Sorbonne Université
Master Mathématiques et Applications
MU4MA015 : Statistique

Lundi 3 janvier 2022
Aucun document autorisé
Durée : 3 heures

Correction Examen

Exercice 1
On considère une variable aléatoire $Y \sim \mathcal{N}(0,1)$ de fonction de répartition notée Φ, un paramètre inconnu $\theta \in \mathbb{R}$, et la variable aléatoire $X = \theta + |Y|$. De plus, on considère $X_1, \ldots, X_n$ i.i.d. et de même loi que X.

1. Pour toute fonction mesurable bornée $\varphi : \mathbb{R} \to \mathbb{R}$, on a
$$\mathbb{E}[\varphi(X)] = \mathbb{E}[\varphi(\theta + |Y|)] = \int_{\mathbb{R}} \varphi(\theta + |y|) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 2\int_0^{+\infty} \varphi(\theta + y) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$
et le changement de variable $x = \theta + y$ donne
$$\mathbb{E}[\varphi(X)] = \int_\theta^{+\infty} \varphi(x) \sqrt{\frac{2}{\pi}} e^{-\frac{(x-\theta)^2}{2}} dx = \int_{\mathbb{R}} \varphi(x) \sqrt{\frac{2}{\pi}} e^{-\frac{(x-\theta)^2}{2}} \mathbf{1}_{x \geq \theta}\, dx,$$
ce qui prouve bien que
$$f_\theta(x) = \sqrt{\frac{2}{\pi}}\, e^{-\frac{(x-\theta)^2}{2}} \mathbf{1}_{x \geq \theta}.$$

2. Il vient
$$\mathbb{E}[|Y|] = \int_{\mathbb{R}} |y| \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 2\int_0^{+\infty} y \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \sqrt{\frac{2}{\pi}} \left[-e^{-\frac{y^2}{2}}\right]_0^{+\infty} = \sqrt{\frac{2}{\pi}}.$$
De plus, puisque $\mathbb{E}[Y^2] = 1$, on en déduit que
$$\mathrm{Var}(|Y|) = \mathbb{E}[|Y|^2] - \mathbb{E}[|Y|]^2 = \mathbb{E}[Y^2] - \mathbb{E}[|Y|]^2 = 1 - \frac{2}{\pi} = \frac{\pi - 2}{\pi}.$$

3. Comme $X = \theta + |Y|$, ce qui précède implique que $\mathbb{E}[X] = \theta + \sqrt{\frac{2}{\pi}}$ et $\mathrm{Var}(X) = \mathrm{Var}(|Y|) = \frac{\pi-2}{\pi}$. La loi des grands nombres assure que $\hat{\theta}_n := \bar{X}_n - \sqrt{\frac{2}{\pi}}$ est un estimateur fortement consistant de θ. En outre, le Théorème Central Limite donne
$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) = \sqrt{n}\left(\overline{X}_n - \left(\theta + \sqrt{\frac{2}{\pi}}\right)\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{\pi - 2}{\pi}\right).$$

4. En notant $q_{1-\alpha/2}$ le quantile d'ordre $(1 - \alpha/2)$ de la loi normale centrée réduite, un intervalle de confiance de niveau asymptotique $(1 - \alpha)$ est
$$IC = \left[\hat{\theta}_n - q_{1-\alpha/2}\sqrt{\frac{\pi - 2}{\pi n}}\ ;\ \hat{\theta}_n + q_{1-\alpha/2}\sqrt{\frac{\pi - 2}{\pi n}}\right].$$

5. Notons F_θ la fonction de répartition de X. Celle-ci est nulle pour $x \leq \theta$, et pour tout $x \geq \theta$ on peut écrire

$$F_\theta(x) = \mathbb{P}(X \leq x) = \mathbb{P}(|Y| \leq x - \theta) = 2\Phi(x-\theta) - 1,$$

donc

$$F_\theta(x) = \frac{1}{2} \iff \Phi(x-\theta) = 3/4 \iff x = \theta + \Phi^{-1}(3/4) = \theta + q_{3/4},$$

c'est-à-dire que $x_{1/2} = \theta + q_{3/4}$.

