Dauphine | PSL — UNIVERSITÉ PARIS

MIDO - L3 Math. Appliquées 2023–2024

Statistical modelling

Examen final du 18 Janvier 2024

DURÉE 2H00 – DOCUMENTS ET CALCULATRICE NON AUTORISÉS

French – English Lexicon

- *i.i.d.* : *independent and identically distributed*
- échantillon : *sample*
- fonction de répartition : *cumulative distribution function*
- fonction de densité : *probability distribution function*
- fonction génératrice des moments : *moment-generating function*
- famille exponentielle : *exponential family*
- espace naturel des paramètres : *natural parameter space*
- vraisemblance : *likelihood*
- statistique exhaustive : *sufficient statistic*
- statistique libre : *ancillary statistic*
- statistique complète : *complete statistic*

Exercise 1

...... /**9**

For the following statements, give the correct answer(s). Incorrect answers and missing justification return zero point while incomplete answers gain partial points.

1. Let $(X_n)_{n\in\mathbb{N}^*}$ be a sequence of independent discrete random variables such that

$$\mathbb{P}[X_n = 0] = \frac{n-1}{n} \quad \text{and} \quad \mathbb{P}[X_n = \sqrt{n}] = \frac{1}{n}.$$

Then, when n goes to $+\infty$,

(a) the sequence converges in L^1 (convergence in mean),

(b) the sequence converges in L^2 (convergence in quadratic mean),

(c) for any continuous function g, $\mathbb{E}[g(X_n)]$ converges to 0,

(d) the sequence converges in distribution,

(e) the sequence does not converge at all.

(a, d) For any continuous and bounded function g, we have

$$\mathbb{E}[g(X_n)] = \frac{n-1}{n} g(0) + \frac{1}{n} g(\sqrt{n}).$$

Since g is bounded, the second term in the above sum converges to 0 when n goes to $+\infty$ and thus

$$\mathbb{E}[g(X_n)] \underset{n\to+\infty}{\longrightarrow} g(0).$$

Then, (X_n) converges in distribution to 0. Moreover, if it converges in L^p, $p \in \mathbb{N}^*$, it is necessarily to 0. We have $\mathbb{E}[X_n] = 1/\sqrt{n}$ and $\mathbb{E}\left[X_n^2\right] = 1$. Thus, X_n converges in L^1 to 0, but not in L^2.

2. Consider the exponential family associated to the Bernoulli distribution with unknown parameter $p \in (0, 1)$. The moment generating function of natural statistic $T(x) = x$ is given for $t \in \Theta \subseteq \mathbb{R}$ by

(a) $(1-p)/(1-p-t)$ (b) $1 + p(\exp(t) - 1)$ (c) $(1-p-t)/(1-p)$ (d) $1/[1+p(\exp(t)-1)]$

(b) To get the canonical form we set $\theta = \log(p) - \log(1-p)$. The canonical form is then

$$f(x \mid \theta) = \frac{1}{1+\exp(\theta)} \exp(\theta x), \quad \theta \in \mathbb{R}.$$

Then the moment generating function is defined for any $t \in \mathbb{R}$ by

$$M(t) = \frac{1+\exp(\theta+t)}{1+\exp(\theta)} = \frac{1+\exp[\log(p)-\log(1-p)+t]}{1+\exp[\log(p)-\log(1-p)]} = 1 + p(\exp(t)-1).$$

3. Consider the density (with respect to the Lebesgue measure on $\mathbb{R}$) parametrised by an **unknown** $(k, \lambda) \in \mathbb{N}^* \times \mathbb{R}_+^*$ and defined by

$$f(x \mid k, \lambda) = \frac{\lambda^k x^{k-1} \exp(-\lambda x)}{(k-1)!} \mathbb{1}_{x \geq 0}.$$

(a) It constitutes a minimal and canonical exponential family.

(b) It constitutes a minimal exponential family but is not in a canonical form.

(c) It constitutes an exponential family that is neither minimal nor canonical.

(d) None of the other answers.

(a) The density writes as

$$f(x \mid k, \lambda) = \frac{\lambda^k}{(k-1)!} \frac{1}{x} \mathbb{1}_{x \geq 0} \exp[k \log(x) - \lambda x].$$

Then it constitutes a canonical exponential family with natural parameter (k, λ) and natural statistic $T(x) = (\log(x), -x)$. Moreover for $(\alpha_1, \alpha_2) \in \mathbb{R}^* \times \mathbb{R}^*$ and $c \in \mathbb{R}$, the set $\{x \in \mathbb{R}_+ \,;\, \alpha_1 \log(x) - \alpha_2 x - c = 0\}$ contains at most 2 elements (maximal number of intersections between an affine function and $x \mapsto \log(x)$) and hence has measure zero (null set) for the Lebesgue measure. The family is minimal.

