Feuille d'exercice 2

Juqu'à la limite

1 Les grands nombres font la loi

1.1 Beber

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi de Bernoulli $\mathcal{B}(p)$ avec $0 < p < 1$. Pour tout $n \geq 1$, on pose

$$S_n = \sum_{k=1}^{n} X_k \qquad \text{et} \qquad \overline{X}_n = \frac{S_n}{n}.$$

Montrer que $\overline{X}_n$ est un estimateur sans biais de p et $\overline{X}_n \to p$ p.s. Afin d'estimer la variance $\sigma^2 = p(1-p)$, on propose d'utiliser $U_n = \overline{X}_n(1-\overline{X}_n)$. Montrer que U_n est un estimateur biaisé de σ^2 et trouver un estimateur V_n sans biais de σ^2 tel que $V_n \to \sigma^2$ p.s. Créer un code Matlab illustrant ces deux LGN sur un n-échantillon de loi de Bernoulli $\mathcal{B}(p)$, où les paramètres n et p sont affectés par l'utilisateur.

1.2 Ampoule

La durée de vie d'une ampoule électrique peut être modélisée par une variable aléatoire X prenant au hasard ses valeurs dans l'intervalle $[0,\theta]$ avec $\theta > 0$. Afin d'optimiser l'agenda d'un réparateur, on cherche à estimer θ à partir d'un n-échantillon $(X_1, X_2, \cdots, X_n)$ de même loi que X. On propose d'estimer θ par

$$\widehat{\theta}_n = \max_{1\leq k\leq n} X_k \qquad \text{ou} \qquad \widetilde{\theta}_n = 2\overline{X}_n.$$

Calculer la fonction de répartition puis la densité de probabilité de $\widehat{\theta}_n$ et en déduire que $\widehat{\theta}_n \to \theta$ p.s. Montrer également que $\widetilde{\theta}_n \to \theta$ p.s. Créer un code Matlab illustrant ces deux LGN sur un n-échantillon de loi uniforme $\mathcal{U}([0,\theta])$, où les paramètres n et θ sont affectés par l'utilisateur.

1.3 Paréto

La loi de Paréto, encore appelée loi de puissance, est utilisée pour modéliser les dépassements d'un seuil. On dit que X suit une loi de Paréto de paramètres $a, \theta > 0$ si $X = \theta \exp(Y)$ avec Y de loi exponentielle $\mathcal{E}(a)$. On propose d'estimer θ, à partir d'un n-échantillon $(X_1, X_2, \cdots, X_n)$ de même loi que X, par

$$\widehat{\theta}_n = \min_{1\leq k\leq n} X_k.$$

Calculer la fonction de répartition puis la densité de probabilité de X. En déduire la loi de probabilité associée à $\widehat{\theta}_n$ et montrer que $\widehat{\theta}_n \to \theta$ p.s. Créer un code Matlab illustrant cette LGN sur un n-échantillon de loi de Paréto $\mathcal{P}(a,\theta)$, où les paramètres n, a et θ sont affectés par l'utilisateur.

1.4 Convergence

Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires i.i.d. de loi la loi de X. Montrer que :

a) Si $E(|X|) < \infty$, alors (X_n/n) converge p.s. vers 0. (indication : X est intégrable ssi $\sum_{n>0} P(|X| \geq an) < \infty$).

b) Si $E(|X|) = \infty$, alors, pour tout A positif, $P(\limsup\{|X_n| \geq An\}) = 1$.

c) Soit $S_n = \sum_{i=1}^{n} X_i$. Déduire de (b) que si $E(|X|) = \infty$, alors $\limsup_n |S_n/n| = +\infty$ p.s.

1.5 Gauss

Soit $(X_k)_{k>0}$ une suite de v.a. indépendantes de loi gaussienne standard $\mathcal{N}(0,1)$. On pose $S_n = \sum_{i=1}^{n} X_i$.

a) Montrer que $\sqrt{2\pi}P(X_0 > a) \sim \frac{1}{a}\exp(-a^2/2)$ quand $a \to +\infty$.

b) Montrer que $S_n/\sqrt{n}$ suit une loi $\mathcal{N}(0,1)$ (calculer la fonction caractéristique de S_n). En déduire que si (a_n) est une suite de réels positifs telle que $a_n/\sqrt{n}$ tende vers $+\infty$. alors S_n/a_n converge vers 0 en probabilité. Peut-on conclure pour la convergence p.s. ? Montrer cependant que si $a_n = \sqrt{n}\log n$, alors S_n/a_n converge p.s. vers 0.

c) Montrer que $\limsup_n (2\log n)^{-1/2} X_n = 1$ p.s. et $\limsup_n (2\log n)^{-1/2}|X_n| = 1$ p.s.

