541A Final Solutions[1]

1. Question 1

Let $X_1, X_2, \ldots$ be real-valued random variables that converge in probability to a constant $a \in \mathbf{R}$.

Let $h\colon \mathbf{R} \to \mathbf{R}$ be a continuous function. (For any $x \in \mathbf{R}$, for any $\varepsilon > 0$, there exists $\delta > 0$ such that, if $y \in \mathbf{R}$ satisfies $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$.)

Show that $h(X_1), h(X_2), \ldots$ converges in probability to $h(a)$.

Solution. Let $\varepsilon > 0$. We are required to show that

$$\lim_{n\to\infty} \mathbf{P}(|h(X_n) - h(a)| \leq \varepsilon) = 1. \qquad (**)$$

Since h is continuous, there exists $\delta > 0$ such that, if $b \in \mathbf{R}$ satisfies $|a - b| \leq \delta$, then $|h(b) - h(a)| < \varepsilon/2$

Since $X_1, X_2, \ldots$ converges in probability,

$$\lim_{n\to\infty} \mathbf{P}(|X_n - a| \leq \delta) = 1. \qquad (*)$$

If $|X_n - a| \leq \delta$, the continuity of h says that $|h(X_n) - h(a)| < \varepsilon/2$. That is,

$$\mathbf{P}(|X_n - a| \leq \delta) \leq \mathbf{P}(|h(X_n) - h(a)| \leq \varepsilon/2).$$

So, $(*)$ implies $(**)$, as desired.

2. Question 2

Let $X_1, \ldots, X_n$ be i.i.d. random variables, so that X_1 has PDF $f_\theta\colon \mathbf{R} \to [0, \infty)$, where $\theta = (\theta_1, \theta_2) \in \mathbf{R}^2$ is an unknown parameter.

Let Y be a statistic (so that Y is a function of $X_1, \ldots, X_n$). In all cases below, as usual, you must **justify your answer**.

(i) Suppose Y is sufficient for θ. Is it true that Y is sufficient for θ_1?
(ii) Suppose Y is sufficient for θ_1, and Y is sufficient for θ_2. Is it true that Y is sufficient for θ?
(iii) Suppose Y is minimal sufficient for θ_1, and Y is minimal sufficient for θ_2. Is it true that Y is minimal sufficient for θ?

Solution. Let $X = (X_1, \ldots, X_n)$. For (i), by assumption the PDF of $X|Y = y$ does not depend on θ. In particular the PDF of $X|Y = y$ does not depend on θ_1. So, yes, Y is sufficient for θ_1.

For (ii), by assumption the PDF of $X|Y = y$ does not depend on θ_1, and the PDF of $X|Y = y$ does not depend on θ_2. Therefore, the PDF of $X|Y = y$ does not depend on θ. So, yes, Y is sufficient for θ.

For (iii), note that Y is sufficient for θ by part (ii). Now, by minimal sufficiency, if Z is sufficient for θ_1, then Y is a function of Z. Now let W be sufficient for θ. We need to show that Y is a function of W. By part (i), W is sufficient for θ_1, so Y is a function of W.

[1]April 25, 2023,

3. Question 3

Find real-valued random variables $X, Y, X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ such that the following holds.

- $X_1, X_2, \ldots$ converges in distribution to X.
- $Y_1, Y_2, \ldots$ converges in distribution to Y.
- $X_1 + Y_1, X_2 + Y_2, \ldots$ does **NOT** converges in distribution to any random variable.

Solution. Let X be a standard Gaussian random variable. Let $X_1, X_2, \ldots$ be the sequence $X, X, X, \ldots$. Then $X_1, X_2, \ldots$ converges in distribution to X. Let $Y_1, Y_2, \ldots$ be the sequence $X, -X, X, -X, \ldots$. Then $Y_1, Y_2, \ldots$ converges in distribution to $X =: Y$ (since $-X$ and X have the same distribution). However,

$$X_1 + Y_1, X_2 + Y_2, \ldots \qquad = \qquad 2X, 0, 2X, 0, 2X, 0, \ldots.$$

So, $X_1 + Y_1, X_2 + Y_2, \ldots$ does not converge in distribution, since its odd subsequence is equal in distribution to two times a fixed Gaussian, whereas its even subsequence is equal to zero.

