II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Solved Problems on Point Estimators 2021-12-17

Problem 1 *A real random variable X on a probability space $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$, which represents the reaction time at some stimulus, has a uniform distribution on an interval $[0, \theta]$, where $\theta > 0$ is a parameter. An investigator wants to estimate θ on the basis of a simple random sample $X_1, \dots, X_n$ of reaction times. Since θ is the largest possible time in the entire population of reaction times, the investigator considers as a first estimator for the parameter θ the largest sample reaction time, that is the statistic*

$$\hat{\theta}_1 \equiv \check{X}_n \equiv \max(X_1, \dots, X_n).$$

1. *Is $\check{X}_n$ unbiased? In case $\check{X}_n$ is not unbiased, is it possible to derive from $\check{X}_n$ an unbiased estimator of θ?*
2. *As a second estimator, the investigator considers the statistic*

$$\hat{\theta}_2 \equiv \bar{X}_n \equiv \frac{1}{n}\textstyle\sum_{k=1}^{n} X_k.$$

Is $\bar{X}_n$ unbiased? In case $\bar{X}_n$ is not unbiased, is it possible to derive from $\bar{X}_n$ an unbiased estimator of θ?

3. *In the investigator's shoes, what estimator would you prefer among those considered?*
4. *Is $\check{X}_n$ consistent in probability? Is $\check{X}_n$ consistent in mean square?*

Solution.

1. Writing $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ for the distribution function of the statistic $\check{X}_n$, we have

$$\begin{aligned} F_{\check{X}_n}(x) &= \mathbf{P}\left(\check{X}_n \leq x\right) = \mathbf{P}(X_1 \leq x, \dots, X_n \leq x) = \textstyle\prod_{k=1}^{n} \mathbf{P}(X_k \leq x) \\ &= \textstyle\prod_{k=1}^{n} \mathbf{P}(X \leq x) = \mathbf{P}(X \leq x)^n = F_X(x)^n, \end{aligned}$$

for every $x \in \mathbb{R}$, where $F_X : \mathbb{R} \to \mathbb{R}$ is the distribution function of X. On the other hand, since X is uniformly distributed on $[0, \theta]$, we know that X is absolutely continuous with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) \stackrel{\text{def}}{=} \frac{1}{\theta} 1_{[0,\theta]}(x), \quad \forall x \in \mathbb{R}.$$

Hence,

$$\begin{aligned} F_X(x) &= \int_{(-\infty,x]} f_X(u)\, d\mu_L(u) = \int_{(-\infty,x]} \frac{1}{\theta} 1_{[0,\theta]}(u)\, d\mu_L(u) = \frac{1}{\theta} \int_{(-\infty,x]\cap[0,\theta]} d\mu_L(u) \\ &= \begin{cases} \frac{1}{\theta} \int_{\varnothing} d\mu_L(u) = 0, & \text{if } x < 0, \\ \frac{1}{\theta} \int_{[0,x]} d\mu_L(u) = \frac{x}{\theta}, & \text{if } 0 \leq x \leq \theta, \\ \frac{1}{\theta} \int_{[0,\theta]} d\mu_L(u) = 1, & \text{if } \theta < x. \end{cases} \end{aligned}$$

Briefly,

$$F_X(x) = \frac{x}{\theta} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

for every $x \in \mathbb{R}$. It then follows,

$$F_{\check{X}_n}(x) = F_X(x)^n = \frac{x^n}{\theta^n} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

for every $x \in \mathbb{R}$. Now, we have

$$F'_{\check{X}_n}(x) = \begin{cases} 0, & \text{if } x < 0, \\ \frac{n x^{n-1}}{\theta^n}, & \text{if } 0 < x < \theta, \\ 0, & \text{if } \theta < x, \end{cases}$$

but $F_{\check{X}_n}$ is not everywhere differentable. Eventually, is not differentiable at the point $x = \theta$. However, considering the function $f_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ given by

$$f_{\check{X}_n}(x) \stackrel{\text{def}}{=} \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x), \quad \forall x \in \mathbb{R},$$

a straightforward computation shows that

$$F_{\check{X}_n}(x) = \int_{(-\infty,x]} f_{\check{X}_n}(u) \; d\mu_L(u),$$

for every $x \in \mathbb{R}$. This implies that $\check{X}_n$ is absolutely continuous with density $f_{\check{X}_n}$. As a consequence,

$$\begin{aligned}\mathbf{E}\left[\check{X}_n\right] &= \int_{\mathbb{R}} x f_{\check{X}_n}(x) \; d\mu_L(x) = \int_{\mathbb{R}} x \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x) \; d\mu_L(x) = \frac{n}{\theta^n} \int_{(0,\theta)} x^n \; d\mu_L(x) \\ &= \frac{n}{\theta^n} \int_0^\theta x^n \; dx = \frac{n}{\theta^n} \left.\frac{x^{n+1}}{n+1}\right|_0^\theta = \frac{n}{n+1}\theta.\end{aligned}$$

We conclude that $\check{X}_n$ is not a unbiased estimator of θ but $\frac{n+1}{n}\check{X}_n$ is an unbiased estimator of θ.

2. We have

$$\begin{aligned}\mathbf{E}\left[\bar{X}_n\right] = \mathbf{E}[X] &= \int_{\mathbb{R}} x f_X(x) \; d\mu_L(x) = \int_{\mathbb{R}} \frac{x}{\theta} 1_{[0,\theta]}(x) \; d\mu_L(x) \\ &= \frac{1}{\theta} \int_{[0,\theta]} x \; d\mu_L(x) = \frac{1}{\theta} \int_0^\theta x \; dx = \frac{1}{\theta} \left.\frac{x^2}{2}\right|_0^\theta = \frac{\theta}{2}.\end{aligned}$$

Hence, $\bar{X}_n$ is not a unbiased estimator of θ but $2\bar{X}_n$ is an unbiased estimator of θ.

3. From 1. and 2. we know that

$$\mathbf{E}\left[\frac{n+1}{n}\check{X}_n\right] = \theta \qquad \text{and} \qquad \mathbf{E}\left[2\bar{X}_n\right] = \theta.$$

Hence, both $\frac{n+1}{n}\check{X}_n$ and $2\bar{X}_n$ are unbiased estimators of the parameter θ. To choose which is preferable between them, we consider

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] \qquad \text{and} \qquad \mathbf{D}^2\left[2\bar{X}_n\right].$$

We have

$$\begin{aligned}\mathbf{E}\left[\check{X}_n^2\right] &= \int_{\mathbb{R}} x^2 f_{\check{X}_n}(x) \, d\mu_L(x) = \int_{\mathbb{R}} x^2 \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x) \, d\mu_L(x) = \frac{n}{\theta^n} \int_{(0,\theta)} x^{n+1} d\mu_L(x) \\ &= \frac{n}{\theta^n} \int_0^\theta x^{n+1} \; dx = \frac{n}{\theta^n} \left.\frac{x^{n+2}}{n+2}\right|_0^\theta = \frac{n}{n+2}\theta^2.\end{aligned}$$

Therefore,

$$\mathbf{D}^2\left[\check{X}_n\right] = \mathbf{E}\left[\check{X}_n^2\right] - \mathbf{E}\left[\check{X}_n\right]^2 = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 = \frac{n}{(n+1)^2(n+2)}\theta^2.$$

As a consequence,

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \mathbf{D}^2\left[\check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \frac{n}{(n+1)^2(n+2)}\theta^2 = \frac{\theta^2}{n(n+2)}.$$

On the other hand,

$$\mathbf{D}^2\left[2\bar{X}_n\right] = 4\mathbf{D}^2\left[\bar{X}_n\right] = \frac{4}{n}\mathbf{D}^2[X] = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

Now, we clearly have

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] < \mathbf{D}^2\left[2\bar{X}_n\right],$$

for every $n > 1$. It follows that the estimator $\frac{n+1}{n}\check{X}_n$ is preferable to $2\bar{X}_n$.

