Corrigé ESSEC II 2009 par Pierre Veuillez

Notations

- Tout au long du sujet $(\Omega, \mathcal{F}, P)$ 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 $E(X)$ et sa variance sera notée $V(X)$.
- Pour un événement A, on notera $P_B(A)$ la probabilité conditionnelle de A sachant B où B est un événement non négligeable.

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

1 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

$$\begin{aligned} f_\theta(x) &= \frac{k+1}{\theta^{k+1}}\ x^k \text{ pour tout } x \in [0,\theta] \\ &= 0 \text{ sinon} \end{aligned}$$

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

1.1. f_θ est continue par morceaux sur $\mathbb{R}$ car k est positif.
f_θ est positive sur $\mathbb{R}$.

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

d'auter part d$\int_{-\infty}^0 f_\theta(x)\,dx = 0$ et $\int_\theta^{+\infty} f_\theta(x)\,dx = 0$
donc $\int_{-\infty}^{+\infty} f_\theta(x)\,dx$ converge et vaut 1
Conclusion : $\boxed{f_\theta \text{ est bien une densité de probabilité.}}$

1.2. Pour l'espérance :

$$\int_0^\theta x f_\theta(x)\,dx = \left[\frac{k+1}{\theta^{k+1}}\frac{1}{k+2}\ x^{k+2}\right]_0^\theta = \frac{k+1}{k+2}\theta$$

Donc X a une espérance et *Conclusion :* $\boxed{E(X) = \frac{k+1}{k+2}\theta}$

1.3. On a $E(\lambda_0 X) = \lambda_0 E(X) = \frac{k+1}{k+2}\lambda_0\theta$
Donc $\lambda_0 X$ estime θ sans biais si et seulement si $\frac{k+1}{k+2}\lambda_0 = 1$ soit
Conclusion : $\boxed{\lambda_0 = \frac{k+2}{k+1}}$

1.4. comme la densité est nulle en dehors de $[0,\theta]$

$$\begin{aligned} E\left(X^2\right) &= \int_0^\theta x^2 f_\theta(x)\,dx = \int_0^\theta \frac{k+1}{\theta^{k+1}}\ x^{k+2}dx \\ &= \frac{k+1}{\theta^{k+1}}\left[\frac{1}{k+3}x^{k+3}\right]_0^\theta \\ &= \frac{k+1}{k+3}\theta^2 \end{aligned}$$

donc

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

1.5. On définit le risque quadratique de T estimateur de θ par

$$R(T,\theta) = E\left((T-\theta)^2\right)$$

on a donc

$$\begin{aligned} E\left((T-\theta)^2\right) &= E\left(T^2 - 2\theta T + \theta^2\right) \\ &= E\left(T^2\right) - 2\theta E(T) + \theta^2 \end{aligned}$$

par linéarité de l'espérance et

$$\begin{aligned} (E(T)-\theta)^2 + V(T) &= E(T)^2 - 2\theta E(T) + \theta^2 + E\left(T^2\right) - E(T)^2 \\ &= R(T,\theta) \end{aligned}$$

1.6. Le biais $E(\lambda_0 X) - \theta$ est nul donc

$$R(\lambda_0 X,\theta) = V(\lambda_0 X) + 0^2 = \lambda_0^2 V(X)$$

et finalement, $R(\lambda_0 X,\theta) = \left(\dfrac{k+2}{k+1}\right)^2 \dfrac{k+1}{(k+2)^2(k+3)}\theta^2 = \dfrac{\theta^2}{(k+1)(k+3)}$

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 $\lambda_0 X$.

1.7. On a

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

donc, avec $Q(\lambda) = \dfrac{k+1}{k+3}\lambda^2 - 2\dfrac{k+1}{k+2}\lambda + 1$ pour tout λ réel, fonction polynôme de degré 2 dont les coefficients ne dépendent que de k on a bien pour tout λ réel

$$R(\lambda X,\theta) = \theta^2 Q(\lambda)$$

1.8. Q est dérivable sur $\mathbb{R}$ et $Q'(\lambda) = 2\dfrac{k+1}{k+3}\lambda - 2\dfrac{k+1}{k+2} = 2\dfrac{k+1}{k+3}\left(\lambda - \dfrac{k+3}{k+2}\right)$ fonction affine, positive après $\lambda^* = \dfrac{k+3}{k+2}$ et négative avant.

