Convergence et théorèmes limites

Préliminaires

Notation. Si $u \in \mathbb{R}^d$ on note par $\|u\|_r$ la norme L^r du vecteur u: $\|u\|_r = (\sum_{i=1}^d |u_i|^r)^{1/r}$. Comme toutes les normes sont équivalentes dans $\mathbb{R}^d$ on prendra $r=1$ et on notera $\|u\| = \|u\|_1 = \sum_{i=1}^r |u_i|$. « iid » est abrégé pour « indépendantes et identiquement distribuées ». On notera $X_1, ..., X_n, ...$ ou $(X_n)_{n\geqslant 1}$ une générique suite (infinie) de v.a.

Lemme 1. *(Inégalité de Bienaymé-Tchebychev) Si X est une v.a. réelle telle que $\sigma^2 = \mathrm{Var}(X) < \infty$ et $\mu = \mathbb{E}[X]$ alors pour tout $\varepsilon > 0$ on a que*

$$\mathbb{P}(|X-\mu| > \varepsilon) \leqslant \frac{\mathbb{E}[(X-\mu)^2]}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

Théorème 2. *(Inégalité de Hölder) Soient X et Y deux v.a. réelles définies sur le même espace de proba $(\Omega, \mathcal{A}, \mathbb{P})$. Si $r, s > 1$ sont tels que $r^{-1} + s^{-1} = 1$ et si $\mathbb{E}[|X|^r] < \infty$ et $\mathbb{E}[|Y|^s] < \infty$ alors*

$$\mathbb{E}[|XY|] \leqslant (\mathbb{E}[|X|^r])^{1/r} (\mathbb{E}[|Y|^s])^{1/s}.$$

Corollaire 3. *Soient $p > 0$ et $p > q > 0$. On suppose que $\mathbb{E}[|X|^p] < \infty$ alors $\mathbb{E}[|X|^q] \leqslant (\mathbb{E}[|X|^p])^{p/q}$ et $\mathbb{E}[|X|^q] < \infty$.*

Modes de convergence d'une suite de v.a.

Convergence en loi (ou en distribution)

Théorème 4. *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. à valeurs dans $\mathbb{R}^d$ et X une v.a. à valeurs dans $\mathbb{R}^d$. Les conditions suivantes sont équivalentes (c-à-d chacune d'entre elles implique toutes les autres):*

1. $\lim_{n\to\infty} \phi_{X_n}(t) = \phi_X(t), \quad \forall t \in \mathbb{R}^d$;
2. $\lim_{n\to\infty} F_X(x) = F(x)$ *pour tout* $x \in \mathbb{R}^d$ *point de continuité de* F_X;
3. $\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$ *pour tout fonction* $f: \mathbb{R}^d \to \mathbb{R}$ *continue et bornée.*

Si une de ces conditions est vérifié (et donc toutes) on dit que $(X_n)_{n\geqslant 1}$ converge en loi (ou en distribution) vers X (et l'on note $X_n \xrightarrow{\mathcal{L}} X$).

Rappel. Dans $\mathbb{R}^d$, $F_X(x) = F_X(x_1, ..., x_d) = \mathbb{P}(X_1 \leqslant x_1, ..., X_d \leqslant x_d)$.

Exemple 5. On considère la suite de v.a. $(X_n)_{n\geqslant 1}$ telle que X_n est une v.a. uniforme discrète à valeurs dans $\{1/n, 2/n, 3/n, ..., (n-1)/n, 1\}$.

$$\phi_{X_n}(t) = \mathbb{E}[e^{itX_n}] = \sum_{k=1}^n \frac{1}{n} e^{itk/n} = \frac{e^{it/n}}{n} \sum_{k=0}^{n-1} e^{itk/n} = \frac{e^{it/n}}{n} \frac{e^{it}-1}{e^{it/n}-1}$$

donc

$$\lim_{n\to\infty}\phi_{X_n}(t)=\lim_{n\to\infty}\frac{e^{it/n}}{n}\frac{e^{it}-1}{e^{it/n}-1}=\frac{e^{it}-1}{it}.$$

Si $X\sim\mathcal{U}([0,1])$ alors

$$\phi_X(t)=\int_0^1 e^{itx}\mathrm{d}x=\frac{e^{it}-1}{it}$$

et donc $X_n\xrightarrow{\mathcal{L}}$X.