4. We have

$$\mathbf{P}\left(\left|\frac{n+1}{n}\check{X}_n-\theta\right|\geq\varepsilon\right)=\mathbf{P}\left(\left|\frac{n+1}{n}\check{X}_n-\mathbf{E}\left[\frac{n+1}{n}\check{X}_n\right]\right|\geq\varepsilon\right)\leq\frac{\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right]}{\varepsilon^2}=\frac{\theta^2}{n\left(n+2\right)\varepsilon^2}.$$

It clearly follows that

$$\frac{n+1}{n}\check{X}_n\xrightarrow{\mathbf{P}}\theta.$$

On the other hand, trivially

$$\frac{n}{n+1}\xrightarrow{\mathbf{P}}1.$$

As a consequence, we have

$$\check{X}_n=\frac{n}{n+1}\cdot\frac{n+1}{n}\check{X}_n\xrightarrow{\mathbf{P}}1\cdot\theta=\theta.$$

Hence, both the estimators $\frac{n+1}{n}\check{X}_n$ and $\check{X}_n$ are consistent in probability. In addition, considering that

$$\mathbf{E}\left[\frac{n+1}{n}\check{X}_n\right]=\theta\qquad\text{and}\qquad\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right]=\frac{\theta^2}{n\left(n+2\right)},$$

we obtain

$$\begin{aligned}\mathbf{E}\left[\left(\check{X}_n-\theta\right)^2\right]&=\left(\frac{n}{n+1}\right)^2\mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n-\frac{n+1}{n}\theta\right)^2\right]\\&=\left(\frac{n}{n+1}\right)^2\mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n-\theta-\frac{\theta}{n}\right)^2\right]\\&=\left(\frac{n}{n+1}\right)^2\mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n-\theta\right)^2-2\left(\frac{n+1}{n}\check{X}_n-\theta\right)\frac{\theta}{n}+\frac{\theta^2}{n^2}\right]\\&=\left(\frac{n}{n+1}\right)^2\left(\mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n-\theta\right)^2\right]-\frac{2\theta}{n}\mathbf{E}\left[\frac{n+1}{n}\check{X}_n-\theta\right]+\mathbf{E}\left[\frac{\theta^2}{n^2}\right]\right)\\&=\left(\frac{n}{n+1}\right)^2\left(\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right]+\frac{\theta^2}{n^2}\right)\\&=\left(\frac{n}{n+1}\right)^2\left(\frac{\theta^2}{n\left(n+2\right)}+\frac{\theta^2}{n^2}\right)\\&=\frac{2\theta^2}{\left(n+1\right)\left(n+2\right)}.\end{aligned}$$

It follows that both the estimators $\frac{n+1}{n}\check{X}_n$ and $\check{X}_n$ are consistent in mean square.

Problem 2 *Let X be a binomially distributed random variable on a probability space $(\Omega,\mathcal{E},\mathbf{P})\equiv\Omega$ with known number of trials parameter m and unknown success parameter p. An investigator wants to estimate p on the basis of a simple random sample $X_1,\dots,X_n$ of size n drawn from X.*

1. *Assume the investigator applies the method of moments. What is the estimator $\hat{p}_n^M$?*
2. *Is $\hat{p}_n^M$ biased? Is $\hat{p}_n^M$ consistent?*
3. *Assume the investigator applies the likelihood method. What is the estimator $\hat{p}_n^{ML}$?*
4. *Given that $m=10$ and we observe a realization $4,4,3,5,6$ of a sample $X_1,\dots,X_5$ of size 5 drawn from X what is the estimate of p by the estimators $\hat{p}_n^M$?*
5. *Can you give an estimate of $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$ by means of the estimator $\hat{p}_n^M$ and the information provided at 3?*

Solution.

1. Write $\mu_1' : (0,1) \to \mathbb{R}$ for the first order population raw moment. We have

$$\mu_1'(p) = \mathbf{E}[X] = mp$$

Replacing p with the estimator $\hat{p}_n^M$, and equating the first order population raw moment to the first order sample moment $\bar{X}_n$, we can write

$$m\hat{p}_n^M = \bar{X}_n.$$

It follows

$$\hat{p}_n^M = \frac{1}{m}\bar{X}_n.$$

2. We have

$$\mathbf{E}\left[\hat{p}_n^M\right] = \mathbf{E}\left[\frac{1}{m}\bar{X}_n\right] = \frac{1}{m}\mathbf{E}\left[\bar{X}_n\right] = \frac{1}{m}\mathbf{E}[X] = p.$$

Hence, the estimator $\hat{p}_n^M$ is unbiased. In addition, we have

$$\begin{aligned}
\mathbf{E}\left[\left(\hat{p}_n^M - p\right)^2\right] &= \mathbf{E}\left[\left(\hat{p}_n^M - \mathbf{E}\left[\hat{p}_n^M\right]\right)^2\right] = \mathbf{D}^2\left[\hat{p}_n^M\right] \\
&= \mathbf{D}^2\left[\frac{1}{m}\bar{X}_n\right] = \frac{1}{m^2}\mathbf{D}^2\left[\bar{X}_n\right] = \frac{1}{m^2}\frac{1}{n}\mathbf{D}^2[X] \\
&= \frac{1}{nm^2}mp(1-p) = \frac{p(1-p)}{nm}.
\end{aligned}$$

It follows,

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{p}_n^M - p\right)^2\right] = \lim_{n\to\infty}\frac{p(1-p)}{nm} = 0$$

foe every $p \in (0,1)$. This means that

$$\hat{p}_n^M \xrightarrow{\mathbf{L}^2} p.$$

That is the estimator $\hat{p}_n^M$ is mean square consistent. A fortiori $\hat{p}_n^M$ is probability consistent.

3. The density function $f_X : \mathbb{N}_0 \times (0,1) \to \mathbb{R}_+$ of a binomial random variable with known number of trials parameter m and unknown success parameter p can be written as

$$f_X(x;p) = \frac{m!}{(m-x)!x!}p^x(1-p)^{m-x}\cdot 1_{\{0,1,\ldots,m\}}(x),$$

for every $x \in \mathbb{N}_0$ and $p \in (0,1)$. Let $X_1,\ldots,X_n$ be a simple random sample of size n drawn from X. Then the likelihood function $\mathcal{L}_{X_1,\ldots,X_n} : (0,1)\times\mathbb{N}_0^n \to \mathbb{R}$ of the sample $X_1,\ldots,X_n$ is given by

$$\begin{aligned}
&\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) \\
&= \prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}p^{x_k}(1-p)^{m-x_k}\cdot 1_{\{0,1,\ldots,m\}}(x_k) \\
&= \left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right)p^{\sum_{k=1}^n x_k}(1-p)^{n\cdot m-\sum_{k=1}^n x_k}1_{\{0,1,\ldots,m\}^n}(x_1,\ldots,x_n)
\end{aligned}$$

for every $p \in (0,1)$ and every realization $(x_1,\ldots,x_n) \in \mathbb{N}_0^n$ of the sample $X_1,\ldots,X_n$. Note that

$$\begin{aligned}
&\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) \\
&= \begin{cases} \left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right)p^{\sum_{k=1}^n x_k}(1-p)^{n\cdot m-\sum_{k=1}^n x_k} > 0, & \text{if } (x_1,\ldots,x_n) \in \{0,1,\ldots,m\}^n, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Therefore,

$$\underset{p\in\mathbb{R}_{++}}{\arg\max}\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) = \underset{p\in\mathbb{R}_{++}}{\arg\max}\left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right)p^{\sum_{k=1}^n x_k}(1-p)^{n\cdot m-\sum_{k=1}^n x_k}$$

Hence, we can consider as the log-likelihood function of the sample $X_1, \ldots, X_n$ the function $\log \mathcal{L}_{X_1,\ldots,X_n} : (0,1) \times \{0,1,\ldots,m\}^n \to \mathbb{R}$ given by

$$\log \mathcal{L}_{X_1,\ldots,X_n}(p; x_1, \ldots, x_n) \overset{\text{def}}{=} \ln\left(\left(\prod_{k=1}^{n} \frac{m!}{(m-x_k)!x_k!}\right) p^{\sum_{k=1}^{n} x_k} (1-p)^{n\cdot m - \sum_{k=1}^{n} x_k}\right),$$
$$\forall\, (p; x_1, \ldots, x_n) \in (0,1) \times \{0,1,\ldots,m\}^n.$$

That is

$$\log \mathcal{L}_{X_1,\ldots,X_n}(p; x_1, \ldots, x_n) = \sum_{k=1}^{n} \ln\left(\frac{m!}{(m-x_k)!x_k!}\right) + \left(\sum_{k=1}^{n} x_k\right) \ln(p) + \left(n\cdot m - \sum_{k=1}^{n} x_k\right) \ln(1-p).$$

