STATISTICS WORKSHEET-4

1. Ans- The Centeral Limit Therome is important for statistics because it allows us to sefely assume that the sampling distribution of the mean will be normal in most cases. this means that we can take advantage of statistical techniques that assums a normal distribution as we will see in the next section.

2. Ans- Sampling is a process in statistical analysis where researchers take a predetermined number of observation from a large population. The method of sampling depends on the type of analysis being performed but it may include simple random sampling or systermatic sampling.

There are many type :-

Staratified random sampling

Clustering Random sampling

Systermatic random sampling

3. Ans- A type 1 error occurs if an investigator rejects a null hypothesis that is actually true in the population a type 2 error occurs if the investigator fails to reject a null hypothesis that is actually false in the population.

4. Ans- A normal distribution is a type of continous probability distribution in which most data points cluster toward the middle of the range while the rest taper off symmetrically

toward either extreme. The middle of the range is also known as the mean of the distribution.

5. Ans-Covariance is a measure to indicate the extent to which two random variables change in tandem. Correlation is a measure used to represent how strongly two random variables are related to each other. Covariance is nothing but a measure of correlation. Correlation refers to the scaled form of covariance

6. Ans-Univariate statistics summarize only one variable at a time. Bivariate statistics compare two variables. Multivariate statistics compare more than two variables.

7. Ans-The sensitivity is calculated by dividing the percentage change in output by the percentage change in input.

8. Ans-n hypothesis testing there are two mutually exclusive hypotheses; the Null Hypothesis (H0) and the Alternative Hypothesis (H1). One of these is the claim to be tested and based on the sampling results (which infers a similar measurement in the population), the claim will either be supported or not.

9. Ans-Quantitative data is numbers-based, countable, or measurable. Qualitative data is interpretation-based, descriptive, and relating to language. Quantitative data tells us how many, how much, or how often in calculations. Qualitative

data can help us to understand why, how, or what happened behind certain behaviors.

10. Ans-The IQR describes the middle 50% of values when ordered from lowest to highest. To find the interquartile range (IQR), first find the median (middle value) of the lower and upper half of the data. These values are quartile 1 (Q1) and quartile 3 (Q3). The IQR is the difference between Q3 and Q1.

11. Ans-A bell curve is a type of graph that is used to visualize the distribution of a set of chosen values across a specified group that tend to have a central, normal values, as peak with low and high extremes tapering off relatively symmetrically on either side.

12. Ans-An outlier is an unusually large or small observation. Outliers can have a disproportionate effect on statistical results, such as the mean, which can result in misleading interpretations.

13. Ans-The p value is a number, calculated from a statistical test, that describes how likely you are to have found a particular set of observations if the null hypothesis were true. P values are used in hypothesis testing to help decide whether to reject the null hypothesis.

14. Ans-The binomial distribution formula is for any random variable X, given by; P(x:n,p) = nCx x px (1-p)n-x Or P(x:n,p) =

nCx x px (q)n-x,

15. Ans-Analysis of variance, or ANOVA, is a statistical method that separates observed variance data into different components to use for additional tests. A one-way ANOVA is used for three or more groups of data, to gain information about the relationship between the dependent and independent variables.