DS8 - Concours blanc - Correction (`ESSEC II 2009`)

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidat·es sont invité·es à ***encadrer*** *dans la mesure du possible les résultats de leurs calculs.*
Aucun document n'est autorisé. ***L'utilisation de toute calculatrice et de tout matériel électronique*** *est* ***interdite****. Seule l'utilisation d'une règle graduée est autorisée.*
Si au cours de l'épreuve, un candidat ou une candidate repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

On suppose, et c'est valable pour toute l'épreuve, que les librairies suivantes sont importées sous leur alias habituel :

- `import numpy as np`
- `import numpy.linalg as al`
- `import numpy.random as rd`
- `import matplotlib.pyplot as plt`
- `import pandas as pd`

Notations

- Tout au long du sujet $(\Omega, \mathcal{F}, \mathbb{P})$ désignera un espace probabilisé et les variables aléatoires utilisées seront toutes définies sur cet espace probabilisé. Sous réserve d'existence, l'espérance mathématique d'une variable aléatoire réelle X sera notée $\mathbb{E}(X)$ et sa variance sera notée $\mathbb{V}(X)$.
- Pour un événement A, on notera $\mathbb{P}_B(A)$ la probabilité conditionnelle de A sachant B où B est un événement non négligeable.

Si T est un estimateur de θ (respectivement de $g(\theta)$), on définit :

- le biais de T par

$$b_\theta(T) = \mathbb{E}(T) - \theta \qquad \big(\text{respectivement,} \qquad b_{g(\theta)}(T) = \mathbb{E}(T) - g(\theta)\big)$$

  sous réserve de l'existence de $\mathbb{E}(T)$,
- et le risque quadratique de T par

$$r_\theta(T) = \mathbb{E}\left((T-\theta)^2\right) \qquad \big(\text{respectivement,} \qquad r_{g(\theta)}(T) = \mathbb{E}\left((T-g(\theta))^2\right)\big)$$

  sous réserve de l'existence de $\mathbb{V}(T)$.

On dit que T est *sans biais* lorsque son biais est nul.
On considérera qu'un estimateur T est meilleur (au sens large) qu'un autre estimateur T' en termes de risque quadratique si le risque quadratique de T est inférieur ou égal à celui de T'.

Le sujet est composé de quatre parties. Les parties I, II, III et IV.1 sont **indépendantes**.
Il s'agit de variations autour de la notion de risque quadratique en théorie de l'estimation.

I. Premier problème d'estimation

Dans ce premier problème d'estimation, on dispose d'une seule observation notée X.
On suppose que X admet pour densité f_θ définie sur $\mathbb{R}$ par :

$$f_\theta : x \mapsto \begin{cases} \dfrac{k+1}{\theta^{k+1}}\, x^k & \text{si } x \in [0,\theta] \\ 0 & \text{sinon} \end{cases}$$

où k est un entier naturel non nul et θ un paramètre réel inconnu strictement positif que l'on souhaite estimer.

1. Montrer que f_θ est bien une densité de probabilité.

Démonstration.

- La fonction f_θ est continue sur $]-\infty, 0[$ et sur $]\theta, +\infty[$ car constante sur ces intervalles ouverts. De plus, elle est continue sur $]0,\theta[$ car elle coïncide sur cet intervalle ouvert avec la fonction $x \mapsto \dfrac{k+1}{\theta^{k+1}}\, x^k$ qui est continue sur $]0,\theta[$ car polynomiale.

$$\boxed{\text{La fonction } f_\theta \text{ est continue sur } \mathbb{R} \text{ sauf éventuellement en 0 et en 1.}}$$

- Comme $k > 0$ et $\theta > 0$, il suit que, pour tout $x \in [0,\theta]$, $f_\theta(x) \geqslant 0$. De plus, pour tout $x \notin [0,\theta]$, $f_\theta(x) = 0$.

$$\boxed{\text{Pour tout } x \in \mathbb{R},\ f_\theta(x) \geqslant 0.}$$

- Tout d'abord,

$$\begin{aligned}
\int_{-\infty}^{+\infty} f_\theta(x)\ dx &= \int_0^\theta f_\theta(x)\ dx && \textit{car } f_\theta \textit{ est nulle en dehors de } [0,\theta] \\
&= \int_0^\theta \frac{k+1}{\theta^{k+1}}\ x^k\ dx \\
&= \frac{k+1}{\theta^{k+1}} \int_0^\theta x^k\ dx
\end{aligned}$$

Ainsi, $\displaystyle\int_{-\infty}^{+\infty} f_\theta(x)\ dx$ est convergente puisque $x \mapsto x^k$ est continue sur $[0,\theta]$.
De plus,

$$\int_0^\theta x^k\ dx = \left[\frac{x^{k+1}}{k+1} \right]_0^\theta = \frac{\theta^{k+1}}{k+1}$$

$$\boxed{\int_{-\infty}^{+\infty} f_\theta(x)\ dx = \frac{k+1}{\theta^{k+1}} \frac{\theta^{k+1}}{k+1} = 1}$$

$\boxed{\text{Ceci permet de conclure que } f_\theta \text{ est une densité de probabilité.}}$

□

2. Calculer $\mathbb{E}(X)$.

Démonstration.

- La variable aléatoire X admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x f_\theta(x)\ dx$ est absolument convergente, ce qui revient à montrer la convergence pour ce calcul de moment.
- Tout d'abord :

$$\begin{aligned}
\int_{-\infty}^{+\infty} x f_\theta(x)\ dx &= \int_0^\theta x f_\theta(x)\ dx && \textit{car } f_\theta \textit{ est nulle en dehors de } [0,\theta] \\
&= \int_0^\theta x \frac{k+1}{\theta^{k+1}}\ x^k\ dx \\
&= \frac{k+1}{\theta^{k+1}} \int_0^\theta x^{k+1}\ dx
\end{aligned}$$

Ainsi, $\displaystyle\int_{-\infty}^{+\infty} x f_\theta(x)\ dx$ est convergente puisque $x \mapsto x^{k+1}$ est continue sur $[0,\theta]$. Il suit que X admet une espérance et

$$\begin{aligned}
\mathbb{E}(X) &= \frac{k+1}{\theta^{k+1}} \left[\frac{x^{k+2}}{k+2} \right]_0^\theta \\
&= \frac{k+1}{\theta^{k+1}} \frac{\theta^{k+2}}{k+2} \\
&= \frac{k+1}{k+2}\theta
\end{aligned}$$

$$\boxed{\mathbb{E}(X) = \frac{k+1}{k+2}\theta}$$

□

3. Déterminer λ_0 un réel dépendant uniquement de k tel que $\lambda_0\, X$ soit un estimateur de θ sans biais.

Démonstration. On pose $\lambda_0 = \dfrac{k+2}{k+1}$ et on note $Y = \lambda_0 X$. La variable aléatoire Y admet une espérance comme transformée affine de X.

$$\begin{aligned}
\mathbb{E}(Y) &= \mathbb{E}(\lambda_0 X) \\
&= \lambda_0 \mathbb{E}(X) && \textit{(par linéarité de l'espérance)} \\
&= \frac{k+2}{k+1}\frac{k+1}{k+2}\theta \\
&= \theta
\end{aligned}$$

$$\boxed{\text{L'expression de } \lambda_0 X \text{ ne dépend pas de } \theta \text{ et } \mathbb{E}(\lambda_0 X) = \theta \text{ donc } \lambda_0 X \text{ est un estimateur de } \theta \text{ sans biais.}}$$

□

4. Calculer $\mathbb{V}(X)$.

Démonstration.

