Scholarship Exam - Statistics Questions

1. Identify the variables that are continuous or discrete? Mark: 1
   a) Time & country are continuous. Weight & colour are discrete
   b) Weight & colour are continuous. Time & country are discrete
   c) Time & country & weight & colour are continuous
   d) Time & weight are continuous. Country & colour are discrete
   Ans – d

2. A histogram is used for what kind of data? Mark: 1
   A. Continuous
   B. Discrete
   Ans – A

3. What method of data representation is best suited to the demonstration of data results if that data is of differing nominal values and needs to represent quantitative data on X axes? Mark: 1
   A. Bar chart
   B. linear chart
   C. pie chart
   D. scatter plot
   Ans - A

4. In a project the business wants to see the relationship between revenue generated and YoY Sales. Which plot is the best? Mark: 1
   A. Scatter plot
   B. Box plot
   C. Density plot
   D. Matrix plot
   Ans - A

5. Millions of Americans work from home during office hours and following is a sample data of individuals who work at home -
   18,54,20,46,25,48,53,27,26,37,40,36,42,25,27,33,28,40,45,25

   Find the Mean and mode? Mark: 2
   Ans - 34.75,25

6. The median age of population of all adults is 36 years. Using the median age obtained in Q5 comment whether the at-home workers tend to be younger or older than the population of adults
   Ans = 34.5 - At home workers are slightly younger Mark: 1

7. Using the sample data from Q5 compute the first and third quartile Mark: 2
   Ans - 25.5,43.5

8. If I replace the 18 with 38 will the standard deviation   Mark :2
    A. Increase
    B. Decrease
    C.remain same?
    Ans – Decrease, With 18 SD = 10.61 and 38 SD = 9.95

9. What is the relation between Standard Deviation and variance when the sample is less than 30?   Mark:1
    a) The SD is directly proportional to variance
    b) The SD is variance/(Sqrt(n))
    c) The SD is the Sqrt(variance)
    d) More than one correct answer
       Ans - C

10. If a population consists of distinct 5000 records between 50 and 5050 and if the last 10 samples are taken for analysis. What is the range of the sample?   Mark:1
    Ans = 10

11. Coefficient of variation Formula?   Mark:1
    Ans - (SD/Mean)*100%

12. What is the range of correlation coefficient?   Mark:1
    A. -1 to +1
    B. -∞ to + ∞
    C. 0 to 1
    D. None of the above
       Ans - A

13. Car rental rates per day for a sample of seven Eastern US cities are as follows

City	Boston	Dallas	Atlanta	Ohio	New York	Miami	Pittsburgh
Rate ($)	43	35	34	58	30	30	36

    Compute the Mean, variance, standard deviation for the car rental rates
    Ans – 38, 97, 9.85

14. A similar sample of seven Western US cities showed a sample mean of $ 38 per day and variance and SD as 93, 9.64. What can you infer from this?   Mark:1
    A. Eastern shows more variation
    B. Western shows more variation
    C. Eastern and western have high variations
    D. All the above
       Ans - Eastern shows more variation

Scholarship Exam - Statistics Questions

15. In an IPL auction a team wants to buy a foreign player who is an all-rounder to balance their side. The team has the budget to spend on only one player and they are looking for a player who could help them in their batting department. The franchise now has the data with the runs scored by player A and player B in their last 20 matches and are confused on their selection

    Who should the Franchise buy and Why?   Mark: 2

    Player X   25,40,38,48,51,35,100,25,15,20,21,25,35,40,7,12,25,40,42,56
    Player Y   0,25,22,38,25,45,51,60,62,93,68,63,42,9,3,5,20,14,18,37

    Ans - SD of X = 19.69
    SD of Y = 25.29

16. What is the Coefficient of variation of X and Y?   Mark: 2
Ans X = 0.562, Y =0.722

17. The sales report about the pharmaceutical company in million $ for the 21 states in US has been provided in the spread sheet   Mark : 5
8408,1374,1872,8879,2459,11413,608,14138,6452,1850,2818,1356,10498,7478,4019,4341,739,2127,3653,5794,8305

    1. Provide the five number summary of the box plot **(Ans - 608,1872,4019,8305,14138)**
    2. Compute the IQR,lower and upper limits**(Ans – 6433,-7777.5,17955)**
    3. Does the data contain any outlier? **(Ans – No)**
    4. Ohio state has the highest sales at $14,138 million. Suppose a data entry error has been made as $ 41,138 million would this been identified as an outlier and corrected
    (Yes, 41,138 would be an outlier)

18. Sample observations were taken between x and y as follows –   Mark: 5

X	6	11	15	21	27
Y	6	9	6	17	12

    a) Compute the covariance **Ans – 26.5**

    b) Compute the correlation coefficient **Ans – 0.693**

    c) What is the relationship between x and y **Ans- Positive linear relationship between x &y**