Conclusion : $\boxed{Q \text{ atteint son minimum en un unique réel } \lambda^* = \dfrac{k+3}{k+2}}$

1.9. Le risque quadratique

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

Les deux risques sont équivalents quand k tend vers $+\infty$. Le gain de précision est est minime.

Conclusion : très classique. fraction à simplifier.

II. Second problème d'estimation

Dans ce second problème d'estimation, on dispose de n observations indépendantes $(n \geq 2)$ notées $X_1, ..., X_n$ 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

$$Y_i = 1 \text{ si } X_i = 0 \quad \text{et} \quad Y_i = 0 \text{ sinon.}$$

Puis on note

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

1.1. Pour tout i élément de $\{1,...,n\}$, Y_i est une variable de Bernoulli et on a $\mathrm{P}(Y_i=1) = \mathrm{P}(X_i=0) = e^{-\theta}\frac{\theta^0}{0!} = e^{-\theta}$ donc $Y_i \hookrightarrow \mathcal{B}\left(e^{-\theta}\right)$

1.2. Les X_i étant indépendantes, les Y_i le sont également et donc $\sum\limits_{i=1}^{n} Y_i \hookrightarrow \mathcal{B}\left(n, e^{-\theta}\right)$.

On a alors $E\left(\overline{Y_n}\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} Y_i\right) = \frac{1}{n}n\exp(-\theta)$

Conclusion : $\boxed{E\left(\overline{Y_n}\right) = \exp(-\theta)}$.

1.3. $V\left(\overline{Y_n}\right) = \dfrac{1}{n^2}V\left(\sum\limits_{i=1}^{n} Y_i\right) = \dfrac{1}{n^2}ne^{-\theta}\left(1-e^{-\theta}\right) = \dfrac{1}{n}e^{-\theta}\left(1-e^{-\theta}\right)$

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

1.4. S_k est une somme variables de Poissons indépendates, donc suit une loi de Poisson de paramètre la somme des paramètres.

Conclusion : $\boxed{S_k \hookrightarrow \mathcal{P}(k\theta)}$

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

$$\varphi(j) = P_{\{S_n=j\}}(X_1=0)$$

1.5. Pour tout j entier naturel

$$\begin{aligned}
\varphi(j) &= \mathrm{P}_{S_n=j}(X_1=0) \\
&= \frac{\mathrm{P}(S_n=j\cap X_1=0)}{\mathrm{P}(S_n=j)}
\end{aligned}$$

On décompose

$$(S_n = j \cap X_1 = 0) = (X_1 = 0) \cap \left(\sum_{i=2}^{n} X_i = j \right)$$

et X_1 étant indépendant de $\sum_{i=2}^{n} X_i$ (qui comme précédemment suit $\mathcal{P}\left((n-1)\theta\right)$)

$$\begin{aligned} \mathrm{P}\left(S_n = j \cap X_1 = 0\right) &= \mathrm{P}\left(X_1 = 0\right) \mathrm{P}\left(\sum_{i=2}^{n} X_i = j \right) \\ &= e^{-\theta} \frac{e^{-(n-1)\theta} \left[(n-1)\theta\right]^j}{j!} \end{aligned}$$

d'où

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

On a donc $\varphi(j)$ indépendant du paramètre θ inconnu.
D'après la question II.5, on peut définir l'estimateur

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

1.6. D'après le théorème de transfert, $\varphi(S_n)$ a une espérance si

$$\sum_{j=0}^{+\infty} \left(1 - \frac{1}{n}\right)^j \mathrm{P}\left(S_n = j\right)$$

converge absolument (ce qui équivaut ici à la convergence simple)

$$\begin{aligned} \sum_{j=0}^{N} \cdots &= \sum_{j=0}^{N} \left(1 - \frac{1}{n}\right)^j \frac{e^{-n\theta}}{j!} (n\theta)^j \\ &= e^{-n\theta} \sum_{j=0}^{N} \left[\left(1 - \frac{1}{n}\right) n\theta\right]^j \frac{1}{j!} \\ &\to e^{-n\theta} e^{\left(1-\frac{1}{n}\right)n\theta} = e^{-\theta} \end{aligned}$$

Conclusion : $\boxed{\varphi(S_n) \text{ admet une espérance et que } E\left(\varphi(S_n)\right) = \exp(-\theta)}$ On dira dans ce cas que $\varphi(S_n)$ est un estimateur sans biais de $\exp(-\theta)$.

