**B.Sc CSIT 2nd Semester.**

**Practical Problems ( Statistics- I) and important questions for final exam**.

1. The following table gives the installation time( in minutes) of hardware on 50 different computers.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Installation time( in minutes) | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | Total |
| Frequency | 4 | - | 10 | - | 10 | 50 |

And average installation time is 30.2 minutes. Find missing frequencies.

1. A manufacturer of electronic components is keenly interested in identifying the lifetime of certain type of battery used in laptops. A random sample, in days of life time is recorded a follows:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 492 | 464 | 488 | 440 | 900 | 504 | 500 | 444 | 472 | 468 |

1. Compute the sample mean and median life time.
2. What feature in the data set would be responsible for the considerable difference between these two averages?
3. Under such scenario, among these two measures, which measure would be preferable?
4. Two computer manufactures A and B compute for profitable and prestigious contract. In their rivalry, each claim that their computer a consistent. For this, it was decided to start execution of the same program simultaneously 0n 50 computer of each company and recorded the time as given below:

|  |  |  |
| --- | --- | --- |
| Time( in sec) | Number of computers |  |
| Computer A | Computer B |
| 0-2 | 5 | 2 |
| 2-4 | 16 | 7 |
| 4-6 | 13 | 12 |
| 6-8 | 7 | 19 |
| 8-10 | 5 | 9 |
| 10-12 | 4 | 1 |

Which company’s computer is more consistent?

1. Nepal Telecommunication ADSL section daily deals huge number of customers who calls for new ADSL internet connections and for ADSL internet complains. The following are the data representing number of applications for new ADSL internet service registered during twelve consecutive days in Nepal Telecommunication office.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Days | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Number of new ADSL application | 37 | 50 | 51 | 58 | 52 | 107 | 53 | 45 | 45 | 44 | 47 | 43 |

1. Calculate arithmetic mean, median, mode, quartile deviation and standard deviation.
2. Evaluate whether there is outlier(s) in the dataset by using 1.5(IQR) rule.
3. If there is outlier(s), delete them, and then re-compute arithmetic mean, median, mode, quartile deviation and standard deviation.
4. Do you find any effect of outlier(s) in descriptive measures? Comment on this issue.
5. A computer technician install the hardware on 100 different computers. The installation time ( in min) to the hardware in given below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Time( in minutes) | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 |
| No. of computer | 10 | 14 | 18 | 24 | 16 | 12 | 6 |

Calculate percentile coefficient of Kurtosis.

1. The probability of discrete random variable is given below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| X=x | -2 | -1 | 0 | 1 | 2 | 3 |
| P(X=x) | 0.1 | K | 0.1 | 0.2 | 0.3 | 0.15 |

Find:

1. Find the value of K.
2. P(X≤0)
3. F(1)
4. E(X), E(X2) and V(X)
5. E(8X+2)
6. V(2X+20)
7. If the values of joint probability distribution of X and Y are shown in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| X Y | 2 | 4 | 5 |
| 1 | 1/12 | 1/24 | 1/24 |
| 2 | 1/6 | 1/12 | 1/8 |
| 3 | 1/4 | 1/8 | 1/12 |

Find:

1. P(X≤2,Y≤4)
2. Marginal pmf of X and Y.
3. P(Y=2/X=1)
4. The joint pdf of X and Y is given below

f(x,y) = e-(x+y) ; x>0,>0

= 0

Find:

1. Marginal distribution of X
2. Marginal distribution of Y
3. Are X and Y are independent?
4. Conditional distribution of X for given Y.
5. Conditional distribution of Y for given X.
6. Joint distribution function of X and Y.
7. Fit a binomial distribution and compute the expected frequency to the following data.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. of seeds germinating | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| No. of rows | 2 | 1 | 2 | 10 | 30 | 35 | 20 |

1. Fit a Poisson distribution and compute the expected frequency to the following data.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No. of calls received per days | 0 | 1 | 2 | 3 | 4 |
| No. of days | 9 | 12 | 5 | 4 | 1 |

1. Observations on the yield of a chemical reaction taken of various temperature(0C) are recorded as follows:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Temp(0C) | 25 | 28 | 35 | 20 | 22 | 30 | 31 | 22 |
| % of yield | 35 | 39 | 48 | 29 | 30 | 38 | 40 | 30 |

1. Compute the Karl Pearsion’s correlation coefficient and interpret your result.
2. Calculate the probable error.
3. Interpret of the value of r is significant or not?
4. Determine the limit within which the population correlation coefficient may be expected to lie.
5. Fit the regression model to describe the given data and also interpret the estimated regression coefficient.
6. What percentage of variation on data rate is explained by the variation on temperature?
7. Predict the % of yield when temperature is 400c.
8. Compute the standard error of estimate.
9. Compute the residual for % of yield when temperature is 300C.