PS1

Samuel Olson

2025-01-27

Overview

- Q1: Base
- Q2: Base
- Q3: Base
- Q4: Base
- Q5: Base

Problem 1

Find the method of moment estimators (MMEs) of the unknown parameters based on a random sample $X_1, X_2, \ldots, X_n$ of size n from the following distributions:

a) Negative Binomial $(3, p)$, unknown p:

The Negative Binomial $(3, p)$ distribution has a mean of $\mu = \frac{3(1-p)}{p}$. Based on the sample $X_1, X_2, \ldots, X_n$, the sample mean is $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Equating the population mean to the sample mean, we have:

$$\frac{3(1-p)}{p} = \bar{X}.$$

Rearranging for p, we get:

$$p = \frac{3}{3+\bar{X}}.$$

Thus, the method of moments estimator for p is:

$$\hat{p} = \frac{3}{3+\bar{X}}.$$

b) Double Exponential (μ, σ), unknown μ and σ:

The Double Exponential distribution has a mean μ and variance $2\sigma^2$. Based on the sample $X_1, X_2, \ldots, X_n$, the sample mean is $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, and the sample variance is $S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$.

Equating the population mean and variance to their sample counterparts, we have:

$$\mu = \bar{X} \quad \text{and} \quad 2\sigma^2 = S^2.$$

Solving for σ, we get:

$$\sigma = \sqrt{\frac{S^2}{2}}.$$

Thus, the method of moments estimators are:

$$\hat{\mu} = \bar{X} \quad \text{and} \quad \hat{\sigma} = \sqrt{\frac{S^2}{2}}.$$

See "Table of Common Distributions" in Casella & Berger (pages 623–623) for the definitions/properties of the above distributions.

Problem 2

Problem 7.1, Casella & Berger:

Hint: For context, there is only one (discrete) data observation X which has possible outcomes as $0, 1, 2, 3, 4$. For a given outcome x of X, the likelihood $(L(\theta) \equiv f(x|\theta)$ is given by the pmf as a function of $\theta \in \Theta \equiv \{1, 2, 3\}$.

One observation is taken on a discrete random variable X with pmf $f(x|\theta)$, where $\theta \in \{1, 2, 3\}$. Find the MLE of θ.

x	$f(x\|1)$	$f(x\|2)$	$f(x\|3)$
0	$\frac{1}{3}$	$\frac{1}{4}$	0
1	$\frac{1}{3}$	$\frac{1}{4}$	0
2	0	$\frac{1}{4}$	$\frac{1}{4}$
3	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{2}$
4	$\frac{1}{6}$	0	$\frac{1}{4}$

After this snippet, you would write the step-by-step process for finding the MLE of θ, leveraging the provided pmf table. Here's what should follow:

To find the maximum likelihood estimator (MLE) of θ, we use the given pmf table. The likelihood function is:

$$L(\theta) = f(x|\theta),$$

where x is the observed value of X, and $\theta \in \{1, 2, 3\}$.

1. Identify the Observed Value x: For a specific observation x, the likelihood function $L(\theta)$ is directly given by $f(x|\theta)$ for each $\theta \in \{1, 2, 3\}$.

2. Extract the Values from the Table: For a given x, use the table to find the corresponding values of $f(x|1)$, $f(x|2)$, and $f(x|3)$.

3. Maximize $f(x|\theta)$: Compare $f(x|1)$, $f(x|2)$, and $f(x|3)$ for the observed x. The MLE $\hat{\theta}$ is the value of θ that maximizes $f(x|\theta)$.

- For $x = 0$ or $x = 1$, $f(x|1) = \frac{1}{3}$ is the largest value, so $\hat{\theta} = 1$.
- For $x = 2$, both $f(x|2)$ and $f(x|3)$ are equal to $\frac{1}{4}$, so $\hat{\theta} = 2$ or 3.
- For $x = 3$, $f(x|3) = \frac{1}{2}$ is the largest value, so $\hat{\theta} = 3$.
- For $x = 4$, $f(x|3) = \frac{1}{4}$ is the largest value, so $\hat{\theta} = 3$.