- La variable aléatoire X admet une variance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^2 f_\theta(x)\ dx$ est absolument convergente, ce qui revient à montrer la convergence pour ce calcul de moment.
- Tout d'abord :

$$\begin{aligned}
\int_{-\infty}^{+\infty} x^2 f_\theta(x)\ dx &= \int_0^\theta x^2 f_\theta(x)\ dx && \textit{car } f_\theta \textit{ est nulle en dehors de } [0,\theta] \\
&= \int_0^\theta x^2 \frac{k+1}{\theta^{k+1}}\ x^k\ dx \\
&= \frac{k+1}{\theta^{k+1}} \int_0^\theta x^{k+2}\ dx
\end{aligned}$$

Ainsi, $\displaystyle\int_{-\infty}^{+\infty} x^2 f_\theta(x)\ dx$ est convergente puisque $x \mapsto x^{k+2}$ est continue sur $[0,\theta]$. Il suit que X admet une variance et

$$\begin{aligned}
\mathbb{E}(X^2) &= \frac{k+1}{\theta^{k+1}} \left[\ \frac{x^{k+3}}{k+3}\ \right]_0^\theta \\
&= \frac{k+1}{\theta^{k+1}} \frac{\theta^{k+3}}{k+3} \\
&= \frac{k+1}{k+3}\theta^2
\end{aligned}$$

- D'après la formule de Koenig-Huygens : $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. D'où :

$$\begin{aligned}
\mathbb{V}(X) &= \frac{k+1}{k+3}\theta^2 - \left(\frac{k+1}{k+2}\theta\right)^2 \\
&= (k+1)\theta^2\left(\frac{1}{k+3} - \frac{k+1}{(k+2)^2}\right) \\
&= (k+1)\theta^2 \frac{(k+2)^2 - (k+1)(k+3)}{(k+3)(k+2)^2} \\
&= (k+1)\theta^2 \frac{k^2+4k+4-(k^2+4k+3)}{(k+3)(k+2)^2} \\
&= \frac{(k+1)}{(k+3)(k+2)^2}\theta^2
\end{aligned}$$

$$\boxed{\mathbb{V}(X) = \frac{(k+1)}{(k+3)(k+2)^2}\theta^2}$$

□

5. Démontrer que pour tout T estimateur de θ admettant une variance :

$$r_\theta(T) = \big(\mathbb{E}(T) - \theta\big)^2 + \mathbb{V}(T)$$

Démonstration. On utilise la linéarité de l'espérance :

$$\begin{aligned}
& \mathbb{V}(T) + (\mathbb{E}(T) - \theta)^2 \\
= \quad & \mathbb{E}((T - \mathbb{E}(T)^2) + (\mathbb{E}(T) - \theta)^2 \\
= \quad & \mathbb{E}(T^2 - 2\mathbb{E}(T)T + (\mathbb{E}(T))^2) + (\mathbb{E}(T))^2 - 2\theta\mathbb{E}(T) + \theta^2 \\
= \quad & \mathbb{E}(T^2) - \cancel{2\mathbb{E}(T)\mathbb{E}(T)} + \cancel{2(\mathbb{E}(T))^2} - 2\theta\mathbb{E}(T) + \theta^2 \\
= \quad & \mathbb{E}(T^2 - 2\theta T + \theta^2) \\
= \quad & \mathbb{E}((T-\theta)^2) \\
= \quad & r_\theta(T)
\end{aligned}$$

□

6. Donner la valeur de $r_\theta(\lambda_0 X)$.
Le but de la fin de cette partie I est de déterminer un estimateur de θ ayant un plus petit risque quadratique que celui de $\lambda_0\, X$.

Démonstration. La variable aléatoire $\lambda_0 X$ admet une variance comme transformée affine de X qui admet une variance, elle admet donc un risque quadratique.

$$\begin{aligned}
r_\theta(\lambda_0 X) &= \mathbb{V}(\lambda_0 X) + (\mathbb{E}(\lambda_0 X) - \theta)^2 && \textit{(d'après la question } \textbf{\textit{5}}\textit{)} \\
&= \mathbb{V}(\lambda_0 X) + (\theta - \theta)^2 && \textit{(d'après la question } \textbf{\textit{3}}\textit{)} \\
&= \lambda_0^2 \mathbb{V}(X) && \textit{(par propriété de la variance)} \\
&= \left(\frac{k+2}{k+1}\right)^2 \frac{(k+1)}{(k+3)(k+2)^2}\theta^2 && \textit{(d'après la question } \textbf{\textit{4}}\textit{)} \\
&= \frac{\theta^2}{(k+1)(k+3)}
\end{aligned}$$

□

7. En utilisant I.5 montrer que pour tout λ réel

$$r_\theta(\lambda X) = \theta^2\ Q(\lambda)$$

où Q est un polynôme de degré 2 dont les coefficients ne dépendent que de k.

Démonstration.
Soit $\lambda \in \mathbb{R}$.

$$\begin{aligned}
r_\theta(\lambda X) &= \mathbb{V}(\lambda X) + \big(\mathbb{E}(\lambda X) - \theta\big)^2 \\
&= \lambda^2 \mathbb{V}(X) + \big(\lambda \mathbb{E}(X) - \theta\big)^2 \\
&= \lambda^2 \frac{(k+1)}{(k+3)(k+2)^2}\theta^2 + \left(\lambda \frac{k+1}{k+2}\theta - \theta\right)^2 \\
&= \theta^2 \left(\frac{(k+1)}{(k+3)(k+2)^2}\lambda^2 + \left(\lambda\frac{k+1}{k+2} - 1\right)^2\right) \\
&= \theta^2 \left(\frac{(k+1)}{(k+3)(k+2)^2}\lambda^2 + \left(\frac{k+1}{k+2}\right)^2 \lambda^2 - 2\frac{k+1}{k+2}\lambda + 1\right) \\
&= \theta^2 \left(\left(\frac{(k+1)}{(k+3)(k+2)^2} + \frac{(k+1)^2}{(k+2)^2}\right)\lambda^2 - 2\frac{k+1}{k+2}\lambda + 1\right)
\end{aligned}$$

Or,

$$\begin{aligned}
\frac{(k+1)}{(k+3)(k+2)^2} + \frac{(k+1)^2}{(k+2)^2} &= (k+1)\frac{1 + (k+1)(k+3)}{(k+3)(k+2)^2} \\
&= (k+1)\frac{k^2 + 4k + 4}{(k+3)(k+2)^2} \\
&= (k+1)\frac{(k+2)^2}{(k+3)(k+2)^2} \\
&= \frac{k+1}{k+3}
\end{aligned}$$

$$\boxed{\text{On peut donc conclure que : } r_\theta(\lambda X) = \theta^2\ Q(\lambda) \text{ où } Q(\lambda) = \frac{k+1}{k+3}\lambda^2 - 2\frac{k+1}{k+2}\lambda + 1}$$

□

8. Montrer que la fonction $\lambda \mapsto Q(\lambda)$ atteint son minimum en un unique réel noté λ^* que l'on exprimera en fonction de k.

Démonstration.
La fonction Q est un trinôme du second degré dont le coefficient dominant est strictement positif. On en déduit que Q admet un minimum global atteint uniquement au point où sa dérivée s'annule.
Soit $\lambda \in \mathbb{R}$.

$$Q'(\lambda) = 2\frac{k+1}{k+3}\lambda - 2\frac{k+1}{k+2} = 2\frac{k+1}{k+3}\left(\lambda - \frac{k+3}{k+2}\right)$$

$$\boxed{\text{La fonction } Q \text{ atteint son minimum global en un unique point qui est } \lambda^* = \dfrac{k+3}{k+2}.}$$

□

9. Conclure sur le but recherché.

Démonstration.

- Tout d'abord, les calculs précédents montrent que $\lambda^* X$ est un estimateur de θ de risque quadratique minimal parmi tous les estimateurs de θ de la forme λX. En particulier :

$$r_\theta(\lambda^* X) \leqslant r_\theta(\lambda_0 X)$$

- Précisons le gain obtenu :

$$\begin{aligned} r_\theta(\lambda^* X) &= \theta^2 Q(\lambda^*) \\ &= \theta^2 \left(\frac{k+1}{k+3} \left(\frac{k+3}{k+2} \right)^2 - 2 \frac{k+1}{k+2} \frac{k+3}{k+2} + 1 \right) \\ &= \theta^2 \left(\frac{(k+1)(k+3)}{(k+2)^2} - 2 \frac{(k+1)(k+3)}{(k+2)^2} + 1 \right) \\ &= \theta^2 \left(1 - \frac{(k+1)(k+3)}{(k+2)^2} \right) \\ &= \frac{\theta^2}{(k+2)^2} \end{aligned}$$

D'où :

$$r_\theta(\lambda^* X) \underset{k \to +\infty}{\sim} \frac{\theta^2}{k^2} \underset{k \to +\infty}{\sim} r_\theta(\lambda_0 X)$$

$$\boxed{\begin{array}{c}\text{On a trouvé un autre estimateur, dont le risque quadratique est plus petit,} \\ \text{mais le gain est assez faible lorsque } k \text{ est grand.}\end{array}}$$

□

II. Second problème d'estimation

Dans ce second problème d'estimation, on dispose de n observations indépendantes $(n \geqslant 2)$ notées $X_1, \ldots, X_n$ de même loi de Poisson de paramètre θ inconnu ($\theta \in]0, +\infty[$). On souhaite estimer le paramètre $\exp(-\theta)$. On définit pour tout i élément de $[\![1, n]\!]$ la variable aléatoire Y_i par :

$$Y_i \quad : \quad \omega \quad \mapsto \quad \begin{cases} 1 & \text{si } X_i(\omega) = 0 \\ 0 & \text{sinon} \end{cases}$$

Puis on note :

$$\overline{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

10. Pour tout i élément de $[\![1, n]\!]$, donner la loi de Y_i.

Démonstration.
Soit $i \in [\![1, n]\!]$.

