CHAPITRE 10

Variables aléatoires

Constantin CARATHÉODORY fut repéré par Felix KLEIN, à Göttingen, et prépara son mémoire d'habilitation avec Hermann MINKOWSKI.

Constantin CARATHÉODORY naît en 1873 à Berlin de parents grecs, puis il grandit à Bruxelles, où son père était ambassadeur de l'Empire ottoman en Belgique. Ses langues maternelles sont le grec et le français, mais il était un érudit en allemand et parlait couramment anglais, italien, turc, grec ancien et latin. Il a fait ses études supérieures à Berlin puis Göttingen, puis devint professeur et enseigna en Allemagne et en Grèce.

Il est l'auteur d'importants travaux en théorie des fonctions à variables réelles, calcul des variations et théorie de la mesure. En 1909, CARATHÉODORY fit œuvre de pionnier dans la formulation axiomatique de la thermodynamique en utilisant une approche purement géométrique.

Le théorème de CARATHÉODORY pour les enveloppes convexes permet d'obtenir des résultats de compacité. La reformulation de la différentiabilité par CARATHÉODORY est également un outil très puissant et montre son intérêt pour l'analyse. Parmi ses théorèmes les plus fameux, on peut citer le théorème d'extension qui permet d'étendre une mesure définie sur un anneau d'ensembles à la σ-algèbre engendrée par cet anneau. Il se base sur la notion de mesure extérieure, qui est une généralisation d'idées mises en place par Henri LEBESGUE pour la théorie de l'intégration. On la définit comme l'infimum de la somme des mesures d'ensembles tels que leur réunion disjointe contienne la partie dont on calcule la mesure extérieure.

Programme

— Espérance d'une variable aléatoire à valeurs dans $\mathbf{R}$, notée $\mathbf{E}(X)$. Linéarité, positivité et croissance de l'espérance sur l'espace des variables aléatoires d'espérance finie définies sur Ω. Si $|X| \leqslant Y$ et si Y est d'espérance finie, alors X est d'espérance finie. Variables centrées.

— Moments. Une variable aléatoire ayant un moment d'ordre 2 est d'espérance finie. Inégalité de Cauchy-Schwarz. Espace des variables aléatoires définies sur Ω admettant un moment d'ordre 2. Variance $\mathrm{V}(X)$, écart type $\sigma(X)$. Variables réduites. Relations $\mathrm{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$ et $\mathrm{V}(aX+b) = a^2\mathrm{V}(X)$. Si $\sigma(X) > 0$, $\dfrac{X - \mathbf{E}(X)}{\sigma(X)}$ est centrée réduite.

— Variance d'une variable aléatoire géométrique, de Poisson.

— Formule de transfert (démonstration non exigible). Inégalité de Markov. Inégalité de Bienaymé-Tchebychev.

Dans ce chapitre X et Y désignent des variables aléatoires réelles discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$.

1 Sommabilité

Définition 10 - 1

Sommabilité – espérance finie

Si X est une variable aléatoire réelle discrète, on dit qu'elle est d'espérance finie (ou qu'elle admet un moment d'ordre 1) si la famille $(\mathbf{P}(X=x)x)_{x\in X(\Omega)}$ est sommable. Autrement dit X est d'espérance finie si et seulement si $\mathbf{E}(|X|)$ est fini.

Exemple 10 - 1

Si X est presque sûrement égal à une constante a, i.e. $\mathbf{P}(X=a)=1$, alors X est d'espérance finie égale à a.

Si X est borné, alors X est d'espérance finie. En effet la famille $(\mathbf{P}(X=x))_{x\in X(\Omega)}$ est sommable de somme 1, par σ-additivité de $\mathbf{P}$ et donc $(\mathbf{P}(X=x)\,|x|)_{x\in X(\Omega)}$ l'est aussi par comparaison avec une famille sommable.

Si X est une variable aléatoire positive, cette notion coïncide avec celle développée dans ce cadre : X est d'espérance finie si et seulement si $\mathbf{E}(X) < +\infty$.

Définition 10 - 2

Espérance

Si X est une variable aléatoire réelle discrète d'espérance finie, on définit l'espérance de X par la formule

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

Exemple 10 - 2

Si $X \sim \mathcal{U}(\llbracket a;b \rrbracket)$, alors X admet un moment d'ordre 1 et $\mathbf{E}(X) = \dfrac{a+b}{2}$.

Notation

Espace $\mathcal{L}^1$ ♠

On note $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbf{P})$ celui des variables aléatoires disrètes sur $(\Omega, \mathcal{A}, \mathbf{P})$ réelles et sommables.

Proposition 10 - 1

L'ensemble $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbf{P})$ est un espace vectoriel.

Démonstration. Soit X et Y deux variables aléatoires réelles avec $X(\Omega)$ et $Y(\Omega)$ au plus dénombrables et α et β des scalaires. En notant $Z = \alpha X + \beta Y$, $Z(\Omega)$ est inclus dans l'image de l'ensemble au plus dénombrable $X(\Omega) \times Y(\Omega)$ par l'application $(x, y) \mapsto \alpha x + \beta y$ et donc Z est à valeurs discrètes. De plus, pour z dans $Z(\Omega)$, on a

$$(Z = z) = \sum_{(x,y)\in I} (X = x, Y = y, Z = z) = \sum_{\substack{(x,y)\in I \\ \alpha x+\beta y=z}} (X = x, Y = y)\,,$$

en notant $I = X(\Omega) \times Y(\Omega)$, et donc Z est une variable aléatoire réelle discrète.

