14
Variables aléatoires discrètes

Images réciproques

1. Notation probabiliste
Soit $f : E \to F$, une application.
1.1 Pour toute partie A de F, l'**image réciproque** de A par f est la partie de E, traditionnellement notée

$$f^{-1}(A) \quad \text{ou} \quad f^*(A)$$

et définie par

$$f^*(A) = \big\{x \in E \,:\, f(x) \in A\big\}.$$

Pour des raisons de commodité, qui apparaîtront peu à peu, l'image réciproque de A par f sera notée en général

$$[f \in A]$$

ou éventuellement $[f(x) \in A]$ (en géométrie notamment).
1.2 Pour tout $y \in F$, l'image réciproque du singleton $\{y\}$ sera notée

$$[f = y] \qquad \text{ou} \qquad [f(x) = y].$$

C'est la **ligne de niveau** y de la fonction f.
1.3 Si f est une fonction à valeurs réelles (c'est-à-dire $F = \mathbb{R}$), on notera

$$[f \geqslant y] \quad \text{au lieu de} \quad \big[f \in [y, +\infty[\big].$$

De même, on notera respectivement

$$[f \leqslant y], \quad [f > y], \quad [f < y], \quad [a \leqslant f \leqslant b] \quad \text{et} \quad [a < f \leqslant b]$$

les images réciproques des intervalles

$$]-\infty, y]\,, \quad]y, +\infty[\,, \quad]-\infty, y[\,, \quad [a, b] \quad \text{et} \quad]a, b]$$

par l'application f.

2. ➝ Propriétés usuelles
Soit $X : \Omega \to E$, une application.
Quelles que soient A et B, deux parties de E,
2.1
$$A \subset B \implies [X \in A] \subset [X \in B].$$
2.2
$$[X \in A \cup B] = [X \in A] \cup [X \in B]$$
2.3
$$[X \in A \cap B] = [X \in A] \cap [X \in B]$$
2.4
$$[X \in A^c] = [X \in A]^c$$

3. ➝ Image réciproque et composée
Soient $X : \Omega \to E$ et $f : E \to F$, deux applications.

$$\forall\, A \subset F, \quad \big[(f \circ X)(\omega) \in A\big] = \big[f(X) \in A\big] = \big[X \in f^*(A)\big]$$

4. Image réciproque d'une tribu
Soient $(E, \mathcal{E})$, un espace mesurable et $X : \Omega \to E$, une application.
4.1 L'ensemble $\sigma(X) = \big\{[X \in A],\ A \in \mathcal{E}\big\} \subset \mathfrak{P}(\Omega)$ est une tribu sur Ω.
4.2 ✍ *La tribu $\sigma(X)$ est appelée **tribu engendrée** sur Ω par l'application $X : \Omega \to E$.*

I
Fondements du calcul mathématique des probabilités

5. La notion de **variable aléatoire** est la notion fondamentale de la théorie mathématique des probabilités.
5.1 ✍ *Soit $(E, \mathcal{E})$, un espace mesurable.*
*On appelle **variable aléatoire** sur un espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$ à valeurs dans E toute application $X : \Omega \to E$ telle que*

$$\forall\, A \in \mathcal{E}, \qquad [X \in A] \in \mathfrak{A}.$$

*On précise que X est une **variable aléatoire réelle** lorsqu'elle prend ses valeurs dans $E = \mathbb{R}$ et que la tribu $\mathcal{E}$ est la tribu borélienne $\mathcal{B}$.*
5.2 Si $X : \Omega \to E$ est une variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P})$, alors l'application $\mu_X : \mathcal{E} \to [0, 1]$ définie par

$$\forall\, A \in \mathcal{E}, \quad \mu_X(A) = \mathbf{P}(X \in A)$$

est une mesure de probabilité sur $(E, \mathcal{E})$.
5.3 ✍ *La mesure de probabilité μ_X sur $(E, \mathcal{E})$ est la **loi (sous $\mathbf{P}$)** de la variable aléatoire X.*
5.4 Point de vue statistique
La loi d'une variable aléatoire X décrit les fréquences d'apparition des différentes valeurs prises par la fonction X, c'est-à-dire la manière dont les valeurs de X sont *distribuées* ou *réparties*.
5.5 Le **support discret** de X est l'ensemble des $x \in E$ tels que

$$\mathbf{P}(X = x) > 0$$

c'est-à-dire l'ensemble des atomes de la mesure μ_X. →[**13.78**]
5.6 ✍ *Une variable aléatoire $X : (\Omega, \mathfrak{A}, \mathbf{P}) \to (E, \mathcal{E})$ est dite **à support fini*** (resp. **dénombrable***, resp. **borné***) lorsqu'il existe une partie finie (resp. dénombrable, resp. bornée) $E_0 \in \mathcal{E}$ telle que*

$$\mathbf{P}(X \in E_0) = 1.$$

6. Hasard et déterminisme
6.1 Le graphe d'une application $f : E \to F$ associe une valeur de l'ensemble d'arrivée F à chaque valeur de l'ensemble de départ E. Autrement dit, chaque valeur x de l'ensemble de départ *détermine* une valeur $y = f(x)$ de l'ensemble d'arrivée : cette valeur y apparaît comme la **conséquence**, ou l'**effet**, de la **cause** x.
6.2 ➝ Théorème fondamental
Quelle que soit la mesure de probabilité μ définie sur un espace mesurable $(E, \mathcal{E})$, il existe un espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$ et une variable aléatoire $X : \Omega \to E$ telle que

$$\forall\, A \in \mathcal{E}, \quad \mathbf{P}(X \in A) = \mu(A).$$

6.3 La démonstration du théorème [**6.2**] importe peu. L'essentiel est de comprendre qu'on peut ainsi représenter n'importe quelle mesure de probabilité μ par une variable aléatoire, c'est-à-dire une fonction X définie on ne sait comment sur un ensemble inconnu Ω.
Cette fonction X est la représentation mathématique d'un phénomène **aléatoire** au sens suivant :
- on sait quelles sont les valeurs que peut prendre X : les éléments de E, ensemble d'arrivée ;
- on connaît la répartition statistique de ces valeurs, décrite par la mesure de probabilité μ ;
- mais chaque valeur prise par cette fonction apparaît comme un effet, $X(\omega)$, dont la cause, ω, reste inconnue.

6.4 Une variable aléatoire est donc une fonction mathématique au sens usuel qui ne sera donc pas étudiée de la manière usuelle : on ne s'intéressera pas au graphe de cette fonction, mais seulement à la probabilité pour que la valeur prise par cette fonction appartienne à un ensemble $A \in \mathcal{E}$, probabilité égale par définition à $\mu(A)$.

6.5 Autrement dit, on s'intéresse seulement à la **loi** d'une variable aléatoire X et aux propriétés de cette loi.

I.1 Variables aléatoires discrètes

7. ✍ *Une **variable aléatoire discrète** sur $(\Omega, \mathfrak{A}, \mathbf{P})$ est une application à valeurs dans un ensemble fini ou dénombrable E :*

$$\forall\, \omega \in \Omega, \quad X(\omega) \in E$$

telle que

$$\forall\, x \in E, \qquad [X = x] \in \mathfrak{A}.$$

8. *Suite de* [6.5] – Pour étudier une variable aléatoire discrète, on doit commencer par expliciter un ensemble E (au plus dénombrable) dans lequel elle prend ses valeurs.

Loi et lois conditionnelles d'une variable discrète

9. Soit X, une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$ qui prend ses valeurs dans un ensemble E fini ou dénombrable.

9.1 La famille

$$\big([X = x]\big)_{x \in E}$$

est un système complet d'événements.

9.2 La famille

$$\big(\mathbf{P}(X = x)\big)_{x \in E'}$$

est une loi discrète sur E.

9.3 La **loi** [5.3] de X est une mesure de probabilité sur l'espace mesurable discret $(E, \mathfrak{P}(E))$, qui est caractérisée [**13.60**] par la famille

$$\big(\mathbf{P}(X = x)\big)_{x \in E}.$$

10. ➞ Réalisation d'une loi discrète
Soit E, un ensemble fini ou dénombrable. Pour toute loi discrète $(p_x)_{x \in E}$ sur E, il existe un espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$ et une variable aléatoire discrète $X : \Omega \to E$ telle que

$$\forall\, x \in E, \qquad \mathbf{P}(X = x) = p_x.$$

11. Soit $X : \Omega \to E$.

11.1 Que X soit une variable aléatoire discrète sur $(\Omega, \mathfrak{A})$ à valeurs dans un ensemble E ne dépend pas de la mesure de probabilité $\mathbf{P}$ sur $(\Omega, \mathfrak{A})$. →[**7**]

11.2 En revanche, les fréquences d'apparition des valeurs de X dépendent de la mesure $\mathbf{P}$. Autrement dit, *la loi de X dépend de* $\mathbf{P}$ et en modifiant la mesure de probabilité sur $(\Omega, \mathfrak{A})$ (par exemple en conditionnant par un événement non négligeable $A \in \mathfrak{A}$), on modifie la loi de X.

12. Soient $(E, \mathcal{E})$, un espace mesurable ; $(\Omega, \mathfrak{A}, \mathbf{P})$, un espace probabilisé et $X : (\Omega, \mathfrak{A}) \to (E, \mathcal{E})$, une variable aléatoire.

12.1 ✍ *Si l'événement $A \in \mathfrak{A}$ n'est pas négligeable :*

$$\mathbf{P}(A) > 0,$$

*alors la **loi conditionnelle de** X **sachant** A est la mesure de probabilité ν_X sur $(E, \mathcal{E})$ définie par*

$$\forall\, B \in \mathcal{E}, \quad \nu_X(B) = \mathbf{P}(X \in B \mid A).$$

12.2 La loi conditionnelle ν_X de X sachant A est en fait la loi de X sous $\mathbf{P}_A$:

$$\forall\, B \in \mathcal{E}, \quad \nu_X(B) = \mathbf{P}_A(X \in B).$$

ce qui revient à considérer X comme une variable aléatoire sur l'espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P}_A)$.

12.3 Si $X : \Omega \to E$ est une variable aléatoire discrète, la loi conditionnelle de X sachant A est caractérisée [**13.60**] par la loi dicrète

$$\big(\mathbf{P}_A(X = x)\big)_{x \in E}.$$

12.4 Le support discret [5.5] de X sous $\mathbf{P}_A$ est contenu dans le support discret de X sous $\mathbf{P}$.

Égalité en loi

13. ✍ *On note $X \sim Y$, ou $X \stackrel{\mathrm{d}}{=} Y$, le fait que les deux variables aléatoires X et Y aient même loi. On dit aussi qu'elles sont **identiquement distribuées**.*

14. Deux variables aléatoires définies sur des espaces probabilisés différents peuvent avoir une même loi : il faut pour cela que l'espace mesurable d'arrivée soit le même pour ces deux fonctions.

14.1 ➞ *Soient X et Y, deux variables aléatoires discrètes respectivement définies sur $(\Omega_1, \mathfrak{A}_1, \mathbf{P})$ et sur $(\Omega_2, \mathfrak{A}_2, \mathbf{Q})$.*
Les deux variables aléatoires X et Y ont même loi si, et seulement si, elles prennent leurs valeurs dans un même ensemble fini ou dénombrable E et si

$$\forall\, x \in E, \quad \mathbf{P}(X = x) = \mathbf{Q}(Y = x).$$

14.2 Si deux variables aléatoires discrètes sur $(\Omega, \mathfrak{A})$ sont presque sûrement égales : $\mathbf{P}(X = Y) = 1$, alors elles ont même loi.
La réciproque est fausse [**18.4**], [**37**].

15. Si deux variables aléatoires X et Y sont définies sur un même espace probabilisé, on peut affiner la comparaison de ces variables en étudiant la loi du **vecteur aléatoire** (X, Y).

I.2 Lois usuelles

16. On présente ici les lois discrètes usuelles les plus simples. Chaque loi est vue, non comme une mesure de probabilité sur un espace mesurable discret, mais comme la loi d'une variable aléatoire discrète X, définie sur un espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$, à valeurs dans $\mathbb{R}$.

17.1 ✍ *Pour $x_0 \in \mathbb{R}$, la variable aléatoire X suit la **loi de Dirac en** x_0 lorsque*

$$\mathbf{P}(X = x_0) = 1.$$

On note alors $X \stackrel{\mathrm{d}}{=} \delta_{x_0}$.

17.2 La variable aléatoire X suit la loi de Dirac en x_0 si, et seulement si, la variable aléatoire Y définie par $X = Y + x_0$ suit la loi de Dirac en 0.

