1 – 52412

Register No. | | | | | | | | |

BE/BTech Degree Examination May 2022

Fourth Semester

Common to Computer Science and Engineering & Information Technology

20MAT42 – PROBABILITY AND STATISTICS

(Regulations 2020)

Use of statistical table is permitted

Time: Three hours Maximum: 100 marks

Answer all Questions

Part – A ($10 \times 2 = 20$ marks)

1. When a die is thrown, X denotes the number that turns up. Find E(X). [CO1,K2]
2. Define n^{th} moment about origin of a random variable X. [CO1,K1]
3. If the probability that a target is destroyed on any one shot is 0.5, what is the probability that it would be destroyed on 6^{th} attempt? [CO2,K2]
4. State Memoryless property of exponential distribution. [CO2,K1]
5. Define Covariance. [CO3,K1]
6. The two lines of regression are $8x - 10y + 66 = 0$ and $40x - 18y - 214 = 0$. Find the mean values of x and y. [CO3,K2]
7. What is meant by Type I and Type II errors in hypothesis testing? [CO4,K1]
8. List any two applications of F-test. [CO4,K1]
9. Name the three basic principles of experimental design. [CO5,K1]
10. Is 2×2 Latin square design possible? Why? [CO5,K1]

Part – B ($5 \times 16 = 80$ marks)

11. a. i) A discrete random variable X has the following probability distribution. (10) [CO1,K3]

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

1) Find the value of 'a'.

2) Find $P(X < 3), P(0 < X < 3)$ and $P(X \geq 3)$.

3) Find the distribution function of X.

ii) For the triangular distribution (6) [CO1,K3]

$$f(x) = \begin{cases} x \ ; & 0 \leq x \leq 1 \\ 2-x; & 1 \leq x \leq 2 \\ 0 \ ; & \text{otherwise} \end{cases}$$

Find the Mean and Variance.

(OR)

b. The density function of a random variable X is given by $f(x) = Kx(2-x); 0 \le x \le 2$. Find K, Mean, Variance and r^{th} moment. (16) [CO1,K3]

12. a. i) Derive M.G.F, Mean and Variance of Binomial distribution. (10) [CO2,K2]

ii) The number of monthly breakdown of a computer is a random variable having a poisson distribution with mean equal to 1.8. Find the probability that this computer will function for a month (6) [CO2,K2]

1) Without a breakdown
2) With only one breakdown and
3) With atlease one breakdown.

(OR)

b. i) If X is uniformly distributed over (0, 5), find the probability that (6) [CO2,K2]
1) $P(X < 2)$, 2) $P(X > 3)$ and 3) $P(2 < X < 5)$

ii) The savings bank account of a customer showed an average balance of Rs.150 and a standard deviation of Rs.50. Assuming that the account balances are normally distributed. (10) [CO2,K2]

1) What percentage of account is over Rs.200?
2) What percentage of account is between Rs.120 and Rs.170?
3) What percentage of account is less than Rs.75?

13. a. i) The joint probability mass function of (X,Y) is given by $P(x, y) = k(2x + 3y)$, $x=0, 1, 2;\ y=1, 2, 3$. Find the Marginal distributions of X and Y. (6) [CO3,K3]

ii) Calculate the correlation coefficient for the following data. (10) [CO3,K3]

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

(OR)

b. i) The joint density function of the R.V's X and Y is given by (6) [CO3,K3]

$$f(x,y) = \begin{cases} 8xy; & 0 \le x \le 1;\ 0 \le y \le x \\ 0 & : \text{otherwise} \end{cases}$$

1) Find the Marginal density functions of X and Y.
2) Find the Conditional density function f (y/x).

ii) Obtain the two regression lines from the following data. (10) [CO3,K3]

X	1	2	3	4	5	6	7
Y	9	8	10	12	11	13	14

14. a. i) In a large city A, 20 percent of a random sample of 900 school boys had a slight physical defect. In another large city B, 18.5 percent of a random sample of 1600 school boys had the same defect. Is the difference between the proportions significant? (8) [CO4,K3]

ii) On the basis of information given below, Use Chi-square test to check whether the new treatment is comparatively superior to the conventional one. (8) [CO4,K3]

	Favourable	Non-favourable	Total
Conventional	40	70	90
New	60	30	110
Total	100	100	200

(OR)

b. The nicotine contents in milligrams in two samples of tobacco were found to be as follows. (16) [CO4,K3]

Sample A	24	27	26	21	25	—
Sample B	27	30	28	31	22	36

Can it be said that two samples come from same normal population.