4. Question 4

Let $X_1, \ldots, X_n$ be i.i.d. random variables, so that X_1 has PDF $f_\theta \colon \mathbf{R} \to [0, \infty)$, where $\theta > 0$ is an unknown parameter and

$$f_\theta(x) := \frac{\theta^2}{2} e^{-\theta^2 |x|}, \qquad \forall\, x \in \mathbf{R}.$$

- Find the MLE Y_n of θ. [Warning: do NOT find the MLE of θ^2.]
- Compute the Fisher information $I_{X_1}(\theta)$.
  (You can freely use without proof that $\mathbf{E} X_1^2 = 2\theta^{-4}$ and $\mathbf{E}\,|X_1| = \theta^{-2}$.)
- Find a random variable Z such that $\sqrt{n}(Y_n - \theta)$ converges in distribution to Z. (You can freely use without proof that $Y_1, Y_2, \ldots$ converges in probability to θ.)

Solution. We have

$$\log \prod_{i=1}^n f_\theta(x_i) = \log \left(\theta^{2n} 2^{-n} e^{-\theta^2 \sum_{i=1}^n |x_i|} \right) = 2n \log \theta - n \log 2 - \theta^2 \sum_{i=1}^n |x_i|.$$

Differentiating in θ gives

$$\frac{d}{d\theta} \log \prod_{i=1}^n f_\theta(x_i) = \frac{2n}{\theta} - 2\theta \sum_{i=1}^n |x_i|. \qquad (*)$$

When $\theta > 0$ is small, the first term is large and positive and the second term is close to zero. So, this derivative is positive when $\theta > 0$ is small. And both terms are decreasing in θ, with the first term going to zero as $\theta \to \infty$ and the second term going to $-\infty$ (unless $x_1 = \cdots = x_n = 0$.) So, the first derivative test implies that the log likelihood increases and then decreases, and it has a unique maximum when

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

That is, (using also $\theta > 0$)

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

In the case $x_1 = \cdots = x_n = 0$, the MLE does not exist, since $f_\theta(x) = \theta^{2n}2^{-n}$ in that case.

Squaring $(*)$ with $n = 1$ and taking the variance,

$$I_{X_1}(\theta) = \mathrm{Var}_\theta(2/\theta - 2\theta\,|X_1|) = 4\theta^2\mathrm{Var}_\theta(X_1) = 4\theta^2[\mathbf{E}X_1^2 - (\mathbf{E}\,|X_1|)^2] = 4\theta^2[\theta^{-4}] = 4\theta^{-2}.$$

Here we used

$$\begin{aligned}\mathbf{E}X_1^2 &= \int_{\mathbf{R}} x^2\frac{\theta^2}{2}e^{-\theta^2|x|}dx = \int_0^\infty x^2\theta^2 e^{-\theta^2 x}dx = \int_0^\infty x^2(-1)\frac{d}{dx}e^{-\theta^2 x}dx = \int_0^\infty 2xe^{-\theta^2 x}dx\\ &= \frac{-1}{\theta^2}\int_0^\infty 2x\frac{d}{dx}e^{-\theta^2 x}dx = \frac{1}{\theta^2}2\int_0^\infty e^{-\theta^2 x}dx = \frac{2}{\theta^4}.\end{aligned}$$

$$\begin{aligned}\mathbf{E}\,|X_1| &= \int_{\mathbf{R}} |x|\,\frac{\theta^2}{2}e^{-\theta^2|x|}dx = \int_0^\infty x\theta^2 e^{-\theta^2 x}dx = \int_0^\infty x(-1)\frac{d}{dx}e^{-\theta^2 x}dx\\ &= \int_0^\infty e^{-\theta^2 x}dx = \frac{1}{\theta^2}.\end{aligned}$$

Finally, we would like to apply the Theorem about the limiting distribution of the MLE. Let us verify a few assumptions. Note that $\{x \in \mathbf{R} \colon f_\theta(x) > 0\} = \mathbf{R}$, and this set does not depend on θ. Also, the joint PDF is twice continuously differentiable in $\theta > 0$. Moreover, for any $\theta > 0$, if we choose $\varepsilon := \theta/2$, then

$$\begin{aligned}\mathbf{E}_\theta \sup_{\theta' \in [\theta-\varepsilon, \theta+\varepsilon]}\left|\frac{d^2}{d[\theta']^2}\log f_{\theta'}(X_1)\right| &= \mathbf{E}_\theta \sup_{\theta' \in [\theta/2, 3\theta/2]}\left|-2[\theta']^{-2} - 2\,|X_1|\right|\\ &= \mathbf{E}_\theta \max\Big(\left|8\theta + 2\,|X_1|\right|, \left|8\theta/9 + 2\,|X_1|\right|\Big)\\ &\leq \mathbf{E}_\theta \left|8\theta + 2\,|X_1|\right| \leq 8\theta + 2\mathbf{E}_\theta\,|X_1| = 8\theta + 2\theta^{-2} < \infty.\end{aligned}$$