Exemple 6. Soient $U_1,U_2,\ldots$ des v.a. iid $\mathcal{U}([0,1])$. On pose $X_n=n\min_{1\leqslant k\leqslant n}U_k$. Montrons que $(X_n)_{n\geqslant 1}$ converge en loi vers une v.a. $X\sim\mathcal{E}(1)$. Soit $x\in\mathbb{R}$

$$F_{X_n}(x)=\mathbb{P}(n\min_{1\leqslant k\leqslant n}U_k\leqslant x)=1-\mathbb{P}(n\min_{1\leqslant k\leqslant n}U_k>x)=1-[\mathbb{P}(U_1>x/n)]^n$$

$$=1-[1-\mathbb{P}(U_1\leqslant x/n)]^n=1-[1-F_{U_1}(x/n)]^n$$

et donc

$$F_{X_n}(x)=\begin{cases}1-[1-(x/n)]^n & \text{si } x/n\in[0,1]\\ 0 & \text{si } x/n<0\\ 1 & \text{si } x/n>1\end{cases}$$

Fixons $x>0$ et choisissons n suffisamment grand tel que $x/n\in[0,1]$. Alors

$$\lim_{n\to\infty}F_{X_n}(x)=\lim_{n\to\infty}1-[1-(x/n)]^n=1-e^{-x}.$$

Donc

$$\lim_{n\to\infty}F_{X_n}(x)=\begin{cases}1-e^{-x} & \text{si } x>0\\ 0 & \text{si } x\leqslant 0\end{cases}=F_X(x)\qquad\forall x\in\mathbb{R}.$$

Exemple 7. Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. discrètes telles que $\mathbb{P}(X_n=1/n)=1$. Alors $X_n\xrightarrow{\mathcal{L}}X$ ou X est la v.a. identiquement nulle $\mathbb{P}(X=0)=1$. On voit bien que $F_{X_n}(0)=0$ pour tout n mais que $F_X(0)=1$. Donc en générale on ne pourrait pas avoir convergence de $F_{X_n}(t)$ vers $F_X(t)$ dans tous les points $t\in\mathbb{R}$.

Exemple 8. Reprenons l'exemple 5 de convergence vers la loi uniforme dans [0, 1]. Montrons que $X_n\xrightarrow{\mathcal{L}}X$ en utilisant le critère *(iii)* du théorème 6. Soit f: $\mathbb{R}\to\mathbb{R}$ une fonction continue et bornée

$$\mathbb{E}[f(X_n)]=\frac{1}{n}\sum_{k=1}^n f(k/n)\underset{n\to\infty}{\longrightarrow}\int_0^1 f(x)\mathrm{d}x=\int_{\mathbb{R}}f(x)\mathbb{I}_{0<x<1}\mathrm{d}x=\mathbb{E}[f(X)].$$

Convergence en probabilité

Définition 9. *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. à valeurs dans $\mathbb{R}^d$ et X une v.a. dans $\mathbb{R}^d$ telles que $(X_n)_{n\geqslant 1}$ et X soient définies sur le même espace de probabilité $(\Omega,\ \mathcal{A},\ \mathbb{P})$. On dit que $(X_n)_{n\geqslant 1}$ converge en probabilité vers X et on note $X_n\xrightarrow{\mathbb{P}}X$ si pour tout $\varepsilon>0$*

$$\lim_{n\to\infty}\mathbb{P}(\|X_n-X\|>\varepsilon)=0.$$

Exemple 10. Soit $U \sim \mathcal{U}([0,1])$. On définit $X_n = \mathbb{I}_{U \in [0,1/n]}$. Montrons que $X_n \xrightarrow{\mathbb{P}} 0$. Soit $\varepsilon > 0$ on doit prouver que $\mathbb{P}(|X_n - 0| > \varepsilon) \to 0$. Mais

$$\mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(X_n > \varepsilon) = \mathbb{P}(\mathbb{I}_{U < 1/n} > \varepsilon) = \mathbb{P}(U < 1/n) = 1/n \to 0$$

pour $n \to \infty$.

