

KATHMANDU UNIVERSITY
End Semester Examination
March/April, 2025

Marks Scored:

Level : B.E.

Year : II

Exam Roll No. :

Time: 30 mins.

Registration No.:

Course : MATH 208

Semester : I

F. M. : 20

Date 01 APR 2025

SECTION "A"

[10 Q. \times 1 = 10 marks]

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

1. For a positively skewed data the difference between third quartile and median, i.e., $Q_3 - M_d$ is greater than the difference between _____.
2. The mean and variance of a set of data are observed to be 14 and 12.25. The coefficient of variation of data is _____.
3. The collection of mutually exclusive events B_1, B_2, \dots, B_k are called _____ if their union is equal to the sample space.
4. If a discrete random variable assumes values 0, 1, 2 and 3 with $p(0) = \frac{1}{8}$; $p(1) = \frac{3}{8}$; $p(2) = \frac{3}{8}$ and $p(3) = \frac{1}{8}$, then the mean of the random variable is _____.
5. If $X \sim B(12, 0.4)$, then $P(X \leq 1) =$ _____.
6. A standard normal distribution is that form of normal distribution which has mean 0 and variance _____.
7. Different statistical values such as mean, median, variance, standard deviation, range, etc. of population are called _____.
8. Chi-square distribution is defined as the probability distribution of sum of squares of _____.
9. _____ is the method of detecting presence of assignable cause of variation in any production process.
10. The predicted value of dependent variable provided by linear model $y = 21.4 + 0.94x$ when the value of independent variable is 7.5 is _____.

SECTION "B"

[10 Q. × 1 = 10 marks]

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

11. The graphical visualization data that uses positions of minimum value, first quartile, median, third quartile and maximum value is called _____.
[stem-and-leaf plot; scatter plot; boxplot; frequency plot]
12. $P(A \cup \bar{A}) =$ _____
[0; 1; $P(A)$; $P(\bar{A})$]
13. If X is a continuous random variable with probability density function $f(x)$, then $E(X) =$ _____.
[$\int f(x) dx$; $\int x \cdot f(x) dx$; $\sum f(x)$; $\sum x \cdot f(x)$]
14. It is known that 3 requests enter into a server on average in 1 second. The probability distribution of number of requests entering into the server can be modeled by _____ distribution.
[binomial; Poisson; normal; chi-square]
15. If $X \sim N(12, 3^2)$, and $Y = X + 4$, then $Y \sim N(\underline{\hspace{2cm}}, \underline{\hspace{2cm}})$.
[$N(16,3)$; $N(16,3^2)$; $N(12,16)$; $N(16, 16)$]
16. The total number of samples of size 3 that can be drawn from a population containing 7 units by simple random sampling without replacement is _____.
[7×3 ; 7^3 ; 7C_3 ; 3^7]
17. If S^2 is the variance of a sample of size n drawn from $N(\mu, \sigma^2)$ population, then the statistic $\frac{(n-1)S^2}{\sigma^2}$ has _____ distribution.
[t-distribution; Z-distribution; chisquare-distribution; F-distribution]
18. An investigator is interested to test whether alloying reduces the resistance of electric wire significantly. Corresponding test is _____ tailed.
[left; right; two; heavy]
19. The error caused by _____ is a factor of random cause of variation.
[faulty equipment; defective raw material; un-experienced operator; minor mistakes in operation]
20. _____ is the study of dependency of one variable on one or more other variables.
[correlation; regression; covariance; variance]

KATHMANDU UNIVERSITY
End Semester Examination
March/April, 2025

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

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

SECTION "C"
[3 Q. × 7 = 21 marks]

1. A researcher is going to establish the fact Linux system is faster than Windows system. He measured time (in nanoseconds) required to execute a batch of program in 10 computers using Linux and another 10 computers of same configuration using Windows system with following result:

System	Time (nanoseconds)									
Windows	47.7	21.5	41.1	53.3	32.9	50.5	42.4	35.9	37.1	29.7
Linux	45	46.3	31	31	54.9	68.1	14.7	2.1	19.5	27.5

- a. First establish the fact that the populations from which these sample measurements are carried have equal variances using F-test at 10% level of significance. [3]
- b. After establishing that population variances are equal, test whether Linux system is significantly faster than the Windows system with respect to the mean time of execution at 5% level of significance. [4]

OR

During the test of hypothesis on difference in means of two populations, in which type of situation is paired t-test applied? Describe with an example. Suppose we want to know whether or not a certain training program is able to increase the maximum vertical jump of college basketball players. Data on maximum vertical jump, in inches, prior to the training program of 10 players and after the training program of the same players are given below-

Maximum Vertical Jump (inch)	Player										
		1	2	3	4	5	6	7	8	9	10
Before		22	20	19	24	25	25	28	22	30	27
After		24	22	19	22	28	26	28	24	30	29

Use paired t-test at 5% level of significance to determine whether the training program is effective. Also, construct 95% confidence interval for actual increase in maximum vertical jump by the training program. [1+4+2]

2. a. Explain following formula of Bayes' rule:

$$P(B_i|A) = \frac{P(B_i) \cdot P(A|B_i)}{\sum_{i=1}^k P(B_i) \cdot P(A|B_i)}$$

An insurance company has insured 4000 doctors, 8000 teachers and 12000 businessmen. The probabilities of a doctor, teacher and businessman dying before the age of 58 are estimated to be 0.01, 0.03 and 0.05, respectively. It is recently reported that one of the insured individuals died before 58, find the probability that he is a doctor. [1+3]

- b. A bag contains 7 red and 5 white balls. Two balls are drawn one after another without replacing. Find probability that (i) the second ball drawn is white given the first ball was also white, (ii) both the balls drawn are white (iii) the second ball drawn is white (whatever may the first ball be). [3]

P.T.O.