So, Theorem 6.53 from the notes applies. We conclude that

$$\sqrt{n}(Y_n - \theta)$$

converges in distribution to a mean zero Gaussian random variable with variance

$$\frac{1}{I_{X_1}(\theta)} = \frac{\theta^2}{4}.$$

Solution 2. The last part of the problem can also be proven using the Delta Method. The random variable $|X_1|$ satisfies $\mathbf{E}\,|X_1| = \theta^{-2}$ and $\mathbf{E}X_1^2 = 2\theta^{-4}$ with $\mathrm{Var}(X_1) = \theta^{-4}$. The CLT then says that

$$\frac{\sum_{i=1}^n |X_i| - n\theta^{-2}}{\sqrt{n}\theta^{-2}}$$

converges in distribution to a mean zero variance one Gaussian as $n \to \infty$. That is,

$$\sqrt{n}\Big(\frac{1}{n}\sum_{i=1}^n |X_i| - \theta^{-2}\Big)$$

converges in distribution to a mean zero variance θ^{-4} Gaussian as $n \to \infty$. So, using $f(t) := t^{-1/2}$ with the Delta Method,

$$\sqrt{n}\left(f\left(\frac{1}{n}\sum_{i=1}^n |X_i|\right) - f(\theta^{-2})\right) = \sqrt{n}\Big(\sqrt{\frac{n}{\sum_{i=1}^n |X_i|}} - \theta\Big)$$

converges in distribution to a mean zero Gaussian with variance

$$\theta^{-4}(f'(\theta^{-2}))^2 = \theta^{-4}(1/4)(\theta^{-2(3/2)})^2 = (1/4)\theta^{-2}.$$

5. Question 5

Let $X_1, \ldots, X_n$ be i.i.d. random variables, so that X_1 has PDF $f_\theta\colon \mathbf{R} \to [0,\infty)$, where $\theta > 0$ is an unknown parameter and

$$f_\theta(x) := \frac{2x}{\theta^2}, \qquad \forall\, 0 \leq x \leq \theta.$$

- Find the method of moments estimator Y_n of θ. Is Y_n unbiased?
- Find constants a_n, b_n such that $a_n(Y_n - b_n)$ converges in distribution as $n \to \infty$ to a mean zero variance one Gaussian random variable.
  (Hint: you should find that $\mathrm{Var}X_1 = \theta^2/18$.)

Solution. We have

$$\mathbf{E}X_1 = \int_0^\theta \frac{2x}{\theta^2} x dx = \theta^{-2}[(2/3)x^3]_{x=0}^{x=\theta} = \theta^{-2}(2/3)\theta^3 = (2/3)\theta.$$

So, $\theta = (3/2)\mathbf{E}X_1$, and the MoM estimator of θ is

$$Y_n := \frac{3}{2}\frac{1}{n}\sum_{i=1}^n X_i.$$

Since $\mathbf{E}Y_n = \frac{3}{2}\mathbf{E}X_1 = \theta$, Y_n is unbiased.

Also,

$$\mathbf{E}X_1^2 = \int_0^\theta \frac{2x}{\theta^2} x^2 dx = \theta^{-2}[(1/2)x^4]_{x=0}^{x=\theta} = \theta^{-2}(1/2)\theta^4 = (1/2)\theta^2.$$

So,

$$\mathrm{Var}(X_1) = (1/2)\theta^2 - [(2/3)\theta]^2 = \theta^2[1/2 - 4/9] = \theta^2/18.$$

From the central limit theorem,

$$\frac{X_1 + \cdots + X_n - n\mathbf{E}X_1}{\sqrt{n}\sqrt{\theta^2/18}} = \frac{X_1 + \cdots + X_n - n(2/3)\theta}{\sqrt{n}\theta/[3\sqrt{2}]} = \sqrt{n}\frac{\left(\frac{1}{n}\sum_{i=1}^n X_i\right) - (2/3)\theta}{\theta/[3\sqrt{2}]}$$

$$= \sqrt{n}\frac{\frac{3}{2}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) - \theta}{\theta/[2\sqrt{2}]} = \sqrt{n}\frac{2\sqrt{2}}{\theta}(Y_n - \theta).$$

converges in distribution to a standard Gaussian random variable (mean zero variance one) as $n \to \infty$.

