Feuille 6 : correction

Exercice 1 (Information de Fisher et inégalité de Cramér-Rao)

1. Remarquons déjà qu'en utilisant la propriété de l'énoncé avec $T(X) = 1$, on obtient $\mathbb{E}_\theta[s_\theta(X)] = 0$. Ainsi,
$$\mathbb{E}_\theta\left[(T(X) - \psi(\theta))\, s_\theta(X)\right] = \mathbb{E}_\theta\left[T(X)s_\theta(X)\right] = \psi'(\theta)\,.$$
En utilisant l'inégalité de Cauchy-Schwarz, on a
$$\psi'(\theta) \leq \sqrt{\mathbb{E}_\theta\left[(T(X) - \psi(\theta))^2\right]\mathbb{E}_\theta\left[s_\theta(X)^2\right]} = \sqrt{\mathrm{Var}_\theta\left(T(X)\right)\mathcal{I}(\theta)}\,,$$
ce qui donne bien la borne de Carmér-Rao.
2. L'EMV $\widehat{\theta}_n(\mathbf{X})$ annule le score : $s_{\widehat{\theta}_n(\mathbf{X})}(\mathbf{X}) = 0$. Comme la fonction $\theta \mapsto s_\theta(\mathbf{X})$ est $\mathcal{C}^1$ par hypothèse, la formule de Taylor-Lagrange garantit l'existence de $\widehat{\eta}_n$ entre θ et $\widehat{\theta}_n$ tel que
$$0 = s_{\widehat{\theta}_n(\mathbf{X})}(\mathbf{X}) = s_\theta(\mathbf{X}) + \left(\widehat{\theta}_n(\mathbf{X}) - \theta\right)\frac{\partial}{\partial\theta}s_{\widehat{\eta}_n}(\mathbf{X})\,.$$
3. On a $s_\theta(\mathbf{X}) = \sum_{i=1}^n s_\theta(X_i)$. Les variables $s_\theta(X_i)$ sont i.i.d., centrées, et de variance $\mathcal{I}(\theta)$. Par le TCL, on a
$$\frac{s_\theta(\mathbf{X})}{\sqrt{n}} \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, \mathcal{I}(\theta)\right)\,.$$
De plus,
$$\frac{1}{n}\frac{\partial}{\partial\theta}s_{\widehat{\eta}_n}(\mathbf{X}) \xrightarrow{\mathbb{P}} \mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}s_\theta(X)\right]\,.$$
Comme $s_\theta(x) = \frac{\frac{\partial}{\partial\theta}f_\theta(x)}{f_\theta(x)}$, on a
$$\mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}s_\theta(X)\right] = \mathbb{E}_\theta\left[\frac{\left(\frac{\partial^2}{\partial\theta^2}f_\theta(X)\right)f_\theta(X) - \left(\frac{\partial}{\partial\theta}f_\theta(X)\right)^2}{f_\theta(X)^2}\right] = \int_E \frac{\partial^2}{\partial\theta^2}f_\theta(x)d\mu(x) - \mathcal{I}(\theta)\,.$$
Par le théorème de dérivabilité de Lebesgue,
$$\int_E \frac{\partial^2}{\partial\theta^2}f_\theta(x)d\mu(x) = \frac{\partial}{\partial\theta}\int_E\frac{\partial}{\partial\theta}f_\theta(x)d\mu(x) = 0\,,$$
car la fonction $\theta \mapsto \int_E \frac{\partial}{\partial\theta}f_\theta(x)d\mu(x)$ est identiquement nulle.
En utilisant le lemme de Slutsky, on obtient donc
$$\sqrt{n}\left(\widehat{\theta}_n(\mathbf{X}) - \theta\right) \overset{\mathcal{L}}{\rightsquigarrow} \frac{1}{\mathcal{I}(\theta)}\mathcal{N}\left(0, \mathcal{I}(\theta)\right) = \mathcal{N}\left(0, \mathcal{I}(\theta)^{-1}\right)\,.$$

Exercice 2 (EMV et régression logistique)

