No of Pages: 3 Course Code: 12XT23 / 12XW21

Roll No:

(To be filled in by the candidate)

PSG COLLEGE OF TECHNOLOGY, COIMBATORE - 641 004 SEMESTER EXAMINATIONS, MAY 2013

MSc – THEORETICAL COMPUTER SCIENCE / SOFTWARE ENGINEERING Semester: 2

12XT23 / 12XW21 PROBABILITY AND STATISTICS

Time : 3 Hours Maximum Marks : 100

INSTRUCTIONS:

- Group I, Group II and Group III questions should be answered in the Main Answer Book.
- 2. Ignore the box titled as "Answers for Group III" in the Main Answer Book.
- Answer ALL questions from GROUP I.
- Answer any 5 questions from GROUP H.
- Answer any ONE question from GROUP III.
- Statistical tables may be permitted.

GROUP - I Marks: 10 × 3 = 30

- You know that if one event A is contained in another event B, then P[A] ≤ P[B]. For any
 two events A and B, what does this imply about the relationship among P[A ∩ B], P[A],
 and P[A ∪ B]?
- Suppose A and B are two events such that P[A ∪ B] = 0.75 and P[A ∩ B] = 0.25. Is it possible to find P[A] and P[B]? Answer the same question if, in addition i) A and B are independent; ii) A and B are mutually exclusive.
- 3. A drunken man tries to open the door of his house. He has a bunch of 6 keys; exactly one of which can open the door. If he tries the keys one after another, what is the expected number of keys he will have to try before the door is opened?
- Assume that X and Y are random variables with joint probability density function f_{XY}(x,y).
 What is the value of E [E [X|Y]]?
- In a certain distribution, the first two moments about the value 2 are known to be 1 and 16.Determine the mean and variance.
- A random variable X has a mean of 4, a variance of 2, and an unknown probability distribution. Find an upper bound for P{|X - 4| ≥ 3}.
- To estimate the population mean μ, we could use sample mean x̄ and sample median M
 as estimators. Which of these will provide a better estimate for μ? Why?
- What is a maximum likelihood estimator? Give an example.
- 9. What do you mean by 'level of significance' in hypothesis testing and 'level of confidence' in interval estimation?
- 10. Suppose the breaking strengths of cables (in kg) is known to have a normal distribution with a standard deviation of 6 kg. How large a sample must be taken so as to be 95% confident that the sample mean breaking strength will not differ from the true mean breaking strength by more than 0.75 kg?

Marks : $5 \times 10 = 50$

No of Pages 13 Course Code : 12XT23 / 12XW21

GROUP-II

11. A simple binary communication channel carries messages by using only two signals, say 0 and 1. We assume that, for a given binary channel, 40% of the time a 1 is transmitted; the probability that a transmitted 0 correctly received is 0.90 and the probability that a transmitted 1 correctly received is 0.95. Determine i) the probability of a 1 being received; ii) given a 1 is received, the probability that 1 was transmitted; iii) the probability for an error.

- 12. A collection of human skull is divided into 3 classes according to the value of length—breadth index X. Skull with X < 75 are classified as short, those with 75 < X < 80 as medium and those with X > 80 as long. The percentages of skulls in the 3 classes in this collection are respectively 58%, 38% and 4%. Find approximately the mean and standard deviation of X on the assumption that X is normally distributed.
- The joint probability density function of two random variables X and Y is given by

$$f(x,y) = \begin{cases} k e^{-(x+y)}, & 0 < x < 1, \ 0 < y < 2 \\ 0 & otherwise \end{cases}$$

- i) Determine the value of k ii) Obtain the marginal probability densities of X and Y
- iii) Are X and Y independent? Why?
- 14. a) The number of weekly breakdowns of a machine is a random variable having Poisson distribution with λ = 0.3. What is the probability that the machine will i) have 3 breakdowns in a week; ii) operate without a breakdown for 2 consecutive days?
 - b) The service life, in years, of a hearing aid battery is a random variable having a Weibull distribution with $\alpha = 1/2$ and $\beta = 2$. i) How long can such a battery be expected to last? ii) What is the probability that such a battery will work after 2 years?
- 15. a) Let $x_1, ..., x_n$ be values of a random sample of size n from a population, show that $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i \bar{x})^2$ is an unbiased estimator of σ^2 , the population variance.
 - b) Consider the following sample of fat content (in percentage) of 10 randomly selected hot dogs: 25.2, 21.3, 22.8, 17.0, 29.8, 21.0, 25.5, 16.0, 20.9, 19.5. Assuming that these were selected from a normal population distribution, construct a 95% confidence interval for the population mean fat content.
- 16. a) A random sample of 150 recent donations at a certain blood bank reveals that 82 were type A blood. Does this suggest that the actual percentage of type A donations differs from 40%, the percentage of the population having type A blood? Carry out a test of the appropriate hypotheses using 1% LOS. Would your conclusion have been different if a 5% LOS had been used?
 - b) Businesses are increasingly placing orders online. Data were collected on the rates of correctly filled electronic orders by industry. Assume a sample of 700 orders provided the results as in Table 1. Using 5% LOS, test whether orders fulfillment is independent of industry.

TABLE 1

~6)	Pharma	Consumer	Computers	Telecom-
Correct	207	136	151	178
Incorrect	30	4	C/C 9	12

Marks : $1 \times 20 = 20$

Course Code: 12X123 / 12XW21

GROUP III

17. a) The owner of multiplex cinema halls in a mall like to estimate weekly gross revenue (Y) as function of Television and Newspaper advertising expenditures. Historical data for a sample of eight weeks are as in Table 2.

				V .
Gross Revenue	Televi		Newsp	aper Adv.
(in '000s)	(in	` '000s) 🧹	O (in	`'000s)_
960 _G	0	50		15-0
900		20		20
950	140	40	185	15
920	SCA	25 <	60	25
950	Ĭ	30 _ (3)		33. 3
940		35		23
940	140	25	145	42
940	60	30 <	60	25
	960 900 950 920 950 940 940	960 900 950 920 950 940 940	960 50 20 20 950 40 950 30 940 950 35 940 25	960 50 20 20 40 25 30 940 35 940 25

TABLE 2

- Develop an estimated regression equation with television adv. expenditure as the independent variable.
- Develop an estimated regression equation with both television adv. expenditure and newspaper adv. expenditure as the independent variables.
- iii) Interpret the regression equation coefficient for television adv. expenditure in each case.
- iv) What is the estimate of the gross weekly revenue when 35000 is spent on television adv. and 1800 is spent on newspaper adv.? [14]
- b) What are the limitations, errors, and caveats of using correlation and regression analysis?
- 18. a) When do we prefer data collection through experimentation? Why design of experiments is important in statistical study? Distinguish between Randomized Block Design (RBD) and Completely Randomized Design (CRD).
 - b) Three varieties of potato are planted each on four homogenous plots of land, each variety is treated with four different fertilizers. The yields (in tons) are as in Table 3.

Varieties

TABLE 3 Fertilizers

-12		V_1	$-V_2$	V_3
1	F_1	164	172	174
	F_2	155	157	147
	F ₃	159	166	158
	F4	158	157	153

Perform an analysis of variance and find whether there is any significant difference in the average yield of potatoes i) due to different fertilizers, and ii) of different varieties. [14]

/END

FD/RI