Statistic

1. Bernoulli random variables take (only) the values 1 and 0.
Ans: a) True

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?
Ans: a) Central Limit Theorem

3. Which of the following is incorrect with respect to use of Poisson distribution?
Ans: b) Modeling bounded count data

4. Point out the correct statement.
Dont Know

5. _____ random variables are used to model rates.
c) Poisson

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

7. 1. Which of the following testing is concerned with making decisions using data?
Ans: b) Hypothesis

8. 4. Normalized data are centered at_____and have units equal to standard deviations of the original data.
Ans: a) 0

9. Which of the following statement is incorrect with respect to outliers?
Ans: c) Outliers cannot conform to the regression relationship

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: A normal distribution is one in which the values are evenly distributed both above and below the mean.

11. How do you handle missing data? What imputation techniques do you recommend?
Ans: Will handle missing data by Imputing the Missing Value

1. Replacing With Arbitrary Value
2. Replacing With Mode
3. Replacing With Median
4. Replacing with previous value – Forward fill
5. Replacing with next value – Backward fill
6. Interpolation
7. Impute the Most Frequent Value

I recommend Average imputation and common-point 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: mean imputation is bad imputation. It does improve power, but your results will be so biased, the improved power won't help much. Sure, your results might be significant, but they're the wrong results.

14. What is linear regression in statistics?
Ans: linear regression is a regression model that estimates the relationship between one independent variable and one dependent variable using a straight line. Both variables should be quantitative.

15. What are the various branches of statistics?
Ans: There are three real branches of statistics: data collection, descriptive statistics and inferential statistics.