1. Conditionnellement à X, la variable Y suit une loi de Bernoulli de paramètre
$$p_{\alpha,\beta}(X) = \frac{e^{\alpha+\beta X}}{1 + e^{\alpha+\beta X}}\,.$$
La vraisemblance s'écrit donc
$$\mathcal{L}(\alpha,\beta) = \prod_{i=1}^n f(X_i)p_{\alpha,\beta}(X_i)^{Y_i}\left(1 - p_{\alpha,\beta}(X_i)\right)^{1-Y_i} = \prod_{i=1}^n \frac{f(X_i)e^{\alpha Y_i + \beta X_i Y_i}}{1 + e^{\alpha+\beta X_i}}\,.$$

Et la log-vraisemblance :

$$\ell(\alpha,\beta) = \sum_{i=1}^{n} \left\{ \ln(f(X_i)) + \alpha Y_i + \beta X_i Y_i - \ln\left(1+e^{\alpha+\beta X_i}\right) \right\} .$$

L'EMV $(\widehat{\alpha},\widehat{\beta})$ doit vérifier

$$\begin{cases} \frac{\partial \ell(\widehat{\alpha},\widehat{\beta})}{\partial \alpha} & = \sum_{i=1}^{n} \left\{ Y_i - \frac{e^{\widehat{\alpha}+\widehat{\beta}X_i}}{1+e^{\widehat{\alpha}+\widehat{\beta}X_i}} \right\} = 0 \\ \frac{\partial \ell(\widehat{\alpha},\widehat{\beta})}{\partial \beta} & = \sum_{i=1}^{n} \left\{ X_i Y_i - \frac{X_i e^{\widehat{\alpha}+\widehat{\beta}X_i}}{1+e^{\widehat{\alpha}+\widehat{\beta}X_i}} \right\} = 0 \end{cases}$$

Ce système n'a pas de solution explicite, mais on peut l'approcher numériquement, par exemple par un algorithme de Newton.

2. La matrice d'information de Fisher s'écrit

$$\begin{aligned} \mathcal{I}(\alpha,\beta) &= \begin{pmatrix} \mathbb{E}\left[(Y - p_{\alpha,\beta}(X))^2\right] & \mathbb{E}\left[X\,(Y - p_{\alpha,\beta}(X))^2\right] \\ \mathbb{E}\left[X\,(Y - p_{\alpha,\beta}(X))^2\right] & \mathbb{E}\left[X^2\,(Y - p_{\alpha,\beta}(X))^2\right] \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{E}\left[p_{\alpha,\beta}(X)\,(1-p_{\alpha,\beta}(X))\right] & \mathbb{E}\left[X p_{\alpha,\beta}(X)\,(1-p_{\alpha,\beta}(X))\right] \\ \mathbb{E}\left[X p_{\alpha,\beta}(X)\,(1-p_{\alpha,\beta}(X))\right] & \mathbb{E}\left[X^2 p_{\alpha,\beta}(X)\,(1-p_{\alpha,\beta}(X))\right] , \end{pmatrix} \end{aligned}$$

qui est inversible dès que X est non-constante.

3. On sait que

$$\sqrt{n}\left(\widehat{\beta}-\beta\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0,\sigma^2_{\alpha,\beta}\right) ,$$

où $\sigma^2_{\alpha,\beta}$ est le deuxième coefficient diagonal de la matrice $\mathcal{I}(\alpha,\beta)^{-1}$. En utilisant le lemme de Slutsky, on a donc

$$\frac{\sqrt{n}\left(\widehat{\beta}-\beta\right)}{\sigma_{\widehat{\alpha},\widehat{\beta}}} \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0,1) .$$

Cela conduit à rejetter H_0 lorsque

$$|\widehat{\beta}| > \frac{\sigma_{\widehat{\alpha},\widehat{\beta}}}{\sqrt{n}} q_\alpha ,$$

avec q_α le quantile d'ordre $1-\alpha/2$ d'une loi normale standard.

4. On trouve $\widehat{\alpha} = 15,04$ et $\widehat{\beta} = -0,2322$. La p-valeur pour le test $\beta = 0$ vaut $0,0320$.
5. On a $\widehat{\alpha} + 31.\widehat{\beta} = 7,84$, ce qui donne une probabilité $\widehat{p}$ de $0,9996$.

