Office Hours

9/3/25

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

fixed

random

$\left.\begin{matrix}\hat{\beta}_0 \\ \hat{\beta}_1\end{matrix}\right\}$ random

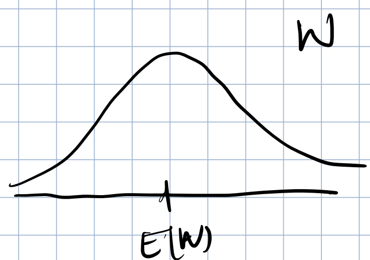
Idea:

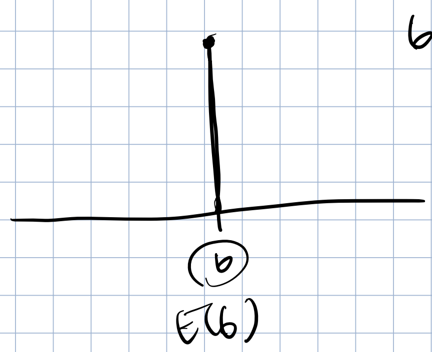
random / fixed

$$E(\ldots) = \underset{\text{fixed}}{\underline{\qquad\qquad}}$$

unknown fixed #

$$E(\beta_0) = \beta_0$$





Unbiasedness

← property of estimators

estimators are random e.g. $\hat{\beta}_0$ or $\hat{\beta}_1$

$$E(\hat{\beta}_0) = \beta_0 \qquad E(\hat{\beta}_1) = \beta_1$$

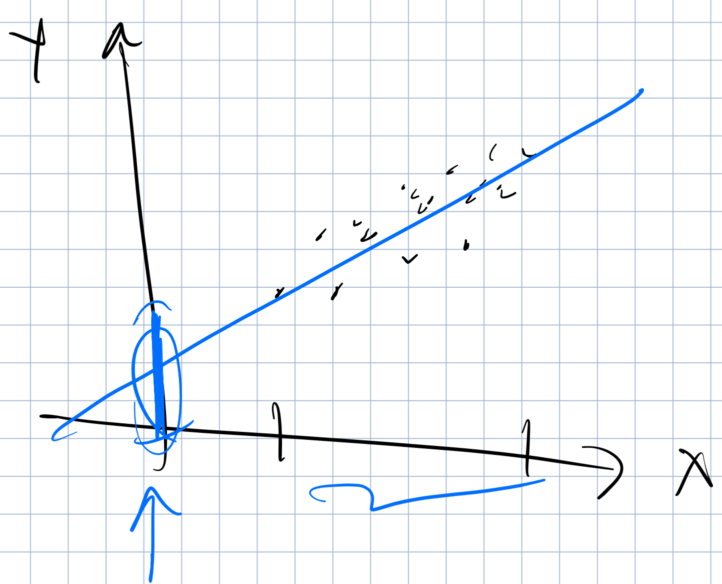
1b. β_0 is the expected sales for a district w/ pop. zero.

$\hat{\beta}_0$ is our best guess for this #

95% CI:

7.4312 w/ [−1.1852, 16.0476]

From the real life context we could imagine some data like this:



5.

$$e_i \sim N\left(0, \sigma^2\left(1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{S_{xx}}\right)\right)$$

$i = 1, \dots, n$

Consider a new out of sample obs $(x_0, \_\_)$

?

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

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

$$e_0 = y_0 - \hat{y}_0$$

What's $E(e_0)$? $Var(e_0)$?

$Var(b + Y) = Var(Y)$

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

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

$$= Var(\varepsilon_0 - \hat{y}_0)$$

$Var(u+v) = Var(u) + Var(v) + 2Cov(u,v)$

$$= Var(\varepsilon_0) + Var(\hat{y}_0) - 2Cov(\varepsilon_0, \hat{y}_0)$$

$\bar{y} - \hat{\beta}_1 \bar{x}$ $\quad \sum_i k_i y_i$

$$= \boxed{} + \boxed{} - 2Cov(\varepsilon_0, \hat{\beta}_0 + \hat{\beta}_1 x_0)$$

$\varepsilon_0 \perp\!\!\!\perp \varepsilon_1, \dots, \varepsilon_n \Rightarrow Cov(\varepsilon_0, \hat{y}_0) = 0$

$$= \boxed{\sigma^2} + \boxed{\sigma^2\left(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{SSX}\right)} + \underline{0}$$

$$= \sigma^2\left(1 + \frac{1}{n} + \cdots\right)$$

Properties of Var

In general for any RVs U & V

$$Var(U+V) = Var(U) + Var(V) + 2Cov(U,V)$$

↳ Def of $Var(U+V) =$

$$Var((U+V)) = E\left(\left[(U+V) - E(U+V)\right]^2\right)$$

$$= \vdots$$

$$= E\left((U-E(U))^2\right) + E\left((V-E(V))^2\right) +$$

$$+ 2\underbrace{E\left((U-E(U))(V-E(V))\right)}_{Cov(U,V)}$$

3. mean response for a given x-level:

$$E(y \mid X = x_0) = \underbrace{\beta_0 + \beta_1 x_0}$$

↑unknown parameter

CI for unknown parameters:

pt. est $\pm$ crit. value $\cdot$ $\widehat{SE}$(pt. est)

in this case:

$$\boxed{\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0 \pm t^*_{1-\alpha/2,\, n-2} \; \widehat{SE}(\hat{y}_0)}$$

