13 APR 2026

Subject Code:24BM11RC06 **R-24** **Reg No:** | | | | | | | | | | |

GAYATRI VIDYA PARISHAD COLLEGE OF ENGINEERING FOR WOMEN

(Autonomous)

(Affiliated to Andhra University, Visakhapatnam)

IIB.Tech. - II Semester Regular Examinations, April – 2026

PROBABILITY AND STATISTICS

[Common to CSE, CSE (AI&ML) & INF]

1. All questions carry equal marks
2. Must answer all parts of the question at one place

Time: 3Hrs. **Max Marks: 70**

UNIT-I

1\. a. Write briefly about box plot. Construct a histogram for the following data: [2M+5M]

Output (units/worker)	500-509	510-519	520-529	530-539	540-549	550-559	560-569
No. of workers	8	18	23	37	47	26	16

b. The frequency distribution of weight in grams of mangoes of a given variety is given below. Calculate the mean and the median. [7M]

Weight in grams	410-419	420-429	430-439	440-449	450-459	460-469	470-479
No. of mangoes	14	20	42	54	45	18	7

OR

2\. a. Find the standard deviation of intelligence quotient (I.Q) of 68 students of the following data: [7M]

I.Q	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of students	5	12	15	20	10	4	2

b. From the following data find out Karl Pearson's coefficient of skewness: [7M]

Measurement	10	11	12	13	14	15
Frequency	2	4	10	8	3	1

UNIT-II

3\. a. Two dice are thrown. Let X assign to each point (a,b) in S the maximum of its numbers i.e. X(a, b) = MAX (a, b). Find the probability distribution. X is a random variable with X(s) = {1, 2, 3, 4, 5, 6}. Also find the mean and variance of the distribution. [7M]

b. If a random variable has the probability density function

$f(x) = k(x^2 - 1), -1 \le x \le 3$
$= 0,\ else\, where$

Find the value of 'k' and $p\left(\frac{1}{2} \le x \le \frac{5}{2}\right)$ [7M]

OR

4\. a. 20% of items produced from a factory are defective. Find the probability that in a sample of 5 chosen at random i) none is defective ii) one is defective [7M]

b. If the masses of 300 students are normally distributed with mean 68 kgs and standard deviation 3 kgs, how many students have masses

(i) Greater than 72 kgs
(ii) Less than or equal to 64 kgs
(iii) Between 65 and 71 kgs inclusive [7M]

UNIT-III

5. a. A population consists of six numbers 4, 8, 12, 16, 20, 24. Consider all possible samples of size two which can be drawn without replacement from this population. Find
   (i) The population means.
   (ii) The population Standard deviation.
   (iii) The mean of the sampling distribution of means.
   (iv) The standard deviation of the sampling distribution of means. [7M]

   b. A random sample of size 100 is taken from an infinite population having the mean $\mu = 76$ and the variance $\sigma^2 = 256$. What is the probability that $\bar{x}$ will be between 75 and 78. [7M]

OR

6. a. The pulse rate of 50 yoga practitioners decreased on the average by 20.2 beats/minute with standard deviation 3.5. If $\bar{x} = 20.2$ is used as a point estimate of the true average decrease in pulse rate then what can we assert with 95% confidence about the maximum error E. [7M]

   b. A random sample of 400 items is found to have mean 82 and Standard deviation 18. Find the maximum error of estimation at 95% confidence interval. [7M]

UNIT-IV

7. a. If in a random sample of 600 cars making a right turn at a certain traffic junction 157 drove in to the wrong line, test whether actually 30% of all drivers make this mistake or not at this given junction. Use 0.05 L.O.S. [7M]

   b. A simple sample of heights of 6400 Englishmen has a mean of 67.85 inches and a S.D. of 2.56 inches while a simple sample of heights of 1600 Australian has a mean of 68.55 inches and a S.D. of 2.52 inches. Do the data indicate that the Australians are on the average taller than the Englishmen at 5% level of significance? [7M]

OR

8. a. A sample of 400 items is taken from a population whose standard deviation is 10. The mean of the sample is 40. Test whether the sample has come from a population with mean 38. [7M]

   b. A sample analysis of examination results of 500 students was made. It was found that 220 students has failed, 170 had secured a third class, 90 were placed in second class and 20 got a first class. Do these figures commensurate with the general examination result which is in the ratio of 4: 3: 2: 1 for the various categories respectively. [7M]

