

Concept Check #6

HST 190: Introduction to Biostatistics

- The project is out-of-sync -- use ``renv::status()`` for details.

This worksheet is open-note. Take 15 minutes to consider the following questions. We will discuss as a group.

1. Consider a hypothetical dataset (X_i, Y_i) with n observations. The regression equation $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ is proposed as a way to relate X and Y . In this regression equation, is β_1 a model, a parameter, or a random variable? Explain.
2. After an introductory biostatistics course, a colleague, considering the same linear regression equation as appears in (1) above, $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, claims that β_1 is observable. Are they correct in their assessment? Yes or no, and explain.
3. After consulting a statistician, your colleague decides to apply linear regression to the dataset (X_i, Y_i) for $n = 200$ (no longer hypothetical); their fitted regression is of the form $Y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$. Is $\hat{\beta}_1$ observable? Yes or no, and explain. What about e_i ?