VIT-AP UNIVERSITY
Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: Fall 2023-24 Semester – CAT-2

Course Code: MAT1011 **Course Title: Applied Statistics**

Set number: 01 **Date of Exam:** 19/10/2023 (FN)(E1)

Duration: 90 Min **Total Marks: 50**

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. A coffee machine is regulated so that it discharges an average of 200 millilitres per cup. If the amount of drink is normally distributed with a standard deviation equal to 15 millilitres, then what is the probability that the cup contains 190 to 210 millilitres. **[08M]**

Q1. The mean and standard deviations sales are 40, 10 and the mean and standard deviation of advertisement expenditures are 4 and 1.5 respectively. When the coefficient of correlation, r = 0.9. Now, calculate the regression coefficient and obtain the lines of regression. **[12M]**

Q4. Find the Spearman Rank Correlation coefficients for the following data set **[10M]**

X (Maths)	78	89	69	59	79	68	62
Y (Physics)	121	72	60	81	87	123	92

Q3. A sample of 10 students from a school has the following scores in an I.Q. test. 89,87,77, 79,83,74,83,75,76 and 90. Do this data support that the mean I.Q. mark of the school students is 80? Test at 5% level of significance. **[10M]**

Q.5 We want to test if the mean daily wage of a party workers is equal to 225 rupees, given that the population standard deviation is 30 rupees. We assume that the daily wages follow a normal distribution. We take a random sample of 40 workers and find that their mean daily wage is 220 rupees. We use a significance level of 0.05 for our hypothesis test. **[10M]**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	3	1,2			08
Q2	4	4	1,2			12
Q3	4	4	1,2			10
Q4	5	5	1,2			10
Q5	5	5	3			10

VIT-AP UNIVERSITY
Apply Knowledge. Improve Life!

Name of the Examination: Fall 2023-24 Semester – CAT-2

Course Code: MAT1011 **Course Title: Applied Statistics**

Set number: 03 **Date of Exam:** 17/10/2023 (AN)(C4)

Duration:90 Min **Total Marks: 50**

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. A certain machine makes electrical resistors having a mean resistance of 40 ohms and a standard deviation of 2 ohms. Assuming that the resistance follows a normal distribution and can be measured to any degree of accuracy,

(a) what percentage of resistors will have a resistance exceeding 43 ohms? and
(b) What will be the strength of the resistors that are less than 10%? **(10M)**

Q2. A study was made by a retail merchant to determine the relation between weekly advertising expenditures and sales.

Advertising Costs ($)	Sales ($)
40	385
20	400
30	395
20	365
50	475

(a) Compute and interpret the sample correlation coefficient between weekly advertising expenditures and sales.
(b) Find the equation of the regression line to predict weekly sales from advertising expenditures. Also estimate the weekly sales when advertising costs are $35. **(20M)**

Q3. Suppose the hospitality of wary otter pops has a population that is normally distributed with a standard deviation of 7. Your friend gets you to sample 58 wary otter pops from this population and obtain a mean hospitality of 65.29 and a standard deviation of 6.0893. Using $\alpha = 0.01$, is this observed mean significantly different than an expected hospitality of 66? **(10M)**

Q4. Based on field experiments, a new variety green gram is expected to give a yield of 12.0 quintals per hectare. The variety was tested on 10 randomly selected farmers' fields. The yield (quintals/hectare) were recorded as 14.3, 12.6, 13.7, 10.9, 13.7, 12.0, 11.4, 12.0, 12.6, 13.1. Do the results conform the expectation at 5% level of significance? **(10M)**

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: Fall 2023-24 Semester – CAT-2

Course Code: MAT 1011 **Course Title: Applied Statistics**

Set number: 05 **Date of Exam:** 19/10/2023 (AN)(E2)

Duration: 90 min **Total Marks: 50**

Instructions:

1. Any assumptions made should be clearly stated.
2. Students are allowed to use Normal distribution tables and t-table endorsed by CoE.

Q1

The average number of acres burned by forest and range fires in a large New Mexico county is 4,300 acres per year, with a standard deviation of 750 acres. The distribution of the number of acres burned is normal. What is the probability that between 2,500 and 4,200 acres will be burned in any given year?

Marks 10

Q2

Find the correlation coefficient for below data for adjacent image pixels (X and Y) and find the regression equation of Y on X.