UNIT-V

9. a. Fit an equation of the form $y = ab^x$ to the following data: [7M]

x	2	3	4	5	6
y	144	172.8	207.4	248.8	298

   b. Find Karl Pearson's coefficient of correlation from the following data: [7M]

Wages	100	101	102	102	100	99	97	98	96	95
Cost of living	98	99	99	97	95	92	95	94	90	91

OR

10. a. For a set of values of x and y, the two regression lines are $31x-37y+5=0$ and $50x-6y-612=0$. Identify the regression line of y on x and that of x on y. Also obtain the mean values of x and y. [7M]

    b. Using the method of least squares, obtain the regression line of X on Y from the following data:

X	10	12	13	16	17	20	25
Y	10	22	24	27	29	33	37

[7M]

Probability & Statistics (24BM11RC06)
April - 2026
Solutions & scheme of valuation

1a) **Box-plot**: It is a standardized way of displaying the distribution of data based on

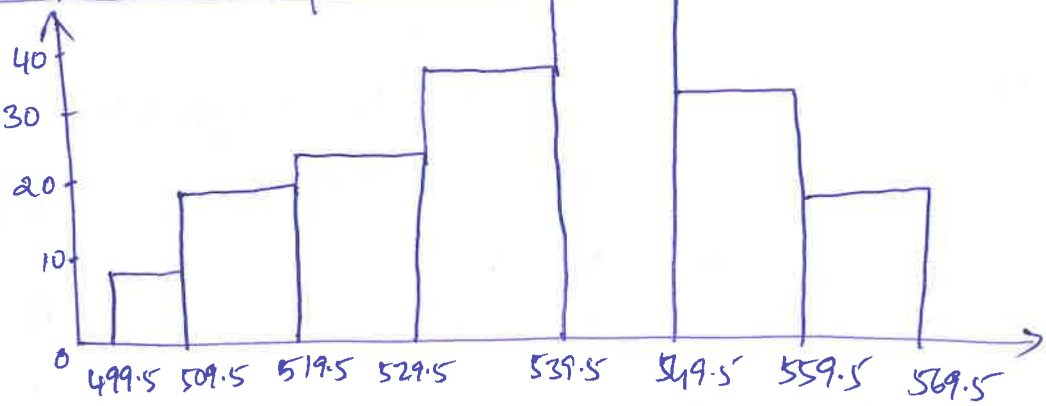
1) Minimum : the lowest data point
2) first quartile (Q_1)
3) Median (Q_2)
4) Third Quartile (Q_3)
5) Maximum : The highest data point

(2M)

$IQR = Q_3 - Q_1$

Continuous	499.5 – 509.5	- - - -				559.5 – 569.5
Frequency	8	18	23	37	47	26	16

— (1M)



— (4M)

1b)

x :	414.5	424.5	434.5	444.5	454.5	464.5	474.5
f :	14	20	42	54 f	45	18	7
cf :	14	34	76(cf)	130	175	193	200

— (2M)

Σfx : 88680 , $N = \Sigma f = 200$

$$\text{Mean} = \frac{\Sigma fx}{N} = 443.4 \text{ grams}$$ — (2M)

$$\text{Median} = L + \left(\frac{\frac{N}{2} - cf}{f}\right) \times h$$

$$= 439.5 + \left(\frac{100 - 76}{54}\right) \times 10 = 443.94 \text{ grams}$$ — (3M)

(2a) x: 15 25 35 45 55 65 75
f: 5 12 15 20 10 4 2

$N = \Sigma f = 68$, $\Sigma fx = 2760$, $\Sigma fx^2 = 125900$ } — (2M)

Mean $(\bar{x}) = \frac{\Sigma fx}{N} = 40.59$ — (2M)

standard deviation $(\sigma) = \sqrt{\frac{\Sigma fx^2}{N} - \left(\frac{\Sigma fx}{N}\right)^2} = 14.28$ — (3M)

(2b)

$\Sigma f = 28$, $\Sigma fx = 348$ — (1M)

$\bar{x} = 12.32 \left(\frac{\Sigma fx}{\Sigma f}\right)$, Mode $= 12$ (highest frequency) — (2M)

SD $(\sigma) = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{N}} = 1.17$ — (2M)

$S_k = \frac{\text{Mean} - \text{Mode}}{\sigma} = 0.27$ — (2M)

3a) total outcomes : (1,1) (1,2) (1,3) - - - - (6,6) — (1M)
(total 36)