6. Question 6

Let $X_1, \ldots, X_n$ be i.i.d. Gaussian random variables with unknown mean $\mu \in \mathbf{R}$ and (known) variance 1.

In this problem, you can freely use that the sample mean $M_n := \frac{1}{n}\sum_{i=1}^n X_i$ is complete and sufficient for μ.

- Find a minimal sufficient statistic for μ.
- Find an unbiased estimator Y_n of the quantity $\mathbf{P}_\mu(X_1 \leq 0)$.

- Using any method you want to use, find the UMVU of the quantity $\mathbf{P}_\mu(X_1 \leq 0)$.

Solution. By Bahadur's Theorem 5.25 in the notes, M_n is minimal sufficient for μ.

Let $Y_n := 1_{X_1 \leq 0}$. Then $\mathbf{E}Y_n = \mathbf{P}(X_1 \leq 0)$. That is, Y_n is unbiased.

Now, we know M_n is complete and sufficient for μ. We claim also that M_n is complete and sufficient for $\mathbf{P}_\mu(X_1 \leq 0) = \int_{-\infty}^{0} e^{-(s-\mu)^2/2}ds/\sqrt{2\pi} = \int_{-\infty}^{\mu} e^{-s^2/2}ds/\sqrt{2\pi} =: \Phi(\mu)$. Note that $\Phi(\mu)$ is strictly increasing in μ, and therefore it is invertible with inverse Φ^{-1}. It follows that M_n is complete and sufficient for $\Phi(\mu)$. (We have $Z|(Y_1, \ldots, Y_n)$ does not depend on μ if and only if $Z|(Y_1, \ldots, Y_n)$ does not depend on $\Phi(\mu)$ since Φ is invertible, hence the sufficiency. Similarly, the condition for completeness of a statistic Z holds for all $\mu \in \mathbf{R}$ if and only if it holds for all $\Phi(\mu)$ with $\mu \in \mathbf{R}$.)

To find the UMVU of $\mathbf{P}(X_1 \leq 0)$, we use the conditioning method, i.e. the Lehmann-Scheffé Theorem 6.14 in the notes to conclude that $\mathbf{E}(1_{X_1 \leq 0}|M_n)$ is the UMVU. To find an explicit formula for the UMVU, we compute this conditional expectation explicitly. For any $t \in \mathbf{R}$, we have

$$
\begin{aligned}
&\mathbf{E}(1_{X_1} \leq 0 \,|\, M_n \leq t) \\
&\quad = \mathbf{P}(X_1 \leq 0 \,|\, X_1 + \cdots + X_n \leq nt) \\
&\quad = \mathbf{P}(X_1 + \cdots + X_n \leq nt \,|\, X_1 \leq 0)\frac{\mathbf{P}(X_1 \leq 0)}{\mathbf{P}(X_1 + \cdots + X_n \leq nt)} \\
&\quad = \int_{-\infty}^{0} \mathbf{P}(X_1 + \cdots + X_n \leq nt \,|\, X_1 = u) f_{X_1}(u)du \cdot \frac{\mathbf{P}(X_1 \leq 0)}{\mathbf{P}(X_1 + \cdots + X_n \leq nt)} \\
&\quad = \int_{-\infty}^{0} \mathbf{P}(X_2 + \cdots + X_n \leq nt - u) f_{X_1}(u)du \cdot \frac{\mathbf{P}(X_1 \leq 0)}{\mathbf{P}(X_1 + \cdots + X_n \leq nt)} \\
&\quad = \int_{-\infty}^{0} \int_{-\infty}^{nt-u} \frac{e^{-(s-\mu)^2/[2(n-1)]}}{\sqrt{2\pi(n-1)}} ds \frac{e^{-(u-\mu)^2/2}}{\sqrt{2\pi}} du \frac{\int_{-\infty}^{0} e^{-(s-\mu)^2/2}\frac{ds}{\sqrt{2\pi}}}{\int_{-\infty}^{nt} e^{-(s-\mu)^2/[2n]}\frac{ds}{\sqrt{2\pi n}}}.
\end{aligned}
$$

