

Leçon 264. Variables aléatoires discrètes. Exemples et applications

Devs :

- Série génératrice et moments
- Marche aléatoire simple sur $\mathbb{Z}$, $\mathbb{Z}^2$ et $\mathbb{Z}^3$

Références :

1. Barbe-Ledoux, Probabilités
2. Ouvrard, Probabilités 1
3. Foata-Fuchs, Calcul des probabilités
4. Candelpergher, Théorie des probabilités
5. Garet, Probabilités et processus stochastiques

Dans tout ce qui suit, on se donne $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(E, \mathcal{B})$ un couple où E est un ensemble et $\mathcal{B}$ une tribu sur E .

1 Généralités

1.1 Définitions et propriétés

Définition 1. On dit qu'une variable aléatoire $X: (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (E, \mathcal{B})$ est discrète si $X(\Omega)$ est fini ou dénombrable.

Proposition 2. La loi d'une variable aléatoire discrète est de la forme $\mathbb{P}_X = \sum_{i \in I} p_i \delta_{x_i}$ avec $I \subset \mathbb{N}$, où δ_{x_i} désigne la masse de Dirac en x_i .

Exemple 3. Si $(E, \mathcal{B}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, toute variable aléatoire $X: (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ est discrète.

Définition 4. On définit la loi d'une variable discrète X comme la probabilité image $\mathbb{P}_X$ définie par $\mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A)) = \mathbb{P}(X \in A)$.

Proposition 5. Soit Λ un espace dénombrable, et g une fonction positive définie sur Λ . On suppose que $\sum_{\lambda \in \Lambda} g(\lambda) = 1$. Alors il existe une unique probabilité $\mathbb{P}$ sur Λ telle que $\mathbb{P}(\{\lambda\}) = g(\lambda)$ pour tout $\lambda \in \Lambda$. L'application g est alors appelée le germe de la probabilité $\mathbb{P}$.

1.2 Exemple de lois discrètes usuelles

Définition 6. 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 Bernoulli $\mathcal{B}(p)$ de paramètre p .

Exemple 7. On effectue n fois, de manière indépendante, une expérience aléatoire ayant deux issues : l'échec, ou le succès, avec une probabilité p de succès. On note X la variable aléatoire qui compte le nombre de succès sur les n lancers. Alors $X \sim \mathcal{B}(n, p)$.

Définition 8. Soit $p \in]0, 1[$. Une variable aléatoire X sur Ω à valeurs dans $\mathbb{N}^*$ est dite de loi géométrique $\mathcal{G}(p)$ si elle est absolument continue par rapport à la mesure de comptage sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ de densité

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

Exemple 9. On effectue plusieurs fois, de manière indépendante, une expérience aléatoire ayant deux issues : l'échec, ou le succès, avec une probabilité p de succès. On s'arrête lorsque l'on obtient le premier succès. On note X la variable aléatoire qui donne le rang du premier succès. Alors $X \sim \mathcal{G}(p)$.

Définition 10. Soit $\lambda \in]0, +\infty[$. Une variable aléatoire X sur Ω à valeurs dans $\mathbb{N}$ est dite de loi de Poisson $\mathcal{P}(\lambda)$ si elle est absolument continue par rapport à la mesure de comptage sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ de densité

$$f(k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Exemple 11. On réalise un prélèvement de n unités dans une population comportant deux sortes d'individus A et B en proportion p et q avec $p + q = 1$. On suppose que n est grand et p proche de zéro, avec $np \in [1, 10]$, par exemple. On note X le nombre d'individus de type A dans le prélèvement. Alors $X \sim \mathcal{P}(\lambda)$ (théoriquement). La loi de Poisson est dite « loi des événements rares ».

Proposition 12. Pour $k \geq 1$ fixé, on a $\lim_{n \rightarrow +\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$. Autrement dit, si $X_n \sim \mathcal{B}(n, \lambda/n)$ et $Y \sim \mathcal{P}(\lambda)$, alors $\lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = \mathbb{P}(Y = k)$ pour tout $k \geq 1$.