1.6 Moment d'ordre 4

Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires i.i.d. et centrées ($E(X_i) = 0$). On suppose que $E(X_i^4) < +\infty$.On pose $S_n = \sum_{i=1}^n X_i$. Montrer que

$$P(|S_n| > n\varepsilon) \leq \frac{1}{(n\varepsilon)^4} E(S_n^4).$$

Calculer $E(S_n^4)$ et en déduire que S_n/n converge p.s. vers 0.

1.7 Borel et ses frères

Montrer que si les v.a X_n sont indépendantes alors X_n converge p.s. vers 0 si et seulement si, pour tout $\epsilon > 0$, $\sum_n P(|X_n| > \epsilon) < \infty$

2 Théorème Limite Centrale.

2.1 Pour les dimanche plus vieux

Utiliser le code Matlab suivant qui permet d'illustrer le TLC sur un n-échantillon de loi Uniforme $\mathcal{U}([0,1])$.

$n = input$('Entrez la taille de l'échantillon n (par exemple 10) : ') ;
$N = input$('Entrez le nombre de réalisations N (par exemple 2000) : ') ;
$d = input$('Préciser la discrétisation d (par exemple 100) : ') ;
$X = rand(n, N); Z = (sum(X) - n/2)/sqrt(n/12);$
$dz = range(Z)/d$; [Effectifs, Classes]=$hist(Z, d)$;
$Abscisses = [min(Z) - dz/2 : dz : max(Z) + dz/2]$;
DS =[0 Effectifs 0]/$(N * dz)$; $bar(Abscisses, DS,' b')$; *hold on*
$DN = dnorm$(Abscisses) ; $plot(Abscisses, DN,' r-')$;
legend('Densite simulée','Loi Normale') ; *hold off*

2.2 Hans

Soit $(\varepsilon_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi de Rademacher $\mathcal{R}(1/2)$. Soit $(X_n)_{n\geq 1}$ la suite définie par $X_n = n^a \varepsilon_n$ avec $a > 0$. On pose $S_n = X_1 + X_2 + \cdots + X_n$. Calculer la transformée de Laplace de S_n donnée, pour tout $t \in \mathbb{R}$, par $L_n(t) = \mathbb{E}[\exp(tS_n)]$. En déduire que

$$\frac{S_n}{n^a\sqrt{n}} \longrightarrow_{\mathcal{L}} \mathcal{N}\Big(0, \frac{1}{2a+1}\Big).$$

Créer un code Matlab illustrant ce TLC sur un n-échantillon de loi de Rademacher $\mathcal{R}(1/2)$, où les paramètres n et a sont affectés par l'utilisateur.

2.3 Poisson

Pour tout $x, \lambda > 0$, on pose

$$L_n(x) = \exp(-n\lambda) \sum_{k=0}^{[nx]} \frac{(n\lambda)^k}{k!}.$$

Montrer à l'aide de la loi des grands nombres et du théorème de la limite centrale que

$$\lim_{n\to} L_n(x) = \begin{cases} 0 & \text{si } x < \lambda, \\ 1/2 & \text{si } x = \lambda, \\ 1 & \text{si } x > \lambda. \end{cases}$$

Vérifier cette convergence avec un code Matlab.

2.4 Pink

Soit X une v.a.r. telle que $P(X = k) = p_k$, $1 \leq k \leq r$, $p_1 + \ldots + p_r = 1$. Soit $(X_1, ..., X_n)$ un échantillon de X. On pose $V_i = (\mathbb{1}_{X_i=1}, ..., \mathbb{1}_{X_i=r})$.

a) Montrer que $S_n = V_1 + \ldots + V_n$ est le vecteur aléatoire de $\mathbb{R}^r$ dont la k-ième composante $S_n^{(k)}$ est égale au nombre d'entiers i de $[1, n]$ tels que $X_i = k$.

b) Montrer que $n^{-1/2}(S_n - E(S_n))$ converge en loi vers un vecteur gaussien dont on déterminera la matrice de covariance.

c) Montrer que $Y_n = n^{-1/2}(p_1^{-1/2}(S_n^{(1)} - E(S_n^{(1)})), ..., (p_r^{-1/2}(S_n^{(r)} - E(S_n^{(r)}))$ converge vers un vecteur gaussien dont on déterminera la matrice de covariance Γ.

d) Soit $(Z_1, ..., Z_r)$ une vecteur gaussien centré de matrice de covariance Γ. Montrer que $Z_1^2 + \ldots + Z_r^2$ suit une loi du χ^2 à $r-1$ degrés de libertés. En déduire que le carré de la norme euclidienne de Y_n converge en loi vers χ^2_{r-1}.

2.5 Ze Wall

Soit X_n un échantillon d'une v.a.r. de loi $N(0,1)$. Soit $Y_n(h) = \sum_{i=1}^n \sin(hX_i)$. Montrer que $\frac{1}{\sqrt{n}}(Y_n(1), ..., Y_n(k))$ converge en loi vers un vecteur gaussien centré dont on donnera la matrice de covariance.