18 Variables aléatoires discrètes

« Ce calcul délicat s'étend aux questions les plus importantes de la vie, qui ne sont en effet, pour la plupart, que des problèmes de probabilité. »

Pierre-Simon, marquis de Laplace (1812)

Plan de cours

I | Variables aléatoires discrètes

A – Définition

On lance simultanément deux dés discernables et on choisit comme univers $\Omega = [\![1,6]\!]^2$, que l'on munit de la probabilité uniforme. Notons alors X la somme des valeurs des dés et Y le maximum des deux valeurs.

- Quelles sont les valeurs possibles pour X et pour Y ?
  X peut prendre toutes les valeurs entières entre 2 et 12 ; Y entre 1 et 6.
- On peut voir X et Y comme des fonctions :

$$X : \left| \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & \omega_1 + \omega_2 \end{array} \right. \qquad \text{et} \qquad Y : \left| \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & \max(\omega_1, \omega_2) \end{array} \right. \qquad \text{où} \quad \omega = (\omega_1, \omega_2)$$

- On peut alors écrire $X(\Omega) = [\![2,12]\!]$ et $Y(\Omega) = [\![1,6]\!]$.
  Tout comme on note plus généralement $f(E)$ l'image de E par f pour $f : E \to F$...
- Notons A l'événement « la somme des valeurs obtenues vaut 5 ».
  On a $A = \{(1,4),(2,3),(3,2),(4,1)\} \in \mathscr{P}(\Omega)$. Donc :

$$\mathbf{P}(A) = \mathbf{P}(X=5) = \frac{4}{36} = \frac{1}{9}$$

  Dans l'égalité précédente, on a noté $(X=5)$ l'événement A. Il convient de voir $(X=5)$ comme une abréviation pratique de $X^{-1}(\{5\}) = \{\omega \in \Omega \mid X(\omega) = 5\}$.

De manière plus générale, $(X = i) = X^{-1}(\{i\}) = \{\omega \in \Omega \mid X(\omega) = i\}$ pour $i \in X(\Omega)$ désigne ici un événement (une partie de Ω) dont on peut calculer la probabilité.

Définition 18.1 : Variable aléatoire discrète

Soit $(\Omega, \mathscr{A})$ un espace probabilisable. On appelle variable aléatoire réelle discrète toute application $X : \Omega \to \mathbb{R}$ telle que :

- $X(\Omega)$ est un ensemble fini ou dénombrable ;
- Si $x \in X(\Omega)$ alors $X^{-1}(\{x\}) = (X = x)$ est un événement, c'est-à-dire :

$$\forall x \in X(\Omega) \quad X^{-1}(\{x\}) \in \mathscr{A}$$

Cette définition appelle plusieurs remarques :

- Malgré son nom, une variable aléatoire n'est pas une *variable* (c'est une fonction) et elle n'est pas *aléatoire*.
- $X(\Omega)$ est l'image (directe) de Ω par X. Si Ω est un univers fini, il en va de même pour $X(\Omega)$.
- Si $A \subset \mathbb{R}$, $X^{-1}(A)$ désigne l'image réciproque de A par X :

$$X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$$

  Notons que la notation $X^{-1}(A)$ ne sous-entend aucunement la bijectivité de X.
- Comme dans l'exemple précédent, on souhaitera calculer la probabilité de l'événement $X^{-1}(A)$ encore noté $(X \in A)$. Encore faut-il que cette partie de Ω soit bien un événement, c'est-à-dire qu'elle soit contenue dans la tribu $\mathscr{A}$. C'est finalement ce que nous assure la définition mais nous reviendrons sur ce point plus tard.

Rappel sur les images réciproques :

Exercice 1

Notons f (resp. g) la fonction définie sur $\mathbb{R}_+^*$ (resp. $\mathbb{R}$) par $f(t) = \ln(t)$ et $g(t) = t^2$.
Déterminer les ensembles suivants :

$$f^{-1}(\mathbb{R})\,; \quad f^{-1}(\mathbb{R}_+)\,; \quad f^{-1}(\mathbb{R}_+^*)\,; \quad f^{-1}([0,1[)\,; \quad g^{-1}([0,a])\,; \quad g^{-1}(]-a,a[)\,; \quad g^{-1}(\mathbb{R}_+)$$

Dans le cadre de ce chapitre, nous aurons recours aux notations suivantes (pour $x \in \mathbb{R}$ et $A \subset \mathbb{R}$) :

$$\begin{aligned}
(X \in A) &= X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\};\\
(X = x) &= X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\};\\
(X \leqslant x) &= X^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leqslant x\};\\
(X < x) &= X^{-1}(]-\infty, x[) = \{\omega \in \Omega \mid X(\omega) < x\};\\
(X \geqslant x) &= X^{-1}([x, +\infty[) = \{\omega \in \Omega \mid X(\omega) \geqslant x\};\\
(X > x) &= X^{-1}(]x, +\infty[) = \{\omega \in \Omega \mid X(\omega) > x\}.
\end{aligned}$$

Remarquons que si le premier ensemble constitue un événement, ce sera le cas de tous les autres.

Exercice 2

Dans le cadre du lancer de dés précédent, expliciter les événements suivants :

$$(X > 10)\,; \quad (X \geqslant 11)\,; \quad (X = 4)\,; \quad (X \leqslant 1)\,; \quad (Y = 3)\,; \quad (Y \in \{1,2\})$$

La définition d'une variable aléatoire impose à $(X = x)$ pour $x \in X(\Omega)$ quelconque d'être un événement. Mais c'est également le cas pour $(X \in A)$ avec $A \subset X(\Omega)$ quelconque.

Proposition 18.2

Soient $(\Omega, \mathscr{A})$ un espace probabilisable et X une variable aléatoire discrète sur cet espace. Soit A un sous-ensemble quelconque de $\mathbb{R}$. L'ensemble $X^{-1}(A)$, noté $(X \in A)$, est un événement.

Démonstration

Tout d'abord, remarquons que $X^{-1}(A) = X^{-1}(A \cap X(\Omega))$ car :

$$X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\} = \{\omega \in \Omega \mid X(\omega) \in A \cap X(\Omega)\}$$

Donc seule l'intersection de A et de $X(\Omega)$ nous intéresse. Or, $A \cap X(\Omega)$ est au plus dénombrable donc on peut le décrire sous la forme :

$$A \cap X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$$

Ainsi, $(X \in A) = \bigcup_{n \in \mathbb{N}} X^{-1}(\{x_n\}) = \bigcup_{n \in \mathbb{N}} (X = x_n)$. $(X \in A)$ est donc la réunion dénombrable d'événements (par définition d'une variable aléatoire), c'est donc un événement. ■

On pourra noter que lorsque Ω est lui-même dénombrable, c'est automatiquement le cas pour $X(\Omega)$. Dans ce cas, on munira alors Ω de sa tribu naturelle $\mathscr{P}(\Omega)$. Toutes les parties de Ω seront alors des événements. Il n'y aura donc pas à « vérifier » que X est bien une variable aléatoire : toute application de Ω dans $\mathbb{R}$ sera une variable aléatoire. Il se peut cependant que Ω ne soit pas dénombrable mais, en pratique, on demandra rarement de prouver que X est bien une variable aléatoire. Ce sera souvent l'énoncé d'un problème qui l'admettra ou bien qui n'abordera même pas la question...

Cette année, toutes les variables aléatoires seront supposées finies ou discrètes. Dans les deux cas de figure, $X(\Omega)$ sera au plus dénombrable ce qui nous autorise désormais à l'écrire sous la forme $X(\Omega)=\{x_n \mid n\in\mathbb{N}\}$.

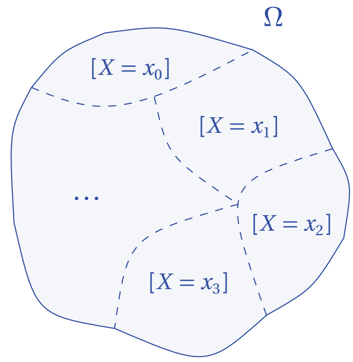
Théorème 18.3

Soient $(\Omega,\mathscr{A})$ un espace probabilisable et X une variable aléatoire discrète sur cet espace.
Avec les notations précédentes, la famille $((X=x_n))_{n\in\mathbb{N}}$ est un système complet d'événements.

Démonstration

Justifions le fait que $\displaystyle\bigsqcup_{n=0}^{+\infty}(X=x_n)=\Omega$.

- Montrons que les événements $[X=x_n]$ sont deux à deux incompatibles, en supposant les x_n deux à deux distincts.
  Soient $i,j\in\mathbb{N}$ avec $i\neq j$ et $\omega\in(X=x_i)\cap(X=x_j)$.
  On a alors $X(\omega)=x_i=x_j$. Absurde, donc $(X=x_i)\cap(X=x_j)=\varnothing$.
- On a $\displaystyle\bigcup_{n=0}^{+\infty}(X=x_n)\subset\Omega$. De plus, si $\omega\in\Omega$ alors il existe $i\in\mathbb{N}$ tel que $X(\omega)=x_i$ et on a donc $\omega\in(X=x_i)$. D'où $\displaystyle\Omega\subset\bigcup_{n=0}^{+\infty}(X=x_n)$, ce qui prouve que $\displaystyle\bigcup_{n=0}^{+\infty}(X=x_n)=\Omega$. ■



$((X=x_n))_{n\in\mathbb{N}}$ est un s.c.e.

Exercice 3

Décrire le système complet d'événements associé à la variable aléatoire Y introduite dans l'exemple précédent.

B – Loi d'une variable aléatoire

On munit désormais l'espace probabilisable $(\Omega,\mathscr{A})$ d'une probabilité $\mathbf{P}$.

Définition 18.4 : Loi d'une variable aléatoire

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$. On appelle loi de probabilité de X (ou plus simplement loi de X) et on note P_X l'application :

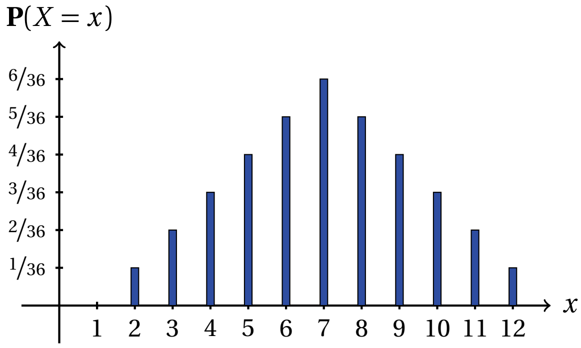
$$P_X : \begin{array}{rcl} X(\Omega)\subset\mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \mathbf{P}(X=x) \end{array}$$

Notons qu'une loi de probabilité est une fonction numérique à variable réelle. On peut étendre P_X à $\mathbb{R}$ car si $x\notin X(\Omega)$, $(X=x)=\varnothing$ qui est de probabilité nulle. On peut représenter une loi de probabilité par un tableau ou par un diagramme en bâtons.

Exemple

Représenter graphiquement la loi de probabilité de la variable aléatoire X dans le cas du lancer de dés. On munira pour cela $(\Omega,\mathscr{P}(\Omega))$ de la probabilité uniforme.

x_i	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}(X=x_i)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$



Faire de même avec la la loi de probabilité de la variable aléatoire Y.

Exercice 4

Une pièce amène pile avec la probabilité p et face avec la probabilité $1-p$, $0<p<1$. On la lance n fois de suite. Soit X le nombre de fois où pile apparaît au cours de ces lancers. Chercher la loi de X.