X	10	12	16	11	15	14	20	22
Y	15	18	23	14	20	17	25	28

Marks 20

Q3

Suppose the courage of psychologists has a population that is normally distributed with a standard deviation of 10. You decide to sample 57 psychologists from this population and obtain a mean courage of 34.81 and a standard deviation of 9.0579. Using an alpha value of $\alpha = 0.05$. Is this observed mean significantly greater than an expected courage of 34?

Marks 10

Q4

A visual research lab has purchased a digital colour blindness test from a company. Before they can use the test in their research, they must ensure it is properly calibrated. To do this they must check that they get the same results as the company when testing participants with no colour deficiencies. The company states that participants with healthy colour vision will score 15 on the test on average. The research lab tests 13 participants with healthy colour vision. On average they score 12 with a standard deviation of 3.6. Is their machine properly calibrated?

Marks 10

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	3	1			10
Q2	4	4	3			20
Q3	5	5	1			10
Q4	5	5	2			10

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!®

QUESTION PAPER

Name of the Examination: Fall 2023-24 Semester – CAT-2

Course Code: MAT 1011 **Course Title: Applied Statistics**

Set number: 06 **Date of Exam:** 17/10/2023 (AN) (C2)

Duration: 90 min **Total Marks: 50**

Instructions:

1. Any assumptions made should be clearly stated.
2. Students are allowed to use Normal distribution tables and t-table endorsed by CoE.

Q1

The length of human pregnancies from conception to birth approximates a normal distribution with a mean of 266 days and a standard deviation of 16 days. What proportion of all pregnancies will last between 240 and 270 days (roughly between 8 and 9 months)?

Marks 10

Q2

Find Karl Pearson's coefficient of correlation between capital employed and profit obtained from the following data and find the equation of true regression line.

Capital Employed (Rs. In Crore)	10	20	30	40	50	60	70	80	90	100
Profit (Rs. In Crore)	2	4	8	5	10	15	14	20	22	50

Marks 20

Q3

Suppose the arousal of hot cats has a population that is normally distributed with a standard deviation of 6. Tomorrow you sample 40 hot cats from this population and obtain a mean arousal of 46.44 and a standard deviation of 5.6968. Using an alpha value of $\alpha = 0.01$, is this observed mean significantly less than the expected arousal of 47?

Marks 10

Q4

Suppose the jewellery of exams has a population that is normally distributed with a standard deviation of 5. You are walking down the street and sample 9 exams from this population and obtain a mean jewellery of 28.95 and a standard deviation of 6.3802. Using an alpha value of $\alpha = 0.01$, is this observed mean significantly different than expected jewellery of 27?

Marks 10

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	3	1			10
Q2	4	4	3			20
Q3	5	5	1			10
Q4	5	5	2			10

VIT-AP UNIVERSITY

Name of the Examination: WINTER 2022-23-CAT-2

Course Code: MAT 1011
Set Number: 2
Duration: 90 minutes

Course Title: Applied Statistics
Date of Examination:
Total Marks: 50 Date: 30/03/2023 (FN) (D1)

1. A manufacture of cotter clips know that 5% of his product is defective. If he sells clips in boxes of 100 and guarantees that not more than 3 clips will be defective. Use Poisson distribution to find an approximate calculation.
   (a) What is the probability that a box will meet the guaranteed quality?
   (b) What is the probability that a box will fail to meet the guaranteed quality? **6+3=9**

2. The resistance in milliohms of a copper cable at a certain temperature is normally distributed with mean 23.8 and variance 1.28.
   (a) What is the probability that a copper cable has a resistance between 24.2 and 24.5?
   (b) What is the resistance level that has a probability 0.95? **5+4=9**

3. Compute and interpret the rank correlation coefficient for the following advertising cost and product sales (in lakhs) of 8 companies are selected at random: **8**

Advertising cost	15	20	28	12	80	20	60	40
Product sales	40	30	50	30	60	30	10	20

4. IC chips production (in thousands) and electricity consumption (in KW/day) for a particular week of the company are given as follows: **12+2=14**

IC chips production	451	358	431	506	499	529	564
Electricity consumption	248	226	247	277	305	583	299

   (a) Estimate the linear regression line.
   (b) Estimate the electricity consumption for the day that company produces 4, 95, 000 IC chips.

