جامعة هواري بومدين للعلوم والتكنولوجيا
U S T H B

Concours d'accès à l'Ecole Doctorale de Recherche Opérationnelle

Pôles USTHB et USDB

Epreuve de : **Mathématiques non déterministes (Probabilités et Statistique)**		
Année 2008 – 2009	Date: 12 – 10 – 2008	Durée: 02 heures

PROBLEME

Soient $X_1, X_2, ..., X_n$ n variables aléatoires (v.a.) indépendantes et identiquement distribuées ($i.i.d.$) de même loi qu'une v.a. X dont la densité de probabilité, $f_X(.,\theta)$, est donnée par

$$f_X(x;\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right), \quad x \in \mathbb{R},$$

(i.e. X suit une loi normale $N(\mu,\sigma^2)$ de paramètres $\mu \in \mathbb{R}$ et $\sigma^2 > 0$). On pose $S_1 = \sum_{i=1}^{n} iX_i$, $S_2 = \sum_{i=1}^{n}(n-i+1)X_i$ et $S_3 = \frac{1}{n+1}(S_1+S_2)$.

Partie I (5 points)

1– (**1.5 points**). En rappelant que la fonction génératrice des moments d'une loi normale $X \sim N\left(\mu,\sigma^2\right)$ est donnée par $M_X(t) = \exp\left(\mu t + \frac{\sigma^2}{2}t^2\right)$, déterminer la fonction génératrice des moments puis en déduire la loi de probabilité de chacune des v.a. S_1, S_2 et S_3 (On pourra utiliser les identités $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ et $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$).

2– (**1.5 points**). Montrer que la loi de probabilité de S_3 peut s'écrire sous la forme

$$f_{S_3}(s;\mu,\sigma^2) = a(s)b(\mu,\sigma^2)\exp\left(c_1(s)d_1(\mu,\sigma^2) + c_2(s)d_2(\mu,\sigma^2)\right),$$

où $a(.)$, $b(.)$, $c_i(.)$ et $d_i(.)$ sont des fonctions réelles. En déduire sans faire de calculs qu'il en est de même pour S_1 et S_2. Que peut-on conclure ?

3– (**1 point**). Soit $T_n = T(X_1, X_2, ..., X_n)$ une v.a. fonction de $X_1, X_2, ..., X_n$, telle que $E(T_n) = \mu$ et $var(T_n) \to 0$ quand $n \to \infty$. Montrer en utilisant l'inégalité de Tchebychev que $(T_n)_{n\in\mathbb{N}}$ converge en probabilité vers μ.

4– (**1 point**). Trouver une v.a. $U = U(S_3, n, \mu, \sigma^2)$, fonction de S_3, n, μ et σ^2, telle que sa loi de probabilité est indépendante de μ et σ^2.

Partie II (10 points)

On suppose que $\sigma^2 = 1$, et on admet dans la suite que $X_1, X_2, ..., X_n$ est un échantillon aléatoire simple, issu d'une population régie par X, et à travers lequel on veut faire des inférences concernant μ.

1− (**2 points**). Rappeler rigoureusement le concept d'exhaustivité, exposer le théorème de factorisation de Neyman-Fisher et l'utiliser pour trouver une statistique exhaustive pour μ.

2− (**1 point**). Donner la définition d'une statistique complète et expliquer son rôle dans l'estimation statistique, puis montrer que les statistiques S_1, S_2 et S_3 sont complètes.

3− (**2 points**). On pose $T_1 = \frac{2}{n(n+1)} S_1$, $T_2 = \frac{2}{n(n+1)} S_2$ et $T_3 = \frac{1}{2}(T_1 + T_2)$. Montrer que les estimateurs T_1, T_2 et T_3 sont sans biais pour θ. Sont-ils convergents (en probabilité)?

4− (**2 points**). Vérifier que $\forall n > 1$, $var(T_3) < \min(var(T_1), var(T_2))$. Peut-on trouver un estimateur sans biais de variance plus petite que celle de T_3 ? En déduire que les statistiques S_1, S_2 ne peuvent être exhaustives.

5− (**2 points**). Expliquer brièvement l'intuition derrière le principe du maximum de vraisemblance. Calculer l'estimateur du maximum de vraisemblance pour μ. Est-il efficace?

6− (**1 point**). Trouver un intervalle de confiance bilatéral pour μ au seuil $1 - \alpha$ ($\alpha \in]0, 1[$).

Partie III (5 points)

On suppose maintenant que $\mu = \theta$ et $\sigma^2 = \theta^2$ ($\theta > 0$).