D'après le théorème précédent, $\displaystyle\sum_{x\in X(\Omega)} \mathbf{P}(X=x) = \sum_{n=0}^{+\infty} \mathbf{P}(X=x_n) = 1.$

Le théorème suivant, dont on admet la preuve, constitue une sorte de réciproque et permet surtout de définir une variable aléatoire par sa loi de probabilité sans avoir à étudier l'expérience aléatoire sous-jacente.

Théorème 18.5

Soient $\{x_n \mid n\in\mathbb{N}\}$ un ensemble infini dénombrable, les x_n étant deux à deux distincts, et $(p_n)_{n\in\mathbb{N}}$ une famille de réels positifs telle que $\displaystyle\sum_{n=0}^{+\infty} p_n = 1$. Alors il existe une variable aléatoire discrète X sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ à valeurs dans $\{x_n \mid n\in\mathbb{N}\}$ telle que : $\quad \forall n\in\mathbb{N} \quad \mathbf{P}(X=x_n)=p_n$

Exemple

À quelle condition sur α les réels $p_n = \alpha\cdot\dfrac{\lambda^n}{n!}$ $(n\geqslant 0)$ sont-ils les coefficients d'une loi de probabilité ?

Comme $\displaystyle\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!} = \mathrm{e}^{\lambda}$, on en conclut que $\displaystyle\sum_{n=0}^{+\infty} p_n = 1$ si et seulement si $\alpha = \mathrm{e}^{-\lambda}$. On a bien alors $p_n \geqslant 0$.

Deux variables aléatoires X et Y ont même loi si $X(\Omega)=Y(\Omega)$ et si pour tout $x\in X(\Omega)$, $\mathbf{P}(X=x)=\mathbf{P}(Y=x)$. Attention, cela ne signifie pas que $X=Y$, c'est-à-dire que $X(\omega)=Y(\omega)$ pour tout $\omega\in\Omega$.

Exercice 5

On considère une pièce équilibrée que l'on lance 4 fois. On note X le nombre de piles obtenus et Y le nombre de faces obtenues. X et Y ont la même loi mais ne sont pas identiques.

C – Opérations sur les variables aléatoires (complément)

Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Posons $X(\Omega)=\{x_n \mid n\in\mathbb{N}\}$.
Considérons maintenant l'application $\varphi : X(\Omega)\to\mathbb{R}$ et l'application $Y=\varphi\circ X$ traditionnellement notée $\varphi(X)$.
D'une part, $Y:\Omega\to\mathbb{R}$. D'autre part, $Y(\Omega)=\varphi(X(\Omega))=\{\varphi(x_n)\mid n\in\mathbb{N}\}$. C'est un ensemble au plus dénombrable.
Il ne reste plus qu'à prouver que Y est une variable aléatoire. En effet,

$$\forall y\in Y(\Omega),\quad (Y=y) = \bigsqcup_{\substack{n\in\mathbb{N}\\ \varphi(x_n)=y}} (X=x_n)$$

car $\omega\in(Y=y)$ si et seulement si $Y(\omega)=y$, c'est-à-dire si et seulement s'il existe $n\in\mathbb{N}$ tel que $X(\omega)=x_n$ et $\varphi(x_n)=y$. Cela revient à dire que $\omega\in(X=x_n)$ avec $\varphi(x_n)=y$.
$(Y=y)$ est donc la réunion au plus dénombrable d'événements, c'est un événement.

Ceci nous assure que toute fonction d'une variable aléatoire est encore une variable aléatoire. Résultat que l'on peut généraliser dans le cas de fonctions à plusieurs variables.

Exemples

Si X et Y sont deux variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, il en va de même pour X^2, $X+Y$, XY, $\sin(Y)$, $\min(X,Y)$.

Remarquons que dans le cas où $Y = \varphi(X)$, on obtient facilement la loi de Y à partir de celle de X :

$$\forall y \in Y(\Omega) \quad \mathbf{P}(Y=y) = P\left(\bigsqcup_{\substack{n\in\mathbb{N}\\ \varphi(x_n)=y}} (X=x_n)\right) = \sum_{\substack{n\in\mathbb{N}\\ \varphi(x_n)=y}} \mathbf{P}(X=x_n)$$

Exercice 6

Déterminer la loi de X^2 où X est une variable aléatoire dont la table de probabilité est :

x_i	-2	-1	0	1	2
$\mathbf{P}(X=x_i)$	0.10	0.35	0.15	0.25	0.15

Exercice 7

Soit X une variable aléatoire réelle discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ telle que $X(\Omega) = \mathbb{N}^*$ et $\mathbf{P}(X = k) = \dfrac{\alpha}{k(k+1)(k+2)}$ pour tout entier naturel non nul k.

1. On pose $u_k = \dfrac{1}{2k} - \dfrac{1}{2(k+1)}$. Montrer que pour tout $k \in \mathbb{N}^*$, on a :

$$u_k - u_{k+1} = \frac{1}{k(k+1)(k+2)}$$

   En déduire la valeur de α.
2. Montrer que $Y = X+1$ est une variable aléatoire réelle discrète et déterminer sa loi.
3. Montrer que $Z = X^2 - 4X + 4$ est une variable aléatoire réelle discrète et déterminer sa loi.

II | Moments d'une variable aléatoire

A – Espérance

L'an dernier, toutes les variables réelles étudiées étaient finies. En notant X une telle variable, on pouvait alors écrire $X(\Omega) = \{x_1, \ldots, x_n\}$. L'espérance de X était ensuite définie par :

$$\mathbf{E}(X) = \sum_{x\in X(\Omega)} x \cdot \mathbf{P}(X=x) = \sum_{k=1}^{n} x_k \cdot \mathbf{P}(X=x_k)$$

L'espérance correspond à la moyenne des valeurs prises par X pondérées par leur probabilité d'apparition.

De plus, si X est à valeurs dans $[\![1,n]\!]$, $X(\Omega) = \{1, \ldots, n\}$ et $\mathbf{E}(X) = \sum_{k=1}^{n} k \cdot \mathbf{P}(X=k)$.

Exemples

Voici deux exemples classiques qui permettent d'illustrer l'espérance comme le gain espéré à un jeu donné.

- On lance une pièce équilibrée et on note X la variable aléatoire qui prend la valeur 1 si l'on obtient pile, 0 si on obtient face.
  On a $X(\Omega) = \{0,1\}$. La pièce étant équilibrée, on munit $\Omega = \{P, F\}$ de la probabilité uniforme.

$$\mathbf{E}(X) = \sum_{x\in X(\Omega)} x_k \mathbf{P}(X=x_k) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

- On lance deux dés non truqués dont les faces sont numérotées de 1 à 6 et on note X la somme des chiffres obtenus. La loi de probabilité de X est donné par le tableau suivant :

x_i	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}(X=x_i)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

L'espérance de X vaut alors :

$$\mathbf{E}(X)=\sum_{x\in X(\Omega)} x_k\mathbf{P}(X=x_k)=\frac{2}{36}+\frac{3}{18}+\frac{4}{12}+\frac{5}{9}+\frac{30}{36}+\frac{7}{6}+\frac{40}{36}+\frac{9}{9}+\frac{10}{12}+\frac{11}{18}+\frac{12}{36}=7$$

Pour donner un sens à $\sum_{x\in X(\Omega)} x\cdot\mathbf{P}(X=x)$, nous sommes confrontés à deux difficultés :

- $X(\Omega)$ étant au plus dénombrable, on peut l'écrire sous la forme $\{x_n \mid n\in\mathbb{N}\}$.
  Dans le cas où $X(\Omega)$ est dénombrable, la série $\sum_{n\in\mathbb{N}} x_n\cdot\mathbf{P}(X=x_n)$ converge-t-elle ?
- La somme ne dépend-elle pas du choix de la bijection entre $X(\Omega)$ et $\mathbb{N}$, autrement dit de l'ordre de sommation ?

La réponse à la deuxième question est non, du moment que la série est supposée absolument convergente : la famille est alors sommable et on peut appliquer le théorème de convergence commutative.

Définition 18.6 : Espérance mathématique

Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$. On pose $X(\Omega)=\{x_n \mid n\in\mathbb{N}\}$. X est dite d'espérance finie si la série $\sum_{n\in\mathbb{N}} x_n\cdot\mathbf{P}(X=x_n)$ converge absolument.
Dans ce cas, on appelle espérance de X et on note $\mathbf{E}(X)$ le réel :

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

On remarquera que toute variable aléatoire finie admet une espérance.

Exercice 8

Dans un concours de saut, la probabilité qu'un sauteur passe la n^{e} barre est $1/n$, et est indépendante des sauts précédents. On note X la dernière barre que le sauteur a franchi avant d'échouer et A_n l'événement « le sauteur a franchi n^{e} barre ».

1. Déterminer la probabilité de l'événement « le sauteur a franchi une infinité de barres ».
   On pose alors $X(\Omega)=\mathbb{N}^*$.
2. Calculer $\mathbf{P}(X\geqslant n)$ puis déterminer la loi de X.
3. En déduire l'espérance de X.

Calculer l'espérance d'une variable aléatoire nécessite de savoir *a priori* déterminer sa loi, ce qui peut s'avérer délicat lorsque la variable est de la forme $\varphi(X)$. Le résultat suivant permet de calculer l'espérance de $\varphi(X)$ à partir de la loi de X.

Théorème 18.7 : Théorème de transfert

Soient X une variable aléatoire discrète sur l'espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$ et $f : X(\Omega)\to\mathbb{R}$. On note $X(\Omega)=\{x_n \mid n\in\mathbb{N}\}$. Alors, $f(X)$ est d'espérance finie si et seulement si $\sum_{n\in\mathbb{N}} f(x_n)\cdot\mathbf{P}(X=x_n)$ converge absolument et, dans ce cas,

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

Théorème 18.8 : Propriétés de l'espérance

Soient X et Y sont deux variables aléatoires discrètes d'espérances finies sur $(\Omega, \mathscr{A}, \mathbf{P})$.

(i) Linéarité de l'espérance
Pour tout $\lambda, \mu \in \mathbb{R}$, $\lambda X + \mu Y$ est d'espérance finie et $\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y)$.

(ii) Positivité de l'espérance
Si $X \geqslant 0$, c'est-à-dire si $X(\omega) \geqslant 0$ pour tout $\omega \in \Omega$, alors $\mathbf{E}(X) \geqslant 0$.

(iii) Croissance de l'espérance
Si $X \leqslant Y$, c'est-à-dire si $X(\omega) \leqslant Y(\omega)$ pour tout $\omega \in \Omega$, alors $\mathbf{E}(X) \leqslant \mathbf{E}(Y)$.

Démonstration

Nous admettons le premier résultat qui découle du théorème de transfert.

(ii) En écrivant $X(\Omega) = \{x_n \mid n \in \mathbb{N}\} \subset \mathbb{R}_+$, il vient $\mathbf{E}(X) = \displaystyle\sum_{n=0}^{+\infty} x_n \cdot \mathbf{P}(X = x_n)$, qui est la somme d'une série à termes positifs donc positive.

(iii) $Y - X$ est une variable positive, d'espérance $\mathbf{E}(Y - X) = \mathbf{E}(Y) - \mathbf{E}(X)$ donc $\mathbf{E}(Y) - \mathbf{E}(X) \geqslant 0$. ■

On notera que $X - \mathbf{E}(X)$ est une variable d'espérance nulle par linéarité de l'espérance. En effet, $\mathbf{E}(X) \in \mathbb{R}$ et $\mathbf{E}(X - \mathbf{E}(X)) = \mathbf{E}(X) - \mathbf{E}(\mathbf{E}(X)) = \mathbf{E}(X) - \mathbf{E}(X) = 0$. On parle alors de variable *centrée*.

