University of Moratuwa

Faculty of Information Technology

Department of Computational Mathematics

Honours Degree of Bachelor of Science in Information Technology & Management

Level 2 - Semester 2 - Examination

CM 2130 - Statistical Distributions and Estimation

Time Allowed: 3 Hours

June 2023

ADDITIONAL MATERIALS

Cumulative Standard Normal Distribution Table

INSTRUCTIONS TO CANDIDATES

1. This paper contains **five** questions on 7 pages (including the cover page and additional materials).
2. The total number of marks obtainable for this examination is 100. The marks assigned for each question & sections thereof are included in square brackets.
3. This examination accounts for 60% of the module assessment.
4. This is a closed-book examination.
5. Answer **ALL** questions.
6. **All the necessary steps** for the answers should be **clearly indicated**.
7. Calculators are NOT ALLOWED.

- END OF INSTRUCTIONS -

Question 01 [Total Marks Allocated: 20 Marks]

(a) A continuous random variable, X has a probability density function, $f(x)$, defined by

$$f(x) = \begin{cases} ax + bx^2, & \text{for } 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

and $E(X) = \frac{1}{2}$. Determine

i. the constants a and b. [5 marks]

ii. the cumulative distribution function, $F(x)$ of X [4 marks]

iii. the variance $\text{Var}(X)$. [5 marks]

(b) City trains on a particular line have a probability of 0.05 failure between two stations. Supposing that the failures are all independent, what is the probability that out of 10 journeys between these two stations more than 8 do not have a breakdown? (No need to simplify the answer.) [6 marks]

Question 02 [Total Marks Allocated: 20 Marks]

In a study of exercise,' a large group of male runners walk on a treadmill for 6 minutes. Their heart rates in beats per minute (bpm) at the end vary from runner according to the $N(104, 12.5)$.

(i) Find z-scores for 120 bpm for a runner. [2 marks]

(ii) What is the median of the heart rates of runners? [2 marks]

(iii) What proportion of the runners has heart rates above 120? [4 marks]

(iv) Find percent of runners with heart rates between 90 and 110. [4 marks]

(v) Find the heart rate with 20% of the runners falling below it. [4 marks]

(vi) Find the heart rate with 5% of the runners falling above it. [4 marks]

Continued.

Question 03 [Total Marks Allocated: 20 Marks]

(a) Briefly explain the basic principles of experimental design. [3 marks]

(b) Why is it important to conduct "randomized comparative experiments"? Explain. [2 marks]

(c) What are the effects of repeated exposure to an advertising message? The answer may depend on the length of the ads and on how often it is repeated. An experiment investigated this question for 120 undergraduate students. All selected students viewed a 40-minute television program that included ads for a digital camera. Some students saw a 30-second commercial; others, a 90-second version. The same commercial was shown 1,3, or 5 times during the program. After viewing, all students answered questions about their recall of the ads, their attitude toward the camera, and their intention to purchase. Identify the following.

i. What is the subject? How many? [2 marks]

ii. How many factors are in the study? What are those? [2 marks]

iii. How many treatments would be in the study? What are the treatments? [3 marks]

iv. What is the response variable(s)? [2 marks]

v. If we decided to use a completely randomized design, draw the outline of this. [3 marks]

vi. Males and females may show different responses for this study. If we have 54 men and 66 women, how do we assign the individuals in to the treatments using a randomized block design? Draw the outline. [3 marks]

Question 04 [Total Marks Allocated: 20 Marks]

a) Bottles of a popular cola are supposed to contain 300 millilitres (ml) of cola. There is some variation from bottle to bottle because the filling machinery is not perfectly precise. The distribution of the contents is Normal with a standard deviation of 3 ml. An inspector suspects that the bottle is under-filling. He measures the contents of six bottles. The results are 299.4, 297.5, 301.0, 298.9, 300.2, and 297.0. Is this convincing evidence that the mean content of cola bottles is less than the advertised 300 ml?

(i) State the hypothesis that you will test. [3 marks]

(ii) Calculate the test statistic. [3 marks]

(iii) Find the p-value. [3 marks]

(iv) Do we reject the null hypothesis at the significance level 5%?
State your conclusion. [3 marks]

b) It was claimed that 75% of all dentists recommend a certain brand of gum for their gum-chewing patients. A consumer group doubled this claim and decided to test $H_0 : p = 0.75$ against the alternative hypothesis $H_1 : p < 0.75$, where p is the proportion of dentists who recommend that brand of gum. A survey of 390 dentists found that 273 recommended the given brand of gum.

(i) Which hypothesis would you accept if the significance level is $\alpha = 0.10$? [4 marks]

(ii) Which hypothesis would you accept if the significance level is $\alpha = 0.01$? [4 marks]

Continued.

Question 05 [Total Marks Allocated: 20 Marks]

(a) A laboratory scale is known to have a standard deviation of $\sigma = 0.001$ gram in repeated weightings. Scale readings in repeated weightings are Normally distributed with a mean equal to the true weight of the specimen. Three weightings of a specimen on this scale give 3.412, 3.416, and 3.414 grams.

i. Calculate a 95% confidence interval for the true weight of this specimen. [4 marks]

ii. In part (a), what would happen to the confidence interval if you want to calculate a 99% confidence interval (no need to calculate)? [2 marks]

iii. Another specimen is weighted eight times on this scale. The average weight is 4.1602 grams. Find a 95% confidence interval for the true weight of this specimen. Assume that σ is remain unchanged. [4 marks]

(b) Let $\{X_1, X_2, \ldots, X_n\}$ be a random sample from a distribution with probability density function

$$f(x) = \begin{cases} \lambda^2 x e^{-\lambda x}, & \text{for } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\lambda > 0$ is unknown. Find the maximum likelihood estimator of λ. [10 marks]

END OF PAPER

Continued.