- Tout d'abord, par définition de Y_i : $Y_i(\Omega) = \{0, 1\}$. Ainsi, Y_i suit une loi de Bernoulli de paramètre $\alpha = \mathbb{P}([Y_i = 1])$.
- Calculons maintenant α.

$$\begin{aligned} \alpha &= \mathbb{P}([Y_i = 1]) \\ &= \mathbb{P}([X_i = 0]) \\ &= \mathrm{e}^{-\theta}\frac{\theta^0}{0!} && \textit{(car } X_i \hookrightarrow \mathcal{P}(\theta)\textit{)} \\ &= \mathrm{e}^{-\theta} \end{aligned}$$

$$\boxed{\text{D'où : } Y_i \hookrightarrow \mathcal{B}\left(\mathrm{e}^{-\theta}\right).}$$

□

11. Donner la loi de $\sum\limits_{i=1}^{n} Y_i$, puis montrer que $\mathbb{E}\big(\overline{Y}_n\big) = \exp(-\theta)$.

Démonstration.

- Les variables aléatoires X_i sont mutuellement indépendantes. Or, pour tout $i \in [\![1, n]\!]$, la variable aléatoire Y_i est une transformée de X_i (plus précisément : $Y_i = \mathbb{1}_{[X_i=0]}$). Ainsi, par lemme des coalitions, les variables aléatoires Y_i sont mutuellement indépendantes.

$$\boxed{\text{D'après la question précédente et la stabilité des lois binomiales : } \sum_{i=1}^{n} Y_i \hookrightarrow \mathcal{B}\left(n, \mathrm{e}^{-\theta}\right)}$$

- La variable aléatoire $\overline{Y}_n$ admet une espérance comme transformée affine de $\sum\limits_{i=1}^{n} Y_i$ qui admet une espérance.

$$\boxed{\text{Par linéarité de l'espérance : } \mathbb{E}\big(\overline{Y}_n\big) = \frac{1}{n} n\mathrm{e}^{-\theta} = \exp(-\theta).}$$

□

On dira dans ce cas que $\overline{Y}_n$ est un estimateur sans biais de $\exp(-\theta)$.

12. Calculer $\mathbb{V}\left(\overline{Y}_n\right)$.

Démonstration. La variable aléatoire $\overline{Y}_n$ admet une variance comme transformée affine de $\sum\limits_{i=1}^{n} Y_i$ qui admet une variance.

$$\begin{aligned} \mathbb{V}\left(\overline{Y}_n\right) &= \mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) \\ &= \frac{1}{n^2}\mathbb{V}\left(\sum_{i=1}^{n} Y_i\right) \\ &= \frac{1}{n^2} n\mathrm{e}^{-\theta}\left(1 - \mathrm{e}^{-\theta}\right) \\ &= \frac{\mathrm{e}^{-\theta}\left(1 - \mathrm{e}^{-\theta}\right)}{n} \end{aligned}$$

$$\boxed{\mathbb{V}\left(\overline{Y}_n\right) = \frac{\mathrm{e}^{-\theta}\left(1 - \mathrm{e}^{-\theta}\right)}{n}}$$

$\square$

Pour tout k élément de $[\![1,n]\!]$ on définit $S_k = \sum\limits_{i=1}^{k} X_i$.

13. Rappeler sans démonstration la loi de S_k pour tout k élément de $[\![1,n]\!]$.

Démonstration.

$$\boxed{\text{Pour tout } k \text{ élément de } [\![1,n]\!] : S_k \hookrightarrow \mathcal{P}(k\theta).}$$

$\square$

On définit jusqu'à la fin de cette partie II pour tout j entier naturel :

$$\varphi(j) = \mathbb{P}_{[S_n=j]}\big(\,[X_1 = 0]\,\big)$$

14. Montrer que pour tout j entier naturel :

$$\varphi(j) = \left(1 - \frac{1}{n}\right)^j$$

Démonstration.
Soit $j \in \mathbb{N}$.

$$\begin{aligned}
\varphi(j) &= \mathbb{P}_{[S_n=j]}\big(\,[X_1 = 0]\,\big) \\
&= \frac{\mathbb{P}([S_n = j] \cap [X_1 = 0])}{\mathbb{P}([S_n = j])} \\
&= \frac{\mathbb{P}\left(\left[\sum\limits_{i=1}^{n} X_i = j\right] \cap [X_1 = 0]\right)}{\mathrm{e}^{-n\theta}\,\dfrac{(n\theta)^j}{j!}} && \textit{(car } S_n \hookrightarrow \mathcal{P}(n\theta)\textit{)} \\
&= \frac{\mathbb{P}\left(\left[\sum\limits_{i=2}^{n} X_i = j\right] \cap [X_1 = 0]\right)}{\mathrm{e}^{-n\theta}\,\dfrac{(n\theta)^j}{j!}}
\end{aligned}$$

Or, les variables aléatoires X_i sont mutuellement indépendantes et donc, par lemme des coalitions, les variables aléatoires X_1 et $\sum\limits_{i=2}^{n} X_i$ sont également indépendantes. Ainsi :

$$\varphi(j) = \frac{\mathbb{P}\left(\left[\sum\limits_{i=2}^{n} X_i = j\right]\right)\mathbb{P}\left([X_1 = 0]\right)}{\mathrm{e}^{-n\theta}\,\dfrac{(n\theta)^j}{j!}}$$

De plus, les variables aléatoires $\sum_{i=2}^{n} X_i$ et S_{n-1} suivent la même loi. Il suit que :

$$\begin{aligned}
\varphi(j) &= \frac{\cancel{e^{-(n-1)\theta}} \dfrac{((n-1)\theta)^j}{\cancel{j!}} \cancel{e^{-\theta}}}{\cancel{e^{-n\theta}} \dfrac{(n\theta)^j}{\cancel{j!}}} \\
&= \frac{(n-1)^j \cancel{\theta^j}}{n^j \cancel{\theta^j}} \\
&= \left(\frac{n-1}{n}\right)^j
\end{aligned}$$

$$\boxed{\text{Finalement : } \varphi(j) = \left(1 - \frac{1}{n}\right)^j.}$$

□

On a donc $\varphi(j)$ indépendant du paramètre θ inconnu.
D'après la question II.13, on peut définir l'estimateur :

$$\varphi(S_n) = \left(1 - \frac{1}{n}\right)^{S_n}$$

15. Montrer que $\varphi(S_n)$ admet une espérance et que $\mathbb{E}\big(\varphi(S_n)\big) = \exp(-\theta)$.

Démonstration.
D'après le théorème de transfert, la variable aléatoire $\varphi(S_n)$ admet une espérance si et seulement si la série $\sum \varphi(k)\mathbb{P}([S_n = k])$ est absolument convergente et si c'est le cas :

$$\mathbb{E}(\varphi(S_n)) = \sum_{k=0}^{+\infty} \varphi(k)\mathbb{P}([S_n = k])$$

Soit $N \in \mathbb{N}$.

$$\begin{aligned}
\sum_{k=0}^{N} \varphi(k)\mathbb{P}([S_n = k]) &= \sum_{k=0}^{N} \left(1 - \frac{1}{n}\right)^k e^{-n\theta} \frac{(n\theta)^k}{k!} \\
&= e^{-n\theta} \sum_{k=0}^{N} \frac{\left(n\theta\left(1 - \frac{1}{n}\right)\right)^k}{k!}
\end{aligned}$$

On reconnaît la somme partielle d'une série exponentielle de paramètre $x = n\theta\left(1 - \frac{1}{n}\right)$. Une telle série est toujours convergente donc $\varphi(S_n)$ admet une espérance et

$$\begin{aligned}
\mathbb{E}(\varphi(S_n)) &= e^{-n\theta} \sum_{k=0}^{+\infty} \frac{\left(n\theta\left(1 - \frac{1}{n}\right)\right)^k}{k!} \\
&= e^{-n\theta} e^{n\theta\left(1 - \frac{1}{n}\right)} \\
&= e^{-n\theta} e^{n\theta - \theta} \\
&= e^{-\theta}
\end{aligned}$$

$$\boxed{\varphi(S_n) \text{ admet une espérance et } \mathbb{E}\big(\varphi(S_n)\big) = \exp(-\theta).}$$

□

On dira dans ce cas que $\varphi(S_n)$ est un estimateur sans biais de $\exp(-\theta)$.

16. Montrer que $\varphi(S_n)$ admet une variance vérifiant :

$$\mathbb{V}\big(\varphi(S_n)\big) = \exp(-2\theta)\ \left(\exp\left(\frac{\theta}{n}\right) - 1\right)$$

Démonstration.