1− (**1 point**). Montrer que la statistique $S = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2 \right)$ est exhaustive pour θ.

2− (**1 point**). Calculer $E\left[2\left(\sum_{i=1}^{n} X_i \right)^2 - (n+1) \sum_{i=1}^{n} X_i^2 \right]$ puis en déduire que la statistique S n'est pas complète.

3− (**2 points**). Trouver par la méthode des moments deux estimateurs différents pour θ.

4− (**1 point**). Déterminer l'estimateur du maximum de vraisemblance pour θ.

جامعة هواري بومدين
للعلوم والتكنولوجيا
U S T H B

Concours d'accès à l'Ecole Doctorale de Recherche Opérationnelle

Pôles USTHB et USDB

Epreuve de : **Mathématiques non déterministes (Probabilités et Statistique)**

Année 2008 – 2009 | Date: 12 – 10 – 2008 | Durée: 02 heures

Corrigé de l'épreuve

Partie I (5 points)

1- (**1.5 points**). *Soit à déterminer la fonction génératrice des moments et la loi de probabilité de chacune des v.a.* S_1*,* S_2 *et* S_3*.*

On note $M_{S_i}(t)$ *la fonction génératrice des moments de la v.a.* S_i $(i = \overline{1,3})$. Alors en exploitant les propriétés de la fonction exponentielle, la propriété *i.i.d.* des $X_1, X_2, ..., X_n$ et l'expression de la fonction génératrice des moments d'une loi normale $N(\mu, \sigma^2)$ qui est donnée par, $M_X(t) = \exp\left(\mu t + \sigma^2 \frac{t^2}{2}\right)$, on trouve:

$$\begin{aligned}
M_{S_1}(t) &= E\left(\exp\left(t\sum_{i=1}^{n} iX_i\right)\right) = \prod_{i=1}^{n} E\left(\exp\left(tiX_i\right)\right) \\
&= \prod_{i=1}^{n} M_{X_i}(it) = \prod_{i=1}^{n} \exp\left(\mu ti + \frac{\sigma^2}{2} i^2 t^2\right) \\
&= \exp\left(\sum_{i=1}^{n} \mu ti + \sum_{i=1}^{n} \frac{\sigma^2}{2} i^2 t^2\right) = \exp\left(\frac{n(n+1)\mu}{2} t + \frac{n(n+1)(2n+1)\sigma^2}{6} \frac{t^2}{2}\right).
\end{aligned}$$

Ainsi on reconnaît la fonction génératrice des moments d'une loi normale de moyenne $\frac{n(n+1)\mu}{2}$ et de variance $\frac{n(n+1)(2n+1)\sigma^2}{6}$. D'où par unicité de la fonction génératrice $S_1 \sim N\left(\frac{n(n+1)\mu}{2}, \frac{n(n+1)(2n+1)\sigma^2}{6}\right)$.
De même,

$$\begin{aligned}
M_{S_2}(t) &= E\left(\exp\left(t\sum_{i=1}^{n} (n-i+1)X_i\right)\right) = \prod_{i=1}^{n} M_{X_i}((n-i+1)t) \\
&= \exp\left(\sum_{i=1}^{n} \mu t(n-i+1) + \frac{\sigma^2}{2}(n-i+1)^2 t^2\right) \\
&= \exp\left(\frac{n(n+1)\mu}{2} t + \frac{n(n+1)(2n+1)\sigma^2}{6} \frac{t^2}{2}\right),
\end{aligned}$$

montrant que S_2 est de même loi que S_1, i.e. $S_2 \sim N\left(\frac{n(n+1)\mu}{2}, \frac{n(n+1)(2n+1)\sigma^2}{6}\right)$.

En fin, de manière similaire, et remarquant que $S_3 = \frac{1}{n+1}(S_1 + S_2) = \sum_{i=1}^{n} X_i$, on trouve

$$M_{S_3}(t) = E\left(\exp\left(t\sum_{i=1}^{n} X_i\right)\right) = \exp\left(n\mu t + \frac{\sigma^2}{2} n t^2\right).$$