1.7. $\varphi(S_n)^2$ a une espérance si $\sum_{j=0}^{+\infty}\left(\left(1-\frac{1}{n}\right)^j\right)^2 \mathrm{P}(S_n=j)$ converge absolument.

$$\begin{aligned}\sum_{j=0}^{N}\cdots &= \sum_{j=0}^{N}\left(1-\frac{1}{n}\right)^{2j}\frac{e^{-n\theta}}{j!}(n\theta)^j \\ &= e^{-n\theta}\sum_{j=0}^{N}\left[\left(1-\frac{1}{n}\right)^2 n\theta\right]^j\frac{1}{j!} \\ &\to e^{-n\theta}e^{\left(1-\frac{1}{n}\right)^2 n\theta} = \exp\left[n\theta\left(-1+1-\frac{2}{n}+\frac{1}{n^2}\right)\right]\end{aligned}$$

Conclusion : $\boxed{\varphi(S_n)^2 \text{ admet une espérance et que } E\left(\varphi(S_n)^2\right)=\exp\left(\left(-2+\frac{1}{n}\right)\theta\right)}$

Donc $\varphi(S_n)$ a une variance qui vaut

$$\begin{aligned}V(\varphi(S_n)) &= E\left(\varphi(S_n)^2\right)-E(\varphi(S_n))^2 \\ &= \exp\left(\left(-2+\frac{1}{n}\right)\theta\right)-\exp(-2\theta) \\ &= \exp(-2\theta)\left[\exp\left(\frac{1}{n}\theta\right)-1\right]\end{aligned}$$

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

a) Pour tout $t\in[0,\theta]$ on a $\exp'=\exp$ donc $\exp'(0)\le\exp'(t)\le\exp'(\theta)$
et comme $0\le\theta$ alors, d'après linégalité des accroissements finis (sans valeur absolue, donc avec l'ordre des termes)

$$(\theta-0)\,1\le\exp(\theta)-\exp(0)\le(\theta-0)\exp(\theta)$$

et donc

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

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

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

h est dérivable sur $[0,1]$ et

$$\begin{aligned}h'(x) &= \exp(\theta)-1-\theta\exp(t\theta) \\ &= \theta\left(\frac{\exp(\theta)-1}{\theta}-\exp(t\theta)\right)\end{aligned}$$

Pour avoir son signe, il faut se creuser la tête et faire apparaître les inégalités précédentes.
Avec $h'(0)=\theta\left(\dfrac{\exp(\theta)-1}{\theta}-1\right)\ge 0$
et $h'(1)=\theta\left(\frac{\exp(\theta)-1}{\theta}-\exp(\theta)\right)\le 0$
h' étant continue et strictement décroissante sur$[0,1]$, elle est bijective de $[0,1]$ sur $[h'(1),h'(0)]$
Et 0 appartenant à cet intervalle, il y a un unique $\alpha\in[0,1]$ tel que $h'(\alpha)=0$
D'où le signe de h'

t	0		α		1
$h'(t)$		$\searrow +$	0	$\searrow -$	
$h(t)$	0	$\nearrow$		$\searrow$	0

c) Donc pour tout $t \in [0,1]$, $h(t) \geq 0$.
en particulier en $t = \frac{1}{n}$ où

