$\circled{1}$

Exercice 1

Partie 1

1) La vraisemblance de $(X_1, \dots, X_n)$ est définie par

$$L(x_1, \dots, x_n ; b) = \prod_{i=1}^{n} \left[\frac{1}{2b} \exp\left(- \frac{|x_i - m|}{b} \right) \right]$$

$$= \frac{1}{(2b)^n} \exp\left[- \frac{1}{b} \sum_{i=1}^{n} |x_i - m| \right]$$

$$\frac{\partial \ln L}{\partial b} \geq 0 \iff -\frac{n}{b} + \frac{1}{b^2} \sum_{i=1}^{n} |x_i - m| \geq 0$$

$$\iff b \leq \frac{1}{n} \sum_{i=1}^{n} |x_i - m|$$

ce qui donne le tableau de variations suivant

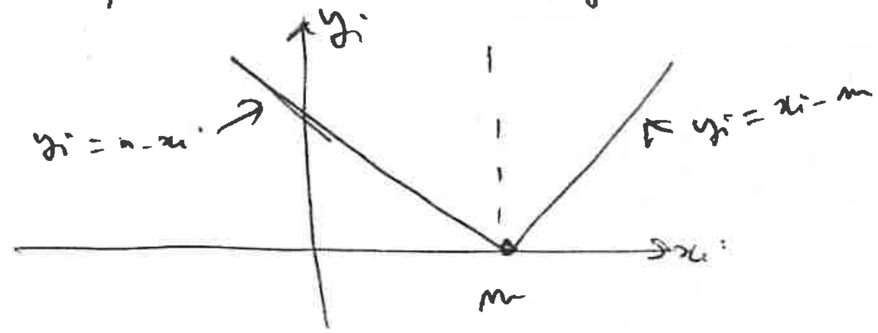
b	0		$\frac{1}{n}\sum_{i=1}^{n} \|x_i - m\|$		$+\infty$
$\frac{\partial \ln L}{\partial b}$		$+$	0	$-$	
L		$\nearrow$		$\searrow$	

2pts

On en déduit que l'estimateur du maximum de vraisemblance du paramètre b est

$$\boxed{\hat{b}_{MV} = \frac{1}{n} \sum_{i=1}^{n} |x_i - m|}$$

2) le changement de variables $Y_i = |X_i - m|$ n'est pas bijectif comme on peut l'observer sur la figure ci-dessous



Ce changement de variables définit deux bijections

$\boxed{1^{ère} \text{ bijection}}$

$$g_1 \quad]m, +\infty[\longrightarrow \mathbb{R}^+$$
$$x_i \longmapsto y_i = |x_i - m| = x_i - m$$

Jacobien $\quad J_1 = \left|\frac{dx_i}{dy_i}\right| = 1$

densité $\quad \pi_1(y_i) = \frac{1}{2b} \exp\left(-\frac{1}{b} y_i\right) \mathbb{1}_{\mathbb{R}^+}(y)$

$\boxed{2^{ème} \text{ bijection}}$

$$g_2 \quad]-\infty, m[\longrightarrow \mathbb{R}^+$$
$$x_i \longmapsto y_i = |x_i - m| = m - x_i$$

2 pts

Jacobien $\quad J_2 = \left|\frac{dx_i}{dy_i}\right| = |-1| = 1$

$$\pi_2(y_i) = \frac{1}{2b} \exp\left(-\frac{1}{b} y_i\right) \mathbb{1}_{\mathbb{R}^+}(y)$$

La densité de Y_i est donc

$$\pi(y_i) = \pi_1(y_i) + \pi_2(y_i) = \boxed{\frac{1}{b} \exp\left(-\frac{y_i}{b}\right) \mathbb{1}_{\mathbb{R}^+}(y_i)}$$

