

10.1: A Recap of Modeling Assumptions

Recall from Chapter 4 that we identified three key assumptions about the error term that is necessary for OLS to provide unbiased, efficient linear estimators; a) errors have identical distributions, b) errors are independent, c) errors are normally distributed.¹⁷

Error Assumptions

- Errors have identical distributions

$$E(\epsilon_i) = 0 \quad E(\epsilon_i^2) = \sigma^2$$

- Errors are independent of XX and other ϵ_i

$$E(\epsilon_i) = E(\epsilon_i | x_i) = 0 \quad E(\epsilon_i) = E(\epsilon_j) = 0$$

and

$$E(\epsilon_i) \neq E(\epsilon_j) \quad E(\epsilon_i) \neq E(\epsilon_j) \text{ for } i \neq j$$

- Errors are normally distributed

$$\epsilon_i \sim N(0, \sigma^2) \quad \epsilon_i \sim N(0, \sigma^2)$$

Taken together these assumptions mean that the error term has a normal, independent, and identical distribution (normal i.i.d.). Figure 10.1.1 shows what these assumptions would imply for the distribution of residuals around the predicted values of YY given XX .

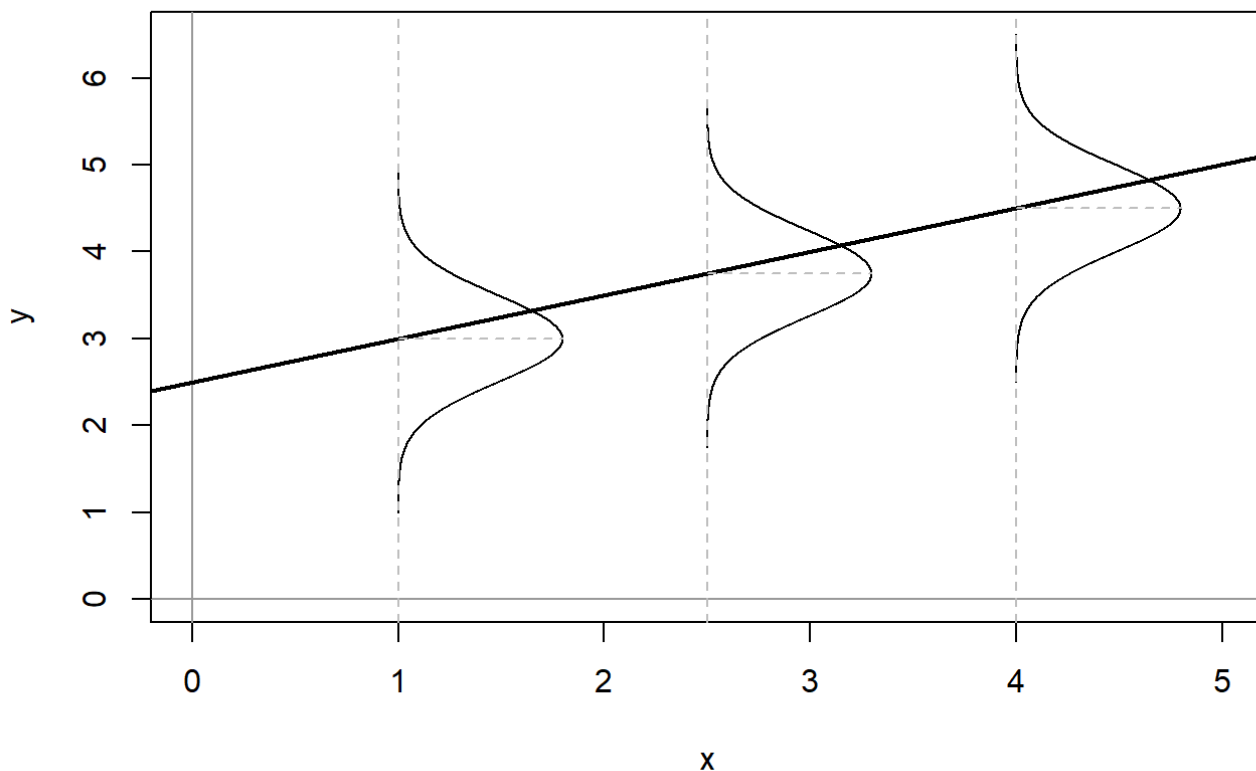


Figure 10.1.1: Assumed Distributions of OLS Residuals

How can we determine whether our residuals approximate the expected pattern? The most straightforward approach is to visually examine the distribution of the residuals over the range of the predicted values for Y . If all is well, there should be no obvious pattern to the residuals – they should appear as a “sneeze plot” (i.e., it looks like you sneezed on the plot. How gross!) as shown in Figure 10.1.2

