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

IS Math314 **Année 2005-2006**

Corrigé du devoir 1

Ex 1.

1) Rappelons l'inégalité de Cauchy Schwarz pour des variables aléatoires :

$$\mathbf{E}|YZ| = \mathbf{E}(|Y||Z|) \leq \big(\mathbf{E}(Y^2)\big)^{1/2}\big(\mathbf{E}(Z^2)\big)^{1/2}. \tag{1}$$

Cette inégalité reste trivialement vérifiée si $\mathbf{E}(Y^2)$ ou $\mathbf{E}(Z^2)$ vaut $+\infty$. D'autre part si $\mathbf{E}|YZ| = +\infty$ (c'est-à-dire si YZ n'a pas d'espérance), comme on a pour tout ω : $|Y(\omega)Z(\omega)| \leq Y^2(\omega) + Z^2(\omega)$, il est facile d'en déduire que $\mathbf{E}(Y^2 + Z^2) = +\infty$. Alors nécessairement l'une au moins des deux espérances de cette somme est infinie et l'inégalité (1) reste vraie. Donc tant que l'on travaille avec des valeurs absolues ou des carrés, on peut appliquer (1) sans se préoccuper de l'existence des espérances.

En appliquant (1) avec $Y = X_iX_j$ et $Z = X_kX_l$, il vient :

$$\mathbf{E}|X_iX_jX_kX_l| \leq \big(\mathbf{E}(X_iX_j)^2\big)^{1/2}\big(\mathbf{E}(X_kX_l)^2\big)^{1/2}. \tag{2}$$

En utilisant à nouveau (1) avec $Y = X_i^2$ et $Z = X_j^2$, on obtient :

$$\mathbf{E}(X_iX_j)^2 = \mathbf{E}(X_i^2X_j^2) \leq \big(\mathbf{E}(X_i^4)\big)^{1/2}\big(\mathbf{E}(X_j^4)\big)^{1/2} \tag{3}$$

et on a une inégalité analogue pour $\mathbf{E}(X_kX_l)^2$. Finalement, en combinant (2) et (3) :

$$\mathbf{E}|X_iX_jX_kX_l| \leq \big(\mathbf{E}(X_i^4)\big)^{1/4}\big(\mathbf{E}(X_j^4)\big)^{1/4}\big(\mathbf{E}(X_k^4)\big)^{1/4}\big(\mathbf{E}(X_l^4)\big)^{1/4}. \tag{4}$$

Comme les X_i ont même loi, elles ont même moment d'ordre 4, le deuxième membre de (4) est donc égal à $\mathbf{E}(X_i^4)$ qui est fini par hypothèse. Ainsi on a bien :

$$\forall (i,j,k,l) \in \{1,\dots,n\}^4, \qquad \mathbf{E}|X_iX_jX_kX_l| < +\infty. \tag{5}$$

Ceci assure l'existence pour tout quadruplet (i,j,k,l) de l'espérance de la variable aléatoire $X_iX_jX_kX_l$.

On remarque alors que :

$$\Big(\sum_{i=1}^n X_i\Big)^4 = \sum_{i,j,k,l} X_iX_jX_kX_l \leq \sum_{i,j,k,l} |X_iX_jX_kX_l|,$$

d'où en utilisant les propriétés de l'espérance :

$$\mathbf{E}\Big(\sum_{i=1}^{n} X_i\Big)^4 \leq \mathbf{E}\Big(\sum_{i,j,k,l} |X_iX_jX_kX_l|\Big) = \sum_{i,j,k,l} \mathbf{E}|X_iX_jX_kX_l|.$$

Ce majorant est la somme de n^4 termes finis d'après (5) donc est lui même fini. Finalement on peut écrire :

$$\mathbf{E}\Big(\sum_{i=1}^{n} X_i\Big)^4 = \sum_{i,j,k,l} \mathbf{E}(X_iX_jX_kX_l) \tag{6}$$

et toutes les espérances intervenant dans cette formule existent.

2) Dans le développement de $\mathbf{E}(\sum_{1\leq i\leq n} X_i)^4$, le quadruplet d'indices (i, j, k, l) décrit tout l'ensemble $\{1, \dots, n\}^4$. Pour simplifier les termes de ce développement, on utilise la propriété $\mathbf{E}(YZ) = \mathbf{E}Y\mathbf{E}Z$ valable pour des variables aléatoires *indépendantes* ayant un moment d'ordre 2.

Cas 1 : un des indice est distinct de tous les autres.
Quitte à faire une permutation d'indices, on se ramène au cas où $i \neq j$, $i \neq k$ et $i \neq l$. La v.a. X_i est indépendante de la v.a. $X_jX_kX_l$. On en déduit :

$$\mathbf{E}(X_iX_jX_kX_l) = \mathbf{E}(X_i)\mathbf{E}(X_jX_kX_l) = 0,$$

car par hypothèse les X_n sont d'espérance nulle.