Loi faible des grandes nombres

Définition 11. *Soit $(X_1, \ldots, X_n)$ un vecteur aléatoire. On définit la moyenne empirique des $(X_i)_{1 \leqslant i \leqslant n}$ par $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$.*

Exemple 12. Soient les X_i des v.a. iid de loi $\mathcal{N}(0,1)$ alors $\overline{X}_n \sim \mathcal{N}(0, 1/n)$. Donc pour tout $\varepsilon > 0$

$$\mathbb{P}(|\overline{X}_n| > \varepsilon) = \mathbb{P}(|Z| > \sqrt{n}\varepsilon) = \mathbb{P}(Z > \sqrt{n}\varepsilon) + \mathbb{P}(Z < -\sqrt{n}\varepsilon) = 2\mathbb{P}(Z < -\sqrt{n}\varepsilon) = 2F_Z(-\sqrt{n}\varepsilon)$$

où $Z \sim \mathcal{N}(0,\, 1)$. Cette quantité est strictement décroissante en n donc converge vers 0 quand $n \to \infty$. Etant donné que $\varepsilon > 0$ est arbitraire cela implique que $\overline{X}_n \xrightarrow{\mathbb{P}} 0$.

Théorème 13. *Soit $(X_n)_{n \geqslant 1}$ une suite iid tel que* $\mathrm{Var}(X_i) = \sigma^2 < +\infty$ *et* $\mu = \mathbb{E}[X_i]$. *On définit $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$ la moyenne empirique des X_j. Alors $\overline{X}_n \xrightarrow{\mathbb{P}} \mu$.*

Démonstration. On a que

$$\mathrm{Var}(\overline{X}_n) = \frac{1}{n^2} \mathrm{Var}(X_1 + \cdots + X_n) = \frac{1}{n^2}[\mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n)] = \frac{\mathrm{Var}(X_1)}{n}.$$

Pour tout $\varepsilon > 0$, par l'inégalité de Tchebychev

$$\mathbb{P}(|\overline{X}_n - \mu| > \varepsilon) \leqslant \frac{\mathrm{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\mathrm{Var}(X_1)}{n\varepsilon^2} \to 0$$

pour $n \to \infty$. Donc $\overline{X}_n \xrightarrow{\mathbb{P}} \mu$. □

Exemple 14. Soient les X_i des v.a. iid de loi $\mathcal{G}(\alpha,\, \beta)$, $\alpha > 0$, $\beta > 0$. $\mathbb{E}[X_i] = \alpha/\beta$ et $\mathrm{Var}(X_i) = \alpha/\beta^2$. Alors $X_1 + \cdots + X_n \sim \mathcal{G}(n\alpha, \beta)$ et $\overline{X}_n \sim \mathcal{G}(n\alpha, n\beta)$ et $\overline{X}_n \xrightarrow{\mathbb{P}} \alpha/\beta$.

Convergence presque sûre

Définition 15. *Soit $(X_n)_{n \geqslant 1}$ une suite de v.a. à valeurs dans $\mathbb{R}^d$ et soit X une v.a. à valeurs dans $\mathbb{R}^d$, telles que X_n et X sont définies sur le même espace de probabilité $(\Omega,\, \mathcal{A},\, \mathbb{P})$. On dit que $(X_n)_{n \geqslant 1}$ converge presque sûrement (ou fortement) vers X et on note $X_n \xrightarrow{p.s.} X$ si $\mathbb{P}(\lim_n X_n = X) = 1$. Autrement dit si l'événement $A = \{\omega \in \Omega \colon \lim_n X_n(\omega) = X(\omega)\}$ est tel que $\mathbb{P}(A) = 1$.*

Exemple 16. Soit $X_n \sim \mathcal{B}\mathrm{in}(n,\, p)$. Montrons que X_n / n^2 converge presque sûrement vers 0. En effet l'ensemble $A = \{0 \leqslant |X_n| \leqslant n \text{ pour tout } n\}$ est tel que $\mathbb{P}(A) = 1$. Donc pour $\omega \in A$ on a que $0 \leqslant |X_n(\omega)/n^2| \leqslant 1/n$ ce qu'implique que $\lim_n X_n(\omega)/n^2 = 0$ pour tout $\omega \in A$ et donc que

$$\mathbb{P}(\lim_n X_n / n^2 = 0) \geqslant \mathbb{P}(A) = 1$$

qui montre la convergence presque sure.