To determine $\arg\max_{p\in(0,1)} \log \mathcal{L}_{X_1,\ldots,X_n}(p; x_1, \ldots, x_n)$, we consider the first order condition

$$\frac{d}{dp} \log \mathcal{L}_{X_1,\ldots,X_n}(p; x_1, \ldots, x_n) = 0,$$

which yields

$$\left(\sum_{k=1}^{n} x_k\right) \frac{1}{p} - \left(n\cdot m - \sum_{k=1}^{n} x_k\right) \frac{1}{1-p} = 0.$$

On account that $p \in (0,1)$, the latter becomes

$$\left(\sum_{k=1}^{n} x_k\right)(1-p) - \left(n\cdot m - \sum_{k=1}^{n} x_k\right) p = 0.$$

That is

$$\sum_{k=1}^{n} x_k - n\cdot m\cdot p = 0,$$

which implies

$$p = \frac{\sum_{k=1}^{n} x_k}{n\cdot m} = \frac{\bar{x}_n}{m}.$$

In addition,

$$\begin{aligned}
\frac{d^2}{dp^2} \log \mathcal{L}_{X_1,\ldots,X_n}(p; x_1, \ldots, x_n) &= -\left(\sum_{k=1}^{n} x_k\right) \frac{1}{p^2} - \left(n\cdot m - \sum_{k=1}^{n} x_k\right) \frac{1}{(1-p)^2} \\
&= \frac{-\left(\sum_{k=1}^{n} x_k\right)(1-p)^2 - \left(n\cdot m - \sum_{k=1}^{n} x_k\right) p^2}{p^2 (1-p)^2} \\
&= \frac{-\sum_{k=1}^{n} x_k + 2\left(\sum_{k=1}^{n} x_k\right) p - n\cdot m\cdot p^2}{p^2 (1-p)^2} \\
&= \frac{-n\bar{x}_n + 2n\bar{x}_n p - n\cdot m\cdot p^2}{p^2 (1-p)^2} \\
&= -\frac{n}{p^2 (1-p)^2} \left(\bar{x}_n - 2\bar{x}_n p + m\cdot p^2\right).
\end{aligned}$$

Now, we have

$$\left(\bar{x}_n - 2\bar{x}_n p + m\cdot p^2\right)_{p=\frac{\bar{x}_n}{m}} = \left(\bar{x}_n - \frac{2}{m}\bar{x}_n^2 + m\cdot \frac{\bar{x}_n^2}{m^2}\right) = \bar{x}_n \left(1 - \frac{1}{m}\bar{x}_n\right).$$

On the other hand, we clearly have

$$\bar{x}_n \le m,$$

for every $(x_1, \ldots, x_n) \in \{0,1,\ldots,m\}^n$. It follows

$$\frac{d^2}{dp^2} \log \mathcal{L}_{X_1,\ldots,X_n}(p; x_1, \ldots, x_n) \le 0$$

which implies that

$$\frac{\bar{x}_n}{m} = \arg\max_{p\in(0,1)} \log \mathcal{L}_{X_1,\ldots,X_n}(p; x_1, \ldots, x_n).$$

As a consequence, we obtain that the maximum likelihood estimator fo p is given by

$$\hat{p}_n^{ML} = \frac{\bar{X}_n}{m}.$$

4. Given that $m = 10$ and we observe a realization $4, 4, 3, 5, 6$ of a sample $X_1, \ldots, X_5$ of size 5 drawn from X, we obtain

$$\hat{p}_n^M(\omega) = \frac{\bar{X}_5(\omega)}{10} = \frac{\frac{1}{5}(4+4+3+5+6)}{10} = 0.44.$$

5. We know that

$$\mathbf{E}[X] = m \cdot p \qquad \text{and} \qquad \mathbf{D}^2[X] = m \cdot p(1-p),$$

where p is the true value of the success parameter, Hence, an estimator $\hat{\mu}_n^M$ [resp. $\hat{\sigma}_n^{2^M}$] of the expectation [resp. variance] of X build from $\hat{p}_n^M$ is given by

$$\hat{\mu}_n^M = m \cdot \hat{p}_n^M \quad \text{and} \quad \hat{\sigma}_n^{2^M} = m \cdot \hat{p}_n^M\left(1 - \hat{p}_n^M\right).$$

An estimate of $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$ by means of the estimator $\hat{p}_n^M$ and the information provided at 3. is the given by

$$\hat{\mu}_X^M(\omega) = m \cdot \hat{p}_n^M(\omega) = 10 \cdot 0.44 = 4.4$$

and

$$\hat{\sigma}_n^{2^M}(\omega) = m \cdot \hat{p}_n^M(\omega)\left(1 - \hat{p}_n^M(\omega)\right) = 10 \cdot 0.44 \cdot (1 - 0.44) = 2.464.$$

This completes the solution.

Problem 3 *Let X be a normally distributed random variable on a probability space $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ with unknown mean μ_X and variance σ_X^2 . An investigator wants to estimate μ and σ^2 on the basis of a simple random sample $X_1, \ldots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the likelihood methods. What are the estimator $\hat{\mu}_n^{LM}$ and $\hat{\sigma}_n^{2^{LM}}$?*
2. *Assume the investigator applies the method of moments. What are the estimators $\hat{\mu}_n^M$ and $\hat{\sigma}_n^{2^M}$? Hint: guess what $\hat{\sigma}_n^{2^M}$ could be and get it!*
3. *Are the estimators $\hat{\mu}_n^{LM}$ and $\hat{\sigma}_n^{2^{LM}}$ unbiased? Are the estimators $\hat{\mu}_n^{LM}$ and $\hat{\sigma}_n^{2^{LM}}$ consistent in probability? Are they consistent in mean square?*

Solution.

1. We know that the joint density function $f_{X_1,\ldots,X_n} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$ of the sample $X_1, \ldots, X_n$ is given by

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n; \mu, \sigma) \stackrel{\text{def}}{=} \prod_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_k-\mu)^2}{2\sigma^2}}, \quad \forall (x_1, \ldots, x_n; \mu, \sigma) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_{++}.$$

That is

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n; \mu, \sigma) = \frac{1}{\sqrt{2^n \pi^n} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n (x_k - \mu)^2}.$$

Hence, the likelihood function $\mathcal{L}_{X_1,\ldots,X_n} : \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}$ of the sample is given by

$$\mathcal{L}_{X_1,\ldots,X_n}(\mu, \sigma; x_1, \ldots, x_n) = \frac{1}{\sqrt{2^n \pi^n} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n (x_k - \mu)^2}, \quad \forall (\mu, \sigma; x_1, \ldots, x_n) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}^n.$$

Thanks to the stucture of $\mathcal{L}_{X_1,\ldots,X_n}$ it is convenient to consider the log-likelihood function $\log \mathcal{L}_{X_1,\ldots,X_n} : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ of the sample which is given by

$$\log \mathcal{L}_{X_1,\ldots,X_n}(\mu, \sigma; x_1, \ldots, x_n) \stackrel{\text{def}}{=} (\log \circ \mathcal{L}_{X_1,\ldots,X_n})(\mu, \sigma; x_1, \ldots, x_n), \quad \forall (\mu, \sigma; x_1, \ldots, x_n) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}^n.$$

That is

$$\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = -n\left(\frac{1}{2}\ln(2\pi)+\ln(\sigma)\right) - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2.$$

Now, to determine $\underset{(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}_{++}}{\arg\max} \log \mathcal{L}_{X_1,\ldots,X_n}$ we consider the first order conditions

$$\frac{\partial}{\partial\mu}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0 \quad \text{and} \quad \frac{\partial}{\partial\sigma}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0.$$

We have

$$\frac{\partial}{\partial\mu}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = \frac{1}{\sigma^2}\sum_{k=1}^{n}(x_k-\mu)$$

and

$$\frac{\partial}{\partial\sigma}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = \frac{1}{\sigma}\left(\frac{1}{\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2 - n\right).$$

Therefore,

$$\begin{aligned}\frac{\partial}{\partial\mu}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0 &\Rightarrow \sum_{k=1}^{n}(x_k-\mu) = 0,\\ &\Rightarrow \mu = \frac{1}{n}\sum_{k=1}^{n}x_k \equiv \bar{x}_n\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial\sigma}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0 &\Rightarrow \frac{1}{\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2 - n,\\ &\Rightarrow \sigma^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k-\mu)^2 \equiv \tilde{s}^2_{X,n}.\end{aligned}$$

