Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IPÉ Math306 **Année 2005-2006**

Corrigé du devoir 2

Dans ce devoir, on utilisera à deux reprises (Ex 1, question 5 et Ex 2, question 2) le résultat suivant :

Soit $(X_n)_n$ une suite de v.a. définies sur un même espace probabilisé $(\Omega, \mathcal{F}, P)$ et à valeurs dans $\mathbb{R}$. On suppose que $(X_n)_n$ converge presque sûrement vers une constante c et que f est une fonction continue au point c. Alors la suite $(f(X_n))_n$ converge presque sûrement vers $f(c)$.

Ex 1. *Comportement asymptotique de variables de loi* $\chi^2(n)$

Dans tout l'exercice, on considère une suite $(X_i)_{i\geq 1}$ *de variables aléatoires indépendantes et de même loi* $\mathfrak{N}(0,1)$. *On pose pour tout* $n \geq 1$,

$$Z_n := \sum_{i=1}^{n} X_i^2.$$

On sait qu'alors Z_n *suit la loi* $\chi^2(n)$.

1) Dans cette question, X *est une v.a. de loi* $\mathfrak{N}(0,1)$.

1.a) *Expliquez pourquoi* X *a des moments de tout ordre.*

Il s'agit d'étudier, pour tout entier $r \geq 1$, la convergence de l'intégrale

$$\int_{-\infty}^{+\infty} |x|^r \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \,\mathrm{d}x.$$

Comme la fonction $g_r \ : \ x \mapsto |x|^r \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}$ est paire et continue sur $\mathbb{R}$, on étudie seulement la convergence de l'intégrale en $+\infty$. Or $\lim_{x\to+\infty} x^2 g_r(x)\,\mathrm{d}x = 0$ implique pour x suffisamment grand,

$$0 \leq g_r(x) \leq \frac{1}{x^2}$$

et la convergence de l'intégrale $\int_1^{+\infty} \frac{1}{x^2}\,\mathrm{d}x$ permet de conclure à celle de l'intégrale $\int_0^{+\infty} g_r(x)\,\mathrm{d}x$. Donc, $\int_{-\infty}^{+\infty} g_r(x)\,\mathrm{d}x$ est convergente et X a des moments de tout ordre.

1.b) *Que valent les* $\mathbf{E}(X^{2n+1})$ *pour* $n \in \mathbb{N}$ *?*
Pour $n \in \mathbb{N}$,

$$\mathbf{E}(X^{2n+1}) = \int_{-\infty}^{+\infty} x^{2n+1} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x,$$

l'intégrale étant bien définie (et finie) par la question précédente. Or la fonction $x \mapsto x^{2n+1} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}$ est impaire et donc son intégrale sur $\mathbb{R}$ est nulle :

$$\forall n \in \mathbb{N}, \quad \mathbf{E}(X^{2n+1}) = 0.$$

1.c) *On pose* $c_n := \mathbf{E}(X^{2n})$ *pour* $n \in \mathbb{N}$*. Montrez en intégrant par parties que*

$$\forall n \in \mathbb{N}, \quad c_n = \frac{1}{2n+1} c_{n+1}.$$

Pour calculer $c_{n+1} = \int_{-\infty}^{+\infty} x^{2n+2} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x$, on effectue une intégration par parties en posant

$$u(x) = x^{2n+1} \quad \text{et} \quad v(x) = -\mathrm{e}^{-\frac{x^2}{2}},$$

qui sont deux fonctions de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Ainsi,

$$\int_{-A}^{A} u(x) v'(x) \, \mathrm{d}x = [uv(x)]_{-A}^{A} + (2n+1) \int_{-A}^{A} x^{2n} \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x.$$

Ce qui permet de conclure que (en faisant tendre A vers $+\infty$),

$$c_{n+1} = \int_{-\infty}^{+\infty} x^{2n+2} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x = (2n+1) \int_{-\infty}^{+\infty} x^{2n} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x = (2n+1) c_n.$$

1.d) *En déduire une formule explicite pour* c_n *et calculez* $\mathbf{E}(X^4)$*.*
Par récurrence, la question précédente permet de montrer que

$$c_n = \prod_{k=0}^{n-1} (2k+1) c_0 = \prod_{k=0}^{n-1} (2k+1)$$

(car $c_0 = 1$). En particulier, $c_2 = \mathbf{E}(X^4) = 3$.

2) *En justifiant votre réponse, calculez* $\mathbf{E} Z_n$ *et* $\operatorname{Var} Z_n$*.*

Comme l'espérance est linéaire, $\mathbf{E} Z_n = \sum_{i=1}^{n} \mathbf{E}(X_i^2) = n$ (car $\mathbf{E}(X_i^2) = 1$ pour tout entier i).