Théorème 17. *(Loi forte des grandes nombres) Soit* $(X_n)_{n\geqslant 1}$ *une suite iid telle que les* X_i *soient intégrables (c-à-d* $\mathbb{E}[|X_i|]<\infty$*). Alors*

$$\overline{X_n}\xrightarrow{p.s.}\mathbb{E}[X_1].$$

Exemple 18. Soient X_1, X_2, ... des v.a. iid $\mathcal{E}(\lambda)$ $(\lambda>0)$. X_1 est intégrable ($\mathbb{E}[|X_1|]=1/\lambda$). Alors

$$\overline{X_n}\xrightarrow{p.s.}1/\lambda.$$

Convergence dans L^r (convergence en moyenne d'ordre r)

Définition 19. *Soit* $(X_n)_{n\geqslant 1}$ *une suite de v.a. à valeurs dans* $\mathbb{R}^d$ *et* X *une v.a. à valeurs dans* $\mathbb{R}^d$*. On suppose que* $\mathbb{E}(\|X_n\|^r)<+\infty$ *pour tout* $n\geqslant 1$ *et que* $\mathbb{E}[\|X\|^r]<+\infty$*. On dit que* $(X_n)_{n\geqslant 1}$ *converge vers* X *dans* L^r *(ou en moyenne d'ordre* r*), et on note* $X_n\xrightarrow{L^r}X$ *(ou* $X_n\xrightarrow{r}X$*) si*

$$\lim_{n\to\infty}\mathbb{E}[\|X_n-X\|^r]=0.$$

En particulier: si $d=1$, $(X_n)_{n\geqslant 1}$ *et* X *sont des v.a. réelles alors* $X_n\xrightarrow{L^r}X$ *si* $\mathbb{E}[|X|^r]<\infty$, $\mathbb{E}[|X_n|^r]<\infty$ *et* $\mathbb{E}[|X_n-X|^r]\to 0$.

Exemple 20. Soit $r>0$. Soit $U\sim\mathcal{U}([0,1])$. On considère $(X_n)_{n\geqslant 1}$ telle que

$$X_n=n\mathbb{I}_{[0,1/n]}(U)$$

Quelle est la condition sur r pour que $X_n\xrightarrow{L^r}0$?

$$\mathbb{E}[|X_n-0|^r]=\mathbb{E}[|X_n|^r]=\mathbb{E}[X_n^r]=\mathbb{E}[n^r\,\mathbb{I}_{[0,1/n]}(U)]=n^r\mathbb{P}(U\leqslant 1/n)=n^r/n$$

et n^{r-1} converge vers 0 ssi $r<1$.

Quelque propriétés de la convergence L^r.

Proposition 21. *Soit* $r>0$ *et* $0<s<r$*. Alors* $X_n\xrightarrow{r}X\Longrightarrow X_n\xrightarrow{s}X$.

Démonstration. Par l'inégalité de Holder on a $\mathbb{E}[|X_n-X|^s]\leqslant(\mathbb{E}[|X_n-X|^r])^{s/r}$. Donc si $X_n\xrightarrow{r}X$ alors $\mathbb{E}[|X_n-X|^r]\to 0$ et $\mathbb{E}[|X_n-X|^s]\to 0$. □

Proposition 22. *Si* $X_n\xrightarrow{1}X$ *alors* $\mathbb{E}[X_n]\to\mathbb{E}[X]$.

