Prediction Intervals

3.15

3.8.2: "CI for a future observation

$y_0 = x_0'\beta + \varepsilon_0$, estimate with $\hat{y}_0 = x_0' b$

zero not $n+1$ — common unfortunate notation

$\varepsilon_1, \ldots, \varepsilon_n, \varepsilon_0$ all independent, yes?

$$t = \frac{z}{\sqrt{w/r}}$$

Theorem: A $(1-\alpha)100\%$ prediction interval for a new observation y_0 is given by

$$x_0' b \pm t_{\alpha/2}\, s \sqrt{1 + x_0'(X'X)^{-1}x_0}$$

compare CI for $\ell'\beta$

$$\ell' b \pm t_{\alpha/2}\, s \sqrt{\ell'(X'X)^{-1}\ell}$$

Proof

3.16

$y_0 \sim N(x_0'\beta, \sigma^2)$, $x_0'b \sim N(x_0'\beta, \sigma^2 x_0'(X'X)^{-1}x_0)$

independent, so

$$y_0 - x_0'b \sim N(0, \sigma^2 + \sigma^2 x_0'(X'X)^{-1}x_0)$$

standardizing, obtaining

$$z = \frac{y_0 - x_0'b}{\sqrt{\sigma^2(1+x_0'(X'X)^{-1}x_0)}} \sim N(0,1) \quad \text{and}$$

$$t = \frac{z}{\sqrt{w/r}} = \frac{\dfrac{y_0 - x_0'b}{\cancel{\sigma}\sqrt{1+x_0'(X'X)^{-1}x_0}}}{\sqrt{\dfrac{e'e}{\cancel{\sigma^2}}\Big/(n-k-1)}}$$

numerator & denominator are independent bec. e is independent of $\begin{pmatrix} b \\ y_0 \end{pmatrix}$ so

$$t = \frac{y_0 - x_0'b}{s\sqrt{1+x_0'(X'X)^{-1}x_0}} \sim t(n-k-1), \text{ and}$$

$$1-\alpha = P_r\left\{ -t_{\alpha/2} < \frac{y_0 - x_0'b}{s\sqrt{1+x_0'(X'X)^{-1}x_0}} < t_{\alpha/2} \right\}$$

And isolate y_0