5. An urban community would like to show that the incidence of breast cancer is higher in their area than in a nearby rural area. If it is found that 20 of 200 adult women in the urban community have breast cancer and 10 of 150 adult women in the rural community have breast cancer. Can we conclude at the 0.05 level of significance that breast cancer is more prevalent in the urban community? **10**

QP Mapping

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
1	3	3	1, 3			9
2	3	3	1, 3			9
3	4	5	3			8
4	4	5	3			14
5	5	4	1, 2			10

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: CAT-2 (WINTER 2022-2023)

Course Code: MAT1011

Course Title: Applied Statistics

Set number: 6

Date of Exam: 29/03/2023 (FN)

Duration: 90 min

Total Marks: 50 (CL)

Q1. What probability model is appropriate to describe a situation where 100 misprints are distributed randomly throughout the 100 pages of a book? For this model, what is the probability that a page observed at random will contain at least 3 misprints? **(10M)**

Q2. The hourly wages of 1,000 workmen are normally distributed around an average of Rs. 70 and with a standard deviation of Rs. 5. Estimate the number of workers whose hourly wages will be (i) between Rs. 69 and Rs. 72 (ii) more than Rs. 75. **(10M)**

Q3. The following table gives the normal weight of a baby during the first six months of life. Find two regression lines and also estimate the weight of a baby at the age of 4 months. **(10M)**

Age in months	0	2	3	5	6
Weight in lbs	5	7	8	10	12

Q4. The owner of a small garment shop is hopeful that his sales are significantly rising week by week. Treating the sales for the previous 5 weeks as a typical example of this rising trend, he recorded them in Rs. 1000's and analysed the results in fitting a second-degree parabola. **(10M)**

Week	1	2	3	4	5
Sales	6	17	34	57	86

Q5. Before an increase in excise duty on tea, 400 people out of a sample of 500 persons were found to be tea drinkers. After an increase in the duty, 400 persons were known to be tea drinkers in a sample of 600 people. Do you think that there has been a significant decrease in the consumption of tea after the increase in the excise duty? **(10M)**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	3	1,2			10
Q2	3	3	1,2			10
Q3	4	5	1,2			10
Q4	4	5	1,2			10
Q5	5	6	1,6			10

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!

QUESTION PAPER

Winter Semester 2022-23: CAT-2

Course Code: MATH1011 Course Name: Applied Statistics

Set Number: 10 Date of Exam: 03/04/2023 (G1) (FN)

Duration: 90 minutes Maximum marks: 50

1. Suppose that a company produces motor vehicle components, out of which 2% come out to be defective. If the same company receives the order to produce 80 components, what is the probability that 3 defective items will be produced? (Solve this problem using both binomial and Poisson distribution.)

(10 marks)

2. The marks of 1000 applicants in an engineering entrance test are normally distributed, with a mean of 125 and standard deviation of 15. If a university decides to consider the applicants who scored at least 100 in the test, how many of the students will be rejected? (Note that marks are recorded to the nearest integers, that is, if an applicant scores 74.6, it will be taken as 75.)

(10 marks)

3. An NGO conducted a survery on job satisfaction and happiness by asking participants to rate these parameters on a scale of 1 to 10. The following data is the collected data for 10 participants.

job satisfaction	5	6	2	1	6	10	4	9	3	7
happiness	6	7	4	3	5	8	7	8	2	9

   (a) Compute the correlation coefficient between their job satisfaction and happiness. Conclude whether the correlation is strong, moderate or weak. **(10 marks)**

   (b) Obtain both regression line equations. Predict the happiness score (on a scale of 1 to 10) of a person whose job satisfaction score is 8. **(10 marks)**

4. In an interview for cricket coach selection, a person claims that if he is selected, he will maintain a 70% winning record. Would you agree with his claim if out of the last 45 matches played under him, teams have won 25 matches. Use a 0.10 level of significance.

(10 marks)

QP Mapping

Question No.	Module No.	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	CO3	PO1,PO2			10
Q2	3	CO3	PO1,PO2			10
Q3	4	CO5	PO1,PO2			20
Q4	5	CO6	PO1,PO6			10

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!

