Roll Number:_____

Thapar Institute of Engineering & Technology, Patiala
Department of Computer Science and Engineering
END SEMESTER EXAMINATION

BE-III: ODD Semester (2020-21)   Course Code: UCS-410
Computer Sc. & Eng.; JAN 2021   Course Name: Probability & Statistics
Time: 2 Hours, M. Marks: 100, Weightage: 50   Name of Faculty: RKS

Note: Attempt any (and only) 5 problems. This is very important to attempt the parts of a problem in one place. You can assume a missing data suitably.

Useful Data (using R-commands): qchisq(0.1, 44) = 32.487, qchisq(0.9, 44) = 56.369, qnorm(0.9) = 1.282, qnorm(0.95) = 1.645, qnorm(0.975) = 1.960, pnorm(0.24) = 0.5948, pnorm(0.6, 0, 1) = 0.7257, qt(0.1, 44) = -1.301, qt(0.9, 44) = 1.301

1(a)	A random variable X may assume four values with probabilities: $(1 + 3x)/4$, $(1 - x)/4$, $(1 + 2x)/4$, and $(1 - 4x)/4$. For what values of x is this a probability distribution?	5
1(b)	Suppose X is a random variable with *CDF*: $F(x) = 0, x < 0; x(2 - x), 0 \le x \le 1; 1, x > 1$. (i) Find $E(X)$ (ii) Find $P(X < 0.4)$.	15
2(a)	Defaulting on a loan means failing to pay it back on time. The default rate among TIET students on their student loans is 1%. As a project you develop a test to predict which students will default. Your test is good but not perfect. It gives 4% false positives, i.e., predicting a student will default who in fact will not. It has a 0% false negative rate, i.e., predicting a student won't default who in fact will. (i) Suppose a random student tests positive. What is the probability that he will truly default? (ii) Someone offers to bet me the student in part (i) won't default. They want me to pay them Rs 100 if the student doesn't default and they will pay me Rs 400 if the student does default. Is this a good bet for me to take? A positive expected winnings means it's a good bet.	10
2(b)	A test is graded on the scale 0 to 1, with 0.55 needed to pass. Student scores are modeled by the density: $f(x) = 4x, 0 < x \le 0.5; 4 - 4x, 0.5 \le x < 1; 0$, elsewhere. (i) What is the probability that a random student passes the examination? (ii) What score is the 87.5 percentile of the distribution (here, you have to find the score such that 87.5% scores are less than or equal to that score)?	10
3(a)	Data was taken on height and weight from the entire population of 700 male students of a school, as given in the adjacent table. Let X encodes the weight, taking the values of a randomly chosen student: 0, 1, 2 for low, average, and high, respectively. Likewise, let Y encodes the height, taking values 0 and 1 for short and tall respectively. | Weight → Height ↓	Low	Average	High
---	---	---	---
Short	170	70	30
Tall	85	190	155 | (i) Determine the joint pmf of X and Y and the marginal pmf's of X and of Y. (ii) Are X and Y independent? (iii) Find the covariance of X and Y. (iv) Find the correlation of X and Y.	16

3(b)	The joint *pdf* of (X, Y) is given by $f(x, y) = (x^3y + xy^3)/2$. Are X and Y dependent or independent? You should support your answer with appropriate reasoning.	4
4(a)	Compute the mean and variance of a random variable whose distribution is uniform on the interval $[a, b]$.	10
4(b)	A poll is taken to determine the fraction p of the students who have a reasonable knowledge of R. There were only 2 options in the poll: (i) Yes, I do, (ii) No, I don't. (i) Assume $p = 0.5$. Use the Central Limit Theorem to estimate the probability that in a poll of 25 people, at least 14 people have a reasonable knowledge of R. (ii) With p unknown and n the number of random students polled, let $\bar{X}_n$ be the fraction of the polled students who have a reasonable knowledge of R. What is the smallest sample size n in order to have a 90% confidence that $\bar{X}_n$ is within 0.05 of the true value of p?	10
5(a)	We toss a coin 3 times. Give the set theoretical representation of the event: "Obtaining exactly two heads". What is the probability of this event?	6
5(b)	Suppose a researcher collects $x_1, x_2, ..., x_n$ *i.i.d.* measurements of the background radiation in Delhi. Further, suppose that these measurements follow a distribution with parameter σ, with *pdf* given by $f(x; \sigma) = x\sigma e^{-\sigma x^2/2}, \sigma > 0, x \geq 0$. Find the maximum likelihood estimate for σ.	14
6(a)	Suppose that X has the standard normal distribution. Find the *pdf* for X^2.	10
6(b)	We independently draw 100 data points from a normal distribution. (i) Suppose we know the distribution is $N(\mu, 4)$ $(4 = \sigma^2)$ and we want to test the null hypothesis H_0: $\mu = 3$ against the alternative hypothesis H_A: $\mu \neq 3$. If we want a significance level of $\alpha = 0.05$. What is our rejection region for H_0? You must clearly state what test statistic you are using. (ii) Suppose the 100 data points have sample mean 5. What is the p-value for this data? Here, you will reject H_0 or accept this.	10
7(a)	Using least square principle, obtain the normal equations for fitting a quadratic curve to the data points $\{(x_i, y_i), i = 1, 2, ..., n\}$	10
7(b)	Data is collected on the execution time of programs (X) at a computer center. We collect a dataset of size 45 with sample mean $\bar{x} = 5.0$ and sample standard deviation $s = 4.0$. Assume the data follows a normal random variable. (i) Find an 80% confidence interval for the population mean μ of X. (ii) Find an 80% Chi-Square confidence interval for the population variance σ^2.	10

Good Luck!