$X = \text{Max}(a,b)$: 1, 2,2,2, 3,3,3,3,3, 4,4,4,4,4,4,4, 5,5,5,5,5,5,5
5,5, 6,6,6,6,6,6,6,6,6,6,6 . — (1M)

x:	1	2	3	4	5	6
$P(x)$:	1/36	3/36	5/36	7/36	9/36	11/36

— (2M)

Mean $= E(X) = \Sigma x P(x) = \frac{161}{36} = 4.47$ — (1M)

$E(x^2) = \Sigma x^2 P(x) = \frac{791}{36} = 21.97$ — (1M)

Variance $(X) = E(x^2) - [E(x)]^2 = \frac{2555}{1296} = 1.97$ — (1M)

3b) WKT, $\int_{-\infty}^{\infty} f(x)\,dx = 1$ — (1M)

$\int_{-1}^{3} k(x^2-1)\,dx = 1$ — (1M)

$k\left(\frac{x^3}{3} - x\right)_{-1}^{3} = 1$

$\Rightarrow k\left(6 - \frac{2}{3}\right) = 1 \Rightarrow k = \frac{3}{16}$ — (2M)

$$P\left(\frac{1}{2} \le x \le \frac{5}{2}\right) = \int_{1/2}^{5/2} \frac{3}{16}(x^2-1)\,dx \quad \text{— 1M}$$

$$= \frac{3}{16}\left[\frac{x^3}{3} - x\right]_{1/2}^{5/2}$$

$$= \frac{3}{16}\left(\frac{125}{24} - \frac{5}{2} - \frac{1}{24} + \frac{1}{2}\right) = \frac{19}{32} \simeq 0.59375 \quad \text{— 2M}$$

4a) This follows a Binomial distribution with parameters — 1M

$$\left.\begin{array}{l} n = 5 \\ p = 20\% = 0.20 \\ q = 1 - p = 0.80 \end{array}\right\} \text{— 2M}$$

$$P(X=k) = {}^nC_k\, p^k q^{n-k}$$

(i) $P(X=0) = {}^5C_0 (0.20)^0 (0.80)^5 = 0.32768$ — 2M

(ii) $P(X=1) = {}^5C_1 (0.20)^1 (0.80)^4 = 0.4096$ — 2M

4b) Given $\mu = 68$, $\sigma = 3$, $N = 300$ — 1M

$$z = \frac{x-\mu}{\sigma}$$

(i) $P(X > 72) = P\left(z > \frac{72-68}{3}\right) = P(z > 1.33)$ — 1M

$= 0.5 - P(0 < z < 1.33) = 0.5 - 0.4082 = 0.0918$ — 1M

No of students $= 300 \times 0.0918 = 27.54 \simeq 28$

(ii) $\left.P(X \le 64) = P\left(z \le \frac{64-68}{3}\right) = P(z \le -1.33)\right\}$ — 1M

$= P(z \ge 1.33) = 0.0918$ — 1M

No of students $= 300 \times 0.0918 = 27.54 \simeq 28$.

(iii) $\left.P(65 < z < 71) = P(-1 < z < 1) = 2 \times P(0 < z < 1)\right\}$ — 1M

$= 2 \times 0.3413 = 0.6826$

No of students $= 300 \times 0.6826 = 204.78 \simeq 205$ — 1M

5a) (i) $\mu = \frac{\Sigma x}{N} = \frac{84}{6} = 14$ — 1M

(ii) $\sigma = \sqrt{\frac{\Sigma (x-\mu)^2}{N}} = \sqrt{\frac{100+36+4+4+36+100}{6}} = \sqrt{\frac{280}{6}} = 6.83$ — 1M

(iii) Samples of size 2 are $= {}^{N}C_{n} = {}^{6}C_{2} = 15$

(4,8), (4,12), (4,16), (4,20), (4,24)
(8,12) (8,16) (8,20) (8,24)
(12,16) (12,20) (12,24)
(16,20) (16,24)
(20,24) — 1M

Mean of each sample is
6 8 10 12 14 10 12 14 16 14 16 18
18 20 22

$\bar{x}$:	6	8	10	12	14	16	18	20	22
f :	1	1	2	2	3	2	2	1	1

— 1M

$\mu_{\bar{x}} = \frac{\Sigma f \bar{x}}{\Sigma f} = 14$ — 1M

(iv) Using $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{6.83}{\sqrt{2}} \sqrt{\frac{6-2}{6-1}} = 4.32$ — 2M

