**MCQs**

1. Which of these distributions is not Continuous Probability Distribution?
   1. **Binomial Distribution**
   2. Normal Distribution
   3. Chi-Square Distribution
   4. F-Distribution
2. For detecting outliers using boxplot related metrics, what do we usually do?
   1. Consider those observations as outliers which lie outside 1.5\*(standard deviation) limits
   2. **Consider those observations as outliers which lie outside 1.5\*(Inter-Quartile Range) limits**
   3. Consider those observations as outliers which lie outside 1.5\*(Mean) limits
   4. Consider those observations as outliers which lie outside 1.5\*(Median) limits
3. We want to consider fitting distribution to the number of customers arriving at a customer service counter within a specific 1-hour time slot. Which of the distributions is suitable for this?
   1. Binomial Distribution
   2. Normal Distribution
   3. Uniform Distribution
   4. **Poisson Distribution**
4. Mutually Exclusive Events are
   1. Events that have everything in common
   2. **Events that have nothing in common**
   3. Events that are independent of each other
   4. Events that can occur concurrently
5. Exhaustive Events are
   1. Set of Events that have everything in common
   2. **Set of Events among which at least one will occur**
   3. Set of Events that are independent of each other
   4. Set of Events that can occur concurrently
6. t-tests are tests for testing which of the following measures?
   1. **Means**
   2. Standard Deviation
   3. Median
   4. Quartiles
7. A medicine is given to a patient for improving the platelet count in blood. The platelet count before giving medicine was measured and then after giving the medicine was measured. This was done for some 20 patients. Which hypothesis test is suitable for testing the effectiveness of that medicine?
   1. Independent Samples t-test
   2. **Paired t-test**
   3. 1 – way ANOVA
   4. One Sample t-test
8. For testing means for multiple populations which hypothesis test can be used?
   1. Independent Samples t-test
   2. Paired t-test
   3. **1 – way ANOVA**
   4. One Sample t-test
9. In Linear Programming Problems, the restrictions in the scenario are expressed mathematically in terms of
   1. Objective Function
   2. Variables
   3. **Constraints**
   4. Divisions
10. For virtually viewing a scenario in a realistic way, we can do
    1. Calculation of Probabilities
    2. **Simulation**
    3. Hypothesis Testing
    4. Machine Learning
11. Monte Carlo Simulation Technique is based on
    1. **Cumulative Probabilities**
    2. Independent Probabilities
    3. Mean
    4. Median
12. Which of the metrics is not classification metrics?
    1. Area under ROC
    2. Log Loss
    3. **Mean Squared Error**
    4. F1-Score
13. For a scenario for any optimization problem, for which cost is the objective function, what type of problem it will be?
    1. Maximization
    2. **Minimization**
    3. Neither Maximization or Minimization
    4. None of Above
14. Correlation Coefficient between variables as “Student Population” and “Sales” is found to be 0.846. What can you comment?
    1. As “Student Population” is increasing “Sales” decreasing with high degree
    2. **As “Student Population” is increasing “Sales” increasing with high degree**
    3. “Student Population” and “Sales” seem to be no way related with each other
    4. “Student Population” and “Sales” seem to be perfectly positively related to each other
15. If Linear Correlation coefficient between two variables is very close to zero then which of the following needs to be done for building a good predictive model?
    1. Build a Linear Regression using Ridge Regression
    2. Build a Linear Regression using Lasso Regression
    3. Build a Linear Regression using ElasticNet Regression
    4. **Try some other model or explore non-linear relationships and try models on them**
16. Ridge Regression considers in its cost(error) function
    1. **Squares of the beta coefficients**
    2. Absolute Values of the beta coefficients
    3. Both Squares and absolute values of the beta coefficients
    4. Correlation coefficient
17. Lasso Regression considers in its cost(error) function
    1. Squares of the beta coefficients
    2. **Absolute Values of the beta coefficients**
    3. Both Squares and absolute values of the beta coefficients
    4. Correlation coefficient
18. Elastic Net Regression considers in its cost(error) function
    1. Squares of the beta coefficients
    2. Absolute Values of the beta coefficients
    3. **Both Squares and absolute values of the beta coefficients**
    4. Correlation coefficient
19. Purpose of Regularized Regression Techniques is to
    1. Reduce the values of beta coefficients for all the variables
    2. **Reduce the values of beta coefficients for only those variables which hardly impact the response variable**
    3. Increase the values of beta coefficients for all the variables
    4. Increase the values of beta coefficients for only those variables which hardly impact the response variable
20. The purpose of Clustering is to
    1. **Form homogeneous groups among the data so that those groups(clusters) can be further analysed separately**
    2. Calculate the means of all the variables
    3. Calculate the variances of all the variables
    4. Scale all the variables
21. Complete Linkage Method for calculating the distance between two clusters considers
    1. **Distance between two farthest points**
    2. Distance between two nearest points
    3. Average Distance of all the points between the two clusters
    4. Distance between two centroids of the two clusters
22. Single Linkage Method for calculating the distance between two clusters considers
    1. Distance between two farthest points
    2. **Distance between two nearest points**
    3. Average Distance of all the points between the two clusters
    4. Distance between two centroids of the two clusters
23. Average Linkage Method for calculating the distance between two clusters considers
    1. Distance between two farthest points
    2. Distance between two nearest points
    3. **Average Distance of all the points between the two clusters**
    4. Distance between two centroids of the two clusters
24. Centroid Linkage Method for calculating the distance between two clusters considers
    1. Distance between two farthest points
    2. Distance between two nearest points
    3. Average Distance of all the points between the two clusters
    4. **Distance between two centroids of the two clusters**
25. For testing the equality of means between two different populations which test can be used?
    1. One Sample t-test
    2. Paired t-test
    3. **Two Independent Samples t-test**
    4. Chi-Squared Test
26. What is the difference between parametric tests and non-parametric tests?
    1. Parametric tests for prediction and non-parametric test are for analysis
    2. **Parametric tests assume a distribution for population and non-parametric tests don not have any such assumption**
    3. Parametric tests are supervised and non-parametric tests are unsupervised
    4. Parametric tests assume Normal Distribution whereas non-parametric tests assume Poisson Distribution