HEC 2010 S 10

Sujet S 10 - Exercice

1) Question de cours : Enoncé du théorème de la limite centrée.
Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$, qui suit une loi de Poisson de paramètre λ inconnu strictement positif.
Pour $n \in \mathbb{N}^*$, on dispose d'un n-échantillon i.i.d. $(X_1, X_2, \ldots, X_n)$ de la loi de X.
On pose : $Y_n = \sum_{i=1}^{n} X_i$ et $\overline{X}_n = \frac{Y_n}{n}$.

2) **a)** Rappeler sans démonstration, la loi, l'espérance et la variance de la variable aléatoire Y_n.
b) Montrer que $\overline{X}_n$ est un estimateur sans biais et convergent du paramètre λ.

3) Soit $n \in \mathbb{N}^*$. Montrer que la loi conditionnelle du vecteur $(X_1, X_2, \ldots, X_n)$ sachant l'événement $[Y_n = k]$ où $k \in \mathbb{N}^*$, ne dépend pas de λ.

4) On pose, pour tout $n \in \mathbb{N}^*$: $T_n = \sqrt{n}\, \frac{\overline{X}_n - \lambda}{\sqrt{\lambda}}$.
a) Quelle est la limite en loi de la suite $(T_n)_{n \in \mathbb{N}^*}$? On note T cette limite en loi.
b) On admet que n est suffisamment grand pour approcher la loi de T_n par la loi de T. On note Φ la fonction de répartition de la loi normale centrée réduite.
Soit α un réel donné vérifiant $0 < \alpha < 1$, et t_α le réel strictement positif tel que :

$$\Phi(t_\alpha) = 1 - \frac{\alpha}{2}.$$

Déterminer un intervalle de confiance pour λ au risque α.

Sujet S 10 - Exercice sans préparation

Soit n un entier supérieur ou égal à deux. Soit $A \in \mathscr{M}_n(\mathbb{R})$.
Montrer que le rang de A est égal au rang de ${}^tA\,A$.

HEC 2010 S 10 Correction de l'exercice

Q1 $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires sur $(\Omega, \mathcal{A}, P)$ indépendantes, ayant même loi, ayant une espérance commune m et une variance commune σ^2 ($\sigma>0$).
On pose $S_n = X_1 + X_2 + \cdots + X_n$ pour tout n dans $\mathbb{N}^*$.
La suite de terme général $S_n^* = \dfrac{S_n - E(S_n)}{\sqrt{V(S_n)}}$ converge en loi vers une variable aléatoire qui suit la loi normale centrée réduite.
Ainsi $\forall x \in \mathbb{R}$, $\lim\limits_{n\to+\infty} P(S_n^* \leq x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$.

Notons que : $\forall n \in \mathbb{N}^*$, $S_n^* = \dfrac{S_n - n m}{\sqrt{n}\,\sigma}$.

Q2 a) Soit $n\in\mathbb{N}^*$. $X_1, X_2, \dots, X_n$ sont indépendantes et suivent la loi de Poisson de paramètre λ. Alors, d'après le cours, $Y_n = \sum_{i=1}^{n} X_i$ suit la loi de Poisson de paramètre $n\lambda$.

Ainsi $E(Y_n) = n\lambda$ et $V(Y_n) = n\lambda$.

b) $\forall n\in\mathbb{N}^*$ $\overline{X}_n = \frac{1}{n} Y_n$ donc $E(\overline{X}_n)$ existe et vaut $\frac{1}{n}(n\lambda)$ donc λ.

$\overline{X}_n$ est donc un estimateur sans biais de λ.

Soit $n\in\mathbb{N}^*$, $V(Y_n)$ existe et vaut $n\lambda$ donc $V(\overline{X}_n)$ existe et vaut $\frac{1}{n^2} n\lambda$ donc $\frac{\lambda}{n}$.

Soit $\varepsilon \in \mathbb{R}_+^*$. L'inégalité de Bienaymé-Tchebychev donne :

$\forall n\in\mathbb{N}^*$, $0 \leq P(|\overline{X}_n - \lambda| \geq \varepsilon) = P(|\overline{X}_n - E(\overline{X}_n)| \geq \varepsilon) \leq \dfrac{V(\overline{X}_n)}{\varepsilon^2} = \dfrac{\lambda}{\varepsilon^2} \times \dfrac{1}{n}$ et $\lim\limits_{n\to+\infty} \left(\dfrac{\lambda}{\varepsilon^2} \times \dfrac{1}{n}\right) = 0$.

