Sorbonne Université
Master Mathématiques et Applications
MU4MA015 : Statistique

Janvier 2023
Aucun document autorisé
Durée : 3 heures

Correction Examen

Exercice 1
Soit $\theta > 0$ un paramètre inconnu et X de densité

$$f_\theta(x) = \frac{1}{\theta}(1-x)^{\frac{1}{\theta}-1}\mathbf{1}_{]0,1[}(x).$$

Dans la suite, on suppose disposer d'un échantillon $(X_1, \ldots, X_n)$ i.i.d. de même loi que X.

1. La fonction de répartition vaut

$$F_\theta(x) = \left(1-(1-x)^{\frac{1}{\theta}}\right)\mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,\infty[}(x).$$

   La médiane $x_{1/2}$ de la loi de X vérifie $F_\theta(x_{1/2}) = 1/2$, ce qui donne $x_{1/2} = 1 - 2^{-\theta} = \varphi(\theta)$ avec φ qui est un C^1-difféomorphisme de $]0, \infty[$ vers $]0, 1[$.

2. Un estimateur $\hat{\theta}_n$ de θ est donc $\hat{\theta}_n = \varphi^{-1}(x_{1/2}(n)) = -\frac{\log(1-x_{1/2}(n))}{\log 2}$ où $x_{1/2}(n)$ désigne la médiane empirique. Puisque $f_\theta(x_{1/2}) > 0$, on sait que

$$\sqrt{n}\left(x_{1/2}(n) - x_{1/2}\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \frac{1}{4f_\theta(x_{1/2})^2}\right),$$

   ce qui revient à dire que

$$\sqrt{n}\left(\varphi(\hat{\theta}_n) - \varphi(\theta)\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \left(\frac{\theta}{2^\theta}\right)^2\right),$$

   Il reste à appliquer la méthode Delta pour obtenir finalement

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \left(\frac{\theta}{\log 2}\right)^2\right).$$

3. En notant $q := \Phi^{-1}(1-\alpha/2)$ le quantile d'ordre $(1-\alpha/2)$ de la gaussienne centrée réduite, on en déduit directement un intervalle de confiance de niveau asymptotique $(1-\alpha)$ pour θ :

$$\left[\frac{\hat{\theta}_n}{1+\frac{q}{\log 2\sqrt{n}}} \,;\, \frac{\hat{\theta}_n}{1-\frac{q}{\log 2\sqrt{n}}}\right].$$

4. La log-vraisemblance s'écrit, pour tout $\theta > 0$,

$$\ell_n(\theta) = -n\log\theta + \left(\frac{1}{\theta} - 1\right)\sum_{i=1}^n \log(1-X_i),$$

donc
$$\ell_n'(\theta) = -\frac{n}{\theta} - \frac{\sum_{i=1}^n \log(1-X_i)}{\theta^2}.$$
Ainsi $\ell_n'(\theta) > 0$ ssi $\theta < -\frac{1}{n}\sum_{i=1}^n \log(1-X_i)$ et $\ell_n'(\theta) < 0$ ssi $\theta > -\frac{1}{n}\sum_{i=1}^n \log(1-X_i)$, ce qui prouve que l'estimateur du maximum de vraisemblance est
$$\tilde{\theta}_n = -\frac{1}{n}\sum_{i=1}^n \log(1-X_i).$$

5. Puisque $0 < X < 1$, la variable $T = -\log(1-X)$ est à valeurs dans $\mathbb{R}_+^*$. Soit donc $h : \mathbb{R}_+^* \to \mathbb{R}$ mesurable bornée, alors le Théorème de Transfert et le changement de variable $t = -\log(1-x)$ donnent
$$\mathbb{E}[h(T)] = \mathbb{E}[h(-\log(1-X))] = \int_0^1 h(-\log(1-x))\frac{1}{\theta}(1-x)^{\frac{1}{\theta}-1}dx = \int_0^\infty h(t)\frac{1}{\theta}e^{-\frac{t}{\theta}}dt,$$
ce qui montre que T suit une loi exponentielle de paramètre $1/\theta$.

