

## **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

**Ans: d)**

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

**Ans: a)**

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

- a) pdf
- b) pmv
- c) pmf
- d) all of the mentioned

**Ans: a)**

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

- a) mode
- b) median
- c) mean
- d) bayesian inference

**Ans: c)**

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

**Ans: a)**

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

**Ans: a)**

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

**Ans: c)**

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

- a) baggyer
- b) bootstrap
- c) jackknife
- d) none of the mentioned

**Ans:b)**

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

- a) frequency
- b) summarized
- c) raw
- d) none of the mentioned

**Ans: b)**

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

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

Ans: 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. The 'whiskers' of a box plot show the least and greatest values in the data set.

A box and whisker plot is defined as a graphical method of displaying variation in a set of data. In most cases, a histogram analysis provides a sufficient display, but a box and whisker plot can provide additional detail while allowing multiple sets of data to be displayed in the same graph.

11. How to select metrics?

Ans- To select metrics-

- Define your primary objective
- Determine cause and effect
- Create relevant activities
- Evaluate periodically

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

Ans- To assess statistical significance, you would use hypothesis testing. The null hypothesis and alternate hypothesis would be stated first. Second, you'd calculate the p-value, which is the likelihood of getting the test's observed findings if the null hypothesis is true. Finally, you would select the threshold of significance (alpha) and reject the null hypothesis if the p-value is smaller than the alpha — in other words, the result is statistically significant.

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

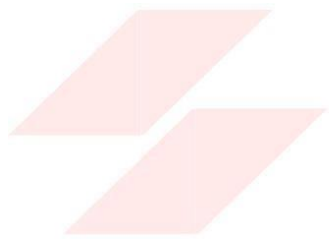
Ans- Exponential distributions do not have a log-normal distribution or a Gaussian distribution. In fact, any type of data that is categorical will not have these distributions as well. Example: Duration of a phone call, time until the next earthquake

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

Ans- Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed. The median indicates that half of all incomes fall below 27581, and half are above it. The mean is used for normal distributions. The median is generally used for skewed distributions. The mean is not a robust tool since it is largely influenced by outliers. The median is better suited for skewed distributions to derive at central tendency since it is much more robust and sensible.

15. What is the Likelihood?

Likelihood is a strange concept, in that it is not a probability, but it is proportional to a probability. The likelihood of a hypothesis (H) given some data (D) is proportional to the probability of obtaining D given that H is true, multiplied by an arbitrary positive constant K. In other words,  $L(H) = K \times P(D|H)$ .



# FLIP ROBO