

Examen

Rappels

- Il sera tenu compte du soin apporté à la présentation et à la rédaction.
- Si vous n'arrivez pas à démontrer un résultat, vous pouvez l'admettre pour la suite de l'exercice.

EXERCICE 1 (10 points)

Soient $\theta > 0$ un paramètre inconnu et X de densité

$$f_{\theta}(x) = \frac{\theta}{x^{\theta+1}} \mathbf{1}_{x \geq 1}.$$

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

1. Déterminer la médiane de la loi de X . En déduire un estimateur $\hat{\theta}_n$ de θ et établir sa normalité asymptotique.

Correction. Pour tout $\theta > 0$, la fonction de répartition

$$F_{\theta}(x) = \left(1 - \frac{1}{x^{\theta}}\right) \mathbf{1}_{x \geq 1}$$

établit une bijection de $[1, +\infty[$ vers $[0, 1[$ et la médiane vaut donc

$$x_{1/2} = F_{\theta}^{-1}(1/2) = 2^{\frac{1}{\theta}} \quad \Longleftrightarrow \quad \theta = \frac{\log 2}{\log(x_{1/2})} =: \varphi(x_{1/2}).$$

En notant $x_{1/2}(n)$ la médiane empirique, on propose donc l'estimateur

$$\hat{\theta}_n := \varphi(x_{1/2}(n)) = \frac{\log 2}{\log(x_{1/2}(n))}.$$

Puisque $f_{\theta}(x_{1/2}) = \theta/2^{1+\frac{1}{\theta}} > 0$, on sait que

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

La méthode Delta donne alors

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

2. Déterminer l'estimateur du maximum de vraisemblance $\tilde{\theta}_n$.

Correction. La vraisemblance s'écrit

$$L_n(\theta) = \prod_{i=1}^n \frac{\theta}{X_i^{\theta+1}} \mathbf{1}_{X_i \geq 1} = \prod_{i=1}^n \frac{\theta}{X_i^{\theta+1}},$$

les indicatrices étant superflues puisque, quel que soit $\theta > 0$, on a $X_i \geq 1$ presque sûrement. On peut donc passer à la log-vraisemblance :

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

puis à sa dérivée :

$$\ell'_n(\theta) = \frac{n}{\theta} - \sum_{i=1}^n \log X_i,$$

or celle-ci est positive si et seulement si

$$\theta \leq \tilde{\theta}_n := \frac{1}{\frac{1}{n} \sum_{i=1}^n \log X_i},$$

qui est donc l'estimateur du maximum de vraisemblance.

3. Montrer que la variable aléatoire $T = \log X$ suit une loi exponentielle dont on précisera le paramètre. En déduire $\mathbb{E}_\theta[\log X]$ et $\text{Var}_\theta(\log X)$.

Correction. La méthode de la fonction muette ou le calcul de la fonction de répartition donnent $T \sim \mathcal{E}(\theta)$, donc $\mathbb{E}_\theta[\log X] = 1/\theta$ et $\text{Var}(\log X) = 1/\theta^2$.

4. Etablir la normalité asymptotique de $\tilde{\theta}_n$.

Correction. Les variables $T_i := \log X_i$ étant donc i.i.d. suivant la loi exponentielle ci-dessus, le TCL donne

$$\sqrt{n} \left(\bar{T}_n - \frac{1}{\theta} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N} \left(0, \frac{1}{\theta^2} \right).$$

Puisque $\tilde{\theta}_n = 1/\bar{T}_n$, il ne reste plus qu'à appliquer la méthode Delta pour obtenir

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

5. D'un point de vue asymptotique, entre $\hat{\theta}_n$ et $\tilde{\theta}_n$, quel estimateur choisissez-vous ?

Correction. Les deux estimateurs convergeant à la même vitesse $1/\sqrt{n}$, on choisit celui ayant la plus petite variance asymptotique, c'est-à-dire $\tilde{\theta}_n$ puisque $0 < \log 2 < 1$.

6. Le modèle $(f_\theta)_{\theta > 0}$ est-il régulier ? Si oui, calculer l'information de Fisher $I(\theta)$ pour une donnée.