6. Puisque $\mathrm{Var}(T) = \theta^2$ et $\tilde{\theta}_n = \bar{T}_n$, le TCL assure que
$$\sqrt{n}\left(\tilde{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \theta^2\right).$$
Il est clair que pour tout $\theta > 0$, on a $\theta^2 < (\theta/\log 2)^2$, donc d'un point de vue asymptotique, entre $\hat{\theta}_n$ et $\tilde{\theta}_n$, on choisira $\tilde{\theta}_n$.

7. Pour tout $1 \leq i \leq n$, puisque T_i suit une loi exponentielle de paramètre $1/\theta$, les variables $T_i/(n\theta)$ suivent une loi $\mathcal{E}(n) = \Gamma(1,n)$. De plus, ces variables sont indépendantes (puisque les X_i le sont), donc leur somme $\tilde{\theta}_n/\theta$ suit une loi $\Gamma(n,n)$.

   Pour décider entre $H_0 : \theta \leq 1$ et $H_1 : \theta > 1$, on cherche un test de la forme $T(\mathbf{X}) = 1$ ssi $\tilde{\theta}_n > c_\alpha$ avec $c_\alpha > 0$ tel que
$$\sup_{0<\theta\leq 1} \mathbb{P}_\theta(T(\mathbf{X}) = 1) = \alpha.$$
   Autrement dit
$$\sup_{0<\theta\leq 1} \mathbb{P}_\theta(\tilde{\theta}_n > c_\alpha) = \alpha \iff \sup_{0<\theta\leq 1} \mathbb{P}_\theta\left(\frac{\tilde{\theta}_n}{\theta} > \frac{c_\alpha}{\theta}\right) = \alpha.$$
   Soit F la fonction de répartition de la loi $\Gamma(n,n)$, celle-ci est strictement croissante et bijective de $[0,\infty[$ vers $[0,1[$ et si on note $q_{1-\alpha}$ son quantile d'ordre $(1-\alpha)$, alors on cherche c_α tel que
$$\sup_{0<\theta\leq 1} 1 - F\left(\frac{c_\alpha}{\theta}\right) = \alpha \iff F(c_\alpha) = 1-\alpha \iff c_\alpha = q_{1-\alpha}.$$
   Ainsi, le test consistant à rejeter H_0 ssi $\tilde{\theta}_n > q_{1-\alpha}$ est de niveau (et même de taille) non asymptotique α.

Exercice 2

On suppose disposer de deux ensembles $T_1, \ldots, T_\ell$ et $U_1, \ldots, U_m$ de variables aléatoires réelles.

On note $T = [T_1, \dots, T_\ell]' \in \mathbb{R}^\ell$ et $U = [U_1, \dots, U_m]' \in \mathbb{R}^m$ les vecteurs correspondants. On suppose que les vecteurs T et U sont observés et donnés par

$$T = V\beta + \sigma\eta \qquad \text{et} \qquad U = W\gamma + \tau\varepsilon$$

où β, $\gamma \in \mathbb{R}^p$ sont inconnus, V et W sont des matrices connues de tailles $\ell \times p$ et $m \times p$ et de rang p (avec $\ell > p$ et $m > p$), σ et τ sont des réels positifs inconnus, et η, ε sont deux vecteurs gaussiens standards indépendants non observés et de dimensions respectives ℓ et m. On désigne par $\hat{\beta}$, $\hat{\gamma}$, $\hat{\sigma}^2$ et $\hat{\tau}^2$ les estimateurs usuels des moindres carrés de β, γ, σ^2 et τ^2 respectivement.