Soit $\boxed{Y_i \sim \Gamma\left(1, \frac{1}{b}\right)}$

D'après les tables $E[Y_i] = b$ et $Var[Y_i] = b^2$ donc

$$E[\hat{b}_{MV}] = \frac{1}{n} \sum_{i=1}^{n} \underbrace{E(Y_i)}_{b} = \boxed{b}$$ l'estimateur $\hat{b}_{MV}$ est non biaisé

$$Var[\hat{b}_{MV}] = \frac{1}{n^2} \sum_{i=1}^{n} \underbrace{Var[Y_i]}_{b^2} = \boxed{\frac{b^2}{n}}$$

$\hat{b}_{MV}$ non biaisé

$Var[\hat{b}_{MV}] \xrightarrow[n \to +\infty]{} 0$

$\Rightarrow$ $\boxed{\hat{b}_{MV} \text{ est un estimateur convergent de } b}$

3) $$\frac{\partial \ln L}{\partial b} = -\frac{n}{b} + \frac{1}{b^2} \sum_{i=1}^{n} |x_i - m|$$

donc $$\frac{\partial^2 \ln L}{\partial b^2} = \frac{n}{b^2} - \frac{2}{b^3} \sum_{i=1}^{n} |x_i - m|$$

donc $E\left[\frac{\partial^2 \ln L}{\partial b^2}\right] = \frac{n}{b^2} - \frac{2n}{b^3} \underbrace{E[\,|x_i - m|\,]}_{E(y_i) = b} = \boxed{-\frac{n}{b^2}}$

On en déduit que la borne de Cramér Rao d'un estimateur non biaisé de b est

$$\boxed{BCR = \frac{b^2}{n}}$$

1pt

$\hat{b}_{MV}$ est donc l'estimateur efficace du paramètre b.

4) Quand $m = 0$, la vraisemblance s'écrit

$$f(x_1, \dots, x_n; b) = \frac{1}{(2b)^n} \exp\left(-\frac{1}{b}\sum_{i=1}^{n} |x_i|\right)$$

- Donc la loi a posteriori de $b | x_1 \dots x_n$ vérifie

$$p(b | x_1, \dots, x_n) \propto p(x_1, \dots, x_n | b)\, p(b)$$

$$\propto \frac{1}{(2b)^n} \exp\left(-\frac{1}{b}\sum_{i=1}^{n}|x_i|\right) \frac{1}{b^{\alpha+1}} \exp\left(-\frac{\beta}{b}\right)$$

$$\propto \frac{1}{b^{n+\alpha+1}} \exp\left[-\frac{1}{b}\left(\beta + \sum_{i=1}^{n}|x_i|\right)\right]$$

on reconnait une loi inverse gamma de paramètres $a = n + \alpha$ et $b = \beta + \sum |x_i|$

donc $\boxed{b | x_1, \dots, x_n \sim IG\left(n+\alpha, \beta + \sum_{i=1}^{n}|x_i|\right)}$

1pt

- l'estimateur MAP maximise $\frac{1}{b^{n+\alpha+1}} \exp\left[-\frac{1}{b}\left(\beta + \sum_{i=1}^{n}|x_i|\right)\right]$

qui a la même forme que la vraisemblance avec les changements

$$n \longrightarrow n+\alpha+1 \quad \text{et} \quad \sum_{i=1}^{n}|x_i| \longrightarrow \beta + \sum_{i=1}^{n}|x_i|$$

On en déduit $\boxed{\hat{b}_{MAP} = \frac{1}{n+\alpha+1}\left[\beta + \sum_{i=1}^{n}|x_i|\right]}$

1pt

On remarque que $\hat{b}_{MAP}$ s'écrit

$$\hat{b}_{MAP} = \frac{\alpha+1}{n+\alpha+1}\;\frac{\beta}{\alpha+1} + \frac{n}{n+\alpha+1}\;\underbrace{\frac{1}{n}\sum_{i=1}^{n}|x_i|}_{\hat{b}_{MV}}$$

d'où $\boxed{b_\alpha(n) = \frac{n}{n+\alpha+1}}$

On voit que $b_\alpha(n) \xrightarrow[n\to+\infty]{} 1$; l'estimateur MAP se comporte asymptotiquement comme l'estimateur du maximum de vraisemblance