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

et donc

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

On reprend alors

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

et

$$V\left(\overline{Y_n}\right) = e^{-\theta}\left(1 - e^{-\theta}\right)$$

donc

$$\begin{aligned} V(\varphi(S_n)) - V\left(\overline{Y_n}\right) &= \exp(-2\theta)\left[\exp\left(\frac{1}{n}\theta\right) - 1\right] - e^{-\theta}\left(1 - e^{-\theta}\right) \\ &= \exp(-2\theta)\left[\exp\left(\frac{1}{n}\theta\right) - \exp(\theta)\right] \end{aligned}$$

et on réintroduit

$$\exp\left(\frac{\theta}{n}\right) \leq \frac{\exp(\theta)}{n} + \frac{n-1}{n} \text{ d'où}$$

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

et

$$\begin{aligned} \frac{\exp(\theta)}{n} + \frac{n-1}{n} - \exp(\theta) &= \left(\frac{1}{n} - 1\right)\exp(\theta) + \frac{n-1}{n} \\ &= \frac{1-n}{n}(\exp(\theta) - 1) \\ &\geq 0 \end{aligned}$$

et finalement $V(\varphi(S_n)) - V\left(\overline{Y_n}\right) \geq 0$

Conclusion : $\boxed{V(\varphi(S_n)) \leq V\left(\overline{Y_n}\right)}$

d) On définit le risque quadratique de T_n estimateur de $\exp(-\theta)$ par

$$R(T_n, \exp(-\theta)) = E\left((T_n - \exp(-\theta))^2\right)$$

On reconnait le risque quadratique $.R(T_n, \exp(-\theta)) = V(T_n) + b(T_n)^2$ (th du cours) où $b(T_n) = E(T_n) - \exp(-\theta)$

Et comme $\varphi(S_n)$ et $\overline{Y_n}$ sont sans biais, $R(\varphi(S_n), \exp(-\theta)) = V(\varphi(S_n))$ et de même pour $\overline{Y_n}$.

De plus $V(\varphi(S_n)) \leq V\left(\overline{Y_n}\right)$

Conclusion : le risque quadratique de $\overline{Y_n}$ est plus grand que celui de $\varphi(S_n)$

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

2 Information de Fisher

A. Cas discret

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

$$P(X = k) = p(\theta, k)$$

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

$$I_X(\theta) = \sum_{k \in X(\Omega)} \left(\frac{\partial}{\partial \theta} \ln p(\theta, k) \right)^2 p(\theta, k)$$

2A.1. Dans cette question 1, on considère X variable aléatoire de loi de Bernoulli de paramètre θ ($\theta \in]0, 1[$).
On a alors $X(\Omega) = \{0, 1\}$, $P(X = 1) = p(\theta, 1) = \theta$, $P(X = 0) = p(\theta, 0) = 1 - \theta$ et

$$I_X(\theta) = \left(\frac{\partial}{\partial \theta} \ln p(\theta, 1) \right)^2 p(\theta, 1) + \left(\frac{\partial}{\partial \theta} \ln p(\theta, 0) \right)^2 p(\theta, 0)$$

On calcule les dérivées (dérivables car $\theta > 0$ et $1 - \theta > 0$) :

$$\frac{\partial}{\partial \theta} \ln p(\theta, 1) = \frac{\partial}{\partial \theta} \ln(\theta) = \frac{1}{\theta}$$
$$\frac{\partial}{\partial \theta} \ln p(\theta, 0) = \frac{\partial}{\partial \theta} (\ln(1 - \theta)) = \frac{-1}{1 - \theta}$$

Donc

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

2A.2. Dans cette question 2, on considère X variable aléatoire de loi binomiale de paramètres N et θ ($N \in \mathbb{N}^*$, $\theta \in]0, 1[$)

a) (on dérive la puissance 0 sous forme $0 \cdot \theta^{-1}$)
On a donc pour tout $k \in [[0, N]]$:

$$\begin{aligned} p(\theta, k) &= \binom{N}{k} \theta^k (1 - \theta)^{N-k} \text{ et } (\theta \text{ et } 1 - \theta > 0) \\ \ln(p(\theta, k)) &= \ln\left(\binom{N}{k} \theta^k (1 - \theta)^{N-k} \right) = \binom{N}{k} [k \ln(\theta) + (N - k) \ln(1 - \theta)] \\ \frac{\partial}{\partial \theta} \ln p(\theta, k) &= \binom{N}{k} \left[\frac{k}{\theta} - \frac{N - k}{1 - \theta} \right] = \binom{N}{k} \frac{k - N\theta}{\theta(1 - \theta)} \end{aligned}$$