1. D'après le cours, on sait que $(\ell - p)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{\ell-p}$, donc son espérance et sa variance valent respectivement $(\ell - p)$ et $2(\ell - p)$. De même, $(m - p)\hat{\tau}^2/\tau^2 \sim \chi^2_{m-p}$. De plus, les variables $\hat{\sigma}^2$ et $\hat{\tau}^2$ sont indépendantes puisque construites à partir des vecteurs η et ε qui le sont. On en déduit que $(\tau^2\hat{\sigma}^2)/(\sigma^2\hat{\tau}^2) \sim \mathcal{F}^{\ell-p}_{m-p}$.

2. Sous l'hypothèse $\tau = \sigma$, on a donc
$$\frac{\tau^2\hat{\sigma}^2}{\sigma^2\hat{\tau}^2} = \frac{\hat{\sigma}^2}{\hat{\tau}^2} =\sim \mathcal{F}^{\ell-p}_{m-p}.$$
Le lien entre intervalle de confiance et test implique que, si on note $q_{\alpha/2}$ et $q_{1-\alpha/2}$ les quantiles d'ordre $\alpha/2$ et $(1 - \alpha/2)$ de la loi $\mathcal{F}^{\ell-p}_{m-p}$, alors le test consistant à rejeter H_0 ssi $\frac{\hat{\sigma}^2}{\hat{\tau}^2} \notin [q_{\alpha/2}; q_{1-\alpha/2}]$ est de niveau α.

3. D'après ce qui a été dit ci-dessus, $(\ell - p)\frac{\hat{\sigma}^2}{\sigma^2} + (m - p)\frac{\hat{\tau}^2}{\tau^2} \sim \chi^2_{\ell+m-2p}$.

4. Introduisons
$$\tilde{\sigma}^2 := \frac{\ell - p}{\ell + m - 2p}\hat{\sigma}^2 + \frac{m - p}{\ell + m - 2p}\hat{\tau}^2.$$
Si $\tau = \sigma$, alors la question précédente montre que
$$\frac{\tilde{\sigma}^2}{\sigma^2} \sim \frac{\chi^2_{\ell+m-2p}}{\ell + m - 2p}.$$
En particulier, $\tilde{\sigma}^2$ (tout comme $\hat{\sigma}^2$) estime sans biais σ^2 et
$$R(\tilde{\sigma}^2, \sigma^2) = \mathrm{Var}(\tilde{\sigma}^2) = \frac{2\sigma^4}{\ell + m - 2p} < \frac{2\sigma^4}{\ell - p} = \mathrm{Var}(\hat{\sigma}^2) = R(\hat{\sigma}^2, \sigma^2).$$

5. On sait que $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2(V'V)^{-1})$, $\hat{\gamma} \sim \mathcal{N}(\gamma, \sigma^2(W'W)^{-1})$ (car $\tau = \sigma$). De plus, les vecteurs $\hat{\beta}$ et $\hat{\gamma}$ sont indépendants puisque construits à partir des vecteurs η et ε qui le sont. Par conséquent,
$$\hat{\beta} - \hat{\gamma} \sim \mathcal{N}\left(\beta - \gamma, \sigma^2((V'V)^{-1} + (W'W)^{-1})\right).$$

6. On sait que
$$(X - m)'\Gamma^{-1}(X - m) \sim \chi^2_d.$$

7. Les matrices $(V'V)^{-1}$ et $(W'W)^{-1}$ étant toutes deux symétriques définies positives, leur somme l'est aussi : on peut donc l'écrire comme l'inverse d'une matrice symétrique définie positive et poser

$$\Gamma := \left((V'V)^{-1} + (W'W)^{-1}\right)^{-1},$$

de sorte que $\hat{\beta} - \hat{\gamma} \sim \mathcal{N}\left(\beta\gamma, \sigma^2\Gamma^{-1}\right)$. Sous l'hypothèse $\beta = \gamma$, il vient donc

$$\frac{1}{p\sigma^2}(\hat{\beta} - \hat{\gamma})'\Gamma(\hat{\beta} - \hat{\gamma}) \sim \frac{\chi_p^2}{p}.$$