Proposition 18.9

Soient X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathscr{A}, \mathbf{P})$.

(i) Si Y est d'espérance finie et $|X| \leqslant Y$, alors X est d'espérance finie.

(ii) Si $|X|$ est d'espérance finie, X aussi et $|\mathbf{E}(X)| \leqslant \mathbf{E}(|X|)$.

Proposition 18.10

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Si X admet une espérance finie, alors :

$$\mathbf{E}(X) = \sum_{n=1}^{+\infty} \mathbf{P}(X \geqslant n)$$

Démonstration

Comme X est à valeurs dans $\mathbb{N}$, $(X \geqslant n) = (X = n) \sqcup (X \geqslant n+1)$ pour tout $n \in \mathbb{N}$. Fixons $p \in \mathbb{N}^*$.

$$\begin{aligned}
\sum_{n=0}^{p} n \cdot \mathbf{P}(X = n) &= \sum_{n=0}^{p} n \cdot \left[\mathbf{P}(X \geqslant n) - \mathbf{P}(X \geqslant n+1)\right] \\
&= \sum_{n=0}^{p} n\mathbf{P}(X \geqslant n) - \sum_{n=0}^{p} (n+1)\mathbf{P}(X \geqslant n+1) + \sum_{n=0}^{p} \mathbf{P}(X \geqslant n+1) \\
&= \sum_{n=0}^{p} \mathbf{P}(X \geqslant n+1) - (p+1)\mathbf{P}(X \geqslant p+1) \qquad \text{(téléscopage)}
\end{aligned}$$

Ainsi, $\displaystyle\sum_{n=0}^{p} n \cdot \mathbf{P}(X = n) = \sum_{n=1}^{p} \mathbf{P}(X \geqslant n) - p\mathbf{P}(X \geqslant p+1)$.

Par ailleurs, $0 \leqslant p\mathbf{P}(X \geqslant p+1) = \displaystyle\sum_{n=p+1}^{+\infty} p\mathbf{P}(X = n) \leqslant \sum_{n=p+1}^{+\infty} n\mathbf{P}(X = n)$ et cette dernière quantité tend vers 0 comme reste d'une série convergente (la variable aléatoire étant d'espérance finie).

Bref, $p\mathbf{P}(X \geqslant p+1) \xrightarrow[p \to +\infty]{} 0$ donc en faisant tendre p vers $+\infty$, il vient $\mathbf{E}(X) = \displaystyle\sum_{n=1}^{+\infty} \mathbf{P}(X \geqslant n)$. ■

B – Variance et écart-type

Définition 18.11 : Moment d'ordre p

Soit X une variable aléatoire discrète et $p \in \mathbb{N}^*$. Si X^p admet une espérance finie, alors on appelle moment d'ordre p de X le réel $\mathbf{E}(X^p)$.

Théorème / Définition 18.12 : Variance

Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2. Alors $(X-\mathbf{E}(X))^2$ est d'espérance finie. On appelle variance de X et on note $\mathbf{V}(X)$ le réel positif :

$$\mathbf{V}(X) = \mathbf{E}\big((X-\mathbf{E}(X))^2\big) = \sum_{x\in X(\Omega)} (x-\mathbf{E}(X))^2 \mathbf{P}(X=x)$$

On appelle écart type de X et on note $\sigma(X)$ le réel $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

Démonstration

Posons $X = \{x_n \mid n \in \mathbb{N}\}$ et supposons que X admette un moment d'ordre 2.

- Montrons tout d'abord que X admet un moment d'ordre 1, c'est-à-dire une espérance finie. Pour $n \in \mathbb{N}$, soit $|x_n| < 1$, soit $|x_n| \geqslant 1$ et donc $|x_n| \leqslant x_n^2$. On a donc $|x_n| \leqslant \max(1, x_n^2) \leqslant 1 + x_n^2$. On obtient alors :
$$\forall n \in \mathbb{N} \quad |x_n| \cdot \mathbf{P}(X=x_n) \leqslant \mathbf{P}(X=x_n) + x_n^2 \mathbf{P}(X=x_n)$$
Comme $\sum \mathbf{P}(X=x_n)$ converge (de somme égale à 1) et que par hypothèse, $\sum x_n^2 \cdot \mathbf{P}(X=x_n)$ converge également, par comparaison de séries à termes positifs, $\sum |x_n| \cdot \mathbf{P}(X=x_n)$ converge. On a bien établi la convergence absolue de $\sum x_n \cdot \mathbf{P}(X=x_n)$.
- $(X-\mathbf{E}(X))^2 = X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2$. Comme X^2 et X admettent une espérance finie, il en va de même pour $(X-\mathbf{E}(X))^2$ (par linéarité de l'espérance).
- Enfin, l'écart type est bien défini par positivité de l'espérance, $\mathbf{V}(X) = \mathbf{E}((X-\mathbf{E}(X))^2) \geqslant 0$. ■

La variance, qui n'est rien d'autre que la moyenne du carrée de la distance entre les valeurs de X et la moyenne de X, mesure la dispersion de X par rapport à son espérance. Il en va de même pour l'écart type qui a l'avantage d'être homogène à X.

Lorsque $\mathbf{E}(X) = 0$ et $\sigma(X) = 1$, la variable aléatoire X est qualifiée de *centrée réduite*.

La proposition suivante permet de calculer plus facilement la variance d'une variable aléatoire.

Proposition 18.13 : Formule de Kœnig-Huygens

Si la variable aléatoire discrète X admet un moment d'ordre 2,

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

Démonstration

On a $\mathbf{V}(X) = \mathbf{E}((X-\mathbf{E}(X))^2) = \mathbf{E}(X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2)$. Par linéarité de l'espérance, il vient

$$\mathbf{V}(X) = \mathbf{E}(X^2) - 2\mathbf{E}(X)^2 + \mathbf{E}(X)^2 = \mathbf{E}(X^2) - \mathbf{E}(X)^2$$

Exemple

Reprenons l'exemple du lancer de dés où l'on a noté X la somme des chiffres obtenus. Que vaut l'écart type de X ? X^2 a pour espérance :

$$\mathbf{E}(X^2) = \frac{4}{36} + \frac{9}{18} + \frac{16}{12} + \frac{25}{9} + \frac{180}{36} + \frac{49}{6} + \frac{320}{36} + \frac{81}{9} + \frac{100}{12} + \frac{121}{18} + \frac{144}{36} = \frac{329}{6}$$

D'où $\mathbf{V}(X) = \dfrac{329}{6} - 7^2 = \dfrac{35}{6}$ et donc $\sigma(X) = \sqrt{\dfrac{35}{6}} \approx 2{,}42$.

Exercice 9

Montrer que la variable aléatoire X définie dans l'exercice 7 n'admet pas de variance.

Proposition 18.14

Soient $\lambda, \mu \in \mathbb{R}$. on a alors $\mathbf{V}(\lambda X+\mu)=\lambda^2\mathbf{V}(X)$.

Démonstration

$\mathbf{V}(\lambda X+\mu)=\mathbf{E}((\lambda X+\mu)^2)-\mathbf{E}(\lambda X+\mu)^2=\mathbf{E}(\lambda^2X^2+2\lambda\mu X+\mu^2)-(\lambda\mathbf{E}(X)+\mu)^2=\lambda^2(\mathbf{E}(X^2)-\mathbf{E}(X)^2)=\lambda^2\mathbf{V}(X)$

On remarquera que si $\mathbf{V}(X)\neq 0$, $\mathbf{V}\left(\dfrac{X}{\sigma(X)}\right)=\dfrac{1}{\sigma(X)^2}\mathbf{V}(X)=1$, donc la variable aléatoire $\dfrac{X}{\sigma(X)}$ est réduite.

La variable $\dfrac{1}{\sigma(X)}(X-\mathbf{E}(X))$ est alors centrée réduite.

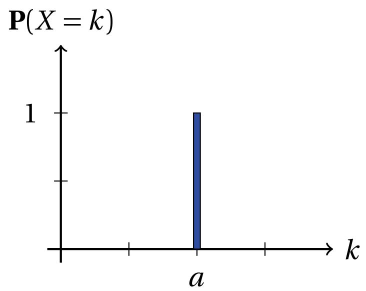
III | Lois usuelles

A – Loi certaine

Définition 18.15 : Loi certaine

On dit qu'une variable aléatoire suit une loi certaine si elle est constante.

On peut alors poser $X(\Omega)=\{a\}$. L'événement $(X=a)$ est certain, on a $\mathbf{P}(X=a)=1$.



LOI DE PROBABILITÉ

$\mathbf{E}(X)=a\cdot\mathbf{P}(X=a)=a$, $\mathbf{V}(X)=\mathbf{E}(X^2)-\mathbf{E}(X)^2=a^2-a^2=0$.

Réciproquement, une variable aléatoire de variance nulle est dite quasi-certaine : elle peut prendre d'autres valeurs que a mais avec une probabilité nulle.

B – Loi uniforme

Définition 18.16 : Loi uniforme

On dit que la variable aléatoire X suit une loi uniforme sur $[\![1,n]\!]$ si :

$$X(\Omega)=[\![1,n]\!] \quad \text{et} \quad \forall k\in[\![1,n]\!],\quad \mathbf{P}(X=k)=\frac{1}{n}$$

Notation : $X\hookrightarrow\mathscr{U}([\![1,n]\!])$.

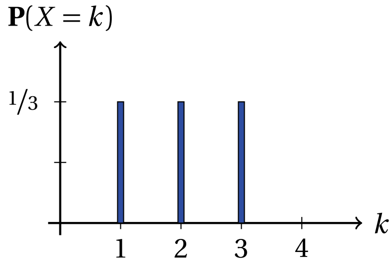
De même, on dira que X suit une loi uniforme sur $[\![a,b]\!]$ avec $a<b$ si :

$$X(\Omega)=[\![a,b]\!] \quad \text{et} \quad \forall k\in[\![a,b]\!],\quad \mathbf{P}(X=k)=\frac{1}{b-a+1}$$

Situation type :

- Lancer d'un dé équilibré ;

- Tirage aléatoire de boules numérotées de 1 à n dans une urne.



LOI DE PROBABILITÉ POUR $n = 3$

Proposition 18.17 : Espérance et variance d'une loi uniforme

Si $X \hookrightarrow \mathscr{U}(\llbracket 1, n \rrbracket)$ alors $\mathbf{E}(X) = \dfrac{n+1}{2}$ et $\mathbf{V}(X) = \dfrac{n^2-1}{12}$.

Démonstration

Rappelons tout d'abord que $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ et $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

$$\mathbf{E}(X) = \sum_{k=1}^{n} k\mathbf{P}(X=k)) = \sum_{k=1}^{n} \frac{k}{n} = \frac{n+1}{2}\ ;$$

$$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \sum_{k=1}^{n} k^2\mathbf{P}(X=k)) - \left(\frac{n+1}{2}\right)^2 = \sum_{k=1}^{n} \frac{k^2}{n} - \left(\frac{n+1}{2}\right)^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$$

Exercice 10

Montrer que si $X \hookrightarrow \mathscr{U}(\llbracket a, b \rrbracket)$ alors on a :

$$\mathbf{E}(X) = \frac{a+b}{2} \quad \text{et} \quad \mathbf{V}(X) = \frac{(b-a)(b-a+2)}{12}$$

On s'intéressera pour cela à la loi suivie par la variable aléatoire $Y = X - a + 1$.