5b) Given $\mu = 76$, $\sigma^2 = 256$, $n = 100$, $z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$ — 2M

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{10} = 1.6$ — 1M

$P(75 < \bar{X} < 78) = P\left(\frac{75-76}{1.6} < z < \frac{78-76}{1.6}\right)$ — 1M

$= P(-0.625 < z < 1.25)$
$= P(z < 1.25) - P(z < -0.625)$ — 2M
$= 0.8944 - 0.2660 = 0.6284$

$\therefore$ The required probability $= 62.84\%$ — 1M

6a) $n = 50$, $\bar{x} = 20.2$, $s = 3.5$, $\alpha = 0.05$ — 2M

$z_{\alpha/2} = 1.96$ — 2M

Max error $(E) = z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$ — 1M

$= 1.96 \times \frac{3.5}{\sqrt{50}}$

$= 0.9701$ — 2M

We can be 95% confident that the maximum error of estimate is 0.97 beats/min

6b) Given $n = 400$, $s = 18$, $\bar{x} = 82$, $\alpha = 0.05$ — 2M

$z_{\alpha/2} = 1.96$ — 2M

Max error $(E) = z_{\alpha/2} \frac{s}{\sqrt{n}}$ — 1M

$= 1.96 \times \frac{18}{\sqrt{400}}$

$= 1.764$ — 2M

7a) Given $n = 600$, $x = 157$ $\Rightarrow$ $\hat{p} = \frac{x}{n} = 0.2617$ — 2M

Let $P = 30\% = 0.30$

1) H_0: $P = 0.30$ (The proportion of drivers who make the mistake is 30%.)

2) H_1: $P \neq 0.30$. (The proportion is not 30%.) — 1M

3) $\alpha = 0.05$ (Two-tailed test) — 1M

4) $z_{\alpha/2} = 1.96$

5) $z_{cal} = \frac{\hat{p} - P}{\sqrt{\frac{P(1-P)}{n}}} = \frac{0.2617 - 0.30}{\sqrt{\frac{0.3 \times 0.7}{600}}} = -2.049$ — 2M

6) Since $|-2.049| > 1.96$

we reject H_0. — 1M

$\therefore$ The actual proportion of drivers making the mistake is not 30%.

7b) Given $n_1 = 6400$, $\bar{x}_1 = 67.85$, $s_1 = 2.56$
$n_2 = 1600$, $\bar{x}_2 = 68.55$, $s_2 = 2.52$ } — 1M

1) H_0: $\mu_1 = \mu_2$ (There is no difference in mean heights)
2) H_1: $\mu_2 > \mu_1$ (Australians are taller than English men) } — 1M
Right tailed test

3) $\alpha = 0.05$
4) $z_\alpha = 1.645$ } — 1M

5) $$z = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{68.55 - 67.85}{\sqrt{\frac{2.56^2}{6400} + \frac{2.52^2}{1600}}} = \frac{0.7}{0.07066} = 9.91$$ — 2M

6) Since $9.91 > 1.645$ — 1M
we reject H_0.
Australians are on average taller than Englishmen.

8a) Given $\mu = 38$, $\bar{x} = 40$, $n = 400$, $\sigma = 10$ — 1M

1) H_0: $\mu = 38$
2) H_1: $\mu \neq 38$ (Two tailed test) } — 1M
3) $\alpha = 5\%$
4) $z_{\alpha/2} = 1.96$ } — 1M

5) $$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{40-38}{10/\sqrt{400}} = 4$$ — 2M

6) Since $4 > 1.96$, we reject H_0. } — 2M
The sample mean of 40 is significantly different from 38

