Statistics Worksheet

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)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

Ans: a)Discrete

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

a) pdf

b) pmv

c) pmf

d) all of the mentioned

Ans : a)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

Ans:c)Mean

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: c)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

Ans: b)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

Ans: c)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

Ans: 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

Ans: 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?

Ans:

Boxplots and histograms are graphical representation for frequency of numerical data values.Histograms are preferred to determine the underlying probability distribution of a data.Boxplot are more useful when comparing between several datasets.

11. How to select metrics?

Ans:

Key metrics should always be closely tied to your primary objective.Good metrics can be improved.Good metrics measure progress which means there need to be room for improvement.Good metrics inspire action.

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 doesnot have a Gaussian distribution, norlog-normal.

Ans:

- 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.

Ans:

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?

Ans;
Likelihood measures the goodness of fit of a statistical model to a sample of data for given values of unknown parameters. But in both frequentist and Bayesian statistics likelihood function plays a fundamental role
It indicates how likely a particular population is to produce an observed sample.