

Statistics Worksheet-1

1. A. True
 2. C. Central limit theorem
 3. B. Modeling bounded count data
 4. D. All of the above mentioned
 5. C. Poisson
 6. b. False
 7. b. Hypothesis
 8. a. 0
 9. c. Outliers can't confirm to regression relationship
 10. Normal distribution is the most common distribution function for the independent and randomly generated variables. The Graph of normal distribution is characterized by two parameters :
 - a. mean , which is the maximum of the graph and about which the graph is always symmetric
 - b. the standard deviation, which determines the amount of dispersion away from the mean
 11. There are multiple ways to handle missing data:
 - a. if the features are **less important** and **more in number** we can simply delete the entire row with missing values by using '**dropna**' attribute
 - b. if the features are important, **quantitative** and we can't afford to delete them we can fill the missing values with the **mean** of the rest of the values .
 - c. if the features are important, **qualitative** and we can't afford to delete them we'll fill them with the most frequent value observed in rest of the cells by using **mode** method
- Different imputation techniques depending on the type of missing data are:
- a. Mean
 - b. Median
 - c. Mode
12. A/B testing is a user experience research methodology. A/B tests consist of a randomized experiment that usually involves two variants, although the concept can be also extended to multiple variants of the same variable.
 13. Mean imputation of missing data is acceptable only in the cases where data to be filled is quantitative data ,if it is qualitative data we can't use mean imputation.
 14. In Statistics ,Linear Regression is a linear approach to model the relationship between dependent and independent variable.
 15. Branches of statistics are:
 - a. descriptive statistics
 - b. Inferential statistics