Quiz -2 for Lecture 3

6 out of 6 correct

1. What is the measure of central tendency that is most influenced by outliers?

- [x] Mean
- [ ] Median
- [ ] Mode
- [ ] Standard Deviation

Explanation: The mean is calculated by summing all the values and dividing by the number of values. Outliers, being extreme values, can significantly affect the sum, resulting in a larger impact on the mean compared to other measures of central tendency.

2. What is the p-value in hypothesis testing?

- [ ] The probability of making a Type I error
- [ ] The probability of making a Type II error
- [ ] The level of significance for the test
- [x] The probability of observing the data given that the null hypothesis is true

Explanation: The p-value represents the likelihood of obtaining the observed data or more extreme data under the assumption that the null hypothesis is true. It helps in determining the strength of evidence against the null hypothesis.

3. What is the standard deviation?

- [ ] The average value in a dataset
- [x] The spread of values around the mean
- [ ] The middle value in a dataset
- [ ] The most frequently occurring value in a dataset

Explanation: The standard deviation measures the dispersion or spread of values in a dataset. It quantifies the average distance between each data point and the mean. A larger standard deviation indicates greater variability in the data.

4. What is the difference between correlation and causation?

- [ ] Correlation indicates a cause-and-effect relationship between variables.
- [ ] Causation indicates a strong association between variables.
- [x] Correlation measures the strength and direction of the linear relationship between variables.
- [ ] Causation measures the statistical significance of the relationship between variables

Explanation: Correlation measures the degree to which two variables are related and how they change together. It does not imply a cause-and-effect relationship. Causation, on the other hand, refers to a cause-and-effect relationship where one variable directly influences the other.

5. What is the formula for calculating the z-score of a data point?

- [x] (Data point - Mean) / Standard Deviation
- [ ] (Data point - Median) / Interquartile Range
- [ ] (Data point - Mode) / Range
- [ ] (Data point - Mean) / Range

Explanation: The z-score measures how many standard deviations a data point is away from the mean. It is calculated by subtracting the mean from the data point and dividing the result by the standard deviation.

6. What is the p-value threshold commonly used for statistical significance?

- [x] 0.05
- [ ] 0.10
- [ ] 0.01
- [ ] 0.001

Explanation: The p-value threshold of 0.05 (or 5%) is commonly used as a criterion for statistical significance. If the p-value is less than 0.05, the result is considered statistically significant, indicating strong evidence against the null hypothesis.

Submit