P. Maurer
ENS Rennes

Leçon 266. Illustration de la notion d'indépendance en probabilités.

Devs :

- Marche aléatoire sur $\mathbb{Z}$, $\mathbb{Z}^2$ et $\mathbb{Z}^3$
- Théorème Central Limite

Références :

- Barbe-Ledoux, Probabilités
- Foata-Fuchs, Calcul des probabilités
- Cadre-Vial, Statistique mathématique
- Garet, Probabilités et processus stochastiques

Le Barbe-Ledoux fait quasiment la leçon à lui tout seul, mais il est trop théorique et manque un peu d'exemples "concrets".

1 Notion d'indépendance

On se donne un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

1.1 Définitions et propriétés

Définition 1. *On dit que deux évènements A et B sont indépendants si*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B).$$

Exemple 2. On jette deux dés, un bleu et un rouge. Les évènements $A = \{$On obtient un nombre ≤ 4 avec le dé rouge$\}$ et $B = \{$On obtient un 6 avec le dé bleu$\}$ sont intuitivement indépendants.

En prenant l'univers $\Omega = \{(i,j) : 1 \leq i, j \leq 6\}$ muni de sa tribu $\mathcal{P}(\Omega)$ et de la probabilité uniforme $\mathbb{P}$ sur Ω, on a $\mathbb{P}(A \cap B) = \frac{1}{9} = \frac{2}{3} \times \frac{1}{6} = \mathbb{P}(A)\,\mathbb{P}(B)$.

Définition 3.

Une famille quelconque d'évènements $(A_i)_{i \in I}$ est dite mutuellement indépendante si pour tout $J \subset I$ fini, on a

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

Une famille de sous-tribus quelconque $(\mathcal{A}_i)_{i \in I}$ de $\mathcal{A}$ est dite mutuellement indépendante si toute famille d'évènements $A_i \in \mathcal{A}_i$, $i \in I$ est mutuellement indépendante.

Exemple 4. Soit $\Omega = [0,1]$ muni de sa tribu borélienne et $\mathbb{P}$ la mesure de Lebesgue sur $[0,1]$. On pose, pour $n \geq 1$,

$$A_n := \bigcup_{1 \leq k \leq 2^{n-1}} \left]\frac{2(k-1)}{2^n}, \frac{2k-1}{2^n}\right].$$

Alors la famille $(A_n)_{n \geq 1}$ et mutuellement indépendante.

Exemple 5. On reprend l'exemple 3, et on considère les évènements

$$A = \{\text{Le résultat du dé rouge est impair}\}$$

$$B = \{\text{Le résultat du dé bleu est impair}\}$$

$$C = \{\text{La somme des deux dés est impaire}\}$$

Alors A, B et C sont indépendants deux à deux mais pas mutuellement indépendants.

Proposition 6. *Soit $\mathcal{C}_1$ et $\mathcal{C}_2$ deux familles indépendantes dans l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Alors les tribus $\sigma(\mathcal{C}_1)$ et $\sigma(\mathcal{C}_2)$ sont indépendantes.*

Définition 7. *Une famille quelconque de variables aléatoires $(X_i)_{i \in I}$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans $(E, \mathcal{B})$ est mutuellement indépendante si la famille des tribus engendrées par les X_i l'est.*

Autrement dit, la famille $(X_i)_{i \in I}$ est indépendante si pour tout $J \subset I$ fini et tout ensemble $B_j \in \mathcal{B}$ mesurable, on a

$$\mathbb{P}\left(\bigcap_{j \in J} \{X_j \in B_j\}\right) = \prod_{j \in J} \mathbb{P}(\{X_j \in B_j\}).$$

Exemple 8. On reprend les notations de l'exemple 4. La suite de variables aléatoires

$$X_n = \mathbf{1}_{A_n} = \sum_{1 \leq k \leq 2^n} \mathbf{1}_{\left]\frac{2(k-1)}{2^n}, \frac{2k-1}{2^n}\right]} \quad , \; n \geq 1$$

de $[0,1]$ dans $\{0,1\}$ est indépendante. On vérifie que pour tout $n \in \mathbb{N}$, $X_n \sim \mathcal{B}(1/2)$.