Soit K un ensemble fini inclus dans I, et K_X et K_Y ses projections sur $X(\Omega)$ et $Y(\Omega)$ respectivement, de sorte qu'on a $K \subset K_X \times K_Y$. Il vient, par positivité des termes,

$$\sum_{(x,y)\in K} |x|\,\mathbf{P}(X = x, Y = y) \leq \sum_{x\in K_X} |x| \sum_{y\in K_Y} \mathbf{P}(X = x, Y = y) \leq \mathbf{E}\,(|X|)$$

puisque $\displaystyle\sum_{y\in K_Y} \mathbf{P}(X = x, Y = y) = \mathbf{P}\left((X = x) \cap \sum_{y\in K_Y} (Y = y)\right) \leq \mathbf{P}(X = x)$ par croissance de $\mathbf{P}$. On en déduit que la famille $(x\mathbf{P}(X = x, Y = y))_{(x,y)\in I}$ est sommable et donc aussi $(y\mathbf{P}(X = x, Y = y))_{(x,y)\in I}$, par symétrie entre les rôles de (X, x) et (Y, y). Comme $\ell^1(I, \mathbf{R})$ est un $\mathbf{R}$-espace vectoriel, on en déduit que la famille $((\alpha x + \beta y)\mathbf{P}(X = x, Y = y))_{(x,y)\in I}$ est sommable. Par σ-additivité, on a

$$\mathbf{P}(Z = z) = \sum_{\substack{(x,y)\in I \\ \alpha x+\beta y=z}} \mathbf{P}(X = x, Y = y)\,.$$

et il résulte donc du théorème de Tonelli que $(z\mathbf{P}(Z = z))_{z\in Z(\Omega)}$ est également sommable, i.e. $\alpha X + \beta Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbf{P})$. □

Remarque 10 - 1

Puisqu'une variable aléatoire peut-être presque sûrement nulle sans être nulle, l'espace $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbf{P})$ n'est pas un espace vectoriel normé par $\mathbf{E}\,(|X|)$. On pourrait définir des classes d'équivalence de variables aléatoires par la relation « est presque sûrement égale à » et munir l'ensemble de ces classes d'équivalence d'une structure d'espace vectoriel normé.

Proposition 10 - 2

Linéarité

L'application qui à une variable aléatoire discrète dans $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbf{P})$ associe son espérance est linéaire, positive et croissante. Plus précisément, soit X et Y dans $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbf{P})$.

1. Si $X \geq 0$ presque sûrement, alors on a $\mathbf{E}\,(X) \geq 0$ et $\mathbf{E}\,(X) = 0$ si et seulement si $X = 0$ presque sûrement ;
2. si $X \leq Y$ presque sûrement, alors on a $\mathbf{E}\,(X) \leq \mathbf{E}\,(Y)$ et $\mathbf{E}\,(X) = \mathbf{E}\,(Y)$ si et seulement si $X = Y$ presque sûrement.

Démonstration. La linéarité résulte du théorème de sommation par paquets : si $Z = \alpha X + \beta Y$, avec X et Y sont deux variables aléatoires réelles discrètes et α et β des scalaires, on a

$$\begin{aligned}
\mathbf{E}(Z) &= \sum_{z\in Z(\Omega)} z\mathbf{P}(Z=z) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} (\alpha x+\beta y)\mathbf{P}(X=x,Y=y) \\
&= \alpha \sum_{x\in X(\Omega)} x \sum_{y\in Y(\Omega)} \mathbf{P}(X=x,Y=y) + \beta \sum_{y\in Y(\Omega)} y \sum_{x\in X(\Omega)} \mathbf{P}(X=x,Y=y) \\
&= \alpha \sum_{x\in X(\Omega)} x\mathbf{P}(X=x) + \beta \sum_{y\in Y(\Omega)} y\mathbf{P}(Y=y) = \alpha\mathbf{E}(X)+\beta\mathbf{E}(Y) \ .
\end{aligned}$$

L'assertion sur la positivité résulte du cas des familles positives appliqué à la famille $(x\mathbf{P}(X=x))_{x\in X(\Omega)}$.

L'assertion relative à la croissance se ramène au cas positif par linéarité. □

♥

La linéarité **ne résulte pas** du cas des familles de nombres réels. En effet l'application qui à une variable aléatoire réelle discrète X associe la famille $(\mathbf{P}(X=x)x)_{x\in X(\Omega)}$ n'est pas linéaire. On somme ici sur les **valeurs** de X et non sur la variable ω. C'est l'application $X \mapsto (X(\omega))_{\omega\in\Omega}$ qui est linéaire, mais ce n'est pas celle-là qui nous intéresse !

Proposition 10 - 3

Inégalité triangulaire

Si $X \in \mathcal{L}^1(\Omega,\mathcal{A},\mathbf{P})$, alors $|\mathbf{E}(X)| \leq \mathbf{E}(|X|)$.