$b_\alpha(n) \xrightarrow[n\to 0]{} 0$ ce qui signifie $\hat{b}_{MAP} \underset{n\to 0}{\sim} \frac{\beta}{\alpha+1}$

qui est le mode (la valeur la plus probable) de la loi a priori

$\boxed{\text{Exercice 2}}$

1) Le test de Neyman - Pearson rejette l'hypothèse H_0 si

$$\frac{f(x_1, \dots, x_n \mid H_1)}{f(x_1 \dots x_n \mid H_0)} > \text{seuil} \iff \frac{\frac{1}{(2b_1)^n} \exp\left[-\frac{1}{b_1}\sum_{i=1}^{n} |x_i - m|\right]}{\frac{1}{(2b_0)^n} \exp\left[-\frac{1}{b_0}\sum_{i=1}^{n} |x_i - m|\right]} > S$$

D'où le test équivalent

$$\text{Rejet de } H_0 \text{ si } \underbrace{\left(\frac{1}{b_0} - \frac{1}{b_1}\right)}_{<0 \text{ car } b_1 < b_0} \sum_{i=1}^{n} |x_i - m| > S$$

1pr

Soit

$$\boxed{\text{Rejet de } H_0 \text{ si } T_n = \sum_{i=1}^{n} |x_i - m| < S_\alpha}$$

2) Pour déterminer la loi de $T_n = \sum_{i=1}^{n} Y_i$, on peut déterminer sa fonction caractéristique

$$E[e^{itT_n}] = E[e^{it\sum_{k=1}^{n} Y_k}]$$

$$= \prod_{k=1}^{n} E[e^{itY_k}]$$

les va Y_k sont indépendants

La loi de Y_k est une loi exponentielle, c'est-à-dire une loi Gamma $\Gamma\left(1, \frac{1}{b}\right)$ dont la fonction caractéristique est donnée dans la table

$$E[e^{itY_k}] = \frac{1}{1 - itb}$$

1pr

On en conclut $E[e^{itT_n}] = \dfrac{1}{(1 - itb)^n}$ d'où $\boxed{T_n \sim \Gamma\left(n, \frac{1}{b}\right)}$

Donc
$$\boxed{\begin{array}{l} \text{sous } H_0, \quad T_n \sim \Gamma(n, \frac{1}{b_0}) \\ \text{sous } H_1, \quad T_n \sim \Gamma(n, \frac{1}{b_1}) \end{array}}$$

Le théorème de la limite centrale nous dit que $\dfrac{T_n - nb}{\sqrt{nb^2}} \xrightarrow[n\to+\infty]{\mathcal{L}} N(0,1)$

avec $nb = E[T_n]$ et $nb^2 = Var[T_n]$.

Donc
$$\boxed{\begin{array}{l} \text{sous } H_0, \quad T_n \approx N(nb_0, nb_0^2) \\ \text{sous } H_1, \quad T_n \approx N(nb_1, nb_1^2) \end{array}}$$

1pt

où $\approx$ signifie que la loi est approchée (et valide lorsque n est "grand").

3) Le risque α est défini par

$$\alpha = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$$
$$= P[T_n < s \mid b = b_0]$$
$$= P\left[\frac{T_n - nb_0}{\sqrt{nb_0^2}} < \frac{s - nb_0}{\sqrt{nb_0^2}} \;\middle|\; \frac{T_n - nb_0}{\sqrt{nb_0^2}} \sim N(0,1)\right]$$

donc
$$\boxed{\alpha = F\left[\frac{s - nb_0}{\sqrt{nb_0^2}}\right]}$$

2pts

De même $\beta = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}]$
$$= P[T_n \geq s \mid b = b_1]$$
$$\boxed{\beta = 1 - F\left[\frac{s - nb_1}{\sqrt{nb_1^2}}\right]}$$

4) Les courbes COR expriment $\pi = 1-\beta$ en fonction de α.

