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

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?

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

4. Point out the correct statement.

5. _____ random variables are used to model rates.

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

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

8. Normalized data are centered at_____and have units equal to standard deviations of the original data.

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

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

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

80%, and the feature is not meaningful, that time we can drop the entire column. I recommend K_Nearest Neighbor Imputation

12. What is A/B testing?

Ans: It is also known as split testing. A/B testing is the process of comparing two variations of a page element, usually by testing users' response to variant A vs. variant B and concluding which of the two variants is more effective.

13. Is mean imputation of missing data acceptable practice?

Ans: No, Mean imputation reduces the variance of the imputed variables. Mean imputation shrinks standard errors, which invalidates most hypothesis tests and the calculation of confidence interval. Mean imputation does not preserve relationships between variables such as correlations.

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

15. What are the various branches of statistics?

Ans: Data collection, descriptive statistics and inferential statistics