Démonstration. Quant aux variables aléatoires, on a $X \leq |X|$ et $-X \leq |X|$, donc $\mathbf{E}(X) \leq \mathbf{E}(|X|)$ et $-\mathbf{E}(X) \leq \mathbf{E}(|X|)$. □

2 Critères de sommabilité

Proposition 10 - 4

Comparaison à une variable aléatoire positive

Si Y est une variable aléatoire discrète positive et d'espérance finie et si $|X| \leq Y$, alors X est d'espérance finie et $|\mathbf{E}(X)| \leq \mathbf{E}(Y)$.

Démonstration. Puisque la famille $(y\mathbf{P}(Y=y))_{y\in Y(\Omega)}$ est sommable, il en va de même, d'après le théorème de sommation par paquets dans le cas positif, de la famille $(y\mathbf{P}(Y=y,X=x))_{(x,y)\in X(\Omega)\times Y(\Omega)}$. Par conséquent, par comparaison dans le cas des familles à termes positifs, la famille $(x\mathbf{P}(Y=y,X=x))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable. Il en résulte, toujours d'après le théorème de sommation par paquets dans le cas positif, que X est d'espérance finie.

La seconde partie résulte de la croissance de l'espérance et de l'inégalité triangulaire. □

Théorème 10 - 1

Formule de transfert

Soit f une fonction de $X(\Omega)$ dans $\mathbf{R}$, alors $f(X)$ est d'espérance finie si et seulement si la famille $(f(x)\mathbf{P}(X=x))_{x\in X(\Omega)}$ est sommable et, dans ce cas, on a

$$\mathbf{E}\left(f(X)\right) = \sum_{x\in X(\Omega)} f(x)\mathbf{P}(X=x)\ .$$

Démonstration. Si $X(\Omega)$ est au plus dénombrable, alors $f(X(\Omega))$ l'est aussi. De plus, pour y dans $f(X(\Omega))$, on a

$$(f(X)=y) = \sum_{x\in f^{-1}(y)} (X=x)$$

et donc le membre de gauche appartient à $\mathcal{A}$. Il en résulte que $f(X)$ est une variable aléatoire réelle discrète.

D'après le théorème de sommation par paquets dans le cas positif, la famille $(f(x)\mathbf{P}(X=x))_{x\in X(\Omega)}$ est sommable si et seulement si, pour tout y dans $f(X(\Omega))$, la famille $(|f(x)|\,\mathbf{P}(X=x))_{x\in f^{-1}(y)}$ l'est et si la famille des sommes de ces familles l'est. Or la première partie de cette condition est toujours vérifiée et on a

$$\sum_{x\in f^{-1}(y)} |f(x)|\,\mathbf{P}(X=x) = |y|\,\mathbf{P}\left(\sum_{x\in f^{-1}(y)} (X=x)\right) = |y|\,\mathbf{P}(f(X)=y)\ .$$

Le critère de sommabilité énonce donc que $f(X)$ admet une espérance si et seulement si $(f(x)\mathbf{P}(X=x))_{x\in X(\Omega)}$ est sommable. La valeur de la somme résulte du théorème de sommation par paquets dans le cas général. $\square$

Exemple 10 - 3

Moment d'ordre 2 d'une variable de Poisson

Soit $X\sim\mathcal{P}(\lambda)$ et f la fonction donnée par $f(x)=x(x-1)$. On s'intéresse à une famille positive donc on peut directement considérer $\sum_{n=0}^{+\infty} n(n-1)\mathbf{P}(X=n)$. Il vient

$$E(f(X)) = \sum_{n=0}^{+\infty} n(n-1)e^{-\lambda}\frac{\lambda^n}{n!} = e^{-\lambda}\lambda^2\sum_{n=2}^{+\infty}\frac{\lambda^{n-2}}{(n-2)!} = \lambda^2$$

et donc aussi, par linéarité, $\mathbf{E}\left(X^2\right) = \mathbf{E}\left(f(X)+X\right) = \lambda(\lambda+1)$.

Corollaire 10 - 1

Inégalité de Markov

Soit f une fonction croissante et positive définie sur un intervalle I. Si X est à valeurs dans I presque sûrement, on a

$$\forall a\in I \qquad f(a)>0 \Longrightarrow \mathbf{P}\left(X\geq a\right)\leq \frac{\mathbf{E}\left(f(X)\right)}{f(a)}\ .$$

Démonstration. Si $f(X)$ a une espérance infinie alors le terme de droite de l'inégalité est infini et l'inégalité est vérifiée. On suppose que $f(X)$ admet une espérance et $a\in I$.

