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?

Ans=d) All of the mentioned

1. Which of the following random variable that take on only a countable number of possibilities?

Ans=c) Continuous

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

Ans=a) pdf

1. The expected value or \_\_\_\_\_\_\_ of a random variable is the center of its distribution

Ans=c) mean

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

Ans=a) variance

1. The \_\_\_\_\_\_\_\_\_ of the Chi-squared distribution is twice the degrees of freedom.

Ans=a) variance

1. The beta distribution is the default prior for parameters between \_\_\_\_\_\_\_\_\_\_\_\_

Ans= b) 1 and 2

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

Ans=b) bootstrap

1. Data that summarize all observations in a category are called \_\_\_\_\_\_\_\_\_\_ data

Ans=b) summarized

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

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

Ans=Histograms and box plots are graphical representations for the frequency of numeric data values.  Histograms are preferred to determine the underlying probability distribution of a data. Box plots on the other hand are more useful when comparing between several data sets.

1. How to select metrics?

Ans=Good metrics are important to your company growth and objectives. Your key metrics should always be closely tied to your primary objective. ...

Good metrics can be improved. Good metrics measure progress, which means there needs to be room for improvement.

Good metrics inspire action.

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

Ans=Statistical significance can be accessed using hypothesis testing: – Stating a null hypothesis which is usually the opposite of what we wish to test (classifiers A and B perform equivalently, Treatment A is equal of treatment B).

Start by looking at the left side of your degrees of freedom and find your variance. Then, go upward to see the p-values. Compare the p-value to the significance level or rather, the alpha. Remember that a p-value less than 0.05 is considered statistically significant.

1. Give examples of data that doesnot have a Gaussian distribution, nor log-normal.

Ans=There are many data types that follow a non-normal distribution by nature. Examples include: Weibull distribution, found with life data such as survival times of a product. Log-normal distribution, found with length data such as heights.

Tools for Normally Distributed Data: Equivalent.

Paired t-test: One-sample sign test.

T-test: Mann-Whitney test; Mood’s median test.

ANOVA: Mood’s median test; Kruskal-Wallis test.

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

Ans=In this case, analysts tend to use the mean because it includes all of the data in the calculations. However, if you have a skewed distribution, the median is often the best measure of central tendency. When you have ordinal data, the median or mode is usually the best choice.

1. What is the Likelihood?

Ans=In [statistics](https://en.wikipedia.org/wiki/Statistics), the likelihood function (often simply called the likelihood) measures the [goodness of fit](https://en.wikipedia.org/wiki/Goodness_of_fit) of a [statistical model](https://en.wikipedia.org/wiki/Statistical_model) to a [sample of data](https://en.wikipedia.org/wiki/Sample_(statistics)) for given values of the unknown [parameters](https://en.wikipedia.org/wiki/Statistical_parameter). It is formed from the [joint probability distribution](https://en.wikipedia.org/wiki/Joint_probability_distribution) of the sample, but viewed and used as a function of the parameters only, thus treating the [random variables](https://en.wikipedia.org/wiki/Random_variable) as fixed at the observed values.