D'où $S_3 \sim N\left(n\mu, n\sigma^2\right)$.
2- (**1.5 points**). *Soit à montrer que la loi de probabilité de S_3 peut se factoriser comme produit d'une fonction exponentielle de la variable et une fonction exponentielle du paramètre, à un facteur de produit séparable près.*
S_3 suivant une loi normale, on a

$$\begin{aligned} f_{S_3}(s) &= \frac{1}{\sqrt{2\pi n}\sigma}\exp\left(-\frac{1}{2n\sigma^2}(s-n\mu)^2\right) \\ &= \frac{1}{\sqrt{2\pi n}\sigma}\exp\left(-\frac{\mu^2}{2\sigma^2}\right)\exp\left(-\frac{1}{2n\sigma^2}s^2+\frac{\mu}{\sigma^2}s\right) \\ &= a(s)b(\mu,\sigma^2)\exp\left(c_1(s)d_1(\mu,\sigma^2)+c_2(s)d_2(\mu,\sigma^2)\right), \end{aligned}$$

avec $a(s)=1$, $b(\mu,\sigma^2)=\frac{1}{\sqrt{2\pi n}\sigma}\exp\left(-\frac{\mu^2}{2\sigma^2}\right)$, $c_1(s)=s^2$, $d_1(\mu,\sigma^2)=-\frac{1}{2n\sigma^2}$, $c_2(s)=s$ et $d_2(\mu,\sigma^2)=\frac{\mu}{\sigma^2}$. Il en est de même pour S_1 et S_2 puisqu'elles suivent également une loi normale. Ainsi, on conclue que les lois de S_1, S_2 et S_3 appartiennent à la famille de lois exponentielles.
3- (**1 point**). *Soit à montrer que sous les hypothèses mentionnées $(T_n)_{n\in\mathbb{N}}$ converge en probabilité vers μ.*
Puisque la v.a. T_n est de variance finie, elle vérifie donc l'inégalité de Tchebychev qui s'énonce comme suit

$$\forall\varepsilon>0,\qquad P\left(|T_n-E\left(T_n\right)|\geq\varepsilon\right)\leq\frac{var\left(T_n\right)}{\varepsilon^2}.$$

Donc par remplacement de la valeur de $E(T_n)$ dans cette dernière inégalité et tout en exploitant la positivité de la probabilité on trouve pour tout $n\in\mathbb{N}$,

$$\forall\varepsilon>0,\qquad 0\leq P\left(|T_n-\mu|\geq\varepsilon\right)\leq\frac{var\left(T_n\right)}{\varepsilon^2},$$

En passant à la limite lorsque $n\to\infty$, alors

$$\forall\varepsilon>0,\qquad 0\leq\lim_{n\to\infty}P\left(|T_n-\mu|\geq\varepsilon\right)\leq\lim_{n\to\infty}\frac{var\left(T_n\right)}{\varepsilon^2}=0.$$

D'où

$$\forall\varepsilon>0,\qquad \lim_{n\to\infty}P\left(|T_n-\mu|\geq\varepsilon\right)=0,$$

signifiant que $T_n\xrightarrow{p}\mu$ quand $n\to\infty$.
4- (**1 point**). *Soit à trouver une v.a. $U=U(S_3,n,\mu,\sigma^2)$, fonction de S_3, n, μ et σ^2, telle que sa loi de probabilité soit indépendante de μ* et σ^2.
D'après I-1), on a $S_3\sim N\left(n\mu,n\sigma^2\right)$. Donc,

$$\frac{S_3-n\mu}{\sqrt{n}\sigma}\sim N\left(0,1\right)$$

Ainsi $U=\frac{S_3-n\mu}{\sqrt{n}\sigma}$.
Partie II (10 points)
On suppose que $\sigma^2=1$, et on admet dans la suite que $X_1,X_2,...,X_n$ est un échantillon aléatoire simple issue d'une population régie par X.
1- (**2 points**). *Soit à rappeler le concept d'exhaustivité, à exposer le théorème de factorisation de Neyman-Fisher et à l'utiliser pour trouver une statistique exhaustive pour μ.*
Soit $\underline{X}=(X_1,X_2,...,X_n)$ un échantillon aléatoire simple de même loi que X, où X est une v.a. de distribution de probabilité connue, fonction d'un paramètre inconnu θ et soit $S(\underline{X})$ une fonction de

l'échantillon dont l'expression est indépendante de θ. Alors formellement, la statistique $S(\underline{X})$ est dite exhaustive pour le paramètre θ si la distribution de probabilité conjointe de l'échantillon $\underline{X}$ conditionnée par cette statistique est indépendante de θ. Intuitivement parlant, une statistique exhaustive pour un paramètre est un résumé de l'échantillon apportant toute l'information concernant le paramètre et contenue dans l'échantillon.
Le théorème de factorisation de Neyman-Fisher donne un moyen simple pour rechercher une statistique S ou pour montrer qu'une statistique donnée soit exhaustive. En effet, une condition nécessaire et suffisante pour qu'une statistique S soit exhaustive pour θ est que la loi de l'échantillon, $f_{\underline{X}}(\underline{x},\theta) = \prod_{i=1}^{n} f_{X_i}(x_i,\theta)$, peut s'écrire comme suit