The MLE $\hat{\theta}$ for each possible observed value x is summarized as follows:

x	$\hat{\theta}$
0	1
1	1
2	2 or 3
3	3
4	3

The MLE $\hat{\theta}$ for any observed x is determined as:

$$\hat{\theta} = \underset{\theta \in \{1,2,3\}}{\operatorname{argmax}} f(x|\theta).$$

At $x = 2$, $f(x|2) = f(x|3) = 1/4$ are both maxima, so both $\hat{\theta} = 2$ or $\hat{\theta} = 3$ are MLEs.

Problem 3

An indicator function $I(A)$ of an event A has the form:

$$I(A) = \begin{cases} 1, & \text{if event } A \text{ holds true,} \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that $A_1, \ldots, A_n$ are n separate events. Show that:

$$\prod_{i=1}^{n} I(A_i) = I(B),$$

where B is the event that $B = \bigcap_{i=1}^{n} A_i$.

The event $B = \bigcap_{i=1}^{n} A_i$ holds true if and only if all events $A_1, A_2, \ldots, A_n$ are true simultaneously.

So we need to prove both directions to conclude.

By the definition of the indicator function:

$$I(B) = I\left(\bigcap_{i=1}^{n} A_i\right) = \begin{cases} 1, & \text{if all } A_i \text{ hold true, i.e., } A_1 \cap A_2 \cap \ldots \cap A_n, \\ 0, & \text{otherwise.} \end{cases}$$

For the product $\prod_{i=1}^{n} I(A_i)$:

$$\prod_{i=1}^{n} I(A_i) = I(A_1) \cdot I(A_2) \cdot \ldots \cdot I(A_n).$$

Each $I(A_i)$ is 1 if A_i is true, and 0 otherwise. The product $\prod_{i=1}^{n} I(A_i)$ will equal 1 if and only if all $I(A_i) = 1$, i.e., all events A_i are true. If any A_i is false, then $I(A_i) = 0$ for that i, making the entire product 0.

Therefore, the product $\prod_{i=1}^{n} I(A_i)$ is therefore 1 if and only if all events $A_1, A_2, \ldots, A_n$ are true, which matches the definition of $I(B)$. If any event A_i is false, the product is 0, again matching the behavior of $I(B)$.

Thus, we have shown that:

$$\prod_{i=1}^{n} I(A_i) = I(B), \quad \text{where } B = \bigcap_{i=1}^{n} A_i.$$

Problem 4

Maximum-Likelihood & Indicator Functions

Given a random sample $X_1, \ldots, X_n$ from a pdf/pmf $f(x|\theta)$, $\theta \in \Theta \subset \mathbb{R}$, we know that the likelihood function will generically be

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta), \quad \theta \in \Theta,$$

but there's one subtle point to again highlight about how to exactly write the likelihood expression depending on the support of $f(x|\theta) > 0$.

- Recall the support or range of $f(x|\theta)$ is a set

$$S_\theta = \{x \in \mathbb{R} : f(x|\theta) > 0\},$$

which could possibly depend on $\theta \in \Theta$. For example, an exponential distribution has a pdf

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & x > 0, \\ 0, & \text{otherwise}, \end{cases}$$

with a parameter $\theta > 0$, and in this case the support $S_\theta = (0, \infty)$ doesn't depend on $\theta \in \Theta = (0, \infty)$.

On the other hand, the pdf (1):

(1)

$$f(x|\theta) = \begin{cases} \frac{2x}{\theta^2}, & 0 < x \leq \theta, \\ 0, & \text{otherwise}, \end{cases}$$

with parameter $\theta > 0$, does have a support $S_\theta = (0, \theta]$ depending on $\theta \in \Theta = (0, \infty)$.

