The Sampling Distribution of $\hat{\beta}$

The $\hat{\beta}$s produced by OLS will differ when a different sample is used for the same model. They are random variables with a **sampling distribution**.

Assumption 7 requires the error term be normally distributed. If the error term is normally distributed then the $\hat{\beta}$s will also be normally distributed, i.e.

$$\hat{\beta} \sim N(E(\hat{\beta}), Var(\hat{\beta})).$$

Properties of Estimators

The Mean

Recall, an estimator is **unbiased** if its expected value equals its true value, i.e. $E(\hat{\beta}) = \beta$.

The fact that an estimator is unbiased does not necessarily mean the single estimate we see will be close to the population mean. However, a single estimate drawn from an unbiased distribution is more likely to be near the true value than one drawn from a biased distribution (assuming the two estimators have sampling distributions with the same variance).

The Variance

So what we know so far is that the sampling distribution of the $\hat{\beta}$s are normally distributed with a mean of β. What about the variance?

Let's work out the variance for the slope coefficient in a simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i.$$

The Gauss-Markov Theorem

If Assumptions 1 – 6 are satisfied, OLS works well! Why? Because of the G-M Theorem.

This theorem states that when assumptions 1 – 6 are true, the OLS estimator of β_j is the minimum variance estimator from among the set of all linear unbiased estimators of β_j for all $j = 0,1, \ldots, k$. That is, the OLS estimator is **BLUE**, the Best Linear Unbiased Estimator.

Best: This just means that $\hat{\beta}_j$ has the smallest variance out of all linear unbiased estimators of β_j. This property is known as **efficiency**.

Linear: The estimator is linear in Y.

Unbiased: $E(\hat{\beta}_j) = \beta_j$ for all $j = 0,1,\ldots,k$

If we add Assumption 7, the OLS estimator becomes **BUE**.