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, properlynormalized, 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: c)

4. Point out the correct statement.
   a) The exponent of a normally distributed random variables follows what is called the log- normaldistribution
   b) Sums of normally distributed random variables are again normally distributed even if the variablesare 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 theoriginal 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, where mean= median=mode , symmetry at the center and 50% of the values less than mean and 50% of the values are greater than mean.

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

Ans. Describing all data and finding missing values (null) we replace the missing values by filling the place with mean of that column. I would recommend KNN imputation technique.

12. What is A/B testing?

Ans. A/B testing is controlled experiment for making decisions, where entire data is experimented into A part and B part.
In A part there will be no significant changes and in B part we make certain changes.
On the basis of collection of feedback, we create hypothesis which part is performing better.

13. Is mean imputation of missing data acceptable practice?

Ans. No as mean imputation reduces the variance of the imputed variables, reduces standard errors resulting in failing most hypothesis tests, does not preserve relationships such as correlation. So it's better to avoid mean imputation practice.

14. What is linear regression in statistics?

Ans. Linear regression is a type of predictive analysis where it shows a relationship between A dependent variable and one or more independent variables.
Mathematically represented as $y = mx + c$.

15. What are the various branches of statistics?

Ans. The three branches of statistics are data collection, descriptive statistics and inferential statistics.
Descriptive statistics has two branches:
1. Measure of central tendency – mean, median, mode.
2. Measure of dispersion – variance, standard deviation.

Inferential statistics is nothing sampling data and inferring the result to describe entire population.