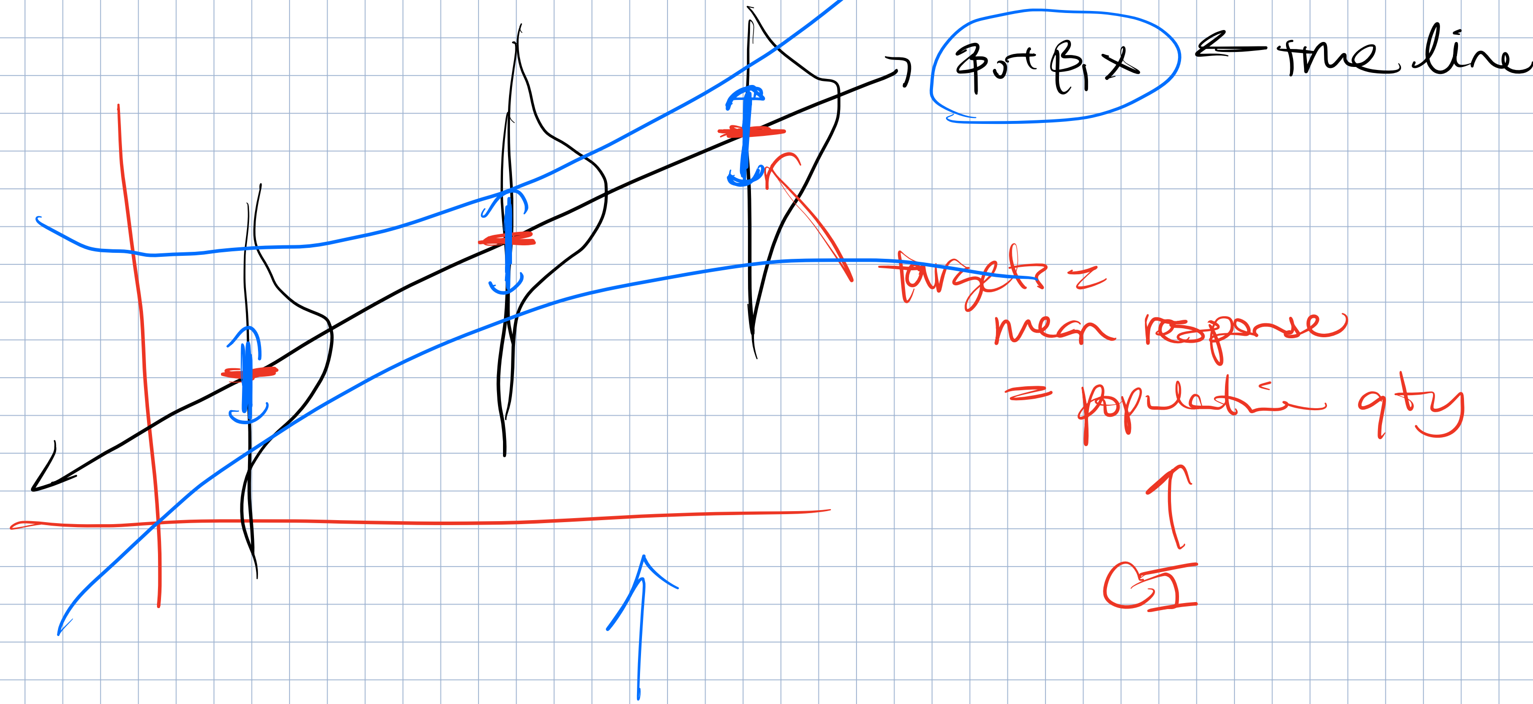
$$\hat{y}_i \sim N\left(\beta_0 + \beta_1 x_i,\ \sigma^2\left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{SSX}\right)\right)$$

