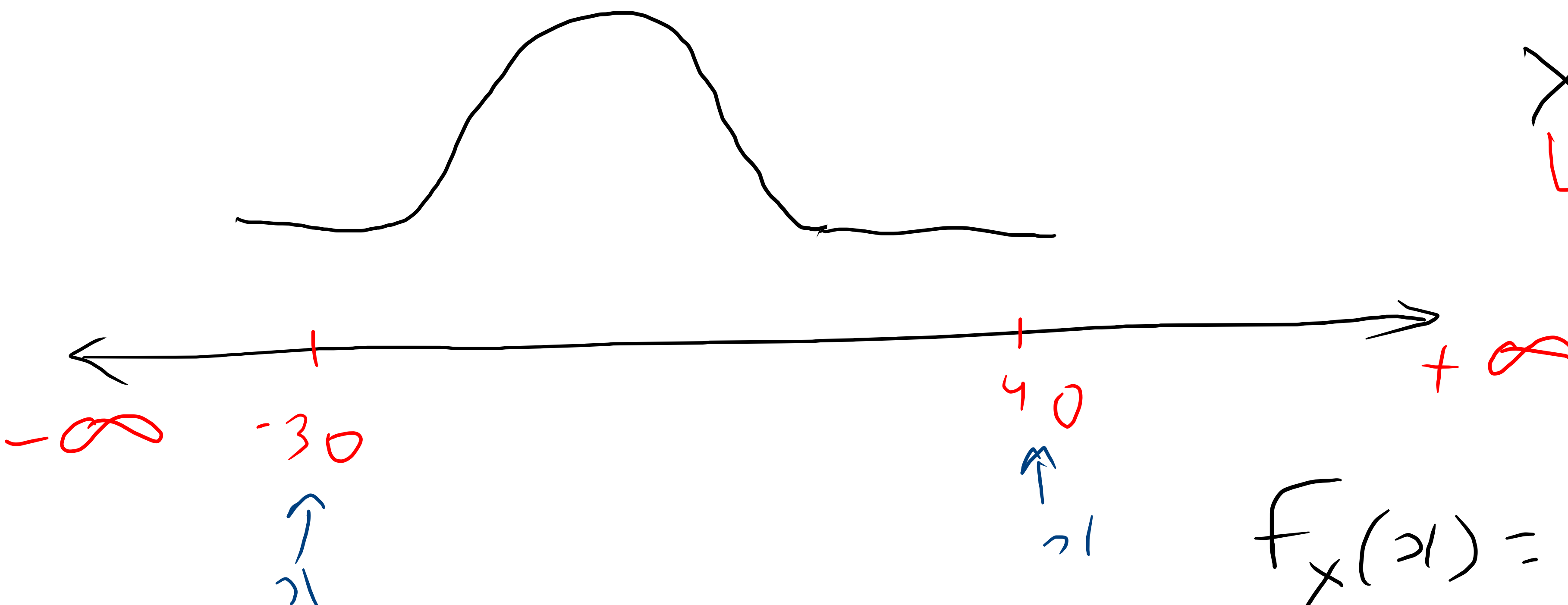
" Unbiased Estimator "

$\beta_0 \longrightarrow \hat{\beta}_0$ $\quad \beta_1 \longrightarrow \hat{\beta}_1$

Normal / Gaussian Distribution:

X : Random Variable

$$\underbrace{X \sim N(\mu, \sigma^2)}$$

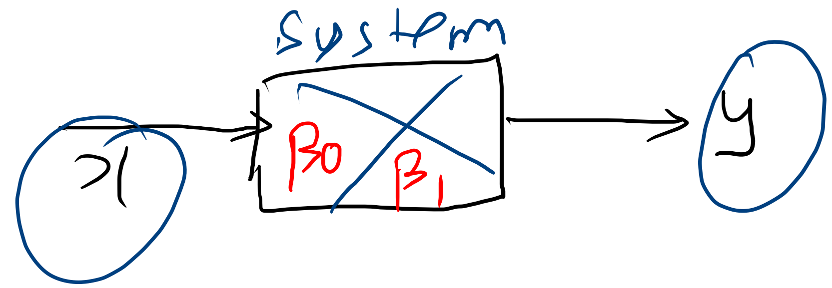


$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$Pr\{x_1 \leq X \leq x_2\} = ??$$

$$\int_{x_1}^{x_2} f_X(x)\, dx$$

LRM:



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

β_0, β_1 : Parameters which are unknown.

(Theoretical Model) $y_i = \beta_0 + \beta_1 x_i$ $\qquad \underbrace{i = 1, 2, \ldots, n}$

$(x_i, y_i) \longrightarrow$ Estimate the unknown Parameter β_0, β_1.

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

(Statistical Model)

✓ $y = 3x + 2$

✗ $5 = m(2) + c$

$\beta_0 \longrightarrow \hat{\beta}_0 \;;\; e_0 = (\hat{\beta}_0 - \beta_0) \;;\; \underbrace{e_0 \to 0}$

$\beta_1 \longrightarrow \hat{\beta}_1 \;;\; e_1 = (\hat{\beta}_1 - \beta_1) \;;\; \underbrace{e_1 \to 0}_{(x_i\,,\,y_i)}$

$\hat{\beta}_0 = \beta_0$

$\hat{\beta}_0 \neq \beta_0$

$\Theta \longrightarrow \hat{\Theta}$

error | Devi. Est.

$= (\hat{\beta}_0 - \beta_0)$

Estimated output: $\hat{\beta}_0 + \hat{\beta}_1 x_i$

Residual Error

r_i

Error $e_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$

r_i^2 Squared Error $e_i^2 = [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2$

$$f(\hat{\beta}_0, \hat{\beta}_1) = \underbrace{\sum_{i=1}^{n} e_i^2}_{\text{Metric}} = \sum_{i=1}^{n} [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2$$

LSE Estimator: Least Squared Error Estimator

$$\frac{\partial f(\cdot)}{\partial \hat{\beta}_0} = 0 \; ; \; \frac{\partial f(\cdot)}{\partial \hat{\beta}_1} = 0.$$

$\hat{\beta}_0 = (\quad)$

$\hat{\beta}_1 = (\quad)$

$$\frac{\partial}{\partial \hat\beta_0} F(\hat\beta_0, \hat\beta_1) = -2 \sum_{i=1}^{n} (y_i - \hat\beta_0 - \hat\beta_1 x_i) = 0$$

$$\Rightarrow \quad \hat\beta_0 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat\beta_1 x_i)$$

$$= \underbrace{\frac{1}{n} \sum_{i=1}^{n} y_i}_{} - \hat\beta_1 \cdot \underbrace{\frac{1}{n} \sum_{i=1}^{n} x_i}_{}$$

$\boxed{(x_i\text{s}, y_i\text{s})}$

$$\boxed{\hat\beta_0 = \bar{y} - \hat\beta_1 \bar{x}}$$

$\hat\beta_0 = (\quad)$ $\qquad\qquad$ $\hat\beta_1 = (\quad)$

$$\frac{\partial}{\partial \hat{\beta}_1} f(\hat{\beta}_0, \hat{\beta}_1) = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0$$

We know, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$$\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2}$$

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

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i - n \bar{y} \bar{x} + n \bar{x}^2 \hat{\beta}_1}{\quad}$$

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

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

ROUGH

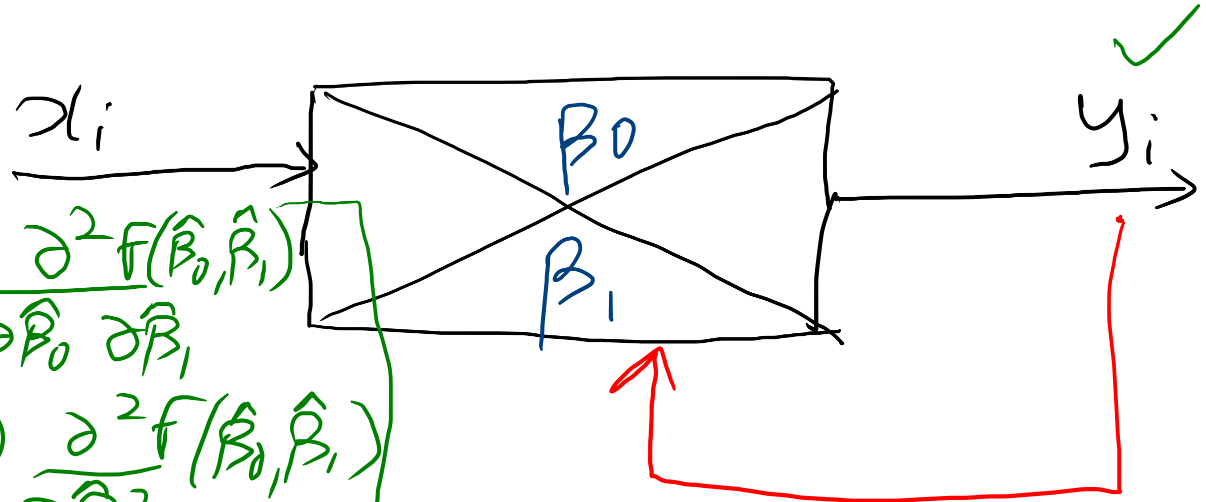
$$\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n}\left[x_i^2+\bar{x}^2-2\bar{x}x_i\right]$$

$$= \sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x}\underbrace{\sum_{i=1}^{n} x_i}$$

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

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

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



$$H = \begin{bmatrix} \dfrac{\partial^2 f(\hat{\beta}_0,\hat{\beta}_1)}{\partial \hat{\beta}_0^2} & \dfrac{\partial^2 f(\hat{\beta}_0,\hat{\beta}_1)}{\partial \hat{\beta}_0\, \partial \hat{\beta}_1} \\ \dfrac{\partial^2 f(\hat{\beta}_0,\hat{\beta}_1)}{\partial \hat{\beta}_1\, \partial \hat{\beta}_0} & \dfrac{\partial^2 f(\hat{\beta}_0,\hat{\beta}_1)}{\partial \hat{\beta}_1^2} \end{bmatrix}$$

$i = 1, 2, \ldots, n$

Data Set

(x_i, y_i)

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

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

$$E\{\hat{\beta}_1\} = \beta_1$$

$$H = \begin{bmatrix} 2n & 2n\bar{x} \\ 2n\bar{x} & 2\sum_{i=1}^{n} x_i^2 \end{bmatrix}$$

$$E\{\hat{\beta}_0\} = \beta_0$$

$$\det(H) = 4n \sum_{i=1}^{n} (x_i - \bar{x})^2 \geq 0$$

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

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) y_i}{S_{xx}} - \bar{y} \frac{\cancel{\sum_{i=1}^{n} (x_i - \bar{x})}}{S_{xx}}$$

0

$$\boxed{\text{Let } \quad c_i = \frac{x_i - \bar{x}}{S_{xx}}} \Rightarrow \boxed{\hat{\beta}_1 = \sum_{i=1}^{n} c_i y_i}$$

Estimator $\hat{\beta}_1$ is a linear function of output/response y_i's

$$\hat{\beta}_1 = \sum_{i=1}^{n} c_i y_i \qquad \text{where } c_i = \frac{x_i - \bar{x}}{S_{xx}}$$

$$\hat{\beta}_1 = \sum_{i=1}^{n} c_i (\beta_0 + \beta_1 x_i + \varepsilon_i)$$

$$\varepsilon_i \overset{IID}{\sim} N(0, \sigma^2)$$

Random Error

$$\boxed{\hat{\beta}_1 = \sum_{i=1}^{n} c_i \beta_0 + \sum_{i=1}^{n} c_i \beta_1 x_i + \sum_{i=1}^{n} c_i \varepsilon_i}$$

$$\sum_{i=1}^{n} c_i \beta_0 = \beta_0 \underbrace{\sum_{i=1}^{n} c_i} = \beta_0 \frac{\overbrace{\sum_{i=1}^{n} (x_i - \bar{x})}}{S_{xx}} = 0$$

$$\sum_{i=1}^{n} c_i \beta_1 x_i = \beta_1 \underbrace{\sum_{i=1}^{n} c_i x_i}_{1} = \beta_1$$

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

ROUGH

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

$$\sum_{i=1}^{n} c_i x_i = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = 1$$

$$\hat{\beta}_1 = \beta_1 + \sum_{i=1}^{n} c_i \varepsilon_i$$

$$E\{\hat{\beta}_1\} = E\left\{\beta_1 + \sum_{i=1}^{n} c_i \varepsilon_i\right\}$$

$$= E\{\beta_1\} + \sum_{i=1}^{n} c_i \underbrace{E\{\varepsilon_i\}}$$

$$\boxed{E\{\hat{\beta}_1\} = \beta_1}$$

Check for unbiasedness of Estimator $\hat{\beta}_1$.

$$\text{Bias}(\hat{\beta}_1, \beta_1) = E(\hat{\beta}_1) - \beta_1 = 0$$

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

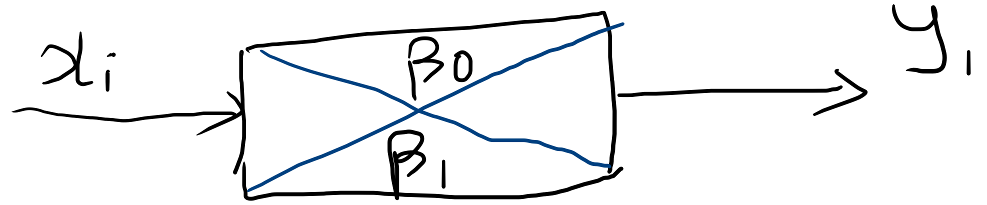
β_0, β_1 : Unknown Parameters

(non-random/ Deterministic)

x_i : Input Values

(non-random)

$\varepsilon_i \overset{IID}{\sim} N(0, \sigma^2)$



(x_i, y_i)

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

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

Check for unbiasedness of $\hat{\beta}_0$

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

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

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

$$= E\{\bar{y}\} - \bar{x}\, \underbrace{E\{\hat{\beta}_1\}}$$

$$\boxed{E\{\hat{\beta}_0\} = \underbrace{E\{\bar{y}\}} - \beta_1 \bar{x}} \Rightarrow \boxed{E\{\hat{\beta}_0\} = \beta_0}$$

Unbiased Estimator

$Bias(\hat{\beta}_0, \beta_0)$

$= E(\hat{\beta}_0) - \beta_0$

$= 0$

$$E\{\bar{y}\} = E\left\{\frac{y_1 + y_2 + \cdots + y_n}{n}\right\}$$

$$= \frac{1}{n}\left[E(y_1) + E(y_2) + \cdots + E(y_n)\right]$$

$$= \frac{1}{n}\left[(\beta_0 + \beta_1 x_1) + (\beta_0 + \beta_1 x_2) + \cdots + (\beta_0 + \beta_1 x_n)\right]$$

$$E(\bar{y}) = \frac{1}{n}\left[n\beta_0 + \beta_1 \cdot n\bar{x}\right] = \beta_0 + \beta_1 \bar{x}$$

ROUGH

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad ; \quad \varepsilon_i \overset{IID}{\sim} N(0, \sigma^2)$$