17.3 Si X suit la loi de Dirac en x_0, alors

$$\forall\, A \subset \mathbb{R}, \qquad \mathbf{P}(X \in A) = \left| \begin{array}{l} 1 \text{ si } x_0 \in A, \\ 0 \text{ si } x_0 \notin A. \end{array} \right.$$

18.1 ✍ *La variable aléatoire X suit la **loi de Bernoulli** de paramètre $0 < p < 1$ lorsque*

$$\mathbf{P}(X = 1) = p \quad \textit{et} \quad \mathbf{P}(X = 0) = 1 - p.$$

Dans ce cas, on note $X \stackrel{\mathrm{d}}{=} \mathcal{B}(p)$ et on pose souvent $q = 1 - p$.

18.2 Variable de Bernoulli associée à un événement
Pour tout événement $A \in \mathfrak{A}$, l'indicatrice

$$\mathbb{1}_A : \Omega \to \{0, 1\}$$

est une variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P})$ qui suit la loi de Bernoulli de paramètre $p = \mathbf{P}(A)$.

18.3 S'il existe deux réels $a \neq b$ tels que

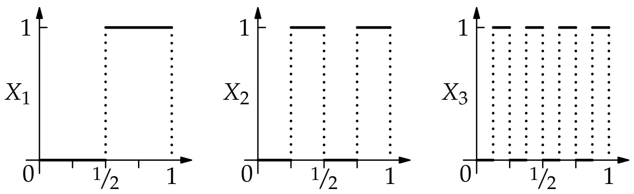
$$\mathbf{P}(X = a) > 0,\ \mathbf{P}(X = b) > 0 \quad \text{et} \quad \mathbf{P}(X = a) + \mathbf{P}(X = b) = 1,$$

alors il existe deux réels α et β et une variable aléatoire Y suivant une loi de Bernoulli tels que $X = \alpha Y + \beta$.

18.4 Admettons que l'ensemble $\Omega = [0,1]$ puisse être muni d'une tribu $\mathcal{B}$, dite **tribu borélienne**, qui contienne tous les intervalles et qu'il existe une mesure de probabilité $\mathbf{P}$ sur $\mathcal{B}$, dite **mesure de Lebesgue**, telle que

$$\forall\, 0 \leqslant a < b \leqslant 1, \quad \mathbf{P}\big([a,b]\big) = b - a.$$

Les fonctions X_1, X_2, X_3 dont les graphes figurent ci-dessous sont des variables aléatoires discrètes sur $(\Omega, \mathcal{B})$ qui suivent toutes la loi de Bernoulli de paramètre $p = 1/2$.



19.1 ✍ *La variable aléatoire X suit la **loi uniforme** sur une famille finie $\mathscr{F} = (x_k)_{0 \leqslant k < n}$ de n réels lorsque*

$$\forall\, 0 \leqslant k < n, \quad \mathbf{P}(X = x_k) = 1/n.$$

On note alors $X \overset{\mathrm{d}}{=} \mathcal{U}(\mathscr{F})$.

19.2 Dans ce cas, il existe une variable aléatoire Y suivant le loi uniforme sur $\{1, 2, \dots, n\}$ et une fonction $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathscr{C}^\infty$ telle que $X = f(Y)$.

20. ✍ *La variable aléatoire X suit la **loi binomiale** de paramètres $n \in \mathbb{N}$ et $0 < p < 1$ lorsque*

$$\forall\, 0 \leqslant k \leqslant n, \quad \mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On note alors $X \overset{\mathrm{d}}{=} \mathcal{B}(n, p)$.

21. Les lois géométriques sont les **lois sans mémoire** [**21.4**].

21.1 ✍ *La variable aléatoire X suit la **loi géométrique** de paramètre $0 < p < 1$ lorsque*

$$\forall\, k \geqslant 1, \quad \mathbf{P}(X = k) = (1-p)^{k-1} p.$$

On note alors $X \overset{\mathrm{d}}{=} \mathcal{G}(p)$ et on pose souvent $q = 1 - p$.

21.2

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X > k) = (1-p)^k.$$

21.3 Caractérisation de la loi géométrique
Une variable aléatoire X à valeurs dans $\mathbb{N}$ suit la loi géométrique de paramètre $p = 1 - q$ si, et seulement si,

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X > k) = q^k.$$

21.4 ➙ *Une variable aléatoire à valeurs dans $\mathbb{N}^*$ suit une loi géométrique si, et seulement si,*

$$\forall\, k, n \in \mathbb{N}, \quad \mathbf{P}(X > n + k \mid X > n) = \mathbf{P}(X > k).$$

22. La loi de Poisson peut être vue comme un cas limite de la loi binomiale.

22.1 ✍ *La variable aléatoire X suit la **loi de Poisson** de paramètre $\lambda > 0$ lorsque*

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

On note alors $X \overset{\mathrm{d}}{=} \mathcal{P}(\lambda)$.

22.2 ➙ **Loi des événements rares**
Soit $(X_n)_{n \in \mathbb{N}}$, une suite de variables aléatoires définies sur $(\Omega, \mathfrak{A}, \mathbf{P})$. Pour tout $n \in \mathbb{N}$, on suppose que X_n suit la loi binomiale $\mathscr{B}(n, p_n)$ et que le produit np_n tend vers $\lambda \in \mathbb{R}_+^*$ lorsque n tend vers $+\infty$. Alors

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X_n = k) \xrightarrow[n \to +\infty]{} e^{-\lambda} \frac{\lambda^k}{k!}.$$

Entraînement

23. Questions pour réfléchir

1. L'application $X : \Omega \to E$ est une variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P})$ si, et seulement si, la tribu $\sigma(X)$ engendrée par X est contenue dans la tribu $\mathfrak{A}$. →[**4**]

2. *Suite de* [**5.6**] –

2.a Si X est une variable aléatoire discrète, alors elle est aussi une variable à support dénombrable.

2.b Avec $\Omega = E = \mathbb{R}$ et $\mathfrak{A} = \mathcal{E} = \mathfrak{P}(\mathbb{R})$, on pose

$$\forall\, \omega \in \Omega, \quad X(\omega) = \omega.$$

Si $\mathbf{P}$ est la mesure de Dirac en 0 [**17.3**], alors X est une variable aléatoire à support fini bien que X ne soit pas une variable aléatoire discrète.

3. Si X et Y sont deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$ et si elles sont égales :

$$\forall\, \omega \in \Omega, \quad X(\omega) = Y(\omega),$$

alors elles ont même loi.

4. Pourquoi exclut-on *a priori* les valeurs $p = 0$ et $p = 1$ de la définition des lois de Bernoulli ?

5.a La loi binomiale de paramètres $n = 0$ et p est une loi de Dirac.

5.b La loi binomiale de paramètres $n = 1$ et p est la loi de Bernoulli $\mathcal{B}(p)$.

6. Si X est une variable aléatoire à valeurs dans $\mathbb{N}$ dont la loi est sans mémoire [**21.4**] et si $0 < \mathbf{P}(X = 0) < 1$, alors la loi conditionnelle de X sachant $[X > 0]$ est une loi géométrique.

7. Si $X \overset{\mathrm{d}}{=} \mathcal{G}(p)$, alors $\mathbf{P}(X > 0) = 1$. D'autre part, plus p est proche de 1, plus l'événement $[X = 1]$ est probable.

8. Si $X \overset{\mathrm{d}}{=} \mathcal{P}(\lambda)$, alors $\lambda = -\ell n\, \mathbf{P}(X = 0)$. Par conséquent, plus λ est proche de 0, plus l'événement $[X = 0]$ est probable.

24. Opérations sur les variables indicatrices [18.2]
Soit $(\Omega, \mathfrak{A}, \mathbf{P})$, un espace probabilisé.

$$\begin{aligned}
\forall\, A \in \mathfrak{A}, &\quad \mathbb{1}_{A^c} = 1 - \mathbb{1}_A. \\
\forall\, A, B \in \mathfrak{A}, &\quad \mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B. \\
\forall\, A, B \in \mathfrak{A}, &\quad \mathbb{1}_{A \cup B} = 1 - (1 - \mathbb{1}_A)(1 - \mathbb{1}_B) \\
&\quad \phantom{\mathbb{1}_{A \cup B}} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B}.
\end{aligned}$$

25. Soit X, une variable aléatoire à valeurs dans $E = \mathbb{N}^*$, dont la loi est caractérisée par

$$\forall\, n \geqslant 1, \quad \mathbf{P}(X = n) = \frac{1}{\zeta(\alpha) n^\alpha}$$

où $\alpha > 1$ est un réel fixé.

1. La probabilité pour que X soit divisible par un entier $p \geqslant 1$ donné est égale à $p^{-\alpha}$.

2. Quelle est la probabilité pour que X soit impaire ?

26. *Suite de* [**18.2**] – Soit $B \in \mathfrak{A}$ tel que $\mathbf{P}(B) > 0$.
La loi conditionnelle de $X = \mathbb{1}_A$ sachant B est la loi de Bernoulli de paramètre $\mathbf{P}(A \mid B)$.

27. Si la variable aléatoire X suit une loi binomiale $\mathcal{B}(n, p)$ avec $n \geqslant 2$, alors la loi conditionnelle de X sachant $A = [X \leqslant 1]$ est la loi de Bernoulli de paramètre

$$\frac{np}{1 + (n-1)p}.$$

28. Taux de défaillance
L'instant au bout duquel un dispositif cesse de fonctionner est modélisé par une variable aléatoire X à valeurs dans $\mathbb{N}^*$. Le **taux de défaillance** de X est la suite de terme général

$$x_n = \mathbf{P}(X = n \mid X \geqslant n).$$

Un taux de défaillance croissant (resp. décroissant) modélise un phénomène d'usure (resp. de rodage).

28.1 Le taux de défaillance est constant si, et seulement si, X suit une loi géométrique (*absence de mémoire*).

28.2 Si la loi de X est caractérisée par

$$\forall n \in \mathbb{N}^*, \quad \mathbf{P}(X = n) = \frac{1}{n(n+1)},$$

alors le taux de défaillance tend vers 0.

28.3 Si X suit la loi de Poisson de paramètre λ, alors

$$\frac{1}{x_n} = e^{\lambda} \int_0^1 nv^{n-1} e^{-\lambda v} \, dv$$

et le taux de défaillance tend vers 1.

29. Fonction de répartition
Soit $X : \Omega \to \mathbb{N}$, une variable aléatoire discrète. On pose

$$r_k = \mathbf{P}(X \leqslant k) \quad \text{et} \quad q_k = \mathbf{P}(X > k)$$

pour tout $k \in \mathbb{N}$.

29.1 La suite $(r_k)_{k \in \mathbb{N}}$ est positive, croissante et tend vers 1.
Comme $\mathbf{P}(X = 0) = r_0$ et que

$$\forall k \geqslant 1, \quad \mathbf{P}(X = k) = r_k - r_{k-1},$$

la loi de X est caractérisée par la suite $(r_k)_{k \in \mathbb{N}}$.

29.2 Soit $(u_k)_{k \in \mathbb{N}}$, une suite positive, croissante et qui tend vers 1. La suite $(v_k)_{k \in \mathbb{N}}$ définie par $v_0 = u_0$ et par

$$\forall k \geqslant 1, \quad v_k = u_k - u_{k-1}$$

est une loi de probabilité discrète sur $\mathbb{N}$ et il existe une variable aléatoire X telle que

$$\forall k \in \mathbb{N}, \quad \mathbf{P}(X \leqslant k) = u_k.$$

29.3 La suite $(q_k)_{k \in \mathbb{N}}$ est positive, décroissante et de limite nulle.
Comme $\mathbf{P}(X = 0) = 1 - q_0$ et que

$$\forall k \geqslant 1, \quad \mathbf{P}(X = k) = q_{k-1} - q_k,$$

la loi de X est aussi caractérisée par la suite $(q_k)_{k \in \mathbb{N}}$.

30. Soit $X : \Omega \to \mathbb{R}$, une variable aléatoire discrète.

1. $[X \leqslant x] \in \mathfrak{A}$ pour tout $x \in \mathbb{R}$.
2. $$\bigcap_{n \in \mathbb{N}} [X \leqslant -n] = \varnothing \qquad \bigcup_{n \in \mathbb{N}} [X \leqslant n] = \Omega$$
3. Pour tout $x \in \mathbb{R}$,

$$[X \leqslant x] = \bigcap_{n \in \mathbb{N}^*} \Big[X \leqslant x + \frac{1}{n}\Big]$$

et

$$[X < x] = \bigcup_{n \in \mathbb{N}^*} \Big[X \leqslant x - \frac{1}{n}\Big] \in \mathfrak{A}.$$

II

Opérations sur les variables discrètes

31. Composition par une fonction
Soit X, une variable aléatoire sur $(\Omega, \mathfrak{A})$, à valeurs dans l'ensemble fini ou dénombrable E muni de sa tribu discrète $\mathcal{E}$.

31.1 ➞ *Si f est une application de E dans un ensemble fini ou dénombrable F, alors*

$$\forall B \subset F, \quad [f(X) \in B] = [X \in f^*(B)]$$

et $f(X)$ est une variable aléatoire discrète sur $(\Omega, \mathfrak{A})$.