6. Par le calcul précédent et compte tenu de la stricte croissance de Φ, la fonction F_θ est strictement croissante en $x_{1/2}$ et de dérivée f_θ strictement positive en ce point, donc l'estimateur $\tilde{\theta}_n = x_{1/2} - q_{3/4}$ est fortement consistant et vérifie

$$\sqrt{n}\left(\tilde{\theta}_n - \theta\right) = \sqrt{n}\left(x_{1/2}(n) - (\theta + q_{3/4})\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{\pi \exp\left(q_{3/4}^2\right)}{8}\right).$$

7. La vraisemblance a pour expression

$$L_n(\theta) = \left(\frac{2}{\pi}\right)^{n/2} e^{-\frac{\sum_{i=1}^n (X_i - \theta)^2}{2}} \mathbf{1}_{\theta \leq X_{(1)}}.$$

Puisque la fonction

$$\theta \mapsto \sum_{i=1}^n (X_i - \theta)^2 = n\left(\theta - \bar{X}_n\right)^2 + \sum_{i=1}^n X_i^2 - n\bar{X}_n^2$$

est décroissante jusqu'en $\theta = \bar{X}_n \geq X_{(1)}$, la fonction $\theta \mapsto L_n(\theta)$ est croissante jusqu'en $X_{(1)}$ et nulle à droite de ce point, ce qui prouve que l'estimateur du maximum de vraisemblance est $X_{(1)}$.

8. le développement limité de Φ à l'ordre 1 en 0 est

$$\Phi(x) = \Phi(0) + \Phi'(0)x + o(x) = \frac{1}{2} + \frac{x}{\sqrt{2\pi}} + o(x).$$

On en déduit que

$$2(1 - \Phi(x)) = 1 - \sqrt{\frac{2}{\pi}}x + o(x).$$

9. Comme $X_{(1)} \geq \theta$ presque sûrement, on commence par noter que, pour tout $x < 0$,

$$\mathbb{P}\left(n\left(X_{(1)} - \theta\right) \leq x\right) = 0.$$

Par ailleurs, pour tout $x \geq 0$, on a

$$\begin{aligned}\mathbb{P}\left(n\left(X_{(1)} - \theta\right) \leq x\right) &= \mathbb{P}\left(X_{(1)} \leq \theta + \frac{x}{n}\right) = 1 - \mathbb{P}\left(X \geq \theta + \frac{x}{n}\right)^n = 1 - \mathbb{P}\left(|Y| \geq \frac{x}{n}\right)^n \\ &= 1 - \left(2\mathbb{P}\left(Y \geq \frac{x}{n}\right)\right)^n = 1 - \left(2(1 - \Phi(x/n))\right)^n.\end{aligned}$$

Or la question précédente donne

$$(2(1-\Phi(x/n)))^n = \exp\left(n \log\left(1-\sqrt{\frac{2}{\pi}}\frac{x}{n}+o(1/n)\right)\right) \xrightarrow[n\to+\infty]{} \exp\left(-\sqrt{\frac{2}{\pi}}x\right),$$

ce qui montre que, pour tout $x \geq 0$,

$$\mathbb{P}\left(n\left(X_{(1)}-\theta\right) \leq x\right) \xrightarrow[n\to+\infty]{} 1-\exp\left(-\sqrt{\frac{2}{\pi}}x\right),$$

c'est-à-dire que $n\left(X_{(1)}-\theta\right)$ tend en loi vers une loi exponentielle de paramètre $\sqrt{\frac{2}{\pi}}$.

10. Puisque $X_{(1)} \geq \theta$ presque sûrement, on a pour tout $c_\alpha > 0$

$$\mathbb{P}\left(X_{(1)}-\frac{c_\alpha}{n} \leq \theta \leq X_{(1)}\right) = \mathbb{P}\left(n\left(X_{(1)}-\theta\right) \leq c_\alpha\right) \xrightarrow[n\to+\infty]{} 1-\exp\left(-\sqrt{\frac{2}{\pi}}c_\alpha\right)$$

et

$$1-\exp\left(-\sqrt{\frac{2}{\pi}}c_\alpha\right) = 1-\alpha \iff c_\alpha = -\sqrt{\frac{\pi}{2}}\ln\alpha.$$

Par conséquent

$$\left[X_{(1)}+\frac{\sqrt{\pi}}{\sqrt{2}n}\ln\alpha\ ;\ X_{(1)}\right]$$

est un intervalle de confiance de niveau asymptotique $(1-\alpha)$.

11. Des trois estimateurs précédents, on choisit $X_{(1)}$ puisqu'il converge à vitesse $1/n$ tandis que les deux autres convergent à vitesse $1/\sqrt{n}$.

12. Pour tout $x \in \mathbb{R}$, la fonction $\theta \mapsto f_\theta(x) = \sqrt{\frac{2}{\pi}}\, e^{-\frac{(x-\theta)^2}{2}}\, \mathbf{1}_{x\geq\theta}$ est discontinue en $\theta = x$ donc le modèle n'est pas régulier.

Exercice 2

1. Pour $\sigma^2 = 1$, la formule de Bayes donne que la densité a posteriori est proportionnelle à

$$\pi(\theta \mid X_1) \propto \exp\left(-\frac{1}{2}(X_1-\theta)^2 - \frac{(\theta-\mu)^2}{2}\right) \propto \exp\left(-\left(\theta-\frac{X_1+\mu}{2}\right)^2\right).$$

On reconnaît la densité d'une loi $\mathcal{N}((\mu+X_1)/2, 1/2)$.