4. We run an experiment where we measure how much time n different customers spend on a specific page of a website. Our observations $x_1, \ldots, x_n$ are stored in a vector `x`. We assume that the underlying statistical model is a Gamma distribution with parameter (α, β). Which one among the following R command lines does return the first quartile of the sample ?

(a) `rgamma(0.25, 1, 2)`

(b) `pgamma(0.25, 1, 2)`

(c) `dgamma(0.25, 1, 2)`

(d) `qgamma(0.25, 1, 2)`

(e) `quantile(0.25, 1, 2)`

(f) `quantile(x, 0.25)`

(f) In order to get the empirical quantiles of a sample we use the function `quantile`. The first argument is the sample, followed by the order of the quantiles we are interested in.

5. Let X be a random variable with density, parametrised by $\lambda \in \mathbb{R}_+^*$, with respect to the composition of the counting measure on $\mathbb{N}$ and the Lebesgue measure on $\mathbb{R}_+^*$:

$$f_X(x) = \begin{cases} \frac{\lambda^x \exp(-\lambda)}{2(x!)}, & \text{if} \quad x \in \mathbb{N}, \\ \frac{\lambda}{2} \exp(-\lambda x), & \text{otherwise.} \end{cases}$$

The likelihood for the sample $(1, 1, 2, 2, 2, x_1, \ldots, x_n)$, with $x_1, \ldots, x_n \notin \mathbb{N}$ is

(a) $\dfrac{\lambda^n}{2^{n+5}} \exp\left(-\lambda \sum_{i=1}^n x_i\right)$

(b) $\dfrac{\lambda^{n+8}}{2^{n+8}} \exp\left[-\lambda\left(\sum_{i=1}^n x_i + 5\right)\right]$

(c) $\dfrac{\lambda^8 \exp(-5\lambda)}{2^{n+8}}$

(d) $\dfrac{\lambda^{n+3}}{2^{n+6}} \exp\left[-\lambda\left(\sum_{i=1}^n x_i + 2\right)\right]$

(b) The likelihood is given by

$$\left[\frac{\lambda}{2} \exp(-\lambda)\right]^2 \left[\frac{\lambda^2 \exp(-\lambda)}{4}\right]^3 \prod_{i=1}^n \frac{\lambda}{2} \exp(-\lambda x_i) = \frac{\lambda^{n+8}}{2^{n+8}} \exp\left[-\lambda\left(\sum_{i=1}^n x_i + 5\right)\right].$$

6. Consider X distributed according to the Binomial distribution $\mathcal{B}(n, p)$, $n \in \mathbb{N}^*$ **known** and $p \in (0, 1)$ **unknown**. If we denote θ the parameter of the canonical form of this exponential family, $I(p)$ and $I(\theta)$ the Fisher information contained in X for p and θ respectively, we have

(a) $I(p) = n/[p(1-p)]$

(b) $I(p) = 1/[p(1-p)]$

(c) $I(\theta) = ne^{-\theta}\left(1 + e^{\theta}\right)^2$

(d) $I(\theta) = ne^{\theta}/\left(1 + e^{\theta}\right)^2$

(a, d) The density $f(\cdot \mid n, p)$ of the Binomial distribution is twice differentiable with respect to p on $(0, 1)$ and

$$\frac{d}{dp} f(x \mid n, p) = \frac{x}{p} - \frac{n-x}{1-p} \quad \text{and} \quad \frac{d^2}{dp^2} f(x \mid n, p) = -\frac{x}{p^2} - \frac{n-x}{(1-p)^2}.$$

Using that $\mathbb{E}_p[X] = np$, we then have

$$I(p) = -\mathbb{E}_p\left[\frac{d^2}{dp^2} f(X \mid n, p)\right] = \frac{np}{p^2} - \frac{n-np}{(1-p)^2} = \frac{n}{p(1-p)}.$$