Correction. Pour tout réel x fixé, la fonction $\theta \mapsto f_\theta(x)$ est C^1 sur $\Theta =]0, \infty[$ donc les deux premiers points de la définition d'un modèle régulier sont vérifiés. De plus, le score vaut $\ell'_\theta(X) = \frac{1}{\theta} - \log X$ et la fonction

$$\theta \mapsto \mathbb{E}_\theta [\ell'_\theta(X)^2] = \mathbb{E}_\theta \left[\left(\frac{1}{\theta} - \log X \right)^2 \right] = \text{Var}(\log X) = \frac{1}{\theta^2}$$

étant continue sur Θ , le modèle est régulier avec $I(\theta) = 1/\theta^2$.

7. Préciser si l'un des estimateurs précédents est asymptotiquement efficace.

Correction. L'EMV $\tilde{\theta}_n$ est asymptotiquement efficace, tandis que $\hat{\theta}_n$ ne l'est pas.

8. Soit $\alpha \in]0, 1[$. Donner un intervalle de confiance de niveau asymptotique $(1 - \alpha)$ pour θ .

Correction. On considère bien entendu $\tilde{\theta}_n$. On peut partir de la convergence en loi

$$\sqrt{n} \left(\frac{\tilde{\theta}_n}{\theta} - 1 \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)$$

pour en déduire l'intervalle de confiance asymptotique

$$\left[\frac{\tilde{\theta}_n}{1 + \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}} ; \frac{\tilde{\theta}_n}{1 - \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}} \right].$$

Une autre solution consiste à appliquer le Lemme de Slutsky pour dire que

$$\sqrt{n} \left(\frac{\tilde{\theta}_n - \theta}{\tilde{\theta}_n} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)$$

dont il découle l'intervalle

$$\left[\tilde{\theta}_n \left(1 - \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}} \right) ; \tilde{\theta}_n \left(1 + \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}} \right) \right].$$

9. Proposer un test de niveau asymptotique α pour décider entre $H_0 : \theta \leq 1$ et $H_1 : \theta > 1$. Déterminer la p-valeur asymptotique $\alpha_0(\mathbf{x})$ associée à une réalisation $\mathbf{x} = (x_1, \dots, x_n)$.

Correction. Vu qu'on a affaire un test unilatère, on part d'un intervalle de confiance unilatère de sens opposé à $\Theta_0 =]0, 1]$, c'est-à-dire de la forme $[A_n, +\infty[$. Les raisonnements sont les mêmes qu'en question précédente. Si l'on part de l'intervalle de confiance asymptotique

$$I_n = \left[\frac{\tilde{\theta}_n}{1 + \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}}} ; +\infty \right],$$

on en déduit que le test $T(\mathbf{X}) = \mathbf{1}_{I_n \cap]0, 1] = \emptyset}$ est de niveau asymptotique α , ce qui s'écrit encore

$$T(\mathbf{X}) = 1 \iff \tilde{\theta}_n > 1 + \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}}.$$

Pour une réalisation $\mathbf{x} = (x_1, \dots, x_n)$, la p-valeur asymptotique associée est

$$\alpha_0(\mathbf{x}) = \inf \left\{ \alpha \in [0, 1], \tilde{\theta}_n(\mathbf{x}) > 1 + \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}} \right\} = 1 - \Phi \left(\sqrt{n}(\tilde{\theta}_n(\mathbf{x}) - 1) \right).$$

Si l'on part plutôt de l'intervalle

$$I'_n = \left[\tilde{\theta}_n \left(1 - \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}} \right) ; +\infty \right],$$

le test $T'(\mathbf{X}) = \mathbf{1}_{I'_n \cap]0, 1] = \emptyset}$ est lui aussi de niveau asymptotique α , ce qui donne la p-valeur asymptotique

$$\alpha_0(\mathbf{x}) = 1 - \Phi \left(\sqrt{n} \left(1 - \frac{1}{\tilde{\theta}_n(\mathbf{x})} \right) \right).$$

Noter que ces deux tests sont en fait de **taille** asymptotique α . Si l'on part des intervalles de confiance de la question précédente, on obtient des tests qui sont bien de niveau asymptotique α mais de taille asymptotique $\alpha/2$.

10. Que se passe-t-il si l'on veut estimer θ par application de la méthode des moments ?

Correction. On a

$$\mathbb{E}_\theta[X] = \int_1^\infty \frac{\theta}{x^\theta} dx,$$

intégrale convergente si et seulement si $\theta > 1$. Or a priori on sait uniquement que $\theta > 0$, donc cette méthode n'est pas adaptée. Clairement, le même problème apparaît pour tout $k \in \mathbb{N}^*$ si l'on souhaite se baser sur le moment d'ordre k , c'est-à-dire $\mathbb{E}_\theta[X^k]$, puisque cette espérance n'est finie que pour $\theta > k$. On peut également noter que l'étude de l'EMV revient en fait à appliquer le même raisonnement que pour la méthode des moments à $\mathbb{E}_\theta[h(X)]$ avec $h(x) = \log x$.