Par croissance de f, on a pour ω dans $(X\in I)$, $X(\omega)\geq a \Longrightarrow f(X(\omega))\geq f(a)$ et donc $(X\geq a)\cap(X\in I)\subset(f(X)\geq f(a))\cap(X\in I)$. Il en résulte par croissance de

$\mathbf{P}$ et puisque $(X \in I)$ est presque sûr, $\mathbf{P}(X \geq a) \leq \mathbf{P}(f(X) \geq f(a))$. Si on a $f(a) > 0$, l'inégalité de MARKOV dans le cas positif permet de conclure $\mathbf{P}(f(X) \geq f(a)) \leq \dfrac{\mathbf{E}(f(X))}{f(a)}$ et le résultat en découle. □

3 Moments

Définition 10 - 3

Soit n un entier naturel, on dit que X admet un moment d'ordre n si X^n admet une espérance, i.e. si la famille $(x^n \mathbf{P}(X = x))_{x \in X(\Omega)}$ est sommable. Son moment d'ordre n est alors défini par la formule

$$\mu_n(X) = \mathbf{E}(X^n) = \sum_{x \in X(\Omega)} x^n \mathbf{P}(X = x) .$$

Exemple 10 - 4

Le moment d'ordre 0 d'une variable aléatoire est toujours défini et vaut 1. Celui d'ordre 1 est l'espérance, comme on l'a déjà signalé avant de définir la notion de moment.

Si X est presque sûrement constant, il admet un moment de tous ordres : si $X = a$ presque sûrement alors $\mu_n(X) = a^n$.

Remarque 10 - 2

Puisque X^2 est une variable aléatoire positive, si $\mu_2(X) = 0$, alors X est presque sûrement égal à 0.

Si X admet un moment d'ordre n, il en va de même pour λX avec λ réel.

Exemple 10 - 5

Variables de POISSON

Si $X \sim \mathcal{P}(\lambda)$, alors X admet des moments de tous ordres en tant que variable positive, i.e. avec $\mathbf{E}(X^n) \in \mathbf{R}_+ \cup \{+\infty\}$. Si H_k est le polynôme de NEWTON $X(X-1)\cdots(X-k+1)$ alors $H_k(X)$ est une variable positive et un calcul direct donne

$$\mathbf{E}(H_k(X)) = \sum_{n=k}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} n(n-1)\cdots(n-k+1) = \lambda^k e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} = \lambda^k$$

de sorte que si $X^n = \sum_{k=0}^{n} a_k H_k$ alors $\mathbf{E}(X^n) = \sum_{k=0}^{n} a_k \lambda^k$. Il vient donc

$$\mu_n(X) = \sum_{k=0}^{n} \frac{\Delta^k(X^n)(0)}{k!} \lambda^k .$$

Proposition 10 - 5

Si X admet un moment d'ordre 2, alors il en admet également un d'ordre 1.

Démonstration. Par convexité, on a $|X| \leq \dfrac{1 + X^2}{2}$. Or, par linéarité et par hypothèse sur X^2, $\dfrac{1 + X^2}{2}$ est positive et admet une espérance, donc par comparaison X

aussi. □

Pour aller plus loin

On montre que si X admet un moment d'ordre n, alors il en admet de tout ordre inférieur à n.

Théorème 10 - 2

Inégalité de Cauchy-Schwarz

Si X et Y admettent des moments d'ordre 2, alors XY en admet un d'ordre 1 et

$$\mathbf{E}(XY)^2 \leq \mathbf{E}\left(X^2\right)\mathbf{E}\left(Y^2\right) .$$

En particulier l'ensemble des variables aléatoires réelles discrètes admettant un moment d'ordre 2 est un espace vectoriel réel.

Démonstration. Par convexité, on a $|XY| \leq \dfrac{X^2+Y^2}{2}$ et donc XY admet un moment d'ordre 1 si X et Y en ont un d'ordre 2. Si on suppose $\mathbf{E}\left(X^2\right) = \mathbf{E}\left(Y^2\right) = 1$, on en déduit $\mathbf{E}(XY) \leq 1$ et le cas général s'en déduit par linéarité de l'espérance et le fait qu'un moment d'ordre 2 n'est nul que si la variable aléatoire est presque sûrement nulle.

Comme $(X+Y)^2 = X^2 + Y^2 + 2XY$, si X et Y admettent des moments d'ordre 2, XY admet un moment d'ordre 1 et donc $X+Y$ admet un moment d'ordre 2. La multiplication par un scalaire ne changeant pas le fait que X admet un moment d'ordre 2, l'ensemble des variables aléatoires réelles discrètes admettant un moment d'ordre 2 est un sous-espace vectoriel réel de celui des variables aléatoires réelles discrètes. □

Pour étudier une variable aléatoire, il est souvent plus instructif de la centrer, i.e. de ramener sa moyenne en 0 en considérant $X - \mathbf{E}(X)$. Cette dernière variable aléatoire existe et admet une espérance si et seulement si X admet une espérance. De plus elle admet un moment d'ordre 2 si et seulement si X en admet un (en fait, c'est vrai pour n'importe quel moment). Le moment d'ordre 2 de cette variable centrée, appelé variance de X, permet de quantifier la dispersion des valeurs que prend X. Une variable constante est de variance nulle. Si X prend les valeurs x et $-x$ avec probabilités $\dfrac{1}{2}$ chacune, sa moyenne est nulle et sa variance est x^2.

Définition 10 - 4

Si X admet un moment d'ordre 2, alors $X - \mathbf{E}(X)$ aussi puisque les constantes ont des moments d'ordre 2, et on dit que X admet une variance. On définit celle-ci par la formule

$$\mathrm{V}(X) = \mathbf{E}\left((X - \mathbf{E}(X))^2\right) .$$

Par positivité de l'espérance $\mathrm{V}(X)$ est positif. On définit alors l'écart-type de X par $\sigma(X) = \sqrt{\mathrm{V}(X)}$.

