

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) 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 random variable?

- 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- a) variance

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) variance

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) jackknife
- d) none of the mentioned

ANS- b) bootstrap

WORKSHEET

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

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 indicate the whole frequency distribution of a variable, whereas the boxplot summarises its most prominent features. These features include median and spread as well as the extent and nature of departures from symmetry, and the possible presence of observations having extreme values (outliers).

11. How to select metrics?

ANS-Choosing the right metrics

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.

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

ANS-Steps in Testing for Statistical Significance

State the Research Hypothesis.

State the Null Hypothesis.

Select a probability of error level (alpha level)

Select and compute the test for statistical significance.

Interpret the results.

13. Give examples of data that doesnot 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 car, 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. For these data, the mean overestimates where most household incomes fall.

15. What is the Likelihood

ANS-The likelihood is the probability that a particular outcome is observed when the true value of the parameter is , equivalent to the probability mass on ; it is not a probability density over the parameter . The likelihood, , should not be confused with , which is the posterior probability of given the data .