DS8 - Concours blanc - Barème (`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.

- **1 pt : pour tout $x \in \mathbb{R}$, $f_\theta(x) \geqslant 0$**
- **1 pt : la fonction f_θ est continue sur $\mathbb{R}$ sauf éventuellement en 0 et en 1**
- **1 pt :**
$$\begin{aligned}\int_{-\infty}^{+\infty} f_\theta(x)\, dx &= \int_0^\theta f_\theta(x)\, dx && \textit{\textbf{car } } f_\theta \textit{\textbf{ est nulle en dehors de }} [0,\theta] \\ &= \int_0^\theta \frac{k+1}{\theta^{k+1}}\, x^k\, dx \\ &= \frac{k+1}{\theta^{k+1}} \int_0^\theta x^k\, dx\end{aligned}$$
- **1 pt : $\displaystyle\int_{-\infty}^{+\infty} f_\theta(x)\, dx$ est convergente puisque $x \mapsto x^k$ est continue sur $[0, \theta]$**

- **1 pt :** $\displaystyle\int_{-\infty}^{+\infty} f_\theta(x)\ dx = \frac{k+1}{\theta^{k+1}}\frac{\theta^{k+1}}{k+1} = 1$

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

- **1 pt :** 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**
- **1 pt :** X **admet une espérance**
- **2 pt :** $\mathbb{E}(X) = \dfrac{k+1}{k+2}\theta$

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

- **1 pt :** $\lambda_0 X$ **admet une espérance comme transformée affine de** X
- **1 pt :** $\lambda_0 = \dfrac{k+2}{k+1}$
- **1 pt : linéarité de l'espérance**

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

- **1 pt :** 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**
- **1 pt :** X **admet une variance**
- **1 pt :** $\mathbb{E}(X^2) = \dfrac{k+1}{k+3}\theta^2$
- **2 pt :** $\mathbb{V}(X) = \dfrac{(k+1)}{(k+3)(k+2)^2}\theta^2$

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

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

- **1 pt : linéarité de l'espérance citée**
- **2 pt : calcul**

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 $\lambda_0\, X$.

- **1 pt :** $\lambda_0 X$ **admet une variance comme transformée affine de** X **qui admet une variance, elle admet donc un risque quadratique**
- **2 pt :** $r_\theta(\lambda_0 X) = \dfrac{\theta^2}{(k+1)(k+3)}$

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.

- **3 pt :** $Q(\lambda) = \dfrac{k+1}{k+3}\lambda^2 - 2\dfrac{k+1}{k+2}\lambda + 1$

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.

- **1 pt :** $Q'(\lambda) = 2\dfrac{k+1}{k+3}\lambda - 2\dfrac{k+1}{k+2} = 2\dfrac{k+1}{k+3}\left(\lambda - \dfrac{k+3}{k+2}\right)$

- **1 pt : Q atteint son minimum global en un unique point qui est $\lambda^* = \dfrac{k+3}{k+2}$**
- **1 pt : explications**

9. Conclure sur le but recherché.

- **1 pt : $\lambda^* X$ est un estimateur de θ de risque quadratique minimal parmi tous les estimateurs de θ de la forme λX**
- **1 pt : $r_\theta(\lambda^* X) \underset{k\to+\infty}{\sim} \dfrac{\theta^2}{k^2} \underset{k\to+\infty}{\sim} r_\theta(\lambda_0 X)$**
- **1 pt : le gain est très faible lorsque k est grand**

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.

- **1 pt : $Y_i(\Omega) = \{0, 1\}$ donc Y_i suit une loi de Bernoulli de paramètre $\alpha = \mathbb{P}([Y_i = 1])$**
- **1 pt : $Y_i \hookrightarrow \mathcal{B}\left(\mathrm{e}^{-\theta}\right)$**

