INDIAN INSTITUTE OF SCIENCE EDUCATION & RESEARCH KOLKATA

Statistics - I (MA3205) – End-Semester Exam

Date: 9th May, 2019 Duration: 2 hours and 30 minutes

Maximum points that you can score is 100. Good luck!

Question 1 (40 points)

Let $x_1, x_2, \ldots, x_n$ be a known and fixed set of real numbers. We have a random sample $Y_1, Y_2, \ldots, Y_n$ that depends on $x_1, x_2, \ldots, x_n$. We model this dependence in the following manner: $Y_i = a + bx_i + \epsilon_i$, where $\epsilon_i,\ i = 1, 2, \ldots, n$ are i.i.d. $\mathrm{N}(0, \sigma^2)$ variables. Here, $\sigma > 0$ is an unknown parameter. The values of the parameters a and b are also unknown, and we wish to estimate them by solving the following optimization

$$(\hat{a}, \hat{b}) = arg \min_{a,b \in \mathbb{R}} n^{-1} \sum_{i=1}^{n} (Y_i - a - bx_i)^2.$$

(i) Without explicitly calculating $\hat{a}$ and $\hat{b}$, prove that $\hat{a}$ and $\hat{b}$ are in fact the MLEs of a and b based on $\{Y_1, Y_2, \ldots, Y_n\}$.

[Hint: Y_i's are independent Normal random variables.]

(ii) Find $\hat{a}$ and $\hat{b}$.

(iii) Show that $\hat{a}$ and $\hat{b}$ are unbiased estimators of a and b, respectively.

(iv) Find the marginal distributions of $\hat{a}$ and $\hat{b}$.

(v) Show that $\hat{a}$ and $\hat{b}$ jointly have a bivariate Normal distribution (no need to find the parameters).

Question 2 (10 points)

Assume that the length of a phone-call (in minutes) of an individual follows an Exponential distribution with an unknown parameter $\lambda > 0$ with density function $f_\lambda(x) = \lambda \exp(-\lambda x),\ x > 0$. However, when the phone company calculates the length of a phone-call, it always considers the nearest integer greater than or equal to the actual length. For example, a 11.07 minutes long phone-call will have a call-length of 12 minutes in the records of the phone company. Suppose you have the data on the lengths of n independent phone-calls $T_1.T_2, \ldots, T_n$ of an individual as reported by the phone company. Based on this data, find a sufficient statistic for λ.

(ANIRVAN CHAKRABORTY)

(DMS)

Please turn over

Question 3 (10 points)

Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from a distribution with pdf f_θ, where $\theta \in \Theta = \{0, 1\}$.
If $\theta = 0$, then

$$f_0(x) = \begin{cases} 1 & \text{if } 0 < x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

while if $\theta = 1$, then

$$f_\theta(x) = \begin{cases} (2\sqrt{x})^{-1} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find the MLE of θ based on $X_1, X_2, \ldots, X_n$.

Question 4 (20 points)

Consider the following data set.

x-values	1	2	3	4	5	6
y-values	18.6	15	10.6	7.2	3.4	-1.1

Find the equation of the line $y = a + bx$ that passes through the point $(5, 3)$ and minimizes

$$\sum_{i=1}^{6} |y_i - a - bx_i|,$$

where (x_i, y_i) denotes the i-th observation for $i = 1, 2, \ldots, 6$.
[Hint: Find a relation between a and b, and substitute this relation in the minimization problem.]

Question 5 (20 points)

Suppose you have only one observation X with probability mass function f_θ, $\theta \in \{0, 1\}$, where f_0 and f_1 are given by

x	1	2	3	4	5
$f_0(x)$	0.01	0.04	0.02	0.84	0.09
$f_1(x)$	0.04	0.20	0.60	0.14	0.01

(i) Find a level-α MP randomized test for $H_0 : \theta = 0$ versus $H_1 : \theta = 1$ when $\alpha = 0.05$.
(ii) Consider the test function

$$\phi_*(X) = \begin{cases} 1 & \text{if } X = 1 \text{ or } 3, \\ 0.5 & \text{if } X = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Show that ϕ_* is also a level-α MP randomized test for the above problem.

– END OF THE EXAM –