In addition,

$$\begin{aligned}\frac{\partial^2}{\partial\mu^2}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) &= -\frac{n}{\sigma^2},\\ \frac{\partial^2}{\partial\mu\sigma}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) &= -2\frac{1}{\sigma^3}\sum_{k=1}^{n}(x_k-\mu),\\ \frac{\partial^2}{\partial\sigma^2}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) &= \frac{1}{\sigma^2}\left(-\frac{3}{\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2+n\right).\end{aligned}$$

Hence,

$$\begin{aligned}&J\left(\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)}\\ &= \begin{pmatrix} \frac{\partial^2}{\partial\mu^2}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) & \frac{\partial^2}{\partial\mu\sigma}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n)\\ \frac{\partial^2}{\partial\mu\sigma}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) & \frac{\partial^2}{\partial\sigma^2}\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n)\end{pmatrix}_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)}\\ &= \begin{pmatrix} -\frac{n}{\sigma^2} & -2\frac{1}{\sigma^3}\sum_{k=1}^{n}(x_k-\mu)\\ -2\frac{1}{\sigma^3}\sum_{k=1}^{n}(x_k-\mu) & \frac{1}{\sigma^2}\left(-\frac{3}{\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2+n\right)\end{pmatrix}_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)}\\ &= \frac{1}{\tilde{s}^2_{X,n}}\begin{pmatrix} -n & 0\\ 0 & -2n\end{pmatrix}\end{aligned}$$

Because,

$$\sum_{k=1}^{n}(x_k-\mu)\Big|_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)} = \sum_{k=1}^{n}x_k - n\mu\Big|_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)} = 0$$

and

$$-\frac{3}{\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2\Big|_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)} = -\frac{3n}{\sigma^2}\frac{1}{n}\sum_{k=1}^{n}(x_k-\mu)^2\Big|_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)} = -\frac{3n}{\tilde{s}^2_{X,n}}\tilde{s}^2_{X,n} = -3n.$$

We then have

$$\det J\left(\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)} = \frac{3n^2}{\tilde{s}^4_{X,n}}$$

and

$$\operatorname{tr} J\left(\log \mathcal{L}_{X_1,\ldots,X_n}\left(\mu,\sigma;x_1,\ldots,x_n\right)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)} = -\frac{3n}{\tilde{s}^4_{X,n}}$$

It follows that the eigenvalues of the Jacobian matrix $J\left(\log \mathcal{L}_{X_1,\ldots,X_n}\left(\mu,\sigma;x_1,\ldots,x_n\right)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)}$ are stictly negative. This imples that

$$\left(\bar{x}_n,\tilde{s}^2_{X,n}\right) = \underset{(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}_{++}}{\arg\max} \log \mathcal{L}_{X_1,\ldots,X_n}.$$

As a consequence, we obtain the maximum likelihood estimators

$$\hat{\mu}_n^{LM} = \bar{X}_n \quad \text{and} \quad \hat{\sigma}_n^{2^{LM}} = \tilde{S}^2_{X,n},$$

where $\bar{X}_n$ [resp. $\tilde{S}^2_n(X)$] is the sample mean [resp. unbiased sample variance] of $X_1,\ldots,X_n$.

2. We know that

$$\mathbf{E}\left[X\right] = \mu \quad \text{and} \quad \mathbf{E}\left[X^2\right] = \mu^2+\sigma^2.$$

Hence, applying the method of moments, the investigator writes

$$\frac{1}{n}\sum_{k=1}^n X_k = \hat{\mu}_n^M \quad \text{and} \quad \frac{1}{n}\sum_{k=1}^n X_k^2 = \left(\hat{\mu}_n^M\right)^2 + \hat{\sigma}_n^{2^M}.$$

The first of the two equations clearly yields

$$\hat{\mu}_n^M = \bar{X}_n.$$

The second equation, on account of the first, yields

$$\hat{\sigma}_n^{2^M} = \frac{1}{n}\sum_{k=1}^n X_k^2 - \bar{X}_n^2 = \tilde{S}^2_{X,n}.$$

Recall that

$$\begin{aligned}
\tilde{S}^2_{X,n} &= \frac{1}{n}\sum_{k=1}^n \left(X_k-\bar{X}_n\right)^2 = \frac{1}{n}\sum_{k=1}^n \left(X_k^2 - 2X_k\bar{X}_n + \bar{X}_n^2\right) \\
&= \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - 2\bar{X}_n\sum_{k=1}^n X_k + \sum_{k=1}^n \bar{X}_n^2\right) = \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - 2n\bar{X}_n^2 + n\bar{X}_n^2\right) \\
&= \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - n\bar{X}_n^2\right) = \frac{1}{n}\sum_{k=1}^n X_k^2 - \bar{X}_n^2.
\end{aligned}$$

3. It is well known that the estimator $\hat{\mu}_n^{ML} = \bar{X}_n$ [resp. $\hat{\sigma}_n^{2^{ML}} = \tilde{S}^2_{X,n}$] is unbiased [resp. biased]. In addition, since

$$\mathbf{D}^2\left[\hat{\mu}_n^{ML}\right] = \mathbf{D}^2\left[\bar{X}_n\right] = \frac{1}{n}\mathbf{D}^2\left[X\right] = \frac{1}{n}\sigma^2$$

we clearly have

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{\mu}_n^{ML}-\mu\right)^2\right] = \lim_{n\to\infty}\mathbf{D}^2\left[\hat{\mu}_n^{ML}\right] = \lim_{n\to\infty}\frac{1}{n}\sigma^2 = 0$$

which means that $\hat{\mu}_n^{ML}$ is consistent in mean square. With regard to the biased estimator $\hat{\sigma}_n^{2^{ML}}$, observe that we can write

$$\begin{aligned}
\mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}}-\sigma\right)^2\right] &= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \left(\frac{n-1}{n}\sigma^2 + \frac{1}{n}\sigma^2\right)\right)^2\right] \\
&= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2 - \frac{1}{n}\sigma^2\right)^2\right] \\
&= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right)^2 - \frac{1}{n}\sigma^2\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right) + \frac{1}{n}\sigma^4\right] \\
&= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right)^2\right] - \frac{1}{n}\sigma^2\mathbf{E}\left[\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right] + \frac{1}{n}\sigma^2 \\
&= \mathbf{E}\left[\left(\tilde{S}^2_{X,n} - \mathbf{E}\left[\tilde{S}^2_{X,n}\right]\right)^2\right] - \frac{1}{n}\sigma^2\mathbf{E}\left[\tilde{S}^2_{X,n} - \mathbf{E}\left[\tilde{S}^2_{X,n}\right]\right] + \frac{1}{n}\sigma^2 \\
&= \mathbf{D}^2\left[\tilde{S}^2_{X,n}\right] + \frac{1}{n}\sigma^2.
\end{aligned}$$

On the other hand,

$$\mathbf{D}^2\left[\tilde{S}^2_{X,n}\right]=\mathbf{D}^2\left[\frac{n}{n+1}S^2_{X,n}\right]=\frac{n^2}{(n+1)^2}\mathbf{D}^2\left[S^2_{X,n}\right]$$
$$=\frac{n^2}{(n+1)^2}\frac{\sigma^4}{n}\left(3-\frac{n-3}{n-1}\right)=\frac{n^2}{(n+1)^2}\frac{\sigma^4}{n}\frac{2n}{n+1}$$
$$=\frac{2n^2}{(n+1)^3}\sigma^4.$$

Therefore,

$$\mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}}-\sigma\right)^2\right]=\frac{2n^2}{(n+1)^3}\sigma^4+\frac{1}{n}\sigma^2$$

It folows

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}}-\sigma\right)^2\right]=\lim_{n\to\infty}\left(\frac{2n^2}{(n+1)^3}\sigma^4+\frac{1}{n}\sigma^2\right)=0,$$

which means that also $\hat{\sigma}_n^{2^{ML}}$ is consistent in mean square. A fortiori, both $\hat{\mu}_n^{ML}$ and $\hat{\sigma}_n^{2^{ML}}$ are consistent in probability.

Problem 4 *Let $\theta>0$ and let X be an uniformly distributed real random variable on the interval $[0,\theta]$. In symbols $X\sim Unif\,(0,\theta)$.*

1. *Write the joint likelihood function of a simple random sample $X_1,\dots,X_n$ of size n drawn from X and determine $\hat{\theta}_n^{ML}$.*
2. *Check whether the MLE is unbiased or biased.*
3. *Determine $\hat{\theta}_n^{M}$, check that $\hat{\theta}_n^{M}$ is unbiased and consistent.*

Solution.