Cas 2 : $\{i, j, k, l\}$ ne comporte que deux éléments, chacun étant présent deux fois.
A une permutation d'indices près, on peut supposer que $i = k$ et $j = l$. L'indépendance de X_i^2 et X_j^2 permet seulement d'écrire $\mathbf{E}(X_i^2X_j^2) = \mathbf{E}X_i^2\mathbf{E}X_j^2$, mais aucun de ces facteurs n'est nul à cause des carrés (sauf dans le cas dégénéré où les X_n sont des v.a. constantes égales à 0).

Cas 3 : $i = j = k = l$.
$\mathbf{E}(X_iX_jX_kX_l) = \mathbf{E}(X_i^4)$ qui n'est pas nul sauf dans le cas dégénéré.

3) En utilisant le fait que les X_i ont même loi, on a pour tout i et pour tout $j \neq i$, $\mathbf{E}X_i^4 = \mathbf{E}X_1^4$ et $\mathbf{E}(X_i^2X_j^2) = \mathbf{E}(X_i^2)\mathbf{E}(X_j^2) = \big(\mathbf{E}X_1^2\big)^2$. Pour calculer le deuxième membre de (6), il ne reste plus qu'à dénombrer les termes relevant des cas 2 et 3. Pour le cas 3 il y en a clairement n. Pour le cas 2, il y en a $C_n^2C_4^2$ puisqu'un tel terme est caractérisé par les deux valeurs distinctes des indices (à choisir parmi les n possibles) et par la façon d'apparier les indices égaux : $i = j$ et $k = l$ ou bien $i = k$ et $j = l$, ou bien $i = l$ et $j = k$. On a donc bien :

$$\mathbf{E}\Big(\sum_{i=1}^{n} X_i\Big)^4 = n\mathbf{E}(X_1^4) + 3n(n-1)\big(\mathbf{E}(X_1^2)\big)^2. \tag{7}$$

Notons $S_n = \sum_{1\leq i\leq n} X_i$. En utilisant l'inégalité de Markov à l'ordre 4, on obtient :

$$\forall \varepsilon > 0, \quad P\Big(\big|\frac{1}{n}S_n\big| > \varepsilon\Big) = P(|S_n| > n\varepsilon) \leq \frac{1}{n^4\varepsilon^4}\mathbf{E}S_n^4.$$

En utilisant (7) et en majorant $n(n-1)$ par n^2, on en déduit :

$$\forall \varepsilon > 0, \quad P\Big(\Big|\frac{1}{n}\sum_{i=1}^{n} X_i\Big| > \varepsilon\Big) \le \frac{\mathbf{E}(X_1^4)}{n^3\varepsilon^4} + \frac{3}{n^2\varepsilon^4}\big(\mathbf{E}(X_1^2)\big)^2. \tag{8}$$

4) On remarque que $\mathbf{E}\big(\mathbf{1}_{A_i}\big) = P(A_i) = p$. On peut alors écrire :

$$F_n - p = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{A_i} - \frac{1}{n}\sum_{i=1}^{n} P(A_i) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Il suffit donc d'appliquer (8) pour obtenir :

$$\forall \varepsilon > 0,\ \forall n \in \mathbb{N}^*, \qquad P\big(|F_n - p| > \varepsilon\big) \le \frac{C}{n^2}, \tag{9}$$

où la constante $C = C(\varepsilon, p)$ peut être choisie égale à :

$$C = \frac{\mathbf{E}(X_1^4)}{\varepsilon^4} + \frac{3}{\varepsilon^4}\big(\mathbf{E}(X_1^2)\big)^2 \le \frac{7}{\varepsilon^4}.$$

En effet, on a $|X_1| \le 1$ d'où $\mathbf{E}X_1^2 \le 1$ et $\mathbf{E}X_1^4 \le 1$.

5) L'événement contraire de $\{\forall n \ge n_0,\ |F_n - p| \le \varepsilon\}$ est :

$$\{\exists n \ge n_0,\ |F_n - p| > \varepsilon\} = \bigcup_{n \ge n_0} \{|F_n - p| > \varepsilon\}.$$

D'où :

$$P(\exists n \ge n_0,\ |F_n - p| > \varepsilon) \le \sum_{n \ge n_0} P(|F_n - p| > \varepsilon) \le \sum_{n \ge n_0} \frac{C}{n^2}.$$

Il suffit donc en suivant l'indication de l'énoncé de choisir n_0 tel que :

$$\sum_{n \ge n_0} \frac{C}{n^2} \le \frac{C}{n_0 - 1} < 0,01,$$

soit encore $n_0 > 100C + 1$ pour obtenir la conclusion souhaitée.