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

- **1 pt : les variables aléatoires Y_i sont mutuellement indépendantes**
- **1 pt : stabilité des lois binomiales**
- **1 pt : $\sum\limits_{i=1}^{n} Y_i \hookrightarrow \mathcal{B}\left(n, \mathrm{e}^{-\theta}\right)$**
- **1 pt : $\overline{Y}_n$ admet une espérance comme transformée affine de $\sum\limits_{i=1}^{n} Y_i$ qui admet une espérance**
- **1 pt : $\mathbb{E}(\overline{Y}_n) = \dfrac{1}{n} n \mathrm{e}^{-\theta} = \exp(-\theta)$**
- **1 pt : linéarité de l'espérance (donné si calcul correct)**

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

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

- **1 pt : $\overline{Y}_n$ admet une variance comme transformée affine de $\sum\limits_{i=1}^{n} Y_i$ qui admet une variance**
- **1 pt : $\mathbb{V}\left(\overline{Y}_n\right) = \dfrac{\mathrm{e}^{-\theta}\left(1 - \mathrm{e}^{-\theta}\right)}{n}$**

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 $[\![1,n]\!]$.

- **1 pt :** $S_k \hookrightarrow \mathcal{P}(k\theta)$

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 :

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

- **2 pt :** $\varphi(j) = \dfrac{\mathbb{P}\left(\left[\sum\limits_{i=2}^{n} X_i = j\right] \cap [X_1=0]\right)}{\mathrm{e}^{-n\theta}\dfrac{(n\theta)^j}{j!}}$
- **1 pt : par lemme des coalitions, les variables aléatoires X_1 et $\sum\limits_{i=2}^{n} X_i$ sont indépendantes**
- **1 pt : les variables aléatoires $\sum\limits_{i=2}^{n} X_i$ et S_{n-1} suivent la même loi**
- **1 pt :** $\varphi(j) = \left(1-\dfrac{1}{n}\right)^j$

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 $\mathbb{E}\big(\varphi(S_n)\big) = \exp(-\theta)$.

- **1 pt : par 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 :**

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

- **1 pt :** $\sum\limits_{k=0}^{N} \varphi(k)\mathbb{P}([S_n=k]) = \sum\limits_{k=0}^{N} \left(1-\dfrac{1}{n}\right)^k \mathrm{e}^{-n\theta}\dfrac{(n\theta)^k}{k!} = \mathrm{e}^{-n\theta}\sum\limits_{k=0}^{N} \dfrac{(n\theta-\theta)^k}{k!}$
- **1 pt : on reconnaît la somme partielle d'une série exponentielle de paramètre $x = n\theta-\theta$**
- **1 pt :** $\mathbb{E}\big(\varphi(S_n)\big) = \exp(-\theta)$

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 :

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

- **1 pt : par 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 :**

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

- **1 pt :** $\sum_{k=0}^{N} \varphi(k)^2 \mathbb{P}([S_n = k]) = \sum_{k=0}^{N} \left(1 - \frac{1}{n}\right)^{2k} \mathbf{e}^{-n\theta} \frac{(n\theta)^k}{k!} = \mathbf{e}^{-n\theta} \sum_{k=0}^{N} \frac{\left(n\theta - 2\theta + \frac{\theta}{n}\right)^k}{k!}$
- **1 pt : on reconnaît la somme partielle d'une série exponentielle de paramètre** $x = n\theta - 2\theta + \frac{\theta}{n}$
- **1 pt :** $\varphi(S_n)^2$ **admet une espérance et** $\mathbb{E}\big(\varphi(S_n)^2\big) = \mathbf{e}^{-2\theta}\mathbf{e}^{\frac{\theta}{n}}$
- **1 pt :** $\mathbb{V}\big(\varphi(S_n)\big) = \exp(-2\theta)\left(\exp\left(\frac{\theta}{n}\right) - 1\right)$

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

a) Démontrer :

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

- **2 pt :** $\mathbf{e}^{\theta} \geqslant \theta + 1$ **par inégalité de convexité (1 pt pour les détails)**
- **2 pt :** $\theta\mathbf{e}^{\theta} - \mathbf{e}^{\theta} + 1 \geqslant 0$ **par étude de fonction**

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.

