Homework 2

ESE 402/542

Due October 1st, 2024 at 11:59pm

- **Type or scan your answers as a single PDF file and submit on Gradescope.**

Problem 1

Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables in a sample with the density function

$$f(x|\sigma) = \frac{1}{2\sigma} \exp\left\{-\frac{|x|}{\sigma}\right\}$$

(a) Use method of moments to estimate σ?

(b) Find the MLE estimate of σ?

(c) What is the asymptotic variance of the MLE?

Problem 2

Consider the following sample data:

$$x_1, x_2, \ldots, x_{30} = 0, 1, 2, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 2, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 2, 0, 1, 0, 0, 1$$

We know that the data is distributed i.i.d., and it can take three values 0, 1, and 2. It takes the value 1 with probability p_1, and it takes the value 2 with probability p_2.

(a) Write down the likelihood function obtained from the sample data.

(b) For the setting of part (a), compute the maximum likelihood estimate for p_1 and p_2. (Your final answers should be real numbers.)

(c) Using the estimate that you obtained in part (b), provide a 95% confidence interval for p_1.

Problem 3

Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. distributed in a sample with pdf

$$f(x|\theta) = \begin{cases} e^{-(x-\theta)}, & \text{if } x \geq \theta \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the method of moments estimate of θ.

(b) Find the MLE of θ. (Hint: Be careful, and don't differentiate before thinking. For what values of θ is the likelihood positive?)

Problem 4

Suppose $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\lambda)$. Given that the random variables are i.i.d., for $\theta = \exp(-\lambda)$:

1. Find an unbiased estimator of θ. (Note that it may not be the best estimator. Any unbiased estimator is fine.)

2. Find the variance of the unbiased estimator you found and compare with the Cramer Rao lower bound.

Problem 5

We have access to a file consisting of $n = 10^4$ numbers. The numbers are either 1, 2, or 3. Moreover, the value 1 appears $n_1 = 2600$ times in the file, the value 2 appears $n_2 = 5200$ times, and the value 3 appears $n_3 = 2200$ times. We know that these numbers are generated i.i.d. according to an unknown distribution.

(a) Let μ denote the mean of the distribution. Estimate the value of μ from sample data provided in the file and provide a 95% confidence interval.

(b) Assume now that the generating distribution of the data has the following form:

$$X = \begin{cases} 1, & \text{with probability } p_1, \\ 2, & \text{with probability } p_2, \\ 3, & \text{with probability } 1 - (p_1 + p_2). \end{cases}$$

We would like to estimate the value of the parameters p_1 and p_2. Consider the following estimator for the value of

$$p_1(X_1, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} 1\{X_i = 1\}$$

(where $1\{A\}$ takes value 1 if A is true, and 0 otherwise.)

Compute the estimate p_1 from the sample data provided in the file. Is this estimator an unbiased estimator for p_1? Justify your answer.

(c) Use the method of moments to estimate the value of p_1 and p_2 (you should compute the estimate from the sample data).

(d) Now, assume that the precise value of p_1 is given as $p_1 = \frac{1}{4}$. As a result, we now know that the distribution of the data has the form:

$$X = \begin{cases} 1, & \text{with probability } \frac{1}{4}, \\ 2, & \text{with probability } p_2, \\ 3, & \text{with probability } \frac{3}{4} - p_2 \end{cases}$$

We would like to estimate the value of the parameter p_2 from data. Find the maximum likelihood estimator for p_2 and provide a 95% confidence interval.

Problem 6

Download data$_H W2.csvandloaditintoPython.Thenumbersareobservationsdrawni.i.d.fromanexponentia$
Include your code in your homework write up.

(a) Compute estimates for the sample mean and sample variance without using inbuilt functions. Compare your answers with inbuilt numpy functions.

(b) Suppose now that the standard deviation is known to be 4. Compute a 90% confidence interval for the population mean. (Python libraries can be used to calculate the confidence interval.)

Problem 7

In this problem, we will use Python for simulation of random experiments.

(a) Let $X_1, X_2, \ldots, X_n$ be independent $N(0,1)$ random variables and $\overline{X}_n$ be their sample mean. For $n = 10, 100,$ and 1000, simulate this problem 2000 times and plot the histogram of the values of $\overline{X}_n$ (you need to plot three histograms; one for each choice of n). How does Var($\overline{X}_n$) relate to the histogram of the values of $\overline{X}_n$?

(b) In a symmetric random walk on a line, at step $n = 0$, we start from point $X_0 = 0$. At each step n, we flip an unbiased coin and set

$$X_{n+1} = \begin{cases} X_n + 1, & \texttt{Heads} \\ X_n - 1, & \texttt{Tails} \end{cases}$$

By running the random walk 1000 times, for $n = \{1, 2, \ldots, 50\}$, estimate Var(X_n) and plot it as a function of n. Do the same for $\mathbb{E}(X_n)$.

(c) For the previous part, compute $\mathbb{E}(X_n)$ and Var(X_n) theoretically and compare them with your plots.