1.2 Loi d'un n-uplet de variables aléatoires et indépendance

Théorème 9. *Soit $d \geq 1$ un entier et $(X_1, \ldots, X_d)$ une famille finie de variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$. La loi $P^{(X_1, \ldots, X_d)}$ du vecteur aléatoire sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ est égale au produit des lois marginales $P^{X_1} \otimes \cdots \otimes P^{X_d}$.*

Exemple 10. Soit (X,Y) un couple de variables aléatoires réelles, de loi absolument continue de densité $x \mapsto f(x)\,g(x)$ par rapport à la mesure de Lebesgue sur $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.

On suppose de plus que $f \in L^1(\mathbb{R})$ et que $\int_{\mathbb{R}} f(x)\,dx = 1$.

Alors X et Y sont indépendantes, absolument continues par rapport à la mesure de Lebesgue sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et de densité respective f et g.

Proposition 11. *Une famille $(X_1,\ldots,X_d)$ de variables aléatoires réelles sur $(\Omega,\mathcal{F},\mathbb{P})$ est indépendante si et seulement si pour tout $(x_1,\ldots,x_d)\in\mathbb{R}^d$, on a*

$$\varphi_{(X_1,\ldots,X_d)}(x_1,\ldots,x_d)=\varphi_{X_1}(x_1)\times\cdots\times\varphi_{X_d}(x_d),$$

où $\varphi_Y:=\mathbb{E}[e^{\langle Y\cdot\rangle}]$ désigne la fonction caractéristique d'une variable aléatoire Y à valeurs dans $\mathbb{R}^k$.

On donne quelques exemples avec des lois classiques.

Définition 12. *Soit $n\in\mathbb{N}$ et $p\in[0,1]$. Une variable aléatoire X sur Ω à valeurs dans $\mathbb{N}$ est dite de loi binomiale $\mathcal{B}(n,p)$ si elle est absolument continue par rapport à la mesure de comptage sur $(\mathbb{N},\mathcal{P}(\mathbb{N}))$ de densité*

$$f(k)=\binom{n}{k}p^k(1-p)^{n-k}.$$

Dans le cas où $n=1$, on dit que X suit une loi de Bernouilli $\mathcal{B}(p)$ de paramètre p.

Proposition 13. *Soit $(X_1,\ldots,X_n)$ un n-uplet de variables aléatoires discrètes de loi $\mathcal{B}(p)$. Alors $X_1+\cdots+X_n\sim\mathcal{B}(n,p)$.*

Définition 14. *Soit $\lambda,p>0$. Une variable aléatoire X sur Ω à valeurs dans $[0,+\infty[$ est dite de loi $\Gamma(p,\lambda)$ si elle est absolument continue par rapport à la mesure de Lebesgue et admet pour densité*

$$f(x)=\frac{\lambda}{\Gamma(p)}e^{-\lambda x}(\lambda x)^{p-1}\mathbf{1}_{\mathbb{R}_+^*}(x).$$

Proposition 15. *Soit (X_1,X_2) est un couple de variables aléatoires indépendantes de lois marginales respectives $\Gamma(p_1,\lambda_1)$ et $\Gamma(p_2,\lambda_2)$, et $\mu>0$. Alors*

- *X_1+X_2 a pour loi $\Gamma(p_1+p_2,\lambda)$.*
- *μX_1 a pour loi $\Gamma\left(p_1,\frac{\lambda}{\mu}\right)$.*

Proposition 16. *Soit $(X_1,\ldots,X_n)$ des variables aléatoires indépendantes, de loi exponentielle de paramètre $\lambda>0$. Alors $X_1+\cdots+X_n$ a pour loi $\Gamma(n,\lambda)$.*