Exercice 3 (EMV et efficacité asymptotique)

1. La log-vraisemblance s'écrit

$$\ell_\theta(\mathbf{X}) = \sum_{i=1}^{n} \left\{ \ln(X_i) - \frac{X_i^2}{2}\ln(\theta) + \ln\ln(\theta) \right\} ,$$

L'annulation de la dérivée donne

$$\widehat{\theta}_n = \exp\left(\frac{2n}{\sum X_i^2}\right) .$$

2. (a) Par la loi des grands nombres,

$$\frac{1}{2n}\sum_{i=1}^{n} X_i^2 \xrightarrow{\text{p.s.}} \mathbb{E}_\theta[X^2] = \frac{1}{\ln(\theta)} ,$$

et par le TCL,

$$\sqrt{n}\left(\frac{1}{2n}\sum_{i=1}^{n} X_i^2 - \frac{1}{\ln(\theta)}\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0,\frac{1}{\ln(\theta)^2}\right) .$$

(b) On utilise la méthode delta avec $g : x \mapsto e^{1/x}$ (sur $\mathbb{R}_+^*$), de dérivée $g'(x) = -\frac{1}{x^2}e^{1/x}$. On obtient

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, \theta^2 \ln(\theta)^2\right) .$$

3. Le score est donné par

$$s_\theta(X) = \frac{1}{\theta \ln(\theta)} - \frac{X^2}{2\theta} .$$

L'information de Fisher vaut donc

$$\mathcal{I}(\theta) = \frac{1}{4\theta^2}\mathbb{E}_\theta\left[\left(X^2 - \frac{2}{\ln(\theta)}\right)^2\right] = \frac{1}{4\theta^2}\operatorname{Var}_\theta(X^2) = \frac{1}{\theta^2 \ln(\theta)^2} .$$

L'estimateur $\widehat{\theta}_n$ est bien efficace.

4. En effectuant le changement de variable $x\sqrt{\ln(\theta)} = z$, on a

$$\begin{aligned}
\mathbb{E}_\theta[X] &= \ln(\theta)\int_0^{+\infty} x^2\theta^{-x^2/2}dx \\
&= \ln(\theta)\int_0^{+\infty} x^2 e^{-\frac{x^2\ln(\theta)}{2}}dx \\
&= \frac{1}{\sqrt{\ln(\theta)}}\int_0^{+\infty} z^2 e^{-z^2/2} \\
&= \sqrt{\frac{\pi}{2\ln(\theta)}} .
\end{aligned}$$

Cela conduit à proposer l'estimateur $\widetilde{\theta}_n = \exp\left(\frac{\pi}{2\overline{X}_n}\right)$. Par la loi des grands nombres et le théorème de continuité, $\widetilde{\theta}_n$ est consistant. Pour la loi limite, le TCL donne

$$\sqrt{n}\left(\overline{X}_n - \sqrt{\frac{\pi}{2\ln(\theta)}}\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, \frac{2-\pi/2}{\ln(\theta)}\right),$$

et la méthode delta avec la fonction $g : x \mapsto \exp\left(\frac{\pi}{2x^2}\right)$, de dérivée $g'(x) = -\pi x^{-3}\exp\left(\frac{\pi}{2x^2}\right)$, donne

$$\sqrt{n}\left(\widetilde{\theta}_n - \theta\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, 8\left(\frac{2}{\pi} - \frac{1}{2}\right)\theta^2\ln(\theta)^2\right) .$$

Comme $8\left(\frac{2}{\pi} - \frac{1}{2}\right) > 1$, l'estimateur $\widetilde{\theta}_n$ a une plus grande variance que $\widehat{\theta}_n$.