2. De même, pour $\sigma^2 > 0$,

$$\begin{aligned}\pi(\theta \mid X_1) &\propto \exp\left(-\frac{1}{2}(X_1-\theta)^2 - \frac{(\theta-\mu)^2}{2\sigma^2}\right) \propto \exp\left(-\frac{1+\sigma^{-2}}{2}\theta^2 + (X_1+\mu\sigma^{-2})\theta\right)\\ &\propto \exp\left(-\frac{1+\sigma^{-2}}{2}\left(\theta-\frac{X_1+\mu\sigma^{-2}}{1+\sigma^{-2}}\right)^2\right).\end{aligned}$$

On reconnaît la densité d'une loi $\mathcal{N}((X_1+\mu\sigma^{-2})/(1+\sigma^{-2}), 1/(1+\sigma^{-2}))$.

3. Il s'agit de la moyenne a posteriori, soit $E[\boldsymbol{\theta} \mid \mathbf{X}] = (X_1 + \mu\sigma^{-2})/(1+\sigma^{-2})$ ou encore

$$\hat{\theta} = \mu + \alpha(X_1 - \mu), \qquad \text{avec } \alpha = \frac{\sigma^2}{\sigma^2+1}.$$

Exercice 3

1. Le risque ponctuel de $T_M = X_1$ vaut $\mathbb{E}_\theta[(X_1-\theta)^2/\theta] = \theta^{-1}\mathrm{Var}[\mathcal{P}(\theta)] = 1$.

2. La formule de Bayes donne que la densité a posteriori est proportionnelle à

$$\pi(\theta \mid \mathbf{X} = X_1) \propto e^{-\theta}\frac{\theta^{X_1}}{X_1!}\theta^{a-1}e^{-b\theta}\mathbf{1}_{\theta>0} \propto e^{-(b+1)\theta}\theta^{a+X_1-1}\mathbf{1}_{\theta>0}.$$

   On reconnaît la densité d'un loi $\Gamma(a+X_1, b+1)$.

3. (a) Comme $\mathbb{E}[\Gamma(a,b)] = a/b$ on obtient $\mathbb{E}[\boldsymbol{\theta} \mid \mathbf{X}] = (a+X_1)/(b+1)$. De même, pour $a>1$, on calcule $\mathbb{E}[\Gamma(a,b)^{-1}] = b/(a-1)$ ce qui donne

$$\mathbb{E}[\boldsymbol{\theta}^{-1} \mid \mathbf{X}] = (b+1)/(a+X_1-1).$$

   (b) Par un théorème du cours, un estimateur de Bayes pour la fonction de perte L et la loi a priori Π peut s'obtenir en minimisant en T le risque a posteriori

$$\begin{aligned}\mathbb{E}\left[\frac{(T(\mathbf{X})-\boldsymbol{\theta})^2}{\boldsymbol{\theta}} \mid \mathbf{X}\right] &= \mathbb{E}[\boldsymbol{\theta}^{-1} \mid \mathbf{X}]T(\mathbf{X})^2 - 2T(\mathbf{X}) + \mathbb{E}[\boldsymbol{\theta} \mid \mathbf{X}] \\ &= \frac{b+1}{a+X_1-1}\left(T(\mathbf{X}) - \frac{a+X_1-1}{b+1}\right)^2 + c(X_1),\end{aligned}$$

   où $c(X_1)$ est une expression qui dépend de X_1 mais pas de T. L'expression précédente est donc minimale pour $T(\mathbf{X}) = T^*_{a,b}(\mathbf{X}) = (a+X_1-1)/(b+1)$.

4. (a) On pose $a=1$ et on évalue $\mathbb{E}_\theta T^*_{1,b}(\mathbf{X}) = \theta/(b+1)$ et $\mathbb{E}_\theta T^*_{1,b}(\mathbf{X})^2 = (\theta^2+\theta)/(b+1)^2$ puis, comme à la question précédente en développant le carré $\mathbf{R}(\theta, T) = \theta^{-1}\mathbb{E}_\theta T^2 - 2\mathbb{E}_\theta T + \theta$ ce qui donne $\mathbf{R}(\theta, T^*_{1,b}) = (b^2\theta+1)/(b+1)^2$.

   (b) Suivant les valeurs de θ dans $\mathbb{R}^\star_+$, le risque à la question précédente est parfois inférieur, parfois supérieur à $1 = \mathbf{R}(\theta, T_M)$, donc aucun estimateur n'est uniformément meilleur qu'un autre.