Scholarship Exam - Statistics Questions

22. 1.The Sum of probabilities of all events is 1   Mark :1

    2 The probability lies between -1 to +1

    3. In a mutually exclusive event P(AnB) = 1

    4. In a mutually exclusive event P(AUB) = 1

From the above please check the appropriate option

   a)   1 & 2 are True
   b)   1 & 3 are True
   c)   1 & 4 are true
   d)   None of the above
     Ans - C

23. When 2 coins are tossed and the probability of getting 2 heads is 0.25 what is the size of sample space?   Mark :1
    Ans = 4

24. A coin is tossed 10 times and below are the observations as follows – T H T H H T H H H T
What is the theoretical probability of obtaining head as success.   Mark :1
Ans = 0.5

25. Twenty four people had a blood test and the results are shown below.
A , B , B , AB , AB , B , O , O , AB , O , B , A, AB , A , O , O , AB , O , O , A , AB , O , B , A

If a person is selected randomly from the group of twenty four people, what is the probability that his/her blood type is not O?   Mark :1
Ans =0.667

26. CSK winning IPL 2018 (0.8 probability) Mark:1

    SRH winning IPL 2018 (0.2 probability)

    CSK winning Champions trophy 2018 (0.6 probability)

    SRH winning Champions trophy 2018 (0.4 probability)

In the above events what are the mutually exclusive events?

Ans (1 & 2 OR 3 & 4)

Scholarship Exam - Statistics Questions

29. In a loan payment case study, the probability that the person is middle aged is 0.690 and probability that the person is not a loan defaulter is 0.786. There is 0.516 probability that person will be a middle aged loan defaulter.

Can you find the probability that if the person is found to be a loan defaulter, what is the chance of him being a middle aged?   Mark:2
Ans = 0.656

30. In a bag of 5 different coloured balls, 2 balls are picked randomly. What are the total number of combinations that the balls can be picked?   Mark:2
Ans = 10 (nCr Formula)

32. Identify the probability distribution

If P is the probability of getting a head during n flips. What is the distribution for estimating number of flips required for getting first head?   Mark :1
Ans – Geometric

33. In a hotel on an average 10 customers visit in an hour. What is the distribution for identifying the probability of 20 customers visiting in an hour   Mark:1
Ans – Poisson

34. In a bag there are 2 coloured balls – red and green. The probability of getting a red ball is 0.4 and if 10 balls are drawn randomly (independent event) then the probability of getting at most 3 red balls. Which distribution?   Mark:1
   Ans – Binomial

35. What is the mean and standard deviation of a standard normal distribution? Mark:1
   Ans = 0,1

36. What is the Empirical Rule for the data distribution in a normal distribution Mark:2
   Ans = (68,95,99.7)

37. Which test to be performed when we have only the mean of population and sample is less than 30?   Mark:1
   Ans = T test

38. In a Normal distribution Mark:1
   a) mean > median > mode
   b) median > mean > mode
   c) mean = median = mode
   d) All the above
      Ans =C

39. A sample of size 50 is drawn from a population of mean 100 and Standard deviation 25. What is the Standard deviation of the sample?   Mark:2
   Ans= 3.53

40 A pharma company manufactures thousands of tablets every day. The manufacturing team gets a complaint stating that the weight of a tablet named zingx has changed from its actual claimed weight of 100mg with population standard deviation 20. The company wants to test this and submit a report to the concerned authority stating the proof of this complaint. The company takes 50 samples with mean 105 to test this. Mark:2

What is the Null and alternate hypothesis?

a) $H_0 = 100$ and $H_a \neq 105$
b) $H_0 = 100$ and $H_a < 105$
c) $H_0 = 100$ and $H_a \neq 100$
d) $H_0 = 105$ and $H_a \neq 105$

Ans C

43. At Ohio University the mean score of scholarship exam for fresh applications is 900 and the population standard deviation is 180. Every year the HOD uses sample applications to determine the change in the examination score. A sample of 200 applications with a sample mean of 935 is used to perform hypothesis test. What is Z value and the result?   Mark:5

Ans = 2.74, Reject the Null hypothesis

44. The California university performs the hypothesis test for the same scenario as Ohio university on 6 samples with the population mean as 900 with samples as 935,925,850,875,945,915

What is the T value and the result?   Mark:5
Ans = T from table 2.571 and T calculated 2.53, Accept the Null hypothesis

45. A pharma company wants to test its 3 formulations of its new relief drug. 27 volunteers were randomly grouped in 3 groups and given a different drug formulation. The participants took the drug when they had an attack and recorded the pain on a scale 1 to 10.

The company wants to understand are these differences with the group or between groups.   Mark:5

Group 1 - 2 5 4 7 3 3 2 5 6

Group 2 - 5 5 2 4 3 2 2 6 5

Group 3 - 5 6 2 4 3 3 6 5 2

Perform ANOVA for a significance level of 0.10 and publish the hypothesis results
Ans F calculated 0.096552, F from table 2.53 (Accept the Null hypothesis)

43. Write an article on Statistics(What all we have learned) and mail me on nursnaaz@gmail.com on or before 31st May 2018.   **Mark:10**