Exercice 4 (EMV et modèle gaussien)

1. La log-vraisemblance s'écrit

$$\ell_{\mu,\sigma^2}(X) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(X-\mu)^2}{2\sigma^2} .$$

On a

$$\frac{\partial \ell_{\mu,\sigma^2}}{\partial\mu}(X) = \frac{X-\mu}{\sigma^2} ,$$

et

$$\frac{\partial \ell_{\mu,\sigma^2}}{\partial\sigma^2}(X) = \frac{1}{2\sigma^2}\left(\left(\frac{X-\mu}{\sigma}\right)^2 - 1\right) .$$

En utilisant que $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$, on obtient

$$\mathcal{I}(\mu,\sigma^2) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix} .$$

2. L'EMV est donné par $\widehat{\mu}_n = \overline{X}_n$ et $\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$. On sait que $\frac{n\widehat{\sigma}_n^2}{\sigma^2} \sim \chi^2(n-1)$. Ainsi $\mathbb{E}[\widehat{\sigma}_n^2] = \left(1-\frac{1}{n}\right)\sigma^2$ et
$$\mathrm{Var}(\widehat{\sigma}_n^2) = \frac{2\sigma^4}{n}\left(1-\frac{1}{n}\right) < \frac{1}{n\mathcal{I}(\sigma^2)}\,.$$
Ici, la variance est plus petite que l'inverse de l'information de Fisher. Cela est possible car il s'agit d'un estimateur biaisé.

3. L'estimateur $\widetilde{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ est sans biais. Son risque quadratique, égal à sa variance, vaut $\frac{2\sigma^4}{n-1} > \frac{1}{n\mathcal{I}(\sigma^2)}$ (comme annoncé par la borne de Cramér-Rao). Quant au risque de $\widehat{\sigma}_n^2$, il vaut
$$\frac{2\sigma^4}{n}\left(1-\frac{1}{n}\right) + \frac{\sigma^4}{n^2} = \frac{\sigma^4(2n-1)}{n^2} < \frac{2\sigma^4}{n-1}\,.$$
L'EMV a donc un plus petit risque quadratique.

Exercice 5 (Test du rapport de vraisemblance)

1. Le test du rapport de vraisemblance s'écrit $\varphi^*(\mathbf{X}) = \mathbb{1}_{\{\Lambda(\mathbf{X})\geq c_\alpha\}}$, où, pour $x \in \{0,1\}^n$,
$$\Lambda(x) = \prod_{i=1}^n \frac{\theta_1^{x_i}(1-\theta_1)^{1-x_i}}{\theta_0^{x_i}(1-\theta_0)^{1-x_i}} = \left(\frac{1-\theta_1}{1-\theta_0}\right)^n \left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)^{\sum_{i=1}^n x_i}\,,$$
et où c_α est tel que $\mathbb{P}_{\theta_0}\left(\Lambda(\mathbf{X}) \geq c_\alpha\right) \leq \alpha$. En prenant le log, on voit que le test s'écrit $\varphi^*(\mathbf{X}) = \mathbb{1}\{n\overline{X}_n \geq q_\alpha\}$ où q_α est tel que $\mathbb{P}_{\theta_0}\left(n\overline{X}_n \geq q_\alpha\right) \leq \alpha$. Pour fixer les idées, on peut prendre q_α comme l'unique entier k tel que $\mathbb{P}_{\theta_0}\left(n\overline{X}_n \geq k\right) \leq \alpha$ et $\mathbb{P}_{\theta_0}\left(n\overline{X}_n \geq k-1\right) > \alpha$ (en particulier, q_α ne dépend pas de θ_1, donc φ^* non plus). Montrons que φ^* est UPP. Soit φ un test tel que $\mathbb{E}_{\theta_0}[\varphi(\mathbf{X})] \leq \mathbb{E}_{\theta_0}[\varphi^*(\mathbf{X})]$. On a
$$\begin{aligned}
\mathbb{E}_{\theta_1}\left[\varphi^*(\mathbf{X}) - \varphi(\mathbf{X})\right] &= \mathbb{E}_{\theta_0}\left[\frac{p_{\theta_1}(\mathbf{X})}{p_{\theta_0}(\mathbf{X})}\left(\varphi^*(\mathbf{X}) - \varphi(\mathbf{X})\right)\right] \\
&= \mathbb{E}_{\theta_0}\left[\left(\frac{p_{\theta_1}(\mathbf{X})}{p_{\theta_0}(\mathbf{X})} - c_\alpha\right)\left(\varphi^*(\mathbf{X}) - \varphi(\mathbf{X})\right)\right] + c_\alpha \mathbb{E}_{\theta_0}\left[\varphi^*(\mathbf{X}) - \varphi(\mathbf{X})\right] \\
&\geq \mathbb{E}_{\theta_0}\left[\left(\frac{p_{\theta_1}(\mathbf{X})}{p_{\theta_0}(\mathbf{X})} - c_\alpha\right)\left(\varphi^*(\mathbf{X}) - \varphi(\mathbf{X})\right)\right] \\
&\geq 0\,,
\end{aligned}$$
car $\left(\frac{p_{\theta_1}(\mathbf{X})}{p_{\theta_0}(\mathbf{X})} - c_\alpha\right)\left(\varphi^*(\mathbf{X}) - \varphi(\mathbf{X})\right) \geq 0$.