$i = 1, 2, \ldots, n$

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

$$\boxed{E(y_i) = \beta_0 + \beta_1 x_i}$$

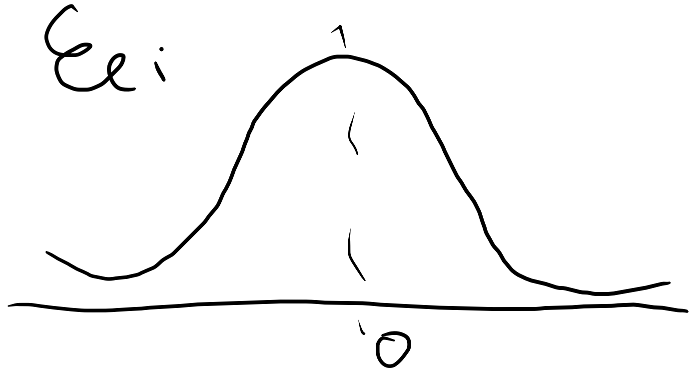
$$Var(y_i) = Var(\beta_0 + \beta_1 x_i + \varepsilon_i) = Var(\beta_0 + \beta_1 x_i) + Var(\varepsilon_i)$$

$$\boxed{Var(y_i) = \sigma^2}$$

$$y_i \overset{InID}{\sim} N(\underbrace{\beta_0 + \beta_1 x_i}_{E(y_i)}, \underbrace{\sigma^2}_{Var(y_i)})$$

InID

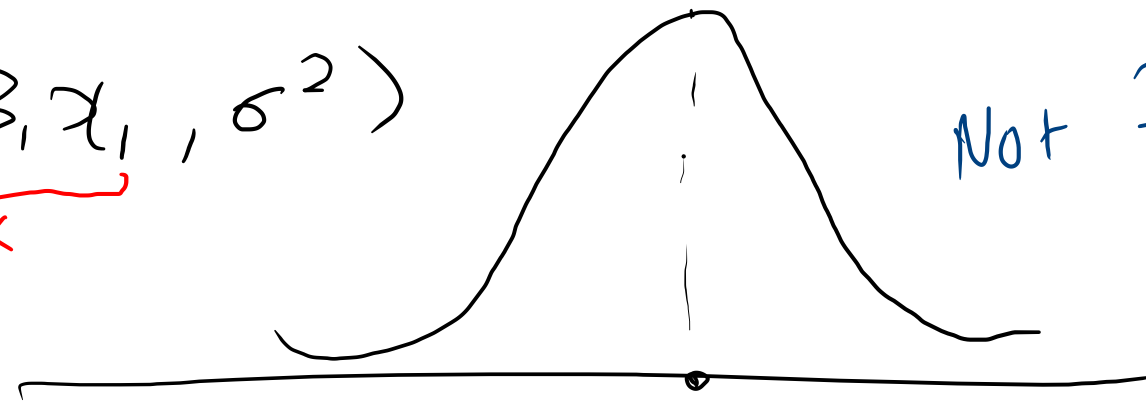
ε_i



ROUGH

$y_1 \sim N(\underbrace{\beta_0+\beta_1 x_1}_{\times}, \sigma^2)$

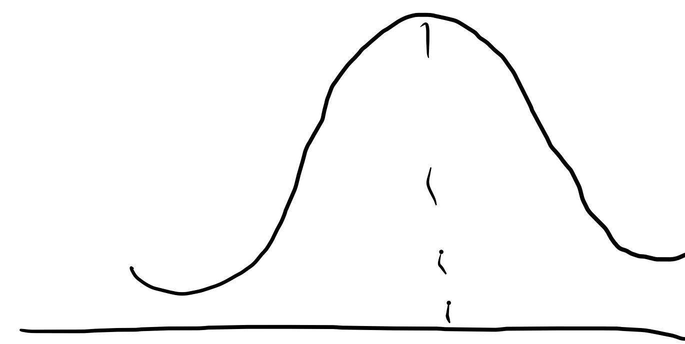
$y_1 = \cancel{\alpha_1} + \varepsilon_1$



Not Identically Distributed.

$y_2 \sim N(\underbrace{\beta_0+\beta_1 x_2}_{\times}, \sigma^2)$

$y_2 = \cancel{\alpha_2} + \varepsilon_2$



$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \sum_{i=1}^{n} c_i y_i$$

$$\begin{aligned}
\mathrm{Var}(\hat{\beta}_1) &= \mathrm{Var}\left(\sum_{i=1}^{n} c_i y_i\right) \\
&= \sum_{i=1}^{n} c_i^2 \,\mathrm{Var}(y_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \underbrace{\mathrm{Cov}(y_i, y_j)}_{0} \\
&= \sum_{i=1}^{n} c_i^2 \,\mathrm{Var}(y_i)
\end{aligned}$$

$$\mathrm{Var}(\hat{\beta}_1) = \sigma^2 \underbrace{\sum_{i=1}^{n} c_i^2} = \frac{\sigma^2}{S_{xx}}$$

$$\sum_{i=1}^{n} c_i^2 = \sum_{i=1}^{n} \left[\frac{x_i - \bar{x}}{S_{xx}} \right]^2$$

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

$$\sum_{i=1}^{n} c_i^2 = \frac{1}{S_{xx}^2} \cdot S_{xx} = \frac{1}{S_{xx}}$$

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

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

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

$$\mathrm{Var}(\hat{\beta}_0) = \underbrace{\mathrm{Var}(\bar{y})}_{\frac{\sigma^2}{n}} + \bar{x}^2 \underbrace{\mathrm{Var}(\hat{\beta}_1)}_{\to \frac{\sigma^2}{S_{xx}}} - 2\bar{x} \underbrace{\mathrm{Cov}(\bar{y}, \hat{\beta}_1)}_{\to 0}$$

$$\boxed{\mathrm{Var}(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right]}$$

$$Cov(\bar{y}, \hat{\beta}_1) = E\left\{ [\bar{y} - E\{\bar{y}\}] \cdot [\hat{\beta}_1 - E\{\hat{\beta}_1\}] \right\}$$

$$Cov(X, Y) = E\left\{ [X - E(X)] \cdot [Y - E(Y)] \right\}$$

$$= E\{XY\} - E\{X\} \cdot E\{Y\}$$

$$Cov(\bar{y}, \hat{\beta}_1) = E\left\{ \underbrace{[\bar{y} - E\{\bar{y}\}]} \cdot \underbrace{[\hat{\beta}_1 - \beta_1]} \right\}$$

$$= E\left\{ \bar{\varepsilon} \cdot \sum_{i=1}^{n} c_i \varepsilon_i \right\}$$

$$\bar{y} - E(\bar{y}) = \frac{1}{n}\sum_{i=1}^{n} y_i - E\left\{\frac{1}{n}\sum_{i=1}^{n} y_i\right\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\beta_0 + \beta_1 x_i + \varepsilon_{ei}\right) - \frac{1}{n}\sum_{i=1}^{n} E\{y_i\}$$

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_{ei}$

$$= \left(\beta_0 + \beta_1 \bar{x} + \bar{\varepsilon}_e\right) - \left(\beta_0 + \beta_1 \bar{x}\right)$$

$$\boxed{\bar{y} - E(\bar{y}) = \bar{\varepsilon}_e}$$

$$\hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} c_i y_i - \beta_1 = \sum_{i=1}^{n} c_i(\beta_0 + \beta_1 x_i + \varepsilon_i) - \beta_1$$

$$= \beta_0 \underbrace{\sum_{i=1}^{n} c_i}_{0} + \beta_1 \underbrace{\sum_{i=1}^{n} c_i x_i}_{1} + \sum_{i=1}^{n} c_i \varepsilon_i - \beta_1$$

$$\boxed{\hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} c_i \varepsilon_i}$$

ROUGH

$$\sum_{i=1}^{n} c_i = c_1 + c_2 + \cdots + c_n$$

$$= \frac{x_1 - \bar{x}}{s_{xx}} + \frac{x_2 - \bar{x}}{s_{xx}} + \cdots + \frac{x_n - \bar{x}}{s_{xx}}$$

$$= \frac{n\bar{x} - n\bar{x}}{s_{xx}} = 0$$

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

ROUGH

$$\sum_{i=1}^{n} c_i x_i = \sum_{i=1}^{n} \frac{(x_i - \bar{x})}{S_{xx}} \cdot x_i$$

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

$$= \frac{\overbrace{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}^{S_{xx}}}{S_{xx}}$$

$$= 1$$

we know,

$$\underbrace{\operatorname{cov}(\bar{y}, \hat{\beta}_1)}_{\parallel} = E\left\{ \bar{\varepsilon}_e \cdot \sum_{i=1}^{n} c_i \, \varepsilon_{e_i} \right\}$$

$$= \sum_{i=1}^{n} c_i \; E\left\{ \bar{\varepsilon}_e \cdot \varepsilon_{e_i} \right\}$$

$$= \sum_{i=1}^{n} c_i \; E\left\{ \frac{(\varepsilon_{e_1} + \varepsilon_{e_2} + \cdots + \varepsilon_{e_n})}{n} \cdot \varepsilon_{e_i} \right\}$$

$$\operatorname{cov}(\bar{y}, \hat{\beta}_1) = \underbrace{\frac{c_1}{n} E\left\{ \varepsilon_{e_1}^2 + \varepsilon_{e_2} \cdot \varepsilon_{e_1} + \cdots + \varepsilon_{e_n} \cdot \varepsilon_{e_1} \right\}}_{} \rightarrow \frac{c_1}{n} \cdot \sigma^2$$

$$+ \underbrace{\frac{c_2}{n} E\left\{ \varepsilon_{e_1} \cdot \varepsilon_{e_2} + \varepsilon_{e_2}^2 + \cdots + \varepsilon_{e_n} \cdot \varepsilon_{e_2} \right\}}_{}$$

$$\frac{c_2}{n} \cdot \sigma^2 \leftarrow$$

$$+ \cdots + \frac{c_n}{n} E\left\{ \varepsilon_{e_1} \cdot \varepsilon_{e_n} + \varepsilon_{e_2} \cdot \varepsilon_{e_n} + \cdots + \varepsilon_{e_n}^2 \right\}$$

$$\frac{c_1}{n}\left[E\{\varepsilon_{e_1}^2\} + E\{\varepsilon_{e_2}\varepsilon_{e_1}\} + \cdots + E\{\varepsilon_{e_n}\cdot\varepsilon_{e_1}\}\right]$$

ROUGH

$$\mathrm{var}(\varepsilon_{e_i}) = E\{\varepsilon_{e_i}^2\} - \left[\cancelto{0}{E\{\varepsilon_{e_i}\}}\right]^2$$

$$\sigma^2 = E\{\varepsilon_{e_i}^2\}$$

$$E\{\varepsilon_{e_i}\cdot\varepsilon_{e_l}\} = \underbrace{E\{\varepsilon_{e_i}\}}_{0}\cdot\underbrace{E\{\varepsilon_{e_l}\}}_{0}$$

$i \neq l$

$$= 0$$

$i = 1$

$l = 2$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1) = \frac{c_1}{n}\sigma^2 + \frac{c_2}{n}\sigma^2 + \cdots + \frac{c_n}{n}\sigma^2$$

$$= \frac{\sigma^2}{n} \underbrace{\left[c_1 + c_2 + \cdots + c_n\right]}_{0}$$

$$\boxed{\operatorname{Cov}(\bar{y}, \hat{\beta}_1) = 0}$$

Summary

(x_i, y_i) Data Set

$i = 1, 2, \dots, n$

$$x_i \longrightarrow \boxed{\begin{array}{c}\text{Linear Model/System} \\ \hline \beta_0,\ \beta_1\end{array}} \longrightarrow y_i$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$E\{\hat{\beta}_1\} = \beta_1$$

$$E\{\hat{\beta}_0\} = \beta_0$$

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

$$\mathrm{Var}(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$$

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

When σ^2 is unknown:

$\hat{\sigma}^2$: Estimator for unknown Parameter σ^2.

$$SS_{res} = \sum_{i=1}^{n} [y_i - \underbrace{(\hat{\beta}_0 + \hat{\beta}_1 x_i)}_{\hat{y}_i}]^2$$

$\hat{y}_i$: Fitted Model
Estimated Regression Line

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

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

$$= \underbrace{\sum_{i=1}^{n} (y_i - \bar{y})^2}_{S_{yy}} + \hat{\beta}_1^2 \underbrace{\sum_{i=1}^{n} (x_i - \bar{x})^2}_{S_{xx}} - 2\hat{\beta}_1 \underbrace{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}_{S_{xy}}$$

$$\boxed{SS_{res} = S_{yy} + \hat{\beta}_1^2 S_{xx} - 2\hat{\beta}_1 S_{xy}}$$

We Know, $\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}} \Rightarrow S_{xy} = \hat{\beta}_1 \cdot S_{xx}$

$$SS_{res} = S_{yy} + \hat{\beta}_1^2 S_{xx} - 2\hat{\beta}_1^2 S_{xx}$$

$$\Rightarrow \boxed{SS_{res} = S_{yy} - \hat{\beta}_1^2 S_{xx}}$$

$\downarrow$ On deriving

$$E\{SS_{res}\} = (n-2)\sigma^2 \Rightarrow \frac{1}{(n-2)} E\{SS_{res}\} = \sigma^2$$

$$\boxed{E\left\{\frac{SS_{res}}{n-2}\right\} = \sigma^2} \longrightarrow \boxed{\hat{\sigma}^2 = \frac{SS_{res}}{n-2}}$$

θ : Unknown Parameter

$\hat{\theta}$: Estimator of unknown Parameter θ

Check for unbiasedness of the estimator $\hat{\theta}$:

$$E\{\underbrace{\hat{\theta}}_{\text{Estimator of } \theta}\} = \theta$$

Analysis

$$SS_{res} = \sum_{i=1}^{n} [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2$$

$$SS_{res} = \sum_{i=1}^{n} [\underbrace{(\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)x_i}_{\alpha_i} + \varepsilon_i]^2$$

$$SS_{res} = \sum_{i=1}^{n} (\alpha_i + \varepsilon_i)^2$$

$$SS_{res} = \sum_{i=1}^{n} (\alpha_i + \varepsilon_{ei})^2$$

$$\alpha_i = [\beta_0 - \hat{\beta}_0] + [\beta_1 - \hat{\beta}_1] x_i$$

$$\varepsilon_{ei} \sim N(0, \sigma^2)$$

$$\frac{SS_{res}}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{\alpha_i + \varepsilon_{ei}}{\sigma} \right)^2$$

$$\frac{SS_{res}}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{[\beta_0 - \hat{\beta}_0] + [\beta_1 - \hat{\beta}_1] x_i + \varepsilon_{ei}}{\sigma} \right)^2$$

