

Mise à niveau en statistique

Correction de la feuille de TD 1 : Outils de probabilités

M2 Mathématiques de l'Aléatoire – Université Paris-Saclay

Exercice 1 – Modes de convergence

1. Supposons que X_n converge vers X en moyenne quadratique. Pour tout $\varepsilon > 0$, l'inégalité de Markov donne

$$\mathbb{P}(|X_n - X| > \varepsilon) \leq \frac{\mathbb{E}[(X_n - X)^2]}{\varepsilon^2} \longrightarrow 0.$$

Ainsi, $X_n \xrightarrow{\mathbb{P}} X$.

2. Supposons que $X_n \xrightarrow{\mathcal{L}} a$, où a est déterministe. Pour $\varepsilon > 0$, l'ensemble fermé $F_\varepsilon = (-\infty, a - \varepsilon] \cup [a + \varepsilon, +\infty)$ ne contient pas a . Le théorème de Portmanteau entraîne

$$\limsup_{n \rightarrow \infty} \mathbb{P}(|X_n - a| \geq \varepsilon) \leq \mathbb{P}(a \in F_\varepsilon) = 0.$$

Par conséquent, $X_n \xrightarrow{\mathbb{P}} a$.

3. D'après la question précédente, $X_n \xrightarrow{\mathbb{P}} a$. Soit $\varepsilon > 0$. Par continuité de h en a , il existe $\eta > 0$ tel que

$$|x - a| < \eta \implies |h(x) - h(a)| < \varepsilon.$$

Il s'ensuit que

$$\mathbb{P}(|h(X_n) - h(a)| \geq \varepsilon) \leq \mathbb{P}(|X_n - a| \geq \eta) \longrightarrow 0.$$

4.
 - Si $X_n = 1/n$, la suite, déterministe, converge vers 0 presque sûrement et dans L^p pour tout $p \geq 1$; elle converge donc aussi en probabilité et en loi.
 - Si $X_n = (-1)^n$, la suite ne converge ni en probabilité ni presque sûrement. Elle ne converge pas davantage en loi, puisque les sous-suites paires et impaires ont pour lois respectives δ_1 et δ_{-1} . Enfin, elle ne converge dans aucun espace L^p , $p \geq 1$.
 - Si $X_n = \mathbf{1}_{A_n}$ et $\mathbb{P}(A_n) \rightarrow 0$, alors

$$\mathbb{E}[X_n^2] = \mathbb{P}(A_n) \longrightarrow 0.$$

Donc $X_n \rightarrow 0$ dans L^2 , en probabilité et en loi. La convergence presque sûre n'est toutefois pas garantie. Par exemple, sur $([0, 1], \mathcal{B}, \text{Leb})$, énumérons par blocs les intervalles dyadiques

$$A_{2^k+j} = [j2^{-k}, (j+1)2^{-k}), \quad 0 \leq j < 2^k, \text{quad } k \geq 0.$$

Alors $\mathbb{P}(A_n) \rightarrow 0$, mais chaque $\omega \in [0, 1)$ appartient à exactement un intervalle de chaque bloc : la suite $\mathbf{1}_{A_n}(\omega)$ prend donc les valeurs 0 et 1 infiniment souvent et ne converge pas.

- Posons $D_n = Z_n \mathbf{1}_{B_n} - Z_n$. Pour tout $\varepsilon > 0$,

$$\mathbb{P}(|D_n| > \varepsilon) \leq \mathbb{P}(B_n^c) \longrightarrow 0,$$

donc $D_n \xrightarrow{\mathbb{P}} 0$. Comme $Z_n \xrightarrow{\mathcal{L}} Z$, le lemme de Slutsky donne $Z_n \mathbf{1}_{B_n} = Z_n + D_n \xrightarrow{\mathcal{L}} Z$. En général, aucune convergence en probabilité vers Z ne peut être déduite de la seule convergence en loi. Par exemple, si $B_n = \Omega$ et si Z_n sont i.i.d. de loi $\mathcal{N}(0, 1)$, alors $Z_n \xrightarrow{\mathcal{L}} Z_1$, mais Z_n ne converge pas en probabilité vers Z_1 .

Exercice 2 – Loi exponentielle et delta-méthode

1. La variable $Z = -\log X$ est positive et, pour $z \geq 0$,

$$\mathbb{P}(Z \leq z) = \mathbb{P}(X \geq e^{-z}) = 1 - e^{-z}.$$

Ainsi, Z suit une loi exponentielle de paramètre 1.

