Bias and Variance of Slope and Intercept

- Slope:

$$E(\hat{\beta}_1) = \beta_1 \qquad V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$$

- Intercept:

$$E(\hat{\beta}_0) = \beta_0 \qquad V(\hat{\beta}_0) = \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$$

Variance Formulae

- Error sum of squares:

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

- Can show that $E(SS_E) = (n-2)\sigma^2$
- Therefore $\hat{\sigma}^2 = \dfrac{SS_E}{n-2}$ is unbiased

Variance Formulae

- Another formula for SS_E

$$SS_E = SS_T - \hat{\beta}_1 S_{xy}$$

  - where $SS_T \equiv \sum_{i=1}^{n} (y_i - \bar{y})^2$

Estimated Standard Errors (definitions)

- Slope:

$$se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- Intercept:

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right]}$$

Use of Regression

- Care should be taken in:
  - selecting variables (strong observed correlation does not always imply a causal relationship)
  - extrapolating beyond the original range of the regressor x