KATHMANDU UNIVERSITY
End Semester Examination
February, 2025

Marks Scored:

Level : B.E./B.Sc./B.Tech.
Year : II

Course : MATH 208
Semester : II

Exam Roll No. :

Time: 30 mins.

F. M. : 20

Registration No.:

Date : 24 FEB 2025

SECTION "A"
[10Q. × 1 = 10 marks]

Fill in the blank space(s) by the most appropriate word(s) or symbol(s).

1. Half of the difference between the third quartile and the first quartile is used as a measure of dispersion in data. It is called _____.

2. A lot contains 7 good and 3 bad items. An item is selected and is found to be good. If another item is selected without replacing the first item selected, then the probability that the second item selected is good _____.

3. If A and B are any two events defined on the same sample space, then probability of occurrence of A only without B is dented as $P(A \cap \bar{B})$ and it is equal to _____.

4. If $x's$ are values assumed by a discrete random variable X having probability mass function $p(x)$, then expectation of X is given by $E(X) =$ _____.

5. Poisson distribution can be considered as a special case of binomial distribution when (i) number of trials (n) is infinite and (ii) the probability of success (p) is infinitesimally small and (iii) _____.

6. If $V(X) = 2.4$, then $V(2X) =$ _____.

7. t-distribution is preferred over Z- distribution to carry inference on mean of a normal population when population variance is unknown and_____.

8. In test of hypothesis process, the complementary statement that is accepted and the null hypothesis is observed to be false is called _____.

9. In a production process, the variations resulted from many minor causes that behave in a random manner which are beyond the control of human hand and which cannot be eliminated are called _____ cause of variation.

10. The main objectives of regression analysis are (i) observe of change in one variable with respect to another (ii) develop model describing relation between two or more variables and (iii) _____

SECTION "B"

[10 Q. × 1 = 10 marks]

Fill in the blank space(s) by selecting the most appropriate answer from among the given ones. ***(DO NOT TICK THE ANSWER)***

11. The most repeated value in a data distribution is seldom used as a measure of central value and it is called _____.
[mean median mode weighted mean]

12. If A and B are _____ events, then $P(A \cap B) = P(A).P(B)$
[mutually exclusive independent dependent equally likely]

13. The set of all possible outcomes of a random experiment is called its _____.
[event trial sample space probability]

14. The probability that a discrete random variable X takes a particular value x is a function of x. It is denoted as $p(x)$ and is called _____function.
[probability mass probability density
probability distribution cumulative distribution]

15. A researcher is interested in the study of time to wait for prescription to be prepared at a clinic. She can achieve her study by using _____ distribution.
[binomial Poisson normal chi-square]

16. If $Y = 2X$, where $X \sim N(15, 9)$, then $Y \sim N($_____, _____$)$.
[$N(30, 9)$ $N(30, 18)$ $N(30, 36)$ $N(30, 81)$]

17. The expected value of sample mean is equal to _____
[sample variance sample standard deviation
population mean population variance]

18. For a left tailed test, the null hypothesis is rejected at 5% level of significance if calculated values of test statistic is _____
[> 1.64 < 1.64 > -1.64 < -1.64]

19. The control chart for proportion of defectives is called _____
[R-chart c-chart p-chart d-chart]

20. The standardized measure of inter-relationship between two variables is called _____.
[correlation regression standard deviation covariance]

KATHMANDU UNIVERSITY
End Semester Examination
February, 2025

Level : B.E./B.Sc./B.Tech.
Year : II
Time : 2 hrs. 30mins.

24 FEB 2025

Course : MATH 208
Semester : II
F. M. : 55

SECTION "C"
[3 Q. × 7 = 21 marks]

1. Following data gives motor fuel octane ratings of several blends of gasoline:

65.2	87.7	78.3	115.2	87.5	124.0	88.6	100.3	63.4	93.3
78.3	91.1	75.6	94.2	106.3	89.9	127.3	87.6	68.1	86.7
76.9	90.8	100.5	98.9	112.8	92.7	69.5	91	67.8	93.4
116.1	90.1	121.9	88.3	75.5	87.9	100.8	90.9	112.5	96.1
68.3	91.8	127.1	90.4	111.8	93.0	67.3	89.9	79.2	89.6
67.5	88.4	72.5	91.2	118.6	94.4	129.4	91.8	111.1	90.4

a. Construct a stem-and-leaf plot for above data. [2]
b. Use the stem-and-leaf plot and part (a) to construct a frequency distribution with class interval of 10. [1]
c. Use the frequency distribution in part (b) to calculate the mean and variance. [3]
d. Calculate coefficient of variation using following formula [1]

