HEC 2012 S 28

Exercice principal S28

Soit p un paramètre réel inconnu vérifiant $0 < p < 1$. Pour $n \in \mathbb{N}^*$, soit $X_1, X_2, \ldots, X_n$, n variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes et de même loi de Bernoulli d'espérance p.

On pose pour tout $n \in \mathbb{N}^*$: $\overline{X}_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$; on note $E(X)$ l'espérance d'une variable aléatoire X et exp la fonction exponentielle.

1. Question de cours : Rappeler comment l'inégalité de Bienaymé-Tchebychev permet de définir à partir de $\overline{X}_n$, un intervalle de confiance de risque α $(0 < \alpha < 1)$ pour le paramètre p.

2. Soit f la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par : $f(t) = -p\,t + \ln(1 - p + p\,e^t)$.

a) Montrer que f est de classe C^2 sur $\mathbb{R}$ et que sa dérivée seconde vérifie : $\forall t \geqslant 0,\ f''(t) \leqslant \dfrac{1}{4}$.

b) Montrer à l'aide d'une formule de Taylor que pour tout $t \geqslant 0$, on a : $f(t) \leqslant \dfrac{t^2}{8}$.

c) En déduire que pour tout $t \geqslant 0$ et pour tout $k \in [\![1, n]\!]$, on a : $E\big(\exp(t(X_k - p))\big) \leqslant \exp\left(\dfrac{t^2}{8}\right)$.

3.a) Montrer que si S est une variable aléatoire discrète finie à valeurs positives et a un réel strictement positif, on a : $P(S \geqslant a) \leqslant \dfrac{E(S)}{a}$.

b) À l'aide des questions 2.c) et 3.a), établir pour tout couple $(t, \varepsilon) \in (\mathbb{R}^+)^2$, l'inégalité :

$$P(\overline{X}_n - p \geqslant \varepsilon) \leqslant \exp\left(-t\varepsilon + \frac{t^2}{8n}\right)$$

c) Montrer que pour tout $\varepsilon \geqslant 0$, on a : $P(|\overline{X}_n - p| \geqslant \varepsilon) \leqslant 2\,\exp(-2n\varepsilon^2)$.

d) En déduire un intervalle de confiance de risque α pour le paramètre p et comparer sa longueur, lorsque α est proche de 0, à celle de l'intervalle de confiance demandé dans la question 1.

Exercice sans préparation S28

Pour $n \in \mathbb{N}^*$, on considère l'espace euclidien $\mathbb{R}^n$ muni de son produit scalaire canonique. Soit $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases orthonormées de $\mathbb{R}^n$. On note $P = (p_{i,j})_{1 \leqslant i,j \leqslant n}$ la matrice de passage de la base $\mathcal{B}_1$ à la base $\mathcal{B}_2$.

1. Exprimer P^{-1} en fonction de P.

2. Établir l'inégalité : $\left|\displaystyle\sum_{1 \leqslant i,j \leqslant n} p_{i,j}\right| \leqslant n$.

Q1 Pour tout $k \in [\![1,n]\!]$, $E(X_k)$ existe et vaut p.

Alors $E(\sum_{k=1}^{n} X_k)$ existe et vaut $\sum_{k=1}^{n} E(X_k)$ donc np.

Alors $E(\overline{X}_n)$ existe et vaut p.

Pour tout $k \in [\![1,n]\!]$, $V(X_k)$ existe et vaut $p(1-p)$.

De plus $X_1, X_2, \dots, X_n$ sont indépendantes. Ainsi $V(\sum_{k=1}^{n} X_k)$ existe et vaut $\sum_{k=1}^{n} V(X_k)$ donc $np(1-p)$

Alors $V(\overline{X}_n)$ existe et vaut $\frac{1}{n^2} V(\sum_{k=1}^{n} X_k)$ donc $\frac{p(1-p)}{n}$.

Notons encore que $\frac{1}{4} - p(1-p) = \frac{1}{4}[1-4p+4p^2] = (2p-1)^2/4 \geq 0$.

Ainsi $0 \leq p(1-p) \leq \frac{1}{4}$. Alors $V(\overline{X}_n) \leq \frac{1}{4n}$.