Pour estimer σ^2, on utilise $\tilde{\sigma}^2$. Cette variable est indépendante de $(\hat{\beta} - \hat{\gamma})$ car c'est une combinaison des variables $\hat{\sigma}^2$ et $\hat{\tau}^2$ qui sont toutes deux indépendantes de $\hat{\beta}$ et de $\hat{\gamma}$. Nous avons donc

$$S := \frac{1}{p\tilde{\sigma}^2}(\hat{\beta} - \hat{\gamma})'\Gamma(\hat{\beta} - \hat{\gamma}) \sim \mathcal{F}^p_{l+m-2p}.$$

Ainsi, en notant q le quantile d'ordre $(1-\alpha)$ de la loi $\mathcal{F}^p_{l+m-2p}$, le test consistant à rejeter H_0 ssi $S > q$ est de niveau α.

Exercice 3

1. La densité est nulle sur $]-\infty, r[$, vaut δ/r en r, puis décroît strictement et tend vers 0 en $+\infty$. De plus, si $\boldsymbol{\theta} \sim \mathcal{P}(\delta, r)$, on a,

$$\mathbb{P}(\boldsymbol{\theta} \leq t) = \begin{cases} 0 & \text{si } t < r, \\ \int_r^t \delta r^\delta t^{-\delta-1} \mathrm{d}t = 1 - \left(\frac{r}{t}\right)^\delta & \text{si } t \geq r. \end{cases}$$

2. Pour tout $\theta \in \mathbb{R}$, on a

$$\begin{aligned} \pi(\theta \mid \mathbf{X}) &\propto \theta^{-\delta-1}\mathbf{1}_{\theta \geq r} \prod_{i=1}^n \left(\frac{1}{\theta}\mathbf{1}_{0 \leq X_i \leq \theta}\right) \\ &\propto \theta^{-\delta-n-1}\mathbf{1}_{\theta \geq \max\{r, X_1, \ldots, X_n\}}. \end{aligned}$$

Ainsi la loi a posteriori $\Pi[\cdot \mid \mathbf{X}]$ est une loi de Pareto $\mathcal{P}(\delta+n, r_{\mathbf{X}})$ avec $r_{\mathbf{X}} = \max\{r, X_1, \ldots, X_n\}$.

3. Comme la densité a posteriori est nulle sur $]-\infty, r_{\mathbf{X}}[$ et positive décroissante sur $[r_{\mathbf{X}}, +\infty[$, trouver la région de plus haute densité a posteriori de niveau $1-\alpha$ revient à trouver $z_\alpha \geq r_{\mathbf{X}}$ tel que

$$\mathbb{P}\left(\boldsymbol{\theta} \in [r_{\mathbf{X}}, z_\alpha] \mid \mathbf{X}\right) = 1 - \alpha,$$

soit $\mathbb{P}(\boldsymbol{\theta} > z_\alpha \mid \mathbf{X}) = \alpha$. On a

$$\mathbb{P}(\boldsymbol{\theta} > z_\alpha \mid \mathbf{X}) = \alpha \;\Leftrightarrow\; \left(\frac{r_{\mathbf{X}}}{z_\alpha}\right)^{\delta+n} = \alpha \;\Leftrightarrow\; z_\alpha = \frac{r_{\mathbf{X}}}{\alpha^{\frac{1}{\delta+n}}}.$$

Ainsi la région de crédibilité de plus haute densité a posteriori au niveau $1-\alpha$ est donnée par l'intervalle $\left[r_{\mathbf{X}}, r_{\mathbf{X}}\alpha^{-\frac{1}{\delta+n}}\right]$.