If $\beta_0 = \hat{\beta}_0$, $\beta_1 = \hat{\beta}_1$,

$$\boxed{\frac{SS_{res}}{\sigma^2} = \sum_{i=1}^{n} \tilde{\varepsilon}_{ei}^{\,2}}$$

n summation

$$\tilde{\varepsilon}_{ei} \sim N(0, 1)$$

$$E\left\{\frac{SS_{res}}{\sigma^2}\right\} = (n-2) \Rightarrow E\left\{\frac{SS_{res}}{(n-2)}\right\} = \sigma^2$$

$$\tilde{\varepsilon}_{ei} = \frac{\varepsilon_{ei}}{\sigma}$$

$$E\{\tilde{\varepsilon}_{ei}\} = \frac{1}{\sigma} \underbrace{E\{\varepsilon_{ei}\}}_{0} = 0 \Rightarrow \boxed{E\{\tilde{\varepsilon}_{ei}\} = 0}$$

$$\mathrm{Var}(\tilde{\varepsilon}_{ei}) = \mathrm{Var}\left(\frac{\varepsilon_{ei}}{\sigma}\right)$$

$$= \frac{1}{\sigma^2} \underbrace{\mathrm{Var}(\varepsilon_{ei})}_{\sigma^2}$$

$$\Rightarrow \boxed{\mathrm{Var}(\tilde{\varepsilon}_{ei}) = 1}$$

$$z \sim \chi^2(n)$$

$$z = z_1^2 + z_2^2 + \cdots + z_n^2$$

$$\boxed{E(z) = n}$$

$$\mathrm{Var}(z) = 2n$$

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

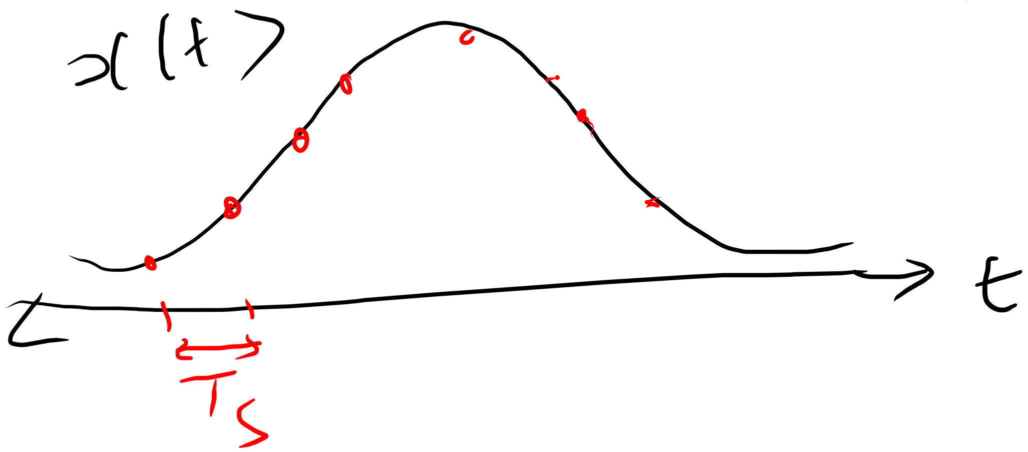
$\varepsilon_i \sim N(0, \underbrace{\sigma^2}_{??})$

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$

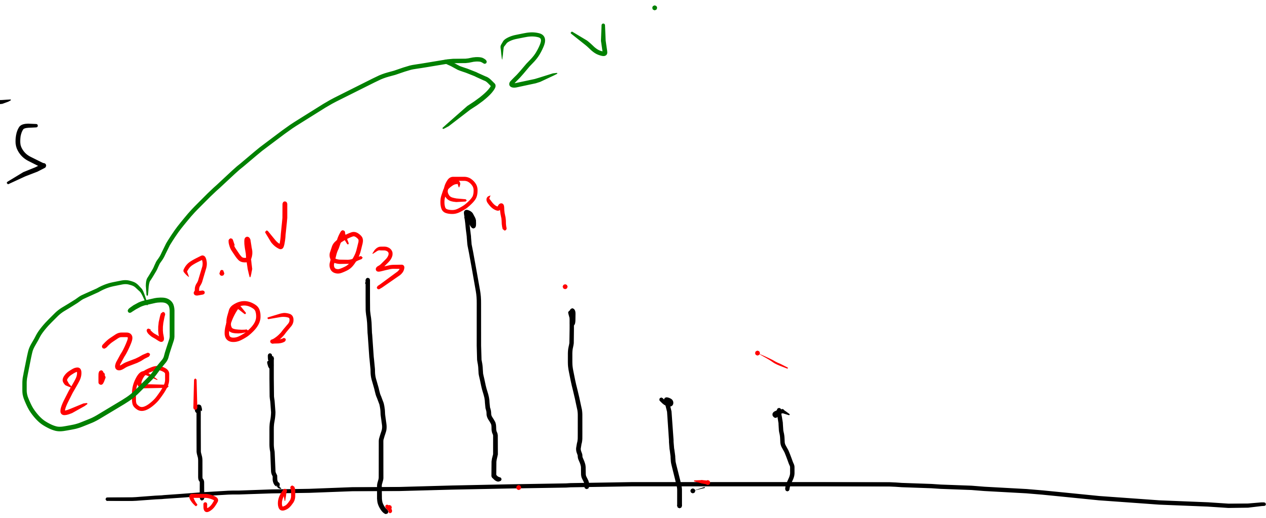
$\boxed{y_i - \hat{y}_i} \to \varepsilon_i \qquad i = 1, 2, \ldots, n$

$$\boxed{\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (\varepsilon_i - \bar{\varepsilon})^2} \; ??$$

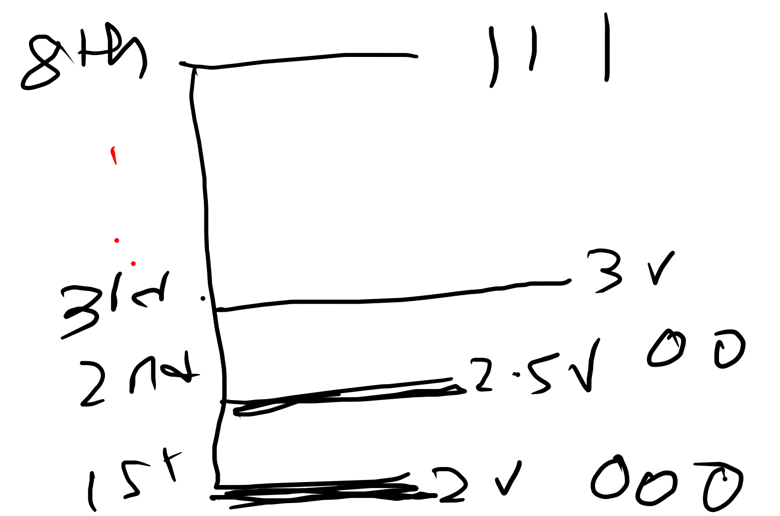
MSE : $E\left\{ (\hat{\beta}_0 - \beta_0)^2 \right\} \downarrow\downarrow$



$t = nT_s$



3 bits → $2^3 = 8$ levels.



Maximum Likelihood Estimation (MLE) framework:

A good estimate (MLE) of the unknown parameter θ would be the value of θ that maximizes the likelihood of getting the data we observed.

$$\boxed{\hat{\theta} = \arg\max_{\theta} \left\{ \mathcal{L}(\theta ; \bar{y}) \right\}}$$

$[y_1, y_2, \ldots, y_n] = \bar{y}$

MLE is a statistical method used to estimate the parameters of a probability distribution that best describe a given dataset. The fundamental idea behind MLE is to find the values of parameters that maximize the likelihood of the observed data, assuming the data are generated by specific distribution.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \varepsilon_i \overset{IID}{\sim} N(0, \sigma^2)$$

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

$$Var(y_i) = \sigma^2 \qquad y_i \overset{InID}{\sim} N(\beta_0 + \beta_1 x_i, \sigma^2)$$

$$f_{y_1, y_2, \ldots, y_n}(y_1, y_2, \ldots, y_n) = f_{y_1}(y_1) \cdot f_{y_2}(y_2) \cdots f_{y_n}(y_n)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[y_1 - (\beta_0 + \beta_1 x_1)]^2}{2\sigma^2}} \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[y_2 - (\beta_0 + \beta_1 x_2)]}{2\sigma^2}} \times \cdots \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[y_n - (\beta_0 + \beta_1 x_n)]^2}{2\sigma^2}}$$

$$f(y_1, y_2, \ldots, y_n; \beta_0, \beta_1, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\beta_0 + \beta_1 x_i)\right]^2}$$

Likelihood function

$$l(\beta_0, \beta_1, \sigma^2; \bar{y}) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\beta_0 + \beta_1 x_i)\right]^2}$$

$y_1, y_2, \ldots y_n$

Log Likelihood function

$[\ln = \log_e]$

$$L(\beta_0, \beta_1, \sigma^2; \bar{y}) = \log_e l(\beta_0, \beta_1, \sigma^2; y)$$

$$\boxed{L(\beta_0, \beta_1, \sigma^2; \bar{y}) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\beta_0 + \beta_1 x_i)\right]^2}$$

$$\frac{\partial L(\cdot)}{\partial \beta_0} = 0 \Rightarrow \beta_0 = (\ ) \rightarrow \hat{\beta}_0^{ML}$$

$$\frac{\partial L(\cdot)}{\partial \beta_1} = 0 \Rightarrow \beta_1 = (\ ) \rightarrow \hat{\beta}_1^{ML}$$

$$\frac{\partial L(\cdot)}{\partial \sigma^2} = 0 \quad \sigma^2 = (\ ) \rightarrow \hat{\sigma}^2_{ML}$$

$$\frac{\partial L(\cdot)}{\partial \beta_0} = \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_i \right)$$

$$\frac{\partial L(\cdot)}{\partial \beta_1} = \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_i \right) x_i$$

$$\frac{\partial L(\cdot)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_i \right)^2 = 0$$

$$\frac{\partial L}{\partial \beta_0} = 0 \Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} y_i - n\beta_0 - \beta_1 n\bar{x} = 0$$

$$\Rightarrow n\beta_0 = n\bar{y} - \beta_1 n\bar{x}$$

$$\Rightarrow \boxed{\beta_0 = \bar{y} - \beta_1 \bar{x}}$$

$$\hat{\beta}_0^{ML} = \bar{y} - \beta_1 \bar{x}$$

$$\frac{\partial L(\cdot)}{\partial \beta_1} = 0 \Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) x_i = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_0 n \bar{x} - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\beta_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i - (\bar{y} - \beta_1 \bar{x}) n \bar{x}$$

$$\Rightarrow \beta_1 \sum_{i=1}^{n} x_i^2 - \beta_1 n \bar{x}^2 = \sum_{i=1}^{n} x_i y_i - n \bar{x} \bar{y}$$

$$\Rightarrow \beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^{n} x_i^2 - n \bar{x}^2}$$

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$\boxed{} \rightarrow y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$

LSE $\leftrightarrow$ MLE

$$\boxed{\hat{\beta}_1^{ML} = \frac{S_{xy}}{S_{xx}}} \quad ; \quad \beta_0 = \bar{y} - \beta_1 \bar{x}$$

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

$$\hat{\sigma}^{2\,ML} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\beta}_0^{ML} - \hat{\beta}_1^{ML} x_i)^2$$

$$\hat{\sigma}^{2\,LS} = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

$$H = \begin{bmatrix} \frac{\partial^2 L(\cdot)}{\partial \beta_0^2} & \frac{\partial^2 L(\cdot)}{\partial \beta_0 \partial \beta_1} \\ \frac{\partial^2 L(\cdot)}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 L(\cdot)}{\partial \beta_1^2} \end{bmatrix} \Rightarrow \mathrm{Det}(H) < 0$$

$$H = \begin{bmatrix} \frac{\partial^2 L(\cdot)}{\partial \beta_0^2} & \frac{\partial^2 L(\cdot)}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 L(\cdot)}{\partial \beta_0 \partial \sigma^2} \\ \frac{\partial^2 L(\cdot)}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 L(\cdot)}{\partial \beta_1^2} & \frac{\partial^2 L(\cdot)}{\partial \beta_1 \partial \sigma^2} \\ \frac{\partial^2 L(\cdot)}{\partial \sigma^2 \partial \beta_0} & \frac{\partial^2 L(\cdot)}{\partial \sigma^2 \partial \beta_1} & \frac{\partial^2 L(\cdot)}{\partial {\sigma^2}^2} \end{bmatrix} \begin{matrix} \beta_0 = \hat{\beta}_0^{ML} \\ \beta_1 = \hat{\beta}_1^{ML} \\ \sigma^2 = \hat{\sigma}^2_{ML} \end{matrix}$$

$$\Rightarrow \mathrm{Det}(H) < 0$$

$$l(\beta_0, \beta_1, \sigma^2; \bar{y}) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - E(y_i)\right]^2}$$

$$\frac{\partial l(\cdot)}{\partial \beta_0} = 0 \Rightarrow \beta_0 = [\ ] \rightarrow \hat{\beta}_0^{ML}$$

β_1

$$f(\hat{\beta}_0, \hat{\beta}_1) = \sum_{i=1}^{n} y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$\beta_0 + \beta_1 x_i + \epsilon_i$

i/p → y_i

$y_1, y_2, \dots, y_n$

β_0, β_1

$\hat{\beta}_0 \quad \hat{\beta}_1$

Data set

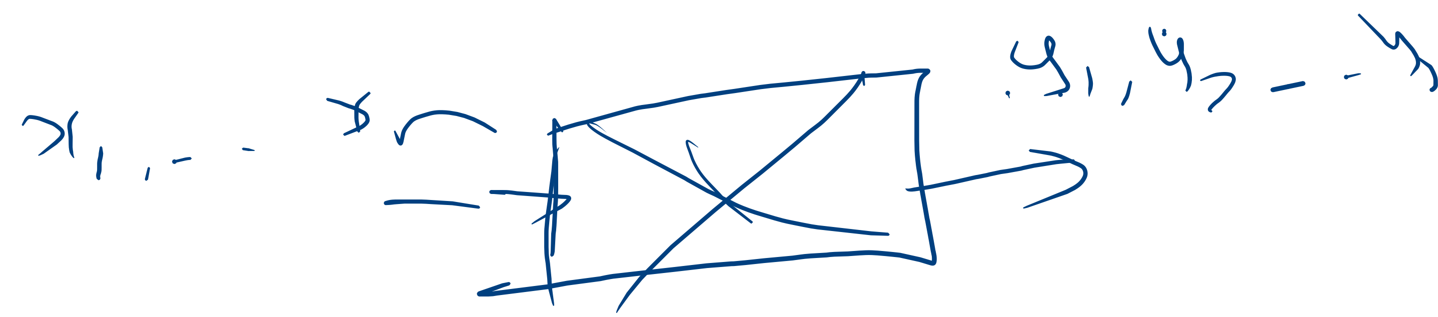
(x_1, y_1)

(x_2, y_2)

$x_n \quad y_n$

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

$$\hat{\beta}_1 = (\quad) \, \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

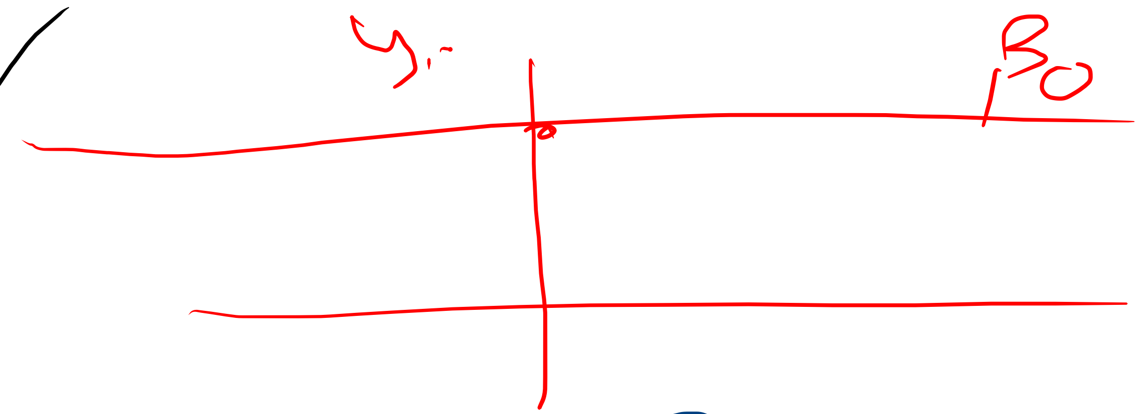


$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$x_i = 0$$

$$y_i = \beta_0 + \text{Random Error}$$

$$y_i = \beta_1 x_i + \varepsilon_i$$



$\hat{\beta}_1$

$$x_i = 0$$

$$y_i = \text{Random Error}$$

$$\frac{\partial^2 f(\cdot)}{\partial \beta_1^2} < 0$$

$y = ax + b$

$$\operatorname{cov}(x, y)$$

$$= \operatorname{cov}(x,\ ax + b)$$

$$\operatorname{cov}(\bar{y}, \hat{\beta}_1) = 0 \qquad \operatorname{cov}\left(\frac{1}{n}\sum_i y_i,\ \sum_i c_i y_i\right)$$

$$\operatorname{cov}(y_i, \hat{y}_i) \to \operatorname{cov}\left[(\beta_0 + \beta_1 x_i + \varepsilon_i),\ (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right]$$

→ Hypothesis Testing

→ Significance level

→ Confidence level.