2. Vérifions déjà que le test φ^* de la question 1 est UPP pour ce problème. Supposons φ^* non UPP. Alors il existe φ tel que $\mathbb{E}_{\theta_0}[\varphi] \leq \mathbb{E}_{\theta_0}[\varphi^*]$ et il existe $\theta_1 > \theta_0$ tel que $\mathbb{E}_{\theta_1}[\varphi] > \mathbb{E}_{\theta_1}[\varphi^*]$. Mais alors φ^* n'est pas UPP pour tester θ_0 contre θ_1, ce qui est en contradiction avec le résultat de la question précédente (rappelons que φ^* ne dépend pas de θ_1). Montrons maintenant que le test basé sur $\frac{\sup_{\theta>\theta_0} p_\theta(\mathbf{X})}{p_{\theta_0}(\mathbf{X})}$ est UPP. Observons que
$$\frac{\sup_{\theta>\theta_0} p_\theta(\mathbf{X})}{p_{\theta_0}(\mathbf{X})} = \mathbb{1}_{\overline{X}_n \leq \theta_0} + \frac{p_{\overline{X}_n}(\mathbf{X})}{p_{\theta_0}(\mathbf{X})}\mathbb{1}_{\overline{X}_n > \theta_0} = 1 + \left(\frac{\overline{X}_n^{n\overline{X}_n}(1-\overline{X}_n)^{n-n\overline{X}_n}}{\theta_0^{n\overline{X}_n}(1-\theta_0)^{n-n\overline{X}_n}} - 1\right)\mathbb{1}_{\overline{X}_n > \theta_0}\,.$$
Vue comme fonction de $\overline{X}_n$, cette fonction vaut 1 pour $\overline{X}_n < \theta_0$ puis croît strictement ensuite. (Remarquons que pour obtenir une erreur de première espèce strictement plus petite que 1, il faut comparer $\overline{X}_n$ à un seuil strictement plus grand que θ_0 et que l'on obtient alors automatiquement une erreur de première espèce inférieure à 1/2.) On tombe donc sur un test de la forme de celui de la quesion 1, qui est UPP.

3. Soit φ^* le test de la question 1. On a $\sup_{\theta\leq\theta_0}\mathbb{P}_\theta(n\overline{X}_n \geq q_\alpha) = \mathbb{P}_{\theta_0}(n\overline{X}_n \geq q_\alpha) \leq \alpha$. De plus, on sait que pour tout test φ et pour tout $\theta_1 > \theta_0$, $\mathbb{E}_{\theta_1}\varphi^* \geq \mathbb{E}_{\theta_1}\varphi$. Donc φ^* est UPP.

4. Montrons qu'il n'existe pas de test UPP pour $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$. Soit φ un test tel que $\mathbb{E}_{\theta_0}\varphi = \alpha$, et, pour $\theta_1 > \theta_0$, soit φ^* le test du rapport de vraisemblance de $\theta = \theta_0$ contre $\theta = \theta_1$ de même niveau α (mais il se peut qu'on ne puisse pas avoir un niveau exactement égal à α mais seulement un

peu en-dessous). Au vu de la preuve de la question 1 (en omettant la différence des erreurs de première espèce qui peut être légèrement négative...), on a $\mathbb{E}_{\theta_1}\varphi^* > \mathbb{E}_{\theta_1}\varphi$, à moins que φ ne soit égal à φ^* (on peut s'arranger pour prendre c_α de telle sorte que la probabilité que le rapport de vraisemblance soit égal à c_α soit nulle). Mais si $\varphi = \varphi^*$, alors pour $\theta_2 < \theta_0$, $\mathbb{E}_\theta\varphi^* < \mathbb{E}_\theta\varphi^{**}$, pour φ^{**} le test du rapport de vraisemblance de $\theta = \theta_0$ contre $\theta = \theta_2$ de même niveau que φ^*.

