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

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

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

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

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

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

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

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

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

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 - In that they both aid in visualising and describing numerical data, histograms and box plots are quite similar to one another. Box plots are better at comparing various data sets than histograms, despite the fact that histograms are better at identifying the underlying distribution of the data. This is because box plots are less detailed and take up less space.

11. How to select metrics?
Answer - For the growth and goals of your business, good metrics are crucial. Your primary goal should always be directly related to your key performance indicators. Excellent metrics may be enhanced. Good metrics track development, therefore there must be space for improvement.
Effective metrics motivate action.

12. How do you assess the statistical significance of an insight?
Answer - These problems are tackled using two different sorts of tools: measures of association for the second and tests for statistical significance for the first. To answer the question: What is the likelihood that what we believe to be a link between two variables is actually just a random occurrence?, tests for statistical significance are utilised. Would the link between these two variables be the same in all of the samples we choose from the same population? Would a population census reveal that this link holds true for the population from whom the sample was taken? Or is our discovery really the result of chance?

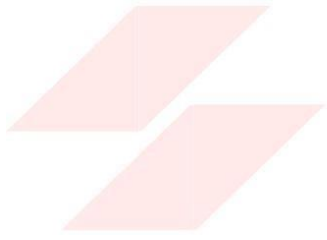13. Give examples of data that doesnot have a Gaussian distribution, nor log-normal.
Answer - There is no Gaussian or log-normal distribution for exponential distributions. In actuality, categorical data of any kind won't have these distributions either. Examples include the length of a phone call, when the next earthquake will occur, etc.

14. Give an example where the median is a better measure than the mean.
Answer - As income has a skewed distribution, it is a prime example of a situation in which the median should be used instead of the mean. According to the median, 50% of all salaries are below 27581 and 50% are above it. The mean overestimates the range of household incomes for these data.

15. What is the Likelihood?
Answer - The likelihood in frequentist inference is a number corresponding to the likelihood of selecting a sample with the observed value x0 from a population with a given value of. The result of a likelihood function, or likelihood, indicates the plausibility of a certain parameter value after seeing a specific occurrence under a given statistical model (the null hypothesis in hypothesis testing).