Department of Mathematics
PROBABILITY & STATISTICS (MATH2331)
TUTORIAL PROBLEMS

Dr.D.Mallikarjuna Reddy

Tutorial Problems-0 (Data averages, Variation, Missing frequencies etc)

1. Find the missing frequencies in the following frequency distribution if the mean is 50.25 and the total no of students in the class is 80.

Marks obt.(%)	25	35	45	55	65	75	85
No. of students	7	23	?	15	10	?	5

2. Mean of 100 observations is found to be 44. If at the time of computation two items are wrongly taken as 30 and 27 in place of 3 and 72. Find the corrected average.

3. In the frequency distribution of 100 families given below, the number of families corresponding to expenditure groups 20—40 and 60—80 are missing from the table. However, the median is known to be 50. Find the missing frequencies.

Expenditure	0—20	20—40	40—60	60—80	80—100
No. of families	14	?	28	?	15

4. The mean wage of 100 workers in a factory, running two shifts of 60 and 40 workers respectively is Rs.38. The mean wage of 60 workers in morning shift is Rs.40. Find the mean wage of 40 workers working in the evening shift.

5. In a distribution of 10 observations, the value of mean and standard deviation are given as 20 and 8. By mistake, two values are taken as 2 and 6 instead of 4 and 8. Find out the value of correct mean and variance.

6. For two firms A and B belonging to same industry, the following details are available :

	Firm A	Firm B
Number of Employees :	100	200
Average wage per month :	Rs. 240	Rs. 170
Standard deviation of the wage per month :	Rs. 6	Rs. 8

Find

 (i) Which firm pays out larger amount as monthly wages?

(ii) Which firm shows greater variability in the distribution of wages?

(iii) Find average monthly wages and the standard deviation of wages of all employees for both the firms.

7. From the following frequency distribution of heights of 360 boys in the age-group 10 – 20 years calculate the :
   (i) arithmetic mean;
   (ii) coefficient of variation; and

(iii)   quartile deviation

Height (cms)	No. of boys	Height (cms)	No. of boys
126 – 130	31	146 – 150	60
131 – 135	44	151 – 155	55
136 – 140	48	156 – 160	43
141 – 145	51	161 – 165	28

11. A set of n values $x_1, x_2, ..., x_n$ has standard deviation σ. The standard deviation of n values $x_1 + k$, $x_2 + k$, ..., $x_n + k$ will be-------

12. Find mean and the standard deviation of the first n natural numbers.

13. . While calculating the mean and variance of 10 readings, a student wrongly used the reading 52 for the correct reading 25. He obtained the mean and variance as and 16 respectively. Find the correct mean and the variance.

14. .Information collected on the average strength of students of Statistics course in two colleges is as follows:

Measure	College A	College B
Mean	150	145
Median	141	152
S.D	30	30

Can we conclude that the two distributions are similar in their variation?

15.Following are the marks obtained, out of 100, by two students Ravi and Hashina in 10 tests.

Ravi   : 25 50 45 30 70 42 36 48 35 60
Hashina :   10 70 50 20 95 55 42 60 48 80
Who is more intelligent and who is more consistent?

Tutorial Problems-1

(Probability, addition law, multiplicative law)

1. A box contains 6 red, 4 white and 5 black balls. A person draws 4 balls from the box at random. Find the probability that among the balls drawn there is at least one ball of each color.

2. If two dice are thrown, what is the probability that the sum is (1) greater than 8, (2) At least 8 (3) neither 7 nor 11

3. The letters in word REGULATIONS is arranged at random. What is the chance that there will be exactly four letters in between R and E.

4. The odds against that person A speaks the truth are 3:2 and the odds that person B speaks the truth 5:3. In what percentage of cases are they likely to contradict each other on an identical point?

5. A problem in statistics is given to three students A, B and C, whose chances of solving it are respectively.1/2 , 1/3 and 1/4 What is the probability that the problem will be solved?

Tutorial Problems-II (Baye's Theorem – Random Variables)

1. In a bolt factory machines A_1, A_2, A_3 manufacture respectively 25%, 35% and 40% of the total output. Of these 5, 4, and 2 percent are defective bolts. A bolt is drawn at random from the product and is found to be defective. What is the probability that it was manufactured by machine A_2

2. The chance that doctor A will diagnose a disease X correctly is 60%. The chance that a patient will die by his treatment after correct diagnosis is 40% and the chance of death by wrong diagnosis is 70%. A patient of the doctor A, who had disease X died. The chance that his disease was diagnosed correctly.