Exercice 6 (Distance de Hellinger pour le modèle de Poisson)

1. La distance de Hellinger entre deux lois de probabilité P et Q, de densités respectives p et q par rapport à une mesure μ, est définie par

$$h(P,Q) = \left(\int_E \left(\sqrt{p(x)} - \sqrt{q(x)}\right)^2 d\mu(x)\right)^{1/2} .$$

En développant, on obtient

$$h^2(P,Q) = 2\left(1 - \int_E \sqrt{p(x)q(x)}d\mu(x)\right) .$$

L'intégrale dans le terme de droite ci-dessus s'appelle l'affinité de Hellinger. Dans le cas de deux lois de Poisson P_θ et P_λ, l'affinité s'écrit

$$\sum_{k\geq 0}\sqrt{\frac{e^{-\theta}\theta^k}{k!}\cdot\frac{e^{-\lambda}\lambda^k}{k!}} = e^{-\frac{\theta+\lambda}{2}}\sum_{k\geq 0}\frac{(\sqrt{\theta\lambda})^k}{k!} = e^{-\frac{1}{2}(\theta+\lambda-2\sqrt{\theta\lambda})} = e^{-\frac{1}{2}(\sqrt{\theta}-\sqrt{\lambda})^2} .$$

2. En utilisant que $\sqrt{1+x} = 1 + \frac{x}{2} + o(x)$ quand $x \to 0$, on a

$$e^{-\frac{1}{2}(\sqrt{\theta+\varepsilon_n}-\sqrt{\theta})^2} = e^{-\frac{\theta}{2}\left(\sqrt{1+\frac{\varepsilon_n}{\theta}}-1\right)^2} = e^{-\frac{\varepsilon_n^2+o(\varepsilon_n^2)}{8\theta}} = 1 - \frac{\varepsilon_n^2}{8\theta} + o(\varepsilon_n^2) .$$

Ainsi $h^2(P_\theta, P_{\theta+\varepsilon_n}) = \frac{\varepsilon_n^2}{4\theta} + o(\varepsilon_n^2)$.

3. On a

$$\begin{aligned}
\ell_{\theta+\varepsilon_n}(\mathbf{X}) - \ell_\theta(\mathbf{X}) &= \sum_{i=1}^n\left\{\ln\left(\frac{e^{-(\theta+\varepsilon_n)}(\theta+\varepsilon_n)^{X_i}}{X_i!}\right) - \ln\left(\frac{e^{-\theta}\theta^{X_i}}{X_i!}\right)\right\} \\
&= \ln\left(1+\frac{\varepsilon_n}{\theta}\right)\sum_{i=1}^n X_i - n\varepsilon_n \\
&= \left(\frac{\varepsilon_n}{\theta} - \frac{\varepsilon_n^2}{2\theta^2} + o(\varepsilon_n^2)\right)\sum_{i=1}^n X_i - n\varepsilon_n \\
&= \frac{n\varepsilon_n}{\theta}(\overline{X}_n - \theta) - (1+o(1))\,\frac{n\varepsilon_n^2}{2\theta^2}\overline{X}_n .
\end{aligned}$$

On voit qu'en prenant $\varepsilon_n = \frac{1}{\sqrt{n}}$, on obtient une limite non-triviale :

$$\overline{X}_n \xrightarrow{\mathbb{P}} \theta \qquad \text{et} \qquad \sqrt{n}(\overline{X}_n - \theta) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0,\theta) .$$

Par le lemme de Slutsky, on obtient

$$\ell_{\theta+1/\sqrt{n}}(\mathbf{X}) - \ell_\theta(\mathbf{X}) \overset{\mathcal{L}}{\rightsquigarrow} \frac{1}{\theta}\mathcal{N}(0,\theta) - \frac{1}{2\theta} = \mathcal{N}\left(-\frac{1}{2\theta}, \frac{1}{\theta}\right) .$$