Comme les variables aléatoires X_i^2 $(1 \leq i \leq n)$ sont indépendantes,

$$\operatorname{Var} Z_n = \sum_{i=1}^{n} \operatorname{Var}(X_i^2).$$

Or $\operatorname{Var}(X_i^2) = \mathbf{E}(X_i^4) - (\mathbf{E} X_i^2)^2 = 2$, et donc $\operatorname{Var} Z_n = 2n$.

3) *Quel est le comportement de la suite de v.a.* $(n^{-1}Z_n)_{n\geq 1}$ *quand* n *tend vers* $+\infty$ *?*

Comme Z_n est une somme de v.a. i.i.d d'espérance $\mathbf{E}(X_1^2) = 1$, la loi forte des grands nombres s'applique et la suite $(n^{-1}Z_n)_{n\geq 1}$ converge presque sûrement vers 1.

4) *Justifiez la convergence*

$$\frac{Z_n - n}{\sqrt{2n}} \xrightarrow[n\to+\infty]{\text{loi}} Z, \tag{1}$$

où Z *suit la loi* $\mathfrak{N}(0,1)$*.*

On peut appliquer le Théorème Limite Centrale à la suite Z_n qui est une somme de n v.a. i.i.d de carré intégrable, pour en déduire la convergence en loi de la v.a. $\dfrac{Z_n - \mathbf{E}(Z_n)}{\sqrt{\operatorname{Var} Z_n}}$ vers Z. Et on conclut en utilisant les valeurs de $\mathbf{E}Z_n$ et de $\operatorname{Var} Z_n$ trouvées à la question 2).

5) *On note* Y *une variable gaussienne de loi* $\mathfrak{N}(0, 2^{-1/2})$*. Prouvez la convergence*

$$\sqrt{Z_n} - \sqrt{n} \xrightarrow[n\to+\infty]{\text{loi}} Y. \tag{2}$$

$$\sqrt{Z_n} - \sqrt{n} = \frac{Z_n - n}{\sqrt{Z_n} + \sqrt{n}} = \frac{Z_n - n}{\sqrt{2n}} \times \frac{\sqrt{2n}}{\sqrt{Z_n} + \sqrt{n}} = U_n V_n.$$

avec $U_n = \dfrac{Z_n - n}{\sqrt{2n}}$ et $V_n = \dfrac{\sqrt{2n}}{\sqrt{Z_n} + \sqrt{n}}$.

D'aprés la question précédente, la suite $(U_n)_{n\geq 1}$ converge en loi vers Z.

D'autre part,

$$V_n = \frac{\sqrt{2}}{\sqrt{n^{-1}Z_n} + 1},$$

et la convergence presque sûre de la suite $(n^{-1}Z_n)_{n\geq 1}$ vers 1 implique celle de $(V_n)_{n\geq 1}$ vers $2^{-1/2}$ (en utilisant la continuité de l'application $x \mapsto \frac{\sqrt{2}}{\sqrt{x}+1}$ au point 1). La convergence presque sûre impliquant la convergence en probabilité, on en déduit que la suite $(V_n)_{n\geq 1}$ converge en probabilité vers la constante $2^{-1/2}$.

Les conditions d'application du Lemme de Slutsky sont satisfaites, ainsi

$$(U_n, V_n) \xrightarrow[n\to+\infty]{\text{loi}} (Z, 2^{-1/2}).$$

Donc, pour toute fonction continue de $\mathbb{R}^2$ dans $\mathbb{R}$, $g(U_n, V_n)$ converge en loi vers $g(Z, 2^{-1/2})$. En particulier, pour $g : (u,v) \mapsto uv$, on obtient la convergence en loi du produit U_nV_n vers $2^{-1/2}Z$. Le résultat est démontré en remarquant que $2^{-1/2}Z$ est une variable aléatoire gaussienne de loi $\mathfrak{N}(0, 2^{-1/2})$.

Ex 2. *Estimation du paramètre d'une loi exponentielle*

1) Soit X une variable aléatoire de loi exponentielle $\operatorname{Exp}(\theta)$*, de densité* $f_\theta : x \mapsto \theta \mathrm{e}^{-\theta x} \mathbf{1}_{\mathbb{R}_+}(x)$*, où* $\theta \in]0, +\infty[$*. Vérifiez que*

$$\mathbf{E}X = \frac{1}{\theta}, \qquad \operatorname{Var} X = \frac{1}{\theta^2}.$$

Par une intégration par partie,

$$\mathbf{E}X = \int_0^{+\infty} x\theta \mathrm{e}^{-\theta x} \,\mathrm{d}x = \frac{1}{\theta},$$

et

$$\mathbf{E}X^2 = \int_0^{+\infty} x^2 \theta \mathrm{e}^{-\theta x} \,\mathrm{d}x = 2 \int_0^{+\infty} x \mathrm{e}^{-\theta x} \,\mathrm{d}x = \frac{2}{\theta} \mathbf{E}(X) = \frac{2}{\theta^2}.$$

Et donc, $\operatorname{Var}(X) = \frac{1}{\theta^2}$.