Denote $g(t) := \mathbf{P}(X_1 \leq 0 \,|\, [X_1 + \cdots + X_n]/n \leq t)$. Using Bayes rule $\mathbf{P}(A|B) = \mathbf{P}(B|A)\mathbf{P}(A)/\mathbf{P}(B)$, we have

$$
g(t) = \frac{\mathbf{P}([X_1 + \cdots + X_n]/n \leq t \,|\, X_1 \leq 0)\mathbf{P}(X_1 \leq 0)}{\mathbf{P}([X_1 + \cdots + X_n]/n \leq t}.
$$

Then, using the quotient rule and Bayes rule again,

$$\begin{aligned}
g'(t) &= \mathbf{P}(X_1 \leq 0)\frac{\mathbf{P}([X_1+\cdots+X_n]/n \leq t)\mathbf{P}([X_1+\cdots+X_n]/n = t \,|\, X_1 \leq 0)}{[\mathbf{P}([X_1+\cdots+X_n]/n \leq t)]^2} \\
&- \mathbf{P}(X_1 \leq 0)\frac{\mathbf{P}([X_1+\cdots+X_n]/n \leq t \,|\, X_1 \leq 0)\mathbf{P}([X_1+\cdots+X_n]/n = t)}{[\mathbf{P}([X_1+\cdots+X_n]/n \leq t)]^2} \\
&= \mathbf{P}(X_1 \leq 0)\frac{\mathbf{P}([X_1+\cdots+X_n]/n = t \,|\, X_1 \leq 0)}{\mathbf{P}([X_1+\cdots+X_n]/n \leq t)} - g(t)\frac{\mathbf{P}([X_1+\cdots+X_n]/n = t)}{\mathbf{P}([X_1+\cdots+X_n]/n \leq t)} \\
&= \mathbf{P}(X_1 \leq 0)\frac{\mathbf{P}([X_1+\cdots+X_n]/n = t \,|\, X_1 \leq 0)}{\mathbf{P}([X_1+\cdots+X_n]/n \leq t)} - g(t)\frac{\mathbf{P}([X_1+\cdots+X_n]/n = t)}{\mathbf{P}([X_1+\cdots+X_n]/n \leq t)} \\
&= \Big[\mathbf{P}(X_1 \leq 0 \,|\, [X_1+\cdots+X_n]/n = t) - g(t)\Big]\frac{\mathbf{P}([X_1+\cdots+X_n]/n = t)}{\mathbf{P}([X_1+\cdots+X_n]/n \leq t)} \\
&= \Big[\mathbf{E}(1_{X_1\leq 0}|M_n = t) - g(t)\Big]\frac{\mathbf{P}([X_1+\cdots+X_n]/n = t)}{\mathbf{P}([X_1+\cdots+X_n]/n \leq t)} \\
&= \Big[\mathbf{E}(1_{X_1\leq 0}|M_n = t) - g(t)\Big]\frac{e^{-(n/2)(t-\mu)^2}}{\int_{-\infty}^{t} e^{-(n/2)(s-\mu)^2}ds}
\end{aligned}$$

So,

$$\mathbf{E}(1_{X_1\leq 0} \,|\, M_n = t) = g(t) + g'(t)\frac{\int_{-\infty}^{t} e^{-(n/2)(s-\mu)^2}ds}{e^{-(n/2)(t-\mu)^2}}$$

$$= \int_{-\infty}^{0}\int_{-\infty}^{nt-u} \frac{e^{-(s-\mu)^2/[2(n-1)]}}{\sqrt{2\pi(n-1)}}ds\frac{e^{-(u-\mu)^2/2}}{\sqrt{2\pi}}du \cdot \frac{\int_{-\infty}^{0} e^{-(s-\mu)^2/2}\frac{ds}{\sqrt{2\pi}}}{\int_{-\infty}^{nt} e^{-(s-\mu)^2/[2n]}\frac{ds}{\sqrt{2\pi n}}}$$

$$+ \frac{\int_{-\infty}^{t} e^{-(n/2)(s-\mu)^2}ds}{e^{-(n/2)(t-\mu)^2}}\frac{d}{dt}\Big[\int_{-\infty}^{0}\int_{-\infty}^{nt-u} \frac{e^{-(s-\mu)^2/[2(n-1)]}}{\sqrt{2\pi(n-1)}}ds\frac{e^{-(u-\mu)^2/2}}{\sqrt{2\pi}}du\frac{\int_{-\infty}^{0} e^{-(s-\mu)^2/2}\frac{ds}{\sqrt{2\pi}}}{\int_{-\infty}^{nt} e^{-(s-\mu)^2/[2n]}\frac{ds}{\sqrt{2\pi n}}}\Big].$$