Définition 17. *Une variable aléatoire sur Ω à valeurs dans $[0,1]$ est dite de loi $B(r,s)$ (loi bêta de paramètres $r>0$ et $s>0$) si elle est absolument continue par rapport à la mesure de Lebesgue et admet pour densité :*

$$f(x)=\frac{1}{B(r,s)}x^{r-1}(1-x)^{s-1}\mathbf{1}_{[0,1]}(x).$$

En particulier, pour $r=1$ et $s=1$, on retrouve la loi uniforme sur $[0,1]$.

Proposition 18. *Soit $(X_1,\ldots,X_n)$ un système de n variables aléatoires indépendantes, toutes de loi uniforme sur $[0,1]$. Alors les variables aléatoires $Y=\min(X_1,\ldots,X_n)$ et $Z=\max(X_1,\ldots,X_n)$ sont de loi respective $B(1,n)$ et $B(n,1)$.*

Définition 19. *Une variable aléatoire X sur Ω à valeurs dans $\mathbb{R}$ est dite de loi normale $\mathcal{N}(\mu,\sigma^2)$ si elle est absolument continue par rapport à la mesure de Lebesgue et admet pour densité*

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx.$$

Proposition 20.

- *Soit X,Y des variables aléatoires réelles indépendantes suivant une loi normale $\mathcal{N}(\mu,\sigma^2)$ et $\alpha>0$. Alors $\alpha X+Y\sim\mathcal{N}(\alpha\mu+\beta,\alpha^2\sigma^2)$.*
- *Soit $(X_1,\ldots,X_n)$ un n-uplet de variables aléatoires réelles indépendantes, tel que pour tout $i\in[\![1,n]\!]$, $X_i\sim\mathcal{N}(\mu_i,\sigma_i)$. Alors la moyenne $\frac{1}{n}\sum_{i=1}^n X_i$ suit une loi normale $\mathcal{N}\left(\frac{\mu_1+\cdots+\mu_n}{n},\frac{\sigma_1^2+\cdots+\sigma_n^2}{n^2}\right)$.*

1.3 Notion de covariance

Définition 21.

On se donne deux variables aléatoires réelles $X,Y\in L^2(\Omega,\mathcal{A},\mathbb{P})$. On définit la covariance de X et Y par

$$\mathrm{Cov}(X,Y):=\mathbb{E}[XY]-\mathbb{E}[X]\,\mathbb{E}[Y]$$

On dit que X et Y sont non corrélées si $\mathrm{Cov}(X,Y)=0$.

Exemple 22. Deux variables indépendantes sont non corrélées.

Exemple 23. La réciproque est fausse en général. Par exemple, si $X\sim\mathcal{N}(0,1)$ et $Y=X^2$, on vérifie que X et Y sont non corrélées, mais $\mathbb{P}(X\geq 1,Y\geq 1)=\mathbb{P}(X\geq 1)\neq\mathbb{P}(X\geq 1)\,\mathbb{P}(Y\geq 1)$ donc X et Y ne sont pas indépendantes.

Proposition 24. *Soit $(X_1,\ldots,X_n)$ des variables non corrélées. Alors on a*

$$\mathrm{Var}(X_1+\cdots+X_n)=\mathrm{Var}(X_1)+\cdots+\mathrm{Var}(X_n).$$

2 Théorèmes limites

2.1 Loi du 0-1 de Kolmogorov et lemme de Borel Cantelli

Définition 25. *Soit $(\mathcal{F}_n)_{n\geq 0}$ une famille de sous-tribus indépendantes sur un espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$. On appelle tribu queue de la suite $(\mathcal{F}_n)_{n\geq 0}$, ou tribu des évènements asymptotiques, la tribu $\mathcal{F}_\infty$ définie par*

$$\mathcal{F}_\infty:=\bigcap_{n\in\mathbb{N}}\sigma(\mathcal{F}_n,\mathcal{F}_{n+1},\ldots).$$

Théorème 26. *(Loi du 0-1)*
Si $\mathcal{A}_\infty$ est une tribu queue, alors pour tout $A\in\mathcal{A}_\infty$, on a $\mathbb{P}(A)=0$ ou $\mathbb{P}(A)=1$.