4. (a) Soit $\theta_0 \in]0, r[$. Comme sous P_{θ_0}, tous les X_i sont dans $[0, \theta_0]$, on a $r_{\mathbf{X}} = r$, P_{θ_0}-presque sûrement. Ainsi $\Pi\left[ ]-\infty, r[\;\middle|\; \mathbf{X} \right] = 0$. En particulier, on peut trouver $\varepsilon > 0$ assez petit (par exemple $\varepsilon = \frac{r-\theta_0}{2}$) tel que

$$\Pi\left[[\theta_0 - \varepsilon, \theta_0 + \varepsilon] \;\middle|\; \mathbf{X} \right] = 0 .$$

La loi a posteriori n'est donc pas consistante en θ_0.

(b) Soit $\theta_0 \geq r$ et $\varepsilon > 0$. Comme $r \leq \theta_0$ et $X_{(n)} \leq \theta_0$ sous P_{θ_0}, on a $r_{\mathbf{X}} \leq \theta_0 \leq \theta_0 + \varepsilon$. Ainsi, par la question 1, on a P_{θ_0}-presque sûrement,

$$\mathbb{P}\left(\boldsymbol{\theta} > \theta_0 + \varepsilon \;\middle|\; \mathbf{X}\right) = \left(\frac{r_{\mathbf{X}}}{\theta_0 + \varepsilon}\right)^{\delta+n} .$$

D'autre part, toujours par la question 1,

$$\mathbb{P}\left(\boldsymbol{\theta} < \theta_0 - \varepsilon \;\middle|\; \mathbf{X}\right) = \begin{cases} 0 & \text{si } \theta_0 - \varepsilon < r_{\mathbf{X}}, \\ 1 - \left(\frac{r_{\mathbf{X}}}{\theta_0 - \varepsilon}\right)^{\delta+n} & \text{si } \theta_0 - \varepsilon \geq r_{\mathbf{X}}. \end{cases}$$

En sommant les deux probabilités, on obtient bien le résultat souhaité.

(c) En utilisant que $\mathbb{P}\left(\boldsymbol{\theta} < \theta_0 - \varepsilon \;\middle|\; \mathbf{X}\right)$ est nulle si $r_{\mathbf{X}} > \theta_0 - \varepsilon$, et inférieure à 1 sinon, on a

$$\mathbb{P}\left(|\boldsymbol{\theta} - \theta_0| > \varepsilon \;\middle|\; \mathbf{X}\right) \leq \mathbf{1}_{\{r_{\mathbf{X}} \leq \theta_0 - \varepsilon\}} + \left(\frac{r_{\mathbf{X}}}{\theta_0 + \varepsilon}\right)^{\delta+n} .$$

Maintenant, comme $r_{\mathbf{X}} \geq X_{(n)}$, l'événement $\{r_{\mathbf{X}} \leq \theta_0 - \varepsilon\}$ est inclus dans l'événement $\left\{X_{(n)} \leq \theta_0 - \varepsilon\right\}$, donc $\mathbf{1}_{\{r_{\mathbf{X}} \leq \theta_0 - \varepsilon\}} \leq \mathbf{1}_{\left\{X_{(n)} \leq \theta_0 - \varepsilon\right\}}$. De plus, $r_{\mathbf{X}} \leq \theta_0$, ce qui donne bien la majoration souhaitée.

(d) On a

$$\begin{aligned} \mathbb{P}_{\theta_0}\left(X_{(n)} \leq \theta_0 - \frac{M_n}{n}\right) &= \prod_{i=1}^{n} \mathbb{P}_{\theta_0}\left(X_i \leq \theta_0 - \frac{M_n}{n}\right) \\ &= \left(1 - \frac{M_n}{n\theta_0}\right)^n \mathbf{1}_{\left\{\frac{M_n}{n} \leq \theta_0\right\}} \\ &\leq \exp\left(-\frac{M_n}{\theta_0}\right) \mathbf{1}_{\left\{\frac{M_n}{n} \leq \theta_0\right\}} \underset{n\to\infty}{\longrightarrow} 0 . \end{aligned}$$

(e) On a

$$\left(\frac{\theta_0}{\theta_0 + \frac{M_n}{n}}\right)^{\delta+n} = \left(1 - \frac{\frac{M_n}{n}}{\theta_0 + \frac{M_n}{n}}\right)^{\delta+n} \leq \exp\left(-\frac{(\delta+n)M_n}{n\left(\theta_0 + \frac{M_n}{n}\right)}\right) \underset{n\to\infty}{\longrightarrow} 0 ,$$

puisque $\frac{\delta+n}{n} \sim 1$ et $\frac{M_n}{\theta_0 + \frac{M_n}{n}} = \frac{1}{\frac{\theta_0}{M_n} + \frac{1}{n}} \underset{n\to\infty}{\longrightarrow} +\infty$.

