Chapitre 14 :

CONVERGENCES ET APPROXIMATIONS EN PROBABILITÉS

Dans ce chapitre, on s'intéresse à la notion de convergence d'une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires réelles, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On dégage dans les deux premiers paragraphes deux notions différentes de convergence avec à la clé dans les deux cas un théorème fondamental de la théorie des probabilités. Dans le troisième paragraphe, on applique ces théorèmes à des problèmes d'approximation.

Dans tout le chapitre, $(\Omega, \mathcal{A}, \mathbb{P})$ désigne un espace de probabilité.

1. Convergence en probabilité

1.1 Définition

Définition 1.1 *Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$ et X une variable aléatoire sur le même espace.*
*On dit que $(X_n)_{n\in\mathbb{N}}$ **converge en probabilité** vers X, et l'on note $X_n \xrightarrow[n\to\infty]{\mathcal{P}} X$, si :*

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

Exemple 1.2 Vérifier que pour tout $n \in \mathbb{N}^*$, la fonction

$$f_n : x \in \mathbb{R} \longmapsto \frac{n}{\pi(1+n^2x^2)}$$

définit une densité de probabilité ; on considère une variable aléatoire X_n à densité f_n.
Étudier la convergence en probabilité de la suite $(X_n)_{n\in\mathbb{N}^*}$.

1.2 Inégalité de Bienaymé-Tchebychev

Théorème 1.3 (Inégalité de Markov) *Soit X une variable aléatoire réelle, discrète ou à densité.*
Si X est positive et admet une espérance, alors :

$$\forall \varepsilon > 0, \quad \mathbb{P}(X \geqslant \varepsilon) \leqslant \frac{\mathbb{E}(X)}{\varepsilon}.$$

Corollaire 1.4 *Soit X une variable aléatoire réelle, discrète ou à densité.*
Si X admet un moment d'ordre 2, alors :

$$\forall \varepsilon > 0, \quad \mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}(X^2)}{\varepsilon^2}.$$

Théorème 1.5 (Inégalité de Bienaymé-Tchebychev) *Soit X une variable aléatoire réelle, discrète ou à densité.*
Si X admet une variance, alors :

$$\forall \varepsilon > 0, \quad \mathbb{P}\big(|X - \mathbb{E}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Exemple 1.6 On considère un dé cubique équilibré. En utilisant l'inégalité de Bienaymé-Tchebychev, déterminer le nombre de lancers qu'il suffit d'effectuer pour pouvoir affirmer avec un risque d'erreur inférieur à 5% que la fréquence d'apparition de l'as est voisine de $\frac{1}{6}$ à la précision 0,01.

1.3 Loi faible des grands nombres

Théorème 1.7 (Loi faible des grands nombres) *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$, discrètes ou à densité, mutuellement indépendantes, admettant une même espérance m et une même variance σ^2.*

On a :

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{\mathcal{P}} m.$$

Plus précisément,

$$\forall \varepsilon > 0, \quad \forall n \in \mathbb{N}^*, \quad \mathbb{P}\big(|\overline{X}_n - m| \geqslant \varepsilon\big) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$$

Remarques 1.8 • Dans l'énoncé précédent, on peut affaiblir l'hypothèse X_n, $n \in \mathbb{N}^*$, mutuellement indépendantes en X_n, $n \in \mathbb{N}^*$, deux-à-deux indépendantes ou même deux-à-deux non corrélées.

• Le théorème précédent porte le nom de loi *faible* des grands nombres car il existe un résultat bien plus précis : on peut montrer (mais cela est bien plus difficile !) que si les variables aléatoires X_n (toujours supposées indépendantes) suivent une même loi et admettent une espérance m, alors $\overline{X}_n$ converge presque sûrement vers la variable certaine égale à m.

Corollaire 1.9 (Théorème d'or de Bernoulli) *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes suivant toutes une loi de Bernoulli de paramètre p.*
La fréquence de succès $\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$ converge en probabilité vers la variable certaine égale à p lorsque $n \to \infty$. Plus précisément,

$$\forall \varepsilon > 0, \quad \forall n \in \mathbb{N}^*, \quad \mathbb{P}\big(|\overline{X}_n - p| \geqslant \varepsilon\big) \leqslant \frac{1}{4n\varepsilon^2}.$$

Remarques 1.10 • On considère une succession d'expériences identiques indépendantes, au cours de chacune desquelles un événement A est susceptible de se produire avec probabilité $p = \mathbb{P}(A)$. Pour tout $n \geqslant 1$, on note X_n la variable aléatoire égale à 1 si l'événement A se réalise au cours de la n-ième expérience et 0 sinon. Les variables aléatoires X_n, $n \geqslant 1$, sont indépendantes et suivent toutes la loi de Bernoulli de paramètre p. Dans ces conditions, la loi des grands nombres affirme que la fréquence $\overline{X}_n$ de réalisation de l'événement A au cours des n premières expériences converge en probabilité vers $p = \mathbb{P}(A)$ (la convergence est même presque sûre d'après la loi forte).
Ainsi *la fréquence statistique de réalisation d'un événement tend vers la probabilité de cet événement.*

• On notera que le majorant dans la dernière inégalité ne dépend pas de p.