1.3 Calcul des moments

Définition 13. Soit X une variable aléatoire réelle discrète. Si $\sum_{x \in \mathbb{R}} |x| \mathbb{P}(X=x) < \infty$, on dit que la variable X admet une espérance, et on note

$$\mathbb{E}[X] = \sum_{x \in \mathbb{R}} x \mathbb{P}(X=x).$$

Dans ce qui suit, on considère une variable aléatoire réelle discrète X admettant une espérance.

Définition 14. Notons $\text{Val}(X)$ le sous espace de Ω défini par $\text{Val}(X) := \{\omega \in \Omega : \mathbb{P}_X(\{\omega\}) \neq 0\}$. Alors on a $\mathbb{E}[X] = \sum_{x \in \text{Val}(X)} x \mathbb{P}(X=x)$.

Si $\text{Val}(X)$ est fini, on peut donc écrire $\mathbb{E}[X] = \sum_{i=1}^n x_i \mathbb{P}(X=x_i)$ avec $\text{Val}(X) = \{x_1, \dots, x_n\}$.

Exemple 15. Si $X \sim \mathcal{B}(n, p)$, $\mathbb{E}[X] = np$.

Proposition 16. (Linéarité de l'espérance). Soit X, Y des variables aléatoires réelles discrètes, et $\lambda \in \mathbb{R}$. On suppose que X et Y admettent une espérance. Alors $X + \lambda Y$ admet une espérance et $\mathbb{E}[X + \lambda Y] = \mathbb{E}[X] + \lambda \mathbb{E}[Y]$.

Proposition 17. (Inégalité triangulaire). L'espérance vérifie $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

Théorème 18. (De transfert). Soit $E = \text{Val}(X)$ et $f: E \rightarrow \mathbb{R}$. Alors $Y = f \circ X$ est une variable aléatoire réelle discrète. Pour que Y admette une moyenne, il faut et il suffit que $\sum_{x \in E} |f(x)| \mathbb{P}(X=x) < \infty$. Dans ce cas, $f(X)$ admet une espérance et $\mathbb{E}[f(X)] = \sum_{x \in E} f(x) \mathbb{P}(X=x)$.

Définition 19. Si $\mathbb{E}[|X|^p] < \infty$, on appelle moment d'ordre p de X le réel $\mathbb{E}[|X|^p]$.

Définition 20. Si $\mathbb{E}[|X|^2] < \infty$, on appelle variance de X , et on note $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$.

Proposition 21. (Formule de König-Huygens). Si $\mathbb{E}[|X|^2] < \infty$, $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Exemple 22. Si $X \sim \mathcal{P}(\lambda)$, $\text{Var}(X) = \lambda$.

Définition 23. On définit la fonction (ou série) génératrice de X comme la série entière $G_X(t) := \sum_{k \geq 0} \mathbb{P}(X=k) t^k$.

Proposition 24. Le rayon de convergence R de la série entière $G_X(t)$ est supérieur ou égal à 1. La fonction $t \mapsto \mathbb{E}[t^X]$ est bien définie et à valeurs finies pour $t \in]-R, R[\cup \{-1, 1\}$, et vérifie alors $\mathbb{E}[t^X] = G_X(t)$.

Proposition 25.

1. La fonction génératrice de X sur $[-1, 1]$ détermine entièrement la loi de X .
2. Si X et Y sont indépendantes, alors $G_{X+Y} = G_X G_Y$.

Développement 1 :

Théorème 26. La variable aléatoire X est d'espérance finie si et seulement si G_X est dérivable à gauche au point 1, et dans ce cas, on a $\mathbb{E}[X] = G'_X(1)$.

