# Office Hours

9/4

$\hat{\beta}_0 + \hat{\beta}_1 x_0$

$$Var(e_0) = Var(y_0 - \hat{y}_0)$$

$$= Var(y_0) + \boxed{Var(\hat{y}_0)} - \boxed{\cancel{2Cov(y_0, \hat{y}_0)}}$$

$\varepsilon_0$

(I)

depends on

$\boxed{y_1 - y_n}$

in sample observations

$$= Var(\cancel{\beta_0 + \beta_1 x_0} + \varepsilon_0) + Var(\hat{y}_0)$$

$$= \sigma^2 + \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SSX}\right)$$

$$= \sigma^2\left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SSX}\right)$$

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

$$\hat{y}(x) = \hat{\beta}_0 + \hat{\beta}_1 \underline{x}$$

$$\hat{y}_0 = \hat{y}(x_0) = \underbrace{\hat{\beta}_0 + \hat{\beta}_1}_{\text{only dependent on } \boxed{y_1 \cdots y_n}} x_0$$

$$e_i \overset{\cancel{iid}}{\sim} N\left(0, \underbrace{\sigma^2\left(1-\frac{1}{n}-\frac{(x_i-\bar{x})^2}{SSX}\right)}\right)$$

iid = ① identically ✗

② independent ✗

you can show

$$\sum_{i=1}^{n} e_i = 0 \Leftarrow \frac{\partial Q}{\partial \beta_0} \overset{!}{=} 0$$

$$e_1 + e_2 + \cdots + e_n = 0$$

$$e_1 = -e_2 - e_3 - \cdots - e_n$$

$$\Rightarrow Cov(e_1, e_2) \neq 0$$

$$P(A \cap B) = P(A)P(B)$$

$$X \perp\!\!\!\perp Y$$

$$f(x,y) = f(x)f(y)$$

$$\hat{y}_0 = \hat{y}(x) = \underline{\underline{\hat{\beta}_0}} + \hat{\beta}_1 x_0$$

$$\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{SSx}\right)\right)$$

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2/SSx)$$

$$\Rightarrow \boxed{\hat{y}_0} \sim N\left(\underbrace{\beta_0 + \beta_1 x_0}_{E(\hat{y}_0) = E(y_0)}, \underbrace{\sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SSx}\right)}_{Var(\hat{y}_0)}\right)$$

$$\boxed{y_0} = \beta_0 + \beta_1 x_0 + \varepsilon_0$$

$$\cancel{Cov(\hat{y}_0, y_0)} = Cov(\hat{\beta}_0 + \hat{\beta}_1 x_0, \underline{\beta_0 + \beta_1} x_0 + \varepsilon_0) \quad 0$$

$$= Cov(\hat{\beta}_0 + \hat{\beta}_1 x_0, \varepsilon_0)$$

$$= \cancel{Cov(\hat{\beta}_0, \varepsilon_0)} + x_0 \cancel{Cov(\hat{\beta}_1, \varepsilon_0)} \quad 0 \quad 0$$

$$\sum_{j=1}^{n} c_j y_j \qquad \sum_{j=1}^{n} k_j y_j$$

$$= \sum_{j=1}^{n} c_j (\beta_0 + \beta_1 x_j + \varepsilon_j)$$

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

$$\Rightarrow \operatorname{cov}(\varepsilon_i, \varepsilon_j) = 0 \quad \text{if } i \neq j$$