And the UMVU is

$$\mathbf{E}(1_{X_1\leq 0} \,|\, M_n) = \int_{-\infty}^{0}\int_{-\infty}^{nM_n-u} \frac{e^{-(s-\mu)^2/[2(n-1)]}}{\sqrt{2\pi(n-1)}}ds\frac{e^{-(u-\mu)^2/2}}{\sqrt{2\pi}}du \cdot \frac{\int_{-\infty}^{0} e^{-(s-\mu)^2/2}\frac{ds}{\sqrt{2\pi}}}{\int_{-\infty}^{nM_n} e^{-(s-\mu)^2/[2n]}\frac{ds}{\sqrt{2\pi n}}}$$

$$\frac{\int_{-\infty}^{M_n} e^{-(n/2)(s-\mu)^2}ds}{e^{-(n/2)(M_n-\mu)^2}}\frac{d}{dt}\Big|_{t=M_n}\Big[\int_{-\infty}^{0}\int_{-\infty}^{nt-u} \frac{e^{-(s-\mu)^2/[2(n-1)]}}{\sqrt{2\pi(n-1)}}ds\frac{e^{-(u-\mu)^2/2}}{\sqrt{2\pi}}du\frac{\int_{-\infty}^{0} e^{-(s-\mu)^2/2}\frac{ds}{\sqrt{2\pi}}}{\int_{-\infty}^{nt} e^{-(s-\mu)^2/[2n]}\frac{ds}{\sqrt{2\pi n}}}\Big].$$

(Even though a bunch of μ terms appear here, this expression should not depend on μ.)

7. Question 7

Let $X_1, \ldots, X_n$ be i.i.d. random variables that are uniformly distributed in $[\theta-1/2, \theta+1/2]$, where $\theta \in \mathbf{R}$ is unknown. Note that $\mathbf{E}X_1 = \theta$, $\mathbf{E}X_1^2 = \theta^2 + 1/12$.

- Give two different method of moments estimators that estimate θ.
- Show that an MLE for θ is not unique. That is, describe two different maximum likelihood estimators for θ.

- Is the Fisher information $I_{X_1}(\theta)$ well-defined? Explain.
- Show that any MLE for θ is consistent.

Solution. Since $\theta = 2\mathbf{E}X_1$, one MoM estimator is $2\frac{1}{n}\sum_{i=1}^{n} X_i$. Since $\theta = \sqrt{|-1/12 + \mathbf{E}X_1^2|}$, another MoM estimator is $\sqrt{\left|-1/12 + \frac{1}{n}\sum_{i=1}^{n} X_i^2\right|}$.

The joint PDF of $X_1, \ldots, X_n$ is $\prod_{i=1}^{n} 1_{X_i \in [\theta-1/2, \theta+1/2]} = 1_{X_{(1)}, X_{(n)} \in [\theta-1/2, \theta+1/2]}$. So, any θ satisfying

$$\theta - 1/2 \leq X_{(1)} \leq X_{(n)} \leq \theta + 1/2$$

is an MLE for θ, since the value of the joint PDF there is 1, and otherwise the value is zero. That is, the MLE can take any value in the open interval $(X_{(n)} - 1/2, X_{(1)} + 1/2)$. To see that this interval can be nonempty, suppose for example that $n = 1$, so that $X_{(1)} = X_{(n)}$. The MLE is then not unique.

