**STATISTICS WORKSHEET- 6**

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

**1. Which of the following can be considered a random variable?**

a) The outcome from the roll of a die



b) The outcome of the flip of a coin



c) The outcome of the exam

d) All of the mentioned

**2. Which of the following random variable 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 random variables 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) baggier

b) bootstrap



c) jackknife

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

**Q10 and Q15 are subjective answer-type questions, Answer them in your own words briefly.**

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

Histograms are a special kind of bar graph that shows a bar for a range of data values instead of a single value. A box plot is a data display that draws a box over a number line to show the interquartile range of the data.

**11. How to select metrics?**

The four key metrics are Deployment Frequency (the frequency at which new releases go to production), Lead Time for Changes (the time until a commit goes to production), Mean Time to Restore (the time it takes to resolve a service impairment in production) and Change Failure Rate.

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

Statistical significance is often calculated with statistical hypothesis testing, which tests the validity of a hypothesis by figuring out the probability that your results have happened by chance.

**13. Give examples of data that does not have a Gaussian distribution, or log-normal.**

Any distribution of money or value will be non-Gaussian. Examples: distributions of income; distributions of house prices; distributions of bets placed on a sporting event.

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

The median better represents the central tendency for the skewed distribution. 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?**

The likelihood function (often simply called the likelihood) represents the probability of random variable realizations conditional on particular values of the statistical parameters.