CSE 312: Foundations of Computing II Autumn 2022

Problem Set 8

Due: Friday, December 2, by 11:59pm.

Instructions

Solutions format, collaboration policy, and late policy. See PSet 1 for further details. The same requirements and policies still apply. Also follow the typesetting instructions from the prior PSets.

Submit your solution via Gradescope as usual.

Task 1 – Lazy Grader [12 pts]

Prof. Lazy decides to assign final grades in CSE 312 by ignoring all the work the students have done and instead using the following probabilistic method: each student independently will be assigned an A with probability θ, a B with probability 2θ, a C with probability $\frac{1}{2}$, and an F with probability $\frac{1}{2} - 3\theta$. When the quarter is over, you discover that only 10 students got an A, 35 got a B, 40 got a C, and 15 got an F.

Find the maximum likelihood estimate for the parameter θ that Prof. Lazy used. Give an exact answer as a simplified fraction.

Task 2 – The Place That's Best [12 pts]

Let $y_1, y_2, ...y_n$ be i.i.d. samples of a random variable from the family of distributions $Y(\theta)$ with densities

$$f(y;\theta) = \frac{1}{2\theta} \exp\left(-\frac{|y|}{\theta}\right),$$

where $\theta > 0$. Find the MLE for θ in terms of $|y_i|$ and n.

Task 3 – Maximum Likelihood Estimators [26 pts]

a) Let $x_1, \ldots, x_n$ be i.i.d samples that follow a Two(θ) distribution with unknown parameter $\theta \in [0, 1]$, where the probabilities from the family are given by

$$\mathbb{P}(x;\theta) = \begin{cases} (1-\theta)^2 & x = 0 \\ 2\theta(1-\theta) & x = 1 \\ \theta^2 & x = 2 \end{cases}$$

Suppose that in the sample there are n_0 0's, n_1 1's, and n_2 2's.

What is the maximum likelihood estimator for θ in terms of n, n_0, n_1, n_2?

b) Let $x_1, \ldots, x_n$ be i.i.d. samples from a random variable that follow a so-called Borel distribution with unknown parameter θ, i.e., a distribution from the family

$$\mathbb{P}(k;\theta) = \frac{e^{-\theta k}(\theta k)^{k-1}}{k!},$$

where $0 < \theta \leqslant 1$ is a real number, and $k \geqslant 1$ is an integer.

What is the maximum likelihood estimator for θ?

c) If the samples from the Borel distribution are 5, 7, 10, 2, 7, 5, 12, 13, 11, what is the maximum likelihood estimator for θ? Give an exact answer as a simplified fraction.

Task 4 – Continuous MLE [24 pts]

a) Let $x_1, x_2, \ldots, x_n$ be independent samples from an exponential distribution with unknown parameter λ. What is the maximum likelihood estimator for λ?

b) Suppose that $x_1, \ldots, x_n$ are i.i.d. realizations (aka samples) from the model

$$f(x; \theta) = \begin{cases} \theta x^{\theta - 1} & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find the maximum likelihood estimate for θ.

Task 5 – Elections [14 pts]

Individuals in a certain country are voting in an election between 3 candidates: A, B and C. Suppose that each person makes their choice independent of others and votes for candidate A with probability θ_1, for candidate B with probability θ_2 and for candidate C with probability $1 - \theta_1 - \theta_2$. (Thus, $0 \leqslant \theta_1 + \theta_2 \leqslant 1$.) The parameters θ_1, θ_2 are unknown.
Suppose that $x_1, \ldots, x_n$ are n independent, identically distributed samples from this distribution. (Let $n_A =$ number of x_i's equal to A, let $n_B =$ number of x_i's equal to B, and let $n_C =$ number of x_i's equal to C.) What are the maximum likelihood estimates for θ_1 and θ_2 in terms of n_A, n_B, and n_C?
(You don't need to check second order conditions.)

Task 6 – (Un)biased Estimation [12 pts]

a) Let $x_1, \ldots, x_n$ be independent samples from the Poisson distribution with parameter θ. In the concept check, we have seen that the MLE estimator for θ is $\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} x_i$. Is this estimator unbiased?

b) Let $x_1, \ldots, x_n$ be independent samples from Unif$(0, \theta)$, the continuous uniform distribution on $[0, \theta]$. Then, consider the estimator $\hat{\theta}_{\text{first}} = 2x_1$, i.e., our estimator ignores the samples $x_2, \ldots, x_n$ and just outputs twice the value of the first sample.

Is $\hat{\theta}_{\text{first}}$ unbiased?