EXERCICE 2 (8 points)

Soit $\sigma^2 > 0$ et soient $x_1, \dots, x_n$ des réels **fixés et connus**. On considère le cadre bayésien suivant :

$$\boldsymbol{\theta} \sim \Pi_{\sigma^2} = \mathcal{N}(0, \sigma^2)$$

$$\mathbf{Y} = (Y_1, \dots, Y_n) \mid \boldsymbol{\theta} = \theta \sim \mathcal{N}(\theta x_1, 1) \otimes \dots \otimes \mathcal{N}(\theta x_n, 1).$$

Autrement dit, la loi de $\mathbf{Y}$ est celle de $\boldsymbol{\theta}x + \varepsilon$, où $x = (x_1, \dots, x_n)$, et où $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \sim \mathcal{N}(0, 1)^{\otimes n}$, avec $\boldsymbol{\theta}$ et ε indépendantes.

1. Montrer que la loi a posteriori est donnée par

$$\Pi_{\sigma^2}(\cdot \mid \mathbf{Y}) = \mathcal{N}\left(\frac{\langle x, \mathbf{Y} \rangle}{\sigma^{-2} + \|x\|^2}, \frac{1}{\sigma^{-2} + \|x\|^2}\right),$$

où $\langle x, \mathbf{Y} \rangle = \sum_{i=1}^n x_i Y_i$ et $\|x\|^2 = \langle x, x \rangle = \sum_{i=1}^n x_i^2$.

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

$$\begin{aligned} \pi(\theta \mid \mathbf{Y}) &\propto e^{-\frac{\theta^2}{2\sigma^2}} \prod_{i=1}^n e^{-\frac{(Y_i - \theta x_i)^2}{2}} \\ &\propto e^{-\frac{1}{2}(\theta^2 \sigma^{-2} + \sum_{i=1}^n (\theta^2 x_i^2 - 2\theta x_i Y_i))} \\ &\propto e^{-\frac{\sigma^{-2} + \|x\|^2}{2} \left(\theta^2 - 2\theta \frac{\langle x, \mathbf{Y} \rangle}{\sigma^{-2} + \|x\|^2} \right)} \\ &\propto e^{-\frac{\sigma^{-2} + \|x\|^2}{2} \left(\theta - \frac{\langle x, \mathbf{Y} \rangle}{\sigma^{-2} + \|x\|^2} \right)^2} \end{aligned}$$

On a donc bien

$$\Pi_{\sigma^2}(\cdot \mid \mathbf{Y}) = \mathcal{N}\left(\frac{\langle x, \mathbf{Y} \rangle}{\sigma^{-2} + \|x\|^2}, \frac{1}{\sigma^{-2} + \|x\|^2}\right).$$

2. Quelle est la moyenne a posteriori $m_{\mathbf{Y}}$? Et la variance a posteriori $v_{\mathbf{Y}}$?

Correction. On a

$$m_{\mathbf{Y}} = \frac{\langle x, \mathbf{Y} \rangle}{\sigma^{-2} + \|x\|^2} \quad \text{et} \quad v_{\mathbf{Y}} = \frac{1}{\sigma^{-2} + \|x\|^2}.$$

3. On considère la perte quadratique $\ell(\theta, t) = (\theta - t)^2$.

- (a) Donner un estimateur de Bayes. On le notera $T^*(\mathbf{Y})$.

Correction. C'est la moyenne a posteriori : $T^*(\mathbf{Y}) = m_{\mathbf{Y}}$.

- (b) Que vaut le risque a posteriori de $T^*(\mathbf{Y})$?

Correction. On a

$$\rho(\Pi, T^*(\mathbf{Y}) \mid \mathbf{Y}) = \mathbb{E}[(\theta - m_{\mathbf{Y}})^2 \mid \mathbf{Y}] = v_{\mathbf{Y}} = \frac{1}{\sigma^{-2} + \|x\|^2}.$$

(c) En déduire la valeur du risque de Bayes $R_B(\Pi_{\sigma^2})$.

