WORKSHEET

STATISTICS WORKSHEET-1

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

1. Bernoulli random variables take (only) the values 1 and 0.

a) True

b) False

Ans:A

2. Which of the following theorem states that the distribution of averages of iid variables, properly normalized, becomes that of a standard normal as the sample size increases?

a) Central Limit Theorem

b) Central Mean Theorem

c) Centroid Limit Theorem

d) All of the mentioned

Ans:A

3. Which of the following is incorrect with respect to use of Poisson distribution?

a) Modeling event/time data

b) Modeling bounded count data

c) Modeling contingency tables

d) All of the mentioned

Ans:B

4. Point out the correct statement.

a) The exponent of a normally distributed random variables follows what is called the log- normal distribution

b) Sums of normally distributed random variables are again normally distributed even if the variables are dependent

c) The square of a standard normal random variable follows what is called chi-squared distribution

d) All of the mentioned

Ans:D

5. \_\_\_\_\_\_ random variables are used to model rates.

a) Empirical

b) Binomial

c) Poisson

d) All of the mentioned

Ans:C

6. 10. Usually replacing the standard error by its estimated value does change the CLT.

a) True

b) False

Ans:B

7. 1. Which of the following testing is concerned with making decisions using data?

a) Probability

b) Hypothesis

c) Causal

d) None of the mentioned

Ans:B

8. 4. Normalized data are centered at\_\_\_\_\_\_and have units equal to standard deviations of the original data.

a) 0

b) 5

c) 1

d) 10

Ans:A

9. Which of the following statement is incorrect with respect to outliers?

a) Outliers can have varying degrees of influence

b) Outliers can be the result of spurious or real processes

c) Outliers cannot conform to the regression relationship

d) None of the mentioned

Ans:C

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.

10. What do you understand by the term Normal Distribution?

Ans: Normal distribution is also known as Gaussian distribution. It is a continuous probability distribution that is symmetrical around its mean, most of the observations cluster around the central peak, and the probabilities for values further away from the mean taper off equally in both directions. In graph form, normal distribution will appear as a bell curve

11. How do you handle missing data? What imputation techniques do you recommend?

Ans: A common technique is to use the mean or median of the non missing observation. This can be useful in cases where the number of missing observations is low. However, for large number of missing values, using mean or median can result in loss of variations in data and it is better to use imputation.

12. What is A/B testing?

Ans: A/B testing is also known as split testing or bucket testing. A/B testing is a user experience research methodology.This is a method of comparing two versions of a webpage or app against each other to determine which one performs better.

13. Is mean imputation of missing data acceptable practice?

Ans: Yes. It is true. Imputing the mean preserves the mean of the observed data , so if the data are missing completely at random, the estimate of the mean remains. And most of the interested in the relationship among variables, mean imputation is not a good solution.

14. What is linear regression in statistics?

Ans: linear regression is a linear approach from modelling the relationship between a scaler response and one or more explanatory variables. The case of one explanatory variable is called simple linear regression and for more than one the process is called multiple linear regression.

15. What are the various branches of statistics?

Ans: Statistic plays main and important role in the field of research, it helps in collection, analysis and presentation of data. And main branches of statistics that is descriptive and inferential statistics.