15. a. An experiment was designed to study the performance of 4 different detergents for cleaning injectors. The following readings were obtained with specially designed equipment for 12 tanks of gas distributed over 3 different models of engines. (16) [CO5,K3]

Detergents	Engine-I	Engine-II	Engine-III
A	45	43	51
B	47	46	52
C	48	50	55
D	42	37	49

Perform the ANOVA, to test whether there are significant differences in the detergents and in the engines.

(OR)

b. Analyse the following results of a Latin square experiments. (16) [CO5,K3]

	1	2	3	4
1	A(12)	D(20)	C(16)	B(10)
2	D(18)	A(14)	B(11)	C(14)
3	B(12)	C(15)	D(19)	A(13)
4	C(16)	B(11)	A(15)	D(20)

The letters A, B, C, D denote the treatments and the figures in brackets denote the observations.

Bloom's Taxonomy Level	Remembering (K1)	Understanding (K2)	Applying (K3)	Analysing (K4)	Evaluating (K5)	Creating (K6)
Percentage	8	21	71	-	-	-

20MAT42 - Probability & Statistics

Answer key

Part - A

1) $E(X) = \dfrac{1+2+3+4+5+6}{6} = \dfrac{7}{2}$.

2) $\mu_r' = \int_{-\infty}^{\infty} x^r f(x)dx$.

3) $P(X=6) = pq^{x-1} = (0.5)(0.5)^5 = 0.0156$

4) If X is exponentially distributed, then $P(X > s+t / X > s) = P(X > t)$, for any $s, t > 0$.

5) If X and Y are RV's then covariance between X & Y is $Cov(X,Y) = E(XY) - E(X)E(Y)$.

6) $\bar{y} = 17$, $\bar{x} = 13$.

7) Type I : Reject H_0 when it is true
Type II : Accept H_0 when it is false.

8) To test if the 2 samples have come from the same population. To test whether there is any significant differen b/w Population variance

9) Randomisation, Replication & Local control.

10) For $n=2$, d.f of SSE $=0$ and hence MSE is not defined.

Part - B

11) a) i)

1) $\sum P_i = 1 \Rightarrow a = \dfrac{1}{81}$ —————— (2)

2) $P(X<3) = P(X=0) + P(X=1) + P(X=2)$
$= 9a = \dfrac{1}{9}$ —————— (2)

$P(0<X<3) = P(X=1) + P(X=2)$
$= 8/81$ —————— (2)

$P(X \geq 3) = 1 - P(X < 3)$

$= 1 - [P(X=0) + P(X=1) + P(X=2)]$

$= 8/9$ ———— (2)

3)

x	0	1	2	3	4	5	6	7	8
$P(X \leq x)$	$1/81$	$4/81$	$9/81$	$\frac{16}{81}$	$\frac{25}{81}$	$\frac{36}{81}$	$\frac{49}{81}$	$\frac{64}{81}$	1

———— (2)

ii)

$E(X) = \int x f(x) dx$

$= 1$ ———— (2)

$E(X^2) = \int x^2 f(x) dx = 7/6$ ———— (2)

$V(X) = E(X^2) - [E(X)]^2 = 1/6.$ ———— (2)

b) $\int_0^2 kx(2-x)dx = 1$

———— (3)

$k = 3/4$

$E(X) = \int_0^2 x f(x) dx = 1$, $E(X^2) = 6/5$

———— (3) ———— (3)