1. The density function $f_X:\mathbb{R}\times\mathbb{R}_+\to\mathbb{R}$, depending of the parameter θ, is given by

$$f_X(x;\theta)\stackrel{\text{def}}{=}\frac{1}{\theta}1_{[0,\theta]}(x),\quad\forall x\in\mathbb{R},\quad\forall\theta\in\mathbb{R}_+.$$

Therefore, the sample likelihood $\mathcal{L}_{X_1,\dots,X_n}:\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}$ can be written as

$$\mathcal{L}_{X_1,\dots,X_n}(\theta;x_1,\dots,x_n)\stackrel{\text{def}}{=}\prod_{k=1}^n\frac{1}{\theta}1_{[0,\theta]}(x_k),\quad\forall\theta\in\mathbb{R}_+,\quad\forall x_1,\dots,x_n\in\mathbb{R}.$$

That is to say

$$\mathcal{L}_{X_1,\dots,X_n}(\theta;x_1,\dots,x_n)=\frac{1}{\theta^n}\prod_{k=1}^n1_{[0,\theta]}(x_k).$$

Now, to enhance the role of θ as a variable, note that

$$1_{[0,\theta]}(x_k)=1_{\mathbb{R}_+}(x_k)\,1_{[x_n,+\infty)}(\theta)$$

for all $\theta\in\mathbb{R}_+$ and all $x_1,\dots,x_n\in\mathbb{R}$. Hence, it may be convenient to write the joint likelihood in the form

$$\mathcal{L}_{X_1,\dots,X_n}(\theta;x_1,\dots,x_n)=\frac{1}{\theta^n}\prod_{k=1}^n1_{\mathbb{R}_+}(x_k)\,1_{[x_k,+\infty)}(\theta)=\frac{1}{\theta^n}\prod_{k=1}^n1_{[x_k,+\infty)}(\theta)\prod_{k=1}^n1_{\mathbb{R}_+}(x_k).$$

Given any realization $x_1,\dots,x_n$ of the random sample $X_1,\dots,X_n$, it follows that

$$\mathcal{L}_{X_1,\dots,X_n}(\theta;x_1,\dots,x_n)=\begin{cases}0 & \text{if } x_k<0,\quad\exists k\in\{1,\dots,n\}\\ \frac{1}{\theta^n}\prod_{k=1}^n1_{[x_k,+\infty)}(\theta) & \text{if } x_k\geq 0,\quad\forall k\in\{1,\dots,n\}\end{cases}.$$

Therefore, under the condition $x_k \geq 0$ for every $k \in \{1, \ldots, n\}$, we have

$$\underset{\theta \in \mathbb{R}_+}{\arg\max} \mathcal{L}_{X_1,\ldots,X_n}(\theta; x_1, \ldots, x_n) = \underset{\theta \in \mathbb{R}_+}{\arg\max} \frac{1}{\theta^n} \prod_{k=1}^n 1_{[x_k,+\infty)}(\theta) = \max\{x_1, \ldots, x_n\}.$$

In fact,

$$\prod_{k=1}^n 1_{[x_n,+\infty)}(\theta) = \begin{cases} 0 & \text{if } \theta < x_k, \quad \exists k \in \{1, \ldots, n\} \\ 1 & \text{if } \theta \geq x_k, \quad \forall k \in \{1, \ldots, n\} \end{cases}.$$

Hence, $\frac{1}{\theta^n} \prod_{k=1}^n 1_{[x_n,+\infty)}(\theta)$ attains its maximum for $\theta \geq x_k$, for every $k = 1, \ldots, n$. That is

$$\max\{x_1, \ldots, x_n\} \leq \underset{\theta \in \mathbb{R}_+}{\arg\max} \frac{1}{\theta^n} \prod_{k=1}^n 1_{[x_k,+\infty)}(\theta).$$

In addition,

$$\frac{1}{\theta^n} \prod_{k=1}^n 1_{[x_k,+\infty)}(\theta) = \begin{cases} \frac{1}{\max\{x_1,\ldots,x_n\}^n} & \text{if } \theta = \max\{x_1, \ldots, x_n\} \\ \frac{1}{\theta^n} < \frac{1}{\max\{x_1,\ldots,x_n\}^n} & \text{if } \theta > \max\{x_1, \ldots, x_n\} \end{cases}.$$

In the end, since $\mathbf{P}(X < 0) = 0$ implies $\mathbf{P}(X_k \geq 0) = 1$ for every $k \in \{1, \ldots, n\}$, we obtain

$$\hat{\theta}_n^{ML} = \max\{X_1, \ldots, X_n\}.$$

2. To check whether $\hat{\theta}_{MLE}$ is unbiased or biased we need check whether

$$\mathbf{E}\left[\hat{\theta}_n^{ML}\right] = \theta$$

or not. Write $\check{X}_n \equiv \max\{X_1, \ldots, X_n\} \equiv \hat{\theta}_n^{ML}$. We will be able to compute $\mathbf{E}\left[\check{X}_n\right]$ if we determine the distribution function $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}_+$ of $\check{X}_n$. We have

$$F_{\check{X}_n}(x) = \mathbf{P}\left(\check{X}_n \leq x\right) = \mathbf{P}(X_1 \leq x, \ldots, X_n \leq x) = \prod_{k=1}^n \mathbf{P}(X_k \leq x) = \mathbf{P}(X \leq x)^n,$$

for every $x \in \mathbb{R}$. On the other hand,

$$\mathbf{P}(X \leq x) = F_X(x) = \frac{x}{\theta} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

Therefore,

$$F_{\check{X}_n}(x) = \frac{x^n}{\theta^n} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

It follows that $\check{X}_n$ is absolutely continuous with density $f_{\check{X}_n} : \mathbb{R} \to \mathbb{R}_+$ given by

$$f_{\check{X}_n}(x) = n \frac{x^{n-1}}{\theta^n} 1_{[0,\theta]}(x),$$

for every $x \in \mathbb{R}$. As a consequence, we can write

$$\begin{aligned} \mathbf{E}\left[\check{X}_n\right] &= \int_{\mathbb{R}} x f_{\check{X}_n}(x)\, d\mu_L(x) = \int_{\mathbb{R}} n \frac{x^n}{\theta^n} 1_{[0,\theta]}(x)\, d\mu_L(x) \\ &= \frac{n}{\theta^n} \int_{[0,\theta]} x^n\, d\mu_L(x) = \frac{n}{\theta^n} \int_0^\theta x^n\, dx = \frac{n}{\theta^n} \frac{1}{n+1}\, x^{n+1}\Big|_0^\theta \\ &= \frac{n}{n+1}\theta. \end{aligned}$$

This proves that $\hat{\theta}_n^{ML}$ is a biased estimator of θ.

3. The first population moment and the first sample moment are given by

$$\mathbf{E}[X] = \frac{\theta}{2} \quad \text{and} \quad \bar{X}_n \equiv \frac{1}{n} \sum_{k=1}^n X_k,$$

respectively. Equating

$$\frac{\theta}{2} = \bar{X}_n,$$

it follows that

$$\hat{\theta}_n^M = 2\bar{X}_n.$$

A straightforward computation yields

$$\mathbf{E}\left[\hat{\theta}_n^M\right] = \mathbf{E}\left[2\bar{X}_n\right] = 2\mathbf{E}\left[\bar{X}_n\right] = 2\mathbf{E}\left[\frac{1}{n}\sum_{k=1}^n X_k\right] = \frac{2}{n}\sum_{k=1}^n \mathbf{E}\left[X_k\right] = \frac{2}{n}\sum_{k=1}^n \mathbf{E}\left[X\right] = 2\mathbf{E}\left[X\right] = \theta,$$

which shows that $\hat{\theta}_n^M$ is unbiased. In addition, from Remark **??** we know that

$$\bar{X}_n \xrightarrow{\mathbf{P}} \mathbf{E}\left[X\right].$$

This implies (see Theorem **??**)

$$2\bar{X}_n \xrightarrow{\mathbf{P}} 2\mathbf{E}\left[X\right] = \theta,$$

which shows that $\hat{\theta}_n^M$ is consistent.

