TD 3 : Observations asymptotiques

Exercice 1 (Temps d'attente d'un bus)**.** Soit $\delta > 0$. Supposons qu'une expérience soit tentée à chaque temps $\delta, 2\delta, 3\delta, \dots$, et qu'à chaque essai le résultat est 1 avec probabilité $p_\delta \in [0,1]$ ou 0 avec probabilité $1 - p_\delta$, indépendamment des autres essais (par exemple, on demande après chaque intervalle de temps de longueur δ si le bus est passé). On appelle T_δ le temps à attendre avant le premier succès (i.e., la première fois que l'on obtient 1).

1) Montrer que pour tout entier $k \geq 0$, $\mathbb{P}(T_\delta > k\delta) = (1 - p_\delta)^k$.

2) Calculer $\mathbb{E}(T_\delta)$.

3) Soit $t \geq 0$. On suppose que $\lim_{\delta \to 0}(p_\delta/\delta) = \lambda > 0$. Montrer que

$$\lim_{\delta \to 0} \mathbb{P}(T_\delta > t) = e^{-\lambda t}.$$

En conclure que T_δ converge lorsque δ tend vers 0, en un sens et vers une limite que l'on précisera.

Exercice 2. Soit $\theta > 0$. Soient $X_1, \dots, X_n, \dots$ une suite de variables aléatoires indépendantes et de même loi uniforme sur $[0, \theta]$. Pour tout $n \in \mathbb{N}$, notons $\widehat{\theta}_n = \max\limits_{1 \leq i \leq n} X_i$. Montrer que $n(\theta - \widehat{\theta}_n) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{E}(1/\theta)$.

Exercice 3 (Méthode de Monte-Carlo)**.** On cherche à approcher l'intégrale d'une fonction g définie sur $[0,1]$ et à valeurs dans $[0,1]$ (par exemple la fonction $x \mapsto \sqrt{1 - x^2}$ définie sur $[0,1]$ qui admet pour intégrale $\pi/4$). Si on trace le carré de côté 1 et la courbe $\{(x, g(x)),\ x \in [0,1]\}$, l'intégrale recherchée est l'aire de la partie du carré qui se situe sous la courbe.

On applique alors l'algorithme suivant : on définit X et Y deux variables aléatoires indépendantes de loi uniforme sur $[0,1]$. On déclare alors qu'une réalisation (x, y) de (X, Y) est réussie si $y \leq g(x)$. On répète l'expérience un grand nombre de fois, chaque expérience étant réalisée indépendamment des autres. On obtient donc une suite de couples $(x_n, y_n)_{n \geq 1}$.

1) Pour tout $n \geq 1$, on définit A_n comme l'événement "la $n^{\text{ième}}$ expérience est réussie", et on pose

$$S_n = \sum_{k=1}^{n} \mathbb{1}_{A_k}.$$

Montrer que S_n/n converge lorsque n tend vers $+\infty$, en un sens et vers une limite que l'on précisera.

2) On suppose que l'on dispose de la possibilité de simuler la réalisation de plusieurs variables aléatoires de loi uniforme sur $[0,1]$ et indépendantes les unes des autres (par exemple avec une fonction RAND). Conclure en proposant une méthode pour obtenir une valeur approchée de l'intégrale de g sur $[0,1]$.

Exercice 4. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes. On suppose que chaque X_n suit une loi de Bernoulli de paramètre p_n, avec $p_n = \frac{1}{3}$ pour n impair et $p_n = \frac{2}{3}$ pour n pair. On note $\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$, étudier la convergence de $\overline{X}_n$.

Exercice 5 (Loi de Cauchy)**.** La loi de Cauchy de paramètre $a > 0$ est la loi d'une variable aléatoire admettant pour densité sur $\mathbb{R}$

$$f_a(x) = \frac{a/\pi}{a^2 + x^2}$$

1) Vérifier que cette définition a un sens. Que dire de l'espérance et de la variance d'une telle loi ?

2) On admet que la fonction caractéristique de X qui suit une loi de Cauchy de paramètre a vaut $\phi_X(u) = e^{-a|u|}$. Montrer que la loi de Cauchy possède la propriété de stabilité suivante : si $X_1, \ldots, X_n$ sont n variables aléatoires indépendantes suivant une loi de Cauchy de paramètre a, alors il en est de même pour $(X_1 + \ldots + X_n)/n$.