- La variable aléatoire $\varphi(S_n)$ admet une variance si et seulement si elle admet un moment d'ordre 2.
- D'après le théorème de transfert, la variable aléatoire $\varphi(S_n)^2$ admet une espérance si et seulement si la série $\sum\ \varphi(k)^2\mathbb{P}([S_n = k])$ est absolument convergente et si c'est le cas :

$$\mathbb{E}(\varphi(S_n)^2) = \sum_{k=0}^{+\infty} \varphi(k)^2\mathbb{P}([S_n = k])$$

Soit $N \in \mathbb{N}$.

$$\begin{aligned}\sum_{k=0}^{N} \varphi(k)^2\mathbb{P}([S_n = k]) &= \sum_{k=0}^{N} \left(1 - \frac{1}{n}\right)^{2k} \mathrm{e}^{-n\theta}\,\frac{(n\theta)^k}{k!} \\ &= \mathrm{e}^{-n\theta}\sum_{k=0}^{N} \frac{\left(n\theta\left(1-\frac{1}{n}\right)^2\right)^k}{k!}\end{aligned}$$

On reconnaît la somme partielle d'une série exponentielle de paramètre $x = n\theta\left(1 - \dfrac{1}{n}\right)^2$. Une telle série est toujours convergente donc $\varphi(S_n)^2$ admet une espérance et

$$\begin{aligned}\mathbb{E}(\varphi(S_n)^2) &= \mathrm{e}^{-n\theta}\sum_{k=0}^{+\infty} \frac{\left(n\theta\left(1-\frac{1}{n}\right)^2\right)^k}{k!} \\ &= \mathrm{e}^{-n\theta}\mathrm{e}^{n\theta\left(1-\frac{1}{n}\right)^2} \\ &= \mathrm{e}^{-n\theta}\mathrm{e}^{n\theta\left(1-\frac{2}{n}+\frac{1}{n^2}\right)} \\ &= \mathrm{e}^{-n\theta}\mathrm{e}^{n\theta-2\theta+\frac{\theta}{n}} \\ &= \mathrm{e}^{-2\theta}\mathrm{e}^{\frac{\theta}{n}}\end{aligned}$$

$$\boxed{\varphi(S_n)^2 \text{ admet une espérance et } \mathbb{E}\big(\varphi(S_n)^2\big) = \mathrm{e}^{-2\theta}\mathrm{e}^{\frac{\theta}{n}}.}$$

- D'après la formule de Koenig-Huygens :

$$\begin{aligned}\mathbb{V}(\varphi(S_n)) &= \mathbb{E}\big(\varphi(S_n)^2\big) - (\mathbb{E}(\varphi(S_n)))^2 \\ &= \mathrm{e}^{-2\theta}\mathrm{e}^{\frac{\theta}{n}} - \mathrm{e}^{-2\theta} \\ &= \mathrm{e}^{-2\theta}\left(\mathrm{e}^{\frac{\theta}{n}} - 1\right)\end{aligned}$$

$$\boxed{\varphi(S_n) \text{ admet une variance vérifiant : } \mathbb{V}\big(\varphi(S_n)\big) = \exp(-2\theta)\ \left(\exp\left(\frac{\theta}{n}\right) - 1\right).}$$

□

17. On souhaite comparer les performances de $\overline{Y}_n$ et $\varphi(S_n)$ en tant qu'estimateurs de $\exp(-\theta)$.

a) Démontrer :

$$1 \leqslant \frac{\exp(\theta) - 1}{\theta} \leqslant \exp(\theta)$$

Démonstration.

- Tout d'abord, remarquons que $\theta > 0$. Ainsi, il s'agit de démontrer que :

$$\mathrm{e}^{\theta} \geqslant \theta + 1 \qquad \text{et} \qquad \theta\mathrm{e}^{\theta} - \mathrm{e}^{\theta} + 1 \geqslant 0$$

- La première inégalité est une inégalité de convexité classique que nous ne détaillons pas ici.
- Soit $f : x \mapsto x\mathrm{e}^x - \mathrm{e}^x + 1$ définie sur $\mathbb{R}^+$. La fonction f est dérivable sur $\mathbb{R}^+$. Soit $x \geqslant 0$.

$$f'(x) = \mathrm{e}^x + x\mathrm{e}^x - \mathrm{e}^x = x\mathrm{e}^x \geqslant 0$$

  Ainsi, la fonction f est croissante sur $\mathbb{R}^+$. De plus, $f(0) = 0$. On en déduit que, pour tout $x \in \mathbb{R}^+$, $f(x) \geqslant 0$. En particulier :

$$f(\theta) = \theta\mathrm{e}^{\theta} - \mathrm{e}^{\theta} + 1 \geqslant 0$$

$$\boxed{1 \leqslant \frac{\exp(\theta) - 1}{\theta} \leqslant \exp(\theta)}$$

□

b) Soit la fonction $h : [0,1] \to \mathbb{R}$ définie par :

$$h(t) = t\exp(\theta) + (1-t) - \exp(t\theta)$$

pour tout $t \in [0,1]$. Étudier les variations de h.

Démonstration.
La fonction h est deux fois dérivable sur $[0,1]$. Soit $t \in [0,1]$.

$$h'(t) = \mathrm{e}^{\theta} - 1 - \theta\mathrm{e}^{t\theta}$$

et

$$h''(t) = -\theta^2\mathrm{e}^{t\theta} < 0$$

Ainsi, la fonction h' est strictement décroissante sur $[0,1]$. De plus, d'après la question précédente :

× $h'(0) = \mathrm{e}^{\theta} - 1 - \theta \geqslant 0$

× $h'(1) = \mathrm{e}^{\theta} - 1 - \theta \mathrm{e}^{\theta} \leqslant 0$

Puisque h' est continue, il suit d'après le théorème de la bijection que h' s'annule une unique fois sur $[0,1]$.

Déterminons maintenant l'unique point où h' s'annule :

$$\begin{aligned} h'(t) = 0 &\iff \mathrm{e}^{\theta} - 1 - \theta \mathrm{e}^{t\theta} = 0 \\ &\iff \mathrm{e}^{t\theta} = \frac{\mathrm{e}^{\theta} - 1}{\theta} \\ &\iff t\theta = \ln\left(\frac{\mathrm{e}^{\theta} - 1}{\theta}\right) \\ &\iff t = \frac{1}{\theta} \ln\left(\frac{\mathrm{e}^{\theta} - 1}{\theta}\right) \end{aligned}$$

Posons $t^* = \dfrac{1}{\theta} \ln\left(\dfrac{\mathrm{e}^{\theta} - 1}{\theta}\right)$. On obtient le tableau de variations suivant :

t	0		t^*		1
Signe de $h'(t)$		$+$	0	$-$	
Variations de h	0	$\nearrow$	$h(t^*)$	$\searrow$	0

□

c) En déduire :

$$\exp\left(\frac{\theta}{n}\right) \leqslant \frac{\exp(\theta)}{n} + \frac{n-1}{n}$$

puis l'inégalité :

$$\mathbb{V}\big(\varphi(S_n)\big) \leqslant \mathbb{V}\big(\overline{Y}_n\big)$$

Démonstration.

- D'après la question précédente : pour tout $t \in [0,1]$, $h(t) \geqslant 0$.
  En particulier, pour $t = \dfrac{1}{n} \in [0,1]$, on obtient :

$$\frac{\exp(\theta)}{n} + 1 - \frac{1}{n} - \exp\left(\frac{\theta}{n}\right) \geqslant 0$$

$$\boxed{\text{D'où : } \exp\left(\frac{\theta}{n}\right) \leqslant \frac{\exp(\theta)}{n} + \frac{n-1}{n}.}$$

- Réécrivons l'inégalité précédente sous la forme :

$$\exp\left(\frac{\theta}{n}\right) - 1 \leqslant \frac{\exp(\theta)}{n} - \frac{1}{n}$$

On a alors :

$$\begin{aligned}\mathbb{V}(\varphi(S_n)) &= \exp(-2\theta)\ \left(\exp\left(\frac{\theta}{n}\right)-1\right)\\ &\leqslant \exp(-2\theta)\frac{\exp(\theta)-1}{n}\\ &= \frac{\exp(-\theta)(1-\exp(-\theta))}{n}\\ &= \mathbb{V}(\overline{Y}_n)\end{aligned}$$

$$\boxed{\text{On a bien : } \mathbb{V}(\varphi(S_n)) \leqslant \mathbb{V}(\overline{Y}_n).}$$

□

d) Comparer les risques quadratiques de $\overline{Y}_n$ et $\varphi(S_n)$ en tant qu'estimateurs de $\exp(-\theta)$.