Par encadrement on obtient : $\lim\limits_{n\to+\infty} P(|\overline{X}_n - \lambda| \geq \varepsilon) = 0$ et ceci pour tout ε dans $\mathbb{R}_+^*$.

Alors $(\overline{X}_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable certaine égale à λ.

$\overline{X}_n$ est donc un estimateur sans biais et convergent de λ.

__Remarque__.. On aurait pu directement utiliser la loi faible des grands nombres pour montrer la convergence en probabilité de $(\overline{X}_n)_{n\in\mathbb{N}^*}$ vers la variable certaine égale à λ

Q3 Soit $n\in\mathbb{N}^*$. Soit $k\in\mathbb{N}^*$. $P(Y_n=k)=\dfrac{(n\lambda)^k}{k!}e^{-n\lambda}\neq 0$. $X_1(\Omega)\times X_2(\Omega)\times\cdots\times X_n(\Omega)=\mathbb{N}^n$.

Soit $(x_1,x_2,\dots,x_n)\in\mathbb{N}^n$.

$$P_{\{Y_n=k\}}(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=\frac{1}{P(Y_n=k)}P(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\}\cap\{Y_n=k\}).$$

$$P_{\{Y_n=k\}}(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=\frac{1}{P(Y_n=k)}P(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\}\cap\{X_1+X_2+\dots+X_n=k\})$$

__1^e^ cas__.. $x_1+x_2+\dots+x_n\neq k$

Alors $\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\}\cap\{X_1+X_2+\dots+X_n=k\}=\emptyset$.

donc $P_{\{Y_n=k\}}(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=0$.

__2^ème^ cas__.. $x_1+x_2+\dots+x_n=k$

Alors $\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\}\cap\{X_1+X_2+\dots+X_n=k\}=\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\}$.

De plus, par indépendance : $P(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=P(X_1=x_1)P(X_2=x_2)\dots P(X_n=x_n)$.

Donc $P(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=\dfrac{\lambda^{x_1}}{x_1!}e^{-\lambda}\times\dfrac{\lambda^{x_2}}{x_2!}e^{-\lambda}\times\dots\times\dfrac{\lambda^{x_n}}{x_n!}e^{-\lambda}=\dfrac{\lambda^{x_1+x_2+\dots+x_n}}{x_1!\,x_2!\dots x_n!}e^{-n\lambda}$.

$$P(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=\frac{\lambda^k}{x_1!\,x_2!\dots x_n!}e^{-n\lambda}.$$

Donc $P(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\}\cap\{X_1+X_2+\dots+X_n=k\})=\dfrac{\lambda^k}{x_1!\,x_2!\dots x_n!}e^{-n\lambda}$

Alors $P_{\{Y=k\}}(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=\dfrac{1}{P(Y=k)}\times\dfrac{\lambda^k}{x_1!\,x_2!\dots x_n!}e^{-n\lambda}$. Or $Y_n\hookrightarrow\mathcal{P}(n\lambda)$.

donc $P_{\{Y=k\}}(\{X_1=x_1\}\cap\{X_2=x_2\}\cap\dots\cap\{X_n=x_n\})=\dfrac{1}{\frac{(n\lambda)^k}{k!}e^{-n\lambda}}\,\dfrac{\lambda^k}{x_1!\,x_2!\dots x_n!}e^{-n\lambda}=\dfrac{k!}{n^k(x_1!\,x_2!\dots x_n!)}$.

Finalement si $(x_1, x_2, \dots, x_n)$ est un élément quelconque de $X_1(\Omega) \times X_2(\Omega) \times \dots \times X_n(\Omega)$:

$$P_{\{Y_n = k\}}(\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \dots \cap \{X_n = x_n\}) = \begin{cases} \dfrac{k!}{n^k (x_1!\, x_2! \dots x_n!)} & \text{si } x_1 + x_2 + \dots + x_n = k \\ 0 & \text{sinon} \end{cases}$$

donc la loi conditionnelle de $(X_1, X_2, \dots, X_n)$ sachant l'événement $\{Y_n = k\}$ ne dépend pas de λ, et ceci pour tout $k \in \mathbb{N}^*$... et même pour tout $k \in \mathbb{N}$.

