STATISTICS WORKSHEET- 6

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

1. Which of the following can be considered as random variable?
   a) The outcome from the roll of a die
   b) The outcome of flip of a coin
   c) The outcome of exam
   d) All of the mentioned

Answer: All of the mentioned

2. Which of the following random variable that take on only a countable number of possibilities?
   a) Discrete
   b) Non Discrete
   c) Continuous
   d) All of the mentioned

Answer: Discrete

3. Which of the following function is associated with a continuous random variable?
   a) pdf
   b) pmv
   c) pmf
   d) all of the mentioned

Answer: pdf

4. The expected value or_____of a random variable is the center of its distribution.
   a) mode
   b) median
   c) mean
   d) bayesian inference

Answer: median

5. Which of the following of a random variable is not a measure of spread?
   a) variance
   b) standard deviation
   c) empirical mean
   d) all of the mentioned

Answer: empirical mean

6. The_____of the Chi-squared distribution is twice the degrees of freedom.
   a) variance
   b) standard deviation
   c) mode
   d) none of the mentioned

Answer: standard deviation

7. The beta distribution is the default prior for parameters between _____
   a) 0 and 10
   b) 1 and 2
   c) 0 and 1
   d) None of the mentioned

Answer: 0 and 1

8. Which of the following tool is used for constructing confidence intervals and calculating standard errors for difficult statistics?
   a) baggyer
   b) bootstrap
   c) jacknife
   d) none of the mentioned

Answer: b) bootstrap

9. Data that summarize all observations in a category are called_____data.
   a) frequency
   b) summarized
   c) raw
   d) none of the mentioned

Answer: b) summarized

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.

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

Answer: Well Box plot and Histogram can be used use Seaborn . where Box plot is Used using .boxplot and Histogram is used using .countplot. Box plot - gives the quartiles and indicate the median data to compare easily

Histogram - gives only the count

11. How to select metrics?

Answer: The key point is to choose metrics that clearly indicate where you are now in relation to your goals. Good metrics can be improved. Good metrics measure progress, which means there needs to be room for improvement. For example, reducing churn by 0.8% or increasing your activation rate by 3%.

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

1. Answer: 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.

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

Answer: The simplest example is the distribution of numbers that show up on the top of a fair die after a large number of throws. Each number from 1 to 6 will occur with approximately equal frequency. Increasing the number of throws will not tend to produce a bell-shaped histogram, in fact the fractional occurrence will approach a constant 1/6 over the possible numbers.

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

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

15. What is the Likelihood?

Answer: the chance that something will happen : probability. There's very little likelihood of that happening