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

$$= \sum_{j=1}^{n} \frac{1}{n} y_j - \bar{x} \sum_{j=1}^{n} k_j y_j$$

$$= \sum_{j=1}^{n} \underbrace{\left(\frac{1}{n} - \bar{x} k_j\right)}_{c_j} y_j$$

$$=$$

$$\rightarrow \hat{\beta}_0 = \sum_{j=1}^{n} c_j y_j \quad \text{where} \quad c_j = \frac{1}{n} - \bar{x} k_j$$

$$\rightarrow \hat{\beta}_1 = \sum_{j=1}^{n} k_j y_j \quad \text{where} \quad k_j = \frac{x_j - \bar{x}}{SSX}$$

$$\boxed{\hat{y}_i} = \hat{\beta}_0 + \hat{\beta}_1 x_i = \sum_{j=1}^{n} c_j y_j + \left(\sum_{j=1}^{n} k_j y_j\right) x_i$$

$$= \sum_{j=1}^{n} (c_j + x_i k_j) y_j$$

$$= \sum_{j=1}^{n} \left(\frac{1}{n} - \bar{x} k_j + x_i k_j\right) y_j$$

$$= \sum_{j=1}^{n} \left( \frac{1}{n} + (x_i - \bar{x}) k_j \right) y_j$$

$$\frac{x_j - \bar{x}}{SSx}$$

$$\mathrm{Var}(\hat{y}_i) = \mathrm{Var}\left( \sum_{j=1}^{n} \boxed{\left( \frac{1}{n} + (x_i - \bar{x}) k_j \right)} y_j \right)$$

$$= \sum_{j=1}^{n} \mathrm{Var}\left( \left( \frac{1}{n} + (x_i - \bar{x}) k_j \right) y_j \right)$$

$y_1 \perp\!\!\!\perp y_2 \cdots \perp\!\!\!\perp y_n$

$\Rightarrow \mathrm{Cov}(y_1, y_2) = 0$ & so on

$$= \sum_{j=1}^{n} \left( \frac{1}{n} + (x_i - \bar{x}) k_j \right)^2 \mathrm{Var}(y_j)$$

$$= \sum_{j=1}^{n} \left( \frac{1}{n} + (x_i - \bar{x}) k_j \right)^2 \sigma^2$$

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

$$= \sigma^2 \left( \frac{1}{n} + \frac{2}{n}(x_i - \bar{x}) \underbrace{\sum_{j=1}^{n} k_j}_{0} + (x_i - \bar{x})^2 \sum_{j=1}^{n} k_j^2 \right)$$

$$= \sigma^2 \left( \frac{1}{n} + 0 + \frac{(x_i - \bar{x})^2}{SSX} \right)$$

$$\sum_{j=1}^{n} k_j = \sum_{j=1}^{n} \frac{(x_j - \bar{x})}{SSX} = \frac{1}{SSX} \underbrace{\sum_{j=1}^{n} (x_j - \bar{x})}_{\left(\sum_{j=1}^{n} x_j - n\bar{x}\right) = 0} = 0$$

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

$$(x_i - \bar{x})^2 \sum_{j=1}^{n} \left( \frac{x_j - \bar{x}}{SSX} \right)^2 =$$

$$(x_i - \bar{x})^2 \sum_{j=1}^{n} \frac{(x_j - \bar{x})^2}{(SSX)^2} = \frac{(x_i - \bar{x})^2}{(SSX)^2} \sum_{j=1}^{n} (x_j - \bar{x})^2$$