L'inégalité de Bienaymé-Tchebychev peut alors être appliquée à $\overline{X}_n$.

Soit $\varepsilon \in \mathbb{R}_+^*$. $P(|\overline{X}_n - E(\overline{X}_n)| \geq \varepsilon) \leq \frac{V(\overline{X}_n)}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$.

Or $P(|\overline{X}_n - E(\overline{X}_n)| \geq \varepsilon) = P(|\overline{X}_n - p| \geq \varepsilon) = P(\{\overline{X}_n - p \geq \varepsilon\} \cup \{\overline{X}_n - p \leq -\varepsilon\})$.

$P(|\overline{X}_n - E(\overline{X}_n)| \geq \varepsilon) = P(\{p \leq \overline{X}_n - \varepsilon\} \cup \{p \geq \overline{X}_n + \varepsilon\}) = P(p \notin [\overline{X}_n - \varepsilon, \overline{X}_n + \varepsilon])$.

Donc $P(p \notin [\overline{X}_n - \varepsilon, \overline{X}_n + \varepsilon]) \leq \frac{1}{4n\varepsilon^2}$; $P(p \in [\overline{X}_n - \varepsilon, \overline{X}_n + \varepsilon]) \geq 1 - \frac{1}{4n\varepsilon^2}$ (passage au complémentaire).

Alors $[\overline{X}_n - \varepsilon, \overline{X}_n + \varepsilon]$ est un intervalle de confiance de p au risque α ou à la confiance $1-\alpha$ dès que $1 - \frac{1}{4n\varepsilon^2} \geq 1 - \alpha$ ou $\alpha \geq \frac{1}{4n\varepsilon^2}$ ou encore $\varepsilon^2 \geq \frac{1}{4n\alpha}$.

<u>Ainsi en posant $\varepsilon = \frac{1}{\sqrt{4n\alpha}}$ on peut dire que $[\overline{X}_n - \varepsilon, \overline{X}_n + \varepsilon]$ est un intervalle de confiance de p au risque α "déduit" de l'inégalité de Bienaymé-Tchebychev.</u>

<u>Remarque</u> : La longueur de cet intervalle de confiance est $\frac{2}{\sqrt{4n\alpha}}$ donc $\frac{1}{\sqrt{n\alpha}}$.

Q2 <u>a)</u> $t \mapsto 1-p+pe^t$ est de classe C^2 sur $\mathbb{R}$, $\ln$ est de classe C^2 sur $\mathbb{R}_+^*$ et $\forall t \in \mathbb{R}$, $1-p+pe^t > 0$. Par composition $t \mapsto \ln(1-p+pe^t)$ est de classe C^2 sur $\mathbb{R}$.

$p \in]0,1[$

Comme $t \mapsto -pt + \ln(1-p+pe^t)$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$, par somme $\underline{\underline{f \text{ est de classe } \mathcal{C}^2 \text{ sur } \mathbb{R}}}$.

$\forall t \in \mathbb{R},\ f'(t) = -p + \dfrac{pe^t}{1-p+pe^t}$ et $f''(t) = p\,\dfrac{e^t(1-p+pe^t) - e^t pe^t}{(1-p+pe^t)^2} = p\,\dfrac{(1-p)e^t}{(1-p+pe^t)^2}$.

Soit $t \in \mathbb{R}$.

$$\frac{1}{4} - f''(t) = \frac{1}{4} - \frac{p(1-p)e^t}{(1-p+pe^t)^2} = \frac{1}{4(1-p+pe^t)^2}\left((1-p+pe^t)^2 - 4p(1-p)e^t\right)$$

Remarque.. $(a+b)^2 - 4ab = (a-b)^2$.

En appliquant cela pour $a = 1-p$ et $b = pe^t$ il vient :

$\dfrac{1}{4} - f''(t) = \dfrac{(1-p-pe^t)^2}{4(1-p+pe^t)^2}$. $\dfrac{1}{4} - f''(t) \geqslant 0$. $f''(t) \leqslant \dfrac{1}{4}$.