- **1 pt :** h **est dérivable sur** $[0,1]$
- **1 pt :** $h'(t) = \mathbf{e}^{\theta} - 1 - \theta\mathbf{e}^{t\theta}$
- **1 pt : tableau de variations**

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

- **1 pt : calcul de** $t^* = \frac{1}{\theta}\ln\left(\frac{\mathbf{e}^{\theta} - 1}{\theta}\right)$
- **1 pt : justification de** $t^* \in [0,1]$

c) En déduire :

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

puis l'inégalité :

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

- **2 pt : pour tout** $t \in [0,1]$**,** $h(t) \geqslant 0$ **et évaluation en** $t = \frac{1}{n} \in [0,1]$ **(1 pt pour préciser que** $t \in [0,1]$**)**
- **2 pt :** $\mathbb{V}\big(\varphi(S_n)\big) \leqslant \mathbb{V}\big(\overline{Y}_n\big)$

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

- **1 pt :** $r_{\mathrm{e}^{-\theta}}\left(\overline{Y}_n\right) = \mathbb{V}\left(\overline{Y}_n\right)$ **et** $r_{\mathrm{e}^{-\theta}}\left(\varphi(S_n)\right) = \mathbb{V}\left(\varphi(S_n)\right)$
- **1 pt :** $r_{\mathrm{e}^{-\theta}}\left(\varphi(S_n)\right) \leqslant r_{\mathrm{e}^{-\theta}}\left(\overline{Y}_n\right)$

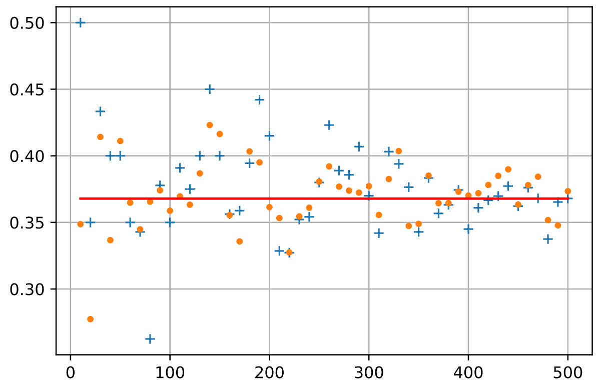
18. a) Compléter la fonction **Python** suivante pour qu'elle simule les estimateurs $\overline{Y}_n$ et $\varphi(S_n)$.

```
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]
```

- **1 pt :** `Xobs = rd.poisson(theta, n)`
- **2 pt :** `if Xobs[k] == 0:`
- **1 pt :** `phi = (1-1/n)**(np.sum(Xobs))`
- **1 pt : bonus si tout est correct**