Point Estimator:

$$\boxed{\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}}$$

Point Estimate: The value yielded by the estimator which is also a single numeric value.

Interval Estimate: The estimator yields the values of the estimate in a range or interval rather than a single numeric value.

Population (N)

5 Lakhs

$X_1, X_2, \ldots, X_N$

$$\mu = \frac{X_1 + X_2 + \cdots + X_N}{N}$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (X_i - \mu)^2$$

Sample (n)

20K

GN → 23 region

(7 regions)

1

2

23

Normal Distribution

Population (N)

"Population Parameters"

(μ, σ^2)

Population Mean

Population Variance

Sample (n)

"Sample Statistics"

$(\bar{X}, S^2)$

Sample Average

Sample Variance

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i \Rightarrow E(\bar{X}) = \mu$$

$$S^2 = \frac{1}{\underbrace{(n-1)}_{\text{"}}} \sum_{i=1}^{n} (x_i - \bar{x})^2 \Rightarrow \boxed{E(S^2) = \sigma^2}$$

*

$X_i \overset{IID}{\sim} N(\mu, \sigma^2)$ — μ: Unknown, σ^2: Known

'n' Samples.

MLE Framework.

$$f_{X_1, X_2, \ldots X_n}(x_1, x_2, \ldots, x_n; \mu, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$L(\mu, \sigma^2; \bar{x}) = \log_e l(\mu, \sigma^2; \bar{x})$$

$$L(\mu, \sigma^2; \bar{x}) = -\frac{n}{2}\log_e(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$\frac{\partial L(\cdot)}{\partial \mu} = 0 \Rightarrow \mu = (\ )$ $\hat{\mu}$

$\frac{\partial L(\cdot)}{\partial \sigma^2} = 0 \Rightarrow \sigma^2 = (\ )$ $\hat{\sigma}^2$

$$\frac{\partial L(\cdot)}{\partial \mu} = 0 \Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

MLE framework

$$\boxed{\hat{\mu}^{ML} = \bar{x}}$$

$$\frac{\partial}{\partial \sigma^2} L(\cdot) = 0 \Rightarrow \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$\hat{\mu} = \bar{x}$$

$$E\{\hat{\mu}\} = E\left\{\frac{1}{n}\sum_{i=1}^{n} x_i\right\}$$

$$E\{\hat{\mu}\} = \mu \quad ✓$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$✗ \quad E\{\hat{\sigma}^2\} = \frac{n-1}{n}\,\sigma^2$$

$$(n-1)$$

$$E\{\hat{\sigma}^2\} = \sigma^2 \quad ✓$$

Significance Level & Confidence Level

(SL) → (α) (CL) → $(1-\alpha)$

$$CL = 1 - SL$$

$\boxed{\text{Normal Distribution}}$

μ : Population Mean (Unknown)
σ^2 : Population Variance (Known)

$\hat{\mu} = \bar{X}$ (Point Estimator)

$\hat{\mu} = [\ \underset{LL}{\ }\ ,\ \underset{UL}{\ }\ ]$ (Interval Estimator)

Test Statistics

$$\boxed{z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}} \sim N(0, 1)$$

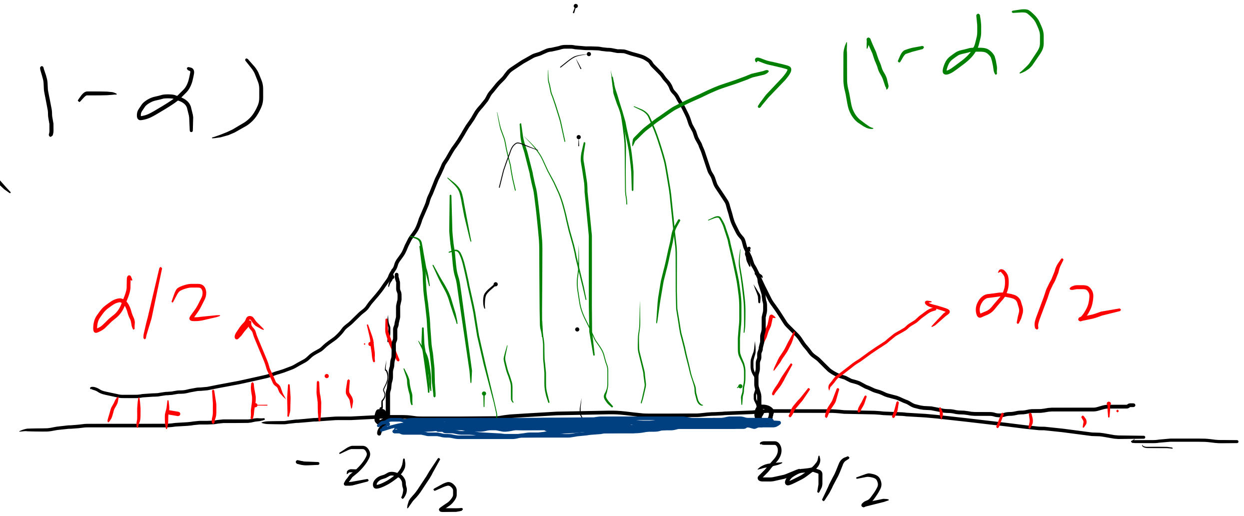
$$E(z) = \frac{1}{\sigma/\sqrt{n}} \left[E(\bar{x}) - E(\mu) \right]$$

$$E(z) = \frac{1}{\sigma/\sqrt{n}} \left[\mu - \mu \right] = 0$$

$$Var(z) = \frac{1}{\sigma^2/n} \underbrace{Var(\bar{x})}_{\sigma^2/n} - \underbrace{\frac{1}{\sigma^2/\sqrt{n}} Var(\mu)}_{0}$$

$$Var(z) = 1$$

$$\Pr\left\{ \underline{-z_{\alpha/2} < Z < z_{\alpha/2}} \right\} = (1-\alpha)$$

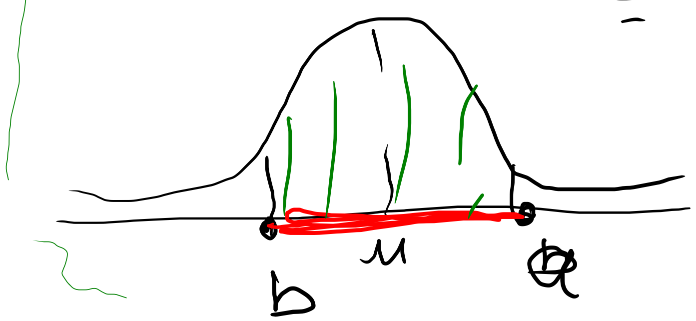


$$\Pr\left\{ -z_{\alpha/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2} \right\} = (1-\alpha)$$

$$X \sim N(\mu, \sigma^2)$$

$$\Pr\left\{ \bar{x} - z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} \right\} = (1-\alpha)$$

$$\Pr\{ b < X < a \} = \int_b^a f_X(x)\,dx$$



$$Pr\left\{ \bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right\} = (1-\alpha)$$

Interval Estimator for unknown Population parameter which is population mean is

$$\left[\underbrace{\bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}}_{\text{Lower Confidence Limit}} \; , \; \underbrace{\bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}}_{\text{Upper Confidence Limit}} \right]$$

let $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$$\hat{\mu} = [\bar{x} - E, \; \bar{x} + E]$$

Confidence level = 80%

$(1-\alpha) = 0.8$

$\Rightarrow \alpha = 0.2$

$\alpha/2 = 0.1$

$Z_{\alpha/2} = \underbrace{Z_{0.1} = 1.28}$

$$\underset{\substack{\downarrow \\ \text{Interval Estimator}}}{\mu} = \left[\bar{x} - 1.28 \times \frac{\sigma}{\sqrt{n}},\ \bar{x} + 1.28 \times \frac{\sigma}{\sqrt{n}}\right]$$

Confidence level = 90%

$(1-\alpha) = 0.9$

$\Rightarrow \alpha = 0.1$

$\alpha/2 = 0.05$

$Z_{\alpha/2} = \underbrace{Z_{0.05} = 1.645}$

$$\mu = \left[\bar{x} - 1.645 \times \frac{\sigma}{\sqrt{n}},\ \bar{x} + 1.645 \times \frac{\sigma}{\sqrt{n}}\right]$$

$\hookrightarrow$ Interval Estimator

Q) Marks of an exam of 400 students are Normally distributed with a population standard deviation of 5.6. A random sample with marks of 40 students is collected, which has a mean of 32.

Estimate the population mean with a confidence of

(a) 80%.

(b) 90%.

(c) 98%.

$N = 400$

$n = 40$.

$n = 40, \quad \bar{x} = 32 \quad \sigma = 5.6$

(a) 80% confidence $(1-\alpha) = 0.8 \Rightarrow \alpha = 0.2$

$\alpha/2 = 0.1 \quad z_{0.1} = 1.28$

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.28 \times \frac{5.6}{\sqrt{40}} = 1.13$$

Average marks of 400 students $= [30.86, 33.13]$

(b) 90% confidence

$$(1-\alpha) = 0.9 \Rightarrow \alpha = 0.1$$

$$\alpha/2 = 0.05$$

$$z_{0.05} = 1.645$$

$$E = z_{0.05} \times \frac{\sigma}{\sqrt{n}} = 1.45$$

$$= [30.55,\ 33.45]$$

(c) $$(1-\alpha) = 0.98 \Rightarrow \alpha = 0.02$$

$$z_{0.01} = 2.33$$

$$\alpha/2 = 0.01$$

$$E = z_{0.01} \times \frac{\sigma}{\sqrt{n}} = 2.33 \times \frac{5.6}{\sqrt{40}} = 2.06 = [29.94, 34.06]$$

80% confidence $[30.87, \; 33.13]$

90% confidence $[30.54, \; 33.46]$

98% confidence $[29.94, \; 34.06]$

As the confidence level increases, the interval for interval estimator increases further.

Recap for LRM:
(β_1)

$$\hat{\beta}_1 \sim N\Big(\underset{E(\hat{\beta}_1)}{\beta_1}, \underset{Var(\hat{\beta}_1)}{\frac{\sigma^2}{S_{xx}}}\Big)$$

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

$$\hat{\beta}_1 = \sum_{i=1}^{n} c_i y_i$$

$$= c_1 y_1 + c_2 y_2 + \cdots + \cdots + c_n y_n$$

$$y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$$

Test Statistics

$$z_1 = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{S_{xx}}}}$$

$$\Rightarrow E(z_1) = \frac{1}{\sqrt{\frac{\sigma^2}{S_{xx}}}}\left[E(\hat{\beta}_1) - E(\beta_1)\right]$$

$$= \frac{1}{\sqrt{\frac{\sigma^2}{S_{xx}}}}\left[\beta_1 - \beta_1\right]$$

$$= 0$$

$$z_1 \sim N(0,1)$$

$$Var(z_1) = \frac{1}{\frac{\sigma^2}{S_{xx}}} Var(\hat{\beta}_1) = \frac{1}{\frac{\sigma^2}{S_{xx}}} \cdot \sigma^2/S_{xx}$$

$$= 1$$

$(1-\alpha)\times 100$ percent Confidence Interval we have:

Test Statistics

$$\Pr\left\{-z_{\alpha/2} < z_1 < z_{\alpha/2}\right\} = (1-\alpha)$$

$$\Rightarrow \Pr\left\{-z_{\alpha/2} < \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{S_{xx}}}} < z_{\alpha/2}\right\} = (1-\alpha)$$

$$\Rightarrow \Pr\left\{\underbrace{\hat{\beta}_1 - z_{\alpha/2}\times\sqrt{\frac{\sigma^2}{S_{xx}}}} < \beta_1 < \underbrace{\hat{\beta}_1 + z_{\alpha/2}\times\sqrt{\frac{\sigma^2}{S_{xx}}}}\right\} = (1-\alpha)$$

$$\beta_1 \xrightarrow{\text{Point Estimator}} \hat{\beta}_1$$

$$\beta_1 \xrightarrow{\text{Interval Estimator}} \left[\hat{\beta}_1 - E,\ \hat{\beta}_1 + E\right]$$

Parameter
(Unknown, Deterministic)

where $E = z_{\alpha/2} \times \sqrt{\dfrac{\sigma^2}{S_{xx}}}$

Recap of LRM

$\beta_0 \longrightarrow \hat{\beta}_0 \sim N\left(\overset{E(\hat{\beta}_0)}{\beta_0}, \overset{Var(\hat{\beta}_0)}{\sigma^2\left(\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right)}\right)$

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

$$= \underbrace{\frac{(y_1+y_2+\cdots+y_n)}{n}}_{\text{Normal}} - \underset{\text{Normal}}{\hat{\beta}_1} \bar{x}$$

$$y_i \sim N(\cdot, \cdot)$$

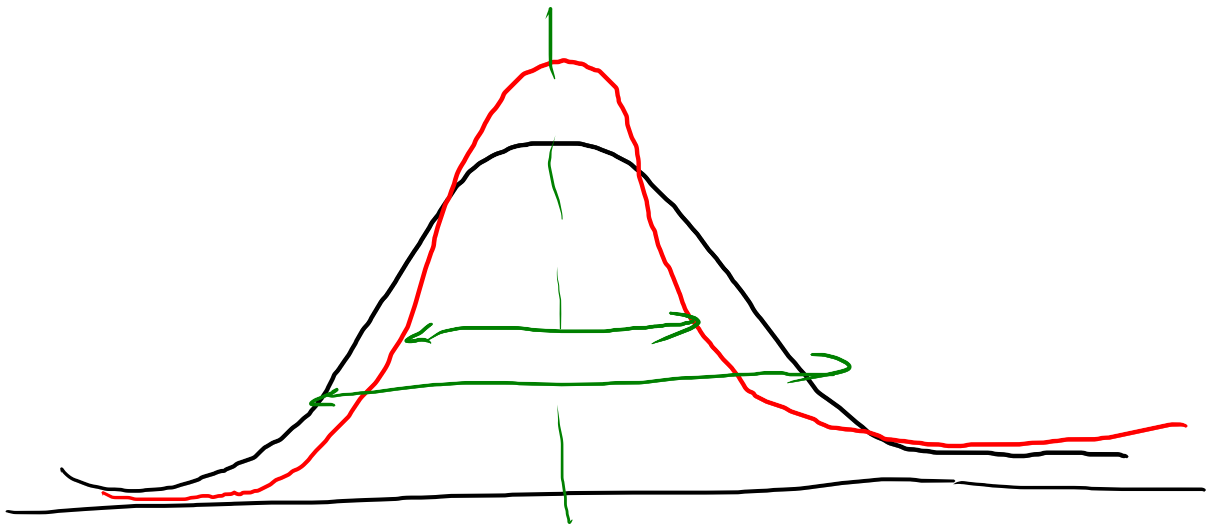
Test Statistics

$$z_0 = \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\sigma^2\left(\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right)}}$$

Interval Estimator for β_0

$$\therefore \left[\underbrace{\hat{\beta}_0 - z_{\alpha/2}\cdot\sqrt{\sigma^2\left(\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right)}}_{\text{Lower Confidence Limit}},\ \underbrace{\hat{\beta}_0 + z_{\alpha/2}\cdot\sqrt{\sigma^2\left(\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right)}}_{\text{Upper Confidence Limit}}\right]$$

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$



Variance $= \dfrac{\sigma^2}{n}$

$n = 30, 50$

$n \uparrow\uparrow$ Variance $\downarrow\downarrow$ $\dfrac{\sigma^2}{n} \downarrow\downarrow$

Estimating the Difference in Means of Two Normal Population: $(\mu_1 - \mu_2)$

=> Let $X_1, X_2, \ldots, X_n$ be a sample of size n from a Normal population having mean μ_1 and variance σ_1^2

=> Let $Y_1, Y_2, \ldots, Y_m$ be a sample of size m from a different Normal population having mean μ_2 and variance σ_2^2.