31.2 La loi de $f(X)$ se déduit de la loi de X : elle est caractérisée par la famille des probabilités

$$\mathbf{P}\big(X \in f^*(\{y\})\big) = \sum_{x \in f^*(\{y\})} \mathbf{P}(X = x)$$

lorsque y parcourt F.

31.3 ➞ *Si X est une variable aléatoire réelle et $\lambda \in \mathbb{R}$, alors λX est une variable aléatoire réelle.*

31.4 Si E et F sont des parties de $\mathbb{R}$, la loi de $f(X)$ peut se déduire de la fonction de répartition **[109]** de X. →**[114.3]**

32. Soient X et Y, deux variables aléatoires discrètes sur $(\Omega, \mathfrak{A})$. On note E_1 et E_2, deux parties finies ou dénombrables de $\mathbb{R}$ telles que

$$\forall \omega \in \Omega, \quad X(\omega) \in E_1 \quad \text{et} \quad Y(\omega) \in E_2.$$

32.1 Lemme fondamental
Pour toute application $f : E_1 \times E_2 \to \mathbb{R}$, l'ensemble

$$F = f_*(E_1 \times E_2)$$

est une partie au plus dénombrable de $\mathbb{R}$ et

$$\forall B \subset F, \quad [f(X,Y) \in B] = \bigsqcup_{(x,y) \in f^*(B)} [X = x] \cap [Y = y]$$

donc $f(X,Y)$ est une variable aléatoire discrète sur $(\Omega, \mathfrak{A})$.

32.2 Somme de deux variables aléatoires
Pour tout $z \in F = \{x_1 + x_2, \, (x_1, x_2) \in E_1 \times E_2\}$,

$$[X + Y = z] = \bigsqcup_{\substack{(x,y) \in E_1 \times E_2 \\ x+y=z}} [X = x] \cap [Y = y]$$

donc $X + Y$ est une variable aléatoire réelle sur $(\Omega, \mathfrak{A})$.

32.3 ➞ *L'ensemble des variables aléatoires réelles presque sûrement bornées est un sous-espace vectoriel, notée $\mathscr{L}^\infty(\Omega)$, de $\mathscr{A}(\Omega, \mathbb{R})$.*

32.4 Produit de deux variables aléatoires
Pour tout $z \in F = \{x_1 x_2, \, (x_1, x_2) \in E_1 \times E_2\}$,

$$[XY = z] = \bigsqcup_{\substack{(x,y) \in E_1 \times E_2 \\ xy=z}} [X = x] \cap [Y = y]$$

donc XY est une variable aléatoire réelle sur $(\Omega, \mathfrak{A})$.

32.5 Quotient de deux variables aléatoires
On suppose que $0 \notin E_2$, de telle sorte que $[Y = 0] = \varnothing$. Alors pour tout $z \in F = \{{}^{x_1}/_{x_2}, \, (x_1, x_2) \in E_1 \times E_2\}$,

$$[{}^{X}/_{Y} = z] = \bigsqcup_{\substack{(x,y) \in E_1 \times E_2 \\ {}^{x}/_{y}=z}} [X = x] \cap [Y = y]$$

donc ${}^{X}/_{Y}$ est une variable aléatoire réelle sur $(\Omega, \mathfrak{A})$.

32.6 Min et max de deux variables aléatoires
Pour tout $z \in F = E_1 \cup E_2$,

$$\begin{aligned}\big[\max(X,Y)=z\big] &= \big([X=z]\cap[Y\leqslant z]\big)\cup\big([X\leqslant z]\cap[Y=z]\big)\\ \big[\min(X,Y)=z\big] &= \big([X=z]\cap[Y\geqslant z]\big)\cup\big([X\geqslant z]\cap[Y=z]\big)\end{aligned}$$

donc $\min(X,Y)$ et $\max(X,Y)$ sont des variables aléatoires réelles sur $(\Omega,\mathfrak{A})$.

33. Arithmétique et variables aléatoires
Soient X et Y, deux variables aléatoires discrètes sur $(\Omega,\mathfrak{A})$ à valeurs dans $\mathbb{N}^*$ et $\mathbb{Z}$ respectivement.
33.1 Pour tout $\omega \in \Omega$, on note $Q(\omega)$ et $R(\omega)$), le quotient et le reste de la division euclidienne de $Y(\omega)$ par $X(\omega)$:

$$Y(\omega) = Q(\omega)X(\omega) + R(\omega) \quad \text{et} \quad 0 \leqslant R(\omega) < X(\omega).$$

Comme

$$\begin{aligned}&\forall\, q\in\mathbb{Z}, \quad [Q=q] = [qX \leqslant Y < (q+1)X]\\ &\forall\, r \in \mathbb{N}, \quad [R=r] = \bigsqcup_{q\in\mathbb{Z}} \big([Y = qX+r]\cap[X>r]\big),\end{aligned}$$

les fonctions Q et R sont des variables aléatoires sur $(\Omega,\mathfrak{A})$.
33.2 Les fonctions $D = \operatorname{pgcd}(X,Y)$ et $M = \operatorname{ppcm}(X,Y)$ sont des variables aléatoires.

34. Processus de branchement
On étudie l'évolution d'une population, génération après génération.
34.1 On modélise par une variable aléatoire N, à valeurs dans $\mathbb{N}$, le nombre d'individus qui constituent une génération donnée de cette population.
Pour tout entier $k \geqslant 1$, on modélise le nombre de descendants du k-ième individu par une variable X_k à valeurs dans $\mathbb{N}$.
Toutes ces variables aléatoires sont définies sur un même espace probabilisé $(\Omega,\mathfrak{A},\mathbf{P})$.
34.2 Pour tout entier $n \geqslant 1$, la somme

$$S_n = \sum_{k=1}^{n} X_k$$

est une variable aléatoire sur $(\Omega,\mathfrak{A},\mathbf{P})$ à valeurs dans $\mathbb{N}$.
34.3 Le nombre total de descendants est alors défini de la manière suivante : pour tout $\omega \in \Omega$,
— si $N(\omega) = 0$, alors on pose $Y(\omega) = 0$;
— si $N(\omega) \geqslant 1$, alors on pose

$$Y(\omega) = \sum_{k=1}^{N(\omega)} X_k(\omega) \in \mathbb{N}.$$

34.4 Pour tout entier $m \geqslant 1$,

$$[Y=m] = \bigsqcup_{n=1}^{+\infty} \big([N=n]\cap[S_n=m]\big) \in \mathfrak{A}$$

et $[Y=0]\in\mathfrak{A}$, donc Y est une variable aléatoire sur $(\Omega,\mathfrak{A},\mathbf{P})$ à valeurs dans $\mathbb{N}$.

Entraînement

35. Questions pour réfléchir
1.a L'ensemble des variables aléatoires définies sur $(\Omega,\mathfrak{A},\mathbf{P})$ et constantes est une droite vectorielle.
1.b L'ensemble des variables aléatoires presque sûrement constantes est un espace vectoriel de dimension infinie.
2. On suppose que $0 \in E_2$ et que $\mathbf{P}(Y=0)=0$. Le quotient X/Y est alors défini sur le complémentaire d'un événement négligeable. Peut-on considérer ce quotient comme une variable aléatoire sur $(\Omega,\mathfrak{A})$?

36. Décomposition d'une variable aléatoire discrète
Soit $(u_k)_{k\in\mathbb{N}}$, une énumération de l'ensemble dénombrable E. Toute variable aléatoire discrète à valeurs dans E est la somme d'une série de variables de Bernoulli :

$$X = \sum_{k=0}^{+\infty} u_k \mathbb{1}_{[X=u_k]} = \sum_{x\in E} x\mathbb{1}_{[X=x]}.$$

37. Soit X, une variable aléatoire sur $(\Omega,\mathfrak{A},\mathbf{P})$ à valeurs dans $E=\{0,1\}$ dont la loi est caractérisée par

$$\mathbf{P}(X=0)=\mathbf{P}(X=1)={}^1\!/_2.$$

1. La fonction $Y : \Omega \to E$ définie par

$$\forall\, \omega\in\Omega, \quad Y(\omega) = 1 - X(\omega)$$

est une variable aléatoire sur $(\Omega,\mathfrak{A},\mathbf{P})$:

$$[Y=0]=[X=1]\in\mathfrak{A}, \quad [Y=1]=[X=0]\in\mathfrak{A}$$

de même loi que X :

$$\forall\, x\in E, \quad \mathbf{P}(Y=x)=\mathbf{P}(X=x)$$

alors que $\mathbf{P}(X=Y)=0$.
2. L'application $(X,Y) : \Omega \to E\times E$ est une variable aléatoire aléatoire discrète sur $(\Omega,\mathfrak{A},\mathbf{P})$.

38. Soient X et Y, deux variables aléatoires réelles sur $(\Omega,\mathfrak{A})$ à valeurs dans $\mathbb{Z}$.
1. L'ensemble

$$[|X|=n] = [X=n]\sqcup[X=-n]$$

est un événement de $\mathfrak{A}$.
2. L'ensemble

$$[X\leqslant Y] = \bigsqcup_{k\in\mathbb{Z}} [X\leqslant k]\cap[Y=k] = \bigsqcup_{k\in\mathbb{Z}}[X=k]\cap[Y\geqslant k]$$

est un événement de $\mathfrak{A}$.
3. Pour tout $n\in\mathbb{Z}$, l'ensemble

$$\begin{aligned}[X+Y=n] &= \bigsqcup_{m\in\mathbb{Z}} [X=m]\cap[Y=n-m]\\ &= \bigsqcup_{m\in\mathbb{Z}} [X=n-m]\cap[Y=m]\end{aligned}$$

est un événement de $\mathfrak{A}$.
4. Pour tout $n\in\mathbb{Z}$ non nul, l'ensemble

$$[XY=n] = \bigsqcup_{\substack{d\in\mathbb{Z}\\ d|n}} [X=d]\cap[Y={}^n\!/_d] = \bigsqcup_{\substack{d\in\mathbb{Z}\\ d|n}} [X={}^n\!/_d]\cap[Y=d]$$

est un événement, de même que $[XY=0]$.

III

Espérance d'une variable discrète

39. L'**espérance** d'une variable aléatoire, lorsqu'elle existe, est une **moyenne** des valeurs prises par cette variable aléatoire.
Cette moyenne est une fonction de la loi de la variable aléatoire seulement. →**[44]**

40. Soit X, une variable aléatoire discrète définie sur un espace probabilisé $(\Omega,\mathfrak{A},\mathbf{P})$, à valeurs dans une partie dénombrable E de $\mathbb{R}$.

40.1 ✍ *On dit que la variable discrète X est **d'espérance finie** lorsque la famille réelle*

$$\big(x\,\mathbf{P}(X=x)\big)_{x\in E}$$

est sommable.
*On dit aussi que X **admet une espérance**, qu'elle admet un **moment d'ordre** 1 ou qu'elle est **intégrable**.*
40.2 Si $(u_k)_{k\in\mathbb{N}}$ est une énumération de E, alors X admet une espérance si, et seulement si, la série $\sum u_k\,\mathbf{P}(X=u_k)$ est *absolument* convergente.
40.3 ✍ *Si X est une variable discrète d'espérance finie, alors son **espérance** est définie par*

$$\mathbf{E}(X)=\sum_{x\in E} x\,\mathbf{P}(X=x).$$

40.4 ✍ *Une variable aléatoire X est **centrée** lorsqu'elle est d'espérance finie et que son espérance est nulle.*
40.5 ✍ *Par extension, si la variable aléatoire discrète X est presque sûrement positive : $\mathbf{P}(X\geqslant 0)=1$, alors son espérance est définie par*

$$\mathbf{E}(X)=\sum_{x\in E} x\,\mathbf{P}(X=x)\in\mathbb{R}_+\cup\{+\infty\}.$$

40.6 La variable aléatoire X est d'espérance finie si, et seulement si, l'espérance de la variable aléatoire positive $|X|$ est finie.
41. Il faut comprendre le sens de l'hypothèse de sommabilité [**40.1**] : si on choisit deux énumérations $(u_k)_{k\in\mathbb{N}}$ et $(v_k)_{k\in\mathbb{N}}$ de l'ensemble E, les séries $\sum u_k\,\mathbf{P}(X=u_k)$ et $\sum v_k\,\mathbf{P}(X=v_k)$ doivent être de même nature et avoir mêmes sommes. →[**5.29.5**]
42. Une variable aléatoire X à valeurs dans $E=\mathbb{N}^*$ dont la loi est caractérisée par

$$\forall\, n\geqslant 1,\quad \mathbf{P}(X=n)=\frac{6}{\pi^2 n^2}$$

n'est pas intégrable (elle est d'espérance infinie).
43. Espérances des lois usuelles
Les variables aléatoires qui suivent l'une des lois usuelles sont toutes d'espérance finie.