Donc

$$\begin{aligned}
I_X(\theta) &= \sum_{k\in X(\Omega)} \frac{\partial}{\partial\theta}\ln p(\theta,k)\, p(\theta,k) \\
&= \\
&= \sum_{k=0}^{N}\left[\frac{k-N\theta}{\theta(1-\theta)}\right]^2 \binom{N}{k}\theta^k(1-\theta)^{N-k} \\
&= \frac{1}{(\theta(1-\theta))^2}\sum_{k=0}^{N}(k-N\theta)^2\binom{N}{k}\theta^k(1-\theta)^{N-k}
\end{aligned}$$

b) On reconnaît

$$\begin{aligned}
\sum_{k=0}^{N}(k-N\theta)^2\binom{N}{k}\theta^k(1-\theta)^{N-k} &= \sum_{k=0}^{N}(k-N\theta)^2\,\mathrm{P}(X=k) \\
&= E\left((X-N\theta)^2\right) \text{ th de transfert} \\
&= E\left[(X-E(X))^2\right] \\
&= V(X)
\end{aligned}$$

Conclusion : $\boxed{I_X(\theta) = \frac{V(X)}{(\theta(1-\theta))^2} = \frac{N\theta(1-\theta)}{(\theta(1-\theta))^2} = \frac{N}{\theta(1-\theta)}}$

2A.3. Dans cette question 3, on considère X une variable aléatoire de Poisson de paramètre θ ($\theta \in]0,+\infty[$). Puisque $X(\Omega)=\mathbb{N}$, on a sous réserve de convergence

$$I_X(\theta) = \sum_{k=0}^{+\infty}\left(\frac{\partial}{\partial\theta}\ln p(\theta,k)\right)^2 p(\theta,k)$$

a) Suivant le même shéma d'étude :

$$\begin{aligned}
p(\theta,k) &= \theta^k\frac{\exp(-\theta)}{k!} \text{ et} \\
\ln p(\theta,k) &= k\ln(\theta) - \theta - \ln(k!) \\
\frac{\partial}{\partial\theta}\ln p(\theta,k) &= \frac{k}{\theta} - 1
\end{aligned}$$

On étudie la somme partielle

$$\begin{aligned}
&\sum_{k=0}^{N}\left(\frac{\partial}{\partial\theta}\ln p(\theta,k)\right)^2 p(\theta,k) \\
&= \sum_{k=0}^{N}\left(\frac{k}{\theta}-1\right)^2\theta^k\frac{\exp(-\theta)}{k!} \\
&= \frac{1}{\theta^2}\exp(-\theta)\sum_{k=0}^{N}\left(k^2-2k\theta+\theta^2\right)\frac{\theta^k}{k!}
\end{aligned}$$

Astuce : on réécrit $k^2 = k(k-1) + k$ pour pouvoir le simplifier avec la factorielle

$$\begin{aligned}
\sum_{k=0}^{N} \cdots &= \sum_{k=0}^{N} \left[k(k-1) + k(1-2\theta) + \theta^2\right] \frac{\theta^k}{k!} \\
&= \sum_{k=2}^{N} \frac{\theta^k}{(k-2)!} + (1-2\theta) \sum_{k=1}^{N} \frac{\theta^k}{(k-1)!} + \theta^2 \sum_{k=0}^{N} \frac{\theta^k}{k!} \text{ réindexé} \\
&= \theta^2 \sum_{k=0}^{N-2} \frac{\theta^k}{k!} + \left(\theta - 2\theta^2\right) \sum_{k=0}^{N-1} \frac{\theta^k}{k!} + \theta^2 \sum_{k=0}^{N} \frac{\theta^k}{k!} \\
&\to \theta \exp(\theta)
\end{aligned}$$

Donc la série de terme général $\left(\frac{\partial}{\partial\theta} \ln p(\theta, k)\right)^2 p(\theta, k)$ converge et vaut

