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

IS Math314 **Année 2005-2006**

Devoir 1

à rendre en TD la semaine du 27 février.

Ex 1. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de même loi telle que $\mathbf{E}X_n = 0$. On suppose de plus que $\mathbf{E}X_n^4 < +\infty$.

1) Montrer que pour tout $(i,j,k,l) \in \{1,\dots,n\}^4$, $\mathbf{E}|X_iX_jX_kX_l| < +\infty$. En déduire en utilisant l'identité :

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

que :

$$\mathbf{E}\Big(\sum_{i=1}^n X_i\Big)^4 < +\infty.$$

2) Montrer que les termes $\mathbf{E}(X_iX_jX_kX_l)$ sont tous nuls à l'exception de ceux qui sont de la forme $\mathbf{E}(X_i^4)$ $(1 \leq i \leq n)$ ou $\mathbf{E}(X_i^2X_j^2)$ $(i \neq j)$.

3) En déduire :

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

puis :

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

4) On considère une suite d'épreuves répétées indépendantes et on note A_i l'événement *succès à la i^{e} épreuve*. On pose $p = P(A_i)$ et $X_i = \mathbf{1}_{A_i} - \mathbf{E}\big(\mathbf{1}_{A_i}\big)$. La fréquence des succès obtenus en n épreuves est la variable aléatoire :

$$F_n = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{A_i}.$$

En utilisant ce qui précède, montrer qu'il existe une constante $C = C(\varepsilon, p)$ telle que :

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

5) En déduire que :

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N}^*, \qquad P\big(\forall n \geq n_0,\ |F_n - p| \leq \varepsilon\big) \geq 0{,}99.$$

Indication : On pourra utiliser après l'avoir justifiée la majoration :

$$\sum_{n=N+1}^{+\infty} \frac{1}{n^2} \leq \frac{1}{N}.$$

Pouvait-on obtenir le même genre de résultat avec l'inégalité de Tchebycheff ?

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

Soit F une fonction de répartition continue et strictement croissante sur $\mathbb{R}$. On sait qu'elle réalise alors une bijection de $\mathbb{R}$ sur $]0,1[$ et que son inverse (au sens classique) F^{-1} est *continue* sur $]0,1[$ et strictement croissante sur cet intervalle. Pour tout $u \in]0,1[$, on définit le quantile associé q par

$$q := F^{-1}(u).$$

On se propose d'estimer q pour *u fixé* à partir d'une suite de variables aléatoires $(X_k)_{k\geq 1}$, définies sur le même espace probabilisé $(\Omega, \mathcal{F}, P)$, indépendantes et de même loi de fonction de répartition F.

On rappelle que la fonction de répartition empirique F_n construite sur l'échantillon $X_1, \dots, X_n$ 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 que pour tout $\omega \in \Omega$, $F_n(\omega, .)$ est une fonction de répartition donc en particulier *croissante* au sens large sur $\mathbb{R}$ et *continue à droite* en tout point de $\mathbb{R}$. Pour t fixé dans $\mathbb{R}$, l'application partielle $F_n(., t)$, notée plus simplement $F_n(t)$ et définie par $F_n(t) : \Omega \to [0,1]$, $\omega \mapsto F_n(\omega, t)$ est une variable aléatoire.

Pour estimer q, on introduit la suite de variables aléatoires Q_n définies par

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

1) Soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires et Y une variable aléatoire, définies sur le même $(\Omega, \mathcal{F}, P)$. On suppose que Y_n converge presque-sûrement vers Y quand n tend vers $+\infty$ et que $g : \mathbb{R} \to \mathbb{R}$ est une fonction continue sur $\mathbb{R}$. Expliquez pourquoi $g(Y_n)$ converge p.s. vers $g(Y)$. Justifiez de même la convergence p.s. de $g(Y_n)$ vers $g(c)$ lorsque Y_n converge p.s. vers la constante c et g est une fonction borélienne [1] $\mathbb{R} \to \mathbb{R}$, continue au point c.

2) Justifiez les inégalités entre variables aléatoires (*i.e.* vraies pour tout $\omega \in \Omega$) :

$$\forall n \geq 1, \quad F_n(Q_n) \geq u, \quad \text{et} \quad F_n\Big(Q_n - \frac{1}{n}\Big) < u.$$

1. On est ainsi assurés que $g(Y_n)$ est une v.a.

3) On définit la variable aléatoire positive Δ_n par

$$\forall \omega \in \Omega, \quad \Delta_n(\omega) := \|F_n(\omega, .) - F\|_\infty = \sup_{t\in\mathbb{R}} |F_n(\omega, t) - F(t)|.$$

Montrez que pour tout $n \geq 1$,

$$F(Q_n) \geq u - \Delta_n \quad \text{et} \quad F\Big(Q_n - \frac{1}{n}\Big) < u + \Delta_n$$

et en déduire l'encadrement (vrai pour tout $n \geq 1$ et tout $\omega \in \Omega$) :

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

Le devoir à rendre s'arrête ici. Lorsque cet exercice a été posé en examen, il comportait deux questions supplémentaires, nécessitant l'emploi du théorème de Glivenko-Cantelli, qui sont données ici à titre informatif :

4) Expliquez pourquoi $F^{-1}(u - \Delta_n)$ converge presque-sûrement vers $q = F^{-1}(u)$ lorsque n tend vers l'infini.

5) Montrez que Q_n converge p.s. vers q lorsque n tend vers l'infini.