The parameter of the canonical form is $\theta = \log(p) - \log(1-p)$, that is $p = \exp(\theta)/[1 + \exp(\theta)] := \psi(\theta)$. We then have

$$I(\theta) = \left(\frac{d}{d\theta}\psi(\theta)\right)^2 I(p) = \frac{\exp(2\theta)}{[1 + \exp(\theta)]^4} \frac{n[1 + \exp(\theta)]^2}{\exp(\theta)} = \frac{n \exp(\theta)}{[1 + \exp(\theta)]^2}.$$

7. Consider a regular and minimal exponential family with natural statistic $T(\cdot)$ and density $f(\cdot \mid \theta)$, $\theta \in \Theta \subseteq \mathbb{R}$. For $X_1, \dots, X_n$ *i.i.d.* random variables distributed according to $f(\cdot \mid \theta)$, we set $S = \sum_{i=1}^n T(X_i)$.

(a) Any bijective transform of S is sufficient for θ.

(b) S is minimal sufficient for θ.

(c) Any sufficient statistic for θ that is a function of S is minimal sufficient for θ.

(d) None of the other answers.

(a, b, c) S is a sufficient statistic for θ. Thus any bijective transform of S is sufficient for θ. For a minimal representation, S is minimal sufficient. Thus S is a function of any other sufficient statistic. Therefore any sufficient statistic R that is a function of S is also a function of any other sufficient statistic. Hence R is also minimal sufficient.

8. Let $X_1, \dots, X_n$ be *i.i.d.* random variables distributed according to the Normal distribution $\mathcal{N}(\mu, 1)$ and denote

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, \quad X_{(1)} = \min(X_1, \dots, X_n) \quad \text{and} \quad X_{(n)} = \max(X_1, \dots, X_n).$$

(a) $X_{(n)}$ - $X_{(1)}$ is independent of $\overline{X}_n$.

(b) $\left(X_{(1)}, X_{(n)}\right)$ is not a complete statistic.

(c) $\left(X_{(1)}, X_{(n)}\right)$ is a sufficient statistic for μ.

(d) $\left(X_1 - \overline{X}_n, \dots, X_n - \overline{X}_n\right)$ is independent of $\overline{X}_n$.

(e) None of the other answers.

(a, b, d) $X_{(n)}$ - $X_{(1)}$ and $(X_1 - \overline{X}_n, \dots, X_n - \overline{X}_n)$ are ancillary statistic. Moreover $\overline{X}_n$ is a complete and sufficient statistic for μ (result on natural statistic associated to an exeponential family). It follows from Basu's theorem that both $X_{(n)}$ - $X_{(1)}$ and $(X_1 - \overline{X}_n, \dots, X_n - \overline{X}_n)$ are independent from $\overline{X}_n$.

$X_{(n)}$ - $X_{(1)}$ is an ancillary statistic that is not constant almost surely and such that $\mathbb{E}_\mu\left[X_{(n)} - X_{(1)}\right] = c < \infty$ is independent of μ. Thus, for the function $\phi : (x, y) \mapsto y - x - c$, we have

$$\mathbb{E}_\mu\left[\phi\left(X_{(1)}, X_{(n)}\right)\right] = 0, \quad \forall \mu \in \mathbb{R}.$$

But $\mathbb{P}_\mu\left[\phi\left(X_{(1)}, X_{(n)}\right) = 0\right] = \mathbb{P}_\mu\left[X_{(n)} - X_{(1)} = c\right] \neq 1$ since $X_{(n)}$ - $X_{(1)}$ is not constant almost surely. Therefore $\left(X_{(1)}, X_{(n)}\right)$ is not a complete statistic.

Exercise 2

...... /15

We first consider a sample $(X_1, \dots, X_n)$, made of $n > 2$ *i.i.d.* random variables distributed according to the Normal distribution $\mathcal{N}(\mu, 1)$ with its mean parameter μ being **unknown**.

1. Express the likelihood function on μ attached with this model. Derive the maximum likelihood estimator of μ, $\hat{\mu}(x_1, \dots, x_n)$, show that it is convergent, asymptotically Normal, and that it meets the Cramèr-Rao lower bound.

This Normal model is a special case of an exponential family distribution, hence the result from the course slides applies, namely that there exists a single maximum likelihood estimator that is solution to $S(x) = \mathbb{E}_\theta[S(X)]$ when using the course slides notations and the natural parameterisation, In the Normal case, this leads to $\hat{\mu}(x_1, \ldots, x_n) = \bar{x}_n$ for which the CLT applies, obviously

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \mathcal{N}(0, 1)$$

In that case, the Fisher information is constant and equal to n for n observations, which is also the inverse of the variance of the MLE.