Consider that the two sample are independent of each other

Interest : To estimate $\mu_1 - \mu_2$ $\xrightarrow{\text{Point estimator}}$ $\hat{\mu}_1 - \hat{\mu}_2 = \boxed{(\bar{x} - \bar{y})}$

Sample (n)

sample statistics

Sample (m)

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

$\bar{y} \sim N\left(\mu_2, \frac{\sigma_2^2}{m}\right)$

$\bar{x} \sim N\left(\mu_1, \frac{\sigma_1^2}{n}\right)$

Test Statistics

$$Z_{test} = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

$Z_{test} \sim N(0, 1)$ $E(Z_{test}) = 0$

$Var(Z_{test}) = 1$

$$\Pr\left\{-z_{\alpha/2} < z_{test} < z_{\alpha/2}\right\} = (1-\alpha)$$

$$\Pr\left\{-z_{\alpha/2} < \frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} < z_{\alpha/2}\right\} = (1-\alpha)$$

$$\Pr\left\{(\bar{x}-\bar{y}) - z_{\alpha/2}\times\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}} < (\mu_1-\mu_2) < (\bar{x}-\bar{y}) + z_{\alpha/2}\times\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\right\} = (1-\alpha)$$

Interval Estimator: $(\bar{x}-\bar{y}) \pm E$ where $E = z_{\alpha/2}\times\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}$

Scenario where population (N) mean μ is unknown & population variance σ^2 is also unknown.

$$X_i \sim N(\mu, \sigma^2)$$

$\mu \rightarrow E(X_i)$, $\sigma^2 \rightarrow Var(X_i)$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Earlier when σ^2 was known: $[\bar{x} \pm E]$ where $E = z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$

σ^2 is unknown $\Rightarrow$ Test statistics $\longrightarrow$ $\dfrac{\bar{x} - \mu}{\sqrt{\hat{\sigma}^2/n}} = \dfrac{\bar{x} - \mu}{\sqrt{S^2/n}}$

$$S^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$E\{S^2\} = \sigma^2 \Rightarrow$ So, $S^2 = \hat{\sigma}^2$

Earlier: Population mean μ : Unknown, Population Variance σ^2 : Known.

Point Estimator for μ : $\hat{\mu} = \bar{x}$

Test Statistics $= \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

Interval Estimator for μ : $\hat{\mu} = [\bar{x} - E, \bar{x} + E]$

where $E = Z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$

Z-score.

1 parameter.

Now: Population mean μ : Unknown, Population Variance σ^2 : Unknown

Test Statistics $= \dfrac{\bar{x} - \mu}{\hat{\sigma}/\sqrt{n}} = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

$S^2 = \dfrac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$E(S^2) = \sigma^2$

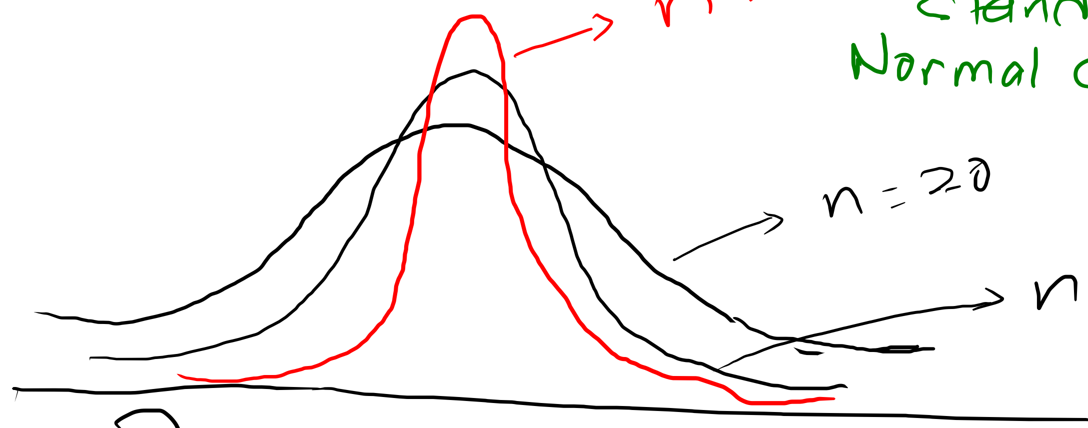
$S^2 = \hat{\sigma}^2$

$\Rightarrow \hat{\sigma} = S$

$S^2 =$

For Degree of Freedom $\to \infty$, the t-distribution converges to Standard Normal distribution

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$



Test Statistics

$$\Pr\left\{ \underbrace{-t_{\alpha/2,(n-1)}} < t < \underbrace{t_{\alpha/2,(n-1)}} \right\} = (1-\alpha)$$

$$\Pr\left\{ -t_{\alpha/2,(n-1)} < \frac{\bar{x} - \mu}{s/\sqrt{n}} < t_{\alpha/2,(n-1)} \right\} = 1-\alpha$$

$$\Pr\left\{ \bar{x} - \frac{s}{\sqrt{n}} \times t_{\alpha/2,(n-1)} < \mu < \bar{x} + \frac{s}{\sqrt{n}} \times t_{\alpha/2,(n-1)} \right\} = (1-\alpha)$$

When σ^2: Unknown, then interval estimator for the population mean μ : $[\bar{x} - E, \ \bar{x} + E]$

where $E = t_{\alpha/2,(n-1)} \times \dfrac{s}{\sqrt{n}}$

Hypothesis Testing:

Hypothesis: A claim/statement made which is subject to verification/testing.

Test Result \ Truth/Actual Scenario	Covid Negative	Covid Positive
Covid Negative	✓	False Negative ✗
Covid Positive	False Positive ✗	✓

Error

Error.

Null Hypothesis

→ Hypothesis which is tested for possible rejection under the assumption that it is TRUE

→ Represented by H_0

Alternate Hypothesis

→ Hypothesis which also contradicts or is complementary to the Null Hypothesis

→ Represented by H_1

~~Truth~~/Actual Scenario

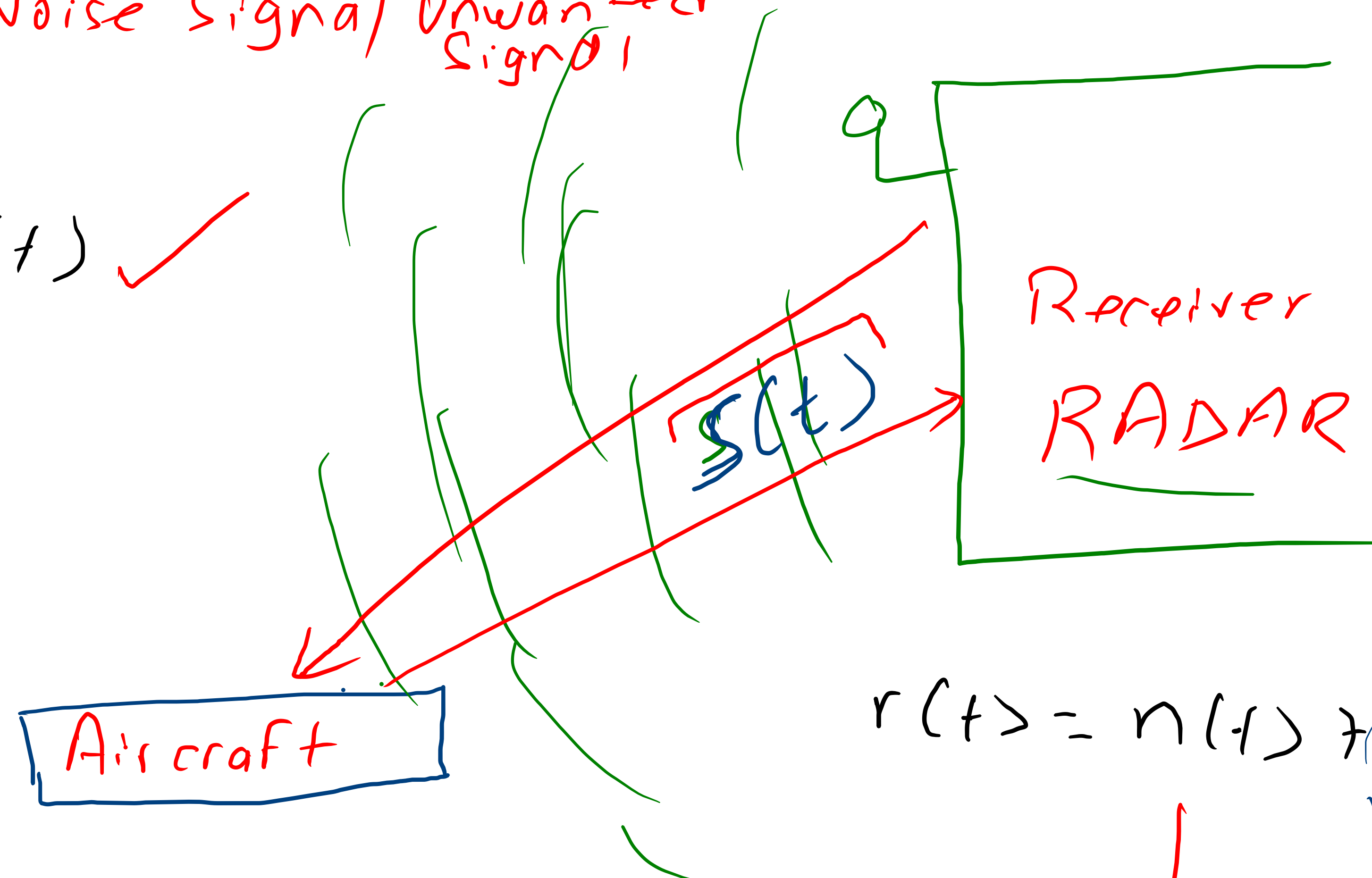
Decision	H_0: True	H_0: False
Reject H_0	Type-1 Error Error ✗	✓
Accept H_0	✓	Type-2 Error Error ✗

False Alarm

Misdetection.

$H_0 : \quad r(t) = n(t)$ → Noise Signal / Unwanted Signal

$H_1 : \quad r(t) = \boxed{s(t)} + n(t)$ ✓



$r(t) = n(t) + s$

↓

H_1 : True

$\theta \sim N(\cdot\,,\cdot)$

$H_0 : \theta \geq -10$

$H_1 : \theta < -10$

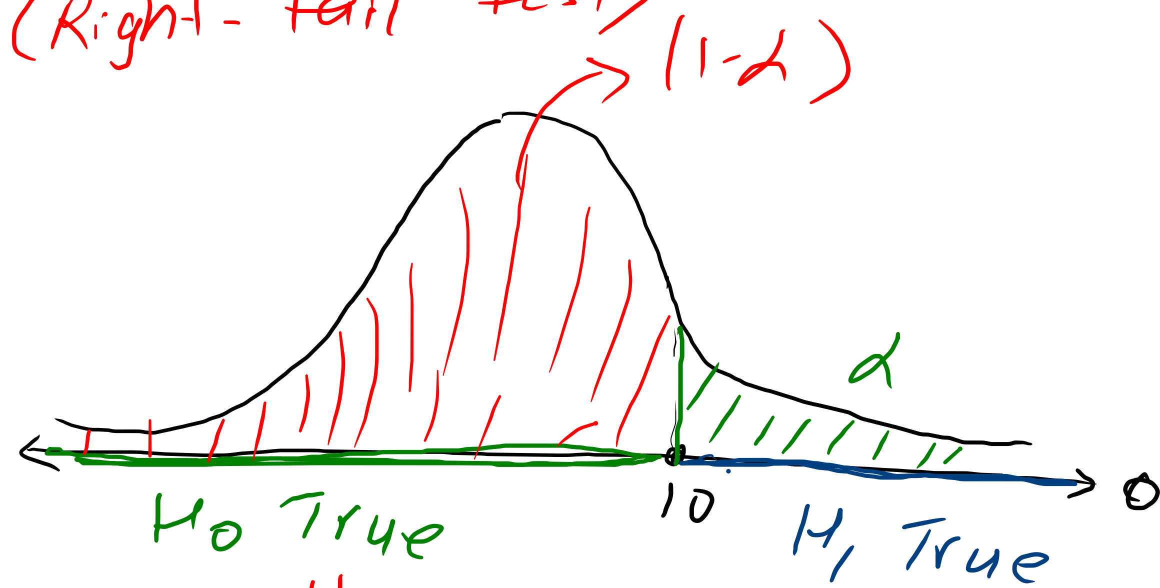
One-Sided Test (Left-Tail test)

IF the value of θ lies in this region H_0 is considered to be true.

One-Sided test (Right-tail test)

$H_0: \quad \theta \leq 10$

$H_1: \quad \theta > 10$



⇓

θ value in this region, H_0 is true.

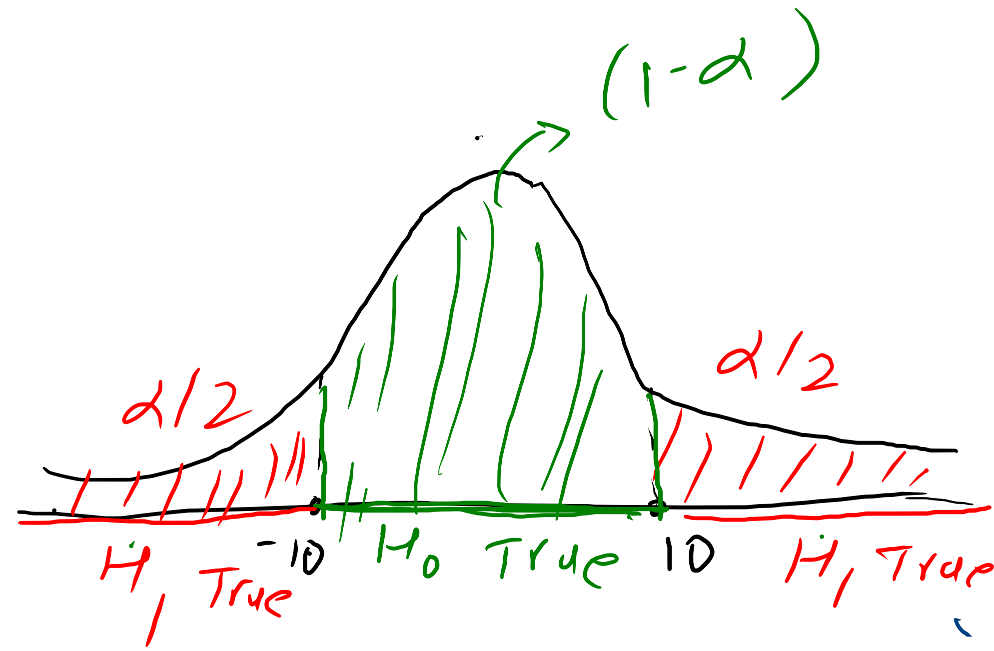
$(1-\alpha)$: Level of Confidence

α: Level of Significance

Two-sided test

$H_0: |\theta| \leq 10$

$H_1: |\theta| > 10$



Q. Manufacturer makes 1200 boxes and claims that the average weight of a box is 1.84 KG

Customer randomly chooses 64 boxes and find sample average weight as 1.88 KG. Considering that population standard deviation is 0.3 KG, use 95% confidence level and test for hypothesis that true average weight of shipment is 1.84 KG.