Name of the Examination: WINTER 2022-2023-CAT-2

Course Code: MAT1011

Course Title: Applied Statistics

Set number: 12

Date of Exam: 28/03/2023 (FN) (B1)

Duration: 90 min

Total marks: 50

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

1. The number of cracks in a ceramic tile has a Poisson distribution with a mean of $\lambda = 2.4$.
   (a) What is the probability that a tile has no cracks?
   (b) What is the probability that a tile has four or more cracks?
   (c) What is the probability that at least one tile has crack? **(10 marks)**

2. The breaking strengths of nylon fibers in dynes are normally distributed with a mean of 12,500 and a variance of 200,000.
   (a) What is the probability that a fiber strength is more than 13,000?
   (b) What is the 95th percentile of the fiber strengths? **(10 marks)**

3. In a biology experiment a number of cultures were grown in the laboratory. The numbers of bacteria, in millions, and their ages, in days, are given below

Age (x)	1	2	3	4	5
No.of Bacteria (y)	35	110	130	175	195

   Calculate and interpret the correlation coefficient. **(10 marks)**

4. The heights and weights of a sample of 5 students are:

Height (x)	1.35	1.44	1.53	1.59	1.6
Weight (y)	50	49	66	62	68

   Calculate the regression line. Estimate the weight of student whose height is 1.55. **(10 marks)**

5. A fuel oil company claims that one-fifth of the homes in a certain city are heated by oil. Do we have reason to believe that fewer than one-fifth are heated by oil if, in a random sample of 1000 homes in this city, 136 are heated by oil? Use 5% level of significance **(10 marks)**

QP Mapping

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
1	3	3	1,2			10
2	3	3	1,2			10
3	4	4	3			10
4	4	4	3			10
5	5	5	1,2			10

VIT-AP UNIVERSITY

Apply Knowledge, Improve Life!

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – CAT-2

Course Code: MAT1011 **Course Title:** Applied Statistics

Set number: 13 **Date of Exam:** 27/03/2023 (FN) (A1)

Duration: 90 minutes **Total Marks:** 50

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. According to *Chemical Engineering Progress* (November 1990), approximately 30% of all pipework failures in chemical plants are caused by operator error. **(5M+5M)**

(a) What is the probability that that out of the random sample of 20 such failures, exactly 5 are due to operator error?

(b) What is the probability that no more than 4 out of 20 such failures are due to operator error?

Q2. A lawyer commutes daily from his suburban home to his midtown office. The average time for a one-way trip is 24 minutes, with a standard deviation of 3.8 minutes. Assume the distribution of trip times to be normally distributed.

(a) What is the probability that a trip will take at least 1/2 hour?

(b) If the office opens at 9:00 A.M. and the lawyer leaves his house at 8:45 A.M. daily, what percentage of the time is he late for work? **(5M+5M)**

Q3. Marks of the 10 randomly chosen students in two different subjects mathematics and statistics are given below:

Mathematics	68	64	75	50	64	80	75	40	55	64
Statistics	62	58	68	45	81	60	68	48	50	70

Obtain the rank correlation coefficient for the given data. **(10M)**

Q4. The following are measurements of the air velocity and evaporation coefficient of burning fuel droplets in an impulse engine:

Air velocity (cm/s) x	20	60	100	140	180
Evaporation coefficient (mm^2) y	0.18	0.37	0.35	0.78	0.56

Fit a straight line to these data by the method of least squares, and use it to estimate the evaporation coefficient of a droplet when the air velocity is 150 cm/s. **(10M)**

Q5. A new radar device is being considered for a certain missile defense system. The system is checked by experimenting with aircraft in which a kill or a no kill is simulated. If, in 300 trials, 250 kills occur, accept or reject, at the 0.05 level of significance, the claim that the probability of a kill with the new system does not exceed the 0.8 probability of the existing device. **(10M)**

VIT-AP UNIVERSITY

Apply Knowledge, Improve Life!

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – CAT-2

Course Code: MAT1011 | **Course Title: Applied Statistics**

Set number: I | **Date of Exam:** 03/04/2023 (AN)

Duration: 90 min | **Total Marks: 50** (G2)

Instructions: -

1) The question paper contains five questions.

2) All questions are compulsory.

3) The number of marks carried by a question/part is indicated against it.