Conclusion : $\boxed{I_X(\theta) = \frac{1}{\theta}}$

b) Par absolue convergence (tous les termes sont positifs grace au carré) le théorème de transfert nous donne

$$\begin{aligned}
E\left(\left(\frac{\partial}{\partial\theta} \ln p(\theta, X)\right)^2\right) &= \sum \left(\frac{\partial}{\partial\theta} \ln p(\theta, k)\right)^2 \mathrm{P}(X = k) \\
&= I_X(\theta)
\end{aligned}$$

B. 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

$$I_X(\theta) = \int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial\theta} \ln(f(\theta, x))\right)^2 f(\theta, x)\, dx$$

2A.1. On a donc

$$\begin{aligned}
f(\theta, x) &= \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2} \text{ et} \\
\ln f(\theta, x) &= -\frac{(x-\theta)^2}{2} - \ln\left(\sqrt{2\pi}\right) \text{ donc} \\
\frac{\partial}{\partial\theta} \ln(f(\theta, x)) &= -(x-\theta) \text{ et donc} \\
\left(\frac{\partial}{\partial\theta} \ln(f(\theta, x))\right)^2 &= (x-\theta)^2
\end{aligned}$$

Conclusion : $\boxed{\begin{array}{l}\text{sous réserve de convergence} \\ I_X(\theta) = \displaystyle\int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta, x)\, dx\end{array}}$

2A.2. On reconnait (th de transfert avec convergence absolue qui équivaut ici à la convergence simple) dans cette quantité l'espérance

$$\begin{aligned}
I_X(\theta) &= E\left[(X-\theta)^2\right] \\
&= E\left[(X-E)X)^2\right] = V(X)
\end{aligned}$$

Or X a une variance qui vaut 1

Conclusion : $\boxed{I_X(\theta) \text{ existe et vaut } 1}$

2A.3. et on peut le lire également sous la forme

$$I_X(\theta) = E\left(\left(\frac{\partial}{\partial\theta}\ln(f(\theta, X))\right)^2\right)$$

IV. Minoration du risque quadratique

A. Inégalité de Cramer-Rao

Dans cette section IV.A, on considère I un intervalle de $\mathbb{R}$, θ un paramètre inconnu de 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 $k \in \{0, ..., N\}$

$$P(X = k) = p(\theta, k)$$

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.A est de démontrer l'inégalité suivante dûe à Cramer et Rao.

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

$$V(f(X)) \geq \frac{(g'(\theta))^2}{I_X(\theta)}$$

2A.1. Pour tout θ élément de I, on a $X(\Omega) = \{0, ..., N\}$ donc

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

et donc

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

2A.2. On a donc

$$\frac{\partial}{\partial\theta}\ln(p(\theta, k)) = \frac{1}{p(\theta, k)}\frac{\partial}{\partial\theta}p(\theta, k)$$

et d'aprèsle théorème de transfert,

$$\begin{aligned} E\left(\frac{\partial}{\partial\theta}\ln(p(\theta, X))\right) &= \sum_{k=0}^{N} \frac{\partial}{\partial\theta}(\ln p(\theta, k))\, p(\theta, k) \\ &= \sum_{k=0}^{N} \frac{\partial}{\partial\theta}p(\theta, k) = 0 \end{aligned}$$

Donc

$$E\left(\frac{\partial}{\partial\theta}\ln(p(\theta, X))\right) = 0 \qquad \text{(E)}$$

2A.3. On repart de

$$\sum_{k=0}^{N}\frac{\partial}{\partial\theta}\left(\ln p\left(\theta,k\right)\right)p\left(\theta,k\right)=0$$

que l'on redérive :

$$\sum_{k=0}^{N}\frac{\partial^{2}}{\partial\theta^{2}}\left(\ln p\left(\theta,k\right)\right)p\left(\theta,k\right)+\frac{\partial}{\partial\theta}\left(\ln p\left(\theta,k\right)\right)\frac{\partial}{\partial\theta}p\left(\theta,k\right)=0$$

donc

$$\sum_{k=0}^{N}\frac{\partial^{2}}{\partial\theta^{2}}\left(\ln p\left(\theta,k\right)\right)p\left(\theta,k\right)=-\sum_{k=0}^{N}\frac{\partial}{\partial\theta}\left(\ln p\left(\theta,k\right)\right)\frac{\partial}{\partial\theta}p\left(\theta,k\right)$$

