



Register Number:

Date: 07-01-2020

**ST. JOSEPH'S COLLEGE (AUTONOMOUS), BANGALORE – 27**  
**M.Sc. DATA ANALYTICS – I SEMESTER**  
**SEMESTER EXAMINATION – JANUARY 2021**  
**BDA1220: THEORY OF PROBABILITY AND STOCHASTIC PROCESS**

**THIS QUESTION PAPER CONTAINS THREE PRINTED PAGE AND ONE PART**

**STUDENTS ARE ALLOWED TO USE SCIENTIFIC CALCULATORS**

**Time: 2 1/2 hrs.**

**Answer any Seven questions**

**7 X 10 = 70**

1. (a) A 5-year study is undertaken to observe cataract in a population of 5000 people 60 years of age and older. A census data reveals that 45% of this population is 60–64 years of age, 28% are 65–69 years of age, 20% are 70–74 years of age, and 7% are 75 or older. Also from the Framingham Eye Study it is found that 2.4%, 4.6%, 8.8%, and 15.3% of the people in these respective age groups will develop cataract over the next 5 years. What percentage of the population in our study will develop cataract over the next 5 years, and how many people with cataract does this percentage represent? (5)
- (b) What is the probability that, in a room of 30 people, there is a pair of people who have the same birthday? (5)

2. (a) Over the course of a season a hockey team plays 40 matches in different conditions with the following results.

		Weather		
		Good	Poor	Total
Results	Win	11	6	19
	Draw	5	3	8
	Lose	7	6	17
	Total	25	15	40

For a match chosen at random, from the seasons

G is the event 'Good weather'

W is the event 'Team wins'

D is the event 'Team draws'

L is the event 'Team losers'

Find the probability in each case:

(i)  $P(G)$  (ii)  $P(G \cap D)$  (c)  $P(G/D)$  (5)

(b) A discrete random variable  $X$  has the probability distribution shown in the table

$X$	8	10	15
$P(X=x)$	0.4	$a$	$0.6-a$

(i) Given that  $E(X)=10.2$  find  $a$

(ii) Find  $\text{Var}(X)$

(iii) Find  $P(\mu - \sigma)$

(5)

3.(a) A quality control agent tests sets of 10 components from a production line which is known to produce 98% defective components

(i) Find the probability that a set chosen at random is free from defects.

(ii) If the quality control agent tests five sets before lunch find the probability that four of these sets were free from defects. (6)

(b) Write a note on Geometric Distribution . (4)

4. (a) An insurance company insures 4000 people against loss of both eyes in a car accident . Based on previous data , the rates were computed on the assumption that on the average 10 persons in 1,00,000 will have car accident each year that result in this type of injury. What is the probability that more than 3 of the insured will collect on their policy in a given year? [given  $e^{-0.4}=0.6703$ ]

(5)

(b) State and prove the Bayes' Theorem (5)

5. (a) In 2021 there will be three candidates for the position of Principal –Mr. Bhatnagar, Mr. Mishra and Mr. Singh whose chance of getting the appointment are in the proportion 4:2:3 respectively. The probability that Mr. Bhatnagar if selected would introduce co-education in the college is 0.3. The probabilities of Mr. Mishra and Mr. . Singh are respectively 0.5 and 0.8

(i) What is the probability that there will be co-education in the college in 2022?

(ii) If there is coeducation in the college in in 2022 what is the probability that Mr . Singh is the principal? (5)

(b) Discuss the Central Limit Theorem. (5)

6. (a) What is the Markovian Property? Explain the idea of one step and two step probability transition matrix .

(2+3=5)

(b) Given the following transition probability matrix

		States of $X_n$			
		0	1	2	3
States of $X_{n-1}$	0	1	0	0	0
	1	0.4	0	0.6	0
	2	0	0.3	0	0.7
	3	0	0	0	1

Compute (i)  $P[X_2 = 1, X_1 = 2 / X_0 = 1]$  (5)

(ii)  $P[X_2 = 1, X_1 = 2, X_0 = 1]$

(iii)  $P[X_3 = 3, X_2 = 2, X_1 = 2, X_0 = 1]$

Given  $P[X_0 = 0] = P[X_0 = 2] = 0.3$  and

$P[X_0 = 1] = P[X_0 = 3] = 0.2$

7. (a) The analysis of time series is of great significance not only to economist and businessman but also to scientist, astronomist, geologist etc. Justify the statement (4)

(b) Calculate the 3-yearly moving averages of the production figures given below and draw the trend :

(6)

Year	1995	1996	1997	1998	1999	2000	2001	2002
Production (m.tonnes)	15	21	30	36	42	46	50	56
Year	2003	2004	2005	2006	2007	2008	2009	
Production (in tonnes)	63	70	74	82	90	95	102	

8. (a) What is a stationary time series? Why is stationarity important? How to identify non stationary time series data? (5)

(b) Explain the importance of normal distribution in decision making. Mention a few properties of Normal distribution.

(5)

9. (a) Twenty-five books are placed at random in a shelf. Find the probability that a particular pair of books shall be (i) Always together (ii) Never together

(5)

(b) X is normally distributed and the mean of X is 12 and S.D is 4. Find out the probability of the following :

(5)

(i)  $P(X \geq 20)$  (ii)  $P(X \leq 20)$  (iii)  $P(0 \leq X \leq 12)$

[Given  $P(0 \leq Z \leq 2) = 0.0228$  and  $P(0 \leq Z \leq 3) = 0.49865$ ]