Problem 5 *Let $\theta > 0$ and let X be an absolutely continuous real random variable with density function $f_X : \mathbb{R} \to \mathbb{R}_+$ given by*

$$f_X(x) \overset{def}{=} \theta x^{\theta-1} 1_{[0,1]}(x), \quad \forall x \in \mathbb{R}.$$

1. *Apply the method of moments to determine the estimator $\hat{\theta}_n^M$ for θ.*
2. *Check whether $\hat{\theta}_n^M$ is unbiased, consistent in probability, and consistent in mean square.*
3. *Apply the method of maximum likelihood to determine the estimator $\hat{\theta}_n^{ML}$ for θ.*
4. *Check whether $\hat{\theta}_n^{ML}$ is unbiased, consistent in probability, and consistent in mean square.*
5. *Use the estimators obtained to build estimators of the mean μ_X and the variance σ_X^2 of the random variable X.*

Solution.

1. Note that

$$\begin{aligned}
\mathbf{P}(0 < X < 1) &= \int_{(0,1)} f_X(x;\theta)\, d\mu_L(x) = \int_{(0,1)} \theta x^{\theta-1} 1_{[0,1]}(x)\, d\mu_L(x) \\
&= \int_{(0,1)} \theta x^{\theta-1} d\mu_L(x) = \int_0^1 \theta x^{\theta-1} dx = \theta \int_0^1 x^{\theta-1} dx \\
&= \theta \left.\frac{x^\theta}{\theta}\right|_0^1 \\
&= 1.
\end{aligned}$$

Hence, considering a simple random sample $X_1, \ldots, X_n$ of size n drawn from X, we have

$$\mathbf{P}(0 < X_k < 1) = 1$$

for every $k = 1, \ldots, n$. Moreover, we clearly have

$$\bigcap_{k=1}^n \{0 < X_k < 1\} \subseteq \left\{0 < \bar{X}_n < 1\right\},$$

which implies

$$\mathbf{P}\left(\bigcap_{k=1}^n \{0 < X_k < 1\}\right) \leq \mathbf{P}\left(0 < \bar{X}_n < 1\right).$$

On the other hand,

$$\mathbf{P}\left(\bigcap_{k=1}^{n}\{0<X_k<1\}\right)=\prod_{k=1}^{n}\mathbf{P}\left(0<X_k<1\right)=1$$

It follows

$$\mathbf{P}\left(0<\bar{X}_n<1\right)=1.$$

Now, we have

$$\begin{aligned}\mathbf{E}\left[X\right]&=\int_{\mathbb{R}}xf_X\left(x;\theta\right)d\mu_L\left(x\right)=\int_{\mathbb{R}}\theta x^{\theta}1_{[0,1]}\left(x\right)d\mu_L\left(x\right)\\&=\int_{[0,1]}\theta x^{\theta}d\mu_L\left(x\right)=\theta\int_0^1x^{\theta}dx=\left.\frac{\theta}{\theta+1}x^{\theta+1}\right|_0^1\\&=\frac{\theta}{1+\theta}\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}\left[X^2\right]&=\int_{\mathbb{R}}x^2f_X\left(x;\theta\right)d\mu_L\left(x\right)=\int_{\mathbb{R}}\theta x^{\theta+1}1_{[0,1]}\left(x\right)d\mu_L\left(x\right)d\\&=\int_{[0,1]}\theta x^{\theta+1}d\mu_L\left(x\right)=\theta\int_0^1x^{\theta+1}dx=\left.\frac{\theta}{\theta+2}x^{\theta+2}\right|_0^1\\&=\frac{\theta}{2+\theta}\end{aligned}$$

It follows,

$$\mathbf{D}^2\left[X\right]=\mathbf{E}\left[X^2\right]-\mathbf{E}\left[X\right]^2=\frac{\theta}{2+\theta}-\frac{\theta^2}{\left(1+\theta\right)^2}=\frac{\theta}{\left(1+\theta\right)^2\left(2+\theta\right)}$$

As a consequence,

$$\mathbf{E}\left[\bar{X}_n\right]=\mathbf{E}\left[X\right]=\frac{\theta}{1+\theta}\quad\text{and}\quad\mathbf{D}^2\left[\bar{X}_n\right]=\frac{1}{n}\mathbf{D}^2\left[X\right]=\frac{\theta}{n\left(1+\theta\right)^2\left(2+\theta\right)}.$$

The estimator $\hat{\theta}_n^M$ for θ is then obtained by solving the equation

$$\frac{\hat{\theta}_n^M}{1+\hat{\theta}_n^M}=\bar{X}_n,$$

which yields

$$\hat{\theta}_n^M=\frac{\bar{X}_n}{1-\bar{X}_n}.$$

Note that, since $\mathbf{P}\left(0<\bar{X}_n<1\right)=1$, the estimator $\hat{\theta}_n^M$ is well defined.

2. To check the properties of the estimator $\hat{\theta}_n^M$, we apply the so called delta method. Considering the function $f:(0,1)\to\mathbb{R}$ given by

$$f\left(x\right)\stackrel{\text{def}}{=}\frac{x}{1-x},\quad\forall x\in(0,1)$$

by virtue of the Taylor formula, fixed any $x_0\in(0,1)$, we can write

$$f\left(x\right)\approx f\left(x_0\right)+f'\left(x_0\right)\left(x-x_0\right),$$

for any $x\in(0,1)$, where

$$f'\left(x_0\right)=\frac{1}{\left(1-x_0\right)^2}.$$

On the other hand, we have

$$\hat{\theta}_n^M=\frac{\bar{X}_n}{1-\bar{X}_n}\equiv f\left(\bar{X}_n\right).$$

Hence, setting

$$\mu_{\bar{X}_n} \equiv \mathbf{E}\left[\bar{X}_n\right] = \mathbf{E}\left[X\right] = \frac{\theta}{1+\theta},$$

the Taylor formula yields

$$f\left(\bar{X}_n\right) \approx f\left(\mu_{\bar{X}_n}\right) + f'\left(\mu_{\bar{X}_n}\right)\left(\bar{X}_n - \mu_{\bar{X}_n}\right),$$

where

$$f\left(\mu_{\bar{X}_n}\right) = \frac{\mu_{\bar{X}_n}}{1-\mu_{\bar{X}_n}} = \frac{\frac{\theta}{1+\theta}}{1-\frac{\theta}{1+\theta}} = \theta$$

and

$$f'\left(\mu_{\bar{X}_n}\right) = \frac{1}{\left(1-\mu_{\bar{X}_n}\right)^2} = \frac{1}{\left(1-\frac{\theta}{1+\theta}\right)^2} = \left(1+\theta\right)^2.$$

We then obtain

$$\hat{\theta}_n^M \approx \theta + \left(1+\theta\right)^2\left(\bar{X}_n - \frac{\theta}{1+\theta}\right).$$

It follows

$$\mathbf{E}\left[\hat{\theta}_n^M\right] \approx \theta + \left(1+\theta\right)^2\left(\mathbf{E}\left[\bar{X}_n\right] - \frac{\theta}{1+\theta}\right) = \theta$$

and

$$\mathbf{D}^2\left[\hat{\theta}_n^M\right] \approx \mathbf{D}^2\left[f\left(\mu_{\bar{X}_n}\right) + f'\left(\mu_{\bar{X}_n}\right)\left(\bar{X}_n - \mu_{\bar{X}_n}\right)\right] = f'\left(\mu_{\bar{X}_n}\right)^2\mathbf{D}^2\left[\bar{X}_n\right] = \frac{\theta\left(\theta+1\right)^2}{n\left(\theta+2\right)}.$$

As a consequence, we have that the estimator $\hat{\theta}_n^M$ is approximatively unbiased. In addition, since

$$\bar{X}_n \xrightarrow{\mathbf{P}} \mathbf{E}\left[\bar{X}_n\right] = \frac{\theta}{1+\theta}, \quad \text{and} \quad \mathbf{P}\left(0 < \bar{X}_n < 1\right) = 1$$

we have that

$$1 - \bar{X}_n \xrightarrow{\mathbf{P}} \frac{1}{1+\theta} \quad \text{and} \quad \frac{\bar{X}_n}{1-\bar{X}_n} \xrightarrow{\mathbf{P}} \frac{\frac{\theta}{1+\theta}}{\frac{1}{1+\theta}} = \theta.$$

Thus, the estimator $\hat{\theta}_n^M$ is consistent in probability. We also have

$$\mathbf{E}\left[\left(\hat{\theta}_n^M - \theta\right)^2\right] \approx \mathbf{E}\left[\left(\hat{\theta}_n^M - \mathbf{E}\left[\hat{\theta}_n^M\right]\right)^2\right] = \mathbf{D}^2\left[\hat{\theta}_n^M\right] \approx \frac{\theta\left(\theta+1\right)^2}{n\left(\theta+2\right)}.$$

It follows

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{\theta}_n^M - \theta\right)^2\right] = 0,$$

that is the estimator $\hat{\theta}_n^M$ is consistent in mean square.