C – Loi de Bernoulli

Définition 18.18 : Loi de Bernoulli

On dit que la variable aléatoire X suit une loi de Bernoulli de paramètre $p \in]0, 1[$ si :

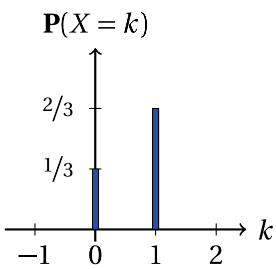
$$X(\Omega) = \{0, 1\} \quad \text{et} \quad \mathbf{P}(X = 1) = p$$

Notation : $X \hookrightarrow \mathscr{B}(p)$ ou $X \hookrightarrow \mathscr{B}(1, p)$.

On notera qu'on a alors $\mathbf{P}(X = 0) = q = 1 - p$.

Situation type :

- Lancer d'une pièce truquée (lancer unique) ;
- Tirage d'une boule dans une urne avec a boules noires et b boules blanches.
  Si on note X le nombre de boules noires tirées, $X \hookrightarrow \mathscr{B}\left(\frac{a}{a+b}\right)$.



LOI DE PROBABILITÉ POUR $p = 2/3$

Proposition 18.19 : Espérance et variance d'une loi de Bernoulli

Si $X \hookrightarrow \mathscr{B}(p)$ alors $\mathbf{E}(X) = p$ et $\mathbf{V}(X) = p(1-p) = pq$.

Démonstration

$\mathbf{E}(X) = 0 \cdot \mathbf{P}(X=0) + 1 \cdot \mathbf{P}(X=1) = p$;
$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \left(0^2 \cdot \mathbf{P}(X=0) + 1^2 \cdot \mathbf{P}(X=1)\right) - p^2 = p - p^2 = pq$. ■

D – Loi binomiale

Définition 18.20 : Loi binomiale

On dit que la variable aléatoire X suit une loi binomiale de paramètres $(n,p) \in \mathbb{N}^* \times]0,1[$ si :

$$\left| \begin{array}{l} X(\Omega) = [\![0,n]\!] \\ \forall k \in [\![0,n]\!],\ \mathbf{P}(X=k) = \binom{n}{k} p^k q^{n-k} \text{ avec } q = 1-p \end{array} \right.$$

Notation : $X \hookrightarrow \mathscr{B}(n,p)$.

Vérifions que nous avons bien défini une variable aléaoire :

- $\forall k \in [\![0,n]\!],\ \mathbf{P}(X=k) = \binom{n}{k} p^k q^{n-k} \geqslant 0$
- $\displaystyle\sum_{k=0}^{n} \mathbf{P}(X=k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1^n = 1$

Pour $n = 1$, on retrouve le cas particulier d'une loi de Bernoulli.

Une loi binomiale de paramètres (n,p) apparaît lorsqu'on s'intéresse à une expérience aléatoire conduisant à une situation de succès (de probabilité p) ou d'échec (de probabilité q), le tout réitéré n fois de manière *indépendante*. En notant X le nombre de succès, on a alors $X \hookrightarrow \mathscr{B}(n,p)$.

Démonstration

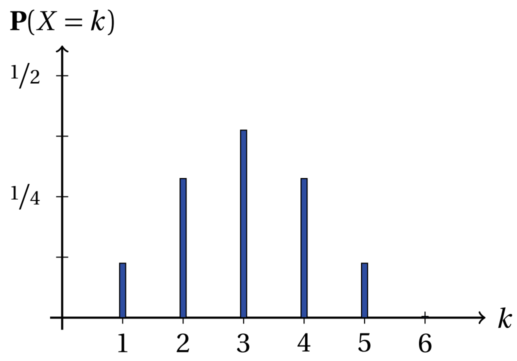
Posons $\Omega = \{S,E\}^n$ et notons que $\mathrm{card}(\Omega) = 2^n$.

- On a $X(\Omega) = [\![0,n]\!]$ (au pire 0 succès, au mieux n).
- Les tirages favorables peuvent être représentés par des mots de la forme $SES \cdots EESEE$, mots contenant k lettres S et donc $n-k$ lettres E. Il y en a précisément $\binom{n}{k}$.
- La probabilité d'apparition de chacun des mots vaut $p^k q^{n-k}$ par indépendance des tirages.
  Donc $\mathbf{P}(X=k) = \binom{n}{k} p^k q^{n-k}$. ■

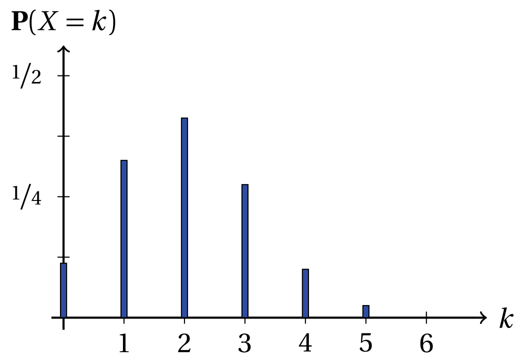
Une loi binomiale de paramètres (n,p) est la somme de n loi de Bernoulli indépendantes, de paramètre p comme nous pourrons le constater dans la prochaine partie.

Situation type :

- Lancers successifs d'une pièce truquée;
- Tirages successifs *avec remise* de n boules dans une urne avec a boules noires et b boules blanches. Si X correspond au nombre de boules blanches tirées, on a $X \hookrightarrow \mathscr{B}\left(n, \frac{a}{a+b}\right)$.



Représentation d'une loi binomiale de paramètres $\left(6, \frac{1}{2}\right)$



Représentation d'une loi binomiale de paramètres $\left(6, \frac{1}{3}\right)$

Proposition 18.21 : Espérance et variance d'une loi binomiale

Si $X \hookrightarrow \mathscr{B}(n,p)$ alors $\mathbf{E}(X) = np$ et $\mathbf{V}(X) = npq$.

Démonstration

Remarquons tout d'abord que $k\binom{n}{k} = n\binom{n-1}{k-1}$ et $k(k-1)\binom{n}{k} = n(n-1)\binom{n-2}{k-2}$.

- Calcul classique pour l'espérance :

$$\begin{aligned}\mathbf{E}(X) &= \sum_{k=0}^{n} k\binom{n}{k}p^k q^{n-k} = \sum_{k=1}^{n} k\binom{n}{k}p^k q^{n-k} = n\sum_{k=1}^{n}\binom{n-1}{k-1}p^k q^{n-k}\\ &= n\sum_{k=0}^{n-1}\binom{n-1}{k}p^{k+1}q^{n-1-k} = np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k q^{n-1-k} = np(p+q)^{n-1} = np\end{aligned}$$

- En ce qui concerne la variance, nous allons calculer $\mathbf{E}(X(X-1))$ plutôt que $\mathbf{E}(X^2)$, en utilisant la formule de transfert :

$$\begin{aligned}\mathbf{E}(X(X-1)) &= \sum_{k=0}^{n} k(k-1)\binom{n}{k}p^k q^{n-k} = \sum_{k=2}^{n} k(k-1)\binom{n}{k}p^k q^{n-k} = n(n-1)\sum_{k=2}^{n}\binom{n-2}{k-2}p^k q^{n-k}\\ &= n(n-1)p^2\sum_{k=0}^{n-2}\binom{n-2}{k}p^k q^{n-2-k} = n(n-1)p^2(p+q)^{n-2} = n(n-1)p^2\end{aligned}$$

On a alors $\mathbf{E}(X(X-1)) = \mathbf{E}(X^2) - \mathbf{E}(X)$ par linéarité, et donc :

$$V(X) = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}(X)^2 = n(n-1)p^2 + np - (np)^2 = npq$$

■

E – Loi géométrique

Définition 18.22 : Loi géométrique

Soit $p \in]0,1[$. On dit que la variable aléatoire X suit une loi géométrique de paramètre p sur $\mathbb{N}^*$ si :

$$\left|\begin{array}{l} X(\Omega) = \mathbb{N}^* \\ \forall n \in \mathbb{N}^*,\ \mathbf{P}(X = n) = q^{n-1}p \text{ avec } q = 1-p \end{array}\right.$$

Notation : $X \hookrightarrow \mathscr{G}(p)$.

Vérifions que l'on a bien défini une variable aléatoire :

$$\forall n \in \mathbb{N}^* \quad \mathbf{P}(X = n) = q^{n-1}p \geqslant 0 \quad \text{et} \quad \sum_{n=1}^{+\infty}\mathbf{P}(X = n) = p\sum_{n=1}^{+\infty}q^{n-1} = p\sum_{n=0}^{+\infty}q^n = \frac{p}{1-q} = 1$$

Une loi géométrique apparaît lorsque l'on s'intéresse à une succession d'épreuves de Bernoulli indépendantes jusqu'à obtenir un premier succès. On parle alors de temps d'attente du premier succès.

Démonstration

Considérons une série infinie de lancers de pièce qui amène pile avec une probabilité $p \in]0,1[$ et face avec une probabilité $1-p$. Notons A_k l'événement « pile apparaît au k^{e} lancer ». Par indépendance des événements,

$$\forall k \geqslant 2 \quad \mathbf{P}(X=k) = \mathbf{P}\left(\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_{k-1}} \cap A_k\right) = \mathbf{P}(\overline{A_1}) \times \cdots \times \mathbf{P}(\overline{A_{k-1}}) \times \mathbf{P}(A_k) = q^{k-1} p$$

De plus, on a bien $\mathbf{P}(X=1) = p$. ■

Remarquons que dans l'exemple précédent, rien ne garantit l'existence d'une telle variable aléatoire X ! En effet, si on désigne par X le rang d'apparition du premier pile, rien ne nous assure que pile va apparaître au cours des tirages. Cependant, un tel événement est de probabilité nulle. En notant B l'événement « pile n'apparaît jamais » et F_n l'événement « face apparaît constamment au cours des n premiers lancers », on a :

$$B = \bigcap_{n=1}^{+\infty} F_n$$

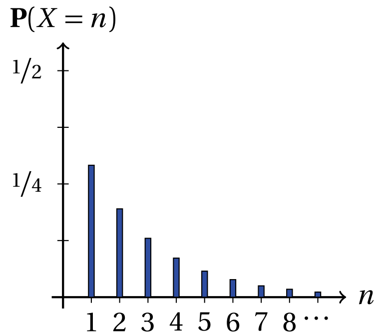
Comme la suite $(F_n)_{n\in\mathbb{N}^*}$ est décoissante, on a d'après le théorème de continuité décroissante :

$$\mathbf{P}(B) = \lim_{n\to+\infty} \mathbf{P}(A_n) = \lim_{n\to+\infty} q^n = 0$$

On peut alors négliger cet événement et considérer X à valeurs dans $\mathbb{N}^*$.

Situation type :

- Premier pile obtenu au cours de lancers successifs d'une pièce truquée;
- Tirages successifs *avec remise* de n boules dans une urne avec a boules noires et b boules blanches jusqu'à obtenir une boule blanche.