Démonstration. Par hypothèse on a que $\mathbb{E}[|X_n|]<\infty$ $\mathbb{E}[|X|]<\infty$ et $\mathbb{E}[|X_n-X|]\to 0$. Donc

$$|\mathbb{E}[X]-\mathbb{E}[X_n]|=|\mathbb{E}[X-X_n]|\leqslant\mathbb{E}[|X-X_n|]\to 0.$$

car $-|X_n-X|\leqslant X_n-X\leqslant|X_n-X|$. □

Proposition 23. $X_n \xrightarrow{2} a \in \mathbb{R}$ *(on di que X_n converge à la constante a en moyenne quadratique) ssi* $\mathbb{E}[X_n] \to a$ *et* $\operatorname{Var}(X_n) \to 0$.

Démonstration. Si $X_n \xrightarrow{2} a \in \mathbb{R}$ alors $\mathbb{E}[|X_n - a|^2] \to 0$. Soit $\mu_n = \mathbb{E}[X_n]$

$$\mathbb{E}[|X_n - a|^2] = \mathbb{E}[|X_n - \mu_n + \mu_n - a|^2] = \mathbb{E}[(X_n - \mu_n)^2] + 2\mathbb{E}[(X_n - \mu_n)](\mu_n - a) + (\mu_n - a)^2$$

$$= \operatorname{Var}(X_n) + (\mu_n - a)^2$$

et donc $\operatorname{Var}(X_n) + (\mu_n - a)^2 \to 0$ ce qui entraîne que $\operatorname{Var}(X_n) \to 0$ et que $\mu_n \to a$. Réciproquement si $\operatorname{Var}(X_n) \to 0$ et $\mu_n \to a$ alors $\mathbb{E}[|X_n - a|^2] = \operatorname{Var}(X_n) + (\mu_n - a)^2 \to 0$. □

Théorème 24. *(de Slusky) Soient $(X_n)_{n\geqslant 1}$, $(A_n)_{n\geqslant 1}$ et $(B_n)_{n\geqslant 1}$ trois suites de v.a.. Soient X une v.a. et $a, b \in \mathbb{R}$. Si $X_n \xrightarrow{\mathcal{L}} X$, $A_n \xrightarrow{\mathbb{P}} a$ et $B_n \xrightarrow{\mathbb{P}} b$ alors*

$$A_n X_n + B_n \xrightarrow{\mathcal{L}} aX + b$$

Théorème 25. *(de continuité) Soit $g: \mathbb{R}^d \to \mathbb{R}^k$ telle que g est continue*

i. $X_n \xrightarrow{\mathcal{L}} X \Longrightarrow g(X_n) \xrightarrow{\mathcal{L}} g(X)$

ii. $X_n \xrightarrow{\mathbb{P}} X \Longrightarrow g(X_n) \xrightarrow{\mathbb{P}} g(X)$

iii. $X_n \xrightarrow{p.s.} X \Longrightarrow g(X_n) \xrightarrow{p.s.} g(X)$

Liens entre les modes de convergence

Proposition 26.

i. La convergence presque sûre entraîne la convergence en probabilité:

$$X_n \xrightarrow{p.s.} X \Longrightarrow X_n \xrightarrow{\mathbb{P}} X$$

ii. La convergence en probabilité entraîne la convergence en loi

$$X_n \xrightarrow{\mathbb{P}} X \Longrightarrow X_n \xrightarrow{\mathcal{L}} X$$

iii. $X_n \xrightarrow{\mathcal{L}} X \Longleftrightarrow X_n \xrightarrow{\mathbb{P}} X$ *si* $X = c \in \mathbb{R}$

iv. La convergence dans L^r entraîne la convergence en probabilité

$$X_n \xrightarrow{r} X \Longrightarrow X_n \xrightarrow{\mathbb{P}} X.$$

Exemple 27. $X_n \sim \mathcal{B}\text{in}(n, p)$

$$\frac{X_n}{n} \xrightarrow{p.s.} 0 \Longrightarrow \frac{X_n}{n} \xrightarrow{\mathbb{P}} 0 \Longleftrightarrow \frac{X_n}{n} \xrightarrow{\mathcal{L}} 0$$

Le théorème central limite (TCL)