Q.1: A gear manufacturing company expects that the chance of a gear being defective is $\frac{1}{200}$. The gears are supplied in boxes of 10 gears. Find the probability that there are two defective gears in a box of 10 gears. Also calculate the number of boxes containing two defective pieces out of 10000 boxes. **(10 M)**

Q.2: A soft drink machine is regulated so that it discharges an average of 200 millilitres per cup. If the amount of drink follows Gaussian distribution with a variance equal to 225,

i) What is probability that a cup contains between 191 to 209 millilitres?
ii) How many cups will probably overflow if 230 millilitre cups are used for the next 1000 drinks?
iii) Below what value do we get the smallest 25% of the drinks?

(10 M)

Q.3: The data on length of panicle (in cm.) and number of grains per panicle of PBW-373 wheat variety is given below:

Length of panicle (X)	7.5	8.5	7.2	9.5	6.4	8.3	10.0	6.8	7.1	7.7
No. of Grains (Y)	34	36	28	46	20	28	50	28	25	25

Calculate the correlation coefficient (up to two decimal) between X and Y. **(10 M)**

Q.4: Two variables have the regression lines

$$3x + 2y = 26 \text{ and } 6x + y = 31$$

Find the correlation coefficient between X and Y variables. Also obtain mean values of X and Y.

(10 M)

Q.5: In a random sample from 100 persons from SAS, 60 are found to be high consumers of Coffee. In another sample of 80 people from SCOPE, 40 are found to be high consumers of Cofee. Does this data reveal a significant difference between the proportions of high wheat consumers in SAS and SCOPE? ($\alpha = 0.05$)

(10 M)

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	Marks
Q1	3	CO3	PO1, PO2	10
Q2	3	CO3	PO1, PO2	10
Q3	4	CO4	PO3	10
Q4	4	CO5	PO1, PO2	10
Q5	5	CO6	PO1,PO6	10

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – CAT-2

Course Code: MAT1011 | **Course Title: Applied Statistics**

Set number: 5 | **Date of Exam:** 29/03/2023 (AN)

Duration: 90 minutes | **Total Marks: 50** (C2)

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. According to a study published by a group of Indian sociologists, approximately 60% of the Valium (a drug used to relieve anxiety) users in our country first took Valium for psychological problems. Find the probability that among the next 8 users from our country who are interviewed,

(a) exactly 3 began taking Valium for psychological problems;

(b) at least 5 began taking Valium for problems that were not psychological. **(10 Marks)**

Q2. The weights of a large number of pugs are approximately normally distributed with a mean of 8 kilograms and a standard deviation of 0.9 kilogram. If measurements are noted for the weights of these dogs, determine the fraction of these pugs having weights

(a) greater than 9.55 kilograms

(b) between 7.25 to 9.15 kilograms (inclusive). **(10 Marks)**

Q3. The price and quantity of five items in a particular shop are as follows:

Price (Rs.)	10	12	14	16	18
Quantity (Units)	20	29	21	22	28

Calculate the Karl Pearson's correlation coefficient and interpret the result. **(10 Marks)**

Q4. An analysis was performed to understand the effect of temperature on the electric power consumed by a nuclear plant in a city of Karnataka. Considering the other factors as constant, the data collected have been tabulated below:

Temperature (^{o}F)	27	45	72	58	31	60	34	74
Electric power (BTU)	250	285	320	295	265	298	267	321

(a) Fit a simple linear regression model between temperature and electric power by finding out the estimates of intercept and slope.

(b) Predict the power consumption for a temperature of 65 ^{o}F. **(10 Marks)**

Q5. Due to the outbreak of a certain type of flu in the month of December 2022, the parents of 2000 babies were surveyed by scientists at Synergy Pharmaceuticals, Vijayawada to determine if the company's new medicine was effective after two days. Among 120 babies who had the flu and were given the medicine, 29 were cured within two days. Among 280 babies who had the flu but were not given the medicine, 56 recovered within two days. Is there any significant indication that supports that the company's claim of the effectiveness of the medicine? Use 5% level of significance. **(10 Marks)**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	CO3	PO1, PO2			10
Q2	3	CO3	PO1, PO2			10
Q3	4	CO5	PO1, PO2			10
Q4	4	CO5	PO1, PO2			10
Q5	5	CO4	PO3			10

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – CAT-2

Course Code: MAT 1011

Course Title: Applied Statistics

Set number: 7

Date of Exam: 31/03/2023 (FN)

Duration: 90 min

Total Marks: 50 (E1)

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. A manufacturer produces light-bulbs that are packed into boxes of 100. If quality control studies indicate that 0.5% of the light-bulbs produced are defective, what percentage of the boxes will contain:

(a) no defective? (b) 2 or more defectives?