2. In this question, we assume that αn of the observations in the above sample $(X_1, \ldots, X_n)$, are missing at random, with $0 < \alpha < 1$ and $\alpha n \in \mathbb{N}$. That is, αn of the indices $1, \ldots, n$ are chosen at random and the corresponding x_i's are removed from the sample before its observation.

Express the likelihood function associated with the new observed sample, denoted as $y_1, \ldots, y_{n-\alpha n}$. Derive the associated maximum likelihood estimator of μ, $\tilde{\mu}(y_1, \ldots, y_{n-\alpha n})$, and show that it is convergent and asymptotically Normal.

The random removal of some of the x_i's does not modify the distribution of the remaining variables, which are still iid Normal. Hence the result of Question 1 applies to this subvector with n replaced by $(1 - \alpha)n$.

3. Given $X_1, \ldots, X_n$ iid with density f and cdf F, show that the distribution of the pair $(Z_1, Z_2) = (X_{(i)}, X_{(j)})$ with $1 \leq i < j \leq n$ has density

$$\frac{n!}{(i-1)!(j-i-1)!(n-j)!} F(z_1)^{i-1} f(z_1)[F(z_2) - F(z_1)]^{j-1-i} f(z_2)[1 - F(z_2)]^{n-j}\, \mathbb{I}_{z_1 \leq z_2}$$

where $X_{(i)}$ denotes the ith *order statistic* associated with the sample, that is, the value of the ith term in the ordered sample. The vector of order statistics $(X_{(1)}, \ldots, X_{(n)})$ is thus such that

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}.$$

The classical results are available and well-explained on Wikipedia.

4. Extend the above derivation of the density to the case of a k- order statistic, $\mathbf{W} = (X_{(i_1)}, \ldots, X_{(i_k)})$ when $1 \leq i_1 < \cdots < i_k \leq n$ and deduce that the density of $\mathbf{Y} = (X_{(n-i+1)}, \ldots, X_{(n)})$ is given by

$$\frac{n!}{(i-1)!} F(y_i)^{i-1} \prod_{j=i}^{n} f(y_j)\, \mathbb{I}_{y_i \leq \cdots \leq y_n}.$$

$$\frac{n!}{(i-1)!} F(y_i)^{i-1} \prod_{j=i}^{n} f(y_j)\, \mathbb{I}_{y_i \leq \cdots \leq y_n}$$

5. Starting from the Normal sample $(X_1, \ldots, X_n)$ described in Question 1, we now remove the αn **smallest** observations from that sample, with $0 < \alpha < 1$ and $\alpha n \in \mathbb{N}$. This means that now only $(Y_1, \ldots, Y_{n-\alpha n}) = (X_{(\alpha n+1)}, \ldots, X_{(n)})$ is observed.
Show that the associated likelihood is

$$\frac{n!}{(\alpha n)!} \Phi(x_{(\alpha n+1)} - \mu)^{\alpha n} \prod_{i=\alpha n+1}^{n} \phi(x_{(i)} - \mu)$$

where Φ and ϕ are the standard Normal cdf and pdf, respectively.
Does this likelihood remain associated with an exponential family? If it does, provide the minimal representation and derive the corresponding maximum likelihood estimator of μ, $\breve{\mu}(Y_1, \ldots, Y_{n-\alpha n})$. Else, explain the main difficulty in computing this maximum likelihood estimator.

From Question 3, the joint density of $(X_{(\alpha n+1)}, \ldots, X_{(n)})$ is indeed given by

$$\frac{n!}{\alpha n!} F(x_{(\alpha n+1)})^{\alpha n} \prod_{i=\alpha n+1}^{n} f(x_{(i)})$$

by taking $i = \alpha n$. Therefore in the Normal case the likelihood function involves the Normal cdf, that is,

$$\Phi(x_{(\alpha n+1)} - \mu)^{\alpha n}$$

a feature that takes the distribution outside exponential families and makes computing the maximum likelihood more involved. It is not possible to find a closed-form expression since the score function involves both exponential and linear terms in μ.

6. Show that the likelihood associated with the sample $X_1, \ldots, X_n$ in Question 1 is the same as the likelihood associated with the sample $X_{(1)}, \ldots, X_{(n)}$. What does this imply on the statistic $(X_{(1)}, \ldots, X_{(n)})$?