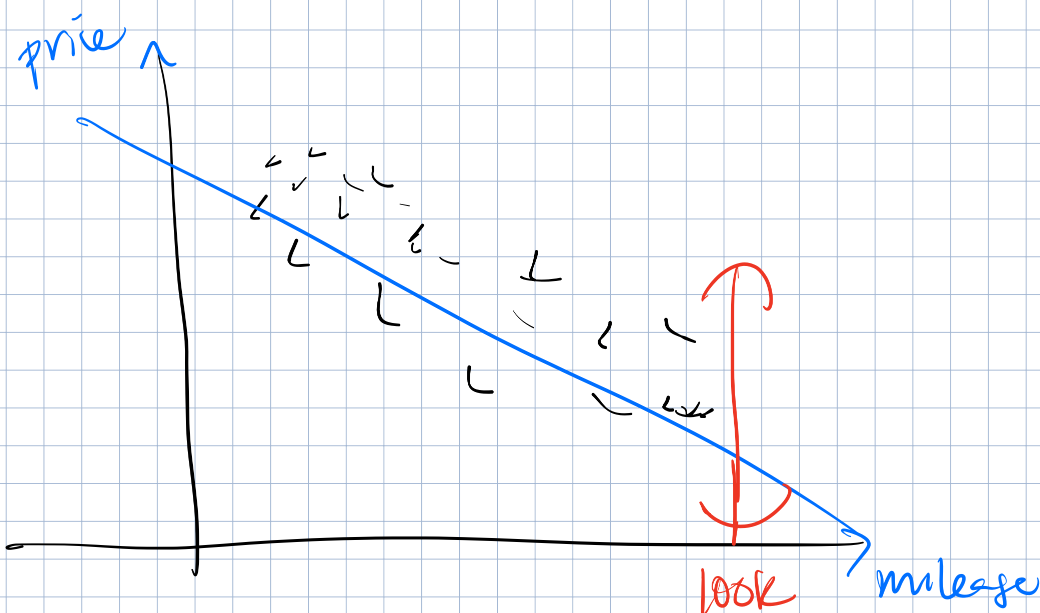
$$Var(\hat{y}_0) = \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SSX}\right)$$

$$SE(\hat{y}_0) = \sigma\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SSX}}$$

$$\widehat{SE}(\hat{y}_0) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SSX}}$$

$$\hat{\beta}_0 + \hat{\beta}_1 x_0 \pm t^*_{\ldots}\ \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SSX}}$$





$i = 1, \ldots, n$

prediction Intervals are for single new obs.

& they are always wider than CI

More randomness / uncertainty in a moving target.

: what's the correct exp for a PI?

$$y_{new} = \beta_0 + \beta_1 X_{new} + \varepsilon_{new}$$

$$\hat{y}_{new} = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 X_{new}}$$

$$\underline{\hat{y}_{new} - y_{new}} \sim N\left(\underline{\ 0\ }, \underline{\sigma^2\left(1 + \frac{1}{n} + \frac{(X_{new} - \bar{X})^2}{SSX}\right)}\right)$$

$$E(\hat{y}_{new} - y_{new}) = E(\hat{y}_{new}) - E(y_{new})$$

$$= \beta_0 + \beta_1 X_{new} - E(\underline{\beta_0 + \beta_1 X_{new}} + \varepsilon_{new})$$

$$= \underline{\beta_0 + \beta_1 X_{new}} - (\underline{\beta_0 + \beta_1 X_{new}} + \cancel{E(\varepsilon_{new})}^{0})$$

$$= 0$$

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

$$- 2\,\overset{③}{Cov(\hat{y}_{new}, y_{new})}$$

$$= \sigma^2\left(\frac{1}{n} + \frac{(X_{new} - \bar{X})^2}{SSX}\right) + \sigma^2$$

$$- 2\,\boxed{0}$$

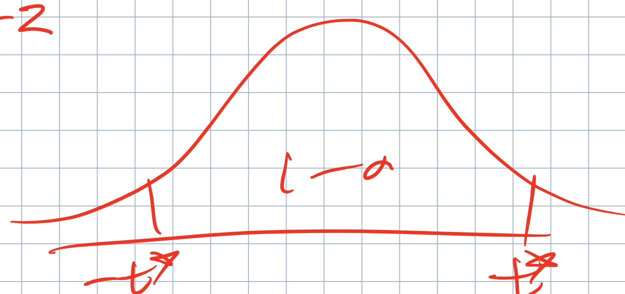
$$= \sigma^2 \left(1 + \frac{1}{n} + \frac{(x_{new}-\bar{x})^2}{SSX}\right)$$

$$\hat{y}_{new} - y_{new} \sim N\left(0, \sigma^2\left(1+\frac{1}{n}+\frac{(x_{new}-\bar{x})^2}{SSX}\right)\right)$$

$$\frac{\hat{y}_{new} - y_{new}}{SE(\hat{y}_{new} - y_{new})} \sim N(0, 1)$$

$$\frac{\hat{y}_{new} - y_{new}}{\sigma\sqrt{1+\frac{1}{n}+\frac{(x_{new}-\bar{x})^2}{SSX}}} \sim N(0,1)$$

$$\frac{\hat{y}_{new} - y_{new}}{\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(x_{new}-\bar{x})^2}{SSX}}} \sim t_{n-2}$$