3. The number of runs made by two cricket clubs 'A' and 'B' in different randomly selected tournaments are presented below:

A	40	47	75	45	82	90	113	78	58	116
B	77	77	86	90	88	78	72	88	87	80

- a. Draw a stem-and-leaf plot to illustrate runs made by the two clubs with common stem and leaves of 'A' on right side and those of 'B' on left side. [2]
- b. Calculate means and standard deviation of the two clubs separately and use following formula to determine which club is more consistent: [3]

$$C.V. = \frac{\text{standard deviation}}{\text{mean}} \times 100\%$$

- c. Calculate the medians of runs of the two clubs and use following expression to calculate skewness of the runs separately for two clubs: [2]

$$S_K = \frac{3(\text{Mean} - \text{Median})}{\text{standard deviation}}$$

SECTION "D"

[6Q. × 4 = 24 marks]

4. The probability density function of a random variable X is given by

$$f(x) = \frac{k}{x^4}; x \geq 1$$

Find (i) the value of k , (ii) mean of X (iii) variance of X (iv) cumulative distribution function of X.

5. An airline knows that 5% of the persons making reservations will not show up at the gate. On a given flight that can seat 50 people, 52 tickets have been sold. Calculate the probability that there will be a seat available for every passenger that will arrive at the gate, i.e., find $P(X \leq 50)$.
6. A manufacturer does not know the mean and standard deviation of the diameters of the ball bearings he is producing. However, a sieving system rejects all bearings larger than 2.4 cm and those under 1.8 cm in diameter. Out of 1000 ball bearings 8% are rejected as too small and 5.5% are rejected as too big. What is the mean and standard deviation of the ball bearings produced?
7. A company manufactures an electronic device to be used in a very wide temperature range. The company knows that increased temperature shortens the life time of the device, and a study is therefore performed in which the life time is determined as a function of temperature. The following data is found:

Temperature in Celcius (x)	10	20	30	40	50	60	70
Life time in hours (y)	420	365	285	220	176	117	69

Develop the linear regression model to establish relationship between temperature and life-time.

8. Based on field experiments, a new variety green gram is expected to give an yield of 12.0 quintals per hectare. The variety was tested on 10 randomly selected farmer's 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? Also find 95% confidence interval for true average production.
9. The thickness of silicon chips produced by a machine are observed. 12 samples each of size 3 are found to have following values of means and ranges of thickness of sample units:

Sample	1	2	3	4	5	6	7	8	9	10	11	12
Mean (\bar{X})	0.258	0.249	0.259	0.268	0.256	0.268	0.26	0.257	0.257	0.245	2.248	0.239
Range (R)	0.038	0.03	0.009	0.024	0.039	0.003	0.026	0.046	0.035	0.017	0.049	0.005

Construct (a) \bar{X} -chart based on range (b) R-chart. Analyze both the charts. [Hints: $A_2 = 1.023, D_3 = 0, D_4 = 2.574$]

SECTION "E"

[5Q. \times 2 = 10 marks]

10. The number of blocked intrusion attempts on each day after changing firewall setting are recorded for 16 days as- 53, 21, 32, 49, 45, 38, 44, 33, 32, 43, 53, 46, 36, 48, 39, 35, 37, 36, 39, 45. If the first, second and third quartiles of these numbers are evaluated to be 35.75, 39.00 and 45.25 respectively, draw a box plot of these values.
11. Components of a personal stereo are normally distributed with mean life of 2400 hours with standard deviation of 300 hours. It is estimated that average user listens for about 1000 hours in one year. What is the probability that a component lasts for more than three years.
12. As part of quality-improvement project focused on the delivery of mail at a department office within a large company, data were gathered on the number of different address that had to be changed so the mail could be redirected to the correct mail stop. The distribution given below describes the number of redirects per delivery

x_i	0	1	2	3	4	5	6
$p(x_i)$	0.17	0.29	0.27	0.16	0.07	0.03	0.01

Find (a) the mean and (b) the standard deviation of number of redirects.

13. It has been observed that, on average, 180 cars per hour pass a specified point on a particular road in the morning rush hour. Due to impending roadworks it is estimated that congestion will occur closer to the city center if more than 5 cars pass the point in any one minute. What is the probability of congestion occurring? Use Poisson distribution.
14. The length of hand and height of five humans are recorded in centimeters as:

Hand	17	15	19	17	21
Height	150	154	169	172	175

Find Pearson's correlation coefficient between length of hand and body height.