95% confidence level

$\Downarrow$

$(1-\alpha) = 0.95$

$N = 1200$ (population)

$\sigma = 0.3$ KG

$n = 64$ (sample)

$\bar{x} = 1.88$ KG.

$\Rightarrow \boxed{\begin{array}{l}\alpha = 0.05 \\ \alpha/2 = 0.025\end{array}}$

$\boxed{Z_{\alpha/2} = Z_{0.025}}$ ✓

$H_0: \mu = 1.84$

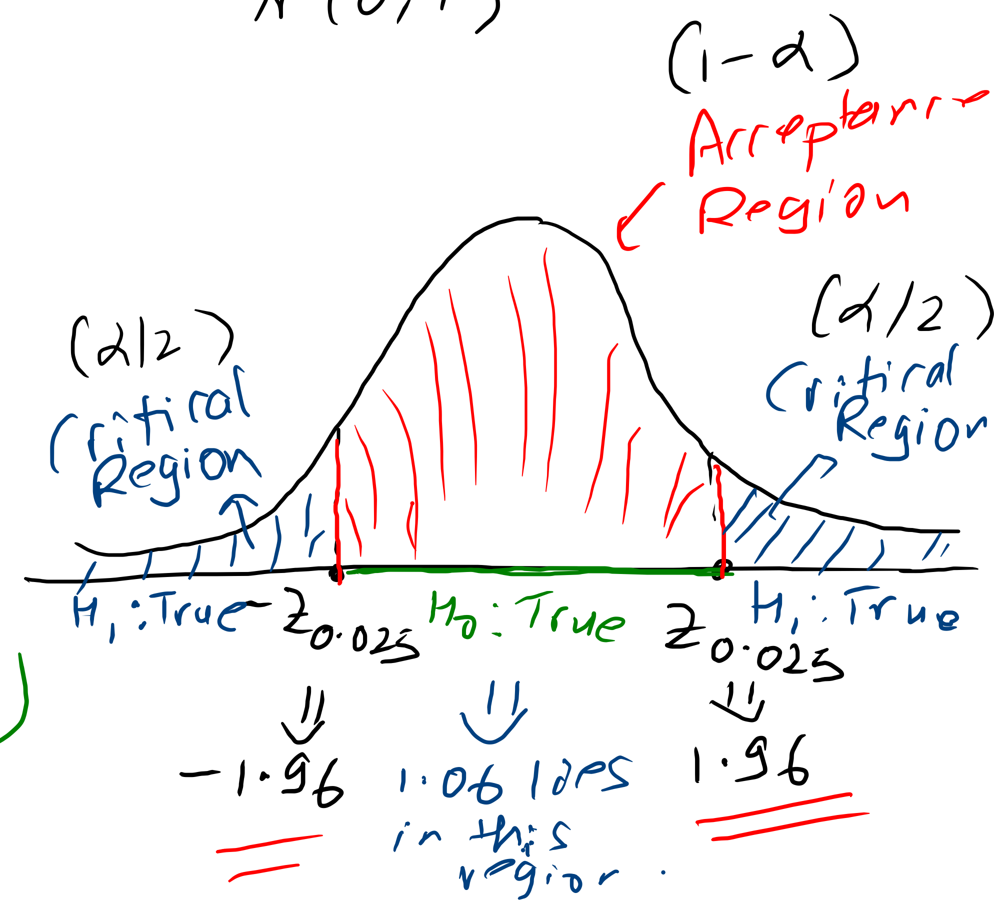
$H_1: \mu \neq 1.84$

Finding Test statistics is the first important step.

(2) Test statistics $= \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$\Rightarrow \quad z = \dfrac{1.88 - 1.84}{0.3/\sqrt{64}}$

$\Rightarrow \quad z = 1.06$



Decision Rule: Accept the Null Hypothesis H_0. (95% confidence level)

Interval Estimator $= \left[\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \, , \, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right]$

$(1-\alpha) \to CL$

LCL

UCL

$\Downarrow$

Test statistics

In LRM:

(β_1)

Test Statictics $= \dfrac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\sigma^2}{S_{xx}}}}$ σ: Known

$\downarrow$

$N(0,1)$

Test Statistics $= \dfrac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}}$ σ: Unknown

$\downarrow$

t . $(n-2)$ Degree of freedom

$$\hat{\sigma}^2 = \frac{SS_{res}}{\underbrace{n-2}}$$

Hypothesis Testing in $\boxed{\text{LRM}}$:

for Parameter β_1:

$$H_0: \beta_1 = \beta_1^{\text{constt.}}$$

$\beta_1^{\text{constt.}}$: Known Specified value.

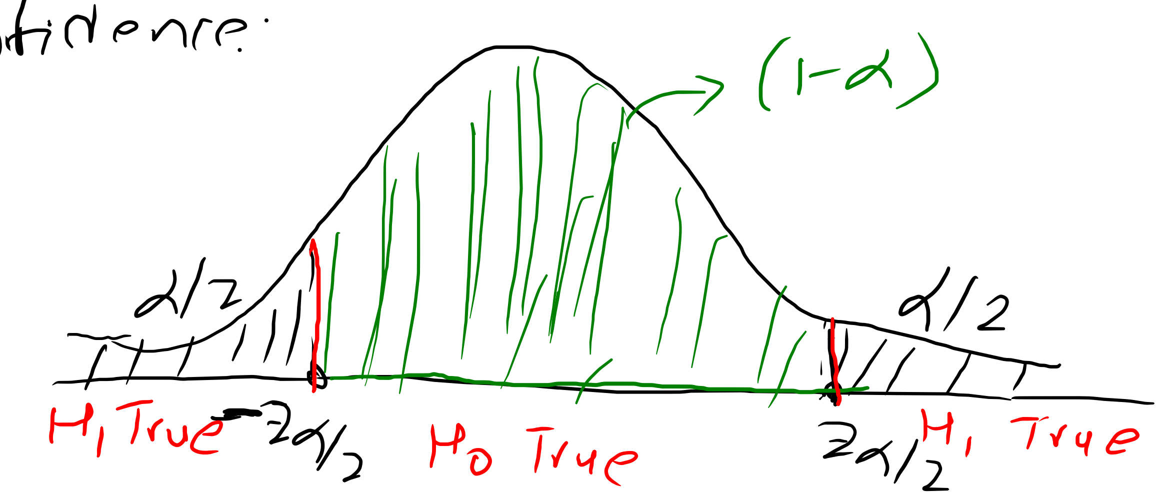
σ^2 : Known

Test Statistics $\boxed{Z_1 = \dfrac{\hat{\beta}_1 - \beta_1^{\text{constt.}}}{\sigma / \sqrt{S_{xx}}}}$

$$E(Z_1) = \frac{1}{\sigma/\sqrt{S_{xx}}}\left[E(\hat{\beta}_1) - E\left(\beta_1^{\text{constt}}\right)\right] = \frac{1}{\sigma/\sqrt{S_{xx}}}\left[\beta_1 - \beta_1\right]$$

$$\Rightarrow \boxed{E(Z_1) = 0}$$

$(1-\alpha)\times 100$ percent level of confidence:



$H_1: \quad \beta_1 \neq \beta_1^{const.}$

Decision Rule: Reject H_0 when $z_1 > z_{\alpha/2}$ or $z_1 < -z_{\alpha/2}$

Reject H_0 when $|z_1| > z_{\alpha/2}$

Other Possible Alternate Hypothesis (One-Sided Hypothesis testing)

$H_1: \quad \beta_1 > \beta_1^{const.}$

Decision Rule: Reject H_0 when $z_1 > z_\alpha$

$H_1: \quad \beta_1 < \beta_1^{const.}$

Decision Rule: Reject H_0 when $z_1 < -z_\alpha$

Hypothesis Testing for β_1 when σ^2 is UNKNOWN:

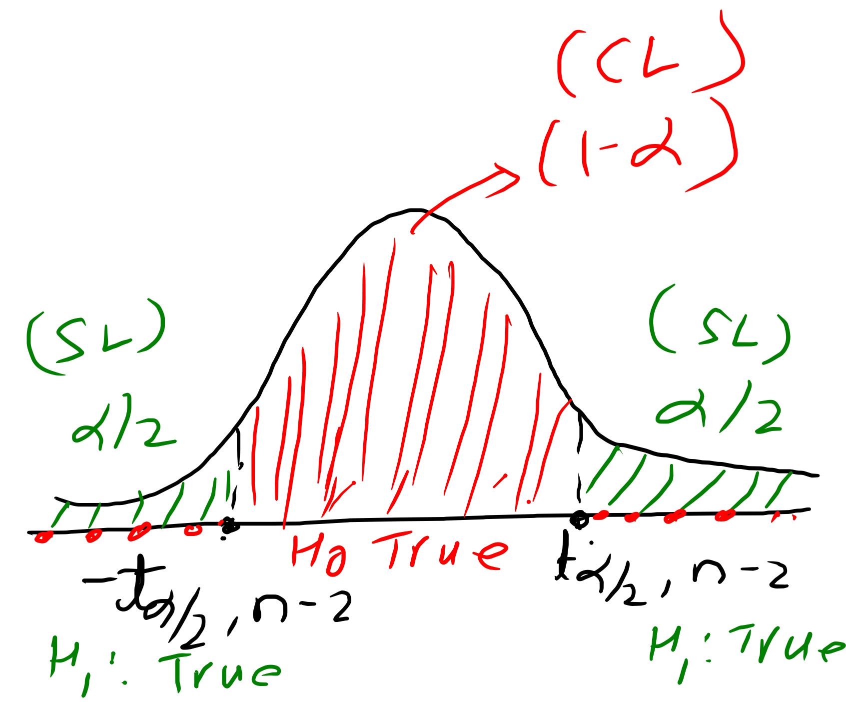
$H_0: \quad \beta_1 = \beta_1^{const.}$

β_1^{const} : Known, Constant, Specified value.

Test Statistics $t_1 = \dfrac{\hat{\beta}_1 - \beta_1^{const.}}{\hat{\sigma}/\sqrt{S_{xx}}} \sim t(n-2)$

$H_1: \quad \beta_1 \neq \beta_1^{const.}$

Decision Rule: Reject H_0 when $|t_1| > t_{\alpha/2,(n-2)}$



Hypothesis Testing for the parameter β_0:

σ^2

σ^2 known

$H_0: \beta_0 = \beta_0^{(const)}$

Test Statistics: $z_0 = \dfrac{\hat{\beta}_0 - \beta_0^{(const)}}{\sqrt{\sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} \sim N(0,1)$

$H_1: \beta_0 \neq \beta_0^{(const)}$

Decision Rule: Reject Null Hypothesis H_0 when $|z_0| > z_{\alpha/2}$

σ^2 Unknown

$H_0: \beta_0 = \beta_0^{(const)}$

Test Statistics: $t_0 = \dfrac{\hat{\beta}_0 - \beta_0^{(const)}}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} \sim t(n-2)$

$H_1: \beta_0 \neq \beta_0^{(const)}$

Decision Rule: Reject H_0 when $|t_0| > t_{\alpha/2,(n-2)}$

Test of Hypothesis for σ^2: (LRM)

$H_0: \sigma^2 = \sigma_0^2$

σ_0^2: known specified constant value.

Test Statistics $C_0 = \dfrac{SS_{res}}{\sigma_0^2} \sim \chi^2(n-2)$

$H_1: \sigma^2 \neq \sigma_0^2$

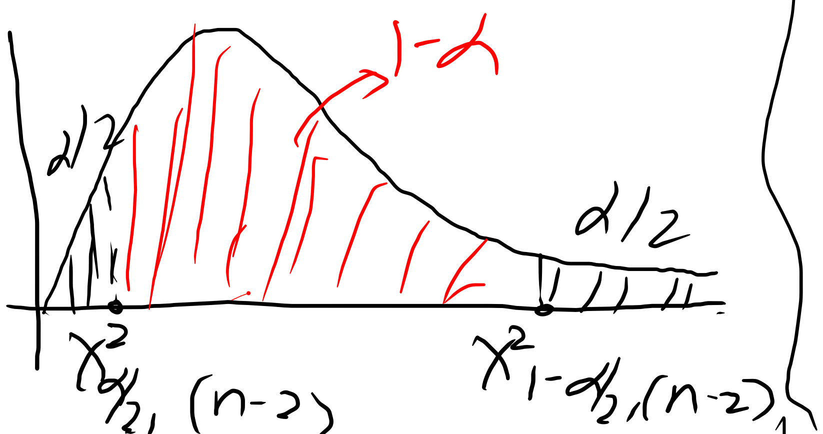
Decision Rule:

Reject H_0 if

$C_0 < \chi^2_{\alpha/2,(n-2)}$

OR

$C_0 > \chi^2_{1-\alpha/2,(n-2)}$



$$SS_{res} = \sum_{i=1}^{n} [y_i - \hat{y}_i]^2$$

$$= \left[(\beta_0 + \beta_1 x_i + \varepsilon_i) - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right]^2$$

$$SS_{res} = \left[\underline{(\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)x_i} + \varepsilon_i\right]^2$$

$$\boxed{\frac{SS_{res}}{\sigma^2}} = [\alpha_i + \tilde{\varepsilon}_i]^2 \sim \boxed{\chi^2(n-2)}$$

where $\alpha_i = \dfrac{(\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)x_i}{\sigma}$

$\tilde{\varepsilon}_i = \varepsilon_i / \sigma$

MLE $\longrightarrow$ $f_X(x) = \frac{e^{-\lambda} \lambda^{x}}{x!}$ $\longrightarrow$ Mean value λ

Poisson Distribution

MLE of λ ??