(5M+5M)

Q2. The length of life of an instrument produced by a machine has a normal distribution with a mean of 12 months and standard deviation of 2 months. Find the probability that an instrument produced by this machine will last

a) less than 7 months.
b) between 7 and 12 months.

(5M+5M)

Q3. From following information find the correlation coefficient between advertisement expenses and sales volume using Karl Pearson's coefficient of correlation method.

Firm	A	B	C	D	E	F	G	H	I	J
Advertise Exp (In Lakhs	11	13	14	16	16	15	15	14	13	13
Sales Volume (In Lakhs)	50	50	55	60	65	65	65	60	60	50

(10M)

Q4. The following data gives the age and blood pressure (BP) of 10 sports persons.

Age (X)	42	36	55	58	35	65	60	50	48	51
BP (Y)	98	93	110	85	105	108	82	102	118	99

Find the linear regression equation of Y on X and estimate the Blood Pressure of a sport person whose age is 45 years.

(10M)

Q5. Suppose the pharma company develops a new drug, designed to prevent colds. The company states that the drug is equally effective for men and women. To test this claim, they choose a simple random sample of 100 women and 200 men from a population of 100,000 volunteers. At the end of the study, 38% of the women caught a cold; and 51% of the men caught a cold. Based on these findings, can we reject the company's claim that the drug is equally effective for men and women? Use a 0.05 level of significance.

(10M)

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	Marks
Q1	3	3	1,2	10
Q2	3	3	1,2	10
Q3	4	4	3	10
Q4	4	4	3	10
Q5	5	5	1,2	10

VIT-AP UNIVERSITY
Apply Knowledge. Improve Life!

Question Paper

Name of the Examination: WIN 2022-2023 – CAT II

Course Code: (MAT1011) Course Title: Applied Statistics

Set number: 8 Date of Exam: 27/03/2023

Time: 90 Mins Max. Marks: 50 (A1) (A2)

Note: All questions are compulsory

1. According to UP state health department, viral fever in Prayag city is its peak in these days. Let the total number $n = 500$ people in the city has a viral infection with the probability $p = \frac{1}{10^3}$. The best known method to detect the infection is the blood test. To make procedure fast blood samples from all $n = 500$ people are pooled (that is mixed up) and tested.
(a) What is the probability under this circumstance, that more than one person has the infection?
(b) Which probability model will be the best explanation of above problem? Which probability model will be best fit for above example if we take $n = 5$ and $p = \frac{1}{5}$ (Justify your answer)?

(10)

2. A randomly chosen IQ test taken obtains a score that is approximately a normal random variable with mean 100 and standard deviation 15. What is the probability that the score of such a person is

(a)Between 90 to 100?

(b) For what value of b probability $P(X < b) = 0.367$. (5+5)

3. A professor in the school of mathematics in a university polled a dozen colleagues about the number of academic workshops they attended in the previous five years (x) and the number of papers they submitted to refereed journals (y) during the same period. The summary of the data are given as follows:

$$n = 12; \bar{x} = 6; \bar{y} = 14$$

$$\sum_{i=1}^{n} x_i^2 = 280; \sum_{i=1}^{n} x_i y_i = 340$$

Fit a simple linear regression model between x and y by finding out the estimates of the intercept and slope. Comment on whether attending more academic workshops was likely to result in publishing more papers. If $\sum_{i=1}^{n} y_i^2 = 240$, then determine the Karl Pearson's correlation coefficient. **(10)**

4. Calculate the Spearman's Rank correlation coefficient from the following data of women beauty context between India and Iran.

India (X)	Iran(Y)
32	40
55	30
49	70
60	20
43	30
37	50
43	72
49	60
10	45
20	25

(10)

5. In "Cavery Bachao Abhiyaan" Government of India has conducted a survey in December, 2022 for checking heavy metal level in the Cavery water. Survey commeetee claims that 70% Cavery water contains heavy metal. Would you agree with this claim if a random survey of Cavery water in different city showed that 8 out of 15 had heavy metal in it? Use a $\alpha = 0.10$ level of significance. **(10)**

Q.No	Module No	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
1	3	3	1,2			10
2	3	3	1,2			10
3	4	4	1,2			10
4	4	4	1,2			10
4	5	5	3			10