Correction. On a $R_B(\Pi_{\sigma^2}) = \mathbb{E} [\rho(\Pi, T^*(\mathbf{Y}) \mid \mathbf{Y})] = \frac{1}{\sigma^{-2} + \|x\|^2}$.

(d) On suppose que $\|x\|^2 > 0$ et l'on considère l'estimateur $T(\mathbf{Y}) = \frac{\langle x, \mathbf{Y} \rangle}{\|x\|^2}$.

i. Déterminer la fonction de risque de T (toujours pour la perte quadratique). On pourra utiliser le fait que sous $\mathbb{P}_\theta$, $\mathbf{Y} \sim x\theta + \varepsilon$ avec $\varepsilon \sim \mathcal{N}(0, 1)^{\otimes n}$.

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

$$\begin{aligned} \mathbf{R}(\theta, T) &= \mathbb{E}_\theta \left[\left(\frac{\langle x, \mathbf{Y} \rangle}{\|x\|^2} - \theta \right)^2 \right] \\ &= \frac{1}{\|x\|^4} \mathbb{E}_\theta \left[(\langle x, x\theta + \varepsilon \rangle - \|x\|^2 \theta)^2 \right] \\ &= \frac{1}{\|x\|^4} \mathbb{E} \left[\left(\sum_{i=1}^n x_i \varepsilon_i \right)^2 \right] \\ &= \frac{1}{\|x\|^4} \sum_{i=1}^n x_i^2 \\ &= \frac{1}{\|x\|^2}, \end{aligned}$$

où l'on a utilisé le fait que, sous $\mathbb{P}_\theta$, $\mathbf{Y} - \theta x \sim \varepsilon \sim \mathcal{N}(0, 1)^{\otimes n}$.

ii. L'estimateur T est-il minimax ?

Correction. Comme la fonction de risque de T est constante, on a $\mathbf{R}_{\max}(T) = \frac{1}{\|x\|^2}$. De plus, en considérant la suite de lois a priori $(\Pi_k = \mathcal{N}(0, k))_{k \geq 1}$, on a

$$\lim_{k \rightarrow +\infty} \mathbf{R}_B(\Pi_k) = \lim_{k \rightarrow +\infty} \frac{1}{\frac{1}{k} + \|x\|^2} = \frac{1}{\|x\|^2} = \mathbf{R}_{\max}(T).$$

Par un résultat du cours, T est minimax.

4. Soit $\theta_0 \in \mathbb{R}$. Dans cette question, on suppose que les réels x_i sont tels que $\|x\|^2$ tend vers $+\infty$ quand $n \rightarrow +\infty$, et l'on s'intéresse à la consistance de la loi a posteriori en θ_0 .

(a) Montrer que pour tout $\delta > 0$, on a

$$\mathbb{P}(|\theta - \theta_0| \geq \delta \mid \mathbf{Y}) \leq \frac{1}{\delta^2} (v_{\mathbf{Y}} + (m_{\mathbf{Y}} - \theta_0)^2).$$

Correction. Par l'inégalité de Markov appliquée au carré, on a

$$\begin{aligned} \mathbb{P}(|\theta - \theta_0| \geq \delta \mid \mathbf{Y}) &\leq \frac{1}{\delta^2} \mathbb{E}[(\theta - \theta_0)^2 \mid \mathbf{Y}] \\ &= \frac{1}{\delta^2} \mathbb{E}[(\theta - m_{\mathbf{Y}})^2 + (m_{\mathbf{Y}} - \theta_0)^2 + 2(\theta - m_{\mathbf{Y}})(m_{\mathbf{Y}} - \theta_0) \mid \mathbf{Y}] \\ &= \frac{1}{\delta^2} (v_{\mathbf{Y}} + (m_{\mathbf{Y}} - \theta_0)^2), \end{aligned}$$

où l'on a utilisé le fait que $\mathbb{E}[(\theta - m_{\mathbf{Y}})(m_{\mathbf{Y}} - \theta_0) \mid \mathbf{Y}] = 0$.

(b) En utilisant l'inégalité de Tchebychev, montrer que

$$\frac{\langle x, \varepsilon \rangle}{\|x\|^2} \xrightarrow{\mathbb{P}} 0.$$

Correction. Soit $\delta > 0$. Comme la variable $\langle x, \varepsilon \rangle$ est centrée, on a, par l'inégalité de Tchebychev,

$$\mathbb{P} \left(\frac{|\langle x, \varepsilon \rangle|}{\|x\|^2} \geq \delta \right) \leq \frac{1}{\delta^2 \|x\|^4} \text{Var} \left(\sum_{i=1}^n x_i \varepsilon_i \right) = \frac{1}{\delta^2 \|x\|^2} \xrightarrow{n \rightarrow \infty} 0,$$

puisque $\|x\|^2 \rightarrow +\infty$ par hypothèse.

