Statistiques appliquées. L3
Interrogation n°2

Questions de cours. 3 points

1) Enoncer le théorème central limite (1 pt).
Si (X_n) est une suite de v.a. indépendantes et de même loi, admettant des moments d'ordre un et deux notés $m = E(X)$ et $\sigma^2 = V(X_n)$, alors en notant $\bar{X}_n = \sum_{i=1}^{n} X_i$:

$\sqrt{n}\frac{\bar{X}_n - m}{\sigma} \rightarrow^{loi}_{n\rightarrow+\infty} N(0,1)$

2) Si $Cov(X,Y) = 0$, alors X et Y sont indépendantes. Vrai ou faux ? Expliquer
Faux. C'est l'implication suivante qui est vraie : si X et Y sont indépendantes, alors $Cov(X,Y) = 0$.

3) Si $X_1, ...X_n \hookrightarrow i.i.d.N(m, \sigma^2)$, quelle est la loi de $\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$?
$\bar{X}_n \hookrightarrow N(m, \frac{\sigma^2}{n})$

4) Si $X_1, ...X_n \hookrightarrow i.i.d.N(0,1)$, quelle est la loi de $X_1^2 + ... + X_n^2$?
Définition de la loi du khi-deux :
Si $X_1, ...X_n \hookrightarrow i.i.d.N(0,1)$
alors $X_1^2 + ... + X_n^2 \hookrightarrow \chi^2_n$ (loi du khi-deux à n degrés de liberté).

5) Si $U \hookrightarrow N(0,1)$ et $Y \hookrightarrow \chi^2_n$, quelle est la loi de $\frac{U}{\sqrt{\frac{Y}{n}}}$?
Définition de la loi de Student :
Si $U \hookrightarrow N(0,1)$ et $Y \hookrightarrow \chi^2_n$
alors $\frac{U}{\sqrt{\frac{Y}{n}}} \hookrightarrow T_n$ (loi de Student à n degrés de liberté).

Exercice 1. 3 points

Soient X une v.a. d'espérance m et de variance σ^2 et $(X_1, ...X_n)$ des observations indépendantes. On dispose de deux estimateurs de m : $\hat{m}_1 = \bar{X}$ et $\hat{m}_2 = \frac{1}{2}(X_3 + X_{n-3})$.

1) Montrer que ces estimateurs sont sans biais.

Définition d'un estimateur sans biais : un estimateur T_n de θ est dit sans biais si pour tout θ de Θ et tout entier positif n : $E_\theta(T_n) = \theta$

Montrons que $E(\hat{m}_1) = m$ et que $E(\hat{m}_2) = m$.
$E(\hat{m}_1) = E(\bar{X}) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$ par linéarité de l'espérance.
En outre $E(X_1) = ... = E(X_n) = m$ donc $E(\hat{m}_1) = \frac{1}{n}.n.m = m$
donc $\hat{m}_1$ est sans biais. La moyenne empirique est un estimateur sans biais de la moyenne théorique.

$E(\hat{m}_2) = E(\frac{1}{2}(X_3 + X_{n-3})) = \frac{1}{2}(E(X_3) + E(X_{n-3}))$ par linéarité de l'espérance.
En outre $E(X_3) = E(X_{n-3}) = m$ donc $E(\hat{m}_2) = \frac{1}{2}.2.m = m$
donc $\hat{m}_2$ est sans biais.

2) Montrer que $\hat{m}_1$ est convergent

Définition d'un estimateur convergent : Un estimateur T_n est convergent si la suite de v.a. (T_n) converge en probabilité vers la valeur du paramètre.

Théorème : Tout estimateur sans biais dont la variance tend vers 0 est convergent.
Ce théorème donne deux conditions suffisantes pour un estimateur sans biais. Ce ne sont pas des conditions nécessaires et suffisantes (il est faux d'écrire que si un estimateur est sans biais et que sa variance ne tend pas vers 0 alors il n'est pas convergent).

On se sert de ce théorème pour montrer que $\hat{m}_1$ est convergent.
En effet, $\hat{m}_1$ est sans biais
et $V(\hat{m}_1) = V(\bar{X}) = V(\frac{1}{n}\sum_{i=1}^{n} X_i) = \dfrac{1}{n^2}V(\sum_{i=1}^{n} X_i) = \dfrac{1}{n^2}\sum_{i=1}^{n} V(X_i)$ car les X_i sont indépendants.
$V(\hat{m}_1) = \frac{1}{n^2}.n.V(X_i) = \frac{1}{n}V(X_i) = \frac{\sigma^2}{n} \to 0$ quand $n \to +\infty$.
Donc $\hat{m}_1$ est convergent. La moyenne empirique est un estimateur convergent de la moyenne théorique (c'est-à-dire de l'espérance).