Loi de X	Espérance de X
Dirac δ_{x_0}	x_0
Bernoulli $\mathcal{B}(p)$	p
Binomiale $\mathcal{B}(n,p)$	np
Uniforme sur $\{x_1,\dots,x_n\}$	$(x_1+\cdots+x_n)/n$
Géométrique $\mathcal{G}(p)$	$1/p$
Poisson $\mathcal{P}(\lambda)$	λ

44. L'espérance d'une variable aléatoire X dépend de la loi de X et *seulement de la loi* de X.
44.1 ➞ *Soient X et Y, deux variables aléatoires discrètes de même loi. Alors X est d'espérance finie si, et seulement si, Y est d'espérance finie et dans ce cas, $\mathbf{E}(X)=\mathbf{E}(Y)$.*
44.2 En conditionnant par un événement $A\in\mathfrak{A}$, on peut modifier la loi de X [**11**] et, par conséquent, son espérance.
45. Espérance conditionnelle sachant un événement
Soient X, une variable aléatoire discrète sur l'espace probabilisé $(\Omega,\mathfrak{A},\mathbf{P})$ et $A\in\mathfrak{A}$, un événement non négligeable.
45.1 Si X est une variable aléatoire d'espérance finie, alors la famille

$$\big(x\,\mathbf{P}_A(X=x)\big)_{x\in E}$$

est sommable.
45.2 ✍ *Soit X, une variable aléatoire d'espérance finie. L'**espérance conditionnelle de** X **sachant** $A\in\mathfrak{A}$ est définie par*

$$\mathbf{E}_A(X)=\sum_{x\in E} x\,\mathbf{P}_A(X=x).$$

Cette espérance est aussi notée $\mathbf{E}(X|A)$.
45.3 L'espérance conditionnelle de X sachant A est l'espérance de X considérée, non plus comme une variable aléatoire sur $(\Omega,\mathfrak{A},\mathbf{P})$, mais comme une variable aléatoire sur $(\Omega,\mathfrak{A},\mathbf{P}_A)$, ce qui modifie, en général, la loi de X [**11**] et son espérance [**44**].
45.4 Selon l'événement A choisi, l'espérance conditionnelle de X peut être supérieure ou inférieure à l'espérance de X. →[**68**]
45.5

$$\mathbf{E}(X\mathbb{1}_A)=\mathbf{E}_A(X)\,\mathbf{E}(\mathbb{1}_A)=\mathbf{E}(X|A)\,\mathbf{P}(A)$$

46. Formule de l'espérance totale
Soit $X:\Omega\to E$, une variable aléatoire discrète d'espérance finie et $(A_k)_{k\in\mathbb{N}}$, un système complet d'événements.
46.1 La famille définie par

$$\forall\, k\in\mathbb{N},\ \forall\, x\in E,\quad u_{k,x}=x\,\mathbf{P}([X=x]\cap A_k)$$

est sommable.
46.2 La famille

$$\big(\mathbf{E}(X|A_k)\,\mathbf{P}(A_k)\big)_{k\in\mathbb{N}}$$

est sommable et

$$\sum_{k\in\mathbb{N}}\mathbf{E}(X|A_k)\,\mathbf{P}(A_k)=\mathbf{E}(X).$$

III.1 Intégrabilité d'une variable aléatoire

47. ➞ *Une variable aléatoire presque sûrement bornée est d'espérance finie et*

$$\mathbf{P}(|X|\leqslant k)=1\implies \big|\mathbf{E}(X)\big|\leqslant k.$$

48. Théorème de comparaison
On considère une variable aléatoire discrète positive Y d'espérance finie et une variable aléatoire discrète réelle X, définie sur le même espace probabilisé $(\Omega,\mathfrak{A},\mathbf{P})$ que Y, telle que

$$\mathbf{P}(|X|\leqslant Y)=1.$$

On note $E\subset\mathbb{R}$ et $F\subset\mathbb{R}_+$, deux parties finies ou dénombrables telles que $\mathbf{P}(X\in E)=\mathbf{P}(Y\in F)=1$.
48.1 La famille

$$\big([X=x,Y=y]\big)_{(x,y)\in E\times F}$$

est un système complet d'événements.
48.2

$$\forall\, x\in E,\quad \mathbf{P}(X=x)=\sum_{y\in F\,:\,y\geqslant|x|}\mathbf{P}(X=x,Y=y).$$

48.3

$$\forall\, y\in F,\quad y\,\mathbf{P}(Y=y)=\sum_{x\in E\,:\,|x|\leqslant y} y\,\mathbf{P}(X=x,Y=y).$$

48.4

$$\forall\, x\in E,\quad |x|\ \mathbf{P}(X=x)\leqslant\sum_{y\in F\,:\,y\geqslant|x|} y\,\mathbf{P}(X=x,Y=y).$$

48.5 ➞ *Soient X et Y, deux variables aléatoires réelles définies sur le même espace probabilisé $(\Omega,\mathfrak{A},\mathbf{P})$. Si Y est une variable aléatoire d'espérance finie et si*

$$\mathbf{P}(|X|\leqslant Y)=1,$$

alors X est une variable aléatoire d'espérance finie et de plus

$$\mathbf{E}(|X|)\leqslant\mathbf{E}(Y).$$

49. Formule de transfert
Soit X, une variable aléatoire discrète sur $(\Omega, \mathfrak{A}, \mathbf{P})$ à valeurs dans un ensemble dénombrable E et $f : E \to F$, une application.
Il n'est pas nécessaire de calculer explicitement la loi de la variable aléatoire $Y = f(X)$ pour vérifier son intégrabilité ou pour calculer son espérance.
49.1

$$\forall\, y \in F, \quad y\,\mathbf{P}(Y = y) = \sum_{\substack{x \in E \\ f(x) = y}} f(x)\,\mathbf{P}(X = x).$$

49.2 ➞ *Soit X, une variable aléatoire discrète sur $(\Omega, \mathfrak{A}, \mathbf{P})$ à valeurs dans un ensemble dénombrable E.*
Alors $f(X)$ est une variable aléatoire d'espérance finie si, et seulement si, la famille

$$\big(f(x)\,\mathbf{P}(X = x)\big)_{x \in E}$$

est sommable et, dans ce cas,

$$\mathbf{E}\big(f(X)\big) = \sum_{x \in E} f(x)\,\mathbf{P}(X = x).$$

49.3 Le théorème [**49.2**] s'applique à *toutes* les variables aléatoires discrètes, qu'elles soient à valeurs réelles ou à valeurs vectorielles.
Par exemple, si X_1 et X_2 sont deux variables aléatoires à valeurs dans $\mathbb{Z}$, alors $X = (X_1, X_2)$ est une variable à valeurs dans $\mathbb{Z}^2$ et on peut appliquer le théorème de transfert à $f(X) = X_1 + X_2$, qui est une variable discrète à valeurs dans $\mathbb{Z}$.

III.2 Propriétés de l'espérance

50. Les propriétés fondamentales de l'espérance sont aussi celle des sommes et des intégrales mais, en outre, l'espérance doit être interprétée comme une moyenne [**51**].

51. ➞ *Une variable aléatoire presque sûrement constante est d'espérance finie et*

$$\mathbf{P}(X = k) = 1 \implies \mathbf{E}(X) = k.$$

Linéarité de l'espérance

52. ➞ *Si X et Y sont deux variables aléatoires réelles d'espérance finie définies sur un même espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$, alors toute combinaison linéaire $Z = \lambda X + Y$ est une variable aléatoire d'espérance finie et*

$$\mathbf{E}(\lambda X + Y) = \lambda\,\mathbf{E}(X) + \mathbf{E}(Y).$$

53. ➞ *L'ensemble $\mathscr{L}^1(\Omega)$ des variables aléatoires réelles d'espérance finie est un sous-espace vectoriel de l'espace $\mathscr{A}(\Omega, \mathbb{R})$ des applications de Ω dans $\mathbb{R}$.*

54. Centrage d'une variable aléatoire
Soit X, une variable aléatoire réelle d'espérance finie.
54.1 La variable aléatoire réelle $aX + b$ est d'espérance finie quels que soient les réels a et b et

$$\mathbf{E}(aX + b) = a\,\mathbf{E}(X) + b.$$

54.2 La variable aléatoire réelle $X - \mathbf{E}(X)$ est centrée.

Positivité de l'espérance

55. ➞ *Soit X, une variable aléatoire presque sûrement positive :*

$$\mathbf{P}(X \geqslant 0) = 1.$$

Alors l'espérance (finie ou non) de X est positive et elle est nulle si, et seulement si, X est presque sûrement nulle :

$$\mathbf{E}(X) = 0 \iff \mathbf{P}(X = 0) = 1.$$

56. ▷ *Soient X et Y, deux variables aléatoires sur $(\Omega, \mathfrak{A}, \mathbf{P})$ qui admettent une espérance et telles que*

$$\mathbf{P}(X \leqslant Y) = 1.$$

Alors $\mathbf{E}(X) \leqslant \mathbf{E}(Y)$ et avec égalité si, et seulement si, X et Y sont presque sûrement égales :

$$\mathbf{E}(X) = \mathbf{E}(Y) \iff \mathbf{P}(X = Y) = 1.$$

57. ▷ *S'il existe deux réels m et M tels que*

$$\mathbf{P}(m \leqslant X \leqslant M) = 1,$$

alors X est une variable aléatoire d'espérance finie et

$$m \leqslant \mathbf{E}(X) \leqslant M.$$

58. ➞ *Si X est une variable aléatoire d'espérance finie, alors*

$$\big|\,\mathbf{E}(X)\big| \leqslant \mathbf{E}(|X|).$$

59. Inégalité de Markov
L'**inégalité de Markov** relie l'espérance d'une variable aléatoire avec sa fonction de répartition [**109**].
59.1 Pour tout $t \in \mathbb{R}_+$ et toute variable aléatoire positive U,

$$\forall\, \omega \in \Omega, \quad t \cdot \mathbb{1}_{(U \geqslant t)}(\omega) \leqslant U(\omega) \cdot \mathbb{1}_{(U \geqslant t)}(\omega) \leqslant U(\omega).$$

59.2 ➞ *Pour toute variable aléatoire réelle X,*

$$\forall\, t > 0, \qquad \mathbf{P}(|X| \geqslant t) \leqslant \frac{\mathbf{E}(|X|)}{t}.$$

59.3 Comme

$$\big[|X| > 0\big] = \bigcup_{n \geqslant 1} \big[|X| \geqslant {}^1\!/_n\big],$$

le cas d'égalité de [**55**] peut être vu comme une conséquence de l'inégalité de Markov.
59.4 Pour toute variable aléatoire réelle, la suite de terme général $\mathbf{P}(X \geqslant n)$ tend vers 0 en décroissant et l'inégalité de Markov donne une estimation de son ordre de grandeur.

Entraînement

60. Questions pour réfléchir
1. Comparer la définition de l'espérance d'une variable aléatoire discrète avec la décomposition [**36**].
2. Comparer la formule [**45.5**] avec [**13.26.5**].
3. Une variable aléatoire qui ne prend qu'un nombre fini de valeurs est une variable d'espérance finie.
4. Si X est une variable aléatoire à valeurs dans $E = \mathbb{N}^*$, alors ${}^1\!/_X$ est une variable aléatoire d'espérance finie.
5. *Suite de* [**55**] – L'inégalité $\mathbf{E}(X) \leqslant \mathbf{E}(Y)$ n'implique pas $\mathbf{P}(X \leqslant Y) = 1$.
6. Si X est une variable aléatoire presque sûrement positive et d'espérance nulle, alors elle suit la loi de Dirac δ_0.

61. *Suite de* [**18.2**] – Les événements A et B sont indépendants si, et seulement si, $\mathbf{E}(\mathbb{1}_A \mathbb{1}_B) = \mathbf{E}(\mathbb{1}_A)\,\mathbf{E}(\mathbb{1}_B)$.

62. *Suite de* [**25**] – La variable aléatoire X est intégrable si, et seulement si, $\alpha > 2$ et dans ce cas

$$\mathbf{E}(X) = \frac{\zeta(\alpha - 1)}{\zeta(\alpha)}.$$

63. Soit X, une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$, à valeurs dans $\mathbb{N}$.
63.1 Les propositions suivantes sont équivalentes.
1. La variable X est une variable aléatoire d'espérance finie.
2. La famille $\big(\mathbf{P}(X = n)\mathbb{1}_{(0 \leqslant k < n)}\big)_{(k,n) \in \mathbb{N}^2}$ est sommable.
3. La série $\sum \mathbf{P}(X > n)$ est convergente.