3. To determine the estimator $\hat{\theta}_n^{ML}$ we start by building the likelihood function of the sample $X_1, \ldots, X_n$. Writing $f_{X_1,\ldots,X_1} : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ for the density function of the sample $X_1, \ldots, X_n$, we clearly have

$$\begin{aligned} f_{X_1,\ldots,X_1}\left(x_1,\ldots,x_n;\theta\right) &= \prod_{k=1}^{n}\theta x_k^{\theta-1}1_{[0,1]}\left(x_k\right) = \theta^n\prod_{k=1}^{n}x_k^{\theta-1}\prod_{k=1}^{n}1_{[0,1]}\left(x_k\right) \\ &= \theta^n\left(\prod_{k=1}^{n}x_k^{\theta-1}\right)1_{[0,1]\times\cdots\times[0,1]}\left(x_1,\ldots,x_n\right), \end{aligned}$$

for every $(x_1,\ldots,x_n;\theta) \in \mathbb{R}^n \times \mathbb{R}_+$. Hence, the likelihood function $\mathcal{L}_{X_1,\ldots,X_n} : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is given by

$$\mathcal{L}_{X_1,\ldots,X_n}\left(\theta, x_1,\ldots,x_n\right) = \theta^n\left(\prod_{k=1}^{n}x_k^{\theta-1}\right)1_{[0,1]\times\cdots\times[0,1]}\left(x_1,\ldots,x_n\right),$$

for every $(\theta; x_1,\ldots,x_n) \in \mathbb{R}^n \times \mathbb{R}_+$. Observing that we have

$$\mathcal{L}_{X_1,\ldots,X_n}\left(\theta, x_1,\ldots,x_n\right) > 0 \qquad \text{[resp. } \mathcal{L}_{X_1,\ldots,X_n}\left(\theta, x_1,\ldots,x_n\right) = 0\text{]}$$

for every $(x_1,\ldots,x_n) \in (0,1] \times \cdots \times (0,1]$ [resp. $(x_1,\ldots,x_n) \notin (0,1] \times \cdots \times (0,1]$] we have

$$\underset{\theta\in\mathbb{R}_+}{\arg\max}\mathcal{L}_{X_1,\ldots,X_n}(\theta, x_1,\ldots,x_n) = \underset{\theta\in\mathbb{R}_+}{\arg\max}\theta^n \left(\prod_{k=1}^n x_k^{\theta-1}\right).$$

Therefore, we rather consider

$$\underset{\theta\in\mathbb{R}_+}{\arg\max}\ln\left(\theta^n\left(\prod_{k=1}^n x_k^{\theta-1}\right)\right) = \underset{\theta\in\mathbb{R}_+}{\arg\max} n\ln(\theta) + (\theta-1)\sum_{k=1}^n \log(x_k)$$

where $(x_1,\ldots,x_n) \in (0,1] \times \cdots \times (0,1]$. Applying the first order condition to the function to be maximized, we obtain

$$\frac{n}{\theta} + \sum_{k=1}^n \log(x_k) = 0$$

which yields

$$\theta = -\frac{n}{\sum_{k=1}^n \log(x_k)}.$$

In addition the second order derivative of the function to be maximized is

$$-\frac{n}{\theta^2} < 0.$$

As a consequence, we can write

$$\underset{\theta\in\mathbb{R}_+}{\arg\max}\mathcal{L}_{X_1,\ldots,X_n}(\theta, x_1,\ldots,x_n) = -\frac{n}{\sum_{k=1}^n \log(x_k)}.$$

It follows

$$\hat{\theta}_n^{ML} = -\frac{n}{\sum_{k=1}^n \log(X_k)}.$$

4. We can also write

$$\hat{\theta}_n^{ML} = -\frac{n}{\log\left(\prod_{k=1}^n X_k\right)}$$

Therefore, considering the function $f:(0,1)\to\mathbb{R}$ given by

$$f(x) \stackrel{\text{def}}{=} -\frac{n}{\log(x)}, \quad \forall x \in (0,1),$$

by virtue of the Taylor formula, fixed any $x_0 \in (0,1)$, we can write

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0),$$

for any $x \in (0,1)$, where

$$f'(x_0) = -\frac{n}{x_0 \log^2(x_0)}.$$

On the other hand, we have

$$\hat{\theta}_n^{ML} = -\frac{n}{\log\left(\prod_{k=1}^n X_k\right)} \equiv f\left(\prod_{k=1}^n X_k\right)$$

Hence, setting

$$\mu_{\prod_{k=1}^n X_k} \equiv \mathbf{E}\left[\prod_{k=1}^n X_k\right] = \prod_{k=1}^n \mathbf{E}[X_k] = \prod_{k=1}^n \mathbf{E}[X] = \frac{\theta^n}{(1+\theta)^n},$$

the Taylor formula yields

$$f\left(\prod_{k=1}^n X_k\right) \approx f\left(\mu_{\prod_{k=1}^n X_k}\right) + f'\left(\mu_{\prod_{k=1}^n X_k}\right)\left(\prod_{k=1}^n X_k - \mu_{\prod_{k=1}^n X_k}\right),$$

where

$$f\left(\mu_{\Pi_{k=1}^{n}X_k}\right) = -\frac{n}{\log\left(\frac{\theta^n}{(1+\theta)^n}\right)} = -\frac{n}{n\left(\log\left(\theta\right) - \log\left(1+\theta\right)\right)} = \frac{1}{\log\left(1+\theta\right) - \log\left(\theta\right)}$$

and

$$f'\left(\mu_{\Pi_{k=1}^{n}X_k}\right) = -\frac{n}{\frac{\theta^n}{(1+\theta)^n}\log^2\left(\frac{\theta^n}{(1+\theta)^n}\right)} = -\frac{n\left(1+\theta\right)^n}{\theta^n n^2\left(\log\left(1+\theta\right) - \log\left(\theta\right)\right)^2} = -\frac{\left(1+\theta\right)^n}{\theta^n n\left(\log\left(1+\theta\right) - \log\left(\theta\right)\right)^2}.$$

It follows

$$\begin{aligned}
\mathbf{E}\left[\hat{\theta}_n^{ML}\right] &\approx f\left(\mu_{\Pi_{k=1}^{n}X_k}\right) + f'\left(\mu_{\Pi_{k=1}^{n}X_k}\right)\left(\mathbf{E}\left[\prod_{k=1}^{n}X_k\right] - \mu_{\Pi_{k=1}^{n}X_k}\right) \\
&= \frac{1}{\log\left(1+\theta\right) - \log\left(\theta\right)} = \frac{1}{\log\left(\frac{1+\theta}{\theta}\right)} = \frac{1}{\log\left(1+\frac{1}{\theta}\right)} \approx \frac{1}{\frac{1}{\theta}} \\
&= \theta,
\end{aligned}$$

for large θ. The stimator $\hat{\theta}_n^{ML}$ is approximatively unbiased for large θ

$$\begin{aligned}
\mathbf{D}^2\left[\hat{\theta}_n^{ML}\right] &\approx f'\left(\mu_{\Pi_{k=1}^{n}X_k}\right)^2 \mathbf{D}^2\left[\prod_{k=1}^{n}X_k\right] \\
&= \frac{\left(1+\theta\right)^{2n}}{\theta^{2n} n^2\left(\log\left(1+\theta\right) - \log\left(\theta\right)\right)^4}\theta^n\frac{\left(\theta+1\right)^{4n} - \theta^{3n}\left(\theta+2\right)^n}{\left(\theta+2\right)^n\left(\theta+1\right)^{4n}} \\
&= \frac{1}{n^2}\frac{\left(1+\theta\right)^{4n} - \theta^{3n}\left(2+\theta\right)^n}{\theta^n\left(1+\theta\right)^{2n}\left(2+\theta\right)^n\left(\log\left(1+\theta\right) - \log\left(\theta\right)\right)^2}
\end{aligned}$$