Remarques :
1) La variance empirique modifiée est un estimateur sans biais de la variance théorique.
2) Les deux estimateurs sont sans biais et $V(\hat{m}_1) \leq V(\hat{m}_2)$ dès lors que $n \geq 4$. Donc l'estimateur $\hat{m}_1$ est préférable à l'estimateur $\hat{m}_2$ au sens de la variance. On dit que $\hat{m}_1$ est plus efficace que $\hat{m}_2$.

Exercice 2. 6 points

Une laiterie produit des fromages dont la masse X en grammes est distribuée selon une loi normale $N(m, \sigma^2)$. On observe les masses de n fromages (indépendantes).

1) On suppose que la variance est connue : $\sigma^2 = 6.25$.
a) Donner un intervalle de confiance bilatéral à 0.95 pour m.
$\sigma = \sqrt{6.25} = 2.5$
$X_1, ...X_n \hookrightarrow i.i.d.N(m, 6.25)$ donc $\bar{X}_n \hookrightarrow N(0, \frac{6.25}{n})$ et $U_n = \sqrt{n}\frac{\bar{X}-m}{2.5} \hookrightarrow N(0,1)$
(on n'utilise pas le Théorème Central Limite ici et U_n suit cette loi même pour n petit).

$P(-u \leq U_n \leq u) = 0.95$. La lecture de la table de la loi $N(0,1)$ donne $u = 1.96$.
Ainsi $P(-1.96 \leq \sqrt{n}\frac{\bar{X}-m}{2.5} \leq 1.96) = 0.95$, soit $P(\bar{X}_n - 1.96 * \frac{2.5}{\sqrt{n}} \leq m \leq \bar{X}_n + 1.96 * \frac{2.5}{\sqrt{n}}) = 0.95$

b) Combien de fromages doit-on peser pour que la longueur de l'intervalle de confiance soit inférieure à 1 ?
La longueur de l'intervalle est $2 * 1.96 * \frac{2.5}{\sqrt{n}}$.
$2 * 1.96 * \frac{2.5}{\sqrt{n}} \leq 1 \Leftrightarrow \frac{9.8}{\sqrt{n}} \leq 1 \Leftrightarrow n \geq 9.8^2 \Leftrightarrow n \geq 96.04$.
On doit peser au moins 97 fromages pour que la longueur de l'intervalle de confiance soit inférieure à 1.

c) Application numérique : pour 17 fromages, on observe que la moyenne de la masse observée est $\bar{x} = 253.5$ grammes. Dans quelle fourchette se situe m ?
L'intervalle de confiance de m est $\bar{X}_n - 1.96 * \frac{2.5}{\sqrt{n}} \leq m \leq \bar{X}_n + 1.96 * \frac{2.5}{\sqrt{n}}$
donc $253.5 - 1.96 * \frac{2.5}{\sqrt{17}} \leq m \leq 253.5 + 1.96 * \frac{2.5}{\sqrt{17}}$ soit $252.31 \leq m \leq 254.69$.

2) On suppose que la variance σ^2 est inconnue.
a) Donner un intervalle de confiance à $0,95$ de m.

Soit $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{n-1}$

$Z_n = \frac{\bar{X}_n - m}{\frac{S}{\sqrt{n}}} \hookrightarrow T_{(n-1)}$

Soit z le fractile d'ordre $0,975$ de la loi de Student à $n-1$ degrés de liberté : $P(-z \leq Z_n \leq z) = 0,95$

Ainsi $P(-z \leq \frac{\bar{X}_n - m}{\frac{S}{\sqrt{n}}} \leq z) = 0,95$ soit $P(\bar{X}_n - z.\frac{S}{\sqrt{n}} \leq m \leq \bar{X}_n + z.\frac{S}{\sqrt{n}}) = 0,95$

b) Application numérique. Pour 17 fromages, la moyenne de la masse observée est $\bar{x} = 253,5$ et la variance est $\frac{1}{17}\sum_{i=1}^{17}(x_i - \bar{x})^2 = 8,01$. Calculer la variance modifiée s^2 et donner l'intervalle dans lequel se situe m.

Indication : le fractile d'ordre $0,975$ de la loi de Student à 16 degrés de liberté T_{16} vaut $2,12$.

