Corrigé : Examen novembre 2017

Solution Exercice 1 1. le modèle géométrique est dominé par la mesure discrète $\mu = \sum_{n\in\mathbb{N}^*} \delta_n$.

— La densité par rapport à cette mesure est la loi donnée dans l'énoncé,

$$p_\theta(x) = \mathbb{1}_{\mathbb{N}^*}(x)\mathbb{P}_\theta\{X = x\} = \theta(1-\theta)^{x-1}, x \in \mathbb{R}.$$

NB : Puisqu'on a choisi μ ne chargeant que $\mathbb{N}^*$, on peut omettre l'indicatrice dans l'expression de p_θ.

— Pour un échantillon i.i.d. de taille n,

$$\log p_\theta^{\otimes n}(x_1,\dots,x_n) = n\log\theta + \sum_{i=1}^n (x_i - 1)\log(1-\theta).$$

2. **max de vraisemblance :** On le trouve en annulant la dérivée de la log-vraisemblance. Pour $x = (x_1,\dots,x_n)$,

$$\frac{\partial}{\partial_\theta}\log p_\theta^{\otimes n}(x) = \frac{n}{\theta} - \frac{\sum x_i - n}{1-\theta} = \frac{n - \theta\sum x_i}{\theta(1-\theta)}$$

cette quantité s'annule en $\theta = n/\sum x_i$, elle est positive à gauche et négative à droite de cette valeur, qui est donc bien un maximiseur de la vraisemblance. Ainsi $\widehat{\theta}_{MV} = n/\sum x_i$.

3. **risque quadratique** pour $g(\theta) = 1/\theta$, on considère $g_n(X) = \frac{1}{n}\sum X_i$.

— biais : on a $\mathbb{E}(g_n(X) = \mathbb{E}(X) = 1/\theta$. L'estimateur est donc non biaisé.

— variance : $\mathbb{V}\mathrm{ar} g_n(X) = \frac{1}{n}\mathrm{var} X = \frac{1-\theta}{n\theta^2}$.

Le risque vaut donc $R(g_n,\theta) = \frac{1-\theta}{n\theta^2}$.

— Efficacité : g_n étant non biaisé, il est efficace si et seulement s'il atteint la borne de Cramér-Rao $B(\theta) = g'(\theta)^2/(nI_1(\theta))$ avec I_1 l'information de Fisher pour 1 observation. On a

$$I_1(\theta) = \mathbb{E}_\theta[\partial_\theta \log p_\theta(X_1)^2]) = \mathbb{V}\mathrm{ar}_\theta[\partial_\theta \log p_\theta(X_1)]$$

car d'après le cours, l'espérance du score est nulle. D'où

$$I_1(\theta) = \mathbb{V}\mathrm{ar}\Big[\frac{1-\theta X_1}{\theta(1-\theta)}\Big] = \frac{\theta^2}{\theta^2(1-\theta)^2}\mathbb{V}\mathrm{ar}(X_1)\Big] = \frac{1}{\theta^2(1-\theta)}$$

de plus $g'(\theta)^2 = 1/\theta^4$, d'où

$$B(\theta) = \partial(1-\theta)n\theta^2 = R(g_n,\theta)$$

g_n est donc efficace.

Approche bayésienne : prior $\pi(\,\cdot\,) = \mathcal{U}_{[0,1]}$.

4. La densité a posteriori peut se calculer à une constante de normalisation près (ne dépendant pas de θ)
$$\pi(\theta|x) \propto \pi(\theta)p_\theta^{\otimes n}(x) = \mathbb{1}_{]0,1[}(\theta)\theta^n(1-\theta)^{\sum x_i - n} \propto \mathrm{beta}(\theta|a_n, b_n)$$
où $\mathrm{beta}(\theta|a,b)$ est la densité de la loi Beta e paramètres a et b, et où
$$a_n = n+1 \ ; \ b_n = \sum x_i - n + 1$$
Ainsi la loi a posteriori est une loi beta de paramètres a_n, b_n comme ci-dessus.
5. l'espérance a posteriori est
$$\mathbb{E}(\theta|X=x) = \mathbb{E}(U)$$
où $U \sim \pi(\,\cdot\,|x) = \mathcal{B}eta(a_n, b_n)$ D'après le résultat de l'encadré sur l'espérance des lois Beta, $\mathbb{E}(U) = a_n/(a_n+b_n)$, d'où
$$d'o\grave{u}\widehat{\theta}_{EP}(x_1, \ldots, x_n) = \mathbb{E}(\theta|X=x) = \frac{n+1}{\sum x_i + 2}$$
6. D'après la loi des grands nombres, si $X_i \overset{\text{i.i.d.}}{\sim} \mathcal{G}(\theta)$, on a, presque sûrement, $\frac{1}{n}\sum_{i=1}^n X_i \to \mathbb{E}_\theta(X_1) = 1/\theta$. Ainsi
$$\widehat{\theta}_{EP}(X_1, \ldots, X_n) = \frac{1+\frac{1}{n}}{\frac{1}{n}\sum X_i + \frac{2}{n}} \to \frac{1}{1/\theta} = \theta \quad \text{presque sûrement.}$$

