ICE Lab

Problem # 1:

The information of 20 persons is given in the following table:

SN	Sex	Religion	Level of Education	SN	Sex	Religion	Level of Education
1	Male	Muslim	Primary	11	Female	Hindu	Primary
2	Female	Hindu	Graduate	12	Male	Christian	Graduate
3	Male	Muslim	Illiterate	13	Male	Others	Secondary
4	Male	Hindu	Graduate	14	Female	Muslim	Secondary
5	Female	Muslim	Primary	15	Male	Hindu	Higher Secondary
6	Female	Christian	Graduate	16	Male	Christian	Others
7	Male	Muslim	Illiterate	17	Female	Muslim	Primary
8	Male	Hindu	Primary	18	Male	Others	Illiterate
9	Female	Muslim	Others	19	Female	Muslim	Higher Secondary
10	Male	Others	Higher Secondary	20	Male	Others	Secondary

(i) Construct the frequency distribution for variables "Religion" and "Level of education".

(ii) Draw pie diagram for the variable "Religion" and comment.

(iii) Draw bar diagram for the variable "Level of Education" and comment.

Problem # 2:

Suppose that 100 students are enrolled in a statistics class and the following are the test scores received by them:

```
77  44  49  33  38  33  76  55  68  39  44  59  36  55  47  61  53  32  65  51
29  41  32  45  83  58  73  47  40  26  59  43  66  44  41  25  39  72  37  55
34  47  66  53  55  58  49  45  61  41  55  92  83  77  45  62  45  36  78  48
54  50  51  66  80  73  57  61  56  50  45  82  71  48  46  69  38  72  56  64
38  45  51  44  41  68  45  92  43  12  37  16  44  57  63  71  40  64  57  51
```

(i) Compute Mean, Median, Mode, Variance and Standard deviation of the above raw data and comment on your results.

(ii) Find the five number summaries.

(iii) Display the box plot. Is there any outlier of the data?

(iv) Select an appropriate class interval and organize the data set into a frequency distribution.

(v) Using the frequency distribution obtained in question (iv) construct a histogram and an ogive. Also approximate the median and mode with the help of ogive and histogram respectively.

(vi) Find the mean, median, mode, skewnwss and kurtosis using the frequency distribution obtained in question (iv).

Problem #3:

The following data represents the ages of the 50 richest people in the world in 2009.

89, 89, 87, 86, 85, 83, 83, 82, 81, 80, 78, 78, 77, 76, 73, 73, 73, 72, 69, 69, 68, 67, 66, 66, 65, 65, 64, 63, 61, 61, 60, 59, 58, 57, 56, 54, 54, 53, 53, 51, 51, 49, 47, 46, 44, 43, 42, 36, 2000.

(i) Find the mean, median and mode of the ages of the 50 richest people. Which measures of central tendency best describes a typical entry of this data set.
(ii) Replace 35 instead of 2000 from the data set then rework (i). Compare these measures of central tendency with those found in (i).
(iii) Construct a frequency distribution using the above data after replacing 35 instead of 2000.
(iv) Construct a relative frequency histogram.
(v) Find the mean, median, mode and variance for grouped data. Comment on the results in the context of the data.

Problem #4:

The grade point average (GPA) in different semesters of two students are shown below:

Student	GPA in semesters							
	1	2	3	4	5	6	7	8
A	2.5	2.5	3.0	3.5	3.5	4.0	3.5	3.5
B	2.5	3.0	4.0	4.0	4.0	2.0	2.5	4.0

Which students would you consider better throughout the courses of studies?

Problem #5:

If X follows binomial distribution with n=50 and p=0.6/0.5/0.2.

i.   Sketch the graph for binomial probability distribution.

ii. Compute

   a. $P(X = 35)$

   b. $P(X \leq 20)$

   c. $P(X > 15)$

   d. $P(20 < X < 45)$

iii. Find first four central moment of the distribution.

iv.   Find the skewness and kurtosis of the distribution.

Problem #6:

Given the following data

0,1,2,3,4,…………………………..,148,149,150.

Show the relation between binomial distribution and Poisson distribution.

Problem #7:

If Z~N(0,1). For the following values of Z

-4.0, -3.9, -3.8, -3.7,………………………..,3.8,3.9,4.0.

(i) Create pdf of Z. Draw standard normal curve and comment the shape characteristics of the distribution.
(ii) Create pdf and cdf of X~N(1000,250000).
(iii) Find
      a) P(X=850)
      b) P(X>1200)
      c) P(1000<X<2000).
(iv) Construct normal density curve and normal cumulative distribution curve. Comment on your results.

❖ **Problem #8:**

The following data gives the number of printing mistakes in a book of five hundred pages.

No. of printing mistakes	No. of pages	No. of printing mistakes	No. of pages
0	150	10	18
1	37	11	17
2	34	12	16
3	30	13	14
4	28	14	11
5	25	15	9
6	24	16	8
7	22	17	7
8	21	18	5
9	20	19 and above	4

i.   Fit Poisson distribution to the above data.

ii.   Sketch the graph for Poisson distribution.

iii. Compute

      a. $P(X = 15)$

      b. $P(X < 10)$

   c. $P(X > 15)$

   d. $P(4 < X < 17)$

iv. Find the skewness and kurtosis of the distribution.

v. Find the probability of Poisson distribution using recurrence relation.

❖ **Problem #9:**

Hypothetical data on weekly family consumption expenditure (Y) and weekly family income (X) are given below:

Y	70	65	90	95	110	115	120	140	155	150
X	80	100	120	140	160	180	200	220	240	260

  i. Construct a Scatter plot of the weekly family consumption expenditure (Y) and weekly family income (X). Do you think are any relationship between X and Y? Do you think a linear model is appropriate for this data?

  ii. Find the coefficient of correlation between weekly family consumption expenditure (Y) and weekly family income (X) and comment on your result.

  iii. Obtain the line of best fit for Y on X.

  iv. How do you interpret the intercept and the slope of the regression line?

  v. If the weekly family income is 175 compute the weekly consumption expenditure.

  vi. Check the goodness of fit of the least squares fit.