MCQ Set

1. **Which of the following options is true for a vector?**  
   a) A vector has only a magnitude.  
   b) A vector has a shape and weight.  
   c) A vector has a magnitude and direction.  
   d) A vector has only direction.  
   **Answer:** c) A vector has a magnitude and direction.
2. **You have a sample of size 20 from a population with unknown mean and standard deviation. You measured that the sample mean ( \bar{X} = 50 ) and the sample standard deviation is ( s = 10 ). A confidence interval of 95% is given by:**  
   • (48.95, 51.05)  
   • (45.32, 54.68)  
   • (45.2, 54.8)  
   • (48.9, 51.1)  
   **Answer:** (45.32, 54.68)
3. **To calculate one confidence interval for the mean of a population with unknown distribution, what assumptions must we ensure (check all that apply)?**  
   • The sample is a not random sample.  
   • We can only have a confidence interval if the population is known as having a Normal distribution.  
   • If the distribution is not Normal, the sample size must be big enough (usually over 30).  
   • The sample must have mean 0 and standard deviation 1.  
   **Answer:** If the distribution is not Normal, the sample size must be big enough (usually over 30).
4. **In statistical hypothesis testing, which of the following statements correctly defines Type I and Type II errors?**  
   • Type I error occurs when we reject a null hypothesis that is true, while Type II error occurs when we do not reject a null hypothesis that is false.  
   • Type I error occurs when we do not reject a null hypothesis that is true, while Type II error occurs when we reject a null hypothesis that is false.  
   • Type I error occurs when we reject a null hypothesis that is false, while Type II error occurs when we do not reject a null hypothesis that is true.  
   • Type I error occurs when we do not reject a null hypothesis that is false, while Type II error occurs when we reject a null hypothesis that is true.  
   **Answer:** Type I error occurs when we reject a null hypothesis that is true, while Type II error occurs when we do not reject a null hypothesis that is false.
5. **Which visualization technique is best for displaying the distribution of a single variable?**  
   • A) Line Chart  
   • B) Scatter Plot  
   • C) Histogram  
   • D) Pie Chart  
   **Answer:** C) Histogram
6. **Which of the following would describe a dataset with a right (positive) skew?**  
   • A) Mean < Median  
   • B) Mean > Median  
   • C) Mean = Median  
   • D) Mean > Mode  
   **Answer:** B) Mean > Median
7. **What does the Central Limit Theorem state?**  
   • A) The distribution of the sample mean approaches a normal distribution as the sample size increases  
   • B) The population mean will always be normally distributed  
   • C) The sample data will always follow a binomial distribution  
   • D) The variance of the sample is always smaller than the population variance  
   **Answer:** A) The distribution of the sample mean approaches a normal distribution as the sample size increases.

Characterizing Data (Descriptive Statistics)

1. **What is data in the context of analytics?**  
   • A) Information collected but not organized  
   • B) Raw facts and figures  
   • C) A set of unchanging facts  
   • D) Information presented in reports  
   **Answer:** B) Raw facts and figures.
2. **Which chart is best for showing trends over time?**  
   • A) Bar Chart  
   • B) Pie Chart  
   • C) Line Chart  
   • D) Scatter Plot  
   **Answer:** C) Line Chart.
3. **What is the primary purpose of a scatter chart?**  
   • A) To compare parts of a whole  
   • B) To show frequency distributions  
   • C) To observe relationships between two variables  
   • D) To show data trends over time  
   **Answer:** C) To observe relationships between two variables.
4. **What is the main difference between a bar graph and a histogram?**  
   • A) Bar graphs show continuous data, histograms show categorical data  
   • B) Histograms display continuous data, bar graphs display categorical data  
   • C) Bar graphs show relationships, histograms show time-based data  
   • D) There is no difference  
   **Answer:** B) Histograms display continuous data, bar graphs display categorical data.
5. **Which chart is most appropriate for showing the proportion of categories in a dataset?**  
   • A) Line Chart  
   • B) Pie Chart  
   • C) Scatter Plot  
   • D) Histogram  
   **Answer:** B) Pie Chart.
6. **Which measure of central tendency is the middle value in an ordered dataset?**  
   • A) Mean  
   • B) Mode  
   • C) Median  
   • D) Standard Deviation  
   **Answer:** C) Median.
7. **What does the standard deviation measure?**  
   • A) The average of the dataset  
   • B) The central value of the dataset  
   • C) The amount of variation or dispersion in a dataset  
   • D) The most frequently occurring value  
   **Answer:** C) The amount of variation or dispersion in a dataset.
8. **Which of the following is a measure of association between two variables?**  
   • A) Correlation  
   • B) Skewness  
   • C) Variance  
   • D) Mean  
   **Answer:** A) Correlation.
9. **What does a negative skew in a dataset indicate?**  
   • A) The tail is on the left side of the distribution  
   • B) The tail is on the right side of the distribution  
   • C) The data is symmetrically distributed  
   • D) The data has no central tendency  
   **Answer:** A) The tail is on the left side of the distribution.

Probability Basics

1. **What is the probability of an event in a uniform distribution?**  
   • A) Depends on the event's frequency  
   • B) The same for all events  
   • C) Higher for frequent events  
   • D) Lower for infrequent events  
   **Answer:** B) The same for all events.
2. **In a binomial distribution, what does "n" represent?**  
   • A) The number of successes  
   • B) The probability of failure  
   • C) The number of trials  
   • D) The expected value  
   **Answer:** C) The number of trials.
3. **What is the mean of a standard normal distribution?**  
   • A) 0  
   • B) 1  
   • C) 2  
   • D) It varies depending on the dataset  
   **Answer:** A) 0.

Predicting from Data (Inferential Statistics)

1. **Which sampling technique ensures every individual in the population has an equal chance of being selected?**  
   • A) Stratified Sampling  
   • B) Random Sampling  
   • C) Cluster Sampling  
   • D) Systematic Sampling  
   **Answer:** B) Random Sampling.
2. **What is the null hypothesis in hypothesis testing?**  
   • A) A statement that no relationship or effect exists  
   • B) A statement that an effect exists  
   • C) A claim to be proven  
   • D) A statement predicting a positive outcome  
   **Answer:** A) A statement that no relationship or effect exists.
3. **According to the Central Limit Theorem, what happens as the sample size increases?**  
   • A) The mean of the sample becomes closer to the population mean  
   • B) The sample variance increases  
   • C) The data becomes more skewed  
   • D) The distribution remains unchanged  
   **Answer:** A) The mean of the sample becomes closer to the population mean.
4. **Which error occurs when the null hypothesis is wrongly rejected?**  
   • A) Type I Error  
   • B) Type II Error  
   • C) Sampling Error  
   • D) Confidence Interval Error  
   **Answer:** A) Type I Error.

Crafting Your Data

1. **What is the first step in data collection?**  
   • A) Cleaning the data  
   • B) Defining the objectives  
   • C) Preparing a report  
   • D) Visualizing the data  
   **Answer:** B) Defining the objectives.
2. **What does data cleaning involve?**  
   • A) Collecting raw data  
   • B) Transforming raw data into a structured form  
   • C) Correcting or removing errors in the data  
   • D) Analyzing the processed data  
   **Answer:** C) Correcting or removing errors in the data.
3. **Which of the following is an example of a data preparation step?**  
   • A) Data storage  
   • B) Data wrangling  
   • C) Model deployment  
   • D) Data archiving  
   **Answer:** B)