(c) En déduire que $m_{\mathbf{Y}} \xrightarrow{\mathbb{P}_{\theta_0}} \theta_0$, et conclure.

Correction. Sous $\mathbb{P}_{\theta_0}$, on a $\mathbf{Y} \sim x\theta_0 + \varepsilon$. Ainsi,

$$m_{\mathbf{Y}} - \theta_0 \sim \frac{\langle x, x\theta_0 + \varepsilon \rangle}{\sigma^{-2} + \|x\|^2} - \theta_0 = \frac{\langle x, \varepsilon \rangle - \sigma^{-2}\theta_0}{\sigma^{-2} + \|x\|^2}.$$

En utilisant la question précédente, le fait que $\|x\|^2$, et la continuité, on obtient bien $m_{\mathbf{Y}} - \theta_0 \xrightarrow{\mathbb{P}} 0$. De plus, $v_{\mathbf{Y}} = \frac{1}{\sigma^{-2} + \|x\|^2} \xrightarrow{n \rightarrow \infty} 0$. Ainsi, l'inégalité de la question 4(a) donne

$$\mathbb{P}(|\theta - \theta_0| \geq \delta \mid \mathbf{Y}) \xrightarrow{\mathbb{P}_{\theta_0}} 0.$$

La loi a posteriori est donc consistante en θ_0 .

EXERCICE 3 (2 points)

On considère le modèle Poisson $\mathcal{P} = (\mathcal{P}(\theta)^{\otimes n})_{\theta > 0}$, et l'on souhaite tester

$$H_0 : \theta = 1 \quad \text{contre} \quad H_1 : \theta \neq 1.$$

Pour cela, on considère la perte équilibrée et la loi a priori

$$\Pi = \alpha \delta_1 + (1 - \alpha) \text{Exp}(1),$$

avec $\alpha \in]0, 1[$.

1. Calculer $\Pi(\{1\} \mid \mathbf{X})$.

Correction.

$$\begin{aligned} \Pi(\{1\} \mid \mathbf{X}) &= \frac{\alpha \prod_{i=1}^n \left(\frac{e^{-1} 1^{X_i}}{X_i!} \right)}{\alpha \prod_{i=1}^n \left(\frac{e^{-1} 1^{X_i}}{X_i!} \right) + (1 - \alpha) \int_{\mathbb{R}_+^* \setminus \{1\}} e^{-\theta} \prod_{i=1}^n \left(\frac{e^{-\theta} \theta^{X_i}}{X_i!} \right) d\theta} \\ &= \frac{\alpha e^{-n}}{\alpha e^{-n} + (1 - \alpha) \int_0^{+\infty} \theta^{n\bar{X}_n} e^{-(n+1)\theta} d\theta} \\ &= \frac{\alpha e^{-n}}{\alpha e^{-n} + (1 - \alpha) \frac{\Gamma(n\bar{X}_n + 1)}{(n+1)^{n\bar{X}_n + 1}}}. \end{aligned}$$

2. Montrer qu'un test de Bayes est donné par

$$\varphi^*(\mathbf{X}) = \mathbb{1} \left\{ \frac{(n\bar{X}_n)! e^n}{(n+1)^{n\bar{X}_n + 1}} \geq c_\alpha \right\},$$

où c_α est une constante dépendant uniquement de α que vous préciserez.

Correction. On sait par le cours qu'un test de Bayes est donné par $\varphi^*(\mathbf{X}) = \mathbb{1}_{\Pi(\{1\}|\mathbf{X}) \leq 1/2}$. Or

$$\begin{aligned}\Pi(\{1\}|\mathbf{X}) \leq \frac{1}{2} &\Leftrightarrow \alpha e^{-n} \leq (1-\alpha) \frac{\Gamma(n\bar{X}_n + 1)}{(n+1)^{n\bar{X}_n+1}} \\ &\Leftrightarrow \frac{(n\bar{X}_n)! e^n}{(n+1)^{n\bar{X}_n+1}} \geq \frac{\alpha}{1-\alpha}.\end{aligned}$$

On a donc bien la forme souhaitée avec $c_\alpha = \frac{\alpha}{1-\alpha}$.