(f) En combinant les résultats précédents, on a

$$\mathbb{E}_{\theta_0}\left[\mathbb{P}\left(|\boldsymbol{\theta} - \theta_0| > \frac{M_n}{n} \;\middle|\; \mathbf{X}\right)\right] \leq \mathbb{P}_{\theta_0}\left(X_{(n)} \leq \theta_0 - \frac{M_n}{n}\right) + \left(\frac{\theta_0}{\theta_0 + \frac{M_n}{n}}\right)^{\delta+n} \underset{n\to\infty}{\longrightarrow} 0 ,$$

ce qui implique $\mathbb{P}\left(|\boldsymbol{\theta} - \theta_0| > \frac{M_n}{n} \;\middle|\; \mathbf{X}\right) \underset{n\to\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0$. Auttrement dit, la loi a posteriori converge à vitesse $1/n$, en tout point $\theta_0 \geq r$.

Exercice 4

Partie 1 On échantillonne ici selon le modèle bayésien :

$$\begin{gathered}\boldsymbol{\theta} \sim \Gamma(a,b), \quad a>0, b>0,\\ \mathbf{X} \mid \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\theta},1)^{\otimes n}, \quad n \in \mathbb{N}^*.\end{gathered}$$

Dans l'illustration numérique, l'objet Python `X` est une réalisation du vecteur aléatoire $\mathbf{X} \mid \boldsymbol{\theta} = \theta_0$, θ_0 étant lui-même une réalisation de la variable aléatoire $\boldsymbol{\theta}$.
On remarque que la loi a priori est bien une loi $\Gamma(a,b)$ puisque la densité de cette dernière est proportionnelle à $x \mapsto x^{a-1}\,\mathrm{e}^{-bx}\,\mathbf{1}_{\mathbb{R}_+}(x)$, où l'on reconnaît le paramètre de dilatation $1/b$.

Partie 2 On calcule ensuite deux estimateurs du paramètre θ_0 : l'estimateur du maximum de vraisemblance (qui est bien $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ puisque nous sommes dans le modèle des lois gaussiennes translatées) et $\mathbb{E}[\boldsymbol{\theta} \mid \mathbf{X}]$, qui est un estimateur de Bayes pour la perte quadratique.
La loi a posteriori n'étant pas une loi classique :

$$\forall \theta \in \mathbb{R}, \quad \pi(\theta \mid \mathbf{X}) \propto \theta^{a-1}\,\mathrm{e}^{-\frac{n}{2}(\theta-(\bar{X}_n-b/n))^2}\,\mathbf{1}_{\mathbb{R}_+}(\theta),$$

on souhaite approcher $\mathbb{E}[\boldsymbol{\theta} \mid \mathbf{X}]$ par une méthode Monte Carlo. Pour ce faire, on remarque que

$$\mathbb{E}[\boldsymbol{\theta} \mid \mathbf{X}] = \psi(\mathbf{X}), \quad \text{où} \quad \forall x \in \mathbb{R}^n, \psi(x) = \frac{\mathbb{E}[\boldsymbol{\theta} p_{\boldsymbol{\theta}}(x)]}{\mathbb{E}[p_{\boldsymbol{\theta}}(x)]},$$

avec, pour tout $\theta \in \mathbb{R}_+$, p_θ la densité de $N(\theta,1)^{\otimes n}$. Ainsi, un estimateur Monte Carlo est obtenu avec $N \in \mathbb{N}^*$ variables aléatoires i.i.d. $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$ de loi $\Gamma(a,b)$ en calculant :

$$\frac{\sum_{i=1}^N \boldsymbol{\theta}_i p_{\boldsymbol{\theta}_i}(\mathbf{X})}{\sum_{i=1}^N p_{\boldsymbol{\theta}_i}(\mathbf{X})}.$$