Démonstration.
D'après la question ***5*** et puisque les estimateurs $\overline{Y}_n$ et $\varphi(S_n)$ sont sans biais, il suit que :

$$r_\theta\left(\overline{Y}_n\right) = \mathbb{V}\left(\overline{Y}_n\right) \qquad \text{et} \qquad r_\theta\left(\varphi(S_n)\right) = \mathbb{V}\left(\varphi(S_n)\right)$$

$$\boxed{\text{D'après la question précédente : } r_\theta\left(\varphi(S_n)\right) \leqslant r_\theta\left(\overline{Y}_n\right).}$$

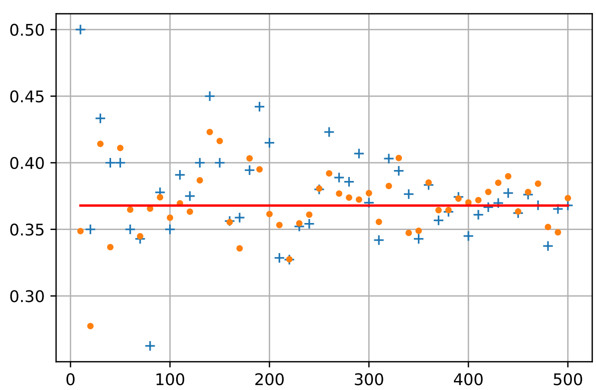
□

18. ***a)*** Compléter la fonction **Python** suivante pour qu'elle simule les estimateurs $\overline{Y}_n$ et $\varphi(S_n)$.

```
def simulEstimateurs(theta, n):
    Xobs =   rd.poisson(theta, n)
    Y = np.zeros(n)
    for k in range(n):
        if   Xobs[k]  == 0  :
            Y[k]  = 1
    Ybar = np.mean(Y)
    phi =    (1-1/n)**(np.sum(Xobs))
    return [Ybar,phi]
```

b) On exécute le script **Python** ci-dessous. Commenter en lien avec la question ***17***.

```
YbarListe=[]; phiListe=[]; theta=1
nListe = [10*i for i in range(1, 51)]
for n in nListe:
  [y,p] = simulEstimateurs(theta, n)
  YbarListe.append(y)
  phiListe.append(p)
plt.plot(nListe, YbarListe, '+')
plt.plot(nListe, phiListe, '.')
plt.plot(nListe,
     [np.exp(-theta) for k in nListe])
plt.show()
```



Démonstration.
On remarque que les $\bullet$ sont légèrement plus concentrés autour de la droite (d'ordonnée égale à $\exp(-\theta)$) que les $+$, ce qui semble confirmer que $\varphi(S_n)$ est un meilleur estimateur que $\overline{Y}_n$. Cependant, la différence de qualité entre les deux estimateurs ne paraît pas très forte sur ce graphique et semble diminuer lorsque la taille de l'échantillon (n) grandit. $\square$

On reprendra à la fin de la partie IV l'étude de $\varphi(S_n)$.

III. Information de Fisher

III.1. Cas discret

Dans cette section III.1, on considère I un intervalle de $\mathbb{R}$, θ un paramètre inconnu appartenant à I et X une variable aléatoire à valeurs dans $\mathbb{N}$ ($X(\Omega) \subset \mathbb{N}$). On suppose qu'il existe une fonction p définie sur $I \times X(\Omega)$ telle que pour tout k élément de $X(\Omega)$:

$$\mathbb{P}\big(\, [X = k] \,\big) = p(\theta, k)$$

et vérifiant pour tout k de $X(\Omega)$, $\theta \mapsto p\,(\theta, k)$ dérivable sur I.
On note de plus : $h : (\theta, k) \mapsto \ln\big(p(\theta, k)\big)$.
On définit enfin, sous réserve d'existence l'**information de Fisher** de X par :

$$I_X(\theta) = \sum_{k \in X(\Omega)} \big(\partial_1(\ln \circ p)(\theta, k)\big)^2 p(\theta, k)$$

19. Dans cette question ***19***, on considère X une variable aléatoire qui suit la loi de Bernoulli de paramètre θ (où $\theta \in \,]0, 1[$).
On a alors $X(\Omega) = \{0, 1\}$, $\mathbb{P}\big(\, [X = 1] \,\big) = p(\theta, 1) = \theta$, $\mathbb{P}\big(\, [X = 0] \,\big) = p(\theta, 0) = 1 - \theta$ et :

$$I_X(\theta) = \big(\partial_1(h)(\theta, 1)\big)^2 p(\theta, 1) + \big(\partial_1(h)(\theta, 0)\big)^2 p(\theta, 0)$$

Montrer :

$$I_X(\theta) = \frac{1}{\theta\,(1 - \theta)}$$

Démonstration.
D'après les formules admises :

$$h(\theta, 1) = \ln(\theta) \qquad \text{et} \qquad h(\theta, 0) = \ln(1 - \theta)$$

Ainsi :

$$\partial_1(h)(\theta, 1) = \frac{1}{\theta} \qquad \text{et} \qquad \partial_1(h)(\theta, 0) = \frac{-1}{1 - \theta}$$

On en déduit :

$$\begin{aligned} I_X(\theta) &= \left(\frac{1}{\theta}\right)^2 \theta + \left(\frac{-1}{1 - \theta}\right)^2 (1 - \theta) \\ &= \frac{1}{\theta} + \frac{1}{1 - \theta} \\ &= \frac{1 - \theta + \theta}{\theta(1 - \theta)} \\ &= \frac{1}{\theta(1 - \theta)} \end{aligned}$$

$$\boxed{I_X(\theta) = \frac{1}{\theta\,(1-\theta)}}$$

□

20. Dans cette question ***20***, on considère X une variable aléatoire qui suit la loi binomiale de paramètres N et θ ($N \in \mathbb{N}^*$, $\theta \in]0,1[$).

a) Montrer :

$$I_X(\theta) = \frac{1}{\big(\theta\,(1-\theta)\big)^2} \sum_{k=0}^{N} (k - N\theta)^2 \binom{N}{k} \theta^k (1-\theta)^{N-k}$$

Démonstration.
Soit $k \in [\![0, N]\!]$.

$$p(\theta, k) = \mathbb{P}([X = k]) = \binom{N}{k} \theta^k (1-\theta)^{N-k}$$

donc

$$h(\theta, k) = \ln\left(\binom{N}{k}\right) + k\ln(\theta) + (N-k)\ln(1-\theta)$$

et donc

$$\begin{aligned}
\partial_1(h)(\theta, k) &= \frac{k}{\theta} - \frac{N-k}{1-\theta} \\
&= \frac{k(1-\theta) - (N-k)\theta}{\theta(1-\theta)} \\
&= \frac{k - k\theta - N\theta + k\theta}{\theta(1-\theta)} \\
&= \frac{k - N\theta}{\theta(1-\theta)}
\end{aligned}$$

Ainsi :

$$\begin{aligned}
I_X(\theta) &= \sum_{k=0}^{N} \big(\partial_1(\ln \circ p)(\theta, k)\big)^2 p(\theta, k) \\
&= \sum_{k=0}^{N} \left(\frac{k - N\theta}{\theta(1-\theta)}\right)^2 \binom{N}{k} \theta^k (1-\theta)^{N-k}
\end{aligned}$$

$$\boxed{\text{Finalement, on a bien : } I_X(\theta) = \frac{1}{\big(\theta\,(1-\theta)\big)^2} \sum_{k=0}^{N} (k - N\theta)^2 \binom{N}{k} \theta^k (1-\theta)^{N-k}.}$$

□

b) En déduire :

$$I_X(\theta) = \frac{\mathbb{V}(X)}{\big(\theta\,(1-\theta)\big)^2}$$

puis donner la valeur de $I_X(\theta)$.

Démonstration. D'après la question précédente, il s'agit de montrer que :

$$\sum_{k=0}^{N} (k - N\theta)^2 \binom{N}{k} \theta^k (1-\theta)^{N-k} = \mathbb{V}(X)$$

(remarquons au passage que $\mathbb{V}(X)$ existe puisque $X \hookrightarrow \mathcal{B}(N, \theta)$)
Par définition de la variance :

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left((X - N\theta)^2\right)$$

puis, par théorème de transfert (valide car X est finie) :

$$\mathbb{V}(X) = \sum_{k=0}^{N} (k - N\theta)^2 \binom{N}{k} \theta^k (1-\theta)^{N-k}$$

Puisque $\mathbb{V}(X) = N\theta(1-\theta)$, il suit que :

$$\boxed{I_X(\theta) = \frac{N}{\theta(1-\theta)}}$$

□

21. Dans cette question ***21***, on considère X une variable aléatoire qui suit la loi de Poisson de paramètre θ ($\theta \in \,]0, +\infty[$). Puisque $X(\Omega) = \mathbb{N}$, on a sous réserve de convergence :

$$I_X(\theta) = \sum_{k=0}^{+\infty} \big(\partial_1(h)(\theta, k)\big)^2 p(\theta, k)$$

a) Montrer que la série de terme général $\big(\partial_1(h)(\theta, k)\big)^2 p(\theta, k)$ converge et calculer sa somme $I_X(\theta)$.

