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

a) True

b) False

ANSWER: 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

ANSWER: 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

ANSWER: 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

ANSWER: d

5. _____ random variables are used to model rates.

a) Empirical

b) Binomial

c) Poisson

d) All of the mentioned

ANSWER: c

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

a) True

b) False

ANSWER: 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

ANSWER: 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

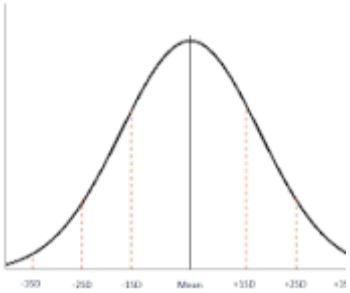
ANSWER: a

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

ANSWER: c

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

- Normal distribution , which is also called as Guassian distribiution is a type of probability distribution that is symmetric that shows data near the mean are more frequent in occurrence than the data far away from the mean.
  Standard normal distribution is usually denoted with Z. Normal distribution appears as a Bell curve in graphical form.



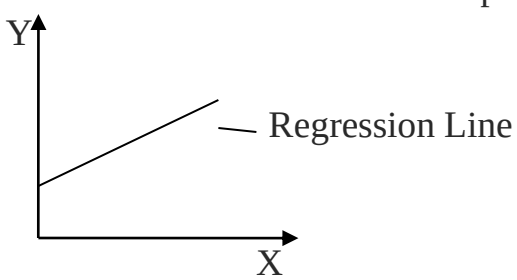
11. How do you handle missing data? What imputation techniques do you recommend?
A) We can handle missing data normally in two ways. one of the easiest way is just ignore the missing values.But its not a good idea which leads to system performance.Another one, Imputation technique.Imputation method develops guesses for missing data which is more reasonable.I recommend Mean,Median,Rounded Mean imputation techniques or by using predictive models for handling missing data.

12. What is A/B testing?
A) A/B Testing is used to optimize web marketing strategies.It allows decision makers to choose the best design for the website by analysing the results obtained with the possible alternatives A and B. A/B Testing is a way to compare multiple variants of the same variable. It includes the application of statistical hypothesis testing.

13. Is mean imputation of missing data acceptable practice?
A) Mean imputation considered as terrible practice due to some drawbacks.

- It will reduce the variance and standard deviation
- Ignores the distribution and correlation of the data
- Creation of unrealistic values

14. What is linear regression in statistics?
A)Linear regression is used to predict the value of a variable based on the value of another variable.The variable you want to predict is called dependent variable where as variable you are using to predict the other variables value is called independent variable.



Here, X-Independent Variables
Y-Dependent Variables

15. What are the various branches of statistics?
A) The two main branches of statistics are:

- Descriptive Statistics
- Inferential Statistics

Descriptive Statistics deals with presentation and collection of data where as, inferential statistics involves drawing the right conclusions performed using descriptive statistics. Both statistics methods are go hand in hand and one cannot exist without other.