Remarque 10 - 3

Si X a une dimension, alors $\sigma(X)$ a la même dimension.

♥

On a $\mathrm{V}(X) = 0$ si et seulement si X est presque sûrement égal à une constante.

Exemples 10 - 6

Variables de Poisson et géométrique

Si $X \sim \mathcal{P}(\lambda)$, on a déjà vu que X admet un moment de tous ordres. On a de plus $\mathbf{E}(X) = \lambda$ et donc $(X - \mathbf{E}(X))^2 = X^2 - 2\lambda X + \lambda^2$. On en déduit $\mathrm{V}(X) = \lambda(\lambda+1) - 2\lambda^2 + \lambda^2$, i.e. $\mathrm{V}(X) = \lambda$.

Si $X \sim \mathcal{G}(p)$, on a, par positivité,

$$\mathbf{E}\left(X^2\right) = \sum_{n=1}^{+\infty} n^2 p(1-p)^{n-1} .$$

En reprenant le calcul fait pour l'espérance, on s'intéresse au produit de Cauchy de $\sum(n+1)x^n$ et $\sum x^n$, ce qui est licite pour x dans $]0;1[$, et on obtient $\displaystyle\sum_{n=0}^{+\infty} \frac{(n+1)(n+2)}{2} x^n = \frac{1}{(1-x)^3}$ ou encore $\displaystyle\sum_{n=1}^{+\infty} n(n+1)x^{n-1} = \frac{2}{(1-x)^3}$. On en déduit

$$\mathbf{E}\left(X^2\right) = \frac{2p}{(1-(1-p))^3} - \mathbf{E}(X) = \frac{2}{p^2} - \frac{1}{p} = \frac{2-p}{p^2}$$

et donc $\mathrm{V}(X) = \mathbf{E}\left(\left(X - \frac{1}{p}\right)^2\right) = \frac{2-p}{p^2} - \frac{2}{p}\mathbf{E}(X) + \frac{1}{p^2}$, i.e. $\mathrm{V}(X) = \frac{1-p}{p^2}$.

Proposition 10 - 6

König-Huyghens

La variable aléatoire X admet une variance si et seulement si elle admet un moment d'ordre 2, et dans ce cas on a $\mathrm{V}(X) = \mathbf{E}\left(X^2\right) - \mathbf{E}(X)^2$.

Samuel König, 1712–1756, disciple de Johann Bernoulli et professeur d'Émilie du Châtelet. Christiaan Huygens, 1629–1695.

Démonstration. Posons $Y = X - \mathbf{E}(X)$. Puisque $\mathbf{E}(X)$ admet un moment d'ordre 2 et qu'on a affaire à un espace vectoriel, Y et X ont simultanément un moment d'ordre 2 ou non. De plus on a $Y^2 = X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2$ et donc, par linéarité, $\mathbf{E}\left(Y^2\right) = \mathbf{E}\left(X^2\right) - 2\mathbf{E}(X)\mathbf{E}(X) + \mathbf{E}(X)^2$, d'où la formule recherchée □

Proposition 10 - 7

Soit a et b deux réels. Si a est non nul $aX+b$ admet une variance si et seulement si X en admet une. De plus $\mathrm{V}(aX+b) = a^2\mathrm{V}(X)$, et cette formule est encore vraie si $a = 0$.

On a également $\sigma(aX+b) = |a|\,\sigma(X)$.

Démonstration. L'équivalence résulte du fait que les variables admettant une variance forment un espace vectoriel contenant les constantes. De plus, par linéarité de l'espérance, on a $\mathbf{E}(aX+b) = a\mathbf{E}(X) + b$ et donc

$$\mathrm{V}(aX+b) = \mathbf{E}\left((aX+b-a\mathbf{E}(X)-b)^2\right) = \mathbf{E}\left(a(X-\mathbf{E}(X))^2\right) = a^2\mathrm{V}(X) .$$

□

Définition 10 - 5

On dit que X est centré s'il admet une espérance et si $\mathbf{E}(X) = 0$.

On dit que X est réduit s'il admet un moment d'ordre 2 et si $\mathrm{V}(X) = 1$.

Exemple 10 - 7

Si $\sigma(X) \neq 0$, alors $\dfrac{X - \mathbf{E}(X)}{\sigma(X)}$ est une variable centrée réduite. Par exemple si $X \sim \mathcal{G}(p)$, alors $\dfrac{pX-1}{\sqrt{1-p}}$ est centré réduit.

Si a et b sont deux réels, avec a non nul, la variable centrée réduite associée à $aX+b$ est au signe de a près celle associée à X, i.e. $\dfrac{(aX+b) - \mathbf{E}(aX+b)}{\sigma(aX+b)} = \operatorname{sgn}(a)\dfrac{X-\mathbf{E}(X)}{\sigma(X)}$.