Démonstration.
Soit $k \in [\![0, N]\!]$.

$$p(\theta, k) = \mathbb{P}([X = k]) = \mathrm{e}^{-\theta} \frac{\theta^k}{k!}$$

donc

$$h(\theta, k) = -\theta + k \ln(\theta) - \ln(k!)$$

et donc

$$\partial_1(h)(\theta, k) = -1 + \frac{k}{\theta}$$

Soit $N \in \mathbb{N}$.

$$\begin{aligned}
\sum_{k=0}^{N} \big(\partial_1(h)(\theta,k)\big)^2 p(\theta,k) &= \sum_{k=0}^{N} \left(\frac{k}{\theta}-1\right)^2 \mathrm{e}^{-\theta}\frac{\theta^k}{k!} \\
&= \mathrm{e}^{-\theta}\sum_{k=0}^{N} \left(\frac{k^2}{\theta^2}-\frac{2k}{\theta}+1\right)\frac{\theta^k}{k!} \\
&= \mathrm{e}^{-\theta}\left(\sum_{k=0}^{N} k^2\frac{\theta^{k-2}}{k!} - 2\sum_{k=0}^{N} k\frac{\theta^{k-1}}{k!} + \sum_{k=0}^{N}\frac{\theta^k}{k!}\right) \\
&= \mathrm{e}^{-\theta}\left(\sum_{k=1}^{N} k\frac{\theta^{k-2}}{(k-1)!} - 2\sum_{k=1}^{N} \frac{\theta^{k-1}}{(k-1)!} + \sum_{k=0}^{N}\frac{\theta^k}{k!}\right) \\
&= \mathrm{e}^{-\theta}\left(\sum_{k=0}^{N-1} (k+1)\frac{\theta^{k-1}}{k!} - 2\sum_{k=0}^{N-1} \frac{\theta^{k}}{k!} + \sum_{k=0}^{N}\frac{\theta^k}{k!}\right) \\
&= \mathrm{e}^{-\theta}\left(\sum_{k=1}^{N-1} \frac{\theta^{k-1}}{(k-1)!} + \sum_{k=0}^{N-1}\frac{\theta^{k-1}}{k!} - 2\sum_{k=0}^{N-1} \frac{\theta^{k}}{k!} + \sum_{k=0}^{N}\frac{\theta^k}{k!}\right) \\
&= \mathrm{e}^{-\theta}\left(\sum_{k=0}^{N-2} \frac{\theta^{k}}{k!} + \frac{1}{\theta}\sum_{k=0}^{N-1}\frac{\theta^{k}}{k!} - 2\sum_{k=0}^{N-1} \frac{\theta^{k}}{k!} + \sum_{k=0}^{N}\frac{\theta^k}{k!}\right)
\end{aligned}$$

On reconnaît quatre sommes partielles de la série exponentielle de paramètre θ. Ainsi, la série $\sum \big(\partial_1(h)(\theta,k)\big)^2 p(\theta,k)$ converge et

$$\boxed{I_X(\theta) = \mathrm{e}^{-\theta}\left(\mathrm{e}^{\theta} + \frac{1}{\theta}\mathrm{e}^{\theta} - 2\mathrm{e}^{\theta} + \mathrm{e}^{\theta}\right) = \frac{1}{\theta}}$$

□

b) Justifier :

$$I_X(\theta) = \mathbb{E}\left(\big(\partial_1(h)(\theta,X)\big)^2\right)$$

Démonstration.
D'après le théorème de transfert, la variable aléatoire $\big(\partial_1(h)(\theta,X)\big)^2$ admet une espérance si et seulement si la série $\sum \big(\partial_1(h)(\theta,k)\big)^2 p(\theta,k)$ converge absolument et si c'est le cas :

$$\mathbb{E}\left(\big(\partial_1(h)(\theta,X)\big)^2\right) = \sum_{k=0}^{+\infty} \big(\partial_1(h)(\theta,k)\big)^2 p(\theta,k)$$

D'après la question précédente, la série $\sum \big(\partial_1(h)(\theta,k)\big)^2 p(\theta,k)$ converge. Or, il s'agit d'une série à termes positifs, donc elle converge absolument.

$$\boxed{\text{On a bien : } \mathbb{E}\left(\big(\partial_1(h)(\theta,X)\big)^2\right) = \sum_{k=0}^{+\infty} \big(\partial_1(h)(\theta,k)\big)^2 p(\theta,k) = I_X(\theta)}$$

□

III.2. Cas d'une variable gaussienne

Soit X une variable aléatoire qui suit la loi normale de moyenne θ ($\theta \in \mathbb{R}$) et de variance 1 dont la densité est notée $x \mapsto f(\theta,x)$. On définit sous réserve de convergence l'**information de Fisher** de X par :

$$I_X(\theta) = \int_{-\infty}^{+\infty} \big(\partial_1(\ln\circ f)(\theta,x)\big)^2 f(\theta,x)\ dx$$

22. Montrer que sous réserve de convergence :

$$I_X(\theta) = \int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta, x)\ dx$$

Démonstration.
Soit $x \in \mathbb{R}$.

$$f(\theta, x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}(x-\theta)^2}$$

donc

$$(\ln \circ f)(\theta, x) = -\ln(\sqrt{2\pi}) - \frac{1}{2}(x-\theta)^2$$

et donc

$$\partial_1(\ln \circ f)(\theta, x) = x - \theta$$

$$\boxed{\text{On a bien, sous réserve de convergence : } I_X(\theta) = \int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta, x)\ dx.}$$

□

23. En déduire l'existence et la valeur de $I_X(\theta)$.

Démonstration.
Soit $x \in \mathbb{R}$. On a :

$$(x-\theta)^2 f(\theta, x) = x^2 f(\theta, x) - 2\theta x f(\theta, x) + \theta^2 f(\theta, x)$$

Puisque $X \hookrightarrow \mathcal{N}(\theta, 1)$, il suit que X admet des moments au moins jusqu'à l'ordre 2 et donc :

× $\displaystyle\int_{-\infty}^{+\infty} f(\theta, x)\ dx$ converge et

$$\int_{-\infty}^{+\infty} f(\theta, x)\ dx = \mathbb{E}(X^0) = 1$$

× $\displaystyle\int_{-\infty}^{+\infty} x f(\theta, x)\ dx$ converge et

$$\int_{-\infty}^{+\infty} x f(\theta, x)\ dx = \mathbb{E}(X) = \theta$$

× $\displaystyle\int_{-\infty}^{+\infty} x^2 f(\theta, x)\ dx$ converge et d'après Koenig-Huygens :

$$\int_{-\infty}^{+\infty} x^2 f(\theta, x)\ dx = \mathbb{E}(X^2) = \mathbb{V}(X) + \mathbb{E}(X)^2 = 1 + \theta^2$$

On en déduit que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta, x)\ dx$ converge et, par linéarité de l'intégrale :

$$\int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta, x)\ dx = 1 + \theta^2 - 2\theta^2 + \theta^2 = 1$$

$$\boxed{\text{D'où : } I_X(\theta) = 1.}$$

□

24. Justifier :

$$I_X(\theta) = \mathbb{E}\left(\big(\partial_1(\ln\circ f)(\theta, X)\big)^2\right)$$

Démonstration.
D'après le théorème de transfert, la variable aléatoire $\big(\partial_1(h)(\theta, X)\big)^2$ admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \big(\partial_1(\ln\circ f)(\theta, x)\big)^2 f(\theta, x)\ dx$ converge absolument et si c'est le cas :

$$\mathbb{E}\left(\big(\partial_1(h)(\theta, X)\big)^2\right) = \int_{-\infty}^{+\infty} \big(\partial_1(\ln\circ f)(\theta, x)\big)^2 f(\theta, x)\ dx$$

D'après la question précédente, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \big(\partial_1(\ln\circ f)(\theta, x)\big)^2 f(\theta, x)\ dx$ converge. Or, il s'agit d'une intégrale d'une fonction positive, donc elle converge absolument.

$$\boxed{\text{On a bien : } \mathbb{E}\left(\big(\partial_1(h)(\theta, X)\big)^2\right) = \int_{-\infty}^{+\infty} \big(\partial_1(\ln\circ f)(\theta, x)\big)^2 f(\theta, x)\ dx = I_X(\theta)}$$

□

IV. Minoration du risque quadratique

IV.1. Inégalité de Cramer-Rao

Dans cette section IV.1, on considère I un intervalle de $\mathbb{R}$, θ un paramètre inconnu appartenant à I et X une variable aléatoire telle que $X(\Omega) = [\![0, N]\!]$ ($N \in \mathbb{N}$). On suppose qu'il existe une fonction p définie sur $I \times X(\Omega)$ telle que pour tout $k \in [\![0, N]\!]$:

$$\mathbb{P}\big(\,[X = k]\,\big) = p(\theta, k)$$

et vérifiant :

- pour tout $k \in [\![0, N]\!]$, $\theta \mapsto p(\theta, k)$ dérivable sur I,
- l'information de Fisher de X notée $I_X(\theta)$ définie dans la partie III est non nulle pour tout $\theta \in I$.