4. Notons $Y_i = \ln\left(\frac{p_{\theta_1}(X_i)}{p_{\theta_0}(X_i)}\right)$. On sait qu'un test optimal est donné par le test du rapport de vraisemblance qui s'écrit

$$\varphi(\mathbf{X}) = \mathbb{1}_{\left\{\sum_{i=1}^n Y_i \geq c_n\right\}} ,$$

pour un seuil c_n à déterminer (on peut déjà dire que c_n doit être positif). Par l'inégalité de Markov, on a, pour tout $\lambda > 0$,

$$\mathbb{P}_{\theta_0}\left(\varphi(\mathbf{X}) = 1\right) \leq e^{-\lambda c_n}\mathbb{E}_{\theta_0}\left[e^{\lambda\sum_{i=1}^n Y_i}\right] = e^{-\lambda c_n}\mathbb{E}_{\theta_0}\left[e^{\lambda Y}\right]^n .$$

On peut remarquer que, pour $\lambda = \frac{1}{2}$,

$$\mathbb{E}_{\theta_0}\left[e^{Y/2}\right] = \mathbb{E}_{\theta_0}\left[\sqrt{\frac{p_{\theta_1}(X_1)}{p_{\theta_0}(X_1)}}\right] = \sum_{k\geq 0}\sqrt{p_{\theta_1}(k)p_{\theta_0}(k)}\,.$$

On reconnaît l'affinité de Hellinger, soit $1 - \frac{h^2(P_{\theta_0},P_{\theta_1})}{2}$. Ainsi,

$$\mathbb{P}_{\theta_0}\left(\varphi(\mathbf{X}) = 1\right) \leq e^{-c_n/2}\left(1 - \frac{h^2(P_{\theta_0},P_{\theta_1})}{2}\right)^n \leq e^{-\frac{1}{2}\left(c_n + nh^2(P_{\theta_0},P_{\theta_1})\right)}\,.$$

Pour l'erreur de seconde espèce, on a, par le même argument,

$$\mathbb{P}_{\theta_1}\left(\varphi(\mathbf{X}) = 0\right) = \mathbb{P}_{\theta_1}\left(-\sum_{i=1}^n Y_i \geq -c_n\right) \leq e^{c_n/2}\mathbb{E}_{\theta_1}\left[e^{-Y/2}\right]^n\,.$$

Or $\mathbb{E}_{\theta_1}\left[e^{-Y/2}\right]$ est aussi égal à l'affinité de Hellinger et l'on obtient

$$\mathbb{P}_{\theta_1}\left(\varphi(\mathbf{X}) = 0\right) \leq e^{c_n/2}\left(1 - \frac{h^2(P_{\theta_0},P_{\theta_1})}{2}\right)^n \leq e^{-\frac{1}{2}\left(nh^2(P_{\theta_0},P_{\theta_1}) - c_n\right)}\,.$$

Pour que les deux erreurs soient majorées par α, il suffit que $nh^2(P_{\theta_0},P_{\theta_1}) - c_n$ et $nh^2(P_{\theta_0},P_{\theta_1}) + c_n$ soient plus grands que $2\ln(1/\alpha)$. Si l'on prend $c_n = \frac{1}{2}nh^2(P_{\theta_0},P_{\theta_1})$, on voit que, dès que $h^2(P_{\theta_0},P_{\theta_1})$ est plus grand que $\frac{4}{n}\ln(1/\alpha)$, c'est bon. Et pour cela, il suffit que ε_n ne décroisse pas plus rapidement que $1/\sqrt{n}$, d'après la question 2.
Inversement, si $\varepsilon_n = o\left(\frac{1}{\sqrt{n}}\right)$, alors, en utilisant le fait que la distance en variation totale est inférieure à la distance de Hellinger, et que l'affinité de Hellinger entre $P_{\theta_0}^{\otimes n}$ et $P_{\theta_1}^{\otimes}$ est égale à l'affinité de Hellinger entre P_{θ_0} et P_{θ_1} à la puissance n, on a, pour tout test φ,