$V(X) = 1/5$ ———— (3)

$\mu_r' = \int_0^2 x^r f(x) dx = \frac{3(2)^{r+1}}{(r+2)(r+3)}$ (or) ———— (4)

$= \frac{6\, 2^r}{(r+2)(r+3)}$

12) (a) i) $M_X(t) = \sum_{x=0}^{n} e^{tx} p(x)$

$= q^n + nc_1 q^{n-1}(pe^t) + nc_2 q^{n-2}(pe^t)^2 + \ldots$

$M_X(t) = (q + pe^t)^n$ —————— (3)

$M_X'(t) = n(q + pe^t)^{n-1} pe^t$

Mean $= E(X) = np$ —————— (3)

$M_X''(t) = n^2p^2 + np(1-p)$

$M_X''(0) = n^2p^2 + npq$

$V(X) = npq$. —————— (4)

ii) Mean $= \lambda = 1.8$.

$P(X=x) = \frac{e^{-1.8}(1.8)^x}{x!}$ —————— (1)

1) $P(X=0) = e^{-1.8} = 0.1652$ —————— (1)

2) $P(X=1) = (1.8)e^{-1.8} = 0.2975$ —————— (2)

3) $P(\text{atleast one}) = P(1) + P(2) + \ldots$

$= 1 - P(0)$

$= 0.8343$ —————— (2)

b) i)

$f(x) = \frac{1}{b-a} = \frac{1}{5-0} = \frac{1}{5}$. —————— (1)

i) $P(X<2) = \int_0^2 \frac{1}{5}dx = 2/5$ —————— (1)

ii) $P(X>3) = \int_3^5 \frac{1}{5}dx = 2/5$ —————— (2)

iii) $P(2<X<5) = \int_2^5 \frac{1}{5}dx = 3/5$. —————— (2)

ii) $\mu = 150, \quad \sigma = 50$

$$z = \frac{X - 150}{50} \quad \text{———— (1)}$$

1) $P(X \geq 200) = P(z \geq 1) = 0.1587$

% of account over Rs 200 is 15.87 ———— (3)

2) $P(120 \leq X \leq 170) = P(-0.6 \leq z \leq 0.4)$

$= 0.2257 + 0.1554$

$= 0.3811$ ———— (3)

% of account b/w 120 & Rs 170 is 38.11.

3) $P(X < 75) = P(z < -1.5)$

$= 0.0668$.

% of account less than Rs 75 is 6.68 % ———— (3)

13) a) i)

Y \ X	0	1	2	P(Y=y)
1	3k	5k	7k	15k
2	6k	8k	10k	24k
3	9k	11k	13k	33k
P(X=x)	18k	24k	30k	72k

$k = 1/72$ ———— (3)

Y \ X	0	1	2	P(Y=y)
1	3/72	5/72	7/72	15/72
2	6/72	8/72	10/72	24/72
3	9/72	11/72	13/72	33/72
P(X=x)	18/72	24/72	30/72	1

—(3)

13a) ii) $\Sigma X = 544$ $\Sigma Y = 552$ } ———— (2)

$\Sigma XY = 37560$, $\Sigma X^2 = 37028$, $\Sigma Y^2 = 38132$

$\bar{X} = 68$, $\bar{Y} = 69$, $\bar{X}\bar{Y} = 4692$ ———— (2)

$\sigma_x = 2.121$ $\sigma_y = 2.345$ } ———— (2)

$$r(X,Y) = \frac{\frac{1}{8}(37560 - 4692)}{(2.121)(2.345)}$$ ———— (3)

$r(X,Y) = 0.6030$ ———— (1)

---

13) b) i)

$f(x) = \int_0^x 8xy\,dy = 4x(1-x^2)$.

$$f(x) = \begin{cases} 4x^3, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$ ———— (2)

$$f(y) = \begin{cases} 4y, & 0 < y < x \\ 0, & \text{otherwise} \end{cases}$$ ———— (2)

$f(x/y) = 2x$, $f(y/x) = \frac{2y}{x^2}$ } ———— (2)

If $y = ax+b$ is the line of regression of y on x

ii) $\Sigma x = 28$, $\Sigma y = 77$ $\Sigma x^2 = 140$ ———— (1)

$\Sigma y^2 = 875$ $\Sigma xy = 334$. ———— (2)

Normal eqns are $77 = 28a + 7b$ & $334 = 140a + 28b$ ———— (1)

$a = 13/14 \qquad b = 51/7$

Regression line is $14y = 13x + 102$. —— (3)

If $x = cy + d$ is line of regression of x & y, then.