$$\prod_{i=1}^{n} f_{X_i}(x_i,\theta) = h\left(\underline{x}\right) g\left(\theta, S\left(\underline{x}\right)\right),$$

où $h\left(\underline{x}\right)$ est une fonction ne dépendant pas de θ et où $g\left(.,.\right)$ est une fonction qui ne dépend de l'échantillon qu'à travers S.
Pour $\sigma^2 = 1$, la loi de l'échantillon s'écrit

$$\begin{aligned}
\prod_{i=1}^{n} f_{X_i}(x_i,\theta) &= \left(\sqrt{2\pi}\right)^{-n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right) \\
&= \left(\sqrt{2\pi}\right)^{-n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}{x_i}^2\right) \exp\left(-\frac{1}{2}\mu^2 + \mu\sum_{i=1}^{n}x_i\right) \\
&= h\left(\underline{x}\right) g\left(\mu, S_3\left(\underline{x}\right)\right),
\end{aligned}$$

Par conséquent, d'après le théorème le théorème de factorisation de Neyman-Fisher, S_3 est exhaustive pour μ.
2- (**1 point**). *Soit à donner la définition d'une statistique complète et expliquer son rôle dans l'estimation statistique, puis à montrer que les statistiques S_1, S_2 et S_3 sont complètes.*
- Une statistique S est dite complète si sa famille de lois $f_S\left(.;\theta\right)$ est complète. Une famille de lois $f_S\left(.;\theta\right)$ est dite complète si pour toute fonction $h\left(S\right)$ mesurable, telle que $E\left(h\left(S\right)\right) = 0\ \forall\theta \in \Theta$, on a $h \equiv 0$ presque sûrement.
- La propriété de complétion est très importante pour l'inférence statistique, notamment pour la recherche de l'estimateur sans biais de variance minimum.
- Comme les lois de S_1, S_2, et S_3 appartiennent à la famille de lois exponentielles (voir question I-2)) il s'ensuit, par le théorème qui stipule que toute statistique dont la famille de lois appartient à la famille de lois exponentielle est complète, que les statistiques S_1, S_2 et S_3 sont complètes.
3- (**2 points**). *Soit à montrer que les estimateurs $T_1 = \frac{2}{n(n+1)}S_1$, $T_2 = \frac{2}{n(n+1)}S_2$ et $T_3 = \frac{1}{2}\left(T_1 + T_2\right)$ sont sans biais (ESB) pour θ.*
- Par un calcul direct on trouve que $E\left(T_1\right) = E\left(T_2\right) = \frac{2}{n(n+1)}\sum_{i=1}^{n} i\mu = \mu$. D'autre part, $T_3 = \frac{1}{2}\left(T_1 + T_2\right)$ n'est rien d'autre que $\frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ qui est un ESB pour μ.
- Puisque T_1, T_2 et T_3 sont ESB pour θ, pour montrer qu'ils sont convergents, il suffit de montrer que leur variances convergent vers 0, lorsque $n \to \infty$ (cf, question I-3). Un calcul simple montre que $var\left(T_1\right) = var\left(T_2\right) = \frac{4}{n^2(n+1)^2}\sum_{i=1}^{n} i^2 = \frac{4}{n^2(n+1)^2}\frac{n(n+1)(2n+1)}{6} \to 0$. De même, $var\left(T_3\right) = \frac{1}{n^2} \overset{n\to\infty}{\to} 0$.
D'où les estimateurs donnés sont convergents.
4- (**2 points**). *Soit à vérifier que $\forall n > 1$ $var(T_3) < \min(var\left(T_1\right), var\left(T_2\right))$.*