Remarque : On aurait pu aussi calculer plus généralement $\mathbf{E}(X^n) = \dfrac{n!}{\theta^n}$.

2) *On se propose d'estimer le paramètre $\theta > 0$ d'une loi exponentielle* $\operatorname{Exp}(\theta)$*, au vu d'un échantillon $X_1, \dots, X_n$ de grande taille. Plus précisément, on considère que les X_i sont des variables aléatoires* strictement *positives, définies sur un modèle statistique* $(\Omega, \mathcal{F}, (P_\theta)_{\theta \in]0,+\infty[})$ *et sont pour tout $\theta \in]0, +\infty[$, P_θ-indépendantes et de même loi* $\operatorname{Exp}(\theta)$ *sous P_θ. On pose pour tout $n \geq 1$,*

$$S_n := \sum_{i=1}^{n} X_i \quad \textit{et} \quad T_n := \frac{n}{S_n}.$$

Montrez que T_n converge presque surement vers θ (on dit que T_n est un estimateur fortement consistant de θ).

La loi forte des grands nombres appliquée à la suite de v.a. i.i.d. $(X_i)_{i \geq 1}$ intégrables, montre que $n^{-1} S_n$ converge presque sûrement vers $\mathbf{E}X_1 = 1/\theta$. Et donc, n/S_n converge vers θ presque sûrement (car l'application $x \mapsto 1/x$ est continue au point $1/\theta$).

3) *Est-ce que $\mathbf{E}(T_n) = \theta$ (si tel est le cas, on dit que l'estimateur T_n est sans biais) ?*

Examinons le cas $n = 1$: la v.a. $T_1 = \frac{1}{X_1}$ est positive et l'intégrale

$$\int_0^{+\infty} \frac{1}{x} \theta \mathrm{e}^{-\theta x} \,\mathrm{d}x = +\infty,$$

car la fonction $x \mapsto \frac{1}{x} \theta \mathrm{e}^{-\theta x}$ n'est pas intégrable en 0. Donc, $\mathbf{E}T_1 = +\infty$ et T_n n'est pas un estimateur *sans biais*.

Remarque : Un calcul d'intégrale multiple permet d'obtenir $\mathbf{E}(T_n) = \dfrac{n}{n-1}\theta$ pour $n \geq 2$.

4) *Justifiez la convergence suivante, où Z désigne une v.a. gaussienne de loi* $\mathfrak{N}(0, 1)$ *:*

$$\frac{\theta S_n - n}{\sqrt{n}} \xrightarrow[n \to +\infty]{\text{loi}} Z. \tag{3}$$

On peut appliquer le Théorème Limite Central à la suite S_n qui est une somme de n v.a. i.i.d de carré intégrable, pour en déduire la convergence en loi de la v.a. $\dfrac{S_n - \mathbf{E}S_n}{\sqrt{\operatorname{Var} S_n}}$ vers la loi $\mathfrak{N}(0,1)$. Et on conclut, en utilisant les valeurs de $\mathbf{E}S_n = n/\theta$ et de $\operatorname{Var} S_n = n/\theta^2$.

5) *En déduire un intervalle de confiance pour θ au niveau de confiance 0,95, en négligeant l'erreur due à l'approximation gaussienne.*

D'après la convergence en loi de $\dfrac{\theta S_n - n}{\sqrt{n}}$,

$$P\left(\sqrt{n}\left|\frac{\theta}{T_n} - 1\right| \leq \epsilon\right) \geq 0,95$$

pour $\epsilon = 1,96$. Or

$$\sqrt{n}\left|\frac{\theta}{T_n} - 1\right| \leq \epsilon \quad \text{équivaut à} \quad \theta \in \left[T_n\left(1 - \frac{\epsilon}{\sqrt{n}}\right), T_n\left(1 + \frac{\epsilon}{\sqrt{n}}\right)\right].$$

On obtient ainsi un intervalle de confiance pour θ au niveau de confiance 0,95 (en négligeant l'erreur due à l'approximation gaussienne) basé sur l'estimateur T_n :

$$\left[T_n\left(1 - \frac{1,96}{\sqrt{n}}\right), T_n\left(1 + \frac{1,96}{\sqrt{n}}\right)\right].$$

Application numérique avec $n = 400$ et $S_{400}(\omega) = 1\ 460$. Donc, $T_{400}(\omega) = 0,274$ et l'intervalle de confiance pour θ au niveau $0,95$ est $[0,247; 0,301]$.