REPRÉSENTATION D'UNE LOI GÉOMÉTRIQUE DE PARAMÈTRE $1/3$

Proposition 18.23 : Espérance et variance d'une loi géométrique

Si $X \hookrightarrow \mathscr{G}(p)$ alors X admet une espérance et une variance et :

$$\mathbf{E}(X) = \frac{1}{p} \qquad \text{et} \qquad \mathbf{V}(X) = \frac{q}{p^2} = \frac{1-p}{p^2}$$

Démonstration

- La série $\sum\limits_{n\in\mathbb{N}^*} npq^{n-1}$ converge d'après le théorème de dérivation terme à terme d'une série entière appliqué à la série $\sum q^n$ avec $(q \in]-1,1[)$. De plus,

$$\mathbf{E}(X) = \sum_{n=1}^{+\infty} n\mathbf{P}(X=n) = \sum_{n=1}^{+\infty} npq^{n-1} = \frac{p}{(1-q)^2} = \frac{1}{p}$$

- Pour la même raison, $\sum_{n\geqslant 1} n(n-1)pq^{n-1}$ converge donc la variance existe et :

$$\mathbf{V}(X)=\mathbf{E}(X(X-1))+\mathbf{E}(X)-\mathbf{E}(X)^2=\sum_{n=1}^{+\infty} n(n-1)pq^{n-1}+\frac{1}{p}-\frac{1}{p^2}=\frac{2pq}{(1-q)^3}+\frac{1}{p}-\frac{1}{p^2}=\frac{1-p}{p^2}$$

Nous avons jusqu'à présent interprété une loi géométrique comme le temps d'attente du 1er succès dans une suite d'épreuves de Bernoulli indépendantes. Si un observateur commence son observation au temps n et qu'il constate que l'événement ne s'est pas encore produit, quelle est la probabilité que l'événement se produise au temps $n+k$? Elle est en fait égale à la probabilité que l'événement se produise au temps k, comme si rien ne s'était produit jusque là. La loi géométrique est qualifiée de « loi sans mémoire ».

Proposition 18.24 : Caractérisation comme loi sans mémoire

Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ telle que $X(\Omega)=\mathbb{N}^*$.
X suit une loi géométrique sur $\mathbb{N}^*$ si et seulement si :

$$\forall (k,n)\in\mathbb{N}^2 \quad \mathbf{P}(X>n+k|X>n)=\mathbf{P}(X>k)$$

Démonstration

Remarquons tout d'abord que si $X\hookrightarrow\mathscr{G}(p)$, $\mathbf{P}(X>n)=1-\sum_{k=1}^{n}\mathbf{P}(X=k)=1-\sum_{k=1}^{n}pq^{k-1}=q^n$.
Ceci s'interprète facilement en disant que pour l'événement se réalise à un temps supérieur à n, cela nécessite n premiers échecs consécutifs.

$\boxed{\Longrightarrow}$ Supposons que X suive une loi géométrique de paramètre $p\in\mathbb{N}^*$.
Pour tout $(k,n)\in\mathbb{N}$, $\mathbf{P}(X>n+k|X>n)=\dfrac{\mathbf{P}(X>n+k)}{\mathbf{P}(X>n)}=\dfrac{q^{n+k}}{q^n}=q^k=\mathbf{P}(X>k)$.

$\boxed{\Longleftarrow}$ Supposons maintenant que pour tout $(k,n)\in\mathbb{N}$, $\mathbf{P}(X>n+k|X>n)=\mathbf{P}(X>k)$.
Ainsi, $\mathbf{P}(X>n+1)=\mathbf{P}(X>n+1|X>n)\mathbf{P}(X>n)=\mathbf{P}(X>1)\mathbf{P}(X>n)$. En posant $q=\mathbf{P}(X>1)$, on voit que $(\mathbf{P}(X>n))_{n\in\mathbb{N}^*}$ est une suite géométrique de raison q.

$$\forall n\in\mathbb{N}^* \quad \mathbf{P}(X>n)=q\mathbf{P}(X>n-1)=\cdots=q^{n-1}\mathbf{P}(X>1)=q^n$$

Ainsi, $\mathbf{P}(X=n)=\mathbf{P}(X\leqslant n)-\mathbf{P}(X\leqslant n-1)=\mathbf{P}(X>n-1)-\mathbf{P}(X>n)=q^{n-1}-q^n=pq^{n-1}$.
Donc $X\hookrightarrow\mathscr{G}(p)$. ■

F – Loi de Poisson

Contrairement aux lois précédentes, il s'avère difficile de donner un modèle simple qui conduise à une loi de Poisson. Celle-ci apparaît plus naturellement comme une loi limite. Cette loi a été introduite en 1838 par Siméon Poisson et sert souvent à modéliser l'apparition d'événements rares.

Définition 18.25 : Loi de Poisson

Soit $\lambda>0$. On dit que la variable aléatoire X suit une loi de Poisson de paramètre λ si :

$$\left|\begin{array}{l} X(\Omega)=\mathbb{N} \\ \forall n\in\mathbb{N},\ \mathbf{P}(X=n)=\dfrac{\lambda^n}{n!}\mathrm{e}^{-\lambda}\end{array}\right.$$

Notation : $X\hookrightarrow\mathscr{P}(\lambda)$.

Vérifions que l'on a bien défini une variable aléatoire :

$$\forall n\in\mathbb{N},\ \mathbf{P}(X=n)=\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}\geqslant 0 \quad \text{et} \quad \sum_{n=0}^{+\infty}\mathbf{P}(X=n)=\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}=\mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!}=\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda}=1$$

Proposition 18.26 : Espérance et variance d'une loi de Poisson

Si $X \hookrightarrow \mathscr{P}(\lambda)$ alors X admet une espérance et une variance et :

$$\mathbf{E}(X)=\lambda \quad \text{et} \quad \mathbf{V}(X)=\lambda$$

Démonstration

- La série de terme général $n\mathbf{P}(X=n)=\lambda e^{-\lambda}\dfrac{\lambda^{n-1}}{(n-1)!}$ converge comme série exponentielle.
  Ainsi, X admet une espérance et :

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

- Pour la même raison, le série de terme général $\lambda^2 e^{-\lambda}\dfrac{\lambda^{n-2}}{(n-2)!}$ converge donc la variance existe et :

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

IV | Vecteurs aléatoires discrets

A – Couples de variables aléatoires réelles discrètes

1 – Définition

Définition 18.27 : Couple de variables aléatoires

Soit $(\Omega,\mathscr{A})$ un espace probabilisable. On appelle couple de variables aléatoires discrètes toute application

$$Z:\left|\begin{array}{ccc}\Omega & \longrightarrow & \mathbb{R}^2\\ \omega & \longmapsto & (X(\omega),Y(\omega))\end{array}\right.$$

où X et Y sont des variables aléatoires discrètes sur $(\Omega,\mathscr{A})$. On note $Z=(X,Y)$ ce couple de variables.

Z est donc une application de Ω dans $\mathbb{R}^2$. Avant de montrer qu'il s'agit bien d'une variable aléatoire [1], insistons sur la différence entre $Z(\Omega)=(X,Y)(\Omega)$ et $X(\Omega)\times Y(\Omega)$.

$$Z(\Omega)=\{(X(\omega),Y(\omega))\mid \omega\in\Omega\}\subset\{(X(\omega_1),Y(\omega_2))\mid(\omega_1,\omega_2)\in\Omega^2\}=X(\Omega)\times Y(\Omega)$$

Exercice 11

(i) Soient $X(\Omega)=\{1,2\}$, $Y(\Omega)=\{1,3\}$ et $\Omega=\{a,b,c\}$ avec $X(a)=X(b)=1$, $X(c)=2$ et $Y(a)=Y(c)=3$, $Y(b)=1$. Déterminer $Z(\Omega)$ et $X(\Omega)\times Y(\Omega)$.

(ii) On lance deux dés. On note X le maximum des valeurs obtenues, Y la valeur du deuxième dé. Que vaut $X(\Omega)$? $Y(\Omega)$? $Z(\Omega)$?

Proposition 18.28

Soient $(\Omega,\mathscr{A})$ un espace probabilisable et X,Y deux variables aléatoires réelles discrètes.
Alors $Z=(X,Y)$ est une variable aléatoire discrète sur $\mathbb{R}^2$.

On parle alors de vecteur aléatoire discret.

1. au sens plus général des variables aléatoires discrètes à valeurs dans $\mathbb{R}^n$. Ici, $n=2$.

Démonstration

- $X(\Omega)$ et $Y(\Omega)$ sont au plus dénombrables. Il en va de même pour $X(\Omega)\times Y(\Omega)$ et donc pour $Z(\Omega)$ qui est inclus dans ce dernier.
- Soit $z=(x,y)\in Z(\Omega)$.

$$(Z=z)=\{\omega\in\Omega \mid X(\omega)=x,\ Y(\omega)=y\}=(X=x)\cap(Y=y)$$

$(Z=z)$ est donc l'intersection de deux événements, c'est un événement ! ■

On pourra utiliser les notations suivantes :

$$(X=x_i)\cap(Y=y_j)=\big((X,Y)=(x_i,y_j)\big)=\big(X=x_i,\ Y=y_j\big)$$

Proposition 18.29

Soit (X,Y) un couple de variables aléatoires réelles discrètes sur $(\Omega,\mathscr{A})$. On pose $X(\Omega)=\{x_i \mid i\in I\}$ et $Y(\Omega)=\{y_j \mid j\in J\}$ avec I et J deux parties de $\mathbb{N}$.
Alors la famille dévénements $\big((Z=(x_i,y_j))\big)_{\substack{i\in I\\ j\in J}}$ est un système complet d'événements.

2 – Loi conjointe

Définition 18.30

Soit (X,Y) un couple de variables aléatoires sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$. On appelle loi conjointe de X et de Y (ou loi de probabilité du couple (X,Y)) l'application :

$$\mathbf{P}_{(X,Y)}:\left|\begin{array}{ccc} X(\Omega)\times Y(\Omega) & \longrightarrow & [0,1]\\ (x,y) & \longmapsto & \mathbf{P}\big((X=x)\cap(Y=y)\big)\end{array}\right.$$

Déterminer la loi conjointe revient à calculer $p_{i,j}=\mathbf{P}((X=x_i)\cap(Y=y_j))$ pour $i\in I$ et $j\in J$.

Exemple

On considère une urne avec 4 boules B_1, B_2, B_3, B_4 et les tirages de 2 boules successifs avec remise. On note X_1 le numéro de la boule au 1[er] tirage et X_2 le numéro de la boule au 2[ème] tirage. On pose enfin $Y=\max(X_1,X_2)$ et $Z=(X_1,Y)$. Quelle est la loi conjointe de X_1 et Y ?
On a $X_1(\Omega)=Y(\Omega)=[\![1,4]\!]$ et :

X_1/Y	1	2	3	4
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
2	0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$
4	0	0	0	$\frac{1}{4}$

$\displaystyle\sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}\mathbf{P}\big(X=x_i,\ Y=y_j\big)=1$ car les $\big(X=x_i,\ Y=y_j\big)$ forment un système complet d'événements.

Exercice 12

Dans une succession de piles ou faces pour laquelle la probabilité d'obtenir pile est $p\in]0,1[$ et d'obtenir face est $q=1-p$, on note X le rang d'apparition du premier pile et Y celui du second. Déterminer la loi conjointe de (X,Y) puis vérifier que la somme des probabilités vaut 1.

3 – Loi marginale

Définition 18.31

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathscr{A}, \mathbf{P})$.
On appelle $1^{\text{ère}}$ loi marginale de (X, Y) la loi de X et $2^{\text{ème}}$ loi marginale de (X, Y) la loi de Y.

On peut facilement déduire les lois marginales de la loi conjointe :

Théorème 18.32

Avec les notations précédentes,

$$\left| \begin{array}{l} \forall i \in I \quad \mathbf{P}(X = x_i) = \sum_{j \in J} \mathbf{P}\big((X = x_i) \cap (Y = y_j)\big) \\ \forall j \in J \quad \mathbf{P}(Y = y_j) = \sum_{i \in I} \mathbf{P}\big((X = x_i) \cap (Y = y_j)\big) \end{array} \right.$$

Il suffit donc de sommer les éléments le long des lignes ou le long des colonnes des tableaux de probabilités pour retrouver les lois marginales.