Théorème 28. *Soit* $(X_n)_{n\geqslant 1}$ *une suite iid tel que* $\mathrm{Var}(X_i)=\sigma^2<\infty$. *Soit* $\mu=\mathbb{E}[X_1]$. *Alors*

$$\sqrt{n}\frac{\bar{X}_n-\mu}{\sigma}\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)$$

Exemple 29. Soient $X_1, X_2, \ldots$ des v.a. iid $\sim\mathcal{E}(\lambda)$. $\mathrm{Var}(X_1)=1/\lambda^2$ et $\mu=\mathbb{E}[X_1]=1/\lambda$. Par le TCL on a

$$\sqrt{n}\frac{\bar{X}_n-1/\lambda}{1/\lambda}\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)\qquad \text{ou}\qquad \sqrt{n}(\overline{X}_n-1/\lambda)\xrightarrow{\mathcal{L}}\mathcal{N}(0,1/\lambda^2).$$

Théorème 30. *Soit* $(X_n)_{n\geqslant 1}$ *une suite iid à valeurs dans* $\mathbb{R}^d$ *tels que la matrice de covariance* Σ *de* X_1 *est finie (c-à-d si* $\Sigma_{ii}<\infty$ *pour* $i=1,\ldots,d$*) alors*

$$\sqrt{n}(\overline{X}_n-\mathbb{E}[X_1])\xrightarrow{\mathcal{L}}\mathcal{N}_d(0,\Sigma).$$

Théorème 31. *(La* δ*-méthode, cas unidimensionnel) Soit* $(Y_n)_{n\geqslant 1}$ *une suite de v.a. réelles. On suppose que* $\sqrt{n}(Y_n-\mu)\xrightarrow{\mathcal{L}}\mathcal{N}(0,\sigma^2)$. *Si* $g\colon\mathbb{R}\to\mathbb{R}$ *est une fonction continûment dérivable au point* μ *(c-à-d* g *est* C^1 *dans un voisinage du point* μ*) alors*

$$\sqrt{n}(g(Y_n)-g(\mu))\xrightarrow{\mathcal{L}}\mathcal{N}(0,(g'(\mu))^2\sigma^2).$$

Exemple 32. Soit $(X_n)_{n\geqslant 1}$ une suite iid $\sim\mathcal{E}(\lambda)$. Soit $Y_n=\overline{X}_n$. Par le TCL on a que $\sqrt{n}(\overline{X}_n-1/\lambda)\xrightarrow{\mathcal{L}}\mathcal{N}(0,1/\lambda^2)$. Soit $g(x)=1/x$. $g'(x)=-1/x^2$ et $g'(1/\lambda)=-\lambda^2$. Donc $(g'(1/\lambda))^2=\lambda^4$ et g est continûment dérivable au point $1/\lambda$. Par la δ-méthode on a que

$$\sqrt{n}\left(\frac{1}{\bar{X}_n}-\lambda\right)\xrightarrow{\mathcal{L}}\mathcal{N}(0,\lambda^2)$$

Exemple 33. (Normalisation de la variance) Soit $(X_n)_{n\geqslant 1}$ une suite iid $\sim$ Bernoulli(p) (avec $p\in]0,1[$), $\sigma^2=\mathrm{Var}(X_1)=p(1-p)$. Par le TCL $\sqrt{n}(\overline{X}_n-p)\xrightarrow{\mathcal{L}}\mathcal{N}(0,p(1-p))$. Peut on trouver une application g: $]0,1[\ \to\mathbb{R}$ (qui ne dépend pas de p) telle que $\sqrt{n}(g(\overline{X}_n)-g(p))\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)$?

Supposons que une telle application existe et qu'elle soit continûment dérivable au point p. Par la δ-méthode on doit avoir que $g'(p)^2p(1-p)=1\Longrightarrow g'(p)^2=1/(p(1-p))$ pour tout $p\in]0,1[$. Une solution possible est

$$g'(p)=\frac{1}{\sqrt{p(1-p)}}\Longrightarrow g(p)=2\arcsin(\sqrt{p})$$

donc on a que

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