63.2 Si X est une variable aléatoire d'espérance finie, alors

$$\mathbf{E}(X) = \sum_{n=0}^{+\infty} \mathbf{P}(X > n).$$

63.3 En déduire l'espérance d'une variable aléatoire suivant la loi géométrique $\mathcal{G}(p)$.

64. Soit X, une variable aléatoire à valeurs dans $\mathbb{N}$. Lorsque n tend vers $+\infty$,

$$\sum_{k=1}^{n} \mathbf{P}(X < k) \sim n$$

et si X est intégrable, alors

$$\sum_{k=1}^{n} \mathbf{P}(X < k) = n - \mathbf{E}(X) + o(1).$$

65. Distribution asymptotique d'une variable intégrable
65.1 Pour toute variable aléatoire réelle X,

$$\lim_{n\to+\infty} \mathbf{P}(|X| \geqslant n) = 0.$$

65.2 On suppose que X est une variable d'espérance finie à valeurs dans l'ensemble (fini ou dénombrable) E. Pour tout entier $n \in \mathbb{N}$, on pose

$$E_n = \{x \in E \,:\, n \leqslant |x| < n+1\} \quad \text{et} \quad \sigma_n = \sum_{x\in E_n} |x|\,\mathbf{P}(X = x).$$

Alors la série $\sum \sigma_n$ est convergente et comme

$$n\,\mathbf{P}(|X| \geqslant n) = n \sum_{k=n}^{+\infty} \mathbf{P}(X \in E_k) \leqslant \sum_{k=n}^{+\infty} \sigma_k,$$

alors

$$\lim_{n\to+\infty} n\,\mathbf{P}(|X| \geqslant n) = 0.$$

Comparer avec l'inégalité de Markov [**59.2**] et avec [**63**].

66. Soit f, une fonction de $\mathbb{R}$ dans $\mathbb{R}$.
66.1 Soit X, une variable aléatoire réelle sur $(\Omega, \mathfrak{A}, \mathbf{P})$.
1. Si le **support** de la loi de X est fini :

$$\exists\, E_0 \in \mathfrak{P}_0(\mathbb{R}), \quad \mathbf{P}(X \in E_0) = 1,$$

alors X et $f(X)$ admettent une espérance.
2. On suppose que la variable aléatoire discrète X est presque sûrement bornée :

$$\exists\, M > 0, \quad \mathbf{P}(|X| \leqslant M) = 1.$$

2.a La variable X admet une espérance.
2.b Si f est continue sur $\mathbb{R}$, alors $f(X)$ admet une espérance.
2.c Si f n'est pas continue sur $\mathbb{R}$, alors $f(X)$ n'admet pas nécessairement une espérance.

66.2 On suppose que

$$\mathbf{E}[f(X)] = f(\mathbf{E}[X])$$

pour toute variable aléatoire X de support fini. Alors

$$\forall\, x \in \mathbb{R},\ \forall\, t \in [0,1], \quad tf(x) + (1-t)f(0) = f(tx)$$

et f est une fonction affine.

67. Inégalité de Jensen
Soient $\varphi : \mathbb{R} \to \mathbb{R}$, une fonction convexe et X, une variable aléatoire d'espérance finie. On suppose que la variable $\varphi(X)$ est d'espérance finie.
Il existe un réel a tel que [**1.27**]

$$\forall\, x \in \mathbb{R}, \quad \varphi(x) \geqslant a\big[x - \mathbf{E}(X)\big] + \varphi\big(\mathbf{E}(X)\big)$$

donc

$$\mathbf{E}\big(\varphi(X)\big) \geqslant \varphi\big(\mathbf{E}(X)\big).$$

En particulier, $\mathbf{E}(X^2) \geqslant [\mathbf{E}(X)]^2$ pour toute variable aléatoire d'espérance finie.

68. *Suite de* [**45.5**] – Soit X, une variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P})$ qui suit une loi de Poisson de paramètre $\lambda = 1$.

68.1 Avec $A = [X \geqslant 2] \in \mathfrak{A}$, on obtient

$$\mathbf{E}(X.\mathbb{1}_A) \geqslant 2\,\mathbf{P}(A) \geqslant \mathbf{E}(X)\,\mathbf{P}(A)$$

c'est-à-dire $\mathbf{E}_A(X) \geqslant \mathbf{E}(X)$.
68.2 Avec $B = [X \leqslant 1] \in \mathfrak{A}$, la loi conditionnelle de X sachant B est une loi de Bernoulli et $\mathbf{E}_B(X) \leqslant \mathbf{E}(X)$.

69.1 Il existe une variable aléatoire X à valeurs dans $\mathbb{Z}^*$ définie sur un espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$, telle que

$$\forall\, n \in \mathbb{N}^*, \quad \mathbf{P}(X = -n) = \mathbf{P}(X = n) = \frac{1}{2n(n+1)}.$$

69.2 Pour tout $n \in \mathbb{N}$, la variable aléatoire X admet une espérance conditionnelle sachant

$$A_n = [X = -n] \cup [X = n] \in \mathcal{A}$$

et la série $\sum \mathbf{E}_{A_n}(X)\,\mathbf{P}(A_n)$ est absolument convergente.
69.3 Cependant, et bien que $(A_n)_{n\in\mathbb{N}}$ soit un système complet d'événements, la variable aléatoire X n'est pas d'espérance finie.

70. Grandes déviations et inégalité de Chernoff
Soit S_n, une variable aléatoire suivant la loi binomiale $\mathscr{B}(n,p)$ avec $n \geqslant 1$ et $0 < p < 1$.
1. Pour tout $a > 0$,

$$\mathbf{P}\Big(\frac{S_n}{n} \geqslant a\Big) \leqslant e^{-n\sigma(a,p)}$$

où $\sigma(a,p) = \sup_{s\in\mathbb{R}_+} sa - \ell\text{n}(q + pe^s)$.
2. Il existe une fonction $h_1 : \mathbb{R}_+^* \to \mathbb{R}_+^*$ telle que

$$\forall\, n \geqslant 1,\ \forall\, \varepsilon > 0, \quad \mathbf{P}\Big(\frac{S_n}{n} \geqslant p + \varepsilon\Big) \leqslant e^{-nh_1(\varepsilon)}.$$

Pour $0 < \varepsilon < q$, on peut choisir

$$h_1(\varepsilon) = \sup_{s>0}\big[s(p+\varepsilon) - \ell\text{n}(q + pe^s)\big].$$

3. Comme $(n - S_n)$ suit la loi binomiale $\mathscr{B}(n,q)$, il existe une fonction $h_2 : \mathbb{R}_+^* \to \mathbb{R}_+^*$ telle que

$$\forall\, n \geqslant 1,\ \forall\, \varepsilon > 0, \quad \mathbf{P}\Big(\frac{S_n}{n} \leqslant p - \varepsilon\Big) \leqslant e^{-nh_2(\varepsilon)}.$$

4. Il existe une fonction $h : \mathbb{R}_+ \to \mathbb{R}_+^*$ telle que

$$\forall\, n \geqslant 1,\ \forall\, \varepsilon > 0, \quad \mathbf{P}\Big(\Big|\frac{S_n}{n} - p\Big| \geqslant \varepsilon\Big) \leqslant 2e^{-nh(\varepsilon)}.$$

5. Soit $\varepsilon > 0$, fixé. Comme la série $\sum e^{-nh(\varepsilon)}$ est convergente, la probabilité pour que

$$\forall\, n \geqslant n_0, \quad p - \varepsilon \leqslant \frac{S_n}{n} \leqslant p + \varepsilon$$

tend vers 1 lorsque n_0 tend vers $+\infty$ [**13.23**].

IV

Variance d'une variable discrète

71. L'*espérance* d'une variable aléatoire X est un **indicateur de position**, qui doit être comprise comme une *valeur moyenne*.
La **variance** de X, qui est aussi définie comme une valeur moyenne, est un **indicateur de dispersion**, qui donne une idée de la manière dont les valeurs prises par X sont réparties de part et d'autre de $\mathbf{E}(X)$.

IV.1 Moments d'une variable aléatoire

72. ✍ *Soit X, une variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P})$.*
72.1 *On dit que X **admet un moment d'ordre** $\alpha \in \mathbb{N}^*$ lorsque la variable X^α est d'espérance finie.*
72.2 *Dans ce cas, le **moment d'ordre** α de X est $\mathbf{E}(X^\alpha)$.*

73. Soit $\alpha \in \mathbb{N}^*$, quelconque.
73.1 Une variable aléatoire presque sûrement bornée admet un moment d'ordre α.
73.2 Une variable aléatoire qui suit une loi géométrique admet un moment d'ordre α.
73.3 Une variable aléatoire qui suit une loi de Poisson admet un moment d'ordre α.

74. Soient $0 < \alpha < \beta$, deux entiers.
74.1

$$\forall\, \omega \in \Omega, \quad 0 \leqslant \big|X(\omega)\big|^\alpha \leqslant 1 + \big|X(\omega)\big|^\beta.$$

74.2 → *Si une variable aléatoire X admet un moment d'ordre $\beta \in \mathbb{N}^*$, alors elle admet un moment d'ordre α pour tout entier $0 < \alpha \leqslant \beta$.*
74.3 En particulier, si X admet un moment d'ordre 2, alors elle admet un moment d'ordre 1. →**[67]**

75.1 ✍ *Pour $n \in \mathbb{N}$, une variable aléatoire X* admet un moment factoriel d'ordre *n lorsque la variable*

$$F_n = X(X-1)\cdots(X-n+1)$$

est d'espérance finie.
*Dans ce cas, le **moment factoriel d'ordre** n de X est égal à*

$$\mathbf{E}[X(X-1)\cdots(X-n+1)].$$

75.2 Pour tout entier $n \geqslant 1$,

$$\forall\, \omega \in \Omega, \quad \big|F_n(\omega)\big| \leqslant (n+1)! + \big|F_{n+1}(\omega)\big|.$$

75.3 Pour tout entier $n \geqslant 1$, une variable aléatoire admet un moment factoriel d'ordre n si, et seulement si, elle admet un moment d'ordre n.
75.4 Les moments factoriels sont souvent plus simples à calculer que les moments eux-mêmes et permettent de trouver ces moments.

$$\mathbf{E}(X^2) = \mathbf{E}[X(X-1)] + \mathbf{E}(X)$$
$$\mathbf{E}(X^3) = \mathbf{E}[X(X-1)(X-2)] + 3\,\mathbf{E}[X(X-1)] + \mathbf{E}(X)\ldots$$

Variables de carré intégrable

76. ✍ *L'ensemble des variables aléatoires réelles sur $(\Omega, \mathfrak{A}, \mathbf{P})$ qui admettent un moment d'ordre 2, dites aussi **variables de carré intégrable** ou **de variance finie**, est noté $\mathscr{L}^2(\Omega)$.*

77. Soient X et Y, deux variables aléatoires de $\mathscr{L}^2(\Omega)$.
77.1

$$\forall\, \omega \in \Omega, \quad \big|X(\omega)Y(\omega)\big| \leqslant \frac{X^2(\omega) + Y^2(\omega)}{2}.$$

77.2 → *L'ensemble $\mathscr{L}^2(\Omega)$ des variables aléatoires réelles de carré intégrable est un sous-espace vectoriel de l'espace $\mathscr{L}^1(\Omega)$ des variables aléatoires réelles d'espérance finie et contient l'espace des variables aléatoires presque sûrement bornées.*

$$\mathscr{L}^\infty(\Omega) \subset \mathscr{L}^2(\Omega) \subset \mathscr{L}^1(\Omega)$$

77.3 Une variable aléatoire X à valeurs dans $\mathbb{N}$ telle que

$$\mathbf{P}(X = n) \sim {}^1\!/_{n^3}$$

lorsque n tend vers $+\infty$ est une variable d'espérance finie qui n'est pas de carré intégrable.

77.4 → **Inégalité de Schwarz**
Si X et Y sont deux variables aléatoires réelles de carré intégrable, alors le produit XY est une variable aléatoire intégrable et

$$\big|\,\mathbf{E}(XY)\big| \leqslant \sqrt{\mathbf{E}(X^2)}\sqrt{\mathbf{E}(Y^2)}.$$

77.5 Soient X et Y, deux variables aléatoires réelles de carré intégrable telles que

$$\big|\,\mathbf{E}(XY)\big| = \sqrt{\mathbf{E}(X^2)}\sqrt{\mathbf{E}(Y^2)}.$$

Alors : ou bien X est presque sûrement nulle, ou bien il existe $a \in \mathbb{R}$ tel que $\mathbf{P}(Y = aX) = 1$.

IV.2 Variance

78. Soit $X \in \mathscr{L}^2(\Omega)$.
78.1 La variable $X - \mathbf{E}(X)$ admet un moment d'ordre 2.
78.2 ✍ *La **variance** de $X \in \mathscr{L}^2(\Omega)$ est définie par*

$$\mathbf{V}(X) = \mathbf{E}\big(\big[X - \mathbf{E}(X)\big]^2\big).$$

78.3 → **Formule de Koenig–Huyghens**

$$\forall\, X \in \mathscr{L}^2(\Omega), \quad \mathbf{V}(X) = \mathbf{E}(X^2) - \big(\mathbf{E}(X)\big)^2.$$

78.4 ✍ *L'**écart type** de $X \in \mathscr{L}^2(\Omega)$ est défini par*

$$\sigma(X) = \sqrt{\mathbf{V}(X)}.$$

78.5 Le seul intérêt de l'écart type est d'avoir la même dimension que la variable aléatoire X (ce qui n'est pas le cas de la variance).