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()
```



- **1 pt : les • 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** $\overline{Y}_n$
- **1 pt : 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**

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 :

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

- **1 pt :** $\partial_1(h)(\theta,1) = \dfrac{1}{\theta}$ **et** $\partial_1(h)(\theta,0) = \dfrac{-1}{1-\theta}$
- **1 pt :** $I_X(\theta) = \dfrac{1}{\theta\,(1-\theta)}$

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 :

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

- **1 pt :** $p(\theta,k) = \mathbb{P}([X=k]) = \dbinom{N}{k}\theta^k(1-\theta)^{N-k}$
- **1 pt :** $h(\theta,k) = \ln\left(\dbinom{N}{k}\right) + k\ln(\theta) + (N-k)\ln(1-\theta)$
- **1 pt :** $\partial_1(h)(\theta,k) = \dfrac{k-N\theta}{\theta(1-\theta)}$
- **1 pt :** $I_X(\theta) = \dfrac{1}{\big(\theta\,(1-\theta)\big)^2} \displaystyle\sum_{k=0}^{N} (k-N\theta)^2 \binom{N}{k}\theta^k(1-\theta)^{N-k}$

b) En déduire :

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

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

- **1 pt :** $\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left((X-N\theta)^2\right)$
- **1 pt : thm de transfert**
- **1 pt :** $I_X(\theta) = \dfrac{N}{\theta(1-\theta)}$

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 $I_X(\theta)$.

- **1 pt :** $\partial_1(h)(\theta,k) = -1 + \dfrac{k}{\theta}$
- **2 pt :** $\sum\limits_{k=0}^{N} \big(\partial_1(h)(\theta,k)\big)^2 p(\theta,k) = \mathrm{e}^{-\theta}\left(\sum\limits_{k=0}^{N-2} \dfrac{\theta^k}{k!} + \dfrac{1}{\theta}\sum\limits_{k=0}^{N-1} \dfrac{\theta^k}{k!} - 2\sum\limits_{k=0}^{N-1} \dfrac{\theta^k}{k!} + \sum\limits_{k=0}^{N} \dfrac{\theta^k}{k!}\right)$
- **1 pt : on reconnaît des sommes partielles de la série exponentielle de paramètre θ**
- **2 pt :** $I_X(\theta) = \mathrm{e}^{-\theta}\left(\mathrm{e}^{\theta} + \dfrac{1}{\theta}\mathrm{e}^{\theta} - 2\mathrm{e}^{\theta} + \mathrm{e}^{\theta}\right) = \dfrac{1}{\theta}$

b) Justifier :

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

- **1 pt : 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 :**

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

- **1 pt : 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**

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 :

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

- **1 pt :** $f(\theta,x) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(x-\theta)^2}$
- **1 pt :** $(\ln\circ f)(\theta,x) = -\ln(\sqrt{2\pi}) - \dfrac{1}{2}(x-\theta)^2$
- **1 pt :** $\partial_1(\ln\circ f)(\theta,x) = x-\theta$

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

- **1 pt : l'intégrale $\displaystyle\int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta,x)\ dx$ converge**
- **3 pt :** $I_X(\theta) = 1$
  - × **soit en faisant le changement de variable $t = x-\theta$**
  - × **soit en développant $(x-\theta)^2 f(\theta,x) = x^2 f(\theta,x) - 2\theta x f(\theta,x) + \theta^2 f(\theta,x)$ et en calculant plusieurs moments de X**

$$\int_{-\infty}^{+\infty} f(\theta,x)\ dx = \mathbb{E}(X^0) = 1$$
$$\int_{-\infty}^{+\infty} x f(\theta,x)\ dx = \mathbb{E}(X) = \theta$$
$$\int_{-\infty}^{+\infty} x^2 f(\theta,x)\ dx = \mathbb{E}(X^2) = \mathbb{V}(X) + \mathbb{E}(X)^2 = 1+\theta^2$$

24. Justifier :

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

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

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

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

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 :

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

25. Montrer que pour tout θ élément de I :

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

- **1 pt :** $([X = k])_{k \in [\![0,N]\!]}$ **est un système complet d'événements**
- **1 pt :** $\displaystyle\sum_{k=0}^{N} p(\theta, k) = 1$
- **1 pt : dérivation par rapport à la première variable**

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

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

- **1 pt :** $\partial_1(h)(\theta, X)$ **est finie car** X **est finie donc** $\partial_1(h)(\theta, X)$ **admet une espérance**
- **1 pt : thm de transfert**

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