Exemple 27. On effectue des jeux de pile ou face successifs indépendants, avec probabilité $p \in]0, 1[$ de faire face, et $1-p$ de faire pile. On s'arrête lorsque l'on obtient deux faces consécutives. En moyenne, il faut $\frac{1+p}{p^2}$ répétitions.

Exemple 28. Si $X \sim \mathcal{G}(\alpha)$, $G_X(t) = (1 - \alpha) \frac{s}{1 - \alpha s}$, et on en déduit $\mathbb{E}[X] = \frac{1}{1 - \alpha}$.

Proposition 29. (Inégalité de Markov). Si X admet une espérance, $\mathbb{P}(|X| \geq \varepsilon) \leq \frac{1}{\varepsilon} \mathbb{E}[|X|]$ pour tout $\varepsilon > 0$.

Corollaire 30. (Inégalité de Tchébychev). Si X admet une variance, $\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \text{Var}(X)$. pour tout $\varepsilon > 0$.

2 Loi d'un n -uplet et théorèmes limites

2.1 Indépendance et loi d'un n -uplet de variables discrètes

Définition 31. Soit $(X_1, \dots, X_n)$ un n -uplet de variables aléatoires discrètes réelles. On appelle loi conjointe la loi $\mathbb{P}_{(X_1, \dots, X_n)}$, et lois marginales les lois $\mathbb{P}_{X_1}, \dots, \mathbb{P}_{X_n}$ de chaque composante du n -uplet.

Remarque 32. La donnée de la loi conjointe permet de retrouver les lois marginales, mais la réciproque est fautive en général. On a cependant le théorème suivant.

Définition 33. Soit $(X_i)_{i \in I}$ une famille de variable aléatoire discrètes. On dit que les X_i sont indépendantes si pour toute famille $(A_i)_{i \in I}$ de boréliens de $\mathbb{R}$, les événements $\{X_i \in A_i\}$ sont indépendants.

Proposition 34. Soit $(X_i)_{1 \leq i \leq n}$ une famille de variables aléatoires discrètes, à valeurs dans des ensembles E_i . Alors les X_i sont indépendantes si et seulement si pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$ on a $\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \times \dots \times \mathbb{P}(X_n = x_n)$.

Remarque 35. Dans le cas où les X_i sont indépendantes, on peut retrouver la loi conjointe d'un n -uplet à partir de ses lois marginales.

2.2 Somme de variables aléatoires discrètes indépendantes

Proposition 36. Soient X et Y deux variables aléatoires discrètes réelles et indépendantes. Alors la variable aléatoire $X + Y$ est discrète et $\mathbb{P}(X + Y = k) = \sum_{\ell \in \mathbb{R}} \mathbb{P}(X = \ell) \mathbb{P}(Y = k - \ell)$.

Exemple 37. Si $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ alors $X + Y \sim \mathcal{P}(\lambda + \mu)$. Si $X \sim \mathcal{B}(n, p)$ et $Y \sim \mathcal{B}(m, p)$ alors $X + Y \sim \mathcal{B}(n + m, p)$.

Définition 38. Soit X une variable aléatoire réelle. On appelle fonction caractéristique de X la fonction $\varphi_X: t \mapsto \mathbb{E}[e^{itX}]$.

Théorème 39. Si X et Y ont la même fonction caractéristique, alors elles ont la même loi.

Théorème 40. Soit $(X_1, \dots, X_n)$ un n -uplet de variables discrètes indépendantes. Alors

$$\varphi_{X_1 + \dots + X_n} = \varphi_{X_1} \cdots \varphi_{X_n}.$$

Exemple 41. Si $X_1, \dots, X_n \sim \mathcal{B}(p)$ sont indépendantes, alors $X_1 + \dots + X_n \sim \mathcal{B}(n, p)$.

2.3 Théorèmes limites pour les variables discrètes

Théorème 42. (Loi forte des grands nombres)

Soit $(X_n)_{n \in \mathbb{N}}$ des variables réelles indépendantes de même loi admettant une espérance. Alors $\frac{X_1 + \dots + X_n}{n}$ converge presque partout vers $\mathbb{E}[X]$.