- It's always true that $f(x|\theta) = f(x|\theta) I(x \in S_\theta)$ for all $x \in \mathbb{R}$ and so always true that (2):

(2)

$$L(\theta) = \prod_{i=1}^{n} [f(x_i|\theta) I(x_i \in S_\theta)] = \left(\prod_{i=1}^{n} f(x_i|\theta) \right) I(x_1, \ldots, x_n \text{ are all in } S_\theta).$$

Questions

(a) If $X_1, \ldots, X_n$ are a random sample from an exponential pdf $f(x|\theta)$, $\theta > 0$ (and so $X_1, \ldots, X_n$ are positive values), show that the likelihood function (2) can be written as

$$L(\theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta},$$

and that the MLE of θ is $\bar{X}_n$. (Message here: The support of an exponential doesn't depend on θ, so we don't have to worry about indicating the support.)

The exponential pdf is given by:

$$f(x|\theta) = \frac{1}{\theta} e^{-x/\theta}, \quad x > 0,\, \theta > 0.$$

The likelihood function for a random sample $X_1, \ldots, X_n$ is:

$$L(\theta) = \prod_{i=1}^n f(x_i|\theta).$$

Substituting the pdf:

$$L(\theta) = \prod_{i=1}^n \left(\frac{1}{\theta} e^{-x_i/\theta} \right).$$

Separate the product:

$$L(\theta) = \left(\prod_{i=1}^n \frac{1}{\theta} \right) \left(\prod_{i=1}^n e^{-x_i/\theta} \right).$$

Simplify:

$$L(\theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta}.$$

This is the likelihood function.

To find the MLE of θ, we maximize the log-likelihood function:

$$\ell(\theta) = \log L(\theta) = -n \log \theta - \frac{\sum_{i=1}^n x_i}{\theta}.$$

Differentiate $\ell(\theta)$ with respect to θ:

$$\frac{\partial \ell(\theta)}{\partial \theta} = -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i}{\theta^2}.$$

Set the derivative to 0 for maximization:

$$-\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i}{\theta^2} = 0.$$

Rearrange:

$$\frac{\sum_{i=1}^n x_i}{\theta^2} = \frac{n}{\theta}.$$

Multiply through by θ^2:

$$\sum_{i=1}^{n} x_i = n\theta.$$

Solve for θ:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X}_n.$$

The likelihood function is:

$$L(\theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^{n} x_i/\theta}.$$

The MLE of θ is:

$$\hat{\theta} = \bar{X}_n.$$

(b) If $X_1, \ldots, X_n$ are a random sample from the pdf

$$f(x|\theta) = \begin{cases} \frac{2x}{\theta^2}, & 0 < x \le \theta, \\ 0, & \text{otherwise}, \end{cases}$$

(and so $X_1, \ldots, X_n > 0$ are less than or equal to θ), show that the likelihood function (2) can be written as

$$L(\theta) = \frac{2^n \prod_{i=1}^{n} x_i}{\theta^{2n}} I\left(\max_{1 \le i \le n} x_i \le \theta\right),$$

and that the MLE of θ is $\max_{1 \le i \le n} X_i$. (Message here: The support in this case depends on θ, so we should think about indicator functions in writing the likelihood.)

The given pdf is:

$$f(x|\theta) = \begin{cases} \frac{2x}{\theta^2}, & 0 < x \le \theta, \\ 0, & \text{otherwise}. \end{cases}$$

The likelihood function for a random sample $X_1, \ldots, X_n$ is:

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta).$$

Substituting the pdf:

$$L(\theta) = \prod_{i=1}^{n} \frac{2x_i}{\theta^2} \cdot I(x_i \le \theta).$$

Simplify the product:

$$L(\theta) = \frac{2^n \prod_{i=1}^n x_i}{\theta^{2n}} \cdot I\left(x_1 \leq \theta, x_2 \leq \theta, \ldots, x_n \leq \theta\right).$$

The indicator function $I(x_1 \leq \theta, \ldots, x_n \leq \theta)$ is equivalent to $I(\max_{1\leq i\leq n} x_i \leq \theta)$ because θ must be greater than or equal to all observed values for the likelihood to be nonzero. Therefore, the likelihood function can be written as:

$$L(\theta) = \frac{2^n \prod_{i=1}^n x_i}{\theta^{2n}} I\left(\max_{1\leq i\leq n} x_i \leq \theta\right).$$

The likelihood function includes the indicator $I(\max_{1\leq i\leq n} x_i \leq \theta)$, which means θ must satisfy $\theta \geq \max_{1\leq i\leq n} x_i$ for $L(\theta) > 0$.