Ainsi $\underline{\underline{\forall t \in \mathbb{R},\ f''(t) \leqslant \frac{1}{4}}}$. donc $\underline{\underline{\forall t \in \mathbb{R}^+,\ f''(t) \leqslant \frac{1}{4}}}$!

b) $\boxed{V1}$ Notons que $f(0) = 0$ et $f'(0) = 0$.

$\forall t \in \mathbb{R}^+,\ f''(t) \leqslant \frac{1}{4}$. Alors $\forall z \in \mathbb{R}^+,\ f'(z) - f'(0) = \displaystyle\int_0^z f''(t)\,dt \underset{z \geqslant 0}{\leqslant} \frac{1}{4}\int_0^z 1\,dt = \frac{1}{4}z$.

donc $\forall z \in \mathbb{R}^+,\ f'(z) \leqslant \frac{1}{4}z$ car $f'(0) = 0$.

$\forall t \in \mathbb{R}^+,\ f(t) = f(t) - f(0) = \displaystyle\int_0^t f'(z)\,dz \underset{t \geqslant 0}{\leqslant} \int_0^t \frac{1}{4}z\,dz = \frac{1}{4}\,\frac{t^2}{2}$. $\underline{\underline{\forall t \in \mathbb{R}^+,\ f(t) \leqslant \frac{t^2}{8}}}$.

$\boxed{V2}$ f est de classe $\mathcal{C}^2$ sur $\mathbb{R}$. Appliquons alors l'inégalité de Taylor-Lagrange à l'ordre 1 pour f.

$$\forall t \in \mathbb{R}^+,\ |f(t) - f(0) - (t-0)f'(0)| \leqslant \frac{t^2}{2} \max_{x \in [0,t]} |f''(x)|.$$

Rappelons que $f(0) = f'(0) = 0$

$\forall x \in \mathbb{R}^+,\ f''(x) = \dfrac{p(1-p)e^x}{(1-p+pe^x)^2} \geqslant 0$ et $f''(x) \leqslant \frac{1}{4}$ donc $\forall x \in \mathbb{R}^+,\ |f''(x)| \leqslant \frac{1}{4}$.

Alors $\forall t \in \mathbb{R}^+,\ |f(t)| \leqslant \dfrac{t^2}{2} \displaystyle\max_{x \in [0,t]} |f''(x)| \leqslant \frac{t^2}{2} \times \frac{1}{4} = \frac{t^2}{8}$.

Ainsi $\forall t \in \mathbb{R}^+,\ f(t) \leqslant |f(t)| \leqslant \dfrac{t^2}{8}$. $\underline{\underline{\forall t \in \mathbb{R}^+,\ f(t) \leqslant \frac{t^2}{8}}}$.

V3 Soit $t \in \mathbb{R}^{+*}$. f est de classe C^2 sur $\mathbb{R}$. Alors la formule de Taylor-Lagrange appliquée à f à l'ordre 1 montre l'existence de c_t dans $]0,t[$ tel que :

$$f(t) = f(0) + (t-0)f'(0) + \frac{(t-0)^2}{2} f''(c_t) \underbrace{=}_{f(0)=f'(0)=0} \frac{t^2}{2} f''(c_t) \underbrace{\leq}_{\frac{t^2}{2} \geq 0 \text{ et } f''(c_t) \leq \frac{1}{4}} \frac{t^2}{8}.$$

$f(t) \leq \frac{t^2}{8}$.

Ceci vaut encore pour $t=0$ car $f(0)=0$. Alors $\forall t \in \mathbb{R}^+,\ f(t) \leq \frac{t^2}{8}$.

Non il n'y aura pas de version 4 !

c) Soit $t \in \mathbb{R}^+$. Soit $k \in [\![1,n]\!]$. $X_k(\Omega) = \{0,1\}$, $P(X_k=1)=p$ et $P(X_k=0)=1-p$.

$e^{t(X_k-p)}$ possède une espérance car c'est une variable aléatoire finie.

Le théorème de transfert donne $E(e^{t(X_k-p)}) = e^{t(0-p)}P(X_k=0) + e^{t(1-p)}P(X_k=1)$.