VIT-AP UNIVERSITY

Name of the Examination: WINTER 2022-23-CAT-2

Course Code: MAT 1011
Set Number: 9
Duration: 90 minutes

Course Title: Applied Statistics
Date of Examination: 30/03/2023 (AN)
Total Marks: 50

(02)

1. The peak temperature T, in degrees Fahrenheit, on a July day in Antarctica is a normal random variable with a variance of 225. With probability 1/2, the temperature T exceeds 10 degrees.
   (a) What is $P[T > 32]$, the probability the temperature is above freezing?
   (b) What is $P[T < 0]$?
   (c) What is $P[40 < T < 50]$?

   3+2+2+3=10

2. The number of buses that arrive at a bus stop in T minutes is a Poisson random variable B with expected value $T/5$.
   (a) What is the PMF of B, the number of buses that arrive in T minutes?
   (b) What is the probability that in a two-minute interval, three buses will arrive?
   (c) What is the probability of no buses arriving in a 10-minute interval?
   (d) How much time should you allow so that with probability 0.99 at least one bus arrives?

   1+3+3+3=10

3. Compute and interpret the correlation coefficient for the following marks of 6 students selected at random:

Mathematics mark	41	45	49	50	38	40
English mark	35	42	39	41	36	42

   8

4. The marks of a class of 6 students on a midterm report (x) and on the final examination (y) are as follows:

x	71	85	89	90	68	70
y	75	92	79	91	86	72

   (a) Estimate the linear regression line.
   (b) Estimate the final examination mark of a student who received a mark of 93 on the midterm.

   10+2=12

5. A vote is to be taken among the residents of a town and the surrounding county to determine whether a proposed chemical plant should be constructed. The construction site is within the town limits, and for this reason many voters in the county believe that the proposal will pass because of the large proportion of town voters who favor the construction. To determine if there is a significant difference

in the proportions of town voters and county voters favoring the proposal, a poll is taken. If 120 of 200 town voters favor the proposal and 240 of 500 county residents favor it, would you agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters? Use an $\alpha = 0.05$ level of significance.

10

QP Mapping

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
1	3	3	1, 2			10
2	3	3	1, 2			10
3	4	4	3			8
4	4	4	1, 2			12
5	5	5	1, 2			10

VIT-AP UNIVERSITY

Apply Knowledge. Improve Life!

Name of the Examination: WINTER 2022-2023 – CAT-2

Course Code: MAT1011 | **Course Title: Applied Statistics**

Set number: II | **Date of Exam:** 28/03/2023 (AN) (B2)

Duration: 90 Mins | **Total Marks: 50**

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. In a certain city district, the need for money to buy drugs is stated as the reason for 75% of all thefts. Find the probability that among the next 5 theft cases reported in this district,

a) exactly 2 resulted from the need for money to buy drugs; **(5M)**

b) at most 3 resulted from the need for money to buy drugs. **(5M)**

Q2. The annual salaries of employees in a large company are approximately normally distributed with a mean of $50,000 and a standard deviation of $20,000.

a) What percent of people earn less than $40,000? **(3M)**

b) What percent of people earn between $45,000 and $65,000? **(4M)**

c) What percent of people earn more than $70,000? **(3M)**

Q3. The scores for nine students in history and algebra are as follows:

History: 35, 23, 47, 17, 10, 43, 9, 6, 28

Algebra: 30, 33, 45, 23, 8, 49, 12, 4, 31

Compute the Spearman rank correlation and use your result to make a conclusion about the students' scores in these two subjects. **(10M)**

Q4. The following table shows the sales of a company (in million dollars). Estimate the sales using the regression line in the year 2020 using method of least squares. **(10 M)**

x	2015	2016	2017	2018	2019
y	12	19	29	37	45

Q5. A researcher claims that Republican Party will win in the next Senate elections, especially in Florida State. Statistical data reported that 23% voted for Republican Party in the last election. To test the claim a researcher surveyed 80 people and found 22 said they voted for Republican Party in the last election. Is there enough evidence at α=0.05 to support this claim? **(10M)**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	CO3	PO1, PO2			10
Q2	3	CO3	PO1, PO2			10
Q3	4	CO4	PO3			10
Q4	4	CO4	PO3			10
Q5	5	CO5	PO1, PO2			10