Confidence Intervals: given $\underline{\underline{X}}$, $\underline{y}$, what are plausible values of β_j?

Is b_j a plausible value?

"Plausible values" are those values not rejected under the hypothesis: $H: \beta_j = b$.

Formally,

a $(1-\alpha)\%$ CI is the set of values b not rejected at level α, under $H: \beta_j = b$.

$$\text{a } (1-\alpha) \text{ CI} = \hat{\beta}_j \pm SE(\hat{\beta}_j) * t_{1-\frac{\alpha}{2}}$$

$$\approx \hat{\beta}_j \pm 2\,SE(\hat{\beta}_j) \text{ if } \alpha = 0.05$$

(95% CI)

know: Conceptually, a CI is a random interval.

β_j is fixed. & unknown.

CI are probability statements about the random interval.

A 95% CI can be written mathematically:

$$Pr\left(\ell(\underline{\underline{X}}, \underline{Y}) < \beta_j < u(\underline{\underline{X}}, \underline{Y}) \mid \beta_j\right) = 0.95$$

$\underline{Ex}$ $\quad \hookrightarrow$ CI random

Pre-experiment : $\Pr(\beta_j \in 95\% \text{ CI} \mid \beta_j) = ?$ $\quad .95$

Post-experiment : $\Pr(\beta_j \in 95\% \text{ CI} \mid \beta_j) = ?$ $\quad 0$ or 1

$\hookrightarrow$ CI fixed

<u>Prediction Interval</u> : given $(\underline{\underline{X}}, \underline{y})$ predict what y_* will look like given $\underline{x}_*$

$x_* \in \mathbb{R}^{p \times 1}$

idea: $y_* = \underline{x}_*^T \underline{\beta} + \varepsilon_*$

$$\approx \hat{y}_* = \underline{x}_*^T \hat{\underline{\beta}} \quad \text{predicted value}$$

How accurate is the prediction?

I care about $\boxed{\hat{\varepsilon}_*}$.

what's $\mathbb{E}[\hat{\varepsilon}_*]$?

$$\mathbb{E}[\hat{\varepsilon}^*] = \mathbb{E}[y_* - \hat{y}_*] = \underline{x}_*^T \underline{\beta} - \mathbb{E}[\underline{x}_*^T \hat{\underline{\beta}}]$$

$$= \underline{x}_*^T \underline{\beta} - \underline{x}_*^T \underline{\beta}$$

$$= 0$$

What about $\text{var}(\hat{\varepsilon}^*)$?

$$\hat{\varepsilon}^* = y_* - \hat{y}_* = \underline{x}_*^T \underline{\beta} + \varepsilon_* - \underline{x}_*^T \hat{\underline{\beta}}$$

$$\text{var}(\hat{\varepsilon}^*) = \underbrace{\text{var}(\varepsilon_*)}_{\sigma^2} + \underbrace{\text{var}(\underline{x}_*^T \underline{\beta} - \underline{x}_*^T \hat{\underline{\beta}})}$$

$$\text{var}(\underline{x}_*^T \hat{\underline{\beta}}) = \text{var}(\underline{x}_*^T (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \ldots$$

σ^2

$x_*^T (X^T X)^{-1} X^T X (X^T X)^{-1} x_*$

$$= \sigma^2 \left(1 + \underline{x}_*^T (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} x_* \right)$$

Prediction interval:

Let $SE(\hat{\varepsilon}^*) = \hat{\sigma} \sqrt{1 + \underline{x}_*^T (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{x}_*}$

then $\dfrac{\hat{\varepsilon}^*}{SE(\hat{\varepsilon}^*)} \sim t_{n-p}$

$\hat{\varepsilon}_* \pm SE(\hat{\varepsilon}_*) * t_{1-\frac{\alpha}{2}}$ is a $1-\alpha$ PI for $\hat{\varepsilon}_*$.

$y_* = \hat{y}_* + \hat{\varepsilon}_*$ so $\hat{y}_* \pm SE(\hat{\varepsilon}_*) \cdot t_{1-\frac{\alpha}{2}}$ is a $(1-\alpha)$ Prediction Interval for y^*.