$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{n}{n-1}(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2) = \frac{17}{16}(\frac{1}{17}\sum_{i=1}^{17}(x_i - \bar{x})^2) = \frac{17}{16}.8,01 = 8,51$

et donc $s = \sqrt{8,51} = 2,92$

On a $\bar{X}_n - z.\frac{S}{\sqrt{n}} \leq m \leq \bar{X}_n + z.\frac{S}{\sqrt{n}}$

et donc $253,5 - 2,12.\frac{2,92}{\sqrt{17}} \leq m \leq 253,5 + 2,12.\frac{2,92}{\sqrt{17}}$ soit $252,00 \leq m \leq 255,00$.

Exercice 3. 3 points. Méthode du maximum de vraisemblance

Cf Lec. ex 13 p.232

Soit X une variable aléatoire admettant pour densité

$$f(x;\theta) = \begin{cases} \frac{1}{2\theta\sqrt{x}} exp(-\frac{\sqrt{x}}{\theta}) & \textbf{si } x > 0 \\ 0 & \textbf{sinon} \end{cases}$$

où θ est un paramètre strictement positif que l'on se propose d'estimer à partir d'un échantillon $(X_1, ...X_n)$ de X. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ.

La vraisemblance est

$$\begin{aligned} l(x_1,...x_n;\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\ &= \prod_{i=1}^{n} \frac{1}{2\theta\sqrt{x_i}} exp(-\frac{\sqrt{x_i}}{\theta}) \\ &= \frac{1}{2\theta\sqrt{x_1}} exp(-\frac{\sqrt{x_1}}{\theta}).\frac{1}{2\theta\sqrt{x_2}} exp(-\frac{\sqrt{x_2}}{\theta})...\frac{1}{2\theta\sqrt{x_n}} exp(-\frac{\sqrt{x_n}}{\theta}) \\ &= (2\theta)^{-n} \prod x_i^{-\frac{1}{2}} exp(-\frac{1}{\theta}\sum_{i=1}^{n}\sqrt{x_i}) \end{aligned}$$

la log vraisemblance s'écrit

$L(x_1,...x_n;\theta) = lnl(x_1,...x_n;\theta) = -nln2 - nln\theta - \frac{1}{2}ln\prod_{i=1}^{n} x_i - \frac{1}{\theta}\sum\sqrt{x_i}$

On cherche la valeur θ qui maximise la vraisemblance. La log-vraisemblance est deux fois dérivables, on applique donc un résultat classique d'optimisation.

Les dérivées s'écrivent :

$\frac{\partial L}{\partial\theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}\sqrt{x_i}$

$\frac{\partial^2 L}{\partial\theta^2} = \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n}\sqrt{x_i} = \frac{n\theta - 2\sum_{i=1}^{n}\sqrt{x_i}}{\theta^3}$

Condition nécessaire d'ordre 1 (CN1) : $\frac{\partial L}{\partial\theta} = 0$

$-\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} \sqrt{x_i} = 0 \Leftrightarrow \theta = \frac{1}{n} \sum_{i=1}^{n} \sqrt{x_i}$.

Condition suffisante d'ordre 2 (CS2) : $\frac{\partial^2 L}{\partial \theta^2} < 0$

En le point $\theta = \frac{1}{n} \sum_{i=1}^{n} \sqrt{x_i}$ c'est-à-dire pour $\sum_{i=1}^{n} \sqrt{x_i} = n\theta$, la dérivée seconde vaut $\frac{n\theta - 2n\theta}{\theta^3} = -\frac{n}{\theta^2}$, ce qui est bien négatif.

L'estimateur du maximum de vraisemblance (emv) est donc : $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} \sqrt{X_i}$

Exercice 4. 5 points. Méthode des moments et méthode du maximum de vraisemblance

Soit X une v.a. de densité $f(x;\theta) = \begin{cases} \frac{2x}{\theta^2} & \textbf{si } 0 \le x \le \theta \\ 0 & \textbf{sinon} \end{cases}$ où θ est un paramètre strictement positif.

$(X_1, ...X_n)$ **est un échantillon de X.**

1) a) Déterminer un estimateur T_n de θ par la méthode des moments.

$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \frac{2}{\theta^2} \int_0^\theta x^2 dx = \frac{2}{\theta^2} [\frac{1}{3}x^3]_0^\theta = \frac{2}{\theta^2} \frac{\theta^3}{3} = \frac{2}{3}\theta$

L'estimateur des moments est la solution de l'équation en θ "moment empirique= moment théorique". Ici on va utiliser le moment d'ordre 1 c'est-à-dire la moyenne. L'équation s'écrit $\bar{X}_n = E(X) \Leftrightarrow \bar{X}_n = \frac{2}{3}\theta$.