Exemple 27. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements indépendants de $(\Omega,\mathcal{A},\mathbb{P})$. Alors

$$\limsup_n A_n \overset{\text{def}}{=} \bigcap_{n\in\mathbb{N}} \bigcup_{m\geq n} A_n = \{A_n \text{ a lieu une infinité de fois}\}$$

est un élément de $\mathcal{A}_\infty$ pour la suite de tribus $\mathcal{F}_n := \sigma(A_n) = \{\varnothing,\Omega,A_n,A_n^c\}$.

Théorème 28. *(Lemme de Borel-Cantelli)*
Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements sur un espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$.

- Si $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n)<\infty$, *alors on a* $\mathbb{P}\left(\limsup_n A_n\right)=0$.
- *Si* $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n)=\infty$ *et que de plus, la suite $(A_n)_{n\in\mathbb{N}}$ est indépendante alors on a* $\mathbb{P}\left(\limsup_n A_n\right)=1$.

Corollaire 29.
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles sur $(\Omega,\mathcal{A},\mathbb{P})$ et X une variable aléatoire réelle sur $(\Omega,\mathcal{F},\mathbb{P})$.

On suppose que pour tout $k\geq 1$, $\sum_{n\in\mathbb{N}} \mathbb{P}\left(|X_n - X| > \frac{1}{k}\right)<\infty$.

Alors $X_n \underset{n\to+\infty}{\overset{\mathbb{P}\text{-}p.s}{\to}} X$.

2.2 Loi(s) des grands nombres

Proposition 30. *(Loi faible des grands nombres)*
On suppose que pour tout $n\in\mathbb{N}$, $X_n\in L^2$ et que $\mathrm{Cov}(X_i,X_j)=0$ pour tout $i,j\geq 1$, où $\mathrm{Cov}(X,Y)\overset{\text{def}}{=}\mathbb{E}[XY]-\mathbb{E}[X]\,\mathbb{E}[Y]$. On suppose de plus que $\mathbb{E}[X_i]=0$ et $\mathrm{Var}(X_i)=\sigma^2$ pour tout $i\geq 1$.

Alors la suite $\left(\frac{1}{n}\sum_{i=1}^n X_i\right)_{n\geq 0}$ *converge en probabilité vers zéro.*

Ce résultat repose sur l'inégalité de Tchébychev, et il est en particulier vrai lorsque les X_i sont indépendantes.

Théorème 31. *(Loi forte des grands nombres)*
Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires intégrables, indépendantes et de même loi. Alors

$$\frac{1}{n}\sum_{i=1}^n X_i \underset{n\to+\infty}{\overset{\mathbb{P}\text{-}p.s}{\to}} \mathbb{E}[X_0].$$

Remarque 32. La preuve est cependant beaucoup plus simple si $X_n\in L^4(\Omega,\mathcal{F},\mathbb{P})$, via l'inégalité de Tchébychev.

Application 33. (En statistiques)
Soit $(X_1,\dots,X_n)$ un n-échantillon issu d'une loi de moyenne μ et de variance σ^2.
Alors l'estimateur des moments $\overline{X_n} := \frac{1}{n}\sum_{i=1}^n X_i$ est sans biais, et fortement convergent.