For $\theta \geq \max_{1\leq i\leq n} x_i$, the likelihood decreases as θ increases because the denominator θ^{2n} grows. To maximize the likelihood, set θ to the smallest value that satisfies the condition $\theta \geq \max_{1\leq i\leq n} x_i$. Thus:

$$\hat{\theta} = \max_{1\leq i\leq n} X_i.$$

The likelihood function is:

$$L(\theta) = \frac{2^n \prod_{i=1}^n x_i}{\theta^{2n}} I\left(\max_{1\leq i\leq n} x_i \leq \theta\right).$$

The MLE of θ is:

$$\hat{\theta} = \max_{1\leq i\leq n} X_i.$$

Problem 5

Problem 7.6(b)-(c), Casella & Berger (Skip part (a).)

Let $X_1, \ldots, X_n$ be a random sample from the pdf

$$f(x|\theta) = \theta x^{-2}, \quad 0 < \theta \leq x < \infty.$$

(b) Find the MLE of θ.

The goal is to find the maximum likelihood estimator (MLE) of θ based on the given pdf:

$$f(x|\theta) = \theta x^{-2}, \quad 0 < \theta \leq x < \infty.$$

The likelihood function for the random sample $X_1, \ldots, X_n$ is:

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta x_i^{-2}.$$

Simplifying:

$$L(\theta) = \theta^n \prod_{i=1}^{n} x_i^{-2}.$$

Since the support depends on θ, the likelihood also includes an indicator function ensuring $\theta \leq x_{(1)}$, where $x_{(1)} = \min(X_1, \ldots, X_n)$. Thus:

$$L(\theta) = \theta^n \prod_{i=1}^{n} x_i^{-2} \cdot I_{[\theta,\infty)}(x_{(1)}).$$

- The term θ^n is increasing in θ, so to maximize $L(\theta)$, we want θ to be as large as possible.
- However, the indicator function $I_{[\theta,\infty)}(x_{(1)})$ ensures $L(\theta) = 0$ for $\theta > x_{(1)}$.

Thus, the maximum likelihood occurs at the largest possible value of θ satisfying $\theta \leq x_{(1)}$.

The MLE of θ is:

$$\hat{\theta} = x_{(1)}.$$

(c) Find the method of moments estimator of θ.

To find the method of moments estimator (MME) of θ, we use the given pdf:

$$f(x|\theta) = \theta x^{-2}, \quad 0 < \theta \leq x < \infty.$$

The first moment (mean) of X is:

$$\mathbb{E}[X] = \int_{\theta}^{\infty} x \cdot f(x|\theta)\, dx = \int_{\theta}^{\infty} x \cdot \theta x^{-2}\, dx = \int_{\theta}^{\infty} \theta x^{-1}\, dx.$$

Simplify the integral:

$$\mathbb{E}[X] = \theta \int_{\theta}^{\infty} x^{-1}\, dx = \theta[\ln x]_{\theta}^{\infty}.$$

Evaluate the bounds of the logarithmic term:

$$\mathbb{E}[X] = \theta(\ln(\infty) - \ln(\theta)).$$

Since $\ln(\infty) \to \infty$, the expected value $\mathbb{E}[X]$ is infinite. This indicates that the first moment does not exist.

Because the first moment does not exist, the method of moments estimator cannot be defined. Thus, the MME for θ does not exist.

Note: This is the Pareto distribution with shape parameter $\alpha = \theta$ and scale parameter $\beta = 1$.