$$C.V. = \frac{s.d.}{mean}$$

2. The speed of cars, in miles per hour and the distances covered to stop, in feet are given in table below-

Speed (x)	8	10	12	14	16	18	20	22	24
Distance (y)	20	26	28	32	36	42	50	62	75

a. Draw a scatterplot to show that there is nearly linear relationship between speed of cars and stopping distance. [1]
b. Calculate values of sum of squares of 'x', sum of squares of 'y' and sum of cross-product of 'x' and 'y' by using [3]

$$S_{xx} = \sum x^2 - \frac{1}{n}\left(\sum x\right)^2, S_{yy} = \sum y^2 - \frac{1}{n}\left(\sum y\right)^2, S_{xy} = \sum xy - \frac{1}{n}\sum x \sum y$$

c. Develop linear regression equation of 'y' on 'x'. [2]
d. Calculate Pearson's correlation coefficient between 'x' and 'y'. [1]

P.T.O.

3. What are different errors that may be associated with test of hypothesis problems? The following random samples are measurements of the heat-producing capacity (in millions of calories per ton) of specimens of coal from two mines:

Mine 1:	826	813	835	807	834	
Mine 2:	795	789	790	814	792	784

Assuming that the population variances of heat-producing capacity of coals from the two mines to be equal, test whether the difference between the means of two samples is significant at 0.05 level of significance. Also obtain 95% confidence interval for difference in true means of two populations from which above samples are drawn.

OR

10 cars were equipped with radial tires and driven over a test course. Then the same 10 cars (with the same drivers) were equipped with regular belted tires and driven over the same course. After each run, the cars' fuel economy (in km/l) was measured. The results are presented below-

Radial tire	4.2	4.7	6.6	7.0	6.8	4.5	5.7	6.0	7.4	4.9
Belted tire	4.1	4.9	6.2	6.9	6.8	4.4	5.7	5.8	6.9	4.7

Is there evidence that radial tires produce better fuel economy? (Assume normality of data, and use paired t-test at $\alpha = 0.05$.) Also, find a 95% confidence interval for the mean difference in fuel economy between radial tires and belted tires.

SECTION "D"

[6Q. × 4 = 24 marks]

4. It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam email detected as spam) is 5%. Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?

5. In a survey, the question, "Do you smoke?" was asked to 100 people. Results are summarized in the following table:

	Yes	No
Male	19	41
Female	12	28

An individual is chosen from them at random. a) What is the probability that the selected individual smokes? b) What is the probability that the selected individual is a male and smokes? c) What is the probability that the selected individual is a male? d) What is the probability of selected individual being a smoker if he was found to be a male?

24 FEB 2025

6. Let X be continuous random variable with probability density function

$$f(x) = \frac{x^3}{4}$$

for an interval $0 < x < c$. What is the value of the constant c that makes $f(x)$ a valid probability density function?

7. Suppose that the data concerning the first-year salaries of engineers is normally distributed with the population mean μ = Rs. 60000 and the population standard deviation σ = Rs.15000. (a) Find the probability of a randomly selected engineer earning is (i) less than Rs.45000 annually. (ii) more than Rs.80000 a year (b) Find the range of annual salaries of the top 15% earners.

8. A professor wants to know if his introductory Statistics class has a good grasp of basic math. Six students are chosen at random from the class and given a proficiency test. The professor wants the class to be able to score above 70 on the test. The six students got scores 62, 92, 75, 68, 83 and 95. Can the professor have 90% confidence that the mean score for the class would be above 70?

9. A team in an accounting group has been working on improving the processing of invoices. The team is trying to reduce the cost of processing invoices by decreasing the fraction of invoices with errors. The team developed the following operational definition for a defective invoice: an invoice is defective if it has incorrect price, incorrect quantity, incorrect coding, incorrect address, or incorrect name. The team decided to pull a random sample of 100 invoices per day. If the invoice had one or more errors it was defective. The data on the number of defective invoices from the last 25 days are 22, 33, 24, 20, 18, 24, 24, 29, 18, 27, 31, 26, 31, 24, 22, 22, 29, 31, 21, 26, 24, 32, 17, 25, 21. Use these data to construct p-chart and interpret the chart.

SECTION "E"

[5Q. × 2=10 marks]

10. What do (i) skewness and (ii) kurtosis of a data distribution signify?

11. A telemarketing executive has determined that 15% of people contacted will purchase the product. 12 people are contacted about this product. Find the probability that among 12 people contacted, at most 2 will buy the product?

12. Mass-produced needles are packed in boxes of 1000. It is believed that 1 needle in 2000 on average is substandard. What is the probability that a box contains 2 or more defectives by using the Poisson distribution?

13. The times, in seconds, required for packing manufactured products by two machines, namely, Machine-1 and Machine-2 for 10 units each are recorded found to have standard deviations of the packing times as 0.683 for Machine-1 and 0.750 for Machine-2. Carry F-test to observe whether packing time of Machine-2 is more varied at 5% level of significance.

14. Write short notes on R-chart.