5. (a) La question 3.(b) correspond au calcul de $\mathbb{E}[L(\boldsymbol{\theta}, T(\mathbf{X})) \mid \mathbf{X}]$: la constante $c(X_1)$ introduite plus haut vaut

$$c(X_1) = \frac{a+X_1}{b+1} - \frac{a+X_1-1}{b+1} = \frac{1}{b+1}.$$

   On peut écrire $\mathbb{E}[L(\boldsymbol{\theta}, T^*_{a,b}(\mathbf{X})) \mid \mathbf{X}] = c(X_1) = 1/(b+1)$, puisque le terme carré de la question 3.(b). s'annule par définition de $T^*_{a,b}$. En reprenant l'espérance on obtient $\mathbf{R}_B(\Pi, T^*_{a,b}) = 1/(b+1)$.

   (b) Comme T_M est de risque constant égal à 1, on a $\mathbf{R}_B(\Pi, T_M) = 1$, qui est donc supérieur à $\mathbf{R}_B(\Pi, T^*_{a,b}) = 1/(b+1)$. C'est attendu puisque $T^*_{a,b}$ est un estimateur de Bayes donc de risque bayésien minimal.

6. Considérons $(\Pi_k)_{k\geq 1}$ la suite de lois a priori avec $a > 1$ fixé et $b = 1/k$. Quand $k \to \infty$, on note que

$$R_B(\Pi_k) = R_B(\Pi_k, T^*_{a,k^{-1}}) = \frac{1}{k^{-1}+1} \to 1.$$

Or T_M est tel que $\mathbf{R}_{\max}(T_M) = 1$. On en déduit que $T_M = X_1$ est minimax.

7. (a) On procède de même qu'à la question 2, avec cette fois une vraisemblance proportionnelle au produit $\prod_{i=1}^n \{e^{-\theta}\theta^{X_i}\} = e^{-n\theta}\theta^{n\bar{X}_n}$ et on obtient la loi a posteriori $\Gamma(a + n\bar{X}_n, b+n)$.

   (b) Il suffit de montrer que $\psi_n(\mathbf{X})$ définie par

   $$\psi_n(\mathbf{X}) := \log \mathbb{E}\left[e^{t\sqrt{n}(\boldsymbol{\theta}-\bar{\mathbf{X}})} \mid \mathbf{X}\right]$$

   converge en probabilité, sous $\mathbb{P}_{\theta_0}$, vers $t^2\theta_0/2$. En utilisant la question précédente et la transformée de Laplace de la loi Gamma donnée dans l'énoncé (qui est bien définie pour n assez grand à t fixé, i.e. dès que $t\sqrt{n} < b+n$), on obtient, en utilisant le développement limité de l'énoncé,

   $$\begin{aligned}\psi_n(\mathbf{X}) &= -(a+n\bar{X}_n)\log\left(1-\frac{t\sqrt{n}}{b+n}\right) - t\sqrt{n}\bar{X}_n \\ &= t\sqrt{n}\bar{X}_n\left(\frac{n}{b+n}-1\right) + O\left(\frac{1}{\sqrt{n}}\right) + \frac{t^2}{2}\frac{n(a+n\bar{X}_n)}{(b+n)^2} + (a+n\bar{X}_n)o(n^{-1}).\end{aligned}$$

   En utilisant que $\bar{X}_n$ converge en probabilité vers $\mathbb{E}_{\theta_0}X_1 = \theta_0$ et le lemme de Slutsky, on obtient que l'expression précédente converge en probabilité vers $t^2\theta_0/2$.

   (c) Pour Z de loi $\mathcal{N}(0,v)$, on calcule $\mathbb{E}\left[e^{vZ}\right] = e^{vt^2/2}$. D'après la question précédente, on peut donc prendre $v = \theta_0$.

   (d) Ce résultat ressemble de très près au théorème de Bernstein–von Mises. En effet, notons que l'estimateur du maximum de vraisemblance est ici $\bar{X}_n$ et l'information de Fisher au point θ_0 égale à $1/\theta_0$ (petit calcul à partir du score du modèle de Poisson). La seule différence est que nous l'avons obtenu ici en termes de convergence en loi, et non de convergence en variation totale.

8. (a) Les méthodes MCMC (Markov Chain Monte Carlo), typiquement l'algorithme de Metropolis–Hastings, permettent de simuler de façon approchée suivant la loi a posteriori. Le principe consiste à simuler une chaîne de Markov dont la loi stationnaire est la loi a posteriori.

   (b) À partir d'une sortie $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$ d'un tel algorithme, on peut poser

   $$\hat{I}_N = \frac{1}{N}\sum_{i=1}^{N} e^{t\sqrt{n}(\boldsymbol{\theta}_i - \bar{X}_n)},$$

   qui fournit une approximation de $\mathbb{E}[e^{t\sqrt{n}(\boldsymbol{\theta}-\bar{X}_n)} \mid \mathbf{X}]$.