Théorème 10 - 3

Inégalité de Bienaymé-Tchebychev

Si X admet une variance, en posant $\mu = \mathbf{E}(X)$ et $\sigma = \sigma(X)$, on a pour tout ε strictement positif

$$\mathbf{P}(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}\ .$$

Irénée-Jules Bienaymé, 1796–1878, et Pafnouti Lvovitch Tchebychev, 1821–1894.

Démonstration. On applique l'inégalité de Markov générale à $|X-\mu|$ pour la fonction carré de $\mathbf{R}_+$ dans lui-même. C'est une fonction croissante et positive. Pour $\varepsilon > 0$, on a $\varepsilon^2 > 0$ et il vient $\mathbf{P}(|X-\mu| \geq \varepsilon) \leq \dfrac{\mathbf{E}\left((X-\mu)^2\right)}{\varepsilon^2}$, ce qui est l'inégalité de Bienaymé-Tchebychev. □

Remarque 10 - 4

L'inégalité a été formulée par Bienaymé, ami et traducteur de Tchebychev. Ce dernier s'en est servi pour préciser les travaux de Siméon Denis Poisson et donner une démonstration de la loi faible des grands nombres.

Néanmoins cette inégalité est en général assez grossière.

Exemple 10 - 8

Variable binomiale

On lance n fois un dé équilibré et on note X_n le nombre de fois où une face donnée a été tirée. On cherche si la probabilité de l'événement $\dfrac{X_n}{n} \leq \dfrac{1}{3}$ est grande : supérieure à $\dfrac{1}{2}$ ou 95%. On a $X_n \sim \mathcal{B}\left(n, \dfrac{1}{6}\right)$ et donc $X_n \leq \dfrac{n}{3}$ s'écrit $|X_n - \mathbf{E}(X_n)| \leq \dfrac{n}{6}$.

Comme $\mathrm{V}(X_n) = n\dfrac{1}{6}\dfrac{5}{6} = \dfrac{5n}{36}$, on a $\mathbf{P}\left(|X_n - \mathbf{E}(X_n)| > \dfrac{n}{6}\right) \leq \dfrac{5}{n}$ d'après l'inégalité de Bienaymé-Tchebychev avec $\sigma^2 = \dfrac{5n}{36}$ et $\varepsilon^2 = \dfrac{n^2}{36}$.

On obtient donc $n = 10$ pour garantir $\mathbf{P}\left(|X_n - \mathbf{E}(X_n)| \leq \dfrac{n}{6}\right) \geq \dfrac{1}{2}$ et $n = 100$ pour $\mathbf{P}\left(|X_n - \mathbf{E}(X_n)| \leq \dfrac{n}{6}\right) \geq 0,95$.

Mais ces estimations sont en fait assez grossières puisque la première inégalité est tout le temps vérifiée et la seconde l'est à partir de $n = 12$ seulement !

Exercices

Variables aléatoires

10 - 1 Ⓢ ★ Espérance

Soit $(A_i)_{i\in I}$ une partition de Ω, avec I au plus dénombrable, et X donné par $X = \sum_{i\in I} \alpha_i \mathbb{1}_{A_i}$ avec $(\alpha_i)_{i\in I}$ une famille de nombres réels, i.e. $\omega \in A_i \Longrightarrow X(\omega) = \alpha_i$. Montrer que X est une variable aléatoire réelle discrète et que, si elle est sommable, alors on a $\mathbf{E}(X) = \sum_{i\in I} \alpha_i \mathbf{P}(A_i)$.

10 - 2 Ⓢ ★ Méthode probabiliste d'Erdös

Un champ est entouré de dix-sept poteaux, dont cinq exactement sont pourris. Calculer l'espérance du nombre de poteaux pourris parmi sept poteaux consécutifs et en déduire qu'il existe sept poteaux consécutifs dont trois au moins sont pourris.

10 - 3 Ⓢ ★ Couples

Parmi les $2n$ personnes formant n couples, on sélectionne m personnes aléatoirement (uniformément et indépendamment). Trouver la probabilité pour qu'un couple ait été sélectionné et en déduire l'espérance du nombre de couples sélectionnés.

10 - 4 Ⓢ ★ Échanges

On place $2n$ personnes dans deux salles, chacune ayant n places. À chaque étape on sélectionne deux personnes, une dans chaque salle, (uniformément et indépendamment) et elles échangent leurs places. Trouver la probabilité pour qu'une personne soit dans sa salle de départ au bout de k étapes et en déduire l'espérance du nombre de personnes n'ayant pas changé de salle après k étapes.

10 - 5 Ⓢ ★ Connecteurs

Soit G un graphe fini. Pour tout ensemble V de sommets de G et toute arête e, on dit que e connecte V et $\overline{V}$ si les sommets de e sont l'un dans V et l'autre non. On note N_V le nombre d'arêtes qui connectent V et $\overline{V}$. On note E le nombre d'arêtes de G. Montrer que, pour un ensemble de sommets V choisi aléatoirement, e connecte V à $\overline{V}$ avec probabilité $\frac{1}{2}$ et en déduire qu'il existe un ensemble V de sommets tel que $N_V \geq \frac{1}{2}E$.