Exemple

X_1/Y	1	2	3	4
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
2	0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$
4	0	0	0	$\frac{1}{4}$

$$\forall k \in \{1, \dots, 4\} \quad \mathbf{P}(X_1 = k) = \frac{1}{4}$$

$$\mathbf{P}(Y = 1) = \frac{1}{16}; \quad \mathbf{P}(Y = 2) = \frac{3}{16};$$
$$\mathbf{P}(Y = 3) = \frac{5}{16}; \quad \mathbf{P}(Y = 4) = \frac{7}{16}.$$

Attention, dans la plupart des cas, on ne peut pas retrouver la loi conjointe à partir des lois marginales.

Exercice 13

Déterminer les lois marginales de (X, Y) définie dans l'exercice 12.

4 – Fonctions de deux variables aléatoires (complément)

Nous avons que pour toutes variable aléatoire X et fonction $f : X(\Omega) \to \mathbb{R}$, $f(X)$ est encore une variable aléatoire. C'est encore le cas pour $Z = g(X, Y)$ avec $g : X(\Omega) \times Y(\Omega) \to \mathbb{R}$. En effet,

$$(Z = z) = \bigsqcup_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y) = z}} \big((X = x) \cap (Y = y)\big)$$

est une réunion dénombrable d'événements. De plus, $Z(\Omega)$ est au plus dénombrable.

Bref, Z est bien une variable aléatoire.

Mais comme la réunion précédente est disjointe, on connaît la loi de Z :

$$\mathbf{P}(Z = z) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y) = z}} \mathbf{P}\big((X = x) \cap (Y = y)\big)$$

Il est cependant très difficile en pratique de déterminer la loi à l'aide de cette simple formule...

B – Indépendance et lois conditionnelles

Définition 18.33

Soit (X, Y) un couple de variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$.

(i) On appelle loi conditionnelle de X sachant $(Y = y)$ pour $\mathbf{P}(Y = y) \neq 0$, l'application :

$$\mathbf{P}_{(Y=y)}((X=\bullet)) : \left| \begin{array}{rcl} X(\Omega) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{\mathbf{P}\big((X=x)\cap(Y=y)\big)}{\mathbf{P}\big(Y=y\big)} = \mathbf{P}_{(Y=y)}(X=x) \end{array} \right.$$

(ii) On appelle loi conditionnelle de Y sachant $(X = x)$ pour $\mathbf{P}(X = x) \neq 0$, l'application :

$$\mathbf{P}_{(X=x)}((Y=\bullet)) : \left| \begin{array}{rcl} Y(\Omega) & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \dfrac{\mathbf{P}\big((X=x)\cap(Y=y)\big)}{\mathbf{P}(X=x)} = \mathbf{P}_{(X=x)}(Y=y) \end{array} \right.$$

Définition 18.34 : Indépendance de deux variables aléatoires

Deux variables aléatoires réelles X et Y sur $(\Omega, \mathscr{A}, \mathbf{P})$ sont dites indépendantes si :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega) \quad \mathbf{P}\big((X=x)\cap(Y=y)\big) = \mathbf{P}(X=x) \times \mathbf{P}\big(Y=y\big)$$

Exercice 14

Comment observer l'indépendance de X et de Y dans une table de probabilités ?

Dans le cas d'indépendance, on peut retrouver la loi conjointe à partir des lois marginales.

Proposition 18.35

Si X et Y sont indépendantes et si $f : X(\Omega) \to \mathbb{R}$, $g : Y(\Omega) \to \mathbb{R}$, alors $f(X)$ et $g(Y)$ sont indépendantes.

Ce résultat est admis.

Exemple

Si X et Y sont indépendantes, alors X^2 et Y^2 le sont également.

Revenons maintenant sur le cas des fonctions de deux variables aléatoires et reprenons la notation précédemment introduite $Z = g(X, Y)$, où X et Y sont deux variables aléatoires discrètes et $g : X(\Omega) \times Y(\Omega) \to \mathbb{R}$. Supposons de plus qu'elles sont indépendantes. On peut alors écrire :

$$\mathbf{P}(Z=z) = \sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega) \\ g(x,y)=z}} \mathbf{P}\big((X=x)\cap(Y=y)\big) = \sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega) \\ g(x,y)=z}} \mathbf{P}(X=x)\cdot\mathbf{P}(Y=y)$$

Exemple

Une urne contient deux boules B et $n \in \mathbb{N}^*$ boules N. On tire toutes les boules une par une.
On note X le rang du tirage de la 1e boule blanche tirée, Y celui de la 2ème.
Quelle est la loi du couple (X, Y) ? la loi de X ? la loi de Y ? la loi de $Y - X$?

- Ω est l'ensemble des arrangements des $n+2$ boules parmi $n+2$.
- $X(\Omega) = [\![1, n+1]\!]$; $Y(\Omega) = [\![2, n+2]\!]$.
- $(X = i) \cap (Y = j) = \varnothing$ pour $i \geqslant j$, et pour $i < j$, on a :

$$\mathbf{P}((X=i)\cap(Y=j)) = \frac{2}{(n+1)(n+2)}$$

- On peut alors obtenir les lois de X et Y :

$$\mathbf{P}(X=i)=\sum_{j=2}^{n+2}\mathbf{P}((X=i)\cap(Y=j))=\sum_{j=i+1}^{n+2}\mathbf{P}((X=i)\cap(Y=)])=\frac{2}{(n+1)(n+2)}(n+2-i)$$

$$\mathbf{P}\big(Y=j\big)=\sum_{i=1}^{n+1}\mathbf{P}((X=i)\cap(Y=j))=\sum_{i=1}^{j-1}\mathbf{P}((X=i)\cap(Y=j))=\frac{2}{(n+1)(n+2)}(j-1)$$

- $(Y-X)(\Omega)=[\![1,n+1]\!]$; les $(X=i)$ formant un système complet d'événements, on trouve :

$$\forall k\in[\![1,n+1]\!]\quad \mathbf{P}(Y-X=k)=\frac{2}{(n+1)(n+2)}(n+2-k)$$

Exemple

Dans le cas d'une somme $X+Y$ avec X et Y prenant des valeurs entières et indépendantes,

$$\forall k\in(X+Y)(\Omega)\quad \mathbf{P}(X+Y=k)=\sum_{\substack{(i,j)\in\mathbb{N}^2\\ i+j=k}}\mathbf{P}(X=i)\mathbf{P}(Y=j)=\sum_{i=0}^{k}\mathbf{P}(X=i)\mathbf{P}(Y=k-i)$$

C – Espérance, variance et covariance

1 – Espérance d'une variable aléatoire fonction de deux variables aléatoires

Le théorème de transfert se généralise aux variables aléatoires fonction de deux variables aléatoires.
Sous certaines conditions de convergence que nous ne préciserons pas ici, si on pose $Z=g(X,Y)$ et $X(\Omega)=\{x_i \mid i\in I\}$ et $Y(\Omega)=\{y_j \mid j\in J\}$ avec I et J deux parties de $\mathbb{N}$, alors :

$$\mathbf{E}(Z)=\sum_{\substack{i\in I\\ j\in J}}g(x_i,y_j)\mathbf{P}\big((X=x_i)\cap(Y=y_j)\big)$$

Exemple

On considère une urne contenant n boules numérotées de 1 à n. On tire successivement deux boules avec remise. On note X le numéro de la $1^{\text{ère}}$ boule, Y le numéro de la 2^{nde}.
Soit $Z=\max(X,Y)$. Quelle est l'espérance de Z ?
$X(\Omega)=Y(\Omega)=Z(\Omega)=[\![1,n]\!]$; $X,Y\hookrightarrow\mathscr{U}([\![1,n]\!])$ et sont indépendantes. De plus,

$$\begin{aligned}
\mathbf{E}(Z)&=\sum_{1\leqslant i,j\leqslant n}\max(i,j)\mathbf{P}\big((X=x_i)\cap(Y=y_j)\big)=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\max(i,j)\\
&=\frac{1}{n^2}\sum_{i=1}^{n}\left[\sum_{j=1}^{i}\max(i,j)+\sum_{j=i+1}^{n}\max(i,j)\right]=\frac{1}{n^2}\sum_{i=1}^{n}\left[\sum_{j=1}^{i}i+\sum_{j=i+1}^{n}j\right]\\
&=\frac{1}{n^2}\sum_{i=1}^{n}\left[i^2+\frac{n(n+1)}{2}-\frac{i(i+1)}{2}\right]=\frac{1}{2n^2}\sum_{i=1}^{n}\left[n(n+1)+i^2-i\right]\\
&=\frac{1}{2n^2}\left(n^2(n+1)+\frac{n(n+1)(2n+1)}{6}-\frac{n(n+1)}{2}\right)=\frac{(n+1)(4n-1)}{6n}
\end{aligned}$$

Que se passe-t-il pour un tirage sans remise ?

On peut déduire la linéarité de l'espérance du résultat précédent.

Exercice 15

Que vaut en moyenne la somme de deux nombres pris au hasard entre 1 et n ?

2 – Covariance

Théorème / Définition 18.36

Soit (X,Y) un couple de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$. Si X et Y admettent un moment d'ordre 2, alors la variable aléatoire $(X-\mathbf{E}(X))(Y-\mathbf{E}(Y))$ admet une espérance. On appelle alors covariance de X et Y le réel noté $\text{cov}(X,Y)$ défini par :

$$\text{cov}(X,Y)=\mathbf{E}((X-\mathbf{E}(X))(Y-\mathbf{E}(Y))$$

La covariance de X et Y est l'espérance du produit des variables centrées.

Théorème 18.37 : Formule de Kœnig-Huygens

Sous les hypothèses précédentes, $\text{cov}(X,Y)=\mathbf{E}(XY)-\mathbf{E}(X)\cdot\mathbf{E}(Y)$.

Démonstration

Il s'agit d'un simple calcul.

$$\begin{aligned}\text{cov}(X,Y)&=\mathbf{E}((X-\mathbf{E}(X))(Y-\mathbf{E}(Y))=\mathbf{E}(XY-\mathbf{E}(X)Y-\mathbf{E}(Y)X+\mathbf{E}(X)\mathbf{E}(Y))\\&=\mathbf{E}(XY)-2\mathbf{E}(X)\mathbf{E}(Y)+\mathbf{E}(X)\mathbf{E}(Y)=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y)\end{aligned}$$

■

On a $\text{cov}(X,X)=\mathbf{V}(X)$ et $\mathbf{E}(XY)=\sum_{\substack{i\in I\\ j\in J}} x_i y_j \mathbf{P}\big((X=x_i)\cap(Y=y_j)\big)$.

Exercice 16

Exemple de l'urne avec 4 boules, $Z=(X_1,Y)$ avec $Y=\max(X_1,X_2)$. Que vaut $\text{cov}(X_1,Y)$?

Proposition 18.38 : Propriétés de la covariance

Si X, Y et Z sont trois variables aléatoires admettant un moment d'ordre 2 et $a,b,c,d\in\mathbb{R}$, alors :

(i) $\text{cov}(X,X)=\mathbf{V}(X)$;

(ii) $\text{cov}(X,Y)=\text{cov}(Y,X)$;

(iii) $\text{cov}(aX+bY,Z)=a\,\text{cov}(X,Z)+b\,\text{cov}(Y,Z)$;

(iv) $\text{cov}(aX+b,cY+d)=ac\,\text{cov}(X,Y)$.

