Week 3 Monday Worksheet

Valentino Aceves

Distribution of $\hat{b}_0$ and $\hat{b}_1$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$$

$$\hat{b}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

Note that $\hat{b}_0$ and $\hat{b}_1$ are functions of random variables, so they are also random. Because they are random, we are interested in their distributions.

$$\hat{b}_0 \sim N(\beta_0, \mathrm{Var}(\hat{b}_0)) \quad \text{where} \quad \mathrm{Var}(\hat{b}_0) = \sigma^2 \frac{\frac{1}{n}\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \approx \sigma^2 \frac{\mathrm{Var}(x) + \bar{x}^2}{n\mathrm{Var}(x)}$$

$$\hat{b}_1 \sim N(\beta_1, \mathrm{Var}(\hat{b}_1)) \quad \text{where} \quad \mathrm{Var}(\hat{b}_1) = \sigma^2 \frac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \approx \sigma^2 \frac{1}{n\mathrm{Var}(x)}$$

where n = number of observations, $\sigma^2 = \mathrm{Var}(\varepsilon_i)$ = variance of errors.

Notice that:

Larger sample size $n \implies$ smaller variances of estimators (i.e. our estimators are more precise).

Larger variance of explanatory variable $Var(x) \implies$ smaller variances of estimators.

Smaller variance of error $\mathrm{Var}(\varepsilon_i) = \sigma^2 \implies$ smaller variances of estimators.

More Stuff:

Properties of $\varepsilon_i = y_i - \hat{b}_0 - \hat{b}_1 x_i$:

$$\sum_{i=1}^{n} \varepsilon_i = 0 \qquad \sum_{i=1}^{n} \varepsilon_i x_i = 0 \qquad \varepsilon \sim N(0, \sigma^2)$$

Another problem: The errors, ε_i, are also a random variable. So how do we estimate the variance of the errors, $\mathrm{Var}(\varepsilon_i) = \sigma^2$?

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \varepsilon_i^2 = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \frac{SSR}{n-2}$$

$$E[Y|X = x] = \beta_0 + \beta_1 x \qquad \Longrightarrow \qquad \frac{\partial E[Y|X = x]}{\partial x} = \beta_1$$

$$\mathrm{Var}(\hat{y}) = \mathrm{Var}(\hat{b}_0 + \hat{b}_1 x) = \mathrm{Var}(\hat{b}_0) + \mathrm{Var}(\hat{b}_1) x^2 + 2x\mathrm{Cov}(\hat{b}_0, \hat{b}_1)$$