Partie 3 On estime ici le risque a posteriori de ces deux estimateurs pour la perte quadratique. Avec un raisonnement identique, on obtient que, pour un estimateur $T(\mathbf{X})$ de θ_0,

$$\mathbb{E}[(\boldsymbol{\theta} - T(\mathbf{X}))^2 \mid \mathbf{X}] = \varphi(\mathbf{X}), \quad \text{où} \quad \forall x \in \mathbb{R}^n, \varphi(x) = \frac{\mathbb{E}[(\boldsymbol{\theta} - T(x))^2 p_{\boldsymbol{\theta}}(x)]}{\mathbb{E}[p_{\boldsymbol{\theta}}(x)]},$$

dont un estimateur Monte Carlo est :

$$\frac{\sum_{i=1}^N (\boldsymbol{\theta}_i - T(\mathbf{X}))^2 p_{\boldsymbol{\theta}_i}(\mathbf{X})}{\sum_{i=1}^N p_{\boldsymbol{\theta}_i}(\mathbf{X})}.$$

On remarque que le risque a posteriori de l'estimateur de Bayes pour la perte quadratique est moindre que celui de l'estimateur du maximum de vraisemblance. Cela était attendu puisque, par construction, l'estimateur de Bayes $\mathbb{E}[\boldsymbol{\theta} \mid \mathbf{X}]$ minimise le risque a posteriori.

Partie 4 On souhaite enfin calculer un dernier estimateur : la médiane a posteriori, qui est un estimateur de Bayes pour la perte en valeur absolue. Puisque cette quantité n'est pas exprimable facilement, nous souhaitons l'estimer. Pour ce faire, on construit un échantillon de la loi a posteriori par la méthode du rejet (avec comme loi instrumentale, la loi a priori), et on calcule la médiane empirique (qui et un estimateur convergent de la médiane dès lors que la densité a posteriori est strictement positive en la médiane).
Comme dans le TP 2, qui traitait du même cadre, on remarque que pour tout candidat $\theta \in \mathbb{R}$,

$$\frac{\pi(\theta \mid \mathbf{X})}{\pi(\theta)} = \frac{p_\theta(\mathbf{X})}{\int_{\mathbb{R}} \pi(t) p_t(\mathbf{X})\, dt} \leq \frac{p_{\bar{X}_n}(\mathbf{X})}{\int_{\mathbb{R}} \pi(t) p_t(\mathbf{X})\, dt}.$$

Ainsi, le ratio d'acceptation vaut tout simplement :

$$\frac{\pi(\theta \mid \mathbf{X})}{\pi(\theta)} \frac{\int_{\mathbb{R}} \pi(t) p_t(\mathbf{X})\, dt}{p_{\bar{X}_n}(\mathbf{X})} = \frac{p_\theta(\mathbf{X})}{p_{\bar{X}_n}(\mathbf{X})} = \mathrm{e}^{-\frac{n}{2}(\bar{X}_n - \theta)^2}.$$

Par ailleurs, on pourra remarquer que le script Python proposé ici est légèrement plus efficace (en pratique, il est largement plus efficace) que celui donné dans le TP 2 puisque chaque tour de boucle réalise des opérations sur des tableaux de nombres (plutôt que sur un seul nombre) et que le taux de rejet, bien qu'inconnu, est pris en compte en échantillonnant à chaque tour de boucle $2m$ candidats.
Comme attendu, ce dernier estimateur est meilleur que les deux autres estimateurs au sens du risque a posteriori associé à la perte en valeur absolue.