Alors $E(e^{t(X_k-p)}) = (1-p)e^{-tp} + p e^{t(1-p)} = e^{-tp}[1-p+pe^t]$.

$$E(e^{t(X_k-p)}) = e^{-tp+\ln(1-p+pe^t)} = e^{f(t)} \leq e^{t^2/8}.$$

$\forall t \in \mathbb{R}^+,\ E(e^{t(X_k-p)}) \leq e^{\frac{t^2}{8}}$.

Q3 a) Soit une variable aléatoire finie à valeurs positives. Posons $S(\Omega) = \{x_1, x_2, \dots, x_r\}$ avec $x_1 < x_2 < \dots < x_r$.

Soit $a \in \mathbb{R}^{+*}$. Posons $I = \{i \in [\![1,r]\!] \mid x_i \geq a\}$.

1er cas.. $I = \emptyset$. Alors $\{S \geq a\} = \emptyset$ donc $P(S \geq a) = 0$.

Or $E(S) = \sum_{i=1}^{r} x_i P(S=x_i) \underbrace{\geq}_{\forall i \in [\![1,r]\!],\ x_i \geq 0 \text{ et } P(S=x_i) \geq 0} 0$ et $a > 0$. Ainsi $P(S \geq a) = 0 \leq \frac{E(S)}{a}$.

2ème cas.. $I \neq \emptyset$ $P(S \geq a) = P\left(\bigcup_{i \in I} \{S = x_i\}\right) = \sum_{i \in I} P(S=x_i)$.

$$E(S) = \sum_{i=1}^{r} x_i P(S=x_i) \underbrace{\geq}_{I \subset [\![1,r]\!] \text{ et } \forall i \in [\![1,r]\!],\ x_i P(S=x_i) \geq 0} \sum_{i \in I} x_i P(S=x_i)$$

Or $\forall i \in I$, $x_i \geq a$ et $P(S=x_i) \geq 0$

donc $\forall i \in I$, $x_i P(S=x_i) \geq a P(S=x_i)$. Donc $E(S) \geq \sum_{i \in I} x_i P(S=x_i) \geq \sum_{i \in I} a P(S=x_i)$.

$E(S) \geq \sum_{i \in I} a\, P(S=x_i) = a \sum_{i\in I} P(S=x_i) = a\,P(S \geq a)$. Comme $a>0$: $P(S\geq a) \leq \dfrac{E(S)}{a}$.

<u>Si S est une variable aléatoire discrète finie à valeurs positives et a un réel strictement positif : $P(S\geq a) \leq \dfrac{E(S)}{a}$.</u> C'est l'inégalité de Markov.

<u>b)</u> Soit $(t,\varepsilon) \in (\mathbb{R}^{+*})^2$. ⚠ <u>Dans toute la suite $n \in \mathbb{N}^*$.</u> Nous ne le redirons pas.

Soit $\omega \in \Omega$. Supposons $(\bar{X}_n - p)(\omega) \geq \varepsilon$. Alors $t(\bar{X}_n-p)(\omega) \geq t\varepsilon$ car $t \in \mathbb{R}^+$.

Donc $e^{t(\bar{X}_n-p)(\omega)} \geq e^{t\varepsilon}$ car $x \mapsto e^x$ est croissante.

Ainsi $\{\bar{X}_n - p \geq \varepsilon\} \subset \{e^{t(\bar{X}_n-p)} \geq e^{t\varepsilon}\}$.

Par croissance de la probabilité P il vient $P(\bar{X}_n - p \geq \varepsilon) \leq P(e^{t(\bar{X}_n-p)} \geq e^{t\varepsilon})$.

Or $e^{t(\bar{X}_n-p)}$ est une variable aléatoire discrète finie à valeurs positives et $e^{t\varepsilon}$ un réel strictement positif. Donc $P(e^{t(\bar{X}_n-p)} \geq e^{t\varepsilon}) \leq \dfrac{1}{e^{t\varepsilon}} E(e^{t(\bar{X}_n-p)})$.

Alors <u>$P(\bar{X}_n - p \geq \varepsilon) \leq e^{-t\varepsilon} E(e^{t(\bar{X}_n-p)})$.</u>

$t(\bar{X}_n - p) = t\left[\frac{1}{n}\sum_{k=1}^n X_k - p\right] = \frac{t}{n}\sum_{k=1}^n X_k - np = \frac{t}{n}\sum_{k=1}^n (X_k-p)$.