- Comme on a vu plus haut (cf, question II- 3)) $var\,(T_1) = var\,(T_2) = \frac{4}{n^2(n+1)^2}\frac{n(n+1)(2n+1)}{6} = \frac{2}{3}\frac{(2n+1)}{n(n+1)}$ et $var\,(T_3) = \frac{1}{n^2}$ On vérifie aisément que dès que $n > 1$ on a $\frac{2}{3}\frac{(2n+1)}{(n+1)} > \frac{1}{n}$, c'est à dire que $var\,(T_3) < var\,(T_1) = var\,(T_2)$ pour tout $n > 1$.
- Puisque $T_3 = \overline{X} = \frac{1}{n}S_3$ est sans biais pour μ et fonction d'une statistique, S_3, exhaustive (par II-1) et complète (par II-2), d'après le théorème de Lehman-Sheffe, il est donc l'unique ESBVM, et on ne peut trouver d'ESB de variance plus petite.
- Si les statistiques S_2 et S_3 étaient exhaustives, alors puisqu'on a montré qu'elles sont complètes (cf, II-2), les ESB T_2 et T_3 fonctions de S_2 et S_3, respectivement, seraient par le théorème de Lehman-Sheffé des estimateurs sont biais de variance minimum. Comme on vient de voir qu'ils ne le sont pas ($var\,(T_3) < var\,(T_1) = var\,(T_2)$), les statistiques S_2 et S_3 ne peuvent donc être exhaustives.
5- (**2 points**). *Soit à expliquer l'intuition derrière le principe du maximum de vraisemblance et à calculer l'estimateur du maximum de vraisemblance (EMV) pour μ.*
- Dans un problème d'estimation paramétrique, le principe du maximum de vraisemblance consiste à trouver par rapport à quelle valeur du paramètre θ l'observation $X_1, ..., X_n$, dont on dispose, est la plus probable. Autrement dit, quelle est la valeur du paramètre θ la plus plausible, la plus vraisemblable à travers laquelle l'observation dont on dispose a été générée.
- Puisque la fonction de vraisemblance $L(\mu; X_1, ..., X_n) = f_{\underline{X}}\,(\underline{x}; \mu)$ est différentiable par rapport à μ sur $\mathbb{R}$, l'EMV est solution de l'équation normale

$$\frac{\partial \log L(\theta; X_1, ..., X_n)}{\partial \mu} = 0,$$

soit,

$$\sum_{i=1}^{n}(X_i - \mu) = 0,$$

qu'on résout par rapport à μ dans l'ensemble des statistiques. Donc l'EMV de μ est égal $\overline{X} = T_3$.
6- (**1 point**). *Soit à trouver un intervalle de confiance bilatéral pour μ au seuil $1-\alpha$ ($\alpha \in]0, 1[$).*
Trouvons d'abord une fonction pivotale pour μ. Puisque d'après I)-4 la loi de $U = \frac{S_3 - n\mu}{\sqrt{n}}$ est indépendante de μ, la fonction U étant monotone en μ et dont la loi ne dépend pas de μ définit une fonction pivotale. Ainsi il suffit de trouver un réel strictement positifs x tel que

$$P(-x \leq \frac{S_3 - n\mu}{\sqrt{n}} \leq x) = 1 - \alpha,$$

soit,

$$P\left(T_3 - \frac{1}{\sqrt{n}}x \leq \mu \leq T_3 + \frac{1}{\sqrt{n}}x\right) = 1 - \alpha.$$

D'où $\left[\overline{X} - \frac{1}{\sqrt{n}}x, \overline{X} + \frac{1}{\sqrt{n}}x\right]$ est un intervalle de confiance bilatéral pour μ au seuil $1-\alpha$.

Partie III (5 points)

On suppose que $\mu = \theta$ et $\sigma^2 = \theta^2$ ($\theta > 0$).

1- (**1 point**). *Soit à montrer que la statistique $S = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$ est exhaustive pour θ.*

$$
\begin{aligned}
f_{\underline{X}}\left(\underline{x};\mu,\sigma^2\right) &= \prod_{i=1}^{n} f_{X_i}\left(x_i;\mu,\sigma^2\right)\\
&= \left(2\pi\theta^2\right)^{-n/2}\exp\left(-\frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i-\theta)^2\right)\\
&= \left(2\pi\theta^2\right)^{-n/2}\exp\left(-n/2\right)\exp\left(-\frac{1}{2\theta^2}\sum_{i=1}^{n}{x_i}^2\right)\exp\left(\frac{1}{\theta}\sum_{i=1}^{n}x_i\right)\\
&= h(\underline{x})g(S(\underline{x}),\theta)
\end{aligned}
$$

où $S(\underline{x}) = \left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2\right)$. D'après le théorème de factorisation de Neyman-Fisher, la statistique $S\left(\underline{X}\right) = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$ est exhaustive pour θ.