As shown in class the full order statistic is a *sufficient statistic* for any iid model.

7. Show that the marginal conditional distribution of $X_{(i)}$ $(1 \le i \le \alpha n)$ conditional on $(X_{(\alpha n+1)}, \ldots, X_{(n)})$ only depends on $X_{(\alpha n+1)}$ and prove that it is a Normal distribution restricted to the interval $(-\infty, X_{(\alpha n+1)})$.

Starting from the joint density of $(X_1, \ldots, X_{(n)})$,

$$n! \prod_{i=1}^{n} f(x_{(i)}) \, \mathbb{I}_{x_{(1)} \le \cdots \le x_{(n)}}$$

it is proportional to

$$\prod_{i=1}^{\alpha n} f(x_{(i)}) \, \mathbb{I}_{x_{(1)} \le \cdots \le x_{(\alpha n+1)}}$$

when fixing $(X_{\alpha n+1}, \ldots, X_{(n)})$. Hence, it only depends on $x_{(\alpha n+1)}$ and by recursive integration over the other variates one can deduce that the marginal conditional density in $X_{(i)}$ is proportional to $f(x_{(i)}) \, \mathbb{I}_{x_{(i)} \le x_{(\alpha n+1)}}$ which is the Normal distribution restricted to the interval $(-\infty, X_{(\alpha n+1)})$.

8. Still based on the iid sample from Question 1, consider the random vectors

$$\mathbf{X} = (X_{(1)}, \dots, X_{(n)}),\ \mathbf{Z} = (X_{(1)}, \dots, X_{(\alpha n)}),\ \mathbf{Y} = (X_{(\alpha n+1)}, \dots, X_{(n)})$$

Denote the likelihood functions associated with $\mathbf{X}$ and $\mathbf{Y}$ as $L^c(\mu|\mathbf{X})$ and $L^o(\mu|\mathbf{Y})$, respectively. Define

$$\log k_\mu(\mathbf{Z}|\mathbf{Y}) = \log L^c(\mu|\mathbf{X}) - \log L^o(\mu|\mathbf{Y}) \qquad (1)$$

and show that $k_\mu(\mathbf{z}|\mathbf{y})$ is a probability density in z for all values of y. In the Normal setting of Question 1, what is the exact expression of $k_\mu(\mathbf{z}|\mathbf{y})$?

This is simply the decomposition of the density of $\mathbf{X}$ as the product of the (marginal) density of $\mathbf{Y}$ times the conditional density of $\mathbf{Z}$ given $\mathbf{Y}$. The marginal density of $\mathbf{Y}$ was produced earlier. As for the conditional density of the αn lowest order statistics $\mathbf{Z}$ given the $(n - \alpha n)$ largest order statistics $\mathbf{Y}$,

$$k_\mu(\mathbf{z}|\mathbf{y}) = (\alpha n)!\, \Phi(y_1 - \mu)^{-\alpha n} \prod_{i=1}^{\alpha n} \varphi(z_i - \mu) \mathbb{I}_{z_1 \le \cdots \le z_{\alpha n} \le y_1}$$

9. Show that, for an arbitrary value $\mu^0 \in \mathbb{R}$,

$$\log L^o(\mu|\mathbf{Y}) = \mathbb{E}_{\mu^0}[\log L^c(\mu|\mathbf{X}) - \log k_\mu(\mathbf{Z}|\mathbf{Y}) \big| \mathbf{Y}] \quad \forall \mu \in \mathbb{R} \qquad (2)$$

where the expectation on the right hand side is taken for the conditional distributions of $\mathbf{X}$ and $\mathbf{Z}$, respectively, given $\mathbf{Y}$, when the parameter value is μ^0, i.e., when associated with iid $X_i \sim \mathcal{N}(\mu_0, 1)$ $(1 \le i \le n)$.

Since the left hand side of

$$\log L^o(\mu|\mathbf{Y}) = \log L^c(\mu|\mathbf{X}) - \log k_\mu(\mathbf{Z}|\mathbf{Y})$$

derived from (1) only depends on $\mathbf{Y}$, the same is true with its right hand side. The right hand side does not depend on Z, which means that it can be integrated in $X(Z)$ and Z conditional on Y, whatever the conditional distribution used for the integration.

