Probability and Statistics

Homework 11

Q1: Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from the following distribution

$$f_X(x) = \begin{cases} \theta\left(x - \frac{1}{2}\right) + 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\theta \in [-2, 2]$ is an unknown parameter. We define the estimator $\hat{\Theta}_n$ as

$$\hat{\Theta}_n = 12\overline{X} - 6$$

to estimate θ.

a. Is $\hat{\Theta}_n$ an unbiased estimator of θ?
b. Is $\hat{\Theta}_n$ a consistent estimator of θ?
c. Find the mean squared error (MSE) of $\hat{\Theta}_n$.

Q2: Let $X_1, \ldots, X_4$ be a random sample from a *Geometric(p)* distribution. Suppose we observed $(x_1, x_2, x_3, x_4) = (2, 3, 3, 5)$. Find the likelihood function using

$$P(X_i = x_i; p) = p(1-p)^{x_i - 1}$$

as the PMF.

Q3: A data scientist is analyzing a large stream of incoming data, $X_1, X_2, \ldots, X_n$, where each X_i is an independent draw from a distribution with mean μ and variance $\sigma^2 < \infty$.

Two different estimators are proposed to estimate the mean μ as the sample size n grows:

(a) **Estimator A (The Sample Mean):** $\hat{\mu}_{A,n} = \frac{1}{n}\sum_{i=1}^{n} X_i$
(b) **Estimator B (A "Weighted" Mean):** $\hat{\mu}_{B,n} = \frac{1}{n}\sum_{i=1}^{n}\left(1 + \frac{1}{i}\right) X_i$

Which estimator, if any, is a consistent estimator for μ? Justify your answer by checking if the limit of its Mean Squared Error (MSE) converges to 0.

Q4: (For conceptual understanding) Often when working with maximum likelihood functions, out of ease we maximize the log-likelihood rather than the likelihood to find the maximum likelihood estimator. Why is maximizing $L(\mathbf{x}; \theta)$ as a function of θ equivalent to maximizing $\log L(\mathbf{x}; \theta)$?

Q5: Let X be one observation from a $N(0, \sigma^2)$ distribution.

a. Find an unbiased estimator of σ^2.
b. Find the log likelihood, $\log(L(x; \sigma^2))$, using

$$f_X(x; \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}$$

as the PDF.

c. Find the Maximum Likelihood Estimate (MLE) for the standard deviation σ, $\hat{\sigma}_{ML}$.

Q6: Consider the simple linear regression model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where ϵ_i's are independent $N(0, \sigma^2)$ random variables. The estimators are:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}$$

where

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2, \quad S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(Y_i - \bar{Y})$$

(a) **Show that $\hat{\beta}_1$ is a normal random variable.**

(b) **Show that $\mathbb{E}[\hat{\beta}_1] = \beta_1$.**

(c) **Show that Var$(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$.**

Q7: Again consider the simple linear regression model from Problem 1.

(a) **Show that $\hat{\beta}_0$ is a normal random variable.**

(b) **Show that $\mathbb{E}[\hat{\beta}_0] = \beta_0$.**

(c) **Show that Cov$(\hat{\beta}_1, Y_i) = \frac{x_i - \bar{x}}{S_{xx}} \sigma^2$.**

(d) **Show that Cov$(\hat{\beta}_1, \bar{Y}) = 0$.**

(e) **Show that Var$(\hat{\beta}_0) = \frac{\sum_{i=1}^{n} x_i^2}{n S_{xx}} \sigma^2$.**

Q8: Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample with unknown mean $\mathbb{E}[X_i] = \mu$, and unknown variance $\mathrm{Var}(X_i) = \sigma^2$. Suppose that we would like to estimate $\theta = \mu^2$. We define the estimator $\hat{\Theta}$ as

$$\hat{\Theta} = (\bar{X})^2 = \left[\frac{1}{n} \sum_{k=1}^{n} X_k \right]^2$$

to estimate θ. Is $\hat{\Theta}$ an unbiased estimator of θ? Why?

Q9: **Estimating the parameter of a uniform random variable I.** We are given i.i.d. observations $X_1, \ldots, X_n$ that are uniformly distributed over the interval $[0, \theta]$. What is the ML estimator of θ? Is it consistent? Is it unbiased or asymptotically unbiased?

Q10: **Estimating the parameter of a uniform random variable II.** We are given i.i.d. observations $X_1, \ldots, X_n$ that are uniformly distributed over the interval $[\theta, \theta + 1]$. Find a ML estimator of θ. Is it consistent? Is it unbiased or asymptotically unbiased?