- **1 pt :** $\partial_1(h)(\theta,k) = \dfrac{\partial_1(p)(\theta,k)}{p(\theta,k)}$
- **1 pt :** $\mathbb{E}\left(\partial_1(h)(\theta,X)\right) = 0$

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

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

- **1 pt : en dérivant partiellement par rapport à θ :**

$$\sum_{k=0}^{N}\Big(\partial^2_{1,1}(h)(\theta,k)p(\theta,k) + \partial_1(h)(\theta,k)\partial_1(p)(\theta,k)\Big) = 0$$

- **1 pt :** $\sum\limits_{k=0}^{N} \partial^2_{1,1}(h)(\theta,k)p(\theta,k) = -\sum\limits_{k=0}^{N} \partial_1(h)(\theta,k)^2 p(\theta,k)$
- **1 pt : thm de transfert**

28. Montrer que pour tout θ élément de I :

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

puis :

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

- **1 pt : par théorème de transfert :**

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

- **1 pt :** $g'(\theta) = \sum\limits_{k=0}^{N} f(k)\partial_1(h)(\theta,k)p(\theta,k)$
- **1 pt :** $\big(f(X)-g(\theta)\big)\big(\partial_1(h)(\theta,X)\big)$ **est finie car X est finie, donc admet une espérance**
- **1 pt : linéarité de l'espérance**
- **1 pt :** $g'(\theta) = \mathbb{E}\Big(\big(f(X)-g(\theta)\big)\big(\partial_1(h)(\theta,X)\big)\Big)$

29. On pose pour tout t réel :

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

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

- **1 pt :** $L(t) = \mathbb{E}\left(\partial_1(h)(\theta,X)^2\right)t^2 + 2\mathbb{E}\Big(\big(f(X)-g(\theta)\big)\partial_1(h)(\theta,X)\Big)t + \mathbb{E}\left(\big(f(X)-g(\theta)\big)^2\right)$
- **1 pt : linéarité de l'espérance**
- **1 pt :** $\mathbb{E}\left(\big(f(X)-g(\theta)\big)^2\right) = \mathbb{E}\left(\big(f(X)-\mathbb{E}(f(X))\big)^2\right) = \mathbb{V}(f(X))$
- **1 pt :** $\mathbb{E}\left(\partial_1(h)(\theta,X)^2\right) = I_X(\theta)$
- **1 pt :** $L(t) = I_X(\theta)t^2 + 2g'(\theta)t + \mathbb{V}(f(X))$

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

- **1 pt :** $\Delta = 4\left((g'(\theta))^2 - I_X(\theta)\mathbb{V}(f(X))\right)$
- **2 pt : 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 :**

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

- **1 pt :** L **ne s'annule au plus qu'une fois**

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

- **1 pt :** $(g'(\theta))^2 \leqslant I_X(\theta)\mathbb{V}(f(X))$
- **1 pt :** $I_X(\theta) \geqslant 0$ **et par hypothèse** $I_X(\theta) \neq 0$ **donc** $I_X(\theta) > 0$

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 $T_n = f(X_1, \ldots, X_n)$ un estimateur sans biais de $g(\theta)$ à savoir tel que $\mathbb{E}\big(f(X_1, \ldots, X_n)\big) = g(\theta)$ 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 $\mathbb{V}\big(\overline{X}_n\big)$.

- **1 pt : $\overline{X}_n$ admet une espérance et une variance en tant que combinaison linéaire de variables aléatoires admettant une espérance et une variance**
- **1 pt : $\mathbb{E}\big(\overline{X}_n\big) = \theta$**
- **1 pt : $\mathbb{V}\big(\overline{X}_n\big) = \dfrac{\theta}{n}$**
- **1 pt : indépendance**

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.

- **1 pt : $\mathbb{E}\big(\overline{X}_n\big) = \theta$ donc $\overline{X}_n$ fait bien partie des estimateurs sans biais de θ**
- **1 pt : choix de $g = \mathrm{id}_{\mathbb{R}}$ pour obtenir $\mathbb{V}(T_n) \geqslant \dfrac{1}{n\, I_{X_1}(\theta)}$**
- **1 pt : $\dfrac{1}{n\, I_{X_1}(\theta)} = \dfrac{\theta}{n} = \mathbb{V}\big(\overline{X}_n\big)$**
- **1 pt : pour les estimateurs sans biais, la variance coïncide avec le risque quadratique**

32. Montrer que pour $g(\theta) = \exp(-\theta)$ où $\theta \in\ ]0, +\infty[$:

$$\mathbb{V}\big(\varphi(S_n)\big) \underset{n\to+\infty}{\sim} \frac{\big(g'(\theta)\big)^2}{n\, I_{X_1}(\theta)}$$

- **2 pt : $\mathbb{V}\big(\varphi(S_n)\big) \underset{n\to+\infty}{\sim} \exp(-2\theta)\dfrac{\theta}{n}$**
- **2 pt : $\dfrac{\big(g'(\theta)\big)^2}{n\, I_{X_1}(\theta)} = \exp(-2\theta)\dfrac{\theta}{n}$**

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

- **2 pt : $\varphi(S_n)$ a asymptotiquement le plus petit risque quadratique parmi tous les estimateurs sans biais de $\exp(-\theta)$**

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 ?

- **2 pt : 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**