Alors $e^{t(\bar{X}_n-p)} = e^{\frac{t}{n}\sum_{k=1}^n (X_k-p)} = e^{\sum_{k=1}^n \frac{t}{n}(X_k-p)} = \prod_{k=1}^n e^{\frac{t}{n}(X_k-p)}$.

Or $X_1, X_2, \dots, X_n$ sont indépendantes donc $e^{\frac{t}{n}(X_1-p)}, e^{\frac{t}{n}(X_2-p)}, \dots, e^{\frac{t}{n}(X_n-p)}$ sont également indépendantes et possèdent une espérance.

Alors $E(e^{t(\bar{X}_n-p)}) = E\left(\prod_{k=1}^n e^{\frac{t}{n}(X_k-p)}\right) = \prod_{k=1}^n E(e^{\frac{t}{n}(X_k-p)})$.

Or d'après Q2c) $\forall k \in [\![1,n]\!]$, $E(e^{\frac{t}{n}(X_k-p)}) \leq e^{(t/n)^2/8} = e^{\frac{t^2}{8n^2}}$.

Mieux $\forall k \in [\![0,n]\!]$, $0 \leq E(e^{\frac{t}{n}(X_k-p)}) \leq e^{\frac{t^2}{8n^2}}$.

Alors $E(e^{t(\bar{X}_n-p)}) = \prod_{k=1}^n E(e^{\frac{t}{n}(X_k-p)}) \leq \prod_{k=1}^n e^{\frac{t^2}{8n^2}} = \left(e^{\frac{t^2}{8n^2}}\right)^n = e^{\frac{t^2}{8n}}$.

<u>$E(e^{t(\bar{X}_n-p)}) \leq e^{\frac{t^2}{8n}}$.</u>

Ainsi $P(\overline{X}_n - p \geqslant \varepsilon) \leqslant e^{-t\varepsilon} E(e^{t(\overline{X}_n - p)}) \leqslant e^{-t\varepsilon} e^{t^2/8n} = e^{-t\varepsilon + \frac{t^2}{8n}}$.

$\forall t \in \mathbb{R}^+, \forall \varepsilon \in \mathbb{R}^+, \ P(\overline{X}_n - p \geqslant \varepsilon) \leqslant e^{-t\varepsilon + \frac{t^2}{8n}}$.

Soit $\varepsilon \in \mathbb{R}_+$

c) $P(|\overline{X}_n - p| \geqslant \varepsilon) = P(\{\overline{X}_n - p \geqslant \varepsilon\} \cup \{\overline{X}_n - p \leqslant -\varepsilon\})$.

$P(|\overline{X}_n - p| \geqslant \varepsilon) = P(\overline{X}_n - p \geqslant \varepsilon) + P(\overline{X}_n - p \leqslant -\varepsilon)$.

$P(\overline{X}_n - p \leqslant -\varepsilon) = P(-\overline{X}_n + p \geqslant \varepsilon) = P(1 - \overline{X}_n - (1-p) \geqslant \varepsilon)$.

Posons $p' = 1-p$. $p' \in]0,1[$. $1 - \overline{X}_n = 1 - \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{1}{n}\sum_{k=1}^{n}(1 - X_k)$.

Posons $\forall k \in [\![1,n]\!]$, $X'_k = 1 - X_k$ et $\overline{X}'_n = \frac{1}{n}\sum_{k=1}^{n} X'_k$.

1) Comme pour tout k dans $[\![1,n]\!]$ X_k suit la loi de Bernoulli de paramètre p, pour tout k dans $[\![1,n]\!]$, X'_k suit la loi de Bernoulli de paramètre $1-p$ donc p'.

2) $X_1, X_2, \dots, X_n$ sont indépendantes donc $1-X_1, 1-X_2, \dots, 1-X_n$ sont indépendantes

Alors $X'_1, X'_2, \dots, X'_n$ sont indépendantes.

Soit $t \in \mathbb{R}^+$. Q3 b) appliquée à $X'_1, X'_2, \dots, X'_n$ donne :

$P(\overline{X}'_n - p' \geqslant \varepsilon) \leqslant e^{-t\varepsilon + \frac{t^2}{8n}}$.

