Name: Priti Chauhan
Internship no: 32
Batch no.: 1845

MACHINE LEARNING -ASSIGNMENT - 6

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-C

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) jacknife
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

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-create a null hypothesis
Create an alternate hypothesis
Determine the significance level
Decide on type of test you need to use
Perform power analysis to find out your sample size
Calculate std deviation
Use std error formula

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

14. Give an example where the median is a better measure than the mean.
Ans-income level researches.
Few millionares can make it look like socio economic status of your sample is higher than iit really is.

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