$$= \frac{(x_i-\bar{x})^2}{(SSX)^2} \cdot SSX$$

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

| $j$ | $x_i$ |
|---|---|
| 1 | 4 |
| 2 | 5 |
| 3 | 6 |

$$\sum_{i=1}^{3} x_i = 4+5+6$$

$$\sum_{j=1}^{3} x_j = 4+5+6$$

$$Var(X+Y) = Var(X) + Var(Y) + \underline{\underline{2Cov(X,Y)}}$$

if $Cov(X,Y) = 0$

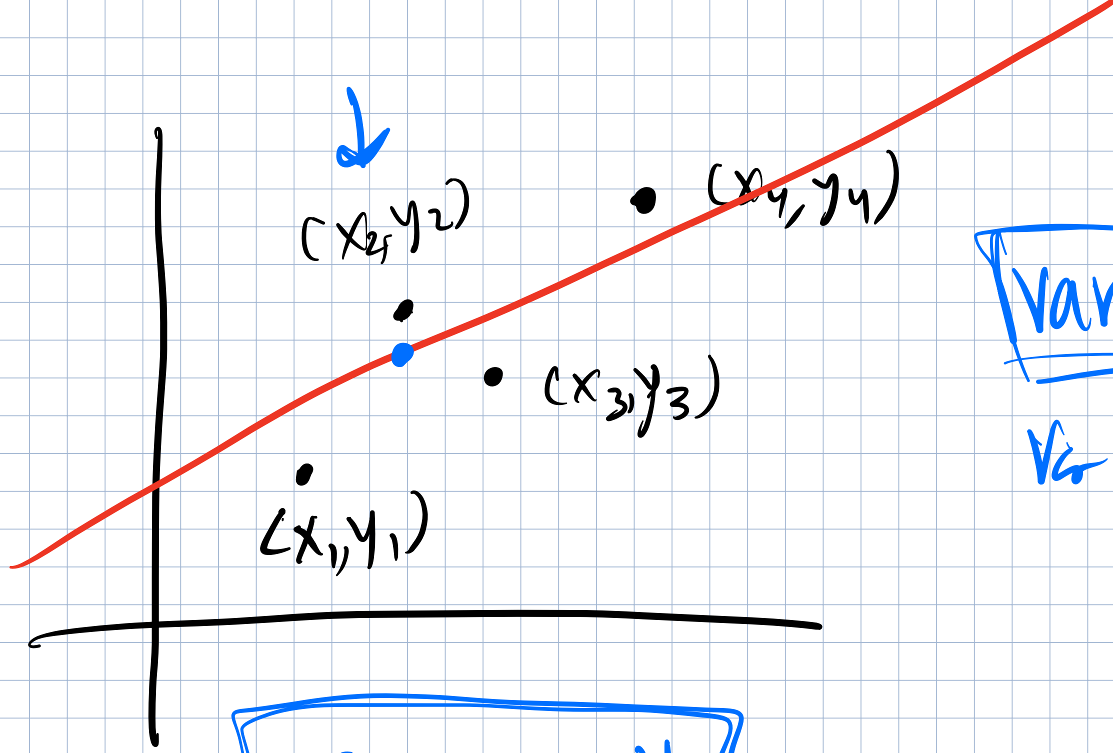
$\Rightarrow$

$$\underline{Var(X+Y)} = \underline{Var(X)} + \underline{Var(Y)}$$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) \quad \text{if } Cov(X_i, X_j) = 0, \ i \neq j$$

$$Var(a_1 \underline{\underline{X_1}} + a_2 \underline{\underline{X_2}} + \cdots + a_n \underline{\underline{X_n}})$$