$\bar{x} = [x_1, x_2, \dots, x_n]$

$x_i \; ; i = 1, 2, \dots, n$ $\longrightarrow$ independent

Joint PMF $\quad$ marginal PMF

$$p_{\bar{X}}(x_1, x_2, \dots, x_n | \lambda) = p_{X_1}(x_1)\, p_{X_2}(x_2) \cdots p_{X_n}(x_n)$$

$$= \frac{e^{-\lambda} \lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda} \lambda^{x_2}}{x_2!} \cdots \frac{e^{-\lambda} \lambda^{x_n}}{x_n!}$$

$$= \left(e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^{n} x_i} \right) \Big/ \prod_{i=1}^{n} x_i!$$

Joint PMF

$$\underbrace{p_{\bar{X}}(x_1, x_2, \ldots, x_n \mid \lambda)}_{} = \frac{e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

$\prod_{i=1}^{n} x_i! \to C$

likelihood function $\ell(\lambda \mid \underbrace{x_1, x_2, \ldots, x_n}_{\bar{x}})$

$\bar{x} : [x_1, x_2 \ldots x_n]$

$$L(\lambda \mid \bar{x}) = \log_e \ell(\lambda \mid \bar{x})$$

$$L(\lambda \mid \bar{x}) = -n\lambda + \sum_{i=1}^{n} x_i \log \lambda - \log C$$

$$\frac{\partial L(\cdot)}{\partial \lambda} = 0 \Rightarrow \lambda = (\quad)$$

$\hookrightarrow \hat{\lambda}$

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

$$\Rightarrow \quad \lambda = \bar{x}$$

$$\boxed{\hat{\lambda}_{MLE} = \bar{x}}$$

Maximum Likelihood Estimate

RV

$X \rightarrow x$

$$\hat{\lambda}_{ML\ Estimator} = \bar{X}$$

$\beta_1^{(0)}$. β_1 : LRM

Test Statistics $= \dfrac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}}$ $\beta_1^{(0)}$

$H_0 : \beta_1 = \beta_1^{(0)}$

LRM

$\hat{\sigma}^2 = \dfrac{SS_{res}}{n-2}$

$$\frac{\hat{\beta}_1 - E(\hat{\beta}_1)}{Var(\hat{\beta}_1)} = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}}$$

$$= \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}}$$

$$\hat{\sigma}^2 = \frac{SS_{res}}{n-2}$$

Test Statistics

$$z = \frac{X - E(X)}{\sqrt{Var(X)}}$$

Population & Sample

(N) (n)

Sample variance

$$S^2 = \hat{\sigma}^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\Rightarrow \hat{\sigma} = S$$

Test Statistics

$$z = \frac{\bar{x} - E(\bar{x})}{\sqrt{Var(\bar{x})}}$$

$$= \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \longrightarrow z = \frac{\bar{x} - \mu}{\hat{\sigma}/\sqrt{n}}$$

$$S \leftarrow \hat{\sigma}$$

Information Theory:

TH ?__ E, Y

C ?_ T A, U, O

CO ?__ ?_ Reduced Uncertainty

OL

AL

AT

COA ?__

L, T

Uncertainty

$\updownarrow$

Information

$\updownarrow$

Entropy

⇒ Uncertainty can vary based on your Knowledge / belief.

→ All elements in the guess-list may not be equally likely.

ENTROPY :

$$H(x) = E(I(x))$$

$$H(x) = \sum_{i=1}^{n} P(x_i) \log_2 \frac{1}{P(x_i)}$$

$\mathcal{X}$ = Set containing all the possible answers

$|\mathcal{X}| = K$ = Number of Possible answer

(likelihood)

$x_i \in \mathcal{X}$: $P(x_i) \Rightarrow$ Probability that the answer x_i is correct

$x_1, x_2, \dots, x_K$; $1 \le i \le K$

$P = [P(x_1) \; P(x_2) \; \dots \; P(x_K)] \Rightarrow$ This is a representation of all the probabilities

$$H(\mathcal{X}) = \sum_{i=1}^{K} P(x_i) \cdot \underbrace{\log_2 \frac{1}{P(x_i)}}_{\rightarrow I(x_i)}$$

$$I(x_i) = \log_2 \frac{1}{P(x_i)}$$

E (Event) : Snowfall in Ahmedabad in Month of April

$P(E) \rightarrow 0$

$I(E) \rightarrow \infty$ (Very high information Value)

?? Information $\propto \frac{1}{\text{Uncertainty}}$??

Machine - 1

(A, B, C, D)

Machine - 2

CCDBAA

BBCCADD

$P(A) = 0.25$

$P(B) = P(C) = P(D) = 0.25$

$\Downarrow$

Equally likely

$P(A) = 0.5$

$P(B) = 0.125 = P(C)$

$P(D) = 0.25$

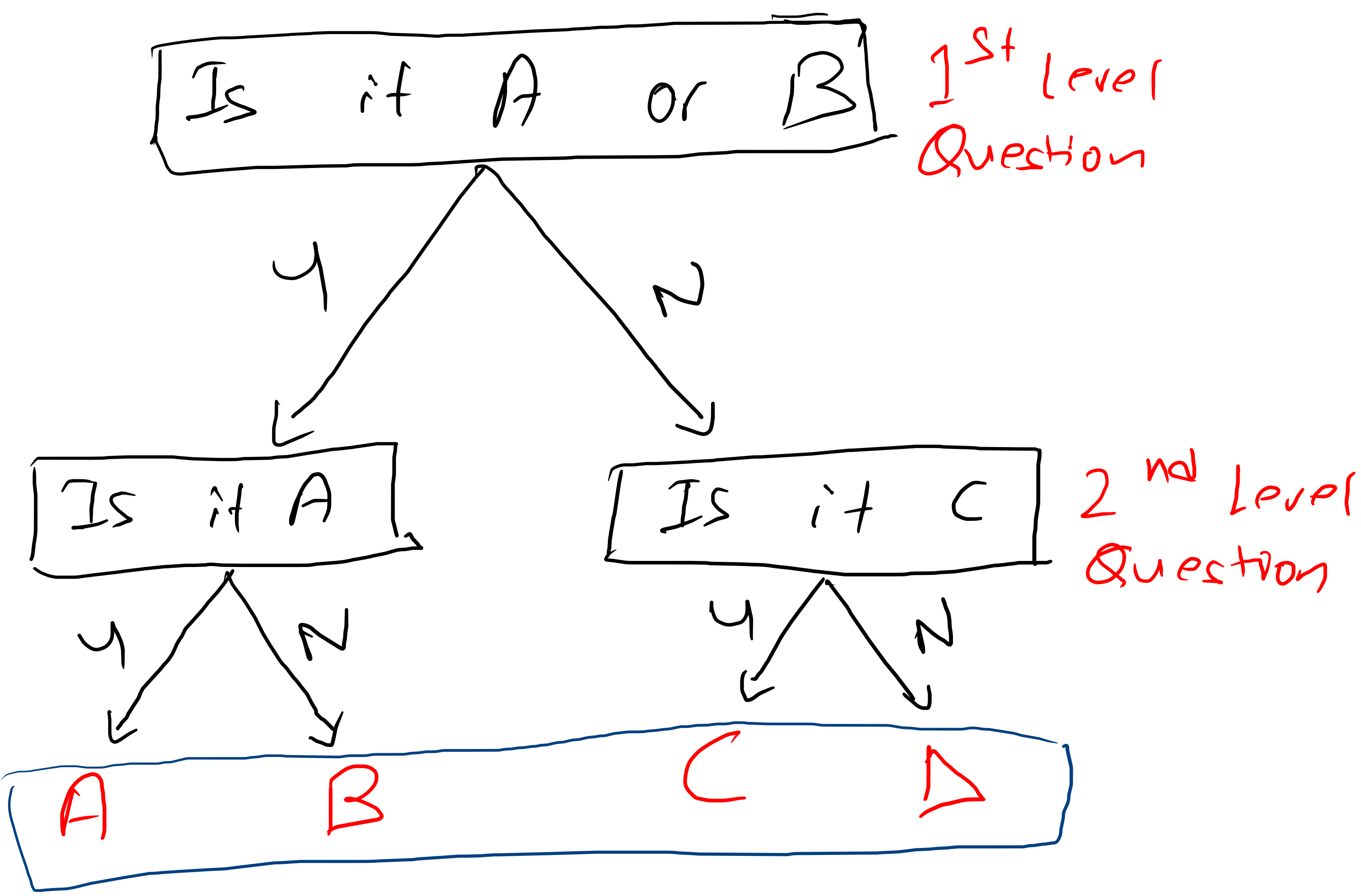
$$H(\text{Machine-1}) = 4\left(\frac{1}{4}\log_2\left(\frac{1}{1/4}\right)\right)$$

$= 2$ Ques / Symbol

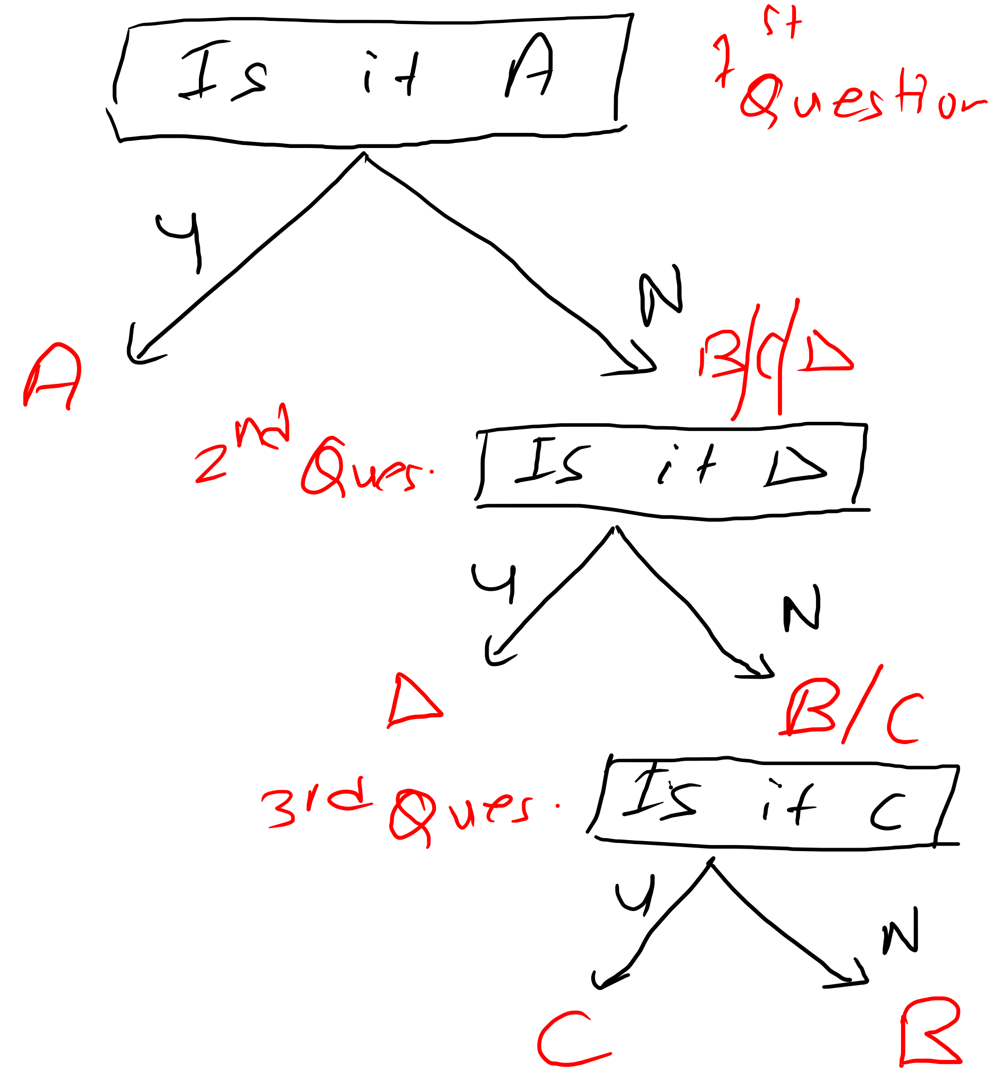
$$H(\text{Machine-2}) = \frac{1}{2}\log_2\left(\frac{1}{1/2}\right) + \left[\frac{1}{8}\log_2\left(\frac{1}{1/8}\right)\right] \times 2 + \frac{1}{4}\log_2\left(\frac{1}{1/4}\right)$$

$= 1.75$ Ques / Symbol

Machine- 1



Machine - 2



If we need to guess 100 alphabets from both the Machines:

Machine - 1

⇒ 200 questions

Machine - 2

⇒ 175 questions

E_1 : CSK/MI winning the IPL

$$P(E_1) = \theta$$

$$I(E_1) = \log_2 \frac{1}{P(E_1)}$$

$$\theta \uparrow\uparrow \longrightarrow P(E_1) \uparrow\uparrow \longrightarrow I(E_1) \downarrow\downarrow$$

E_2 : RCB winning the IPL

$$P(E_2) = \delta$$

$$\delta \downarrow\downarrow \longrightarrow P(E_2) \downarrow\downarrow \longrightarrow I(E_2) \uparrow\uparrow$$

⇒ Events which have Deterministic outcomes contain NO information.

⇒ When a low-probability event occurs, the events carries more information ("Surprisal") than when a high-probability event occurs.

Source coding : Data Compression

Shannon Fano Encoding Scheme.

⇒ Divide symbols into 2 subsets with equal probability at each stage.

→ Higher the probability of a symbol, shorter is the codeword length of the symbol.

⇒ Arrange the symbol in decreasing order of probability

⇒ Among the 2 sub set which we create, the 1st subset is assigned BIT '0' & BIT '1' to 2nd subset.

Q

$S = \{S_1, S_2, S_3, S_4\}$ $\qquad H = 1.75$ bits/symbol

$P = \{1/2 \quad 1/4 \quad 1/8 \quad 1/8\}$ $\qquad H = \sum_i P_i \log_2 \frac{1}{P_i}$

Symbols	Prob.	1st L	2nd L	3rd L	Codeword	length (l_i)
S_1	1/2	0			0	1
S_2	1/4	1	0		1 0	2
S_3	1/8	1	1	0	1 1 0	3
S_4	1/8	1	1	1	1 1 1	3

Average Codeword length

$$\bar{L} = \sum_i P_i\, l_i$$

$$\bar{L} = \left(\frac{1}{2} \times 1\right) + \left(\frac{1}{4} \times 2\right) + \left(\frac{1}{8} \times 3\right) + \left(\frac{1}{8} \times 3\right)$$

$\bar{L} = 1.75$ bits / symbol ✓

Efficiency $\quad \eta = \dfrac{H}{\bar{L}\,\underbrace{\log_2 r}_{=1 \,\to\, 1}} = 1$

$r = 2$: Binary Encoding Scheme.

Redundancy $\quad R_e = 1 - \eta = 0$

code-word

l_1	s_1	0
l_2	s_2	10
l_3	s_3	110
l_4	s_4	111

$\Rightarrow$ Prefix-free code

Kraft Inequality

$$\sum_{i=1}^{n} 2^{-l_i} \leq 1$$

Codes will be Prefix free

$$1 \leq i \leq n$$

l_i : codeword length of i^{th} symbol.

$$\mathcal{X} = \{ \overset{0}{x_1} \quad \overset{01}{x_2} \quad \overset{1}{x_3} \quad \overset{10}{x_4} \}$$

$$01 \to x_2$$

$$01 \to x_1, x_3$$

Sense of confusion

Symbol/Message $X = \{x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7\}$

Probability $P = \{0.4 \quad 0.2 \quad 0.12 \quad 0.08 \quad 0.08 \quad 0.08 \quad 0.04\}$

Entropy = 2.4205 bits/Symbol

Two scenario of subset classification [No equal probability subset is Possible]

Symbols	(P_i) Prob.	1st L	2nd L	3rd L	4th L	Codeword	length (l_i)
x_1	0.4	0				0	1
x_2	0.2	1	0 (0.32)	0		1 0 0	3
x_3	0.12	1	0	1		1 0 1	3
x_4	0.08	1	1 (0.28)	0	0	1 1 0 0	4
x_5	0.08	1	1	0	1	1 1 0 1	4
x_6	0.08	1	1	1	0	1 1 1 0	4
x_7	0.04	1	1	1	1	1 1 1 1	4

Average Codeword Length

$$\bar{L} = \sum_{i=1}^{7} P_i l_i = 2.48 \text{ bits/symbols}$$

$$\eta = \frac{H}{\bar{L}} = \frac{2.4205}{2.48} = 0.9760$$

$$Re = 1-\eta = 0.0240 \ldots$$

Symbols	P_i	1^{st} L	2^{nd} L	3^{rd} L	4^{th} L	Codeword	l_i
x_1	0.4	0	0 ×			0 0	2
x_2	0.2	0	1 ×			0 1	2
x_3	0.12	1	0	0 ×		1 0 0	3
x_4	0.08	1	0	1 ×		1 0 1	3
x_5	0.08	1	1	0 ×		1 1 0	3
x_6	0.08	1	1	1	0	1 1 1 0	4
x_7	0.04	1	1	1	1	1 1 1 1	4

$\bar{L} = 2.52$ bits/Symbol ✓

$$\eta = \frac{H}{\bar{L}} = 0.9605$$

$$R_e = 0.0395$$

Huffman Coding Algorithm:

(1) ⇒ Arrange given symbols/messages in decreasing order of Probabilities.