2. Puisque $\mathbb{E}[Z_1] = \text{Var}(Z_1) = 1$, le théorème central limite donne

$$\sqrt{n}(\bar{Z}_n - 1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

3. Comme

$$Y_n = \left(\prod_{i=1}^n X_i \right)^{-1/n} = \exp\left(-\frac{1}{n} \sum_{i=1}^n \log X_i\right) = e^{\bar{Z}_n},$$

on applique la delta-méthode à $g(z) = e^z$ en $z = 1$. Puisque $g'(1) = e$,

$$\sqrt{n}(Y_n - e) \xrightarrow{\mathcal{L}} \mathcal{N}(0, e^2).$$

Exercice 3 – Inégalité de Hoeffding simplifiée

Notons $S_n = \sum_{i=1}^n X_i$.

1. La fonction $u \mapsto e^{\lambda u}$ est convexe. Tout $x \in [-M, M]$ s'écrit

$$x = \frac{M-x}{2M}(-M) + \frac{x+M}{2M}M.$$

Par convexité,

$$e^{\lambda x} \leq \frac{M-x}{2M}e^{-\lambda M} + \frac{x+M}{2M}e^{\lambda M}.$$

2. En appliquant cette inégalité à X_i puis en prenant l'espérance, et puisque $\mathbb{E}[X_i] = 0$,

$$\mathbb{E}[e^{\lambda X_i}] \leq \frac{1}{2}(e^{-\lambda M} + e^{\lambda M}) = \cosh(\lambda M) \leq e^{\lambda^2 M^2/2}.$$

La dernière inégalité vaut pour tout réel u sous la forme $\cosh(u) \leq e^{u^2/2}$.

3. Pour $\lambda > 0$, l'inégalité de Markov et l'indépendance donnent

$$\begin{aligned} \mathbb{P}\left(\frac{S_n}{n} > x\right) &= \mathbb{P}(e^{\lambda S_n} > e^{\lambda n x}) \\ &\leq e^{-\lambda n x} \mathbb{E}[e^{\lambda S_n}] = e^{-\lambda n x} \prod_{i=1}^n \mathbb{E}[e^{\lambda X_i}] \\ &\leq \exp\left(n\left(\frac{\lambda^2 M^2}{2} - \lambda x\right)\right). \end{aligned}$$

4. Le membre de droite est minimal pour $\lambda = x/M^2$. On obtient

$$\mathbb{P}\left(\frac{S_n}{n} > x\right) \leq \exp\left(-\frac{nx^2}{2M^2}\right).$$

En appliquant la même borne aux variables $-X_i$ puis en réunissant les deux queues,

$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| > x\right) \leq 2 \exp\left(-\frac{nx^2}{2M^2}\right).$$

Exercice 4 – Inégalités classiques

1. Inégalité de Cauchy–Schwarz

(a) Appliquons l'inégalité de Cauchy–Schwarz dans L^2 aux variables $|X|$ et $|Y|$:

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]} \sqrt{\mathbb{E}[Y^2]}.$$

On peut aussi la démontrer en observant que $\mathbb{E}[(t|X| - |Y|)^2] \geq 0$ pour tout $t \in \mathbb{R}$ et que le discriminant de ce trinôme est donc négatif ou nul.

(b) En appliquant cette inégalité à X_n^2 et Y_n^2 ,

$$\mathbb{E}[(X_n Y_n)^2] \leq \sqrt{\mathbb{E}[X_n^4]} \sqrt{\mathbb{E}[Y_n^4]} \longrightarrow 0.$$

Ainsi, $X_n Y_n \rightarrow 0$ en moyenne quadratique.

2. Inégalité de Hölder

(a) Pour $p, q > 1$ conjugués, la concavité du logarithme donne, lorsque $a, b > 0$,

$$\log(ab) = \frac{1}{p} \log(a^p) + \frac{1}{q} \log(b^q) \leq \log\left(\frac{a^p}{p} + \frac{b^q}{q}\right).$$

En passant à l'exponentielle, on obtient l'inégalité de Young

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Les cas $a = 0$ ou $b = 0$ sont immédiats.

(b) Supposons d'abord $\|X\|_p \|Y\|_q > 0$ et appliquons l'inégalité de Young à $a = |X|/\|X\|_p$ et $b = |Y|/\|Y\|_q$. Après intégration,

$$\frac{\mathbb{E}[|XY|]}{\|X\|_p \|Y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

On en déduit

$$\mathbb{E}[|XY|] \leq (\mathbb{E}|X|^p)^{1/p} (\mathbb{E}|Y|^q)^{1/q}.$$

Si l'une des deux normes est nulle, le résultat est immédiat.

(c) Prenons $p = r/2$ et $q = r/(r-2)$, qui sont conjugués puisque $2 < r < 4$. En appliquant Hölder à $|X_n|^2$ et $|Y_n|^2$,

$$\mathbb{E}[X_n^2 Y_n^2] \leq (\mathbb{E}|X_n|^r)^{2/r} (\mathbb{E}|Y_n|^{2r/(r-2)})^{(r-2)/r} \longrightarrow 0.$$

Donc $X_n Y_n \rightarrow 0$ en moyenne quadratique.