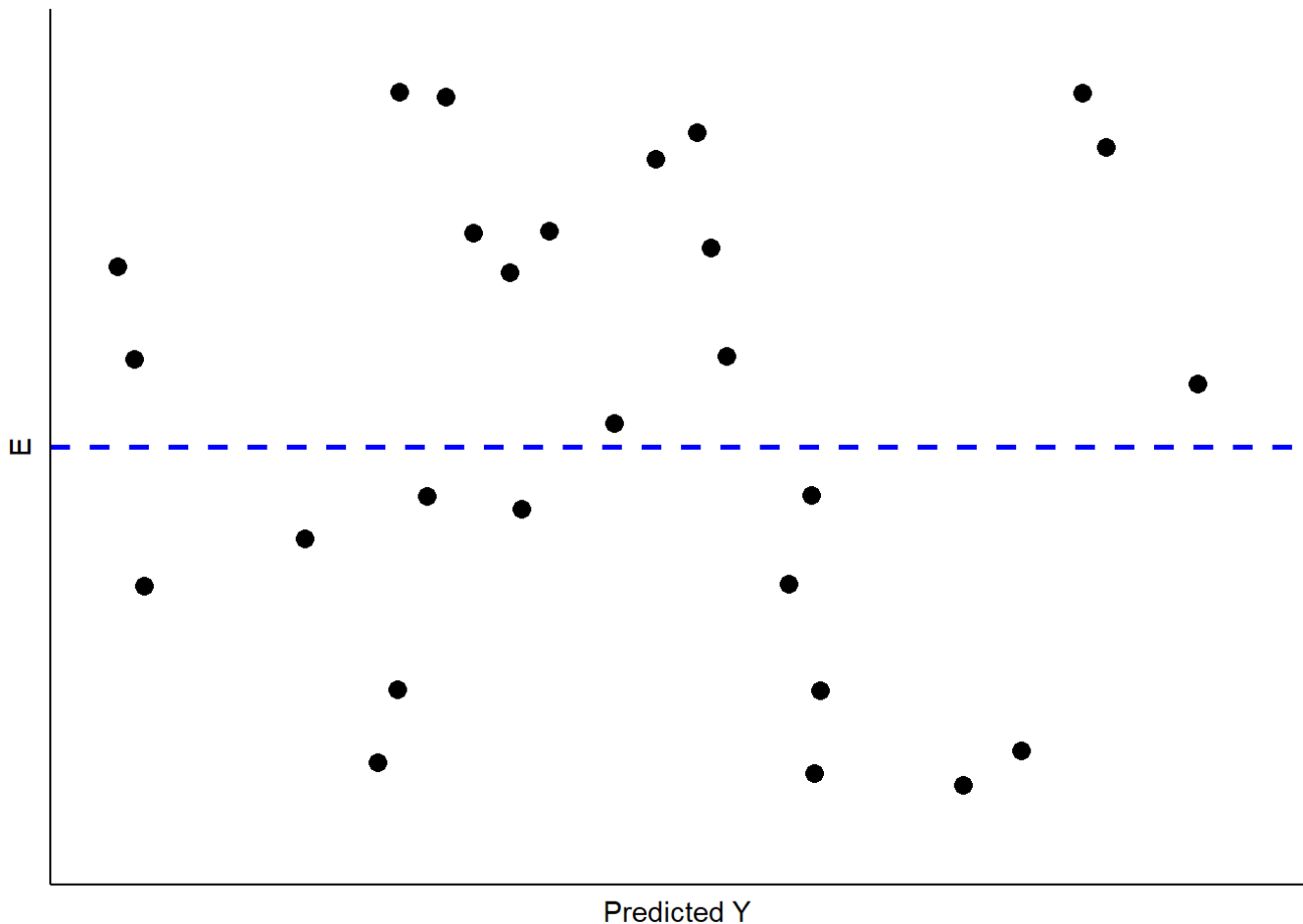


Figure 10.1.2: Ideal Pattern of Residuals from a Simple OLS Model

Generally, there is no pattern in such a sneeze plot of residuals. One of the difficulties we have, as human beings, is that we tend to look at randomness and *perceive* patterns. Our brains are wired to see patterns, even where they are none. Moreover, with random distributions, there will in some samples be clumps and gaps that do *appear* to depict some kind of order when in fact there is none. There is the danger, then, of over-interpreting the pattern of residuals to see problems that aren't there. The key is to know what kinds of patterns to look for, so when you do observe one you will know it.

This page titled [10.1: A Recap of Modeling Assumptions](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Jenkins-Smith et al. \(University of Oklahoma Libraries\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.