Lecture Summary: Feb. 1, 2023

- Basic diagnostic graphical tools:

Dot plot, stem-and-leaf, boxplot

- Residuals

$$\hat{\epsilon}_i = Y_i - \hat{Y}_i, \ \ i = 1, \ldots, n.$$

- Types of residual plots:

1. against the predictor variable, x;

2. against the fitted values, $\hat{Y}$;

3. box plot of residuals; and

4. normal probability plot of residuals.

- What is a good pattern?

No pattern. The residual plot should be roughly evenly around the zero line; no obvious pattern of any kind. The box plot of the residuals should be approximately symmetric; the normal probability plot should be roughly a straight line.

- The residuals can be used to check if

1. the regression function is not linear (e.g., pp. 105–106);

2. the variance of the errors is not constant (e.g., p. 107);

3. the errors are not independent (e.g., p. 109);

4. outliers (e.g., p. 109);

5. the errors are not normal (e.g., p. 112); and

6. some important predictors are missing.