Q4 a) Soit $n \in \mathbb{N}^*$. $T_n = \sqrt{n}\, \dfrac{\bar{X}_n - \lambda}{\sqrt{\lambda}} = \dfrac{n \bar{X}_n - n\lambda}{\sqrt{n}\sqrt{\lambda}} = \dfrac{Y_n - E(Y_n)}{\sqrt{n\lambda}} = \dfrac{Y_n - E(Y_n)}{\sqrt{V(Y_n)}}$ et

$Y_n = X_1 + X_2 + \dots + X_n$.

$(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires sur $(\Omega, \mathcal{B}, P)$ indépendantes, de même loi (la loi de Poisson de paramètre λ) ayant une espérance commune λ et une variance commune $(\sqrt{\lambda})^2$ ($\sqrt{\lambda} > 0$). Le théorème de la limite centrée montre alors que la suite de terme général $\dfrac{Y_n - E(Y_n)}{\sqrt{V(Y_n)}}$ converge en loi vers une variable aléatoire qui suit la loi normale centrée réduite.

$(T_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire qui suit la loi normale centrée réduite.

b) Nous supposerons ici n suffisamment grand pour confondre les fonctions de répartition de T_n et T.

$\Phi(t_\alpha) = 1 - \dfrac{\alpha}{2}$. $\alpha = 2(1 - \Phi(t_\alpha)) = 1 - \Phi(t_\alpha) + 1 - \Phi(t_\alpha) = \Phi(-t_\alpha) + 1 - \Phi(t_\alpha)$.

Alors $\alpha = P(T_n \leq -t_\alpha) + P(T_n \geq t_\alpha) = P(|T_n| \geq t_\alpha)$.

$$\alpha = P\left(\left|\sqrt{n}\, \frac{\bar{X}_n - \lambda}{\sqrt{\lambda}}\right| \geq t_\alpha\right) \underset{\substack{\uparrow \\ t_\alpha \geq 0}}{=} P\left(n \frac{(\bar{X}_n - \lambda)^2}{\lambda} \geq t_\alpha^2\right) = P\left(\bar{X}_n^2 - 2\lambda \bar{X}_n + \lambda^2 \geq \frac{\lambda}{n} t_\alpha^2\right).$$

$$\alpha = P\left(\lambda^2 - \left(2\overline{X}_n + \frac{t_\alpha^2}{n}\right)\lambda + \overline{X}_n^2 \geqslant 0\right) = P\left(\lambda^2 - 2\left(\overline{X}_n + \frac{t_\alpha^2}{2n}\right)\lambda + \overline{X}_n^2 \geqslant 0\right).$$

$$\alpha = P\left(\left(\lambda - \overline{X}_n - \frac{t_\alpha^2}{2n}\right)^2 - \left(\overline{X}_n + \frac{t_\alpha^2}{2n}\right)^2 + \overline{X}_n^2 \geqslant 0\right).$$

$$\alpha = P\left(\left(\lambda - \overline{X}_n - \frac{t_\alpha^2}{2n}\right)^2 - \left(\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}\right) \geqslant 0\right).$$

$$1-\alpha = P\left(\left(\lambda - \overline{X}_n - \frac{t_\alpha^2}{2n}\right)^2 - \left(\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}\right) < 0\right)$$

$$1-\alpha = P\left(-\sqrt{\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}} < \lambda - \overline{X}_n - \frac{t_\alpha^2}{2n} < \sqrt{\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}}\right)$$

$$1-\alpha = P\left(\underbrace{\overline{X}_n + \frac{t_\alpha^2}{2n} - \sqrt{\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}}}_{U_n} < \lambda < \underbrace{\overline{X}_n + \frac{t_\alpha^2}{2n} + \sqrt{\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}}}_{V_n}\right)$$

Donc $1-\alpha = P(\lambda \in]U_n, V_n[)$. Or $\{\lambda \in]U_n, V_n[\} \subset \{\lambda \in [U_n, V_n]\}$.