et en reprenant par la droite :

$$E\left(\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)^{2}\right)=\sum_{k=0}^{N}\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,k\right)\right)\right)^{2}p\left(k,\theta\right)$$

dans laquelle on conserve une $\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,k\right)\right)$ intacte pour retrouver l'expression précédente

$$\begin{aligned}E\left(\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)^{2}\right)&=\sum_{k=0}^{N}\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,k\right)\right)\left(\frac{1}{p\left(\theta,k\right)}\frac{\partial}{\partial\theta}p\left(\theta,k\right)\right)p\left(k,\theta\right)\\&=\sum_{k=0}^{N}\frac{\partial}{\partial\theta}\left(\ln p\left(\theta,k\right)\right)\frac{\partial}{\partial\theta}p\left(\theta,k\right)\end{aligned}$$

et finalement,

$$E\left(\frac{\partial^{2}}{\partial\theta^{2}}\ln\left(p\left(\theta,X\right)\right)\right)=-E\left(\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)^{2}\right)$$

2A.4. On a $g\left(\theta\right)=E\left(f\left(X\right)\right)$ donc

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

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

et comme

$$\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,k\right)\right)\right)\ p\left(\theta,k\right)=\frac{\partial}{\partial\theta}\left(p\left(\theta,k\right)\right)$$

Conclusion : $\boxed{g'\left(\theta\right)=\sum_{k=0}^{N}f\left(k\right)\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,k\right)\right)\right)\ p\left(\theta,k\right)}$

et comme $E\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)=0$ alors $E\left(g\left(\theta\right)\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)=0$ puis

$$\begin{aligned}E\left(\left(f\left(X\right)-g\left(\theta\right)\right)\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)&=E\left(f\left(X\right)\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)-E\left(g\left(\theta\right)\frac{\partial}{\partial\theta}\ln\left(p\left(\theta,X\right)\right)\right)\\&=g'\left(\theta\right)\end{aligned}$$

2A.5. On pose pour tout t réel

$$L(t) = E\left(\left((f(X) - g(\theta)) + t\frac{\partial}{\partial\theta}[\ln p(\theta, X)]\right)^2\right)$$

a) On développe par rapport à t

$$\begin{aligned}
L(t) &= E\left(t^2\left(\frac{\partial}{\partial\theta}[\ln p(\theta, X)]\right)^2 + 2t\frac{\partial}{\partial\theta}[\ln p(\theta, X)](f(X) - g(\theta)) + (f(X) - g(\theta))^2\right) \\
&= E\left(\left[\frac{\partial}{\partial\theta}[\ln p(\theta, X)]\right]^2\right)\cdot t^2 + 2E\left(\frac{\partial}{\partial\theta}[\ln p(\theta, X)][f(X) - g(\theta)]\right)\cdot t \\
&+ E\left([f(X) - g(\theta)]^2\right)
\end{aligned}$$

b) $L(t)$ étant l'espérance d'un carrée, elle est positive ou nulle.
Les polynômes du second degré positif ou nul sur $\mathbb{R}$ ont au plus une racine et ont donc un discriminant négatif (aucune solution) ou nul (une seule) .
On le calcule :

$$\begin{aligned}
\Delta &= 4E\left(\frac{\partial}{\partial\theta}[\ln p(\theta, X)][f(X) - g(\theta)]\right)^2 - 4E\left(\left[\frac{\partial}{\partial\theta}[\ln p(\theta, X)]\right]^2\right)E\left([f(X) - g(\theta)]^2\right) \\
&= 4\left[E\left(\frac{\partial}{\partial\theta}[\ln p(\theta, X)][f(X) - g(\theta)]\right)^2 - E\left(\left[\frac{\partial}{\partial\theta}[\ln p(\theta, X)]\right]^2\right)E\left([f(X) - g(\theta)]^2\right)\right] \\
&\leq 0
\end{aligned}$$

c) On se souvient que

$$\begin{aligned}
4)\ g'(\theta) &= E\left((f(X) - g(\theta))\frac{\partial}{\partial\theta}\ln(p(\theta, X))\right) \\
I_X(\theta) &= E\left(\left[\frac{\partial}{\partial\theta}[\ln p(\theta, X)]\right]^2\right) \quad \text{et} \\
V(f(X)) &= E\left([f(X) - E(f(X))]^2\right) \\
&= E\left([f(X) - g(\theta)]^2\right)
\end{aligned}$$