Le but de la section IV.1 est de démontrer l'inégalité suivante dûe à Cramer et Rao.

Théorème 1. *(de Cramer-Rao)*
Soit $f(X)$ un estimateur sans biais de $g(\theta)$ à savoir tel que $\mathbb{E}\big(f(X)\big) = g(\theta)$ où g est dérivable sur I. On a alors :

$$\mathbb{V}\left(f(X)\right) \geqslant \frac{\big(g'(\theta)\big)^2}{I_X(\theta)}$$

25. Montrer que pour tout θ élément de I :

$$\sum_{k=0}^{N} \partial_1(p)(\theta, k) = 0$$

Démonstration.
Soit $\theta \in I$. La famille $([X = k])_{k\in[\![0,N]\!]}$ est un système complet d'événements donc

$$\sum_{k=0}^{N} \mathbb{P}([X = k]) = 1$$

ce qui se réécrit :

$$\sum_{k=0}^{N} p(\theta, k) = 1$$

On dérive alors par rapport à la première variable, ce qui donne, 1 étant une constante :

$$\boxed{\sum_{k=0}^{N} \partial_1(p)(\theta, k) = 0}$$

□

26. En déduire que pour tout θ élément de I :

$$\mathbb{E}\left(\partial_1(h)(\theta, X)\right) = 0 \qquad (E)$$

Démonstration.
Soit $\theta \in I$. Tout d'abord, la variable aléatoire $\partial_1(h)(\theta, X)$ est finie car X est finie donc $\partial_1(h)(\theta, X)$ admet une espérance. Il suit, par théorème de transfert :

$$\mathbb{E}\left(\partial_1(h)(\theta, X)\right) = \sum_{k=0}^{N} \partial_1(h)(\theta, k)\mathbb{P}([X = k]) = \sum_{k=0}^{N} \partial_1(h)(\theta, k)p(\theta, k)$$

Ensuite :

$$h(\theta, k) = \ln(p(\theta, k))$$

et donc

$$\partial_1(h)(\theta, k) = \frac{\partial_1(p)(\theta, k)}{p(\theta, k)}$$

Il vient alors :

$$\begin{aligned}
\mathbb{E}\left(\partial_1(h)(\theta, X)\right) &= \sum_{k=0}^{N} \frac{\partial_1(p)(\theta, k)}{p(\theta, k)} p(\theta, k) \\
&= \sum_{k=0}^{N} \partial_1(p)(\theta, k) \\
&= 0 && \textit{(d'après la question } \textbf{\textit{25}}\textit{)}
\end{aligned}$$

$$\boxed{\mathbb{E}\left(\partial_1(h)(\theta, X)\right) = 0}$$

□

27. En dérivant partiellement par rapport à θ les deux membres de l'égalité (E), montrer que pour tout θ élément de I :

$$\mathbb{E}\left(\partial_{1,1}^2(h)(\theta, X)\right) = -\mathbb{E}\left(\left(\partial_1(h)(\theta, X)\right)^2\right)$$

Démonstration.
Soit $\theta \in I$. Réécrivons l'égalité précédente sous la forme :

$$\sum_{k=0}^{N} \partial_1(h)(\theta, k)p(\theta, k) = 0$$

Dérivons partiellement par rapport à θ :

$$\sum_{k=0}^{N} \left(\partial_{1,1}^2(h)(\theta, k)p(\theta, k) + \partial_1(h)(\theta, k)\partial_1(p)(\theta, k)\right) = 0$$

D'où :

$$\begin{aligned}
\sum_{k=0}^{N} \partial_{1,1}^2(h)(\theta, k)p(\theta, k) &= -\sum_{k=0}^{N} \partial_1(h)(\theta, k)\partial_1(p)(\theta, k) \\
&= -\sum_{k=0}^{N} \partial_1(h)(\theta, k)\partial_1(h)(\theta, k)p(\theta, k) && \textit{(voir calcul question } \textbf{\textit{26}}\textit{)} \\
&= -\sum_{k=0}^{N} \partial_1(h)(\theta, k)^2 p(\theta, k)
\end{aligned}$$

De plus, comme les variables aléatoires $\partial^2_{1,1}(h)(\theta, X)$ et $\partial_1(h)(\theta, X)^2$ sont finies, on peut réécrire l'égalité précédente à l'aide du théorème de transfert :

$$\boxed{\mathbb{E}\left(\partial^2_{1,1}(h)(\theta, X)\right) = -\mathbb{E}\left(\left(\partial_1(h)(\theta, X)\right)^2\right)}$$

□

28. Montrer que pour tout θ élément de I :

$$g'(\theta) = \sum_{k=0}^{N} f(k)\ \left(\partial_1(h)(\theta, k)\right)\ p(\theta, k)$$

puis :

$$g'(\theta) = \mathbb{E}\left(\left(f(X) - g(\theta)\right)\left(\partial_1(h)(\theta, X)\right)\right)$$

Démonstration.
Soit $\theta \in I$. Par hypothèse :

$$g(\theta) = \mathbb{E}(f(X))$$

ce qui se réécrit, par théorème de transfert :

$$g(\theta) = \sum_{k=0}^{N} f(k)p(\theta, k)$$

Dérivons partiellement par rapport à θ :

$$\begin{aligned} g'(\theta) &= \sum_{k=0}^{N} f(k)\partial_1(p)(\theta, k) \\ &= \sum_{k=0}^{N} f(k)\partial_1(h)(\theta, k)p(\theta, k) \end{aligned}$$

La variable aléatoire $\left(f(X) - g(\theta)\right)\left(\partial_1(h)(\theta, X)\right)$ est finie car X est finie, donc elle admet une espérance. Par linéarité :

$$\begin{aligned} \mathbb{E}\left(\left(f(X) - g(\theta)\right)\partial_1(h)(\theta, X)\right) &= \mathbb{E}\left(f(X)\partial_1(h)(\theta, X)\right) - \mathbb{E}\left(g(\theta)\partial_1(h)(\theta, X)\right) \\ &= \mathbb{E}\left(f(X)\partial_1(h)(\theta, X)\right) - g(\theta)\mathbb{E}\left(\partial_1(h)(\theta, X)\right) \end{aligned}$$

Or, d'après ce qui précède et par théorème de transfert :

$$\mathbb{E}\left(f(X)\left(\partial_1(h)(\theta, X)\right)\right) = g'(\theta)$$

et d'après (E) :

$$\mathbb{E}\left(\partial_1(h)(\theta, X)\right) = 0$$

$$\boxed{\text{D'où : } g'(\theta) = \mathbb{E}\left(\left(f(X) - g(\theta)\right)\left(\partial_1(h)(\theta, X)\right)\right).}$$

□

29. On pose pour tout t réel :

$$L(t) = \mathbb{E}\left(\left(\left(f(X) - g(\theta)\right) + t\left(\partial_1(h)(\theta, X)\right)\right)^2\right)$$

a) Développer le polynôme L suivant les puissances décroissantes de t.