Proposition 18.39

Si X et Y sont deux variables alatoires admettant un moment d'ordre 2 et $a,b\in\mathbb{R}$, alors :

$$\mathbf{V}(aX+bY)=a^2\mathbf{V}(X)+b^2\mathbf{V}(Y)+2ab\,\text{cov}(X,Y)$$

Pour $a=b=1$, $\mathbf{V}(X+Y)=\mathbf{V}(X)+\mathbf{V}(Y)+2\text{cov}(X,Y)$.

Démonstration

On a tout simplement :

$$\begin{aligned}\mathbf{V}(aX+bY)&=\text{cov}(aX+bY,aX+bY)=a\,\text{cov}(X,aX+bY)+b\,\text{cov}(Y,aX+bY)\\&=a^2\text{cov}(X,X)+b^2\text{cov}(Y,Y)+2ab\,\text{cov}(X,Y)=a^2\mathbf{V}(X)+b^2\mathbf{V}(Y)+2ab\,\text{cov}(X,Y)\end{aligned}$$

Théorème 18.40

Si X et Y sont deux variables aléatoires admettant un moment d'ordre 2 et indépendantes, alors :

(i) $\mathbf{E}(XY)=\mathbf{E}(X)\mathbf{E}(Y)$;

(ii) $\text{cov}(X,Y)=0$;

(iii) $\mathbf{V}(X+Y)=\mathbf{V}(X)+\mathbf{V}(Y)$.

La démonstration s'appuie encore une fois sur le théorème de transfert et la sommation par tranche...
La réciproque est fausse !

Exemple

Considérons la variable aléatoire réelle X sur $(\Omega, \mathscr{A}, \mathbf{P})$ telle que :

x_i	-2	-1	0	1	2
$\mathbf{P}(X = x_i)$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$

et posons $Y = X^2$. On vérifie que $\operatorname{cov}(X, Y) = 0$ bien que X et X^2 ne sont pas indépendantes.

3 – Coefficient de corrélation linéaire

Rappelons que $\sigma(X) = \sqrt{\mathbf{V}(X)}$ avec $\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.
De plus, $\sigma(X) = 0$ si et seulement si la variable X suit une loi certaine.

Définition 18.41

Soient X et Y deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et admettant un moment d'ordre 2. On suppose de plus que leur écart type est non nul.
On appelle coefficient de corrélation linéaire de X et Y le réel $\rho(X, Y) = \dfrac{\operatorname{cov}(X, Y)}{\sigma(X) \cdot \sigma(Y)}$.

On notera que $\rho(X, Y)$ est une grandeur sans unité.

Proposition 18.42

Si X et Y sont deux v.a.r. d'écart-types non nuls, alors :

(i) $\rho(X, Y) \in [-1, 1]$;

(ii) $\rho(X, Y) = \pm 1$ si et seulement s'il existe $a, b \in \mathbb{R}$ tels que $Y = aX + b$.

Démonstration

$\mathbf{V}(\lambda X + Y) = \lambda^2 \mathbf{V}(X) + \mathbf{V}(Y) + 2\lambda \operatorname{cov}(X, Y)$. C'est un trinôme en λ car $\mathbf{V}(X) \neq 0$.
De plus, $\mathbf{V}(\lambda X + Y) \geqslant 0$ donc le discriminant Δ est négatif ou nul.
Or $\Delta = 4\operatorname{cov}(X, Y)^2 - 4\mathbf{V}(X)\mathbf{V}(Y)$ donc :

$$|\operatorname{cov}(X, Y)| \leqslant \sqrt{\mathbf{V}(X)\mathbf{V}(Y)} = \sigma(X)\sigma(Y) \quad \text{ce qui donne} \quad |\rho(X, Y)| \leqslant 1$$

Il y a égalité en présence d'une racine double ; il existe donc λ tel que $\mathbf{V}(\lambda X + Y) = 0$, c'est-à-dire que $\lambda X + Y$ est une variable certaine. Il existe donc $\mu \in \mathbb{R}$ tel que $\lambda X + Y = \mu$. ■

D – Exemples de sommes de variables aléatoires

Lemme 18.43 : Formule de Vandermonde

On considère trois entiers naturels m, n, p avec $p \leqslant n + m$. Alors,

$$\sum_{k=0}^{p} \binom{n}{k} \binom{m}{p-k} = \binom{n+m}{p}$$

Démonstration

Donnons une preuve non combinatoire de ce résultat. Posons pour cela $P = (X+1)^n$ et $Q = (X+1)^m$.

- Le terme de degré p du polynôme $PQ = (X+1)^{n+m}$ a pour expression $\binom{n+m}{p} X^p$.

- En utilisant la formule du produit de deux polynômes, on trouve également :

$$\sum_{i+j=p}\binom{n}{i}\binom{m}{j}X^{i+j}=\sum_{k=0}^{p}\binom{n}{k}\binom{m}{p-k}X^{p}$$

D'où l'égalité par identification. ■

Théorème 18.44 : Somme de variables suivant une loi binomiale

(i) Si X et Y sont deux variables indépendantes définies sur $(\Omega,\mathscr{A},\mathbf{P})$ suivant des lois binomiales de paramètres respectifs (n_1,p) et (n_2,p) alors $X+Y$ suit une loi binomiale de paramètre (n_1+n_2,p).

(ii) Plus généralement, si $X_1,\dots,X_n$ sont n variables aléatoires mutuellement indépendantes suivant des lois binomiales de paramètres respectifs $(m_1,p),\dots,(m_n,p)$ alors $X_1+\dots+X_n$ suit une loi binomiale de paramètre $(m_1+\dots+m_n,p)$.

Démonstration

(i) Supposons donc que $X\hookrightarrow\mathscr{B}(n_1,p)$ et $Y\hookrightarrow\mathscr{B}(n_2,p)$ sont deux variables indépendantes.
Alors $X(\Omega)=[\![0,n_1]\!]$, $Y(\Omega)=[\![0,n_2]\!]$. En posant $Z=X+Y$, $Z(\Omega)=[\![0,n_1+n_2]\!]$ et en vertu du résultat précédent, pour tout $k\in[\![0,n_1+n_2]\!]$,

$$\mathbf{P}(Z=k)=\sum_{i=0}^{k}\mathbf{P}(X=i)\mathbf{P}(Y=k-i)=\sum_{i=0}^{k}\binom{n_1}{i}\binom{n_2}{k-i}p^kq^{n_1+n_2+k}=\binom{n_1+n_2}{k}p^kq^{n_1+n_2-k}$$

On a bien $Z\hookrightarrow\mathscr{B}(n_1+n_2,p)$.

(ii) On raisonne par récurrence en utilisant l'indépendance des variables $X_1+\dots+X_n$ et de X_{n+1}. ■

En conséquence, la somme de n variables indépendantes qui suivent une loi de Bernoulli de paramètre p est une variable aléatoire qui suit une loi binomiale de paramètre (n,p).

Corollaire 18.45

Si $X\hookrightarrow\mathscr{B}(n,p)$ alors $\mathbf{E}(X)=np$ et $\mathbf{V}(X)=npq$.

Théorème 18.46 : Somme de variables suivant une loi de Poisson

Si X et Y sont deux variables indépendantes de $(\Omega,\mathscr{A},\mathbf{P})$ suivant des lois de Poisson de paramètres respectifs λ et μ alors $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.
De manière générale, la somme de n variables indépendantes suivant une loi de Poisson de paramètres respectifs $\lambda_1,\dots,\lambda_n$ est une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda_1+\ldots+\lambda_n$.

Démonstration

Supposons donc que $X\hookrightarrow\mathscr{P}(\lambda)$ et $Y\hookrightarrow\mathscr{P}(\mu)$ sont deux variables indépendantes.
Alors $X(\Omega)=\mathbb{N}$, $Y(\Omega)=\mathbb{N}$. Alors, en posant $Z=X+Y$, $Z(\Omega)=\mathbb{N}$. On a de plus :

$$\begin{aligned}\forall n\in\mathbb{N}\quad \mathbf{P}(Z=n)&=\sum_{i=0}^{n}\mathbf{P}(X=i)\mathbf{P}(Y=n-i)\\&=\sum_{i=0}^{n}\mathrm{e}^{-\lambda}\frac{\lambda^i}{i!}\mathrm{e}^{-\mu}\frac{\mu^{n-i}}{(n-i)!}=\frac{\mathrm{e}^{-(\lambda+\mu)}}{n!}\sum_{i=0}^{n}\binom{n}{i}\lambda^i\mu^{n-i}=\mathrm{e}^{-(\lambda+\mu)}\cdot\frac{(\lambda+\mu)^n}{n!}\end{aligned}$$

■

V | Fonctions génératrices

Dans toute cette partie, on considérera une variable aléatoire discrète X définie sur un univers probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ que l'on supposera de plus à valeurs dans $\mathbb{N}$.

Définition 18.47 : Fonction génératrice

La fonction génératrice de la variable X est définie par :

$$G_X : t \mapsto \mathbf{E}\left(t^X\right) = \sum_{n=0}^{+\infty} \mathbf{P}(X=n) t^n$$

On appelle série génératrice la série entière associée. On rappelle que G est définie sur un intervalle de la forme

$$]-R,R[, \quad]-R,R], \quad [-R,R[\quad \text{ou bien} \quad [-R,R]$$

où R est le rayon de convergence de la série entière de coefficients $\mathbf{P}(X=n)$.

De plus, cette série converge pour $t=1$ et on a : $G_X(1) = \sum_{n=0}^{+\infty} \mathbf{P}(X=n) = 1$ car la famille $((X=n)_{n\in\mathbb{N}})$ forme un système complet d'événements. D'où le résultat suivant.

Proposition 18.48

Le rayon de convergence de la série entière $\sum_{n\in\mathbb{N}} \mathbf{P}(X=n) t^n$ est supérieur ou égal à 1.

Dans le cas où la variable aléatoire est finie, la fonction génératrice est polynomiale et $R = +\infty$.

Nom	Notation	$G(t)$
Bernoulli	$\mathscr{B}(p)$	$q+pt$
Binomiale	$\mathscr{B}(n,p)$	$(q+pt)^n$
Uniforme	$\mathscr{U}([\![1;n]\!])$	$\dfrac{t-t^{n+1}}{n(1-t)}$
Géométrique	$\mathscr{G}(p)$	$\dfrac{pt}{1-qt}$
Poisson	$\mathscr{P}(\lambda)$	$\mathrm{e}^{\lambda(t-1)}$

FONCTION GÉNÉRATRICE DES LOIS USUELLES

La fonction G_X est de classe $\mathscr{C}^\infty$ sur l'intervalle ouvert $]-R,R[$ et on peut dériver terme à terme la somme de la série. On sait de plus que G_X est la somme de sa série de Taylor :

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

La connaissance de G_X nous permet donc de retrouver la loi de X. On dit alors que la loi de X est caractérisée par sa série génératrice. Il sera dans certains cas intéressant de passer par la fonction génératrice afin de déterminer la loi de probabilité d'une variable donnée.

C'est notamment le cas pour la somme de deux variables aléatoires indépendantes.