10 - 6 Ⓢ ★★ Stop ou encore

On lance un dé à six faces équilibré. Si on obtient un 1 on doit arrter. Sinon on peut choisir de continuer. Le gain est indiqué par la valeur du dé au moment de l'arrêt.

a. On note $S(r)$ le gain moyen si on choisit de ne s'arrêter volontairement que si la valeur du dé est supérieure à r. Calculer $S(r)$ pour r dans $[\![1;6]\!]$ et en déduire la meilleure stratégie dans ce jeu.

b. On modifie le jeu en imposant un coût à chaque lancer de dé. Quelle est la meilleure stratégie si ce coût est $\frac{1}{3}$? s'il est 1 ?

c. On modifie le premier jeu en donnant comme gain le carré de la valeur du dé. Quelle est alors la meilleure stratégie ?

10 - 7 Ⓢ C 2015 ★★ Un équivalent

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbf{P})$ admettant une espérance et telle que $X(\Omega) = \mathbf{N}$. Soit $S_n = \sum_{k=1}^{n} \mathbf{P}(X < k)$. Montrer $S_n \sim n$ et préciser cette propriété asymptotique.

10 - 8 Ⓢ M 2017 ★★ Variable géométrique

a. On considère X une variable aléatoire discrète à valeurs dans $\mathbf{R}_+^*$, et on pose la variable aléatoire $Y = X + \frac{1}{X}$. On suppose que Y est d'espérance finie. Montrer $\mathbf{E}(Y) \geq 2$.

b. On suppose que X suit une loi géométrique de paramètre p dans $]0;1[$.

i. Rappeler la définition d'une loi géométrique, et donner son espérance.

ii. Montrer que Y est d'espérance finie et calculer son espérance.

iii. Quelle est la limite de $\mathbf{E}(Y)$ quand p tend vers 1 ?

iv. On note ℓ cette limite. Calculer un équivalent de $\mathbf{E}(Y) - \ell$ en 1.

10 - 9 Ⓢ M 2015 ★★★ Loi ζ

Sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, soit X une variable aléatoire à valeurs dans $\mathbf{N}^*$ telle que $\mathbf{P}(X = n) = \frac{1}{\zeta(a)n^a}$, où a est un réel strictement plus grand que 1.

a. Justifier la cohérence de la définition de la loi de probabilité.

b. À quelle condition X admet-elle une espérance ? La calculer et en donner un équivalent quand a tend vers $+\infty$.

c. Pour i dans $\mathbf{N}^*$, on note $A_i = \{n \in \mathbf{N}^* \mid i \text{ divise } n\}$. Calculer $\mathbf{P}(X \in A_i)$. À quelle condition deux A_i sont-ils indépendants ?

d. Soit r dans $\mathbf{N}^*$, $p_1, \ldots, p_r$ des nombres premiers distincts et

$$C_r = \{n \in \mathbf{N}^* \mid \forall i \in [\![1;r]\!] \ p_i \text{ ne divise pas } n\} \ .$$

Calculer $\mathbf{P}(X \in C_r)$. Que dire quand r tend vers $+\infty$?

Moments

10 - 10 Ⓢ ★ Paradoxe de la médiane

Justifier que $99,9\%$ des gens ont un nombre de jambes strictement supérieur à la moyenne.

10 - 11 Ⓢ ★ Dé

Calculer la variance du nombre obtenu lors d'un lancer d'un dé à six faces équilibré.

10 - 12 Ⓢ TPE 2017 ★ Variable très réduite

Soit X une variable aléatoire réelle discrète, admettant un moment d'ordre 4, telle que $E(X^2) = E(X^4) = 1$. On note x son espérance. Montrer $|x| \leq 1$ et déterminer la loi de X.

(Couplé avec 214.)

10 - 13 Ⓢ ★ Questions pour ... maximiser

Lors d'un jeu, on pose deux questions. On suppose qu'on connaît les probabilités p_1 et p_2 de répondre à ces deux questions respectivement. On choisit de commencer le jeu par l'une des questions. Si on répond correctement on a alors le droit de tenter de répondre à la seconde, mais sinon le jeu s'arrête. Chaque réponse correcte rapporte un gain égal à g_1 et g_2 respectivement. Par quelle question doit-on commencer pour maximiser l'espérance de gain ?

10 - 14 Ⓢ CCP 2015 ★ Jeu

Dans un jeu, on tire un nombre entier n supérieur ou égal à 1 avec une probabilité donnée par $\mathbf{P}(n) = \dfrac{1}{2^n}$. Si n est pair, on gagne n €; sinon on perd n €.

a. Quelle est la probabilité de gain ?

b. Calculer l'espérance et la variance de ce gain.

10 - 15 Ⓢ ★ Gestion de stocks

Un produit de saison rapport un bénéfice b par vente et une perte p par invendu. On note X le nombre d'unités de ce produit achetées dans un magasin, où X est à valeurs dans $\mathbf{N}^*$ et on cherche à savoir comment ce magasin doit optimiser son stock afin de maximiser son espérance de gain. On note s la taille du stock et Y_s le gain associé.

a. Montrer qu'on a $Y_s = sb$ si $X > s$ et $Y_s = bX - (s - X)p$ sinon.

b. En déduire $\mathbf{E}(Y_s) = sb + (b+p)\displaystyle\sum_{k=0}^{s}(k-s)\mathbf{P}(X = k)$.

c. Montrer

$$\mathbf{E}(Y_{s+1} - Y_s) = b - (b+p)\sum_{k=0}^{s}\mathbf{P}(X = k) \ .$$

d. Conclure que s doit être choisi minimal parmi les entiers k tels que $\mathbf{P}(X \leq k) \geq \dfrac{b}{b+p}$.

