# PYTHON WORKSHEET 1

**1.Which of the following operators is used to calculate reminder in a division?**

©%

**2. In python 2//3 is equal to?**

(B)0

**3.In python,6<<2 is equal to?**

©24

**4.In python, 6&2 will give**

(A)2

**5.In python,6|2 will give**

(D)6

**6. What does the finally keyword denotes in python?**

©the finally block will be executed no matter if the try block raises an error or not

**7.What does raise keyword is used for in python?**

**(**A) It is used to raise an exception

**8.Which of the following is a common case of yield keyword in python?**

**©in defining a generator**

**10.Which of the following keywords in python?**

(A)yield

(B)raise

**9.Which of the following are the valid variable names?**

(A)_abc

©abc2

# SATISTICS WORKSHEET 1

**1.Bernoulli random variables take the values 1 and 0**

(a)True

**5.___random variables are used to model rates.**

©Poisson

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

(B)Hypothesis

**8. Normalized data are centered at___and have units equal to standard diviations of the original data?**

(A)0

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

©outlayers cannot conform to the regression relationship

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

(B)False

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

(B)Modeling bounded count data

**4. Point out correct statement.**

(B) Sums of normally distributed random variables are again normally distributed even if the variables are dependent

**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 simple size increases?**

(A)Central Limit Theorem

11.NORMAL DISTRIBUTION: A mean with a standard deviation, or a continuous random variable when threre probabilities are explained by the normal distribution with the help of mean and standard deviation is called a normal distribution.

12. We handle missing data by two methods imputation or the removal of data.

.Mean or Median Imputation: When data is missing at random, we can use pair wise deletion of the missing observation.

13. A/B is an analytical method which is making decisions that estimate population parameters based on simple statistics, Basically statistical hypothesis testing or in the other word statistical inference.

14. Linear regression which is a kind of statistical analysis that attempts to show a relationship between two variables.

15. There are two types of statistics 1) Descriptive statistics and 2) Inferential statistics. Descriptive statistics have two types 1) central Tendency and 2) Spread Data. In the central tendency there are mean, median and mode or in the spread data there are range, variance, standard deviation and percentile.

# MACHINE LEARNING

**1.Which of the following methods do we use to find the best fit line for data in Linear Regression?**

(A) Least Square Error

**2.Which of the following statement is true about outlayers in Linear regression?**

(A) Linear regression is sensitive to outliers

**3. A line falls left to right if the slope is____**

(B) negative

**4.Which of the following will have symmetric relation between dependent and independent?**

©both of them

**5. Which of the following is the reason for over condition?**

© Low bias and high variance

**6.If output involves label then that model is called as:**

(B)Predictive label

**7.Lasso and Ridge regression techniques belong to_____**

(D)Regularization

**8.To overcome with imbalance dataset which techniques can be used?**

(D)Smote

**9.The AUCROC curve is an evaluation metric for binary classification problems. It uses___to make graphs?**

© Sensitivity and Specificity

**10.AUCROC curve is an evaluation metric for the better model area under the curve should be less**

(B)False

**11.Pick the feature extraction from below:**

(b)Apply PCA to project high dimensional data

**12.Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?**

(B)It becomes show when number of features is very large

©we need to iterate

13.Regularizations are techniques which is used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.

14.LASSO,Ridge,Elastic-Net regression are used for regularization.

15.A error term which is represents the margin of error within a statistical model. That is refers to the sum of the deviations within the regression

Line, that can provide an explanation for the difference between the theoretical value of the model.