$$Var(a_1 X_1) + Var(a_2 X_2) + \cdots + Var(a_n X_n)$$




$\boxed{Var(\hat{y}_i)}$

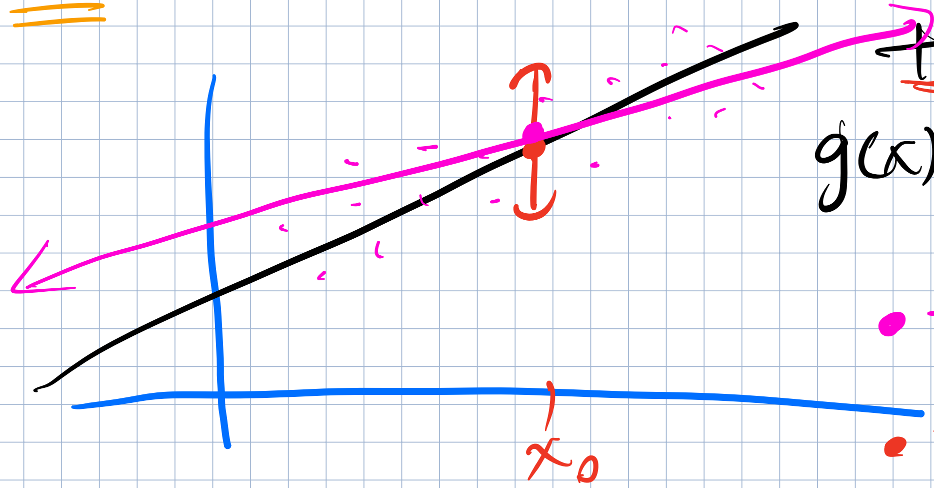
$Var(\hat{y}_2)$

$\boxed{j = 1, \ldots, 4}$

$$\hat{y}_i \sim N\left(\beta_0 + \beta_1 x_i,\ \sigma^2\left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{SSx}\right)\right)$$

CI vs PI

CI targets a pop. parameter




true line

$g(x) = \beta_0 + \beta_1 x$

$\bullet = \hat{y}(x_0)$

$\bullet = E(Y \mid X = x) = g(x)$

$\hat{y}(x_0)$ estimates $g(x_0)$

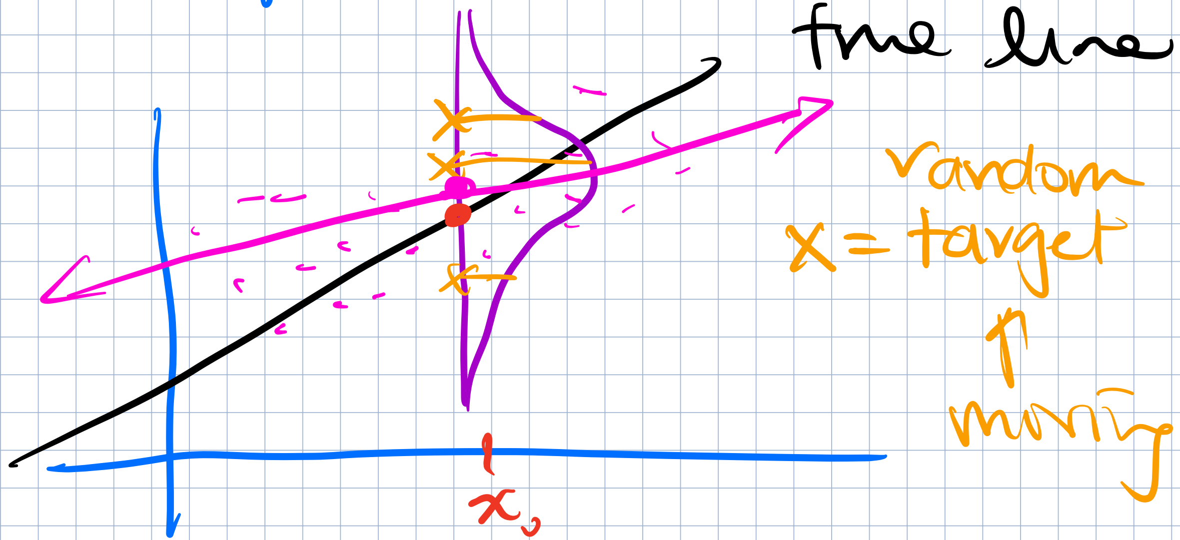
$= \underbrace{\hat{\beta}_0 + \hat{\beta}_1 x_0}_{\text{est}} \longrightarrow = \underbrace{\beta_0 + \beta_1 x_0}_{\text{truth}}$

fixed / deterministic

$\Downarrow$

CI

PI targets a random observation




$\hat{y}(x_0)$ predicts $g(x) + \varepsilon_0 = y_0$

$= \hat{\beta}_0 + \hat{\beta}_1 x_0$ predicts $\beta_0 + \beta_1 x_0 + \varepsilon_0$

new observation