79. Comme l'espérance, la variance ne dépend que de la loi d'une variable aléatoire.
79.1 Si deux variables aléatoires X et Y ont même loi, alors X admet un moment d'ordre 2 si, et seulement si, Y admet un moment d'ordre 2 et dans ce cas, $\mathbf{V}(X) = \mathbf{V}(Y)$.
79.2 En conditionnant par un événement $A \in \mathfrak{A}$, on modifie en général la loi de X et sa variance.
79.3 Si X admet un moment d'ordre 2 en tant que variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P})$, alors X admet un moment d'ordre 2 en tant que variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P}_A)$.
79.4 ✍ *Si X admet un moment d'ordre 2 et si l'événement A n'est pas négligeable, la **variance conditionnelle de X sachant** A est définie par*

$$\mathbf{V}_A(X) = \mathbf{E}_A\big([X - \mathbf{E}_A(X)]^2\big).$$

79.5 Il s'agit de la variance de X considérée comme une variable aléatoire sur $(\Omega, \mathfrak{A}, \mathbf{P}_A)$ et la formule de Koenig–Huyghens est encore vraie :

$$\mathbf{V}_A(X) = \mathbf{E}_A(X^2) - \big(\mathbf{E}_A(X)\big)^2.$$

80. Variances des lois usuelles
Les variables aléatoires qui suivent l'une des lois usuelles admettent toutes un moment d'ordre deux.
Comme de coutume, on pose ici $q = 1 - p$.

Loi de X	Variance de X
Dirac δ_{x_0}	0
Bernoulli $\mathcal{B}(p)$	pq
Binomiale $\mathcal{B}(n,p)$	npq
Géométrique $\mathcal{G}(p)$	${}^q\!/_{p^2}$
Poisson $\mathcal{P}(\lambda)$	λ

81. Réduction d'une variable aléatoire

81.1 → *Soit $X \in \mathscr{L}^2(\Omega)$. Quels que soient les réels a et b,*

$$\mathbf{V}(aX+b) = a^2\,\mathbf{V}(X).$$

81.2 ✍ *La variable aléatoire $X \in \mathscr{L}^2(\Omega)$ est **réduite** lorsque sa variance est égale à 1.*

81.3 Si $\sigma(X) > 0$, alors la variable aléatoire

$$X^* = \frac{X - \mathbf{E}(X)}{\sigma(X)}$$

est une variable aléatoire de carré intégrable, centrée et réduite. →[**83.3**]

82. → **Inégalité de Bienaymé–Tchebychev**

Si X est une variable aléatoire réelle de carré intégrable, alors

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

83. Caractérisation des variables aléatoires constantes

83.1 L'écart type d'une variable aléatoire presque sûrement constante est nul.

83.2 Si X est une variable aléatoire d'espérance finie, alors

$$\mathbf{P}\big(X = \mathbf{E}(X)\big) = \lim_{\varepsilon\to 0} \mathbf{P}\big(\big|X - \mathbf{E}(X)\big| < \varepsilon\big).$$

83.3 → *Si X est une variable aléatoire réelle de carré intégrable dont la variance est nulle, alors X est presque sûrement constante.*

$$\mathbf{V}(X) = 0 \implies \mathbf{P}\big(X = \mathbf{E}(X)\big) = 1$$

IV.3 Covariance

84. La **covariance** de deux variables aléatoires réelles X et Y donne une idée de la manière dont ces variables fluctuent l'une par rapport à l'autre autour de leurs moyennes respectives.

84.1 Si X et Y sont deux variables aléatoires réelles de carré intégrable, alors elles admettent toutes deux une espérance et les produits XY et $[X - \mathbf{E}(X)][Y - \mathbf{E}(Y)]$ admettent aussi une espérance.

84.2 ✍ *Si X et Y sont deux variables aléatoires réelles qui admettent un moment d'ordre 2, leur **covariance** est définie par*

$$\mathbf{Cov}(X,Y) = \mathbf{E}\big[\big(X - \mathbf{E}(X)\big)\big(Y - \mathbf{E}(Y)\big)\big].$$

En particulier, $\mathbf{Cov}(X,X) = \mathbf{V}(X)$.

84.3 ✍ *Deux variables aléatoires réelles sont **corrélées** lorsque leur covariance n'est pas nulle.*

84.4 S'il existe deux réels a et b tels que $Y = aX + b$, alors

$$\mathbf{Cov}(X,Y) = a\,\mathbf{V}(X).$$

Dans ce cas, la covariance est du signe de a.

85. → **Formule de Koenig–Huyghens**

Si X et Y sont deux variables aléatoires réelles de carré intégrable,

$$\mathbf{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\,\mathbf{E}(Y).$$

86. L'application $\mathbf{Cov}$ est une forme bilinéaire symétrique sur l'espace des variables aléatoires de carré intégrable.

$$\mathbf{Cov}\Big(\sum_{i=1}^n a_i X_i, \sum_{j=1}^p b_j Y_j\Big) = \sum_{i=1}^n \sum_{j=1}^p a_i b_j\,\mathbf{Cov}(X_i, Y_j)$$

87. Inégalité de Schwarz et cas d'égalité

Soient X et Y, deux variables aléatoires réelles de carré intégrable.

87.1
$$\big|\,\mathbf{Cov}(X,Y)\big| \leqslant \sigma(X)\sigma(Y)$$

87.2 Si la variable aléatoire X n'est pas presque sûrement constante et si

$$\big|\,\mathbf{Cov}(X,Y)\big| = \sigma(X)\sigma(Y),$$

alors il existe $(a,b) \in \mathbb{R}^2$ tels que $\mathbf{P}(Y = aX + b) = 1$. →[**84.4**]

88. La covariance est la *forme polarisée* de la variance considérée comme une *forme quadratique*.

88.1 → *Si $X_1, \ldots, X_n$ sont des variables aléatoires réelles de carré intégrable, alors*

$$\mathbf{V}(X_1 + \cdots + X_n) = \sum_{k=1}^n \mathbf{V}(X_k) + 2 \sum_{1\leqslant k<\ell\leqslant n} \mathbf{Cov}(X_k, X_\ell).$$

88.2 ▷ *Si $X_1, \ldots, X_n$ sont des variables aléatoires réelles de carré intégrable deux à deux non corrélées, alors*

$$\mathbf{V}(X_1 + \cdots + X_n) = \sum_{k=1}^n \mathbf{V}(X_k).$$

88.3 Quels que soient les réels $a_1, \ldots, a_n$, b,

$$\mathbf{V}\Big(\sum_{k=1}^n a_k X_k + b\Big) = \sum_{k=1}^n a_k^2\,\mathbf{V}(X_k) + 2 \sum_{1\leqslant k<\ell\leqslant n} a_k a_\ell\,\mathbf{Cov}(X_k, X_\ell).$$

89. Matrice de covariance

Soit $X = (X_k)_{1\leqslant k\leqslant n}$, une famille de variables aléatoires de carré intégrable définies sur un même espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$.

89.1 ✍ *La **matrice de covariance** du vecteur aléatoire X est définie par*

$$\mathbf{Cov}(X) = \big(\mathbf{Cov}(X_i, X_j)\big)_{1\leqslant i,j\leqslant n} \in \mathfrak{M}_n(\mathbb{R}).$$

89.2 La matrice de covariance $\mathbf{Cov}(X)$ est symétrique réelle.

89.3 Quelle que soit la matrice $A \in \mathfrak{M}_{p,n}(\mathbb{R})$, la variable aléatoire

$$Y = AX$$

est un vecteur aléatoire de carré intégrable et

$$\mathbf{E}(Y) = A\,\mathbf{E}(X), \qquad \mathbf{Cov}(Y) = A\,\mathbf{Cov}(X)\,{}^tA.$$

89.4 Pour toute matrice colonne $U = (u_i)_{1\leqslant i\leqslant n} \in \mathfrak{M}_{n,1}(\mathbb{R})$,

$${}^tU\,\mathbf{Cov}(X)U = \mathbf{V}\Big(\sum_{i=1}^n u_i X_i\Big) \geqslant 0.$$

89.5 Les valeurs propres de la matrice de covariance $\mathbf{Cov}(X)$ sont positives.

89.6 D'après le théorème spectral, la matrice $\mathbf{Cov}(X)$ admet 0 pour valeur propre si, et seulement si, il existe des scalaires u_1, …, u_n non tous nuls tels que la variable aléatoire

$$Y = u_1 X_1 + \cdots + u_n X_n$$

soit presque sûrement constante.

Entraînement

90. Questions pour réfléchir

1. Retrouver facilement l'équivalence [**75.3**] lorsque X est une variable aléatoire discrète à valeurs dans $\mathbb{N}$.
2. Si X est une variable aléatoire réelle de carré intégrable, alors

$$\mathbf{V}(X) = \min_{a\in\mathbb{R}} \mathbf{E}\big[(X-a)^2\big].$$

3. Si la variable aléatoire X admet un moment d'ordre deux, alors la variable aléatoire $|X|$ admet un moment d'ordre deux et

$$\mathbf{V}(|X|) \leqslant \mathbf{V}(X).$$

Comment justifier intuitivement cette inégalité ?

4. Si X est une variable aléatoire réelle de carré intégrable, alors

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

5. Soit X, une variable aléatoire strictement positive qui admet une espérance.
Si $1/X$ admet une espérance et si $\mathbf{E}(1/X) = 1/\,\mathbf{E}(X)$, alors la variable X est presque sûrement constante. →[77.5]
6. $\mathbf{Cov}(X,X) = \mathbf{V}(X)$
7. Si X et Y sont deux variables aléatoires de carré intégrable telles que $\mathbf{V}(X) = \mathbf{V}(Y)$, alors les variables $X+Y$ et $X-Y$ ne sont pas corrélées.
8. Si $\mathbf{Cov}(X,Y) < 0$, alors $\mathbf{V}(X+Y) < \mathbf{V}(X-Y)$.
9. Deux variables aléatoires X et Y ne peuvent pas vérifier les propriétés suivantes :

$$\mathbf{E}(X) = 3,\ \mathbf{E}(Y) = 2,\ \mathbf{E}(X^2) = 10,\ \mathbf{E}(Y^2) = 29,\ \mathbf{E}(XY) = 0.$$

91. Si X est une variable aléatoire à valeurs dans $\{1,\dots,n\}$, alors la loi de X est déterminée par les valeurs de $\mathbf{E}(X^k)$ pour $1 \leqslant k < n$.

92. Soit X, une variable aléatoire définie sur $(\Omega, \mathfrak{A}, \mathbf{P})$, qui suit la loi de Poisson de paramètre λ. Pour tout $\omega \in \Omega$, on pose

$$Y(\omega) = \left|\begin{array}{ll} X(\omega)/2 & \text{si } X(\omega) \text{ est pair,} \\ 0 & \text{sinon.} \end{array}\right.$$

Alors Y est une variable aléatoire discrète sur $(\Omega, \mathfrak{A}, \mathbf{P})$ et

$$\mathbf{E}(Y) = \frac{\lambda}{2} e^{-\lambda} \operatorname{sh} \lambda,$$
$$\mathbf{V}(Y) = \frac{\lambda e^{-\lambda}}{4} \big[\lambda \operatorname{ch} \lambda + \operatorname{sh} \lambda - \lambda e^{-\lambda} \operatorname{sh}^2 \lambda\big].$$

93. Si la variable aléatoire S_n suit la loi binomiale $\mathcal{B}(n,p)$, alors

$$\mathbf{P}\Big(\Big|\frac{S_n}{n} - p\Big| \geqslant \varepsilon\Big) \leqslant \frac{1}{4n\varepsilon^2}$$

pour tout $\varepsilon > 0$.

94. La covariance ne met pas en évidence toutes les formes de liaison entre deux variables aléatoires. Par exemple, une variable aléatoire X dont la loi est caractérisée par

$$\mathbf{P}(X = -1) = \mathbf{P}(X = 0) = \mathbf{P}(X = 1) = 1/3$$

et la variable aléatoire définie par $Y = X^2$ ne sont pas corrélées bien que Y soit une fonction de X.

95. Soient $X_1, \dots, X_n$, des variables aléatoires de carré intégrable. On suppose qu'il existe deux réels σ et γ tels que

$$\forall\, 1 \leqslant k \leqslant n, \quad \mathbf{V}(X_k) = \sigma^2$$
$$\text{et} \quad \forall\, 1 \leqslant k < \ell \leqslant n, \quad \mathbf{Cov}(X_k, X_\ell) = \gamma.$$

Alors $(n-1)\gamma \geqslant -\sigma^2$.