Alors $P(\overline{X}_n - p \leqslant -\varepsilon) = P(1 - \overline{X}_n - (1-p) \geqslant \varepsilon) = P(1 - \overline{X}_n - p' \geqslant \varepsilon)$

Or $1 - \overline{X}_n = 1 - \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{1}{n}\sum_{k=1}^{n}(1 - X_k) = \frac{1}{n}\sum_{k=1}^{n} X'_k = \overline{X'_n}$ ... comme nous l'avions déjà vu.

Donc $P(\overline{X}_n - p \leqslant -\varepsilon) = P(\overline{X}'_n - p' \geqslant \varepsilon) \leqslant e^{-t\varepsilon + \frac{t^2}{8n}}$.

Donc $P(|\overline{X}_n - p| \geqslant \varepsilon) = P(\overline{X}_n - p \geqslant \varepsilon) + P(\overline{X}_n - p \leqslant -\varepsilon) \leqslant 2e^{-t\varepsilon + \frac{t^2}{8n}}$.

Posons $\forall t \in \mathbb{R}^+$, $\varphi(t) = -t\varepsilon + \frac{t^2}{8n}$. φ est dérivable sur $\mathbb{R}^+$ et $\forall t \in \mathbb{R}^+$, $\varphi'(t) = -\varepsilon + \frac{t}{4n}$.

φ est décroissante sur $[0, 4n\varepsilon]$ et croissante sur $[4n\varepsilon, +\infty[$.

Alors φ admet en $t_0 = 4n\varepsilon$ un minimum égal à $\varphi(4n\varepsilon)$.

Notons que $\varphi(4n\varepsilon) = -(4n\varepsilon)\varepsilon + \frac{(4n\varepsilon)^2}{8n} = -4n\varepsilon^2 + \frac{16n^2\varepsilon^2}{8n} = -4n\varepsilon^2 + 2n\varepsilon^2 = -2n\varepsilon^2$.

$\forall t \in \mathbb{R}^+$, $P(|\bar{X}_n - p| \geq \varepsilon) \leq 2e^{-t\varepsilon + \frac{t^2}{8n}} = 2e^{\varphi(t)}$

En particulier $P(|\bar{X}_n - p| \geq \varepsilon) \leq 2e^{\varphi(t_0)} = 2e^{-2n\varepsilon^2}$.

Finalement pour tout ε dans $\mathbb{R}^+$, $P(|\bar{X}_n - p| \geq \varepsilon) \leq 2e^{-2n\varepsilon^2}$.

d) Alors $1 - P(|\bar{X}_n - p| \geq \varepsilon) \geq 1 - 2e^{-2n\varepsilon^2}$; $P(|\bar{X}_n - p| < \varepsilon) \geq 1 - 2e^{-2n\varepsilon^2}$.

Or $\{|\bar{X}_n - p| < \varepsilon\} \subset \{|\bar{X}_n - p| \leq \varepsilon\}$.

Donc $P(|\bar{X}_n - p| \leq \varepsilon) \geq P(|\bar{X}_n - p| < \varepsilon) \geq 1 - 2e^{-2n\varepsilon^2}$.

$P(|X_n - p| \leq \varepsilon) = P(-\varepsilon \leq \bar{X}_n - p \leq \varepsilon) = P(\bar{X}_n - \varepsilon \leq p \leq \bar{X}_n + \varepsilon) = P(p \in [\bar{X}_n - \varepsilon, \bar{X}_n + \varepsilon])$.

Soit $\alpha \in]0,1[$. Pour avoir $P(p \in [\bar{X}_n - \varepsilon, \bar{X}_n + \varepsilon]) \geq 1-\alpha$ il suffit alors d'avoir $1 - 2e^{-2n\varepsilon^2} \geq 1 - \alpha$.

$1 - 2e^{-2n\varepsilon^2} \geq 1 - \alpha \iff \alpha \geq 2e^{-2n\varepsilon^2} \iff \ln\frac{\alpha}{2} \geq -2n\varepsilon^2 \iff \varepsilon^2 \geq \frac{-\ln\frac{\alpha}{2}}{2n}$.