Démonstration.
Soit $t \in \mathbb{R}$.

$$\begin{aligned} L(t) &= \mathbb{E}\left(\big(\big(f(X)-g(\theta)\big)+t\big(\partial_1(h)(\theta,X)\big)\big)^2\right) \\ &= \mathbb{E}\left(\partial_1(h)(\theta,X)^2t^2+2\big(f(X)-g(\theta)\big)\partial_1(h)(\theta,X)t+\big(f(X)-g(\theta)\big)^2\right) \\ &= \mathbb{E}\left(\partial_1(h)(\theta,X)^2\right)t^2+2\mathbb{E}\left(\big(f(X)-g(\theta)\big)\partial_1(h)(\theta,X)\right)t+\mathbb{E}\left(\big(f(X)-g(\theta)\big)^2\right) \quad \textit{(par linéarité)} \end{aligned}$$

Or, puisque $f(X)$ est sans biais :

$$\mathbb{E}\left(\big(f(X)-g(\theta)\big)^2\right)=\mathbb{E}\left(\big(f(X)-\mathbb{E}(f(X))\big)^2\right)=\mathbb{V}(f(X))$$

et, d'après la question ***28***,

$$\mathbb{E}\left(\big(f(X)-g(\theta)\big)\big(\partial_1(h)(\theta,X)\big)\right)=g'(\theta)$$

Enfin, par théorème de transfert :

$$\mathbb{E}\left(\partial_1(h)(\theta,X)^2\right)=I_X(\theta)$$

$$\boxed{\text{D'où : } L(t)=I_X(\theta)t^2+2g'(\theta)t+\mathbb{V}(f(X)).}$$

□

b) Calculer le discriminant Δ de L et justifier : $\Delta \leqslant 0$.

Démonstration.

- Tout d'abord :

$$\begin{aligned} \Delta &= 4(g'(\theta))^2-4I_X(\theta)\mathbb{V}(f(X)) \\ &= 4\Big((g'(\theta))^2-I_X(\theta)\mathbb{V}(f(X))\Big) \end{aligned}$$

- Ensuite, par positivité de la variable aléatoire $\big(\big(f(X)-g(\theta)\big)+t\big(\partial_1(h)(\theta,X)\big)\big)^2$ et par croissance de l'espérance :

$$\forall t \in \mathbb{R}, L(t) \geqslant 0$$

  Ceci implique que L s'annule au plus une fois et donc

$$\boxed{\Delta \leqslant 0}$$

□

c) En déduire l'inégalité de Cramer-Rao.

Démonstration.
D'après ce qui précède :

$$(g'(\theta))^2-I_X(\theta)\mathbb{V}(f(X)) \leqslant 0$$

d'où

$$(g'(\theta))^2 \leqslant I_X(\theta)\mathbb{V}(f(X))$$

De plus, $I_X(\theta) \geqslant 0$ et par hypothèse $I_X(\theta) \neq 0$ donc $I_X(\theta) > 0$.

$$\boxed{\text{On en déduit : } \mathbb{V}\left(f(X)\right) \geqslant \frac{\big(g'(\theta)\big)^2}{I_X(\theta)}}$$

□

IV.2. Extension du théorème de Cramer-Rao

On reprend dans cette section IV.2 les notations et hypothèses de la partie II. On admet que, dans ce contexte, le théorème de Cramer-Rao peut se généraliser comme suit :

Théorème 2. *(de Cramer-Rao)*
Soit $T_n = f(X_1, \dots, X_n)$ un estimateur sans biais de $g(\theta)$ à savoir tel que $\mathbb{E}\big(f(X_1, \dots, X_n)\big) = g(\theta)$ où g est dérivable sur $]0, +\infty[$. On a alors :

$$\mathbb{V}(T_n) \geqslant \frac{\big(g'(\theta)\big)^2}{n\, I_{X_1}(\theta)}$$

où $I_{X_1}(\theta)$ est l'information de Fisher d'une variable aléatoire de loi de Poisson de paramètre θ définie et calculée à la partie III.

Il s'agit dans cette section d'exploiter cette nouvelle inégalité de Cramer-Rao. On note :

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

30. Calculer $\mathbb{E}\big(\overline{X}_n\big)$ et $\mathbb{V}\big(\overline{X}_n\big)$.

Démonstration.
La variable aléatoire $\overline{X}_n$ admet une espérance et une variance en tant que combinaison linéaire de variables aléatoires admettant une espérance et une variance.

$$\begin{aligned}
\mathbb{E}\big(\overline{X}_n\big) &= \mathbb{E}\Big(\frac{1}{n} \sum_{i=1}^{n} X_i\Big) \\
&= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i) && \textit{(par linéarité)} \\
&= \frac{1}{n} \sum_{i=1}^{n} \theta \\
&= \frac{1}{n} n\theta \\
&= \theta
\end{aligned}$$

et

$$\begin{aligned}
\mathbb{V}\big(\overline{X}_n\big) &= \mathbb{V}\Big(\frac{1}{n} \sum_{i=1}^{n} X_i\Big) \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{V}(X_i) && \textit{(par indépendance)} \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \theta \\
&= \frac{1}{n^2} n\theta \\
&= \frac{\theta}{n}
\end{aligned}$$

$$\boxed{\mathbb{E}\big(\overline{X}_n\big) = \theta \text{ et } \mathbb{V}\big(\overline{X}_n\big) = \frac{\theta}{n}}$$

□

31. Déduire de la généralisation de Cramer-Rao, que $\overline{X}_n$ a le plus petit risque quadratique parmi les estimateurs sans biais de θ.

Démonstration.

- Tout d'abord, $\mathbb{E}\big(\overline{X}_n\big) = \theta$ donc $\overline{X}_n$ fait bien partie des estimateurs sans biais de θ.
- Soit $T_n = f(X_1, \dots, X_n)$ un estimateur sans biais de θ (ici $g = \mathrm{id}_{\mathbb{R}}$). D'après la généralisation de l'inégalité de Cramer-Rao :
$$\mathbb{V}(T_n) \geqslant \frac{1}{n\, I_{X_1}(\theta)}$$
(car $g'(\theta) = 1$).
- D'après la question ***21.a)***,
$$I_{X_1}(\theta) = \frac{1}{\theta}$$
d'où
$$\mathbb{V}(T_n) \geqslant \frac{1}{n\, I_{X_1}(\theta)} = \frac{\theta}{n} = \mathbb{V}\big(\overline{X}_n\big)$$

On a montré que l'estimateur $\overline{X}_n$ a la plus petite variance parmi les estimateurs sans biais de θ. Or, pour les estimateurs sans biais, la variance coïncide avec le risque quadratique.

$$\boxed{\text{L'estimateur } \overline{X}_n \text{ a le plus petit risque quadratique parmi les estimateurs sans biais de } \theta.}$$

□

32. Montrer que pour $g(\theta) = \exp(-\theta)$ où $\theta \in \,]0, +\infty[$:
$$\mathbb{V}\big(\varphi(S_n)\big) \underset{n\to+\infty}{\sim} \frac{\big(g'(\theta)\big)^2}{n\, I_{X_1}(\theta)}$$

Démonstration.

- Rappelons le résultat de la question ***16*** :
$$\mathbb{V}\big(\varphi(S_n)\big) = \exp(-2\theta)\, \left(\exp\left(\frac{\theta}{n}\right) - 1\right)$$
D'où :
$$\mathbb{V}\big(\varphi(S_n)\big) \underset{n\to+\infty}{\sim} \exp(-2\theta)\frac{\theta}{n}$$
(on a utilisé l'équivalent usuel $\mathrm{e}^u - 1 \underset{u\to 0}{\sim} u$)
- D'autre part, si $g : \theta \mapsto \mathrm{e}^{-\theta}$, il vient, pour tout $\theta \in \,]0, +\infty[$:
$$g'(\theta) = -\mathrm{e}^{-\theta}$$
et donc
$$\big(g'(\theta)\big)^2 = \exp(-2\theta)$$
Or,
$$I_{X_1}(\theta) = \frac{1}{\theta}$$
et donc
$$\frac{\big(g'(\theta)\big)^2}{n\, I_{X_1}(\theta)} = \exp(-2\theta)\frac{\theta}{n}$$

$$\boxed{\text{On a bien : } \mathbb{V}\big(\varphi(S_n)\big) \underset{n\to+\infty}{\sim} \frac{\big(g'(\theta)\big)^2}{n\, I_{X_1}(\theta)}.}$$

$\square$

33. Que prouve ce résultat en terme d'optimalité de $\varphi(S_n)$ dans l'estimation de $\exp(-\theta)$?

Démonstration.
Le résultat précédent prouve que la variance de $\varphi(S_n)$ est asymptotiquement minimale parmi les estimateurs sans biais de $\exp(-\theta)$ (d'après la généralisation de l'inégalité de Cramer-Rao). Autrement dit, $\varphi(S_n)$ a asymptotiquement le plus petit risque quadratique parmi tous les estimateurs sans biais de $\exp(-\theta)$. $\square$

34. À la lumière de la partie II, peut-on conclure que lorsque n est grand $\varphi(S_n)$ est le meilleur estimateur de $\exp(-\theta)$ en terme de risque quadratique ?

Démonstration.
Il est possible qu'un estimateur biaisé de $\exp(-\theta)$ ait un plus petit risque quadratique que $\varphi(S_n)$. Ce n'est donc pas nécessairement le meilleur estimateur de $\exp(-\theta)$ en terme de risque quadratique. $\square$