Sa solution est $T_n = \frac{3}{2}\bar{X}_n$

b) Pourquoi T_n est-il convergent ? Calculer $V(T_n)$. Est-ce que T_n est un estimateur efficace ?

Définition d'un estimateur convergent : Un estimateur T_n est convergent si la suite de v.a. (T_n) converge en probabilité vers la valeur du paramètre (ici θ).

On nous demande pourquoi T_n converge en probabilité vers θ.

Théorème (loi des grands nombres) : Si (X_n) est une suite de v.a. mutuellement indépendantes qui admettent les mêmes moments d'ordres un et deux, c'est-à-dire avec pour tout entier i, $E(X_i) = m$ et $V(X_i) = \sigma^2$, alors quand $n \to \infty$: $\bar{X}_n \to_p m$

La méthode des moments se justifie par les propriétés de convergence des moments empiriques vers les moments théoriques.

On vérifie dans le cas présent qu'il y a bien convergence.

D'après la loi des grands nombres, $\bar{X}_n \to_p m = E(X) = \frac{2}{3}\theta$, donc $T_n = \frac{3}{2}\bar{X}_n \to_p \frac{3}{2}.\frac{2}{3}\theta = \theta$

Ainsi T_n est convergent.

$V(X) = E(X^2) - [E(X)]^2$

$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \frac{2}{\theta^2} \int_0^\theta x^3 dx = \frac{2}{\theta^2} [\frac{1}{4}x^4]_0^\theta = \frac{2}{\theta^2} \frac{\theta^4}{4} = \frac{1}{2}\theta^2$

donc $V(X) = \frac{1}{2}\theta^2 - [\frac{2}{3}\theta]^2 = \frac{1}{18}\theta^2$

$V(T_n) = V(\frac{3}{2}\bar{X}_n) = V(\frac{3}{2}.\frac{1}{n} \sum_{i=1}^{n} X_i) = V(\frac{3}{2n} \sum_{i=1}^{n} X_i) = \frac{9}{4n^2} \sum_{i=1}^{n} V(X_i)$ car les X_i sont mutuellement indépendantes.

$V(T_n) = \frac{9}{4n^2} \sum_{i=1}^{n} V(X)$ car $V(X_1) = V(X_2) = ... = V(X_n) = V(X)$

$V(T_n) = \frac{9}{4n^2}.n.V(X) = \frac{9}{4n}.\frac{1}{18}\theta^2 = \frac{1}{n}\theta^2$

La question de l'efficacité ne se pose pas parce que toutes les hypothèse de Cramer-Rao ne sont pas vérifiées ; en effet $X(\Omega) = [0, \theta]$ n'est pas indépendant du paramètre à estimer θ.

2) Déterminer l'estimateur du maximum de vraisemblance (noté M_n) de θ.
La vraisemblance de l'échantillon en θ s'écrit :
$l(x_1, ...x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$

$$\begin{aligned} l(x_1, ...x_n; \theta) &= (\frac{2}{\theta^2})^n \prod_{i=1}^{n} x_i \text{ si } 0 \leq minx_i \leq maxx_i \leq \theta \\ &= 0 \text{ sinon} \end{aligned}$$

ou encore
$l(x_1, ...x_n; \theta) = (\frac{2}{\theta^2})^n \prod_{i=1}^{n} x_i.\mathbf{1}_{[0 \leq minx_i \leq maxx_i \leq \theta]} = \mathbf{1}_{[0 \leq minx_i]}.(\frac{2}{\theta^2})^n \prod_{i=1}^{n} x_i.\mathbf{1}_{[\theta \geq maxx_i]}$

$l(x, \theta)$ n'est pas différentiable en θ, donc la recherche du maximum de vraisemblance ne peut pas se faire en utilisant les résultats classiques d'optimisation.

On cherche pour quelle valeur de θ, fonction des observations $(x_1, ...x_n)$, on aura l maximale. Lorsque $\theta < maxx_i$, l est nulle. Lorsque $\theta \geq maxx_i$, l est décroissante. Ainsi il apparaît clairement que l est maximale pour $\theta = maxx_i$. Finalement $M_n = maxX_i$ est l'estimateur du maximum de vraisemblance pour θ.