96. Convergence en probabilité
Soient X et $(X_n)_{n\in\mathbb{N}}$, des variables aléatoires réelles définies sur un même espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$.
96.1 ✍ *La suite $(X_n)_{n\in\mathbb{N}}$* ***converge en probabilité*** *vers X lorsque*

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

96.2 Si $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers le réel m et si f est une fonction continue au point m, alors $f(X)$ converge en probabilité vers $f(m)$.
96.3 Si $\mathbf{V}(X_n)$ tend vers 0 et si $\mathbf{E}(X_n)$ tend vers m, alors la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers la variable aléatoire constante égale à m.

97. Soit X et $(X_n)_{n\in\mathbb{N}}$, des variables aléatoires de carré intégrable définies sur un même espace probabilisé $(\Omega, \mathfrak{A}, \mathbf{P})$.
97.1 ✍ *La suite $(X_n)_{n\in\mathbb{N}}$* ***converge en moyenne quadratique*** *vers X lorsque*

$$\lim_{n\to+\infty} \mathbf{E}[(X_n - X)^2] = 0.$$

97.2 La suite $(X_n)_{n\in\mathbb{N}}$ converge en moyenne quadratique vers un réel m si, et seulement si, $\mathbf{E}(X_n)$ tend vers m et $\mathbf{V}(X_n)$ tend vers 0.
97.3 Si la suite $(X_n)_{n\in\mathbb{N}}$ converge en moyenne quadratique vers un réel m, alors elle converge également vers m en probabilité [96.1].
97.4 On suppose que, pour tout entier $n \geqslant 1$,

$$\mathbf{P}(X_n = 1/n) = 1 - \frac{1}{n^2} \quad \text{et} \quad \mathbf{P}(X_n = n) = \frac{1}{n^2}.$$

Alors la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers 0 mais ne converge pas en moyenne quadratique vers 0.

98. Ajustement linéaire
Deux grandeurs sont modélisées par les variables aléatoires X et Y. On cherche deux réels a et b tels que la relation affine

$$Y = aX + b$$

soit aussi précise que possible et pour cela, on cherche à minimiser l'**écart quadratique**

$$\varphi(a,b) = \mathbf{E}\big[(Y - [aX+b])^2\big].$$

On pose

$$\widehat{a} = \frac{\mathbf{Cov}(X,Y)}{\mathbf{V}(X)}$$

et $\widehat{b} = \mathbf{E}(Y) - \widehat{a}\,\mathbf{E}(X)$.
98.1 L'écart quadratique $\varphi(a,b)$ est minimal si, et seulement si, la variance $\mathbf{V}(Y - aX)$ est minimale et $\mathbf{E}(Y) = a\,\mathbf{E}(X) + b$.
Par conséquent, l'écart quadratique $\varphi(a,b)$ est minimal si, et seulement si, $(a,b) = (\widehat{a}, \widehat{b})$.
98.2 Les variables $\widehat{a}X$ et $Y - \widehat{a}X$ sont décorrélées et

$$\mathbf{V}(Y) = \mathbf{V}(\widehat{a}X) + \mathbf{V}(Y - \widehat{a}X).$$

Le terme $\mathbf{V}(\widehat{a}X)$ est la **variance expliquée**, c'est-à-dire la dispersion des valeurs de Y qui se déduit de la dispersion des valeurs de X (on considère Y comme une fonction affine de X).
Le terme $\mathbf{V}(Y - \widehat{a}X)$ est la **variance résiduelle**, qui est nulle si, et seulement si, Y est exactement une fonction affine de X.
98.3 Le signe de $\mathbf{Cov}(X,Y)$ indique dans quel sens la variable Y a tendance à évoluer en fonction de la variable X. →[84.4]
98.4 Le **coefficient de corrélation linéaire** est défini par

$$r = \frac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{V}(X)\,\mathbf{V}(Y)}}.$$

Il est compris entre -1 et 1 et indique dans quelle mesure Y est une fonction affine de X en reliant la variance résiduelle et la variance de Y :

$$\frac{\mathbf{V}(Y - \widehat{a}X)}{\mathbf{V}(Y)} = 1 - r^2.$$

98.5 Plus généralement, l'**espérance conditionnelle de Y sachant** X est une variable aléatoire $f_0(X)$ qui minimise l'écart quadratique

$$\mathbf{E}\big(|Y - f(X)|^2\big)$$

parmi l'ensemble de *toutes* les fonctions f possibles (et non plus seulement parmi l'ensemble des fonctions *affines* f).

V

Fonctions génératrices

99. Dans ce paragraphe, on considère exclusivement des variables aléatoires discrètes à valeurs dans $\mathbb{N}$.

100.1✍ *La **série génératrice** d'une variable aléatoire X définie sur $(\Omega, \mathfrak{A}, \mathbf{P})$ et à valeurs dans $\mathbb{N}$ est la série entière $\sum \mathbf{P}(X = n)t^n$.*

100.2✍ *La somme de cette série entière, définie par*

$$G_X(t) = \sum_{n=0}^{+\infty} \mathbf{P}(X = n)t^n,$$

*est la **fonction génératrice** de X.*

100.3 Si X est une variable aléatoire à valeurs dans $\mathbb{N}$, alors t^X est une variable aléatoire discrète pour tout $t \in \mathbb{R}$ et

$$\forall\, t \in [-1,1], \quad G_X(t) = \mathbf{E}(t^X)$$

et en particulier $G_X(1) = 1$.

101. Fonction génératrice des lois usuelles
Dans les expressions suivantes, on note $q = 1 - p$.

101.1 Si X suit la loi de Bernoulli de paramètre p, alors

$$\forall\, t \in \mathbb{R}, \quad G_X(t) = q + pt.$$

101.2 Si X suit la loi binomiale de paramètres n et p, alors

$$\forall\, t \in \mathbb{R}, \quad G_X(t) = [q + pt]^n.$$

101.3 Si X suit la loi géométrique de paramètre p, alors

$$\forall\, t \in \left]-1/q, 1/q\right[, \quad G_X(t) = \frac{pt}{1 - qt}.$$

101.4 Si X suit la loi de Poisson de paramètre λ, alors

$$\forall\, t \in \mathbb{R}, \quad G_X(t) = \exp[\lambda(t - 1)].$$

102. Caractérisation de la loi de X
La fonction génératrice G_X caractérise la loi de X.

102.1 Le rayon de convergence de la série entière $\sum \mathbf{P}(X = n)z^n$ est supérieur à 1.

102.2 La série entière $\sum \mathbf{P}(X = n)z^n$ converge normalement sur le disque unité fermé $\mathbb{D} \subset \mathbb{C}$ et

$$\forall\, n \in \mathbb{N}, \quad \mathbf{P}(X = n) = \frac{1}{2\pi} \int_0^{2\pi} G_X(e^{it})e^{-int}\, \mathrm{d}t.$$

102.3 La fonction génératrice G_X est continue sur $[-1,1]$, croissante et convexe sur $[0,1]$.

102.4 La fonction génératrice G_X est de classe $\mathscr{C}^\infty$ sur $]-1,1[$ et

$$\forall\, n \in \mathbb{N}, \quad \mathbf{P}(X = n) = \frac{G_X^{(n)}(0)}{n!}.$$

102.5➙ *Si X et Y sont deux variables aléatoires à valeurs dans $\mathbb{N}$ dont les fonctions génératrices G_X et G_Y coïncident sur $[0,1]$, alors X et Y ont même loi.*

Régularité et moments factoriels

103. De même que la formule de Taylor en $t = 0$ permet de déduire la loi de X de sa fonction génératrice, la formule de Taylor en $t = 1$ permet de calculer les moments de X par l'intermédiaire du calcul des moments factoriels **[75]**.

103.1 Si le rayon de convergence de la série génératrice est strictement supérieur à 1, alors X admet un moment d'ordre n pour tout $n \in \mathbb{N}$ et

$$\forall\, n \geqslant 1, \quad \mathbf{E}[X(X-1)\cdots(X-n+1)] = G_X^{(n)}(1).$$

103.2➙ *Si X admet un moment d'ordre $n \geqslant 1$, alors G_X est de classe $\mathscr{C}^n$ sur $[0,1]$ et*

$$\forall\, 1 \leqslant k \leqslant n, \quad \mathbf{E}[X(X-1)\cdots(X-k+1)] = G_X^{(k)}(1).$$

104. On considère une série convergente $\sum a_k$ de terme général positif et on pose

$$w_n = \sum_{k=n+1}^{+\infty} a_k$$

pour tout entier $n \in \mathbb{N}$.

104.1 La série $\sum w_n$ est convergente si, et seulement si, la série $\sum k a_k$ est convergente.
Dans ce cas,

$$\sum_{n=0}^{+\infty} w_n = \sum_{k=0}^{+\infty} k a_k.$$

104.2 La fonction φ définie par

$$\varphi(t) = \sum_{n=0}^{+\infty} w_n t^n$$

est définie au moins pour $t \in [0,1[$.

1. Si la série $\sum w_n$ converge, alors la fonction φ est définie et continue sur $[0,1]$.
2. Si la série $\sum w_n$ diverge, alors la fonction φ tend vers $+\infty$ au voisinage de 1.

104.3 La fonction φ tend vers une limite finie au voisinage de 1 si, et seulement si, la série $\sum w_n$ converge. Dans ce cas,

$$\lim_{t \to 1} \varphi(t) = \sum_{n=0}^{+\infty} w_n$$

104.4 La fonction F définie par

$$F(t) = \sum_{k=0}^{+\infty} a_k t^k$$

est définie pour tout $t \in [0,1]$.
Elle est dérivable sur $[0,1[$ au moins et

$$\forall\, 0 \leqslant t < 1, \quad F'(t) = \sum_{k=0}^{+\infty} k a_k t^{k-1}.$$

D'autre part,

$$\forall\, 0 \leqslant t < 1, \quad \frac{F(1) - F(t)}{1 - t} = \sum_{n=0}^{+\infty} w_n t^n.$$

104.5 La fonction F est dérivable en $t = 1$ si, et seulement si, la série $\sum k a_k$ est convergente. Dans ce cas,

$$F'(1) = \sum_{n=0}^{+\infty} w_n = \sum_{k=0}^{+\infty} k a_k.$$

105. Nous allons déduire du lemme précédent une réciproque partielle de **[103.2]**.

105.1 On considère une variable aléatoire $X : \Omega \to \mathbb{N}$ et sa fonction génératrice G_X définie par

$$\forall\, t \in [0,1], \quad G_X(t) = \sum_{k=0}^{+\infty} \mathbf{P}(X = k)t^k.$$

105.2➙ *Une variable aléatoire $X : \Omega \to \mathbb{N}$ est d'espérance finie si, et seulement si, sa fonction génératrice G_X est dérivable en 1. Dans ce cas,*

$$\mathbf{E}(X) = G_X'(1).$$

105.3 On suppose que X est d'espérance finie. La fonction génératrice G_X est de classe $\mathscr{C}^1$ sur le segment $[0,1]$ et

$$\forall\, t \in [0,1], \quad G'_X(t) = \sum_{k=0}^{+\infty}(k+1)\,\mathbf{P}(X = k+1)t^k.$$

105.4→ *Une variable aléatoire $X : \Omega \to \mathbb{N}$ est de carré intégrable si, et seulement si, sa fonction génératrice G_X est deux fois dérivable en 1. Dans ce cas,*

$$\mathbf{V}(X) = G''_X(1) + G'_X(1)\big[1 - G'_X(1)\big].$$

Entraînement

106. Questions pour réfléchir

1. Si np_n tend vers $\lambda \in \mathbb{R}_+^*$ lorsque n tend vers $+\infty$, alors

$$\forall\, t \in \mathbb{R}_+, \quad \big((1-p_n) + p_n t\big)^n \xrightarrow[n\to+\infty]{} \exp[\lambda(t-1)].$$

Comparer avec **[22.2]**.

2. Quelle que soit la variable aléatoire X, toutes les dérivées de G_X sont croissantes sur $[0,1[$.
3. La fonction génératrice G_X est strictement croissante sur $[0,1]$ si, et seulement si, la variable aléatoire X n'est pas presque sûrement nulle.
4. La fonction génératrice de X est polynomiale si, et seulement si, X est une variable aléatoire à support fini **[5.6]**.