$$\begin{aligned}
\mathbb{P}_{\theta_1}\left(\varphi(\mathbf{X}) = 1\right) - \mathbb{P}_{\theta_0}\left(\varphi(\mathbf{X}) = 1\right) &\leq d_{\mathrm{VT}}(P_{\theta_0}^{\otimes n}, P_{\theta_1}^{\otimes n}) \\
&\leq h(P_{\theta_0}^{\otimes n}, P_{\theta_1}^{\otimes n}) \\
&= \sqrt{2\left(1 - \left(1 - \frac{h^2(P_{\theta_0},P_{\theta_1})}{2}\right)^n\right)} \\
&\leq \sqrt{nh^2(P_{\theta_0},P_{\theta_1})} = o(1)\,.
\end{aligned}$$

Il ne sera donc pas possible d'obtenir un seuil $\alpha < 1/2$ pour les deux types d'erreurs.

Exercice 7 (Tableau de contingence)

1. La vraisemblance sous $\alpha = (\alpha_1, \ldots, \alpha_K)$ et $\beta = (\beta_1, \ldots, \beta_K)$ s'écrit

$$\mathcal{L}_{\alpha,\beta}(\mathbf{X},\mathbf{Y}) = \prod_{i=1}^n\prod_{k=1}^K \alpha_k^{\mathbb{1}_{X_i=k}}\beta_k^{\mathbb{1}_{Y_i=k}} = \prod_{k=1}^K \alpha_k^{n_k}\beta_k^{m_k}\,,$$

d'où pour la log-vraisemblance

$$\ell_{\alpha,\beta}(\mathbf{X},\mathbf{Y}) = \sum_{k=1}^K\left\{n_k\ln(\alpha_k) + m_k\ln(\beta_k)\right\}\,.$$

En écrivant que $\alpha_K = 1 - \alpha_1 - \cdots - \alpha_{K-1}$, l'annulation des dérivées partielles par rapport à $\alpha_1, \ldots, \alpha_{K-1}$ donne

$$\frac{\alpha_1}{n_1} = \cdots = \frac{\alpha_K}{n_K}\,,$$

Comme $\sum_{k=1}^K n_k = n$, on obtient finalement que l'EMV pour α est $\widehat{\alpha} = \left(\frac{n_1}{n}, \ldots, \frac{n_K}{n}\right)$. De même, l'EMV pour β est $\widehat{\beta} = \left(\frac{m_1}{n}, \ldots, \frac{m_K}{n}\right)$.

2. Dans le cas général, la vraisemblance sous $\theta = (\theta_{k\ell})$ s'écrit
$$\mathcal{L}_\theta(\mathbf{X}, \mathbf{Y}) = \prod_{i=1}^{n} \prod_{k=1}^{K} \prod_{\ell=1}^{K} \theta_{k\ell}^{\mathbb{1}_{(X_i,Y_i)=(k,\ell)}} = \prod_{k=1}^{K} \prod_{\ell=1}^{K} \theta_{k\ell}^{n_{k\ell}} ,$$
d'où pour la log-vraisemblance
$$\ell_\theta(\mathbf{X}, \mathbf{Y}) = \sum_{k=1}^{K} \sum_{\ell=1}^{K} n_{k\ell} \ln(\theta_{k\ell}) .$$
L'EMV pour $\theta_{k\ell}$ est alors donné par $\widehat{\theta}_{k\ell} = \frac{n_{k\ell}}{n}$.
3. Sous H_0 (indépendance), on peut montrer que la statistique
$$S = \sum_{k=1}^{K} \sum_{\ell=1}^{K} \frac{1}{\widehat{\alpha}_k \widehat{\beta}_\ell} \left(\sqrt{n} \left(\widehat{\theta}_{k\ell} - \widehat{\alpha}_k \widehat{\beta}_\ell \right) \right)^2$$
converge en loi vers une loi $\chi^2\left((K-1)^2\right)$. Ainsi le test consistant à rejeter H_0 quand S est plus grande que le quantile d'ordre $1-\alpha$ d'une $\chi^2\left((K-1)^2\right)$ est de niveau asymptotique $1-\alpha$.