Solution Exercice 2 [paramètre de translation]

1. l'espérance de X_1 vaut
$$\mathbb{E}_\theta(X_1) = \int_\theta^{+\infty} x e^{\theta - x}\mathrm{d}x \overset{\text{calcul simple}}{=} \theta + 1.$$
En prenant $\widehat{\theta}_n(X) = \frac{1}{n}\sum(X_i) - 1$, on a bien par linéarité de l'espérance
$$\mathbb{E}_\theta(\widehat{\theta}_n(X)) = \frac{1}{n}\sum_i \mathbb{E}(X_i) - 1 = \theta.$$
Ainsi $\widehat{\theta}_n$ est non biaisé.
2. Risque quadratique : Pour $\theta > 0$, il est donnée par
$$R(\theta, \widehat{\theta}_n) = \mathbb{E}\left[(\widehat{\theta}_n(X) - \theta)^2\right]$$
et puisque $\mathbb{E}(\widehat{\theta}_n(X)) = \theta$, on a $R(\theta, \widehat{\theta}_n) = \mathbb{V}\mathrm{ar}_\theta(\widehat{\theta}_n(X))$. De plus, pour une observation $X_1 \sim \mathrm{P}_\theta$,
$$\begin{aligned}\mathbb{V}\mathrm{ar}_\theta(X_1) &= \int_\theta^\infty (x-\theta-1)^2 e^{\theta-x}\mathrm{d}x \\ &= \int_0^\infty (t-1)^2 e^{-t}\mathrm{d}t \\ &= \left[-(t-1)^2e^{-t}\right]_0^\infty + 2\underbrace{\int_0^\infty (t-1)e^{-t}}_{0} \\ &= 1.\end{aligned}$$

Ainsi, par indépendance des X_i,

$$\mathbb{V}\mathrm{ar}_\theta(\widehat{\theta}_n(X)) = \frac{1}{n}\mathbb{V}\mathrm{ar}_\theta(X_1) = \frac{1}{n}.$$

3. On considère $\tilde{\theta}_n(X) = \min_{i=1}^n X_i$.
   Remarque : l'idée de choisir $\tilde{\theta}_n$ vient du fait que θ est la borne inférieur des $\{x : p_\theta(x) > 0\}$. Ainsi, pour $x > \theta$ on a $\mathbb{P}_\theta(X_1 \leq x) > 0$ alors que pour $x < \theta, \mathbb{P}_\theta(X \leq x) = 0$. Dans ce sens, θ est la plus petite valeur « possible » pour les X_i.
   Pour calculer la loi de $\tilde{\theta}_n(X)$, il est dans ce cas plus facile de calculer sa fonction de répartition que sa densité (à supposer que cette dernière existe). En effet, si on appelle $\tilde{F}_\theta$ cette fonction de répartition,

$$\tilde{F}_\theta(x) = \mathbb{P}_\theta(\tilde{\theta}_n(X) \leq x),$$

   on a

$$\begin{aligned} 1 - \tilde{F}_\theta(x) &= \mathbb{P}_\theta(\tilde{\theta}_n(X) > x) \\ &= \mathbb{P}_\theta\left[\cap_{i=1}^n X_i > x\right] \\ &= \prod_{i=1}^n \mathbb{P}_\theta(X_i > n) \qquad \text{(indépendance)} \\ &= \mathbb{P}_\theta(X_1 > x)^n. \end{aligned}$$