On a

$$\pi = 1-\beta = F\left[\frac{s - nb_1}{\sqrt{nb_1^2}}\right] \quad \text{avec} \quad s = F^{-1}(\alpha)\sqrt{nb_0^2} + nb_0$$

donc

$$\pi = F\left[F^{-1}(\alpha)\,\frac{b_0}{b_1} + \frac{nb_0 - nb_1}{\sqrt{nb_1^2}} \right]$$

soit

$$\boxed{\pi = F\left[F^{-1}(\alpha)\,\frac{b_0}{b_1} + \sqrt{n}\,\frac{b_0 - b_1}{b_1} \right]}$$

On a $\frac{b_0}{b_1} > 1$. Quand $\frac{b_0}{b_1} \nearrow$, $F^{-1}(\alpha)\frac{b_0}{b_1}$ et $\sqrt{n}\left(\frac{b_0}{b_1} - 1\right)$ augmentent tous les deux donc la puissance du test augmente, ce qui est naturel

Quand $n \nearrow$, comme $b_0 > b_1$, on a $\frac{b_0 - b_1}{b_1} > 0$ donc $\pi \nearrow$ donc la puissance du test augmente, ce qui est naturel

Ex3

1) Pour centrer et réduire les données, il suffit de construire les observations

$$y_i = \frac{x_i - \bar{x}}{S_n}$$

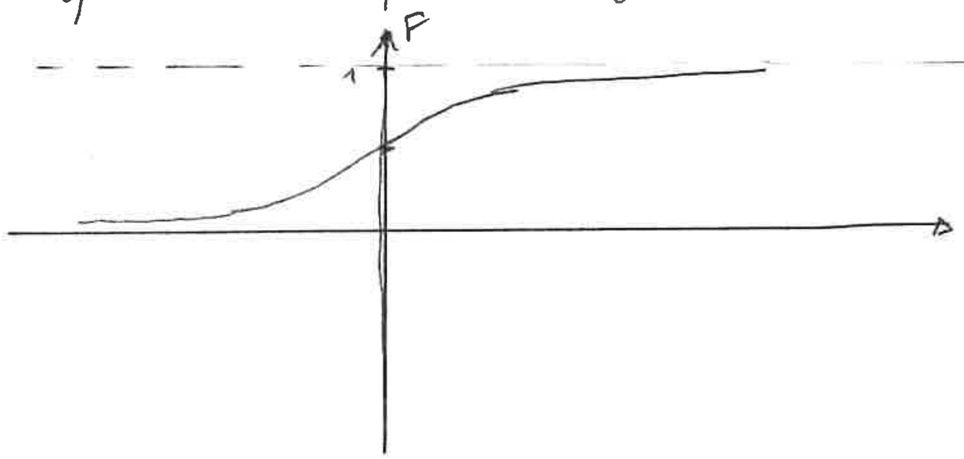
1pt

avec $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ et $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ (estimateur non biaisé et convergent de la variance des x_i)

2) La fonction de répartition d'une loi de Laplace est

$$F_0(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2}} e^{-\sqrt{2}|u|} du = \begin{cases} \frac{1}{2} e^{\sqrt{2}x} & x<0 \\ 1 - \frac{1}{2} e^{-\sqrt{2}x} & x \geq 0 \end{cases}$$

qui admet la représentation graphique suivante



3) Comme vu en cours, la statistique du test de Kolmogorov est

$$D_n = \max_{i \in \{1,\dots,n\}} \max\{E_i^+, E_i^-\}$$

avec $E_i^+ = \left| \frac{i}{n} - F_0(x_i) \right|$ et $E_i^- = \left| \frac{i-1}{n} - F_0(x_i) \right|$

La région critique du test est définie par

Rejet de H_0 si $D_n > S_\alpha$

4) on a $\alpha = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$

$= P[D_n > S_\alpha \mid D_n \sim \text{loi de Kolmogorov}]$

$= 1 - G_n[S_\alpha]$ où G_n est la fonction de répartition de la loi de Kolmogorov

d'où $\boxed{S_\alpha = G_n^{-1}(1-\alpha)}$