(2) ⇒ Group the least probable r (here r = 2, binary encoding scheme) messages and assign them with BIT 0 & BIT 1.

(3) ⇒ Add the probabilities of grouped messages and place them as high as possible. Rearrange them in decreasing order again. This process is called Reduction

(4) ⇒ Repeat Step (2) and (3) till only r (here r = 2) or less than r probabilities/Symbols remain.

(5) Obtain the codeword for each symbol by tracing the probabilities / bit sequences in reverse direction, i.e, reverse traversal.

$$S = \{S_1, S_2, S_3, S_4, S_5\}$$

$$P = \{0.4, 0.2, 0.2, 0.1, 0.1\}$$

$$H(S) = \sum_{i=1}^{5} P_i \log_2 \left(\frac{1}{P_i}\right) = 2.1216 \text{ bits/symbol.}$$

Symbols	Prob.				Codeword	l_i
S_1	0.4	0.4	0.4	0.6 (0)	$S_1 \rightarrow 00$	(2)
S_2	0.2	0.2	0.4 (0)	0.4 (1)	$S_2 \rightarrow 10$	(2)
S_3	0.2	0.2 (0)	0.2 (1)		$S_3 \rightarrow 11$	(2)
S_4	0.1 (0)	0.2 (1)			$S_4 \rightarrow 010$	(3)
S_5	0.1 (1)				$S_5 \rightarrow 011$	(3)

Average Codeword Length

$$\bar{L} = \sum_{i=1}^{5} P_i l_i = 2.2 \text{ bits/Symbol.}$$

$$\eta = \frac{H(f)}{\bar{L}\underbrace{\log_2 r}_{1}} = \frac{2.1216}{2.2} = 96.4\,\%$$

$$S = \{S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8\}$$

$$P = \{1/4, 1/4, 1/8, 1/8, 1/16, 1/16, 1/16, 1/16\}$$

Use shannon Fano source coding scheme & obtain the codeword for each symbol.

$$H(S) = \sum_{i=1}^{8} P_i \log_2\left(\frac{1}{P_i}\right) = 2.75 \text{ bits/Symbol}$$

Symbols	Prob.	1st L	2nd L	3rd L	4th L	Codeword	l_i
S_1	1/4	0	0			$S_1 \to 00$	(2)
S_2	1/4	0	1			$S_2 \to 01$	(2)
S_3	1/8	1	0	0		$S_3 \to 100$	(3)
S_4	1/8	1	0	1		$S_4 \to 101$	(3)
S_5	1/16	1	1	0	0	$S_5 \to 1100$	(4)
S_6	1/16	1	1	0	1	$S_6 \to 1101$	(4)
S_7	1/16	1	1	1	0	$S_7 \to 1110$	(4)
S_8	1/16	1	1	1	1	$S_8 \to 1111$	(4)

Average Codeword Length $\bar{L} = \sum_{i=1}^{8} P_i l_i = 2.75$ bits/symbol

$$\eta = \frac{H(S)}{\bar{L}} = 1.$$

$$Cov(y_i, \hat{\beta}_1) = E\{[y_i - E(y_i)][\hat{\beta}_1 - \underbrace{\cancel{E(\hat{\beta}_1)}}_{\beta_1}]\}$$

$$= E\{\varepsilon_i \underbrace{(\hat{\beta}_1 - \beta_1)}\} \qquad \left[\hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} c_i \varepsilon_i\right]$$

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

$$Cov(y_i, \hat{\beta}_1) = E\left\{\varepsilon_i \cdot \sum_{j=1}^{n} c_j \varepsilon_j\right\}$$

$j = i$:–

$$= \sum_{i=1}^{n} c_i \underbrace{E\{\varepsilon_i^2\}}_{\sigma^2} = \sigma^2 \underbrace{\sum_{i=1}^{n} c_i}_{0}$$

$$\underbrace{Cov(y_1, \hat{\beta}_1)}_{0} = E\{\boxed{\varepsilon_1 \cdot (c_1 \varepsilon_1} + c_2 \varepsilon_2 + \dots + c_n \varepsilon_n)\}$$

$$= c_1 E(\varepsilon_1^2)$$

$$= c_1 \sigma^2$$

$$\operatorname{Cov}(y_i, \hat\beta_1) = \operatorname{Cov}\left(y_i, \sum_{i=1}^{n} c_i y_i\right)$$

$$= \operatorname{Cov}(y_i, c_1 y_1 + c_2 y_2 + \dots + c_i y_i + \dots + c_n y_n)$$

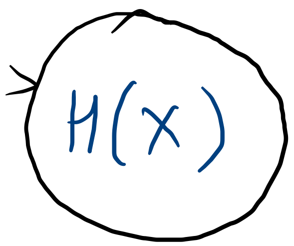
$$= \operatorname{Cov}(y_i, c_i y_i)$$

$$= c_i \underbrace{\operatorname{Cov}(y_i, y_i)}_{\operatorname{Var}(y_i)}$$

$$= \sigma^2 c_i$$

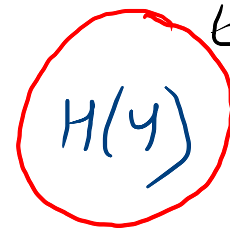
X

x_1 (9am) x_2 (10am) - - - (5pm)



$I(X;Y) = 0$

Y

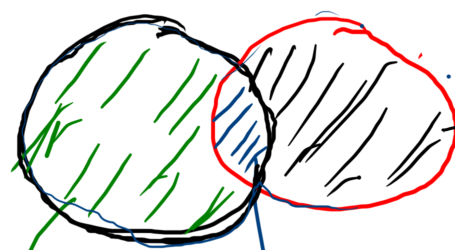


X, Y: Statistically Independent.

$H(X|Y) = H(X)$

$H(Y|X) = H(Y)$

X   Y



X, Y: Statistically Dependent

$I(X;Y) = H(X) - H(X|Y)$

$= H(Y) - H(Y|X)$

$= H(X) + H(Y) - H(X,Y)$

$I(X;Y) = 8$

$I(X;Y) = H(X,Y) - H(X|Y) - H(Y|X)$

Discrete Random variable X with possible outcomes x_i ; $i = 1, 2, \ldots, n$

The self information of the event $X = x_i$ is:

$$I(x_i) = \log_2 \frac{1}{P(x_i)} = -\log_2 P(x_i)$$

MUTUAL INFORMATION:

Discrete RV $X \in \{x_i \; ; i = 1, 2, \ldots, n\}$

$Y \in \{y_j \; ; j = 1, 2, \ldots, m\}$

Suppose we observe outcome $Y = y_j$ and we want to determine the amount of information this event provides about the event $X = x_i$

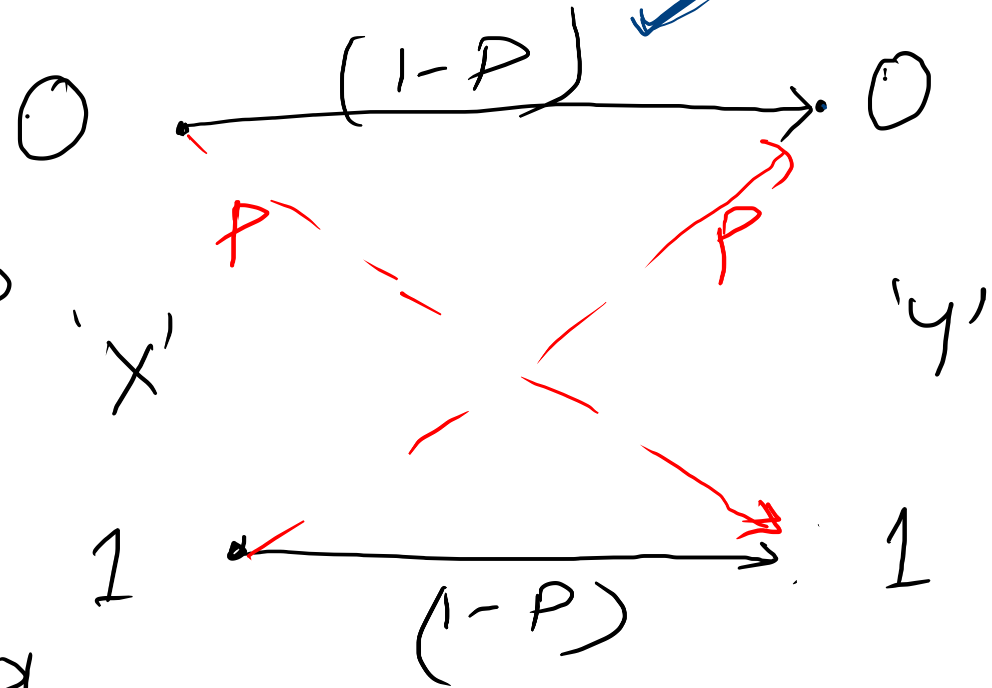
In case X and Y are independent, then $Y = y_j$ provides NO INFORMATION about $X = x_i$

In case X and Y are dependent, the occurence of $Y = y_j$ determines the occurrence of the event $X = x_i$

$$\frac{P(x_i|y_i)}{P(x_i)} = \frac{P(x_i|y_j)}{P(x_i)}\,\frac{P(y_j)}{P(y_j)} = \frac{P(x_i, y_j)}{P(x_i)\,P(y_j)} = \frac{P(y_j|x_i)\,\cancel{P(x_i)}}{\cancel{P(x_i)}\,P(y_j)}$$

$$\Rightarrow \frac{P(x_i|y_j)}{P(x_i)} = \frac{P(y_j|x_i)}{P(y_j)}$$

$$I(x_i, y_j) = \log\left[\frac{P(x_i|y_j)}{P(x_i)}\right] = \log\left[\frac{P(y_j|x_i)}{P(y_j)}\right] = I(y_j; x_i)$$



$X=0 \Rightarrow x_0$

$X=1 \Rightarrow x_1$

$Y=0 \Rightarrow y_0$

$Y=1 \Rightarrow y_1$

$$\frac{P(y_0|x_0)}{P(y_0)}$$

$$I(x_0; y_0) = I(X=0; Y=0) = \log_2\left[\frac{P(Y=0|X=0)}{P(Y=0)}\right]$$

$$I(x_0; y_0) = \log_2\left[\frac{1-P}{0.5}\right] = \log_2 2(1-P)$$

$$P(Y=0) = 0.5$$

$$= P(Y=0 \mid X=0) \cdot P(X=0) + P(Y=0 \mid X=1) \cdot P(X=1)$$

$$= (1-P) \times 0.5 + P \times 0.5 = 0.5$$

$$P(Y=1) = 0.5$$

$$I(x_0) = I(X=0) = \log_2 \frac{1}{P(X=0)} = \log_2 \frac{1}{1/2}$$

$$I(x_0) = 1$$

$\Rightarrow$ In case $P=0$, i.e., it is ideal noiseless channel.

$$I(x_0, y_0) = \log_2 2(1-P) = 1 \text{ bit}$$

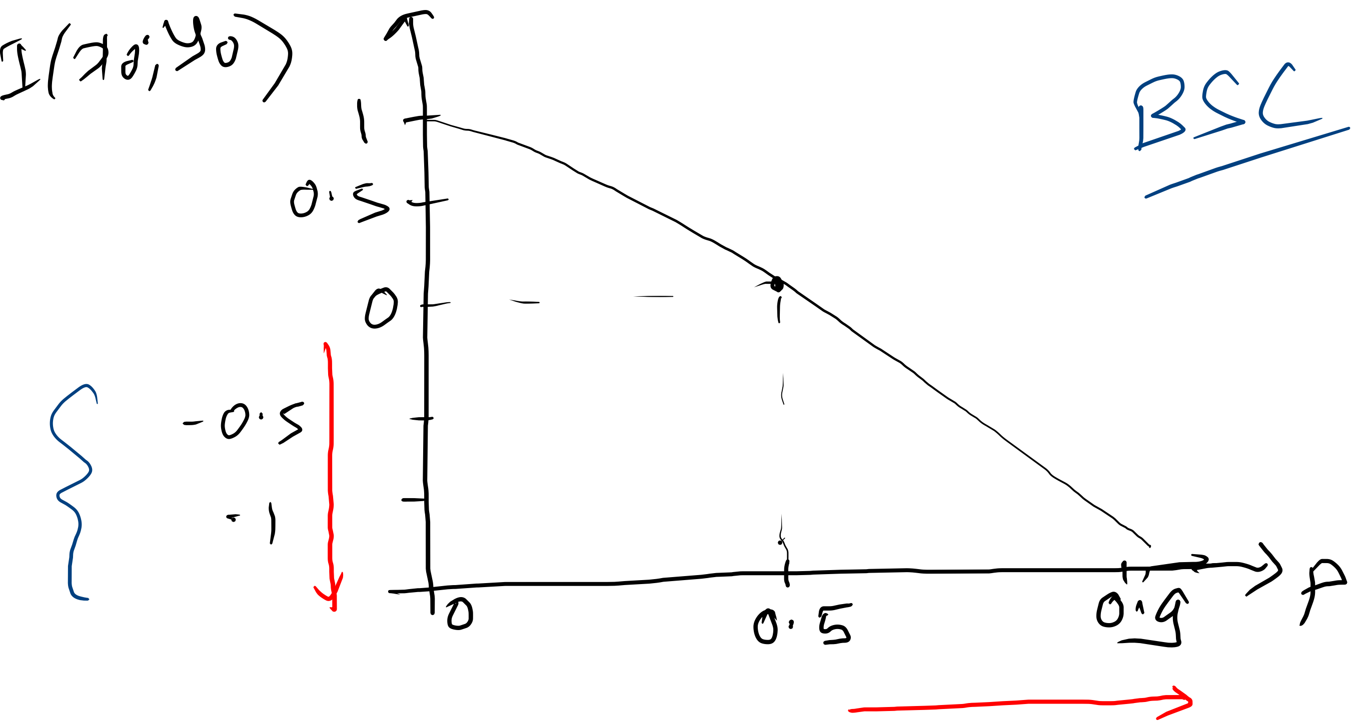
whatever uncertainty at receiver side we observe now is because of uncertainty of the transmitter side.

If $p = 0.5$,

$$I(x_0; y_0) = I(x=0; y=0) = \log_2 2\cdot(1-p) = 0$$

This implies that having observed the output, we have no information about what was transmitted.

This channel conveys no information, instead we can toss a fair coin at the receiver in order to estimate what was sent from transmitter.



$I(x_0; y_0) = \log_2 2(1-p)$

$p = 0.6$ पर

$\log_2 2(0.4) = \log_2 0.8$
$= -ve$

$\log_2(1) = 0$

$\log_2(1.1) = +ve$

$\log_2(0.99) = -ve$