3. A random variable has the following probability distribution

X	0	1	2	3	4	5	6	7	8	9
P(x)	a	3a	5a	7a	9a	11a	13a	15a	17a	19a

   (1) Determine 'a' (2) Find (i) P(x < 3) (ii) P(x ≤ 3) (iii) P(x >7) (iv) P(2 ≤ x ≤ 5),
   (v) Mean and variance of the distribution.

4. A random variable X has the following probability distribution

X	-3	6	9
P(X=x)	1/6	1/2	1/3

   Find (1) Mean and variance (2) Find E[Y], Var[Y], given that Y=2X+1

5. A continuous random variable X has the probability density law $f(x) = Ax^2, 0 \leq x \leq 1$. Determine A (ii) a, P(X ≤ a) = P(X>a) (iii) Mean, SD

6. The diameter X of an electric cable is assumed to be a continuous random variable with pdf $f(x) = 6x(1-x), 0 \leq x \leq 1$. (a) Show that it is pdf (b) Determine b such that Pr(X ≤ b) = Pr(X > b). (c) Mean SD

Tutorial Problems-III (Binomial-Poisson distribution)

1. A multiple choice questionnaire has 12 questions with 5 options A-E. a student is completely un prepared and wrote the test. If test follows binomial probability law find (1) Exactly two answers are correct 2) At least two answers are correct 3) More than 7 answers are correct.

2. With the usual notation find p for binomial random variable X if n = 6 and 9P(X = 4) = P(X = 2)

3. The mean and variance of a Binomial variate X with parameters n and p are 16 and 8. Find P (X = 1).
4. A typist makes on average 2 mistakes per page. What is the probability of a particular page having no errors on it?

5. Components are packed in boxes of 20. The probability of a component being defective is 0.1. What is the probability of a box containing 2 defective components?

Tutorial Problems-IV (Normal distribution)

1. X is a normally distributed with mean $\mu = 30$ and SD $\sigma = 4$. Find a) $P(x < 40)$ b) $P(x > 21)$ c) $P(30 < x < 35)$.
2. In a normal distribution 10.03% of the items are under 25 kilogram and 89.97% of the items are under 70 kilogram.

3. What are the mean and standard deviation of the distribution A large group of students took a test in Physics and the final grades have a mean of 70 and a standard deviation of 10. If we can approximate the distribution of these grades by a normal distribution, what percent of the students a) scored higher than 80? b) Should pass the test (grades≥60)? c) Should fail the test (grades<60)?

Tutorial Problems-V (Curve Fitting, Correlation –Regression)

1. Using least square method, fit a curve of the form $Y=ax^2+bx+c$ following data:

   x: 1   3   5   6
   y: 12 18 25 35

2. Fit an exponential cure $Y=ab^X$ for the following data

X	40	65	90	5	30	10	80	85	70	25
Y	30	20	10	80	40	65	15	15	20	50

3. Calculate the correlation coefficient for the following heights (in inches) of fathers (X) and their sons (Y):

   X : 65 66 67 67   68 69 70 72
   Y : 67 68 65 68   72 72 69 71
4. Find the spearman rank correlation coefficient to the following data:

X:	11	12	43	84	15
Y:	8	15	30	60	12

5. Fit a linear regression equation of Y on X to the following data:

X:	5	8	7	6	4
Y:	3	4	5	2	1

6. Estimate the production for the year 2010, by fitting regression line to the following data:

Year:	2003	2004	2005	2006	2007
Production:	5	8	14	12	13

(in thousand Qt).

Estimate the production in the year of 2008.

Tutorial Problems-VII (Large Sample Tests) (Z-test means)

1. A sample of 900 members has a mean 3.4 cms and s.d 2.61 cms. Is the sample from a large population of mean 3.25 cms and s.d 2.61 cms? 5% level.

2. The means of two large samples of 1000 and 2000 items are 67.5 cms and 68.0cms respectively. Can the samples be regarded as drawn from the population with standard deviation 2.5 cms. Test at 5% level of significance.

