Regression

9/4/25

Announcements:

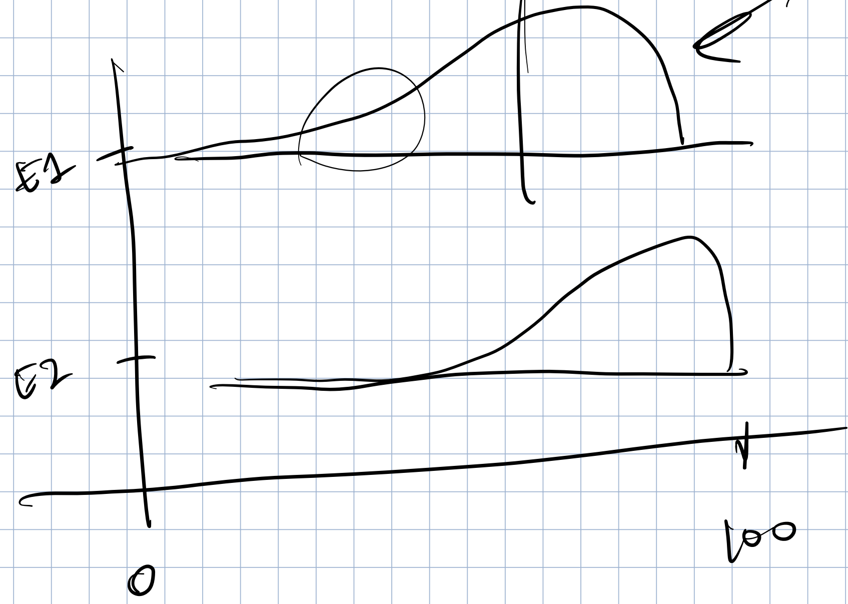
- Final Proj Sign Up on Slack Channel
  - ↳ you need my approval - DM me

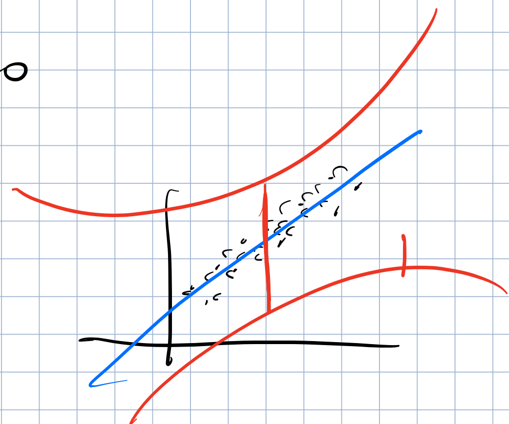
- Final Proj:
  - "Can we use LLMs?"
    - ↳ Yes for coding the app/graphic.
    - ↳ No for writing the actual language on the app/blog!!

      I want your voice & your perspective. Not some LLM's.
    - ↳ this is Academic Honesty!!

First Quiz Next Wk: Thurs 8:50 - 9:50AM

- Closed notes ← I provide the formula sheet
  - ↳ posted on Canvas/Slack
- 60 min
- no calculators
- Practice Exam up on Canvas





Recap

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

SLR model + assumptions

LS Estimators $\boxed{\hat{\theta} \pm \text{crit} \cdot \widehat{se}(\hat{\theta})}$

	$E(\cdot)$	$Var(\cdot)$	CIs	Hyp Testing
$\hat{\beta}_0$	β_0	$\sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{SSX}\right)$	...	...
$\hat{\beta}_1$	β_1	..		
$\hat{y}_i$	$\beta_0 + \beta_1 x_i$	$\sigma^2\left(\frac{1}{n} + \frac{(x_i-\bar{x})^2}{SSX}\right)$		
e_i	0		X	X

CI for $E(Y|X=x) = \boxed{\beta_0 + \beta_1 x}$

$$\hat{y}_i \pm t^*_{1-\frac{\alpha}{2}, n-2} \cdot \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(x_i-\bar{x})^2}{SSX}}$$

$$\frac{\sum_{i=1}^{n} e_i^2}{\sigma^2} \sim \chi^2_{n-p} \qquad p=2 \text{ for SLR}$$

↳ unbiased est of σ^2:

$$MSE = \hat{\sigma}^2 = \frac{\sum_i e_i^2}{n-2} = \frac{SSE}{n-2}$$

Sum of Squares Decomp & F-test:

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

$$SST = SSE + SSR$$

$$F = \frac{SSR/1}{SSE/n-2} \overset{H_0}{\sim} F_{1,n-2}$$

If $F^*_{1,n-2}(.95) < F$ then reject H_0 ⟺ X is a significant pred. of Y.

~~Diagnostics (?)~~

$$SE(\hat{y}(x=80) - y_{new}(x=80)) = \sqrt{Var(\hat{y}(x=80) - y_{new}(x=80))}$$

...

$$Var(\hat{y}(x=80) - y_{new}(x=80)) = \overset{①}{Var(\hat{y}(x=80))} + \overset{②}{Var(y_{new}(x=80))}$$

$$\cancel{-2\,Cov(\boxed{\hat{y}(x=80)}, \boxed{y_{new}(x=80)})} \to 0$$

$$\hat{y}(x=80) = \hat{\beta}_0 + \hat{\beta}_1 \cdot 80$$

$$① \quad Var(\hat{y}(x=80)) = \sigma^2\left(\frac{1}{n} + \frac{(80-\bar{x})^2}{SSX}\right)$$

$$② \quad Var(\underline{y_{new}(x=80)}) = Var(\underline{\beta_0 + \beta_1(80)} + \varepsilon_{new})$$

$$= Var(\varepsilon_{new}) = \sigma^2$$

$$Var(\hat{y}(x=80) - y_{new}(x=80)) = \sigma^2\left(\frac{1}{n} + \frac{(80-\bar{x})^2}{SSX}\right) + \sigma^2$$

$$= \sigma^2\left(1 + \frac{1}{n} + \frac{(80-\bar{x})^2}{SSX}\right)$$

$$SE(\hat{y}(x=80) - y_{new}(x=80)) = \sigma\sqrt{\underline{1} + \frac{1}{n} + \frac{(80-\bar{x})^2}{SSX}}$$

Pred interval for Y_{new} when $X = x$:

$$\hat{y}(x) \pm t^{*}_{1-\frac{\alpha}{2}, n-2} \, \hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{SSX}}$$