Donc $P(\lambda \in [U_n, V_n]) \geqslant P(\lambda \in]U_n, V_n[) = 1-\alpha$. $P(\lambda \in [U_n, V_n]) \geqslant 1-\alpha$.

Ainsi $[U_n, V_n]$ est un intervalle de confiance de λ à la confiance $1-\alpha$ ou au risque α. Pour finir donnons une "autre forme" à U_n et V_n.

$$V_n = \overline{X}_n + \frac{t_\alpha^2}{2n} + \sqrt{\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}} = \overline{X}_n + \frac{t_\alpha^2}{2n} + \frac{t_\alpha}{\sqrt{n}}\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}}.$$

$$V_n = \frac{t_\alpha^2}{4n} + \overline{X}_n + \frac{t_\alpha^2}{4n} + 2\,\frac{t_\alpha}{2\sqrt{n}}\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} = \left(\frac{t_\alpha}{2\sqrt{n}} + \sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}}\right)^2.$$

$$U_n = \overline{X}_n + \frac{t_\alpha^2}{2n} - \sqrt{\frac{t_\alpha^2}{n}\overline{X}_n + \frac{t_\alpha^4}{4n^2}} = \overline{X}_n + \frac{t_\alpha^2}{2n} - 2\times\frac{t_\alpha}{2\sqrt{n}}\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} = \frac{t_\alpha^2}{4n} + \overline{X}_n + \frac{t_\alpha^2}{4n} - 2\,\frac{t_\alpha}{2\sqrt{n}}\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}}.$$

$$U_n = \left(\frac{t_\alpha}{2\sqrt{n}} - \sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}}\right)^2.$$

Donc $\left[\left(\frac{t_\alpha}{2\sqrt{n}} - \sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}}\right)^2, \left(\frac{t_\alpha}{2\sqrt{n}} - \sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}}\right)^2\right]$ est un intervalle de confiance de λ au risque α.

Remarque. -- Retrouve Q4 b dans une ESPHEC 2009.

R..

Question 10 HEC 2010 $\boxed{\text{F 1}}$ W. MARTUCCI

n est un entier supérieur ou égal à deux. A est une matrice de $\mathcal{M}_n(\mathbb{R})$.

Montrer que le rang de A est égal au rang de ${}^tAA$.

Posons $E=\mathbb{R}^n$. Soit $\mathcal{B}$ la base canonique de $\mathbb{R}^n$. Soit f (resp. g) l'endomorphisme de E de matrice A (resp. ${}^tAA$) dans la base $\mathcal{B}$.

Soit u un élément de E de matrice X dans $\mathcal{B}$.

* Supposons que u est élément de $\operatorname{Ker} f$. Alors $AX=0_{\mathcal{M}_{n,1}(\mathbb{R})}$.

${}^tAAX={}^tA\,0_{\mathcal{M}_{n,1}(\mathbb{R})}=0_{\mathcal{M}_{n,1}(\mathbb{R})}$; donc $g(u)=0_E$; $u\in\operatorname{Ker} g$.

* Supposons que u est élément de $\operatorname{Ker} g$. Alors ${}^tAAX=0_{\mathcal{M}_{n,1}(\mathbb{R})}$.

Donc ${}^tX\,{}^tAAX={}^tX\,0_{\mathcal{M}_{n,1}(\mathbb{R})}=0$. ${}^t(AX)AX=0$. $\langle AX,AX\rangle=0$.

$\|AX\|^2=0$. $\|AX\|=0$. $AX=0_{\mathcal{M}_{n,1}(\mathbb{R})}$. $u\in\operatorname{Ker} f$.

Finalement $\operatorname{Ker} f=\operatorname{Ker} g$.

Alors $\operatorname{rg} A=\operatorname{rg} f=\dim E-\dim\operatorname{Ker} f=\dim E-\dim\operatorname{Ker} g=\operatorname{rg} g=\operatorname{rg}{}^tAA$.

$\operatorname{rg} A=\operatorname{rg}{}^tAA$.

Exercice.. Et si $A\in\mathcal{M}_{n,p}(\mathbb{R})$?