$N = 500$

8b) Expected ratio $= 4:3:2:1$ (Total parts $4+3+2+1 = 10$) — 1M

Fail (4 parts) : $\frac{4}{10} \times 500 = 200$

Third (3 parts) : $\frac{3}{10} \times 500 = 150$

Second (2 parts) : $\frac{2}{10} \times 500 = 100$

First (1 part) : $\frac{1}{10} \times 500 = 50$

$E_1 = 200$, $E_2 = 150$, $E_3 = 100$, $E_4 = 50$ — 2M

Observed frequencies are — 1M

$O_1 = 220,\ O_2 = 170,\ O_3 = 90,\ O_4 = 20$

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 23.67$$ — 1M

Here $\alpha = 5\%$ $\Rightarrow \chi^2_{\alpha} = 7.815$ with 3 dof — 1M

$\therefore \chi^2_{cal} > \chi^2_{tab}$ $(23.67 > 7.815)$

we reject H_0 — 1M

$\therefore$ The sample data does not fit $4:3:2:1$ distribution

9a) $y = ab^x$

$\ln y = \ln a + x \ln b$

$Y = A + Bx$ where $\ln y = Y$, $\ln a = A$, $\ln B = B$ — 1M

x	y	$Y = \ln y$	x^2	xY
2	144	2.1584	4	4.3168
3	172.8	2.2375	9	6.7125
4	207.4	2.3168	16	9.2672
5	248.8	2.3959	25	11.9795
6	298.6	2.4751	36	14.8506
20	~~11.5837~~	11.5837	90	47.1266

— 2M

Normal eqns are $\Sigma Y = nA + B\Sigma x \Rightarrow 11.5837 = 5A + 20B$

$\Sigma xY = A\Sigma x + B\Sigma x^2 \Rightarrow 47.1266 = 20A + 90B$ — 2

Solving we get, $A = 2$, $B = 0.07918$ — 1M

$a = 10^2 = 100$

$b = 10^{0.07918} = 1.2$

$\therefore y = 100(1.2)^x$ — 1M

9b)

x	y	x^2	y^2	xy
100	98	10000	9604	9800
101	99	10201	9801	9999
102	99	10404	9801	10098
102	97	10404	9409	9894
100	95	10000	9025	9500
99	92	9801	8464	9108
97	95	9409	9025	9215
98	94	9604	8836	9212
96	90	9216	8100	8640
95	91	9025	8281	8645
990	950	98064	90346	94111

— (3M)

Karl Pearson's Coefficient of Correlation $(r) = \dfrac{n\Sigma xy - \Sigma x \, \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2]}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$ — (2M)

here $n = 10$ observations

$$r = \frac{10(94111) - (990)(950)}{\sqrt{10(98064) - (990)^2}\sqrt{10(90346) - (950)^2}}$$

$$= \frac{610}{\sqrt{518400}} = 0.847$$ — (2M)

10a) Regression lines always intersect at mean point $(\bar{x}, \bar{y})$ — (1M)

so, $31\bar{x} - 37\bar{y} = -5$
$50\bar{x} - 6\bar{y} = 612$ $\Big\}$ — ($\frac{1}{2}$M)

on solving, we get $\bar{x} = 13.626$, $\bar{y} = 11.552$ — ($\frac{1}{2}$M)

Case ① Assume $31x - 37y + 5 = 0$ is y on x

$\Rightarrow y = \frac{31}{37}x + \frac{5}{37}$ $\Rightarrow$ slope $b_{yx} = \frac{31}{37} = 0.838$ $\Big\}$ — (2M)

Case ② Assume $50x - 6y - 612 = 0$ is x on y

$x = \frac{6}{50}y + \frac{612}{50}$ $\Rightarrow$ slope $b_{xy} = \frac{6}{50} = 0.12$ $\Big\}$ — (2M)

Verification: $r^2 = b_{xy} \cdot b_{yx} = 0.101$

Since $0 \le r^2 \le 1$, this assumption is correct — (1M)

10b)

x	y	xy	y^2
10	10	100	100
12	22	264	484
13	24	312	576
16	27	432	729
17	29	493	841
20	33	660	1089
25	37	925	1369
113	182	3186	5188

— 2M

$x = a + by$ — 1M

where b is the regression coefficient and a is the intercept

$$b_{xy} = b = \frac{n\Sigma xy - \Sigma x \Sigma y}{n(\Sigma y^2) - (\Sigma y)^2} = \frac{7(3186) - 113(182)}{7(5188) - (182)^2}$$ — 2M

$= 0.5439$

Formula for a is $a = \bar{x} - b\bar{y}$

here $\bar{x} = \frac{\Sigma x}{n} = 16.1429$ — 1M

$\bar{y} = \frac{\Sigma y}{n} = \frac{182}{7} = 26$

$a = 16.1429 - 0.5439(26) = 2.0015$

$\therefore$ $x = a + by$ is $x = 2.0015 + 0.5439y$ — 1M

Prepared by

VSSSD Prakash

Verified by

A. K

Dr. A. Kameswara Rao