10. Show that $\mathbb{E}_{\mu^0}[\log L^c(\mu|\mathbf{X})|\mathbf{Y} = (y_1, \dots, y_{n-\alpha n})]$ can be written as

$$-{}^{n}\!/_{2} \log(2\pi) - {}^{\alpha n}\!/_{2} \int_{-\infty}^{y_1} (x-\mu)^2 \frac{e^{-(x-\mu^0)^2/2}}{\sqrt{2\pi}\Phi(y_1 - \mu^0)}\, dx - {}^{1}\!/_{2} \sum_{i=1}^{n-\alpha n} (y_i - \mu)^2$$

where $\Phi(x)$ is the standard Normal cdf.

Since

$$\log L^c(\mu|\mathbf{X}) = -{}^{n}\!/_{2} \log(2\pi) - {}^{1}\!/_{2} \sum_{i=1}^{n} (X_{(i)} - \mu)^2$$

and

$$\sum_{i=1}^{n} (X_{(i)} - \mu)^2 = \sum_{i=1}^{\alpha n} (X_{(i)} - \mu)^2 + \underbrace{\sum_{i=\alpha n}^{n} (X_{(i)} - \mu)^2}_{\sum_{i=1}^{n-\alpha n} (Y_i - \mu)^2}$$

the second sum term is fixed when $\mathbf{Y}$ is fixed and the first sum term is made of iid terms by Question 6, since all $X_{(i)}$ are conditionaly Normal $\mathcal{N}(\mu^0, 1)$ truncated to $(-\infty, y_1)$.

11. Show that

$$\int_{-\infty}^{y_1} x \frac{e^{-(x-\mu^0)^2/2}}{\sqrt{2\pi}\Phi(y_1-\mu^0)}\,dx = \mu^0 - \frac{e^{-(y_1-\mu^0)^2/2}}{\sqrt{2\pi}\Phi(y_1-\mu^0)}$$

and deduce that the argument of

$$\max_{\mu} \mathbb{E}_{\mu^0}[\log L^c(\mu|\mathbf{X})|\mathbf{Y} = (y_1,\dots,y_{n-\alpha n})]$$

is

$$\mu^1 = \frac{\alpha n}{n}\left\{\mu^0 - \frac{e^{-(y_1-\mu^0)^2/2}}{\sqrt{2\pi}\Phi(y_1-\mu^0)}\right\} + \frac{1}{n}\sum_{i=1}^{n-\alpha n} y_i \qquad (3)$$

The integral can be solved by observing that

$$xe^{-(x-\mu^0)^2/2} = -\frac{d}{dx}e^{-(x-\mu^0)^2/2} + \mu^0 e^{-(x-\mu^0)^2/2}$$

By further taking the derivative in μ of the expression in Question 9,

$$\mu^1 = \frac{\alpha n}{n}\mathbb{E}[X_{(1)}|Y_1 = y_1] + \frac{1}{n}\sum_{i=1}^{n-\alpha n} y_i$$

12. (Bonus question) Show that, when μ^1 is defined in (3)

$$L^o(\mu^1|\mathbf{Y}) \geq L^o(\mu^0|\mathbf{Y})$$

by using Jensen inequality.

This is the argument behind the EM algorithm, as explained in the slides of the course.

Exercise 3

...... /6

1. We consider a sample $(X_1,\dots,X_n)$, made of n i.i.d random variables distributed according to a Laplace distribution Laplace (θ, b) with $\theta \in \mathbb{R}$ unknown and $b \in \mathbb{R}_+^*$ fixed. The objective is to construct an estimator of θ.

The density of a Laplace (θ, b) against the Lebesgue measure on $\mathbb{R}$ is defined as

$$f(x \mid \theta) = \frac{1}{2b}\exp\left(-\frac{|x-\theta|}{b}\right), \quad \text{for} \quad x \in \mathbb{R}.$$

For this density, $\mathbb{E}[X_1] = \theta$ and $\mathbb{V}[X_1] = 2b^2$.

