**STATISTICS-WORKSHEET1**

1. A)True
2. A)Central Limit Theorem
3. B) Modelling bounded count data
4. D)All of the mentioned
5. C) Poisson
6. B) False
7. B) Hypothesis
8. A) 0
9. C) Outliers cannot conform to the regression relationship
10. What do you understand by the term Normal Distribution?

* Normal distribution, also known as Gaussian distribution, the most familiar distribution function for independent, randomly generated variables, it follows the bell shaped curve. The normal distribution is a probability function that tells us how the values of a variable are distributed. It is a symmetric distribution where most of the observations cluster around the central value that is mean and the probabilities for values further away from the [mean](https://statisticsbyjim.com/glossary/mean/) taper off equally in both directions. Extreme values in both tails of the distribution are similarly unlikely.

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

* First of all we need to understand what kind of data it is and how much impact will it make of the overall dataset and how important is that particular parameter with respect to predicting the model. I can simply delete the particular column from the dataset and handle the missing data.
* See if it is a continuous type of variables it is better to calculate to mean/median/mode depends on the dataset and do the linear regression model.
* If it is a categorical data then I change the level as ‘NA’ and can do the prediction for logistic regression model. Another thing is that we do not required to calculate mode that will create bias in your model in case of categorical data.

12. What is A/B testing?

* A/B testing is simply a randomized control experiment. It is a way to contrast the two versions of a variable to find out which variable performs better in a controlled environment. Let’s say for example you have two different kind of products in your company, for example X and Y. Now let’s say you have done some changes in Product Y. Now you can check the sales figure for product Y for a particular period and see how it performs also please note you have not made any changes to product X. To prepare the outside we need to group the data and make a hypothesis to find out the outcome. A/B testing is actually statistical method to understand & test product in a control environment.

13. Is mean imputation of missing data acceptable practice?

* I would say if it is a numerical variables then you can consider mean value and impute the missing data. But it is also depends how large the dataset it, if the data has outliers and spread across then mean does not make any sense if you apply mean for the imputation of missing data.

14. What is linear regression in statistics?

* Linear regression is mainly used to find the relationship between an input variables and the single output variables.
* Y=M1x1+M2x2+M3X3+…+C, here M is the coefficient, C is constant, X1, X2, X3 are input variables and Y is the output variable.
* For example let’s say we have a numerical data points to understand whether we have diabetes/non diabetes, so here diabetes is the output variable and let’s say height, weight, Blood pressure (BP) level, daily food consumption (in calories) are input variables, so we need to predict each input variables from the data points perspective and if above certain level if we have the output value then the patient is diabetic or else non diabetic.

15. What are the various branches of statistics?

* We have mainly two braches one is descriptive statistics and the other inferential statistics, the other concept is also there that is prescriptive statistics.
* For example let’s say you went for health checkup, doctor is finding out you have lung cancer or not, so he will ask some relevant questions and based on the response he will infer something also he will check on symptoms (inference statistics) and the data he can check from the description of your test results (descriptive statistics) which you have done earlier and based on that doctor is going to prescribe medicine/treatment (prescriptive statistics) to cure the disease.