STATISTICS WORKSHEET-1

1. 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

ANS – (d) All of the mentioned

5. _____ random variables are used to model rates.

ANS – (c) Poisson

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

ANS – (a) True

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

ANS – (b) Hypothesis

8. 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

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

ANS – Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean.

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

ANS - Best techniques to handle missing data

* Use deletion methods to eliminate missing data. The deletion methods only work for certain datasets where participants have missing fields.
* Use regression analysis to systematically eliminate data. ...
* We can also use data imputation techniques.

12. What is A/B testing?

ANS - Essentially, A/B testing eliminates all the guesswork out of website optimization and enables experience optimizers to make data-backed decisions. In A/B testing, A refers to 'control' or the original testing variable. Whereas B refers to 'variation' or a new version of the original testing variable.

13. Is mean imputation of missing data acceptable practice?

ANS - True, imputing the mean preserves the mean of the observed data. So if the data are missing completely at random, the estimate of the mean remains unbiased. ... Since most research studies are interested in the relationship among variables, mean imputation is not a good solution.

14. What is linear regression in statistics?

ANS - In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called linear models

15. What are the various branches of statistics?

ANS - The two main branches of statistics are descriptive statistics and inferential statistics Both of these are employed in scientific analysis of data and both are equally important