Exemple 43. Si $X_1, \dots, X_n$ sont i.i.d de loi $\mathcal{B}(p)$, alors $\frac{X_1 + \dots + X_n}{n} \sim \mathcal{B}(n, \frac{p}{n})$ converge presque sûrement vers p .

Définition 44. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles discrètes. On dit que $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une variable aléatoire X si $\lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$ pour tout $k \in \text{Val}(X)$.

Exemple 45. D'après la proposition 12, si $(X_n)_{n \in \mathbb{N}}$ est une suite i.i.d de variables de loi $\mathcal{B}(n, \lambda)$, alors $(X_n)_{n \in \mathbb{N}}$ converge en loi vers $X \sim \mathcal{P}(\lambda)$.

Théorème 46. (Central limite). Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables réelles indépendantes de même loi admettant une variance. Alors $\frac{X_1 + \dots + X_n - n\mathbb{E}[X_1]}{\sqrt{n}}$ converge en loi vers une variable $N \sim \mathcal{N}(0, \text{Var}(X_1))$.

3 Chaînes de Markov à espace d'état fini ou dénombrable

3.1 Définitions et propriétés

Définition 47. Soit E un espace d'état. On appelle noyau de transition sur E une fonction

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

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

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

Définition 48. On appelle chaîne de Markov sur l'espace d'état E de noyau de transition P un quintuplet $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, (\mathbb{P}_x)_{x \in E}, X = (X_n)_{n \geq 0})$, tel que

- La suite X est adapté à la filtration $(\mathcal{F}_n)_{n \geq 0}$.
- $\forall x \in E, \quad X_0 = x \quad \mathbb{P}_x$ -presque sûrement.
- $\forall n \geq 0 \quad \forall y \in E \quad \forall x \in E \quad \mathbb{P}_x(X_{n+1} = y | \mathcal{F}_n) = p(X_n, y)$.

Exemple 49. (Marche aléatoire sur $\mathbb{Z}^d$)

On prend $E = \mathbb{Z}^d, d \geq 1$. On se donne μ une probabilité sur $\mathbb{Z}^d$ et une suite $(\xi_i)_{i \geq 1}$ de variables aléatoires i.i.d sur un espace probabilisable $(\Omega, \mathcal{F})$, à valeurs dans $\mathbb{Z}^d$ et de loi μ . X_0 est à valeurs dans $\mathbb{Z}^d$ et pour tout $n \geq 1, X_n = \xi_1 + \dots + \xi_n$.

Alors $(\Omega, \mathcal{F}, (\sigma(X_n))_{n \in \mathbb{N}}, (\mathbb{P}_x)_{x \in \mathbb{Z}^d}, X = (X_n)_{n \geq 0})$ 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 $x \in \mathbb{Z}^d$.

Proposition 50. On note $\mathcal{C}_n = \sigma(\widehat{X}_1, \dots, \widehat{X}_n)$, où $\widehat{X}_i: \begin{cases} E^{\mathbb{N}} & \rightarrow E \\ (\omega_n)_{n \in \mathbb{N}} & \mapsto \omega_n \end{cases}$.

Alors $\bigcup_{n \geq 0} \mathcal{C}_n$ est un π -système, et une fonction $F: (\Omega, \mathcal{F}) \rightarrow (E^{\mathbb{N}}, \mathcal{C})$ est mesurable si et seulement si

$$\forall n \geq 0 \quad \widehat{X}_n \circ F \text{ est mesurable.}$$

Définition 51. On définit l'opérateur de décalage θ sur $E^{\mathbb{N}}$ par $\theta((\omega_n)_{n \in \mathbb{N}}) = (\omega_{n+1})_{n \in \mathbb{N}}$. On note $\theta_0 = \text{Id}_{E^{\mathbb{N}}}$ et $\theta_n = \theta \circ \dots \circ \theta = \theta^n$.