   Reste à calculer cette dernière quantité en intégrant la densité de X_1, ce qui donne

$$\forall x \geq \theta, \mathbb{P}_\theta(X_1 > x) = \int_x^\infty e^{\theta - t}\mathrm{d}t = e^{\theta - x},$$

   et pour $x \leq \theta$, $\mathbb{P}_\theta(X_1 > x) = 1$. Finalement $\mathbb{P}_\theta(X_1 > x) = e^{-\max(x-\theta,0)}$. D'où

$$\tilde{F}_\theta(x) = 1 - e^{-n\max(x-\theta,0)}.$$

4. Puisque $\tilde{F}_\theta(\theta) = 0$, on a, avec probabilité 1, $\tilde{\theta}_n(X) \geq \theta$. Ainsi,la variable aléatoire $Z = \tilde{\theta}_n(X) - \theta$ est presque sûrement positive. Comme elle n'est pas constante, sont espérance est strictement positive. D'où : $\tilde{\theta}_n$ est biaisé.
5. Le risque quadratique de $\tilde{\theta}_n$ est $R(\theta, \tilde{\theta}_n) = \mathbb{E}[(\theta - \tilde{\theta}_n(X))^2]$. On utilise le fait que pour une variable aléatoire $Y : \Omega \to \mathbb{R}^+$, $\mathbb{E}(Y) = \int_{t=0}^{+}\infty\mathbb{P}(Y > t)\mathrm{d}t$. On pose $Y = (\tilde{\theta}_n - \theta)^2$. Alors, pour $t \geq 0$, $\mathbb{P}_\theta(Y > t) = \mathbb{P}(\tilde{\theta}_n - \theta > \sqrt{t})$ car avec probabilité

1, $\tilde{\theta}_n - \theta \geq 0$. Ainsi,

$$\begin{aligned}
R(\theta, \tilde{\theta}_n) &= \mathbb{E}(Y) \\
&= \int_{t=0}^{\infty} \mathbb{P}_\theta(Y > t)\mathrm{d}t \\
&= \int_{t=0}^{\infty} \mathbb{P}_\theta(\tilde{\theta}_n - \theta > \sqrt{t})\mathrm{d}t \\
&= \int_{t=0}^{\infty} \mathbb{P}_\theta(\tilde{\theta}_n - \theta > \sqrt{t})\mathrm{d}t \\
&= \int_{t=0}^{\infty} e^{-n\sqrt{t}}\mathrm{d}t \quad \text{(cf question précédente)} \\
&= \int_{u=0}^{\infty} e^{-nu} 2u\mathrm{d}u \\
&= \frac{2}{n^2} \int_{r=0}^{\infty} e^{-r} r\mathrm{d}r \\
&= \frac{2}{n^2}
\end{aligned}$$

6. Lorsque n est plus grand que 2, on a

$$\forall \theta, \quad R(\theta, \tilde{\theta}_n) = 2/n^2 < 1/n = R(\theta, \widehat{\theta}_n).$$

Le risque du deuxième estimateur (qui est pourtant biaisé) est uniformément plus faible que celui du premier estimateur. $\tilde{\theta}_n$ est donc préférable à $\widehat{\theta}_n$.

7. le modèle n'est pas régulier car l'ensemble des points x tels que $p_\theta(x) > 0$ dépend de θ (c'est ensemble est $[\theta, +\infty[$).

8. D'après la question 3., on a $\mathbb{P}_\theta[\tilde{\theta}_n > x] = e^{-n(x-\theta)}$ pour $x > \theta$. En inversant la relation $\delta = e^{-n(x-\theta)}$ on obtient

$$x(n, \theta, \delta) = \theta + \frac{1}{n} \log(1/\delta).$$

Ainsi

$$\mathbb{P}_\theta[\tilde{\theta}_n > \theta + \frac{1}{n} \log(1/\delta)] = \delta, \quad \text{d'où}$$
$$\mathbb{P}_\theta[\theta > \tilde{\theta}_n - \frac{1}{n} \log(1/\delta)] = 1 - \delta$$

d'autre part on a $\mathbb{P}_\theta(\tilde{\theta}_n \geq \theta) = 1$. Ainsi en posant $I = [a_n, \tilde{\theta}_n]$ avec

$$a_n = \tilde{\theta}_n - \frac{1}{n} \log(1/\delta),$$

I est un intervalle de confiance pour θ de niveau $1 - \delta$. Avec $\delta = 0.05$, on obtient $a_n = \tilde{\theta}_n - log(1/0.05)/10 \approx \tilde{\theta}_n - 0.3$.