Justification : La majoration du reste de la série donnée par l'énoncé s'obtient par comparaison avec l'intégrale de x^{-2} : il suffit de remarquer (faites le dessin !) que :

$$\forall n > 1, \quad \frac{1}{n^2} \le \int_{n-1}^{n} \frac{dx}{x^2} \quad \text{d'où} \quad \sum_{n=N+1}^{+\infty} \frac{1}{n^2} \le \int_{N}^{+\infty} \frac{dx}{x^2} = \frac{1}{N}.$$

Commentaire : L'inégalité de Tchebycheff nous aurait donné à la place de (9) une inégalité du type $P\big(|F_n - p| > \varepsilon\big) \le C'/n$. Ce majorant étant le terme général d'une série *divergente*, la méthode de la question 5) n'aurait pas fonctionné. D'ailleurs le résultat obtenu ici est plus profond que l'inégalité de Tchebycheff. Cette inégalité peut nous fournir un n_1 ($n_1 \ge \frac{0,01}{C'}$) tel que F_{n_1} soit dans $[p-\varepsilon, p+\varepsilon]$ avec une probabilité d'au moins $0,99$. Elle ne nous dit rien sur le comportement ultérieur de F_n (pour $n > n_1$). Par contre, le résultat du problème nous permet d'affirmer avec une probabilité d'au moins $0,99$ que F_n *ne sortira plus jamais* de $[p-\varepsilon, p+\varepsilon]$ au delà du rang n_0.

Ex 2. *Estimation de quantiles (examen mai 2005).*

1) Rappelons que la continuité de la fonction g au point x_0 est caractérisée par le fait que pour toute suite $(u_n)_n$ tendant vers x_0, on a $\lim_{n\to+\infty} g(u_n) = g(x_0)$.

Supposons d'abord que $(Y_n)_{n\geq 1}$ converge presque surement vers Y et que la fonction g est continue sur tout $\mathbb{R}$. Cela implique qu'il existe $A \in \mathcal{F}$ tel que $P(A) = 1$ et

$$\forall \omega \in A \qquad \lim_{n\to+\infty} Y_n(\omega) = Y(\omega)$$

Cela implique aussi que pour tout $\omega \in \Omega$ la fonction g est continue au point $Y(\omega)$. On en déduit que :

$$\forall \omega \in A \qquad \lim_{n\to+\infty} g(Y_n(\omega)) = g(Y(\omega))$$

Comme $P(A) = 1$, ceci prouve que $g(Y_n)$ converge p.s. vers $g(Y)$.

On se place maintenant dans le cas où Y_n converge p.s. vers la constante c. Il existe $B \in \mathcal{F}$ tel que $P(B) = 1$ et

$$\forall \omega \in B \qquad \lim_{n\to+\infty} Y_n(\omega) = c$$

ce qui par continuité de la fonction g au point c implique que

$$\forall \omega \in B \qquad \lim_{n\to+\infty} g(Y_n(\omega)) = g(c)$$

donc $g(Y_n)$ converge p.s. vers $g(c)$.

2) Dans toute cette question, $n \in \mathbb{N}$ est fixé.

La fonction de répartition empirique F_n est définie par

$$\forall \omega \in \Omega, \quad t \mapsto F_n(\omega, t) := \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{]-\infty, X_k(\omega)]}(t)$$

et $Q_n(\omega)$ a été défini pour chaque $\omega \in \Omega$ par

$$Q_n(\omega) = F_n^{-1}(\omega, u) := \inf\{t \in \mathbb{R};\ F_n(\omega, t) \geq u\}.$$

ce qui par définition de l'inf se traduit par :

$$\begin{array}{ll} \forall t \in \mathbb{R} & F_n(\omega, t) \geq u \Rightarrow t \geq Q_n(\omega) \\ \text{et} & \exists (t_k)_k \text{ suite décroissant vers } Q_n(\omega) \text{ t.q. } \forall k \in \mathbb{N}\ \ F_n(\omega, t_k) \geq u. \end{array}$$

Remarquons que la suite $(t_k)_k$ ci-dessus dépend de ω.
Pour tout $\omega \in \Omega$, la fonction de répartition $F_n(\omega, \cdot)$ est continue à droite en tout point, donc en particulier au point $Q_n(\omega)$. Puisque $(t_k)_k$ est décroissante ceci implique que

$$\lim_{k\to+\infty} F_n(\omega, t_k) = F_n(\omega, \lim_{k\to+\infty} t_k) = F_n(\omega, Q_n(\omega))$$

et comme $\forall k \in \mathbb{N}$ $F_n(\omega, t_k) \geq u$ on obtient $F_n(\omega, Q_n(\omega)) \geq u$, pour chaque $\omega \in \Omega$. C'est la première des deux inégalités à établir.