Définition 52.

- Soit $x \in E$ et $X: (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{P}(E))$ une variable aléatoire. On note $\mathbb{E}_x[X] := \int_{\Omega} X d\mathbb{P}_x$ l'espérance sous la mesure $\mathbb{P}_x$.
- Si $Z: (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{P}(E))$ est une variable aléatoire et Y une variable aléatoire positive sur $(\Omega, \mathcal{F})$, on note $\mathbb{E}_Z[Y] := \sum_{x \in E} \mathbb{E}_x[Y] \mathbf{1}_{\{Z=x\}}$.

Théorème 53. (Propriété de Markov faible)

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, (\mathbb{P}_x)_{x \in E}, X = (X_n)_{n \geq 0})$ une chaîne de Markov d'espace d'état E et de noyau de transition P . Soit $F: (\Omega, \mathcal{F}) \rightarrow (E^{\mathbb{N}}, \mathcal{C})$.

Alors

$$\forall n \geq 0 \quad \forall x \in E \quad \mathbb{E}_x[F(\theta_n(X)) \mid \mathcal{F}_n] = \mathbb{E}_{X_n}[F(X)] \quad \mathbb{P}_x\text{-presque sûrement.}$$

Théorème 54. (Propriété de Markov forte)

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, (\mathbb{P}_x)_{x \in E}, X = (X_n)_{n \geq 0})$ une chaîne de Markov d'espace d'état E et de noyau de transition P . Soit $F: (\Omega, \mathcal{F}) \rightarrow (E^{\mathbb{N}}, \mathcal{C})$. Soit T un temps d'arrêt relativement à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$.

Alors

$$\forall n \geq 0 \quad \forall x \in E \quad \mathbb{E}_x[F(\theta_T(X)) \mathbf{1}_{\{T < \infty\}} \mid \mathcal{F}_n] = \mathbb{E}_{X_n}[F(X)] \cdot \mathbf{1}_{\{T < \infty\}} \quad \mathbb{P}_x\text{-presque sûrement.}$$

3.2 Classification des états

Dans ce qui suit, on se donne un espace d'état E , un noyau de transition P et une chaîne de Markov $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, (\mathbb{P}_x)_{x \in E}, X = (X_n)_{n \geq 0})$. On se donne un élément $x \in E$.

Définition 55. On définit le nombre N_x de visites en x et le premier temps T_x de retour en x par

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

Proposition 56. 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, $\mathbb{E}_x[N_x] = \frac{1}{\mathbb{P}_x(T_x = +\infty)}$. On dit que l'état x est transient.

Exemple 57.

Dans la marche aléatoire simple sur $\mathbb{Z}$, l'état zéro est récurrent.

Définition 58. Soit $x, y \in E$. On dit que x mène à y et on note $x \rightarrow y$ si $\mathbb{E}_x[N_x] > 0$.

La relation $\rightarrow$ est réflexive et transitive.

Proposition 59. Soit $x, y \in E$. On suppose que $x \rightarrow y$ et que x est récurrent. Alors y est récurrent, et $y \rightarrow x$.

Définition 60. On dit que la chaîne de Markov, ou le noyau de transition P est irréductible si

$$\forall x, y \in E \quad x \rightarrow y.$$

Théorème 61. (Classification des états d'une chaîne irréductible)

Supposons la chaîne irréductible. Alors une et une seule des deux situations suivantes a lieu :

- Tous les états sont récurrents, et $\forall x \in E \quad \mathbb{P}_x(\forall y \in E, N_y = +\infty) = 1$.
- Tous les états sont transients, et $\forall x \in E \quad \mathbb{P}_x(\forall y \in E, N_y < \infty) = 1$.

Si E est fini, alors on est toujours dans la première situation.

Exemple 62. (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 $\mathbb{Z}^3$ est transiente.