



Register Number:

DATE:

ST. JOSEPH'S COLLEGE (AUTONOMOUS), BENGALURU-27
M.Sc. (BIG DATA ANALYTICS) – I SEMESTER
SEMESTER EXAMINATION - OCTOBER 2021
(Examination conducted in January-March 2022)
BDA 1221: PROBABILITY & STOCHASTIC PROCESS

TIME 2.5 HOURS

MAXIMUM MARKS 70

This Question Paper Contains FIVE Printed Papers And THREE Parts

PART A

Answer ALL questions (MCQs)

20 X 1 = 20

1. Which of the following is a discrete distribution?
 - i) Pareto
 - ii) Lognormal
 - iii) Poisson
 - iv) Uniform

2. Which of the following is always true?
 - i) $P(AB) = P(A) \cdot P(B)$
 - ii) $P(AB) = P(A/B) \cdot P(B)$
 - iii) $P(BA) = P(A/B) \cdot P(A)$
 - iv) $P(A \cup B) = P(A) + P(B)$

3. If A and B are disjoint events, then:
 - i) $P(A \cup B) = P(A) + P(B)$
 - ii) $P(AB) = 1$
 - iii) $P(A/B) = P(B)$
 - iv) $P(B/A) = P(B)$

4. An urn contains 3 red and 2 black balls. Three balls are drawn with replacement. What is the probability that they are all red?
 - i) $3/5$
 - ii) $27/125$
 - iii) 0
 - iv) $9/25$

5. If the p-value equals 0.02, and we want to test at 5% level of significance, then:
 - i) Reject the null hypothesis
 - ii) Don't reject the null hypothesis
 - iii) Don't have enough information to decide
 - iv) Always accept the null hypothesis

6. After looking at the numbers (4, 7, 9, 4, 0), Mrs Rao picks 4. She has picked:
- The arithmetic mean
 - The mode
 - The harmonic mean
 - The geometric mean
7. Let $X = \{1,2,3\}$ and $Y = \{4,5,6\}$. Then:
- $V(X) < V(Y)$
 - $V(X) > V(Y)$
 - $V(X) = V(Y) = 0$
 - $V(X) / V(Y) = 1$
8. If A and B are independent, $P(B/A) = 1/3$, $P(A/B) = 1/3$, then $P(A)$ equals
- 1/3
 - 2/3
 - 1/9
 - Data is insufficient to say anything
9. An urn contains 2 red and 2 white balls. Two balls are drawn with replacement. The chance that we draw one ball of each colour is:
- 1/4
 - 1/2
 - 1/8
 - 2/3
10. When someone talks of the "loss of memory" property, the name that springs to the mind is:
- Bayes
 - Fisher
 - Pearson
 - Markov
11. Just before dying, someone scribbled the number 1.05 and wrote that it is a statistical index. Which index is he NOT talking about?
- Mean deviation
 - Median
 - Correlation coefficient
 - Covariance
12. To test the null hypothesis of equality of three means, it is best to use:
- Z test
 - T test
 - F test
 - Paired T test

13. Which of these distributions have only one parameter?
- i) Normal
 - ii) Chi square
 - iii) Binomial
 - iv) Poisson
14. The probability of passing a test is $\frac{3}{4}$. If three students take the test, what is the probability that at least one student will fail?
- i) $\frac{3}{4}$
 - ii) 0
 - iii) $\frac{1}{4}$
 - iv) $\frac{37}{64}$
15. The set of all rational numbers is:
- i) Finite
 - ii) Uncountable
 - iii) Countably finite
 - iv) Countably infinite
16. A biased coin is thrice as likely to show "tails" than "heads". If this coin is tossed twice, what's the probability that both tosses yield "heads"?
- i) $\frac{1}{16}$
 - ii) $\frac{9}{16}$
 - iii) $\frac{1}{8}$
 - iv) 0
17. There are 24 students assembled in a classroom. Each student is asked to indicate his or her birth date. Then the probability that at least two students have the same birthday is between:
- i) 0.5 – 0.6
 - ii) 0.4 - 0.5
 - iii) Less than 0.4
 - iv) More than 0.6
18. Winston Churchill's statement: '*Never in the field of human conflict was so much been owed by so many to so few*' is indicative of which probability distribution?
- i) Uniform
 - ii) Weibull
 - iii) Exponential
 - iv) Pareto
19. If the R_0 value in the spread of an epidemic is 1.3, then the infection will
- i) Become stable after 13 days
 - ii) Rise in the days to come
 - iii) Fall in the days to come
 - iv) Can't say anything for sure
20. The Naïve-Bayes method is useful for:
- i) Error detection
 - ii) Stability estimation
 - iii) Classification
 - iv) Clustering

PART B

Answer ANY SIX questions

6 X 5 = 30

21. Old man Ramaiah is likely to live for five more years with a probability of $\frac{3}{4}$. His wife Sitamma, who has cancer, only has a $\frac{1}{4}$ chance of surviving these five years. What's the chance that both will be dead after five years? What's the chance that at least one of them will be dead after five years? [2+3]
22. Distinguish between a 'discrete' and 'continuous' distribution. Give an example each of a discrete and continuous distribution, mentioning in each case the associated parameters [3+2].
23. State the Binomial theorem after carefully listing its assumptions. If X is the outcome of 20 independent trials with a probability 0.4 of success, what is its mean and standard deviation?
24. A, B and C are playing ball. A is twice as likely to throw the ball to B than to C. B always throws the ball to C. C is equally likely to throw to A or B. Write down the probability transition matrix for this ball game.
25. Describe the one-way ANOVA test. To test the equality of three means, why is it better to use the F test, instead of a succession of 't' tests? [4+1]
26. Given that $P(A \cup B) = P(A) + P(B) - P(AB)$, what could be the expansion of $P(A \cup B \cup C)$?
27. State the central limit theorem (CLT). What are the implications of CLT in testing of hypotheses? [4+1]
28. Let $\{A_1, A_2 \dots A_n\}$ be n events that are "mutually exclusive" and "collectively exhaustive". Explain these two concepts, and then state (don't prove!) the Bayes theorem [1+1+3].

PART C

Answer ANY TWO questions

2 X 10 = 20

29. Carefully explain (with examples and illustrations, if needed) *any two* of the following three concepts or ideas:
 - i) Time series analysis and the idea of stationarity
 - ii) Using scatter plots to understand the correlation between two random variables X and Y , and obtaining the best regression line to estimate Y given X
 - iii) If X is a random variable, explain how to compute and interpret $E(X)$ and $V(X)$ [5+5]
30. Explain the idea of a Markov chain using the classical example of 'sunny', 'rainy' and 'cloudy' states. Discuss transition probabilities and the idea of a probability transition matrix (PTM). Introduce the idea of a two-step transition, and explain how a one-step PTM is related to a two-step PTM. [2+4+4]
31. The length of a chalk piece is supposed to be 3 inches. If the chalk piece is either too small or too large, the quality department will immediately reject it. Indicate a step-by-step procedure to test the hypothesis that the chalk piece length equals 3 inches. What will be the null and alternate hypotheses, and what will be the role of the p -value in testing the hypothesis.