$28 = 77c + 7d$ & $334 = 875c + 77d$.

$c = \frac{13}{14}$, $d = \frac{-87}{14}$.

regression of x on y is $14x = 13y - 87$. —— (3)

14) a) i)

$P_1 = 0.2$, $P_2 = 0.185$. —— (2)

$H_0 : P_1 = P_2$

$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = 0.1904$ —— (4)

$Z = 0.9375$. —— (2)

$|z| < 1.96$,

$\therefore$ Accept H_0.

ii) H_0: No difference b/w new and conventional treatm

O	E
60	45
30	45
40	55
70	55

—— (4)

$\chi^2 = \frac{\Sigma(O-E)^2}{E} = 18.18$. —— (4)

χ^2 for 1 Dof at 5% is 3.841.

$\chi^2_{cal} > \chi^2_{0.05}$

Reject H_0 at 5% LOS.

b) To test i) equality of Variances by using F-test. 4

ii) Equality of means by t-test.

$\bar{x} = 24.6, \quad \bar{y} = 29.$ ——— (2)

$s_1^2 = 5.3, \quad S_2^2 = 21.6.$ ——— (2)

$F = \frac{S_2^2}{S_1^2} = 4.07.$ ——— (4)

$F_{cal} < F_{0.05}(5,4)$. $\therefore$ accept. H_0 at 5% LOS

(ii) $H_0 : \mu_1 = \mu_2$.

$S = 3.78.$ ——— (2)

$t = \frac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = -1.92.$ ——— (4)

$t_{0.05}(9) = 2.26.$ ——— (2)

$t_{cal} < t_{0.05}$.

$\therefore$ Accept H_0 at 5% LOS.

15a) By subtracting 50 from each value, we proceed,

$N = 12. \quad T = -35$ ——— (2)

$\frac{T^2}{N} = 102.08.$ ——— (2)

$TSS = 264.92.$ ——— (2)

$SSC = \frac{(-18)^2}{4} + \frac{(-24)^2}{4} + \frac{7^2}{4} - 102.08$

$= 135.17.$ ——— (2)

$SSR = 110.91$ ——— (2)

SSE = 18.84.

Source of var	Sum of sqrs	D.f	Mean sqrs	Variance test
B/W columns	SSC = 135.17	2	MSC = 67.58	$F_C = 21.52$
B/W rows	SSR = 110.91	3	MSR = 36.97	$F_R = 11.77$.
Residual	SSE = 18.84	6	MSE = 3.14	
Total	TSS = 264.92	11		

—— (6)

$F_C(2,6) = 5.14$ F_C & F_R reject H_0 at 5% LOS

~~FR(3,6) = 4.76~~

b). H_0: There is no significant difference b/w rows, columns & treatment

T = −4., —— (2)

Correction factor = 1 —— (2)

TSS = 157 —— (2)

Source of Variation	Sum of squares	D.f	Mean squares	F-ratio
Rows	SSR = 3.5	3	MSR = 1.167	$F_R = 1.081$
Columns	SSC = 2.5	3	MSC = 0.87	$F_C = 1.24$
Treatments	SST = 144.5	3	MST = 48.17	$F_T = 44.60$
Residual	SSE = 6.5	6	MSE = 1.08	

$F_R < F_{tab}$ Accept H_0, $F_C < F_{tab}$, Accept H_0.

$F_T > F_{tab}$ reject H_0 —— (10)