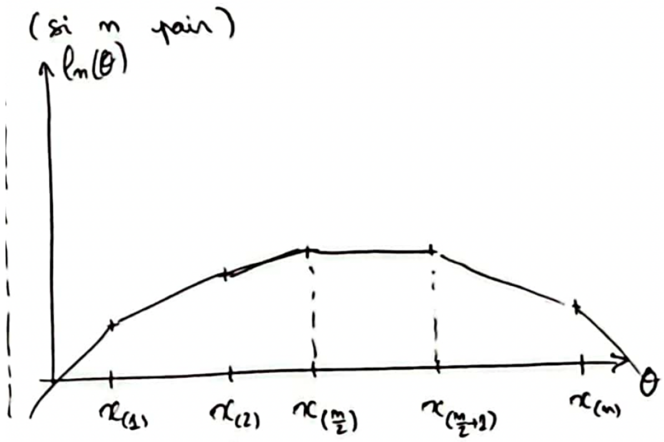
Derive the Fisher information of θ using the following formula $\mathcal{I}_{X_1}(\theta) = \mathbb{E}_\theta\left[\left(\frac{\partial \log f(X_1 \mid \theta)}{\partial\theta}\right)^2\right]$.

Given that $\frac{\partial \log(f(x \mid \theta))}{\partial\theta} = \frac{1}{b}\mathbb{1}_{x-\theta>0} + \frac{1}{b}\mathbb{1}_{x-\theta<0}$, $\mathcal{I}_{X_1}(\theta) = \frac{1}{b^2}\mathbb{E}\left[\mathbb{1}_{X-\theta\neq 0}\right] = \frac{1}{b^2}$.

2. Give the expression of the log-likelihood $l_n(\theta)$ of the n-sample.

$l_n(\theta) = -n\log(2b) - \sum_{i=1}^n |x_i - \theta|$.

3. Consider the case when n is even (i.e., $n = 2k$ for $k \in \mathbb{N}$). Explain or illustrate with a drawing why a MLE of θ is $\hat{\theta}_{\text{MLE}} = X_{(\frac{n}{2})}$.



4. Let $0 < \alpha < 1$. For n even and large enough, give two confidence intervals of asymptotic level $1 - \alpha$, one based on the estimator $\bar{X}_n$ and the other on $\hat{\theta}_{\text{MLE}}$.

The sequence$(X_1, \dots, X_n)$ is a i.i.d sample of finite variance. The CLT states :

$$\sqrt{n}\left(\bar{X}_n - \theta\right) \xrightarrow[n]{\mathcal{L}} \mathcal{N}\left(0, 2b^2\right).$$

We deduce the confidence interval of level $1-\alpha$: $IC_1(\alpha) = \left[\bar{X}_n - \frac{q_{1-\frac{\alpha}{2}}(\sqrt{2}b)}{\sqrt{n}}, \bar{X}_n + \frac{q_{1-\frac{\alpha}{2}}(\sqrt{2}b)}{\sqrt{n}}\right]$ where $q_{1-\frac{\alpha}{2}}$ is the quantile of level $1 - \frac{\alpha}{2}$ of the law $\mathcal{N}(0,1)$.
From the asymptotic normality of the MLE, $\sqrt{n}\left(\hat{\theta}_{\text{MLE}} - \theta\right) \xrightarrow[n]{\mathcal{L}} \mathcal{N}\left(0, \mathcal{I}_{X_1}(\theta)^{-1}\right) \equiv \mathcal{N}\left(0, b^2\right)$.

We deduce a second confidence interval for θ, $IC_2(\alpha) = \left[\hat{\theta}_{\text{MLE}} - \frac{q_{1-\frac{\alpha}{2}} b}{\sqrt{n}}, \hat{\theta}_{\text{MLE}} + \frac{q_{1-\frac{\alpha}{2}} b}{\sqrt{n}}\right]$.

5. For n large enough and even, which confidence interval would you prefer and why ?

The confidence interval based on $\hat{\theta}_{\text{MLE}}$ is smaller thus one should select IC_2.

6. Write an R code to compute the MLE given a n-sample $x = (x_1, \dots, x_n)$.

```
compute_theta = function(x){
x_sort = sort(x)
return(x_sort[n/2])
}
```

7. Using parametric bootstrap with $B = 1000$ bootstrap samples and the MLE as estimate of θ, write an R code to compute a 66% empirical bootstrap confidence interval. (The entries `alpha`, `b` and `n` are known and already defined in the R environment.)

```
theta_hat = compute_theta(x)
B = 1000
theta_star = rep(0, B)
for (i in seq_len(B)) {
theta_star[i] = compute_theta(rlaplace(n, theta_hat, b))
```

```
}
#lower bound
alpha = 1 - 0.66
b_low  = 2*theta_hat - quantile(theta_star, 1 - alpha/2)
#upper bound
b_up = 2*theta_hat - quantile(theta_star, alpha/2)
```