Exercice 6. Soit $(X_\lambda)_{\lambda>0}$ une famille de variables aléatoires, où le terme général X_λ suit une loi de Poisson de paramètre λ. Montrer que $(X_\lambda - \lambda)/\sqrt{\lambda}$ converge en loi vers une variable aléatoire de loi $\mathcal{N}(0,1)$, lorsque λ tend vers l'infini.

Exercice 7. 1) Montrer que

$$e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \underset{n\to\infty}{\longrightarrow} \frac{1}{2}.$$

2) Soit $0 < p < 1$, $p + q = 1$. Montrer que si $p > q$, alors

$$\sum_{k=\lfloor n/2 \rfloor + 1}^{n} \binom{n}{k} p^k q^{n-k} \underset{n\to\infty}{\longrightarrow} 1.$$

Exercice 8. Soit $(X_n)_{n\geq 1}$ une famille de variables aléatoires i.i.d avec $\mathbb{E}[X_1^2] < \infty$. On note $m = \mathbb{E}[X_1]$, $\sigma^2 = \mathrm{Var}(X)$.

1) On pose $\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$. Montrer que

$$\sigma_n^2 \xrightarrow[n\to\infty]{\mathbb{P}} \sigma^2.$$

2) Montrer la version du TCL avec variance inconnue :

$$\sqrt{n}\frac{(\overline{X}_n - m)}{\sigma_n} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

Exercice 9 (Majoration de la transformée de Laplace)**.** Par définition, la transformée de Laplace d'une v.a. X est la fonction

$$\phi(\lambda) = \mathbb{E}[e^{\lambda X}] \qquad (\lambda \in \mathbb{R}).$$

Le but de cet exercice est de montrer le lemme suivant :

Lemme. Soit X une v.a. telle que $\mathbf{0 \leq X \leq 1}$. Alors $\forall \lambda > 0$

$$\mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}] \leq \exp\left(\frac{\lambda^2}{8}\right).$$

a) On introduit $\psi(\lambda) = \log \phi(\lambda)$: calculer les dérivées première et seconde de ψ.

b) Montrer que

$$\psi''(\lambda) \leq \theta - \theta^2, \quad \text{où } \theta = \frac{\mathbb{E}[Xe^{\lambda X}]}{\mathbb{E}[e^{\lambda X}]}.$$

c) En déduire une majoration de ψ'', puis montrer que pour tout $\lambda > 0$,

$$\log \mathbb{E}[e^{\lambda X}] \leq \lambda \mathbb{E}[X] + \frac{\lambda^2}{8}.$$

Conclure.

Exercice 10 (L'inégalité de Hoeffding)**.** Soient $X_1, \ldots, X_n$ des v.a. indépendantes telles que $\forall i, \quad 0 \leq X_i \leq 1$. On note

$$Z = \sum_{i=1}^{n} \Big(X_i - \mathbb{E}[X_i] \Big).$$

a) Démontrer que pour tout $\lambda > 0$,

$$\mathbb{P}\Big(Z > u \Big) \leq \mathbb{E}[e^{\lambda(Z-u)}].$$

b) Soit $\lambda > 0$. Démontrer que

$$\mathbb{E}[e^{\lambda Z}] \leq \exp\Big(n \frac{\lambda^2}{8} \Big).$$

c) En déduire l'inégalité de Hoeffding,

$$\mathbb{P}\Big(\sum_{i=1}^{n} \big(X_i - \mathbb{E}[X_i] \big) > u \Big) \leq \exp\Big(- \frac{2u^2}{n} \Big).$$

d) En changeant X_i en $1 - X_i$, montrer que

$$\mathbb{P}\Big(\Big| \sum_{i=1}^{n} \big(X_i - \mathbb{E}[X_i] \big) \Big| > u \Big) \leq 2 \exp\Big(- \frac{2u^2}{n} \Big).$$

Dans le cours figurent deux estimations de l'écart d'une moyenne empirique par rapport à son espérance : une estimation exacte, l'inégalité de Tchebychev, et une estimation asymptotique, le Théorème Central Limite. L'inégalité de Hoeffding fournit une deuxième estimation exacte, souvent plus précise que l'inégalité de Chebyshev. C'est le premier résultat de ce qu'on appelle en mathématiques la théorie de la *concentration de la mesure*, très utilisée en statistiques.

Ces estimations peuvent servir par exemple à évaluer la précision de l'estimateur usuel de l'espérance dans le modèle de Bernoulli, correspondant pour un sondage à la proportion de personnes qui vont voter pour un candidat donné.