10 - 16 Ⓢ ★ Paradoxe de Saint-Petersbourg

Un jeu consiste à lancer une pièce équilibrée jusqu'à ce que *pile* apparaisse pour la première fois. Si c'est au n^{e} coup, on gagne 2^n euro. Le prix pour pouvoir jouer est de 1 million d'euro. On note X la variable aléatoire correspondant au gain à ce jeu, pour une partie.

a. Calculer $\mathbf{E}(X)$.

b. Seriez-vous prêt(e) à jouer une partie ?

c. Seriez-vous prêt(e) à jouer autant de parties que vous le souhaitez en vous acquittant des droits d'entrée à l'arrêt du jeu ?

10 - 17 Ⓢ ★ Roue de la fortune

Un jeu invite à miser une certaine somme S sur un nombre entre 1 et 6, puis à lancer trois dés. Si aucun des dés ne tombe sur le nombre choisi, la mise est perdue. Sinon la mise est récupérée et de plus elle est gagnée autant de fois que le nombre apparaît sur les dés. Ce jeu est-il favorable à la personne qui mise ?

10 - 18 Ⓢ ★ Variable de Poisson

Soit X avec $X \sim \mathcal{P}(\lambda)$. Montrer $\mathbf{P}\left(X \leq \dfrac{\lambda}{2}\right) \leq \dfrac{4}{\lambda}$ et $\mathbf{P}(X \geq 2\lambda) \leq \dfrac{1}{\lambda}$.

10 - 19 Ⓢ ★ Inverse d'une loi binomiale

Soit X avec $X \sim \mathcal{B}(n,p)$. Calculer $\mathbf{E}\left(\dfrac{1}{X+1}\right)$.

Indication : On pourra utiliser $\dfrac{1}{k+1}\dbinom{n}{k} = \dfrac{1}{n+1}\dbinom{n+1}{k+1}$. *Une autre interprétation sera donnée en utilisant l'espérance conditionnelle.*

10 - 20 Ⓢ ★ Inverse d'une variable géométrique

Soit X avec $X \sim \mathcal{G}(p)$. Calculer $\mathbf{E}\left(\dfrac{1}{X}\right)$.

10 - 21 Ⓢ ★★ Variance

Soit X admettant un moment d'ordre 2.

a. Montrer $\min\limits_{a \in \mathbf{R}} \mathbf{E}\left((X-a)^2\right) = \mathrm{V}(X)$ et donner les points où ce minimum est atteint.

b. On appelle écart moyen absolu de X le réel $d(X)$ défini par $d(X) = \mathbf{E}(|X - \mathbf{E}(X)|)$. Montrer $d(X) \leq \sigma(X)$.

c. On suppose $X \sim \mathcal{B}\left(2n, \frac{1}{2}\right)$. Montrer $d(X) = n\binom{2n}{n}2^{-2n}$ et en déduire qu'on a $d(X) \sim \sqrt{\frac{2}{\pi}}\sigma(X)$ lorsque n tend vers l'infini.

10 - 22 Ⓢ ★★ Moments d'une variable entière

Soit X une variable aléatoire à valeurs dans $\mathbf{N}$. Montrer $\mathbf{E}\left(X\right) = \sum_{n\in\mathbf{N}} \mathbf{P}(X > n)$ et $\mathbf{E}\left(X^2\right) = \sum_{n\in\mathbf{N}}(2n+1)\mathbf{P}(X > n)$.

10 - 23 Ⓢ ★★ Inégalité de Hölder

Soit X et Y des variables aléatoires réelles et discrètes sur un même espace probabilisé. Soit p un réel supérieur à 1 et q donné par $\frac{1}{p} + \frac{1}{q} = 1$. Montrer $\mathbf{E}\left(|XY|\right) \leq \mathbf{E}\left(|X|^p\right)^{1/p}\mathbf{E}\left(|Y|^q\right)^{1/q}$.

10 - 24 Ⓢ ★★ Distribution de Yule-Simmons

Soit X une variable aléatoire à valeurs dans $\mathbf{N}^*$ dont la loi est donnée par $\mathbf{P}(X = n) = \dfrac{4}{n(n+1)(n+2)}$.

a. Montrer qu'on a bien affaire à une loi de probabilité.

b. Calculer $\mu_1(X)$ et $\mu_2(X)$.

10 - 25 Ⓢ ★★ Loi binomiale négative

On reprend l'exercice 1 - 51. Soit X et Y suivant des lois binomiales négatives de paramètres r et p, et $r+1$ et p respectivement.

a. Pour k dans $\mathbf{N}$, montrer $\mu_k(X) = \frac{r}{p}\mu_{k-1}(Y-1)$.

b. En déduire l'espérance et la variance de X.

c. Calculer l'espérance et la variance du nombre de lancers de dé nécessaires pour obtenir quatre fois 1 sur un dé à six faces.