1.4 Théorèmes opératoires

Remarque 1.11 Bien que ce ne soit pas au programme, on peut démontrer (cf. TD) que si (X_n) et (Y_n) sont des suites de variables aléatoires convergeant en probabilité respectivement vers X et Y alors, pour tout $\lambda \in \mathbb{R}$, $(\lambda X_n + Y_n)$ converge en probabilité vers $\lambda X + Y$.

Le seul résultat au programme est le suivant, que l'on admet.

Proposition 1.12 *Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles, X une variable aléatoire réelle et $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue.*
Si (X_n) converge en probabilité vers X, alors $\big(f(X_n)\big)$ converge en probabilité vers $f(X)$.

Remarque 1.13 Lorsque la variable $X = c$ est constante, il suffit que la fonction f soit continue au point c pour que la suite $\big(f(X_n)\big)$ converge en probabilité vers $f(X) = f(c)$.

2. Convergence en loi

2.1 Définition

Définition 2.1 *Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$ et X une variable aléatoire sur le même espace. On note F_{X_n} (resp. F_X) la fonction de répartition de la variable aléatoire de X_n, $n \in \mathbb{N}$ (resp. de X).*

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ ***converge en loi*** *vers* X, *et l'on note* $X_n \xrightarrow[n\to\infty]{\mathcal{L}} X$, *si en tout point* $x \in \mathbb{R}$ *de continuité de* F_X, *on a :*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x).$$

Exemple 2.2 (rôle de la restriction sur x) La suite de terme général $X_n = \frac{1}{n}$, $n \geqslant 1$, converge en loi vers $X = 0$.
Plus généralement (exercice), si $(x_n)_{n\in\mathbb{N}}$ est une suite réelle convergeant vers ℓ, alors la suite de terme général $X_n = x_n$, $n \in \mathbb{N}$, converge en loi vers $X = \ell$.

Proposition 2.3 *Lorsque les variables aléatoires* X_n, $n \in \mathbb{N}$, *et* X *sont à valeurs dans* $\mathbb{Z}$, *la convergence en loi de* $(X_n)_{n\in\mathbb{N}}$ *vers* X *équivaut à :*

$$\forall k \in \mathbb{Z}, \quad \lim_{n\to\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k).$$

Exemple 2.4 Pour la plupart des chaînes de Markov étudiées en informatique, on a observé et démontré la convergence en loi de la chaîne (X_n) vers sa loi stationnaire Π.

Exemple 2.5 Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires où, pour tout $n \in \mathbb{N}^*$, X_n suit une loi uniforme sur l'ensemble $\{k/n\}_{1\leqslant k\leqslant n}$. Étudier la convergence en loi de la suite $(X_n)_{n\in\mathbb{N}^*}$.

Remarque 2.6 La convergence en loi d'une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires vers une variable aléatoire X ne fait intervenir que *la loi* des variables aléatoires. Il n'y a donc pas unicité de la limite.

Le théorème suivant n'est pas explicitement au programme.

Théorème 2.7 *Soient* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires réelles et* X *une variable aléatoire réelle.*
Si $(X_n)_{n\in\mathbb{N}}$ *converge en probabilité vers* X, *alors* $(X_n)_{n\in\mathbb{N}}$ *converge en loi vers* X. *La réciproque est fausse.*

2.2 Théorème limite central

Théorème 2.8 *Soit* $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires réelles sur* $(\Omega, \mathcal{A}, \mathbb{P})$, *mutuellement indépendantes et de même loi, admettant une espérance m et une variance* σ^2 $(\sigma > 0)$.
La variable aléatoire centrée réduite associée à $\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$:

$$\overline{X}_n^* = \frac{\sqrt{n}}{\sigma}(\overline{X}_n - m)$$

converge en loi lorsque $n \to \infty$ *vers (une variable aléatoire suivant) la loi normale* $\mathcal{N}(0, 1)$.
En d'autres termes, pour tous $a < b$ *(éventuellement infinis),*

$$\lim_{n\to\infty} \mathbb{P}(a < \overline{X}_n^* \leqslant b) = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-t^2/2}\,\mathrm{d}t.$$

Remarque 2.9 La conclusion peut également être énoncée sur $S_n = X_1 + \cdots + X_n = n\overline{X}_n$:

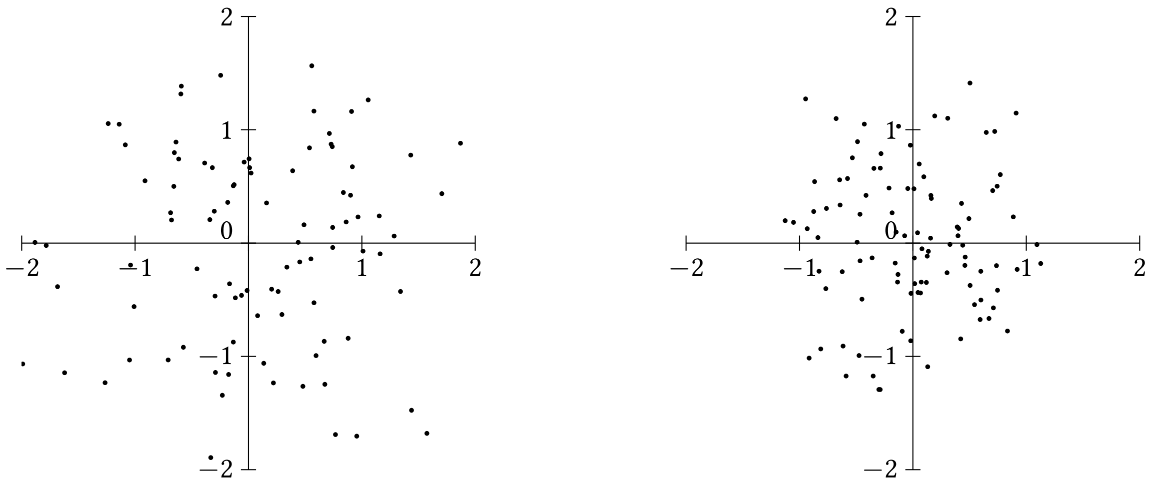
$$\overline{X}_n^* = S_n^* = \frac{\overline{X}_n - nm}{\sigma\sqrt{n}}$$

converge en loi lorsque $n \to \infty$ vers (une variable aléatoire suivant) la loi normale $\mathcal{N}(0, 1)$.

Remarque 2.10 Outre les approximations qui en découlent (cf. paragraphe suivant) et les applications en statistiques (cf. chapitre suivant), le théorème limite central a plusieurs interprétations théoriques.

- Tout d'abord, il précise les fluctuations de $\overline{X}_n$ autour de sa limite m (la loi faible des grands nombres énonce une convergence en probabilité, la loi forte donnerait une convergence *presque sûre*) : $\overline{X}_n - m$ suit approximativement une loi normale $\mathcal{N}(0, \sigma^2/n)$ lorsque n est assez grand.

- Par ailleurs, le théorème exprime le fait que la somme d'un grand nombre de variables indépendantes, de même loi et admettant une variance a une distribution à peu près gaussienne. La *loi des erreurs* généralise ce fait à la somme de variables indépendantes, admettant une variance, et dont aucune n'est prépondérante. Cela met en évidence l'universalité de la loi normale, et explique pourquoi de nombreuses variables aléatoires intervenant en modélisation sont souvent supposées suivre des lois normales.
  On peut illustrer ce fait par l'exemple du tireur à la carabine. La figure ci-dessous représente $n = 100$ réalisations de couples (X_k, Y_k), $1 \leqslant k \leqslant n$, de variables gaussiennes centrées indépendantes, de variances $\sigma^2 = 1$ puis $\sigma^2 = 0{,}4$. Elle correspond bien aux résultats obtenus à un tir à la carabine sur une cible, ce qu'on peut expliquer ainsi : si (X_k, Y_k) sont les coordonnées du k-ième tir, X_k et Y_k sont des sommes de petites variables indépendantes (erreur de visée, tremblement, défaut de concentration, recul, perturbation par un élément extérieur comme un contrejour ou un cri, …). D'après la loi des erreurs, il est donc raisonnable de supposer que (X_k, Y_k) est un couple de variables gaussiennes. L'écart-type de ces gaussiennes mesure l'adresse du tireur : plus il est petit et plus le tireur est adroit.



2.3 Théorèmes opératoires

Remarque 2.11 Attention, on ne dispose pas des théorèmes opératoires classiques pour la convergence en loi. Par exemple, la convergence en loi de (X_n) vers X et de (Y_n) vers Y n'entraîne pas la convergence en loi de $(X_n + Y_n)$ vers $X + Y$.

On admet les deux résultats suivants.