2.3 Théorème central limite et applications

Proposition 34. *Soit X une variable aléatoire définie sur $(\Omega,\mathcal{F},\mathbb{P})$ suivant une loi normale centrée réduite. Alors la fonction caractéristique φ_X de X vérifie*

$$\forall t\in\mathbb{R} \quad \varphi_X(t) = e^{-\frac{t^2}{2}}.$$

Développement 1 :

Théorème 35. *(Théorème central limite)*
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées, définies sur $(\Omega,\mathcal{F},\mathbb{P})$. On suppose que $X_0\in L^2(\Omega,\mathcal{A},\mathbb{P})$ et on note $\sigma^2=\mathrm{Var}(X_0)$ et $\mu=\mathbb{E}[X_0]$. On suppose de plus que $\sigma^2>0$.
Par hypothèse, on a donc $X_n\in L^2(\Omega,\mathcal{F},\mathbb{P})$, $\mathrm{Var}(X_n)=\sigma^2$ et $\mathbb{E}[X_n]=\mu$pour tout $n\in\mathbb{N}$.
On note également $S_n=X_1+\cdots+X_n$ pour tout $n\geq 1$.
On a alors la convergence en loi

$$\frac{S_n - n\mu}{\sqrt{n}} \underset{n\to+\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}(0,\sigma^2).$$

Application 36. (En statistiques)
On considère le modèle statistique $(\{0,1\}^n, \{\mathcal{B}(\theta)^{\otimes n}\}_{\theta\in]0,1[})$ du jeu de pile ou face. Alors l'estimateur des moments $\overline{X_n}$ construit avec le n-échantillon $(X_1,\dots,X_n)$ est asymptotiquement normal, de vitesse $\sqrt{n}$, car le théorème central limite donne :

$$\sqrt{n}\,(\overline{X_n}-\theta) \underset{n\to+\infty}{\to} \mathcal{N}(0,\theta(1-\theta)).$$

3 D'autres applications de l'indépendance

3.1 Vecteurs gaussiens

Définition 37. *Une variable aléatoire $X=(X_1,\dots,X_d)\colon(\Omega,\mathcal{A},\mathbb{P})\to(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ est dite gaussienne (on parle de vecteur gaussien) si pour tout $\alpha=(\alpha_1,\dots,\alpha_d)\in\mathbb{R}^d$, il existe $\mu\in\mathbb{R}$ et $\sigma>0$ tel que*

$$\langle\alpha,X\rangle = \sum_{1\leq i\leq d}\alpha_i X_i \sim \mathcal{N}(\mu,\sigma^2).$$

Proposition 38. *Un vecteur gaussien $X=(X_1,\dots,X_d)$ est entièrement caractérisé par le vecteur des espérance $m=(\mathbb{E}[X_1],\dots,\mathbb{E}[X_d])$ et la matrice de variance-covariance $\Gamma=(\mathrm{Cov}(X_i,X_j))_{1\le i,j\le n}$*

On dit que $X\sim\mathcal{N}(m,\Gamma)$.

Exemple 39. Si $G_1,\dots,G_n$ sont des variables aléatoires indépendantes de loi normale centrée réduite $\mathcal{N}(0,1)$, alors $G=(G_1,\dots,G_n)$ est un vecteur gaussien, de loi $\mathcal{N}(0,I_n)$.

Sa loi a pour densité $f(x)=(2\pi)^{-d/2}\exp(-\|x\|^2/2)$ par rapport à la mesure de Lebesgue dx sur $\mathbb{R}^d$.

Proposition 40. *La matrice de variance-covariance Γ est symétrique définie positive. Aussi, il existe toujours une matrice carrée A telle que $\Gamma=A\,{}^tA$.*

Théorème 41. *Soit $X=(X_1,\dots,X_d)$ un vecteur gaussien centré, de matrice de variance-covariance égale à $\Gamma=A\,{}^tA$. Alors X a la même loi que AG, où G suit une loi $\mathcal{N}(0,I_d)$.*

Théorème 42. *Soit $X=(X_1,\dots,X_d)$ un vecteur gaussien de matrice de covariance Γ. Si les composantes de X sont deux à deux non corrélées (i.e si Γ est diagonale), alors la famille $(X_1,\dots,X_d)$ est mutuellement indépendante.*

