Statistics_Worksheet_6

QUES 1: Which of the following can be considered as random variable?

Answer. D) All of the mentioned.

QUES 2: Which of the following random variable that take on only a countable number of possibilities?

Answer. A) Discrete.

QUES 3. Which of the following function is associated with a continuous random variable?

Answer. A) PDF.

QUES 4. The expected value or _____ of a random variable is the center of its distribution.

Answer. C) Mean.

QUES 5. Which of the following of a random variable is not a measure of spread?

Answer. A) Variance.

QUES 6. The _____ of the Chi-squared distribution is twice the degrees of freedom.

Answer. B) Standard Deviation.

QUES 7. The beta distribution is the default prior for parameters between _____

Answer. C) 0 and 1.

QUES 8. Which of the following tool is used for constructing confidence intervals and calculating standard errors for difficult statistics?

Answer. B) Bootstrap.

QUES 9. Data that summarize all observations in a category are called _____ data.

Answer. B) Summarized.

QUES 10. What is the difference between a boxplot and histogram?

Answer. **Boxplots**: Boxplots may also depict values that are far outside of the normal range of responses (referred to as outliers). They represents outliers. It typically provides the median, 25th and 75th percentile, min/max that is not an outlier and explicitly separates the points that are considered outliers.

Histogram: A histogram is a graphical representation of the spread of data points. Histogram is a type of bar chart to visualize the distribution of a dataset.

QUES 11. How to select metrics?

Answer. Key steps for evaluating Metrics:

1. **Classification.** This algorithm will predict data type from defined data arrays. For example, it may respond with yes/no/not sure.
2. **Regression.** The algorithm will predict some values. For example, weather forecast for tomorrow.
3. **Ranking.** The model will predict an order of items.

QUES 12. How do you assess the statistical significance of an insight?

Answer. We can assess the statistical significance of an insight by the following steps:

1. State the Research Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level)
4. Select and compute the test for statistical significance.
5. Interpret the results.

QUES 13. Give examples of data that does not have a Gaussian distribution, nor log-normal?

Answer. Examples of data that does not have a Gaussian distribution, nor log-normal are:

Exponential distributions do not have a log-normal distribution or a Gaussian distribution.

Any type of data that is categorical will not have these distributions as well.

Example: Duration of a phone car, time until the next earthquake, etc.

QUES 14. Give an example where the median is a better measure than the mean.

Answer. **MEAN:** The mean is used for normal distributions. The mean is not a robust tool since it is largely influenced by outliers.

MEDIAN: The median is generally used for skewed distributions. The median is better suited for skewed distributions to derive at central tendency since it is much more robust and sensible.

Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed.

QUES 15. What is the Likelihood?

Answer. The likelihood is the probability that a particular outcome is observed when the true value of the parameter is, equivalent to the probability mass on, it is not a probability density over the parameter. The likelihood, should not be confused with, which is the posterior probability of given the data.

……
……