PROPOSITION 2.12 (SLUTSKY) *Soient (X_n), (Y_n) deux suites de variables aléatoires, X une variable aléatoire et c un réel.*
Si (X_n) converge en loi vers X et si (Y_n) converge en probabilité vers une constante c, alors $(X_n + Y_n)$ converge en loi vers $X + c$ et (X_nY_n) converge en loi vers cX.

PROPOSITION 2.13 *Soient (X_n) une suite de variables aléatoires, X une variable aléatoire et $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue.*
Si (X_n) converge en loi vers X, alors $\big(f(X_n)\big)$ converge en loi vers $f(X)$.

3. Approximations

3.1 Approximations de la loi binomiale

On est souvent amené à travailler sur des lois binomiales. Il en existe des tables numériques, usuellement pour $n \leqslant 50$ et $0 \leqslant p \leqslant 0{,}5$. Pour les valeurs de n supérieures, on envisage deux méthodes.

Si n est grand alors que p est petit (cas des événements rares), on utilise l'approximation par la loi de Poisson décrite ci-dessous.

Théorème 3.1 *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires où, pour tout $n \in \mathbb{N}^*$, X_n suit une loi binomiale $\mathcal{B}(n, p_n)$.*
Si np_n admet une limite finie $\lambda > 0$ lorsque $n \to \infty$ (en particulier lorsque $p_n = \lambda/n$), alors

$$\forall k \in \mathbb{N}, \quad \lim_{n\to\infty} \mathbb{P}(X_n = k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

et la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers (une variable alétoire suivant) la loi de Poisson $\mathcal{P}(\lambda)$.

Remarque 3.2 En pratique, on considère que l'approximation d'une loi binomiale $\mathcal{B}(n, p)$ par une loi de Poisson $\mathcal{P}(np)$ est satisfaisante si $n \geqslant 30$, $p \leqslant 0{,}1$ et $np \leqslant 15$.

Exemple 3.3 Soit X une variable aléatoire de loi binomiale $\mathcal{B}(100; 0{,}05)$. Comparer la valeur approchée de $\mathbb{P}(X = 2)$ obtenue par approximation par une loi de Poisson avec la valeur exacte.

Si n est grand alors que p reste fixe, on utilise l'approximation normale décrite ci-dessous.

Théorème 3.4 *Soient $p \in]0, 1[$, $q = 1 - p$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires où, pour tout $n \in \mathbb{N}^*$, X_n suit une loi binomiale $\mathcal{B}(n, p)$.*
La suite des variables aléatoires centrées réduites de terme général

$$X_n^* = \frac{X_n - np}{\sqrt{npq}}, \quad n \in \mathbb{N},$$

converge en loi vers (une variable aléatoire suivant) la loi normale centrée réduite $\mathcal{N}(0, 1)$.

Remarque 3.5 Soit $q = 1 - p$. En pratique, on considère que l'approximation d'une loi binomiale $\mathcal{B}(n, p)$ par une loi normale $\mathcal{N}(np, npq)$ est satisfaisante si $n \geqslant 30$, $np \geqslant 5$ et $nq \geqslant 5$.

Exemple 3.6 Reprendre l'exemple 1.6 en utilisant une approximation gaussienne de la loi binomiale.

Remarque 3.7 Le résultat précédent conduit à approcher une loi discrète par une loi continue. On effectue donc souvent une *correction de continuité* en approchant $\mathbb{P}(B = k)$ par $\mathbb{P}(k - 0{,}5 < N < k + 0{,}5)$ où B suit une loi binomiale et N suit la loi normale appropriée.

3.2 Approximation d'une loi de Poisson par une loi normale

Théorème 3.8 *Soient $\alpha > 0$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires où, pour tout $n \in \mathbb{N}^*$, X_n suit une loi de Poisson $\mathcal{P}(n\alpha)$.*
La suite des variables aléatoires centrées réduites de terme général

$$X_n^* = \frac{X_n - n\alpha}{\sqrt{n\alpha}}, \quad n \in \mathbb{N},$$

converge en loi vers (une variable aléatoire suivant) la loi normale centrée réduite $\mathcal{N}(0, 1)$.

Remarques 3.9
- En pratique, on considère que l'approximation d'une loi de Poisson $\mathcal{P}(\lambda)$ par une loi normale $\mathcal{N}(\lambda, \lambda)$ est satisfaisante dès que $\lambda \geqslant 18$.
- Il s'agit ici encore d'approcher une loi discrète par une loi continue, et il pourra donc être intéressant d'effectuer une correction de continuité.

Exemple 3.10 Soit X une variable aléatoire suivant la loi de Poisson $\mathcal{P}(16)$. Comparer les valeurs approchées de $\mathbb{P}(X \leqslant 21)$ et $\mathbb{P}(X = 16)$ obtenues par approximation gaussienne avec correction de continuité aux valeurs exactes.