Section 11/18

1   Question 12.1

Recall that the intercept of the regression line is given by "the average of Y minus the slope times the average of x". That is, $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$. Is $\hat{\beta}_0$ an unbiased estimator of β_0?

2   Question 12.2

The fitted value of the response for the ith individual is the height of the regression line at x_i:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Show that $\hat{Y}_i$ is an unbiased estimator of $\beta_0 + \beta_1 x_i$, the height of the true line at x_i.

3   Residual Practice

Assume that unknown to a set of researchers, the true distribution of error in their regression was as follows:

$$\epsilon_i | (X_i = x) \sim N(\frac{(x - E(X))^2}{100}, \sigma^2)$$

a) Draw an example of what a residual plot for this regression might look like.

b) Would linear regression be a good idea for a regression like this? Why or why not?

c) Repeat parts a and b, but now using:

$$\epsilon_i | (X_i = x) \sim N(0, x\sigma^2)$$