107. Fonction génératrice des moments
La **fonction génératrice des moments** d'une variable aléatoire réelle X est définie par

$$\psi(t) = \mathbf{E}(e^{tX}).$$

S'il existe un réel $t_0 > 0$ tel que $\psi(t_0)$ soit défini, on dit que la variable aléatoire X admet des **moments exponentiels**.
107.1 Si X est une variable aléatoire presque sûrement bornée, alors ψ est définie sur $\mathbb{R}$.
107.2 Si X est une variable aléatoire à valeurs dans $\mathbb{N}$ et s'il existe $0 < a < 1$ tel que

$$\mathbf{P}(X = k) = o(a^k)$$

lorsque k tend vers $+\infty$, alors le rayon de convergence de sa série génératrice est strictement supérieur à 1, la variable X admet des moments de tout ordre et il existe $\alpha > 0$ tel que ψ soit définie au moins sur $]-\infty, \alpha[$ et

$$\forall\, t \in\,]-\alpha, \alpha[, \quad \psi(t) = \sum_{n=0}^{+\infty} \frac{\mathbf{E}(X^n)}{n!}\, t^n.$$

107.3 Sous les hypothèses précédentes, pour tout $k \in \mathbb{N}$,

$$\mathbf{P}(X > k) \leqslant \inf_{|t|<\alpha} \psi(t)e^{-kt}.$$

108. *Suite de* **[104]** – Soit X, une variable aléatoire à valeurs dans $\mathbb{N}$, de fonction génératrice G.
108.1 La fonction définie par

$$H(t) = \sum_{n=0}^{+\infty} \mathbf{P}(X > n)t^n$$

est définie au moins sur $]-1,1[$ et

$$\forall\, t \in\,]-1,1[, \quad H(t) = \frac{1-G(t)}{1-t}.$$

108.2 La fonction H est prolongeable en une fonction continue sur $[0,1]$ si, et seulement si, X est une variable aléatoire d'espérance finie et dans ce cas, $\mathbf{E}(X) = H(1)$.
108.3 On suppose que X est une variable aléatoire d'espérance finie. La fonction H est alors dérivable en 1 si, et seulement si, la série $\sum k\,\mathbf{P}(X > k)$ est convergente.
Dans ce cas, X est une variable aléatoire de carré intégrable et

$$H'(1) = \frac{1}{2}\,\mathbf{E}\big((X-1)X\big).$$

Questions, exercices & problèmes

Approfondissement

109.1✍ *La **fonction de répartition*** d'une variable aléatoire réelle X définie sur $(\Omega, \mathfrak{A}, \mathbf{P})$ est la fonction $F_X : \mathbb{R} \to \mathbb{R}$ définie par*

$$\forall\, a \in \mathbb{R}, \quad F_X(a) = \mathbf{P}(X \leqslant a).$$

109.2 La fonction de répartition d'une variable aléatoire réelle X est une fonction croissante, continue à droite, qui tend vers 0 au voisinage de $-\infty$ et vers 1 au voisinage de $+\infty$.
109.3 La fonction de répartition F_X est discontinue en x_0 si, et seulement si, $\mathbf{P}(X = x_0) > 0$.
109.4 Deux variables aléatoires discrètes réelles ont même fonction de répartition si, et seulement si, elles ont même loi.
109.5 Si la fonction de répartition F_X est de classe $\mathscr{C}^1$ sur $\mathbb{R}$, alors il existe une, et une seule, fonction f_X continue et intégrable sur $\mathbb{R}$ telle que

$$\forall\, a \leqslant b, \quad \mathbf{P}(a \leqslant X \leqslant b) = \int_a^b f_X(t)\,\mathrm{d}t.$$

La fonction f_X est une **densité** de X.
110. Soit X, une variable réelle discrète sur $(\Omega, \mathfrak{A}, \mathbf{P})$. On suppose que le support de la loi de X est **bien-ordonné**, au sens où il existe une suite strictement croissante $(x_n)_{n\in\mathbb{N}}$ telle que

$$\sum_{k=0}^{+\infty} \mathbf{P}(X = x_k) = 1,$$

et que cette variable est d'espérance finie.

1. Pour tout $x \in \mathbb{R}$, on pose

$$G(x) = \mathbf{P}(X > x).$$

La fonction $G : \mathbb{R} \to \mathbb{R}$ ainsi définie est la **queue** de la fonction de répartition **[109]** de X.
1.a La fonction G est décroissante et tend vers 0 en $+\infty$.
1.b Pour tout $N \in \mathbb{N}$,

$$x_N\ \mathbf{P}(X \geqslant x_N) \leqslant \mathbf{E}(X\,\mathbb{1}_{(X \geqslant x_N)})$$

et la suite de terme général $x_N G(x_{N-1})$ tend vers 0 lorsque N tend vers $+\infty$.

2.
$$\begin{aligned}\mathbf{E}(X) &= x_0\big[1 - G(x_0)\big] + \sum_{k=1}^{+\infty} x_k\big[G(x_{k-1}) - G(x_k)\big] \\ &= \sum_{k=0}^{+\infty}(x_{k+1} - x_k)\,\mathbf{P}(X > x_k)\end{aligned}$$

3. Comparer avec **[63]**.

111. Soient $(\Omega, \mathfrak{A}, \mathbf{P})$, un espace probabilisé et $(E_n)_{n\in\mathbb{N}}$, une suite d'événements tels que la série $\sum \mathbf{P}(E_n)$ converge.

1. Reconnaître la loi de la variable aléatoire $\mathbb{1}_{E_n}$.

2. Pour tout entier n, la somme

$$Z_n = \sum_{k=0}^{n} \mathbb{1}_{E_k}$$

est une variable aléatoire discrète d'espérance finie :

$$\mathbf{E}(Z_n) \leqslant \sum_{k=0}^{+\infty} \mathbf{P}(E_k).$$

3. On note F, l'ensemble des $\omega \in \Omega$ tels que l'ensemble d'indices

$$\{n \in \mathbb{N} \,:\, \omega \in E_n\}$$

soit fini. Comme

$$F = \bigcup_{n\in\mathbb{N}} \bigcap_{k=n}^{+\infty} E_k^c,$$

alors F est un événement presque sûr.

4.a La fonction $Z : \Omega \to \mathbb{N}$ définie par

$$\forall\, \omega \in F, \quad Z(\omega) = 0$$

et par

$$\forall\, \omega \in F^c, \quad Z(\omega) = \sum_{n=0}^{+\infty} \mathbb{1}_{E_n}(\omega)$$

est une variable aléatoire discrète sur $(\Omega, \mathfrak{A})$.

4.b

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(Z > k) = \lim_{n\to+\infty} \mathbf{P}(Z_n > k)$$

4.c Quels que soient les entiers $m \leqslant n$,

$$\sum_{k=0}^{m} \mathbf{P}(Z_n > k) \leqslant \sum_{k=0}^{n} \mathbf{P}(Z_n > k) = \mathbf{E}(Z_n) \leqslant \sum_{k=0}^{+\infty} \mathbf{P}(E_k)$$

donc la série $\sum \mathbf{P}(Z > k)$ est convergente et Z est une variable aléatoire d'espérance finie **[63]**.

112. *Suite de* **[25]** –

112.1 Pour tout entier $i \geqslant 1$, on note A_i, l'ensemble des entiers naturels divisibles par i. Les événements

$$[X \in A_{i_1}], \dots, [X \in A_{i_n}]$$

sont indépendants si, et seulement si, les entiers $i_1, \dots, i_n$ sont deux à deux premiers entre eux.

112.2 Soit $(p_n)_{n\in\mathbb{N}}$, une énumération de l'ensemble des entiers premiers. Pour tout entier $n \in \mathbb{N}$, on note C_n, l'ensemble des entiers qui ne sont divisibles ni par p_0, ni par $p_1, \dots$, ni par p_n.

1. Calculer $\mathbf{P}(X \in C_n)$.

1.a Que dire lorsque n tend vers $+\infty$?

Pour aller plus loin

113. Soit X, une variable aléatoire à valeurs dans $\mathbb{N}$ telle que

$$\forall\, n \in \mathbb{N}, \quad \mathbf{P}(X = n) = \frac{1}{2^{n+1}}.$$

Il existe une suite $(X_k)_{k\in\mathbb{N}}$ de variables aléatoires de Bernoulli telles que

$$\forall\, \omega \in \Omega, \quad X(\omega) = \sum_{k=0}^{+\infty} 2^k X_k(\omega)$$

et, pour tout $k \in \mathbb{N}$, le paramètre de X_k se déduit de l'égalité suivante.

$$[X_k = 1] = \bigsqcup_{\substack{q\in\mathbb{N} \\ 2^d \leqslant r < 2^{d+1}}} [X = 2^{d+1}q + r]$$

114. Variables aléatoires réelles

La **tribu borélienne** sur $\mathbb{R}$ est la plus petite tribu sur $\mathbb{R}$ qui contienne les intervalles de $\mathbb{R}$ (on admettra l'existence de cette tribu).

114.1 Si $X : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathcal{B})$ est une variable aléatoire, alors :

1. Quels que soient $a < b$, les parties suivantes de Ω sont des événements :

$$[a \leqslant X \leqslant b], \quad [a < X \leqslant b], \quad [a \leqslant X < b], \quad [a < X < b].$$

2. Pour tout $a \in \mathbb{R}$, les parties suivantes de Ω sont des événements :

$$[X = a], \quad [X \leqslant a], \quad [X > a], \quad [X < a].$$

114.2 On admettra que $X : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathcal{B})$ est une variable aléatoire si, et seulement si, $[X \leqslant a] \in \mathfrak{A}$ pour tout réel a.

114.3 Soit $X : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathcal{B})$, une variable aléatoire réelle.

3. Soit $f : \mathbb{R} \to \mathbb{R}$, une fonction monotone.

3.a Alors $f(X)$ est une variable aléatoire réelle. →**[114.2]**

3.b Quels que soient $y \in \mathbb{R}$, $m \in \mathbb{R}$ et $\sigma > 0$,

$$\Big[\frac{X - m}{\sigma} \leqslant y\Big] = [X \leqslant m + \sigma y].$$

4. Si $y < 0$, alors $[X^2 \leqslant y] = \varnothing$ et

$$\forall\, y \geqslant 0, \quad [X^2 \leqslant y] = [-\sqrt{y} \leqslant X \leqslant \sqrt{y}] \in \mathfrak{A}.$$

4.a Quels que soient $t > 0$ et $y > 0$,

$$[\exp(tX) \leqslant y] = \Big[X \leqslant \frac{\ell n\, y}{t}\Big] \in \mathfrak{A}.$$

4.b

$$[\cos X \geqslant 0] = \bigsqcup_{k\in\mathbb{Z}} [-\pi/2 + 2k\pi \leqslant X \leqslant \pi/2 + 2k\pi] \in \mathfrak{A}$$

114.4 Soit $(X_n)_{n\in\mathbb{N}}$, une suite de variables aléatoires réelles sur $(\Omega, \mathfrak{A}, \mathbf{P})$. On pose

$$\forall\, \omega \in \Omega, \quad Y(\omega) = \inf_{n\in\mathbb{N}} X_n(\omega) \quad \text{et} \quad Z(\omega) = \sup_{n\in\mathbb{N}} X_n(\omega).$$

Alors, quel que soit $x \in \mathbb{R}$,

$$[Y \geqslant x] = \bigcap_{n\in\mathbb{N}} [X_n \geqslant x] \quad \text{et} \quad [Z \leqslant x] = \bigcap_{n\in\mathbb{N}} [X_n \leqslant x]$$

donc Y et Z sont des variables aléatoires réelles. →**[114.2]**

115. Autre définition des variables discrètes
On note ici $\mathcal{B}$, une tribu sur $\mathbb{R}$ qui contient tous les intervalles.
La **tribu borélienne** [114] est un exemple d'une telle tribu.
On convient ici qu'une variable aléatoire réelle

$$X : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathcal{B})$$

est **discrète** lorsqu'il existe une partie $E \subset \mathbb{R}$ finie ou dénombrable telle que $\mathsf{P}(X \in E) = 1$.

1. Comparer cette définition avec [7].
2. Comme la tribu $\mathcal{B}$ contient les singletons, les parties E et E^c appartiennent à $\mathcal{B}$.
3. Que dire de $[X \in E^c]$?
4. On suppose que la tribu $\mathfrak{A}$ est complète [**13.24**]. Comme

$$\forall\, B \in \mathfrak{P}(\mathbb{R}), \quad [X \in B] = [X \in B \cap E] \sqcup [X \in B \cap E^c],$$

alors X est une variable aléatoire de $(\Omega, \mathfrak{A})$ dans $(\mathbb{R}, \mathfrak{P}(\mathbb{R}))$.

116. Convergence presque sûre
Soient X et $(X_n)_{n\in\mathbb{N}}$, des variables aléatoires réelles définies sur un même espace probabilisé $(\Omega, \mathfrak{A}, \mathsf{P})$.

116.1 L'ensemble des $\omega \in \Omega$ tels que la suite de terme général $X_n(\omega)$ converge vers $X(\omega)$ est l'événement

$$C = \bigcap_{p\in\mathbb{N}^*} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geqslant N} \big[|X_n - X| \leqslant 1/p\big].$$

116.2✍ *On dit que la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ **converge presque sûrement** vers X lorsque $\mathsf{P}(C) = 1$.*

116.3 Si la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X, alors elle converge aussi en probabilité [**96.1**] vers X.