The Fisher information is not well-defined since the derivative of the joint PDF (viewed as a function of θ) is not well-defined (since the region where the PDF is nonzero does depend on θ). In particular, the derivative of the joint PDF does not exist

Finally, note that the joint PDF of $X_1, \ldots, X_n$ is $\prod_{i=1}^{n} 1_{X_i \in [\theta-1/2, \theta+1/2]} = 1_{X_{(1)}, X_{(n)} \in [\theta-1/2, \theta+1/2]}$. So, any θ satisfying

$$\theta - 1/2 \leq X_{(1)} \leq X_{(n)} \leq \theta + 1/2$$

is an MLE for θ. As $n \to \infty$, $X_{(1)}$ converges in probability to $\theta - 1/2$, and $X_{(n)}$ converges in probability to $\theta + 1/2$, so as $n \to \infty$ any MLE for θ converges to θ. (For example, we know that $\mathbf{P}(X_{(n)} \leq t) = \mathbf{P}(X_1 \leq t)^n$ which converges to $1_{t \leq \theta + 1/2}$ as $n \to \infty$.)

Note that we cannot use the consistency theorem for MLEs that we discussed in class, since the set of parameters $\Theta = \mathbf{R}$ is not compact. Similarly, Theorem 9.11 in the Keener book is not applicable since the PDF is not continuous in θ.

8. Question 8

Prove the Cramér-Rao inequality:

Let $X : \Omega \to \mathbf{R}^n$ be a random variable with distribution from a family of multivariable PDFs $\{f_\theta : \theta \in \Theta\}$ with $\Theta \subseteq \mathbf{R}$. Let $t : \mathbf{R}^n \to \mathbf{R}$ and let $Y := t(X)$ be statistic. For any $\theta \in \Theta$ let $g(\theta) := \mathbf{E}_\theta Y$. Then

$$\mathrm{Var}_\theta(Y) \geq \frac{|g'(\theta)|^2}{I_X(\theta)}, \qquad \forall\, \theta \in \Theta.$$

Moreover, if $I_X(\theta) = 0$, then $g'(\theta) = 0$.

(You are allowed to differentiate under any integral in your proof. Also, we assume that $\{x \in \mathbf{R}^n : f_\theta(x) > 0\}$ does not depend on θ, and for a.e. $x \in \mathbf{R}^n$, $(d/d\theta) f_\theta(x))$ exists and is finite, and the Fisher information satisfies any identity we have ever shown it to satisfy in this course.)

Solution. Before the proof, we prove two facts

Fact 1. $\mathbf{E}\frac{d}{d\theta} \log f_\theta(X) = 0$.

Proof of Fact 1.

$$\mathbf{E}\frac{d}{d\theta}\log f_\theta(X) = \int_{\mathbf{R}^n} \frac{\frac{d}{d\theta} f_\theta(x)}{f_\theta(x)} f_\theta(x)\,dx = \int_{\mathbf{R}^n} \frac{d}{d\theta} f_\theta(x)\,dx$$
$$= \frac{d}{d\theta}\int_{\mathbf{R}^n} f_\theta(x)\,dx = \frac{d}{d\theta}(1) = 0.$$

Fact 2. If $\mathbf{E}W = 0$, then $\mathbf{E}(WZ) = \mathrm{cov}(W, Z)$.

Proof of Fact 2. For the first equality, note that, since $\mathbf{E}W = 0$,

$$\mathrm{cov}(W, Z) = \mathbf{E}(W - \mathbf{E}W)(Z - \mathbf{E}Z) = \mathbf{E}(WZ) - \mathbf{E}W\mathbf{E}Z = \mathbf{E}WZ.$$

Fact 3. $\mathrm{cov}(W, Z) \leq \sqrt{\mathrm{Var}(W)}\sqrt{\mathrm{Var}(Z)}$. This inequality follows from the Cauchy-Schwarz Inequality.

$$|g'(\theta)| = \left|\frac{d}{d\theta}\int_{\mathbf{R}^n} f_\theta(x)t(x)dx\right| = \left|\int_{\mathbf{R}^n} \frac{d}{d\theta}\log f_\theta(x)t(x)f_\theta(x)dx\right| = \left|\mathbf{E}_\theta\frac{d}{d\theta}\log f_\theta(X)t(X)\right|$$
$$\overset{(1)\wedge(2)}{=} \left|\mathrm{Cov}_\theta\Big(\frac{d}{d\theta}\log f_\theta(X), t(X)\Big)\right| \overset{(3)}{\leq} \sqrt{\mathrm{Var}_\theta\Big(\frac{d}{d\theta}\log f_\theta(X)\Big)\mathrm{Var}_\theta(t(X))}$$
$$= \sqrt{I_X(\theta)\mathrm{Var}_\theta(t(X))}.$$

The last equality uses the definition of Fisher Information.