In fact,

$$\begin{aligned}
\mathbf{D}^2\left[\prod_{k=1}^{n}X_k\right] &= \mathbf{E}\left[\left(\prod_{k=1}^{n}X_k\right)^2\right] - \mathbf{E}\left[\prod_{k=1}^{n}X_k\right]^2 \\
&= \mathbf{E}\left[X_k^2\right] - \left(\prod_{k=1}^{n}\mathbf{E}\left[X_k\right]\right)^2 \\
&= \prod_{k=1}^{n}\mathbf{E}\left[X_k^2\right] - \left(\prod_{k=1}^{n}\mathbf{E}\left[X_k\right]\right)^2 \\
&= \prod_{k=1}^{n}\mathbf{E}\left[X^2\right] - \left(\prod_{k=1}^{n}\mathbf{E}\left[X\right]\right)^2 \\
&= \mathbf{E}\left[X^2\right]^n - \mathbf{E}\left[X\right]^{2n} \\
&= \frac{\theta^n}{\left(2+\theta\right)^n} - \frac{\theta^{4n}}{\left(1+\theta\right)^{4n}} \\
&= \theta^n\frac{\left(1+\theta\right)^{4n} - \theta^{3n}\left(2+\theta\right)^n}{\left(2+\theta\right)^n\left(1+\theta\right)^{4n}}
\end{aligned}$$

As a consequence, for large θ

$$\begin{aligned}
\mathbf{E}\left[\left(\hat{\theta}_n^{ML} - \theta\right)^2\right] &\approx \mathbf{E}\left[\left(\hat{\theta}_n^{ML} - \mathbf{E}\left[\hat{\theta}_n^{ML}\right]\right)^2\right] = \mathbf{D}^2\left[\hat{\theta}_n^{M}\right] \\
&\approx \frac{1}{n^2}\frac{\left(1+\theta\right)^{4n} - \theta^{3n}\left(2+\theta\right)^n}{\theta^n\left(1+\theta\right)^{2n}\left(2+\theta\right)^n\left(\log\left(1+\theta\right) - \log\left(\theta\right)\right)^2}.
\end{aligned}$$

It follows

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{\theta}_n^{ML} - \theta\right)^2\right] = 0,$$

that is the estimator $\hat{\theta}_n^M$ is consistent in mean square. In particular, the estimator $\hat{\theta}_n^M$ is consistent in probability.

5. We have

$$\mathbf{E}[X] = \frac{\theta}{1+\theta} \quad \text{and} \quad \mathbf{D}^2[X] = \frac{\theta}{(1+\theta)^2(2+\theta)}.$$

Hence, writing $\hat{\mu}_{X,n}^M$ [resp. $\hat{\mu}_{X,n}^{ML}$] for an estimator of size n of $\mathbf{E}[X]$ built from $\hat{\theta}_n^M$ [resp. $\hat{\theta}_n^{ML}$], we have

$$\hat{\mu}_{X,n}^M = \frac{\hat{\theta}_n^M}{1+\hat{\theta}_n^M} = \frac{\frac{\bar{X}_n}{1-\bar{X}_n}}{1+\frac{\bar{X}_n}{1-\bar{X}_n}} = \bar{X}_n$$

and

$$\hat{\mu}_{X,n}^{ML} = \frac{\hat{\theta}_n^{ML}}{1+\hat{\theta}_n^{ML}} = \frac{-\frac{n}{\sum_{k=1}^n \log(X_k)}}{1-\frac{n}{\sum_{k=1}^n \log(X_k)}} = \frac{n}{n-\sum_{k=1}^n \log(X_k)}.$$

Similarly, writing $\hat{\sigma}_{X,n}^{2,M}$ [resp. $\hat{\sigma}_{X,n}^{2,ML}$] for an estimator of size n of $\mathbf{D}^2[X]$ built from $\hat{\theta}_n^M$ [resp. $\hat{\theta}_n^{ML}$], we have

$$\hat{\sigma}_{X,n}^{2,M} = \frac{\hat{\theta}_n^M}{\left(1+\hat{\theta}_n^M\right)^2\left(2+\hat{\theta}_n^M\right)} = \frac{\frac{\bar{X}_n}{1-\bar{X}_n}}{\left(1+\frac{\bar{X}_n}{1-\bar{X}_n}\right)^2\left(2+\frac{\bar{X}_n}{1-\bar{X}_n}\right)} = \frac{\frac{\bar{X}_n}{1-\bar{X}_n}}{\frac{1}{(1-\bar{X}_n)^2}\frac{2-\bar{X}_n}{1-\bar{X}_n}} = \frac{\bar{X}_n\left(1-\bar{X}_n\right)^2}{2-\bar{X}_n}$$

and

$$\begin{aligned}\hat{\sigma}_{X,n}^{2,ML} &= \frac{\hat{\theta}_n^{ML}}{\left(1+\hat{\theta}_n^{ML}\right)^2\left(2+\hat{\theta}_n^{ML}\right)} = \frac{-\frac{n}{\sum_{k=1}^n \log(X_k)}}{\left(1-\frac{n}{\sum_{k=1}^n \log(X_k)}\right)^2\left(2-\frac{n}{\sum_{k=1}^n \log(X_k)}\right)} \\ &= \frac{-\frac{n}{\sum_{k=1}^n \log(X_k)}}{\left(\frac{\sum_{k=1}^n \log(X_k)-n}{\sum_{k=1}^n \log(X_k)}\right)^2\left(\frac{2\sum_{k=1}^n \log(X_k)-n}{\sum_{k=1}^n \log(X_k)}\right)} = \frac{n\left(\sum_{k=1}^n \log(X_k)\right)^2}{\left(n-\sum_{k=1}^n \log(X_k)\right)\left(n-2\sum_{k=1}^n \log(X_k)\right)}.\end{aligned}$$

Problem 6 *Let X a random variable representing a characteristic of a certain population. Assume that X has a density $f_X : \mathbb{R} \to \mathbb{R}$ given by*

$$f_X(x) \stackrel{def}{=} \frac{1}{\theta} e^{-\frac{x-3}{\theta}} 1_{[3,+\infty)}(x), \quad \forall x \in \mathbb{R},$$

where θ is a positive parameter.

1. *Apply the method of moments to find the estimator $\hat{\theta}_M$ of the parameter θ.*
2. *Apply the maximum likelihood method to find the estimator $\hat{\theta}_{ML}$ of the parameter θ.*
3. *Use the estimators $\hat{\theta}_M$ and $\hat{\theta}_{ML}$ to build estimators for $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$.*

Solution. .

Problem 7 *Let $\theta > 0$ and let X be an absolutely continuous real random variable with density function $f_X : \mathbb{R} \to \mathbb{R}_+$ given by*

$$f_X(x) \stackrel{def}{=} \frac{1}{2} e^{-|x-\theta|}, \quad \forall x \in \mathbb{R}.$$

1. *Apply the method of moments to determine the estimator $\hat{\theta}_n^M$ for θ.*
2. *Check whether $\hat{\theta}_n^M$ is unbiased, consistent in probability, and consistent in mean square.*
3. *Can you "guess" the result of the method of maximum likelihood to determine the estimator $\hat{\theta}_n^{ML}$ for θ?*

   Hint: recall that an estimator $\hat{\theta}_n$ for the true value of a parameter θ is said to be consistent in probability [resp. in mean square] if

$$\hat{\theta}_n \xrightarrow{\mathbf{P}} \theta \quad \text{[resp. } \hat{\theta}_n \xrightarrow{\mathbf{L}^2} \theta\text{]},$$

   as $n \to \infty$.

Solution. .

Problem 8 *Let* $\lambda > 0$ *and let* X *be a Poisson real random variable with rate parameter* λ*, in symbols* $X \sim Poiss(\lambda)$*. Consider a simple random sample* $X_1, \dots, X_n$ *of size* n *drawn from* X*.*

1. *Let* Z_n *be the sample sum* $X_1, \dots, X_n$*, namely* $Z_n \equiv \sum_{k=1}^{n} X_k$*. Write the distribution of* Z_n *and compute* $\mathbf{E}[Z_n]$ *and* $\mathbf{D}^2[Z_n]$*.*

2. *Consider the sample mean* $\bar{X}_n \equiv \frac{1}{n} Z_n$ *and the unbiased sample variance* $S_n^2 \equiv \frac{1}{n-1} \sum_{k=1}^{n} \left(X_k - \bar{X}_n\right)^2$*. Might they both be used to estimate* λ*? Which would perform better?*

Solution. .