2- (**1 point**). Soit à calculer $E\left[2\left(\sum_{i=1}^{n}X_i\right)^2 - (n+1)\sum_{i=1}^{n}X_i^2\right]$ puis à en déduire que la statistique S n'est pas complète.

i) $E\left[2\left(\sum_{i=1}^{n}X_i\right)^2 - (n+1)\sum_{i=1}^{n}X_i^2\right] =$

$$
\begin{aligned}
& 2E\left(\left(\sum_{i=1}^{n}X_i\right)^2\right) - (n+1)\sum_{i=1}^{n}E\left(X_i^2\right)\\
&= 2var\left(\sum_{i=1}^{n}X_i\right) + 2\left(E\left(\sum_{i=1}^{n}X_i\right)\right)^2 - (n+1)\sum_{i=1}^{n}\left(var\left(X_i\right)+\left(E\left(X_i\right)\right)^2\right)\\
&= 2\sum_{i=1}^{n}var\left(X_i\right) + 2\left(\sum_{i=1}^{n}E\left(X_i\right)\right)^2 - (n+1)\sum_{i=1}^{n}\left(\theta^2+\theta^2\right)\\
&= 2\left(n\theta^2+n^2\theta^2\right) - 2n(n+1)\theta^2\\
&= 0,\ \text{pour tout } \theta > 0.
\end{aligned}
$$

ii) Prenons la fonction à deux variables $h\left(S_1,S_2\right) = 2\left(S_1\right)^2 - (n+1)\,S_2$. Alors d'après i) $E\left(h\left(S\left(\underline{X}\right)\right)\right) = 0\ \forall\theta > 0$ sans que $h\left(.\right)$ ne soit une fonction identiquement nulle presque surement. D'où la statistique S n'est pas complète.

3- (**2 points**). *Soit à trouver par la méthode des moments deux estimateurs différents pour θ.*

i) Un estimateur des moments $\widehat{\theta}_M$ de θ, basé sur $X_1, X_2, ..., X_n$, est solution de l'équation

$$E\left(X\right) = \overline{X},$$

qu'on résout par rapport à θ dans l'ensemble des statistiques fonctions du moment empirique d'ordre 1, où $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Or, puisque $X \sim \mathcal{N}\left(\theta,\theta^2\right)$ et donc $E(X) = \theta$, la dernière équation s'écrit tout simplement

$$\theta = \overline{X}.$$

D'où

$$\widehat{\theta}_M = \overline{X}.$$

ii) Un autre estimateur des moments de θ peut être obtenu, en raison de la dépendance entre la moyenne θ et la variance θ^2, comme solution de l'équation du second moment

$$E\left(X^2\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^2,$$

ou encore

$$var(X) + E\left(X\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

D'où il suffit de résoudre l'équation au second degré suivante

$$\theta^2 + \theta^2 - \frac{1}{n}\sum_{i=1}^{n} X_i^2 = 0,$$

dont la solution est

$$\widehat{\theta}_M = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}.$$

(l'autre solution négative ne sera pas retenue puisque $\theta > 0$.
4- (**1 point**). *Soit à déterminer l'estimateur du maximum de vraisemblance $\widehat{\theta}_{MV}$ pour θ.*
La fonction de vraisemblance $L\left(\theta; x_1, ..., x_n\right)$ s'écrit avec $\mu = \theta$ et $\sigma^2 = \theta^2$ comme suit

$$L\left(\theta; x_1, ..., x_n\right) = \left(2\pi\theta^2\right)^{-n/2} \exp\left(-\frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \theta)^2\right), \quad \theta > 0.$$

Cette fonction est différentiable sur $]0, +\infty[$, et donc l'EMV sera la solution de l'équation normale

$$\frac{\partial \log L(\theta; X_1, ..., X_n)}{\partial\theta} = 0,$$

soit

$$\frac{-n}{\theta} + \frac{1}{\theta^3}\sum_{i=1}^{n}(X_i - \theta)^2 + \frac{1}{\theta^2}\sum_{i=1}^{n}(X_i - \theta) = 0,$$

qui se réduit à l'équation du second degré suivante

$$-n\theta^2 - \sum_{i=1}^{n} X_i\theta + \sum_{i=1}^{n} X_i^2 = 0,$$

dont l'unique solution positive, $\widehat{\theta}_{MV}$, (l'autre elle est négative et ne sera pas retenue) est donnée par

$$\widehat{\theta}_{MV} = \frac{\sqrt{\sum\limits_{i=1}^{n} X_i + 4n\sum\limits_{i=1}^{n} X_i^2} - \sum\limits_{i=1}^{n} X_i}{2n}.$$