Exercice 5 – Lemme de Slutsky

Comme la limite de Y_n est la constante c , la convergence $Y_n \xrightarrow{\mathcal{L}} c$ implique $Y_n \xrightarrow{\mathbb{P}} c$ d'après l'exercice 1.

Fixons $(s, t) \in \mathbb{R}^2$. On a

$$\begin{aligned} & \left| \mathbb{E}[e^{i(sX_n + tY_n)}] - e^{itc} \mathbb{E}[e^{isX_n}] \right| \\ & \leq \mathbb{E}[|e^{itY_n} - e^{itc}|] \longrightarrow 0. \end{aligned}$$

En effet, $|e^{itY_n} - e^{itc}| \rightarrow 0$ en probabilité et cette quantité est bornée par 2, donc son espérance tend vers 0. Par ailleurs, $X_n \xrightarrow{\mathcal{L}} X$ implique

$$e^{itc} \mathbb{E}[e^{isX_n}] \longrightarrow e^{itc} \mathbb{E}[e^{isX}] = \mathbb{E}[e^{i(sX + tc)}].$$

Les fonctions caractéristiques de (X_n, Y_n) convergent donc vers celle de (X, c) . Le théorème de Lévy permet de conclure que

$$(X_n, Y_n) \xrightarrow{\mathcal{L}} (X, c).$$

Le caractère déterministe de la limite est essentiel : deux convergences marginales en loi ne déterminent pas la dépendance limite. Par exemple, si $X \sim \mathcal{N}(0, 1)$, $X_n = X$ et $Y_n = (-1)^n X$, alors $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} X$, mais la loi du couple alterne entre celles de (X, X) et $(X, -X)$ et ne converge pas.

Exercice 6 – Delta-méthode

1. La suite $a_n(U_n - U)$ converge en loi, donc elle est tendue. Comme $a_n \rightarrow +\infty$, il en résulte que $U_n - U \xrightarrow{\mathbb{P}} 0$, c'est-à-dire $U_n \xrightarrow{\mathbb{P}} U$.

Définissons, pour $x \neq U$,

$$R(x) = \frac{\ell(x) - \ell(U)}{x - U} - \ell'(U), \quad R(U) = 0.$$

La dérivabilité de ℓ en U assure que R est continue en U . Ainsi, $R(U_n) \xrightarrow{\mathbb{P}} 0$. Or

$$a_n[\ell(U_n) - \ell(U)] - \ell'(U)a_n(U_n - U) = a_n(U_n - U)R(U_n).$$

Le premier facteur est tendu et le second converge en probabilité vers 0; leur produit converge donc en probabilité vers 0.

2. Par hypothèse et continuité de la multiplication par la constante $\ell'(U)$,

$$\ell'(U)a_n(U_n - U) \xrightarrow{\mathcal{L}} \ell'(U)V.$$

En combinant cette convergence avec la question précédente, le lemme de Slutsky donne

$$a_n[\ell(U_n) - \ell(U)] \xrightarrow{\mathcal{L}} \ell'(U)V.$$

3. Pour un échantillon i.i.d. de loi de Poisson de paramètre λ ,

$$\sqrt{n}(\bar{X}_n - \lambda) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \lambda).$$

La delta-méthode donne donc, pour une fonction g dérivable en λ ,

$$\sqrt{n}(g(\bar{X}_n) - g(\lambda)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \lambda[g'(\lambda)]^2).$$

Il suffit de choisir g telle que $\sqrt{\lambda} g'(\lambda)$ soit constante. Par exemple, pour $g(x) = 2\sqrt{x}$,

$$\sqrt{n}(2\sqrt{X_n} - 2\sqrt{\lambda}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Le choix $g(x) = \sqrt{x}$ convient également et conduit à la limite $\mathcal{N}(0, 1/4)$.

4. Supposons désormais $\sqrt{n}(U_n - U) \xrightarrow{\mathcal{L}} V$, $\ell'(U) = 0$ et $\ell''(U) \neq 0$. Un développement de Taylor en U (déterministe d'abord) s'écrit

$$\ell(x) - \ell(U) = \frac{\ell''(U)}{2}(x - U)^2 + o((x - U)^2).$$

Comme $U_n \xrightarrow{\mathbb{P}} U$, on peut l'évaluer en U_n :

$$n[\ell(U_n) - \ell(U)] = \frac{\ell''(U)}{2}(\sqrt{n}(U_n - U))^2 + o_{\mathbb{P}}(1).$$

Le théorème de continuité puis le lemme de Slutsky donnent finalement

$$n[\ell(U_n) - \ell(U)] \xrightarrow{\mathcal{L}} \frac{\ell''(U)}{2}V^2.$$

La bonne renormalisation est donc n . La limite est non dégénérée dès que V^2 ne l'est pas; par exemple, si $V \sim \mathcal{N}(0, \sigma^2)$, elle vaut $(\ell''(U)\sigma^2/2)\chi_1^2$.