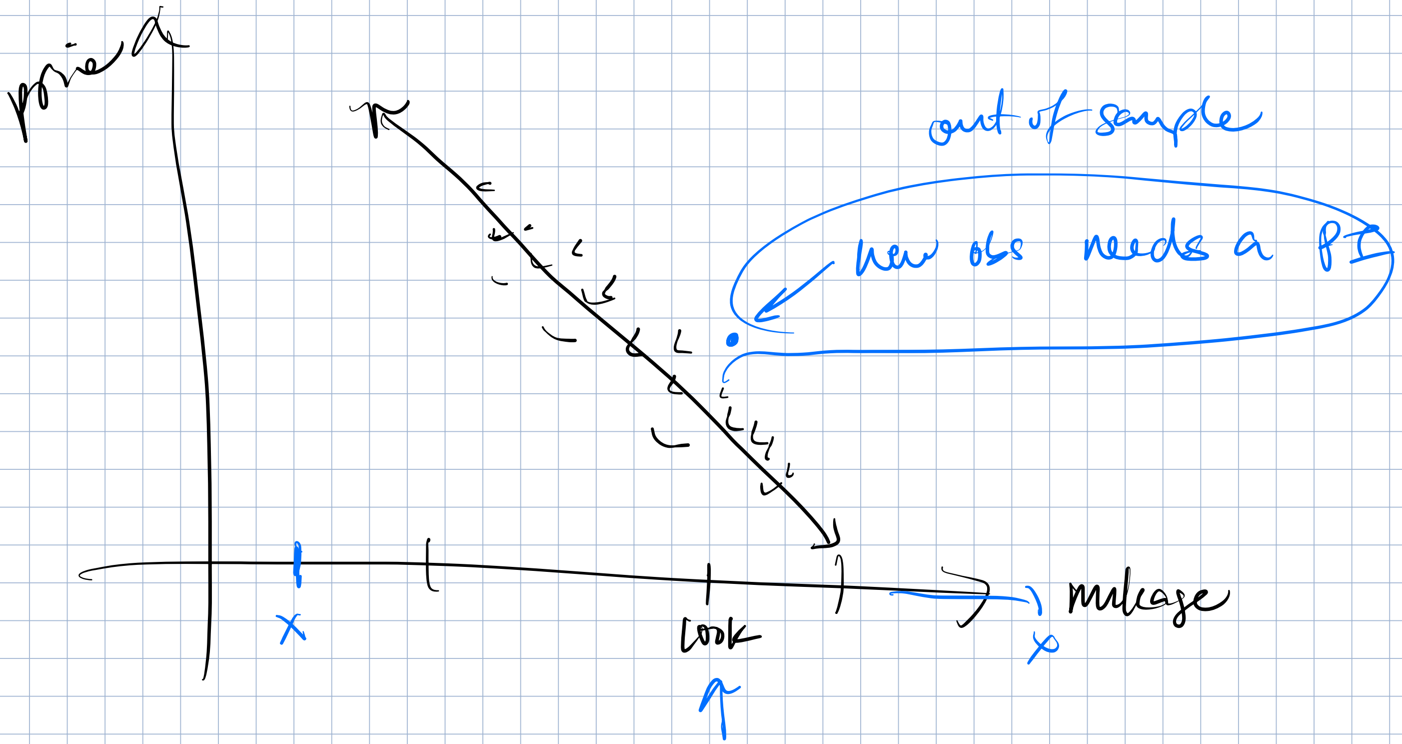


$$P\left(-t^*_{1-\alpha/2,n-2} \le \frac{\hat{y}_{new}-y_{new}}{\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(x_{new}-\bar{x})^2}{SSX}}} \le t^*_{1-\alpha/2,n-2}\right)=1-\alpha$$

↑ solve for target y_{new}

⇓

$$P\left(y_{new} \in \underbrace{\left[\hat{y}_{new} \pm t^*_{1-\alpha/2,n-2}\,\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(x_{new}-\bar{x})^2}{SSX}}\right]}_{PI}\right)=1-\alpha$$



How can we show $\sum_{i=1}^n x_i e_i = 0$?

Recall: $Q(\beta_0,\beta_1)=\sum_{i=1}^n (y_i-\beta_0-\beta_1 x_i)^2$

$$\frac{\partial Q}{\partial \beta_1} = -2\sum_{i=1}^n x_i(y_i-\beta_0-\beta_1 x_i) \overset{!}{=} 0$$

$$\left.\frac{\partial Q}{\partial \beta_1}\right|_{(\hat\beta_0,\hat\beta_1)} = -2\sum_{i=1}^{n} x_i(\underbrace{y_i-\hat\beta_0-\hat\beta_1 x_i}_{e_i}) \overset{!}{=} 0$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} x_i e_i = 0$$