On a finalement

$$g'(\theta)^2 \leq I_X(\theta)\cdot V(f(X))$$

d'où l'inégalité de Cramer-Rao.

$$V(f(X)) \geq \frac{g'(\theta)^2}{I_X(\theta)}$$

B. Extension du théorème de Cramer-Rao

On reprend dans cette section IV.B 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 de Cramer-Rao
Soit $T_n = f(X_1, ..., X_n)$ un estimateur sans biais de $g(\theta)$ à savoir tel que

$E\left(f\left(X_1,...,X_n\right)\right) = g\left(\theta\right)$ **où** g **est dérivable sur** $]0,+\infty[$**. On a alors**

$$V\left(T_n\right) \geq \frac{\left(g'\left(\theta\right)\right)^2}{nI_{X_1}\left(\theta\right)}$$

où $I_{X_1}\left(\theta\right)$ **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

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

2A.1. On a

$$\begin{aligned} E\left(\overline{X_n}\right) &= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) \\ &= \theta \end{aligned}$$

et

$$\begin{aligned} V\left(\overline{X_n}\right) &= \frac{1}{n^2}\sum_{i=1}^{n} V\left(X_i\right) \text{ par indépendance} \\ &= \frac{\theta}{n} \end{aligned}$$

2A.2. Pour un estimateur sans biais, le risque quadratique est l'espérance.
On estime $g\left(\theta\right) = \theta$. Donc $g'\left(\theta\right) = 1$
et l'information de Fischer est

$$I_{X_1}\left(\theta\right) = \frac{1}{\theta}$$

donc pour tout estimateur sans biais T_n son risque quadratique est (généralisation de Cramer-Rao)

$$V\left(T_n\right) \geq \frac{\theta}{n} = V\left(\overline{X_n}\right)$$

Conclusion : $\boxed{\overline{X_n} \text{ a le plus petit risque quadratique parmi les estimateurs sans biais de } \theta.}$

2A.3. pour $g\left(\theta\right) = \exp\left(-\theta\right)$ où $\theta \in]0,+\infty[$ on a $g'\left(\theta\right) = -\exp\left(-\theta\right)$
et $I_{X_1}\left(\theta\right) = \dfrac{1}{\theta}$
enfin d'après II 7. $V\left(\varphi\left(S_n\right)\right) = \exp\left(-2\theta\right)\left[\exp\left(\dfrac{1}{n}\theta\right) - 1\right]$
et comme $\exp\left(x\right) - 1 \sim x$ quand $x \to 0$ alors $\exp\left(\dfrac{1}{n}\theta\right) - 1 \sim \dfrac{\theta}{n}$ quand $n \to +\infty$
alors

$$\begin{aligned} V\left(\varphi\left(S_n\right)\right) &= \exp\left(-2\theta\right)\left[\exp\left(\frac{1}{n}\theta\right) - 1\right] \\ &\underset{n\to+\infty}{\sim} \left[-\exp\left(-\theta\right)\right]^2 \frac{\theta}{n} \end{aligned}$$

Conclusion : $\boxed{V\left(\varphi\left(S_n\right)\right) \underset{n\to+\infty}{\sim} \dfrac{\left(g'\left(\theta\right)\right)^2}{nI_{X_1}\left(\theta\right)}}$

2A.4. Assymptotiquement, $\varphi\left(S_n\right)$ donne donc le risque quadratique minimal parmi les estimateurs sans biais de $\exp\left(-\theta\right)$

2A.5. Mais un estimateur biaisé peut donner un risque quadratique moindre.
Donc, lorsque n est grand $\varphi\left(S_n\right)$ n'est pas focément le meilleur estimateur de $\exp\left(-\theta\right)$.