3.2 Chaînes de Markov : exemple des marches aléatoires simples

Définition 43. *Soit E un espace d'état. On appelle noyau de transition sur E une fonction*

$$P:\begin{cases} E\times E & \to & [0,1]\\ (x,y) & \mapsto & p(x,y)\end{cases},$$

tel que pour tout $x\in E$ on ait

$$\sum_{y\in E} p(x,y)=1.$$

Définition 44. *On appelle chaîne de Markov sur l'espace d'état E de noyau de transition P un quituplet $(\Omega,\mathcal{F},(\mathcal{F}_n)_{n\in\mathbb{N}},(\mathbb{P}_x)_{x\in E},X=(X_n)_{n\ge0})$, tel que*

- *La suite X est adapté à la filtration $(\mathcal{F}_n)_{n\ge0}$.*
- *$\forall x\in E,\quad X_0=x$ $\mathbb{P}_x$-presque sûrement.*
- *$\forall n\ge0\quad \forall y\in E\quad \forall x\in E\quad \mathbb{P}_x(X_{n+1}=y|\mathcal{F}_n)=p(X_n,y)$.*

Exemple 45. (Marche aléatoire sur $\mathbb{Z}^d$)

On prend $E=\mathbb{Z}^d$, $d\ge1$. On se donne μ une probabilité sur $\mathbb{Z}^d$ et une suite $(\xi_i)_{i\ge1}$ de variables aléatoires $i.i.d$ sur un espace probabilisable $(\Omega,\mathcal{F})$, à valeurs dans $\mathbb{Z}^d$ et de loi μ.

On pose $X_0=0$ et pour tout $n\ge1$, $X_n=\xi_1+\cdots+\xi_n$.

Alors $(\Omega,\mathcal{F},(\sigma(X_n))_{n\in\mathbb{N}},(\mathbb{P}_x)_{x\in\mathbb{Z}^d},X=(X_n)_{n\ge0})$ est une chaîne de Markov sur $\mathbb{Z}^d$ de noyau de transition $p(x,y)=\mu(y-x)$, avec $\mathbb{P}_x=\delta_0$ pour tout $x\in\mathbb{Z}^d$.

Dans ce qui suit, on se donne un espace d'état E, un noyau de transition P et une chaîne de Markov $(\Omega,\mathcal{F},(\mathcal{F}_n)_{n\in\mathbb{N}},(\mathbb{P}_x)_{x\in E},X=(X_n)_{n\ge0})$. On se donne un élément $x\in E$.

Définition 46. *On définit le nombre N_x de visites en x et le premier temps T_x de retour en x par*

$$N_x:=\sum_{n\in\mathbb{N}}\mathbf{1}_{\{X_n=x\}}\quad \textit{et}\quad T_x:=\inf\{N\ge1:\ X_N=x\}.$$

Proposition 47. *Une et une seule des deux situations suivantes a lieu :*

- *$\mathbb{P}_x(T_x<\infty)=1$. Dans ce cas, $N_x=+\infty$ $\mathbb{P}_x$-p.s. On dit que l'état x est récurrent.*
- *$\mathbb{P}_x(T_x<\infty)<1$. Dans ce cas, $N_x<\infty$ $\mathbb{P}_x$-p.s, et de plus, $\mathbb{E}_x[N_x]=\dfrac{1}{\mathbb{P}_x(T_x=+\infty)}$. On dit que l'état x est transient.*

Développement 2 :

Exemple 48. (Récurrence de la marche aléatoire simple sur $\mathbb{Z}^d$)

La marche aléatoire simple sur $\mathbb{Z}$ et sur $\mathbb{Z}^2$ est récurrente. La marche aléatoire simple sur $\mathbb{Z}^d$ avec $d\ge3$ est transiente.