STATISTICS WORKSHEET

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

a) True

b) False

Ans: B.

7. 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. 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

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

Ans: Normal distribution is a type of continuous probability distribution for a real valued random variable. It is also called as bell curve.

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

Ans: In Statistics, missing data are the missing values which occurs when no data value is stored for the variable in an observation. When dealing with missing data we can use two primary methods: Imputation and removal of data.

The various imputation techniques are: Regression imputation; predictive mean matching; hot deck imputation.

12. What is A/B testing?

Ans: A/B testing is a basic randomized control experiment. It is a way to compare the two versions of a variable to find out which performs better in a controlled environment.

13. Is mean imputation of missing data acceptable practice?

Ans: Yes, it is an acceptable practice. Imputing the mean preserves the mean of the observed data.

14. What is linear regression in statistics?

Ans: In statistics, linear regression is the linear approach for modelling the relationship between a scalar response and one or more explanatory variables.

15. What are the various branches of statistics?

Ans: The two main branches of statistics are:

- Descriptive statistics
- Inferential statistics