3. A random sample of 400 students is found to have a mean height of 171.38 cms. Can it be reasonably regarded as a sample from a lalrge population with mean height 171.17 cms. and standard deviation 3.30 cms. (Test at 5% level of significance).

4. A test of the breaking strengths of two different types of cables was conducted using samples of $n_1 = n_2 = 100$ pieces of each type of cable.

	Cable I	Cable II
	$\bar{x} = 1925$	$\bar{y} = 1905$
	$\sigma_1 = 40$	$\sigma_2 = 30$

Do the data provide sufficient evidence to indicate a difference between the mean breaking strengths of the two cables at 10% level?

5. An insurance agent has claimed that the average age of policy holders who issue through him is less than the average for all agents which is 30.5 years. A random sample of 100 policy holders who had issued through him gave the following age distribution?

Age in years:	16-20	21-25	26-30	31-35	36-40
No of persons:	12	22	20	30	16

Test the significant difference at 5% level of significance. Tabulated value at 5% = 1.645.

Tutorial Problems-VIII (Z-test Proportions)

1. A machine puts out 16 imperfect articles in a sample of 500. After machine is overhauled, it puts out 3 imperfect articles in a batch of 100. Has the machine improved?

2. In a sample of 1000 people in a state, 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat eaters are equally popular in this state at 1% level of significance? (*table value=2.58*).

3. Before an increase in excise duty on tea, 800 persons out of a sample of 1000 persons were found to be tea drinkers. After an increase in excise duty, 800 persons were tea drinkers in a sample of 1200 persons. State whether there is a significant decrease in the consumption of tea after the increase in excise duty? (Test at 5% level of significance).

4. A random sample of 500 apples was taken from a large consignment and 60 were found bad. Obtain the 98% confidence limits for the percentage of bad apples in the consignment (given z = 2.33)

5. 1000 articles from a factory A are examined and found to have 3% defectives. 1500 similar articles from a second factory B are found to have only 2% defectives. Can it be reasonably concluded that the product of the first factory is inferior to the second?

Tutorial Problems-IX : (Small Sample Tests)(t-test)

1. A sample of 20 items has mean 42 units and S.D 5 units. Test the hypothesis that it is a random sample from a normal population with mean 45 units.

2. A sample of 10 boys has the I.Q's 70, 120, 110,101, 88, 83, 95, 98, 107 and 100. Test the mean I.Q of the students is 100 at 0.05 level of significance.

3. The heights of 10 males of a given locality are found to be 70,67,62,68,61,68,70,64,64,66 inches. Is it reasonable to believe that the average height is greater than 64 inches? Test at 5% significance level assuming that for 9 degrees of freedom $P(t>1.83) = 0.05$

4. Two independent samples of 8 and 7 items respectively had the following values:

   Sample1: 9   11   13   11   15   9   12   14
   Sample2: 10   12   10   14   9   8   10
   Is the difference between the means of samples significant? Test at 1% level of significance.

5. A pair of 27 sample observations has the correlation coefficient 0.67. Test the hypothesis that the correlation coefficient of the population at 95% confidence level.

Tutorial Problems-X : (Small Sample Tests) (Chi-square & F-Test)

1. The following figures show the distribution of digits in numbers chosen at random from a telephone directory.

Digits:	0	1	2	3	4	5	6	7	8	9	Total
Frequency:	1026	1107	997	966	1075	933	1107	972	964	853	10000

   Test whether the digits in the directory are equal frequently occurs at 0.05 level.

2. A survey of 320 families with 5 children each, revealed the following distribution. Is the result consistent with the hypothesis that male and female births are equally probable at 0.01 significance level?

No. of Boys	5	4	3	2	1	0
No. of Girls	0	1	2	3	4	5
No. of families	14	56	110	88	40	12

3. Fit a Poisson distribution to the following data and test the goodness of fit:

No. of accidents:	0	1	2	3	4	5	6
No. of days :	150	65	45	34	10	6	2

4. Two random samples gave the following results.

Sample	size	Sample mean	Sum of squares of Deviations from the mean
1	12	14	108
2	10	15	90

Test whether the two independent samples have been drawn from the same normal population (i) the equality of population means (ii) the equality of population variances.

5. Obtain the equation of the normal curve that may be fitted to the data and test the goodness of fit.

Class:	60-65	65-70	70-75	75-80	80-85	85-90	90-95	95-100
Frequency:	3	21	150	335	326	135	26	4