

1. Which of the following scenarios can be classified under explanatory modeling?

☐ Observational studies that seek to describe how changes in some variables are caused by changes in others.

☐ Observational studies that seek to produce a "best guess" of what the response will be at a future set of predictors.

☐ Data analysis problems where the goal is to learn about regression parameters that have a physical meaning.

☐ Data analysis problems where researchers pay special attention to the interpretation of parameter values.

☒ Experimental studies where regression parameters are given a causal interpretation.

2. It can be shown that prediction and explanation are equivalent.

☐ True

☐ False

3. The best point prediction for the response at a new set of predictors is equivalent to the best point estimate for the mean response at those same predictors.

☐ True

☐ False

4. All else equal (e.g., the confidence level, sample size), a prediction interval for the response at a new set of predictors is equivalent to a confidence interval for the mean response at those same predictors.

☐ True

☐ False

5. The mean squared prediction error may be used for comparing linear regression models.

☐ True

☐ False

6. Consider a dataset that has been split into a training set and a test set. The mean squared prediction error is computed using the:

☐ training set data.

☐ test set data.

7. All else equal (e.g., the confidence level, sample size), a prediction interval for the response at a new set of predictors will always be two units wider than a confidence interval for the mean response at those same predictors.

☐ True

☐ False

8. Prediction intervals account for both the variation in estimating the population mean (regression line/surface), and also the random variation of the individual response values.

☐ True

☐ False

9. A 90% prediction interval can be interpreted as follows: for a fixed set of predictors, if one were to resample the response and refit a model many times, and use each of those models to make a prediction interval at a fixed set of new predictors, 90% of those intervals would contain the true value of the new response.

☐ True

☐ False

10. In cities with 100,000 to 500,000 residents, the sales ( $y$ ) of a product are recorded, along with TV ( $x_1$ ), radio ( $x_2$ ), and newspaper advertising ( $x_3$ ). A linear regression model is fit with  $y$  as the response and the remaining variables as predictors. Diagnostics are performed and there is no evidence of any violations of assumptions.

Suppose that a market opens for a different product in a city with 200,000 people. The market has advertising characteristics  $x_1^*$ ,  $x_2^*$ , and  $x_3^*$ . Predicting sales of the new product in the new market,  $y^*$ , using  $x_1^*$ ,  $x_2^*$ , and  $x_3^*$  and the model fit in the previous paragraph is an example of a qualitative extrapolation.

☐ True

☐ False

3 / 3 points

3 / 3 points

3 / 3 points

1 / 1 point

3 / 3 points