Notons que $\frac{-\ln\frac{\alpha}{2}}{2n} > 0$ car $\alpha \in]0,1[$. Alors $1 - 2e^{-2n\varepsilon^2} \geq 1 - \alpha \iff \varepsilon \geq \sqrt{-\frac{1}{2n}\ln\frac{\alpha}{2}}$.

Si $\varepsilon = \sqrt{-\frac{1}{2n}\ln\frac{\alpha}{2}}$: $[\bar{X}_n - \varepsilon, \bar{X}_n + \varepsilon]$ est un intervalle de confiance de p à la confiance $1-\alpha$ ou au risque α dont la longueur est $2\sqrt{-\frac{1}{2n}\ln\frac{\alpha}{2}}$.

Notons que dans Q1, Bienaymé-Tchebychev nous a donné un intervalle de confiance de p au risque α de longueur $\frac{2}{\sqrt{4n\alpha}}$.

Or $\lim_{\alpha \to 0^+} \left(\frac{2\sqrt{-\frac{1}{2n}\ln\frac{\alpha}{2}}}{2 \times \frac{1}{\sqrt{4n\alpha}}} \right) = \lim_{\alpha \to 0^+} \sqrt{-2\alpha \ln\frac{\alpha}{2}} = 0$. La longueur de l'intervalle de confiance de Q3 est négligeable devant la longueur de l'intervalle de confiance de Q1 lorsque α tend vers 0. Lorsque α est proche de 0 le second intervalle de confiance est meilleur que le premier.

R..

Question 8 **HEC 2012-8-S28** $\boxed{\textbf{F 2}}$

Pour n dans $\mathbb{N}^*$, on considère l'espace vectoriel euclidien $\mathbb{R}^n$ muni de son produit scalaire canonique. Soit $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases orthonormées de $\mathbb{R}^n$. On note $P = (p_{i,j})$ la matrice de passage de la base $\mathcal{B}_1$ à la base $\mathcal{B}_2$.

Q1. Exprimer P^{-1} en fonction de P.

Q2. Établir l'inégalité $\left|\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}\right| \leqslant n$.

Question de cours. *Soit p un paramètre réel inconnu vérifiant $0 < p < 1$. Pour n dans $\mathbb{N}^*$, soit X_1, X_2, ..., X_n n variable aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes et de même loi de Bernoulli d'espérance p. On pose pour tout n dans $\mathbb{N}^*$: $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$.*

Rappeler comment l'inégalité de Bienaymé-Tchebychev permet de définir à partir de $\overline{X_n}$, un intervalle de confiance de risque α $(0 < \alpha < 1)$ pour le paramètre p.

Q1 P est la matrice de passage d'une base orthonormée $\mathcal{B}_1$ de $\mathbb{R}^n$ à une base orthonormée $\mathcal{B}_2$ de $\mathbb{R}^n$ donc P est une matrice orthogonale.
Alors P est inversible et $P^{-1} = {}^tP$.

Q2 Soit U l'élément de $\mathcal{M}_{n,1}(\mathbb{R})$ dont tous ~~les~~ coefficients sont égaux à 1.

Posons $V = PU = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$. $\forall i \in [\![1,n]\!]$, $v_i = \displaystyle\sum_{j=1}^{n} p_{i,j} \times 1 = \sum_{j=1}^{n} p_{i,j}$.

Notons $<\cdot,\cdot>$ le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ et $\|\cdot\|$ la norme associée.

$<U, PU> = \displaystyle\sum_{i=1}^{n} 1 \times v_i = \sum_{i=1}^{n} v_i = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}$.

$\left|\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}\right| = |<U, PU>| \leqslant \|U\|\,\|PU\|$ d'après Cauchy-Schwarz.

$\|PU\|^2 = <PU, PU> = {}^t(PU)PU = {}^tU\,{}^tPPU = {}^tU I_n U = {}^tUU = \|U\|^2$ donc $\|PU\| = \|U\|$.

Alors $\|U\|\,\|PU\| = \|U\|^2 = \displaystyle\sum_{i=1}^{n} 1^2 = n$.

Ainsi $\left|\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}\right| \leqslant n$.