D'autre part, s'il existait un $\omega \in \Omega$ pour lequel $F_n\Big(\omega, Q_n(\omega) - \frac{1}{n}\Big) \geq u$, alors le nombre $Q_n(\omega) - \frac{1}{n}$ appartiendrait à l'ensemble $\{t \in \mathbb{R};\ F_n(\omega, t) \geq u\}$ et donc l'inf de cet ensemble serait plus petit que $Q_n(\omega) - \frac{1}{n}$. Ceci est impossible puisque l'inf de cet ensemble est $Q_n(\omega)$. Par conséquent, $F_n\Big(\omega, Q_n(\omega) - \frac{1}{n}\Big) < u$ pour tous les $\omega \in \Omega$.

3) Sur tout Ω, on a $F(Q_n) = F_n(Q_n) - \Big(F_n(Q_n) - F(Q_n)\Big)$.
Or par définition de Δ_n

$$F_n(Q_n) - F(Q_n) \leq |F_n(Q_n) - F(Q_n)| \leq \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| = \Delta_n$$

donc $-(F_n(Q_n) - F(Q_n)) \geq -\Delta_n$ et

$$F(Q_n) \geq F_n(Q_n) - \Delta_n \geq u - \Delta_n$$

d'aprés l'inégalité $F_n(Q_n) \geq u$ établie à la question précédente.

La fonction de répartition F étant continue et strictement croissante sur $\mathbb{R}$, elle a un inverse F^{-1} continu et strictement croissant sur $]0, 1[$. La croissance de F^{-1} assure que

$$u - \Delta_n \leq F(Q_n) \quad \Longrightarrow \quad F^{-1}(u - \Delta_n) \leq F^{-1}(F(Q_n)) = Q_n$$

donc on a bien $F^{-1}(u - \Delta_n) \leq Q_n$.

Attention, les inégalités données ici sont des inégalités entre variables aléatoires, valables sur Ω. Vérifiez que vous pouvez traduire ce raisonnement pour n'utiliser que des inégalités entre nombres réels !

De même, on remarque que $F(Q_n - \frac{1}{n}) = F_n(Q_n - \frac{1}{n}) + \Big(F(Q_n - \frac{1}{n}) - F_n(Q_n - \frac{1}{n})\Big)$ et que

$$F(Q_n - \frac{1}{n}) - F_n(Q_n - \frac{1}{n}) \leq \left|F_n(Q_n - \frac{1}{n}) - F(Q_n - \frac{1}{n})\right| \leq \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| = \Delta_n$$

ce qui donne $F(Q_n - \frac{1}{n}) \leq F_n(Q_n - \frac{1}{n}) + \Delta_n$. Comme $F_n(Q_n - \frac{1}{n}) < u$ ceci implique que $F(Q_n - \frac{1}{n}) < u + \Delta_n$ et la stricte croissance de F^{-1} permet d'en déduire que $Q_n - \frac{1}{n} < F^{-1}(u + \Delta_n)$.

On obtient l'encadrement

$$F^{-1}(u - \Delta_n) \leq Q_n < \frac{1}{n} + F^{-1}(u + \Delta_n).$$

pour tout $n \geq 1$ et tout $\omega \in \Omega$.

4) Le théorème de Glivenko-Cantelli affirme que la variable aléatoire Δ_n tend p.s. vers 0 quand n tend vers $+\infty$:

$$P(E) = 1 \text{ où } E = \{\ \omega \in \Omega,\ \lim_{n \to +\infty} \Delta_n(\omega) = 0\ \}$$

Pour tout $\omega \in E$, les suites $(u - \Delta_n(\omega))_n$ et $(u + \Delta_n(\omega))_n$ convergent vers u. La continuité de la fonction F^{-1} au point u implique alors que $F^{-1}(u - \Delta_n(\omega))$ et $F^{-1}(u + \Delta_n(\omega))$ convergent vers $F^{-1}(u)$. Par conséquent, $F^{-1}(u - \Delta_n)$ et $F^{-1}(u + \Delta_n)$ convergent presque-sûrement vers $q = F^{-1}(u)$.

5) Comme on a prouvé que

$$\forall \omega \in \Omega \quad F^{-1}(u - \Delta_n(\omega)) \leq Q_n(\omega) < \frac{1}{n} + F^{-1}(u + \Delta_n)$$

le théorème des gendarmes assure que pour $\omega \in E$ la suite des $Q_n(\omega)$ converge vers $F^{-1}(u)$. En se rappelant que $P(E) = 1$, on obtient bien que Q_n converge p.s. vers $F^{-1}(u)$ lorsque n tend vers l'infini.