Théorème 18.49 : Fonction génératrice de la somme de deux variables indépendantes

Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{N}$.
Notons G_X et G_Y leurs fonctions génératrices de rayon de convergence respectifs R_X et R_Y.
Si X et Y sont indépendantes, la fonction génératrice G_{X+Y} de $X+Y$ est définie sur au moins $]-R,R[$ avec $R \geqslant \min(R_X, R_Y)$ et :

$$\forall t \in \mathscr{D}_X \cap \mathscr{D}_Y \quad G_{X+Y}(t) = \mathbf{E}\left(t^{X+Y}\right) = \mathbf{E}\left(t^X\right)\mathbf{E}\left(t^Y\right) = G_X(t) G_Y(t)$$

Démonstration

Par produit de Cauchy de séries absolument convergentes,

$$\begin{aligned} G_X(t)G_Y(t) &= \left(\sum_{n=0}^{+\infty} \mathbf{P}(X=n)t^n\right)\cdot\left(\sum_{n=0}^{+\infty} \mathbf{P}(Y=n)t^n\right) \\ &= \sum_{n=0}^{+\infty}\sum_{k=0}^{n} \mathbf{P}(X=k)\mathbf{P}(Y=n-k)t^n = \sum_{n=0}^{+\infty} \mathbf{P}(X+Y=n)t^n = G_{X+Y}(t) \end{aligned}$$

Exercice 17

Retrouver la fonction génératrice d'une loi binomiale.

Exercice 18

Retrouver la loi de la somme de deux variables aléatoires indépendantes suivant une loi de Poisson.

Théorème 18.50 : Fonction génératrice et moments

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{N}$.

(i) La variable aléatoire X admet une espérance $\mathbf{E}(X)$ si et seulement si G_X est dérivable en 1. Si tel est le cas, $\mathbf{E}(X) = G'_X(1)$.

(ii) La variable aléatoire X admet une variance si et seulement si G_X est deux fois dérivable en 1.

Démonstration

Nous admettrons l'équivalence demandée. Supposons que $R > 1$. Alors, d'après le théorème de dérivation terme à terme de la somme d'une série entière, G_X est dérivable en 1 et :

$$G'_X(1) = \sum_{n=1}^{+\infty} n\mathbf{P}(X=n) = \sum_{n=0}^{+\infty} n\mathbf{P}(X=n) = \mathbf{E}(X)$$

De plus, en dérivant une nouvelle fois,

$$G''_X(1) = \sum_{n=0}^{+\infty} n(n-1)\mathbf{P}(X=n) = \mathbf{E}(X(X-1))$$

Comme $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}(X)^2$, $\mathbf{V}(X) = G''_X(1) + G'_X(1)(1 - G'_X(1))$. ■

L'expression de la variance en fonction de $G'_X(1)$ et de $G''_X(1)$ n'est pas à retenir.

VI | Convergence et approximations

A – Approximation d'une loi binomiale

Théorème 18.51 : Approximation d'une loi binomiale par une loi de Poisson

Si pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathscr{B}(n, p_n)$ et si $\lim\limits_{n\to+\infty} np_n = \lambda$, alors on a :

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

Démonstration

Pour tout $k \in \mathbb{N}$,

$$\mathbf{P}(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n!}{k!(n-k)!} \times \frac{(np_n)^k}{n^k} \times (1-p_n)^{n-k}$$
$$= \frac{1}{k!}\left[\prod_{i=0}^{k-1}(n-i)\right]\frac{1}{n^k}(np_n)^k(1-p_n)^{n-k} = \frac{1}{k!}\left[\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)\right](np_n)^k(1-p_n)^{n-k}$$

k est une constante et $p_n \xrightarrow[n\to+\infty]{} 0$. Ainsi, quand $n \longrightarrow +\infty$,

$$\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right) \xrightarrow[n\to+\infty]{} 1 \quad \text{et} \quad (np_n)^k \xrightarrow[n\to+\infty]{} \lambda^k$$

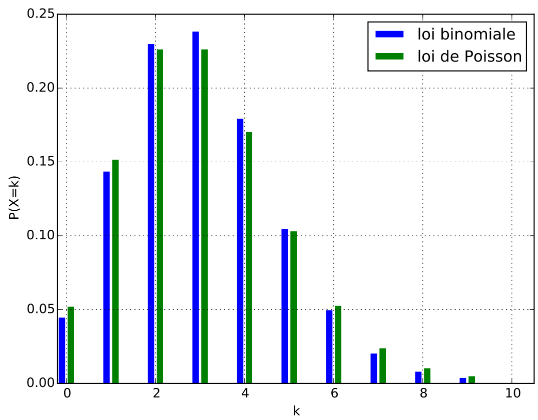
De plus, comme $(n-k)\ln(1-p_n) \underset{n\to+\infty}{\sim} -np_n$,

$$(1-p_n)^{n-k} = \mathrm{e}^{(n-k)\ln(1-p_n)} \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\lambda}$$

On a donc bien $\mathbf{P}(X_n = k) \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}$. ■

On peut donc approcher numériquement toute loi binomiale de paramètre (n, p) par une loi de Poisson de paramètre np mais la qualité de l'approximation dépendra des valeurs de n et p. Dans la pratique, on considère l'approximation valable pour $n \geqslant 30$ et $p \leqslant 0.1$.

```
import numpy as np
import matplotlib.pyplot as plt

p = 0.1
n = 30

def binom(n, k):
    return np.math.factorial(n)/(np.math.factorial(k)*np.math.
      factorial(n-k))

def B(k):
    return binom(n, k)*p**k*(1-p)**(n-k)

def \P(k):
    return np.exp(-n*p)*(n*p)**k/np.math.factorial(k)

for k in range(10):
    plt.plot([k-0.1, k-0.1], [0, B(k)], 'b', linewidth = 6)
    plt.plot([k+0.1, k+0.1], [0, \P(k)], 'g', linewidth = 6)
plt.axis([-0.2,10.5,0,0.25])
plt.grid(True)
plt.xlabel('k')
plt.ylabel('\P(X=k)')
plt.legend(['loi binomiale', 'loi de Poisson'])
plt.show()
```



Approximation d'une loi binomiale par une loi de Poisson

B – Loi faible des grands nombres

Lemme 18.52 : Inégalité de Markov

Soit X une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, à valeurs positives et admettant une espérance.

$$\forall a > 0 \quad \mathbf{P}(X \geqslant a) \leqslant \frac{\mathbf{E}(X)}{a}$$

Démonstration

Soit $a > 0$. Posons $D = \{x \in X(\Omega) \mid x \geqslant a\}$.

$$\mathbf{E}(X) = \sum_{x\in X(\Omega)} x\mathbf{P}(X=x) = \sum_{x\in D} x\mathbf{P}(X=x) + \sum_{x\in\overline{D}} x\mathbf{P}(X=x) \geqslant \sum_{x\in D} x\mathbf{P}(X=x) \geqslant a\sum_{x\in D}\mathbf{P}(X=x) = a\mathbf{P}(x \geqslant a)$$

Proposition 18.53 : Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2.

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

Démonstration

On applique l'inégalité de Markov à la variable $(X - \mathbf{E}(X))^2$, positive et d'espérance finie, en prenant $a = \varepsilon^2$.

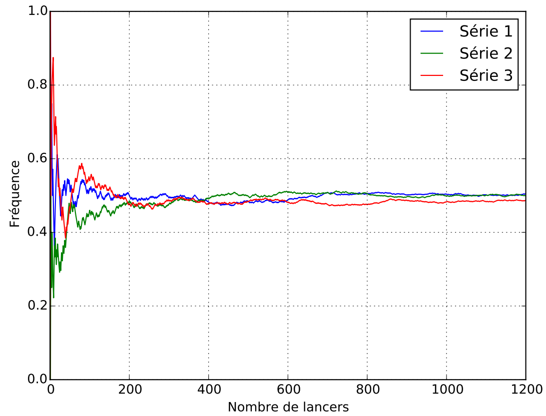
Cette inégalité, bien que grossière comme nous pourrons le constater dans l'exemple ci-dessous, majore la probabilité qu'une variable aléatoire s'écarte de sa moyenne.

Exercice 19

- Comparer $\mathbf{P}(|X - \mathbf{E}(X)| \geqslant 1)$ et $\mathbf{V}(X)$ dans le cas où X représente la somme des chiffres obtenus lors du lancer de deux dés.
- Trouver $n \in \mathbb{N}$ tel que $\mathbf{P}(X \leqslant n) \leqslant \frac{1}{2}$.

Cette inégalité a avant tout un intérêt théorique. Nous allons l'utiliser pour prouver la loi faible des grands nombres.

Mais avant de présenter le résultat, intéressons-nous aux lancers successifs d'une pièce équilibrée. Intuitivement, la fréquence d'apparition de « Pile » au cours d'un très grand nombre de lancers devrait être proche de $\frac{1}{2}$. De manière plus générale, pour déterminer la probabilité d'un événement A, on cherchera à répéter l'expérience aléatoire sous-jacente un maximum de fois dans l'espoir que la fréquence de réalisation soit proche de la probabilité recherchée. Fol espoir ? Non, même si rien ne nous garantit qu'au cours d'une série particulière de 10 000 lancers d'une pièce équilibrée, la fréquence d'apparition soit proche de $1/2$!



FRÉQUENCE D'APPARITION DE « PILE » AU COURS DE 1200 LANCERS

Considérons une suite de variables aléatoires $(X_n)_{n\geqslant 1}$ indépendantes et de même loi. On suppose qu'elles admettent toutes un moment d'ordre 2. Posons :

$$M_n = \frac{X_1 + \cdots + X_n}{n}$$

M_n s'interprète naturellement comme la moyenne des résultats obtenus au cours des n premiers « tirages ». Remarquons qu'il n'y a aucune raison que M_n suive une loi semblable aux variables aléatoires X_i. Nous allons cependant montrer que la probabilité que M_n s'écarte de l'espérance commune $\mathbf{E}(X_i)$ tend vers 0 quand n tend vers l'infini. C'est la loi faible des grands nombres !

Théorème 18.54 : Loi faible des grands nombres

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On suppose toutes les variables indépendantes et de même loi, admettant une espérance m et un écart type σ. Posons $S_n = \sum_{i=1}^{n} X_i$. Alors,

$$\forall \varepsilon \geqslant 0, \quad \mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \xrightarrow[n \to +\infty]{} 0$$

Démonstration

D'après l'inégalité de Bienaymé-Tchebychev et par indépendance des variables aléatoires,

$$\forall \varepsilon > 0 \quad \mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbf{V}\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$

L'espérance d'une variable aléatoire correspond donc au résultat moyen obtenu par une succession d'expériences identiques répétées un grand nombre de fois.

VII | Lois usuelles – Synthèse

Nom	**Notation**	$X(\Omega)$	$\mathbf{P}(X = k)$	$\mathrm{E}(X)$	$\mathrm{V}(X)$	$G(t)$
Bernoulli	$\mathscr{B}(p)$	$\{0;1\}$	$\begin{cases} p \text{ si } k = 0 \\ 1-p \text{ si } k = 1 \end{cases}$	p	pq	$1-p+pt$
Binomiale	$\mathscr{B}(n,p)$	$[\![0;n]\!]$	$\binom{n}{k} p^k q^{n-k}$	np	npq	$(1-p+pt)^n$
Uniforme	$\mathscr{U}([\![1;n]\!])$	$[\![1;n]\!]$	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{t-t^{n+1}}{n(1-t)}$
Géométrique	$\mathscr{G}(p)$	$\mathbb{N}^*$	$q^{k-1}p$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pt}{1-qt}$
Poisson	$\mathscr{P}(\lambda)$	$\mathbb{N}$	$\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ	$\mathrm{e}^{\lambda(t-1)}$