Variables aléatoires finies

« On appelle ça, un peu obscurément, la loi des grands nombres. Par quoi l'on peut dire à peu près que, si un homme se tue pour telle raison et un autre pour telle autre, dès qu'on a affaire à un très grand nombre, le caractère arbitraire et personnel de ces motifs disparaît, et il ne demeure... précisément, qu'est-ce qui demeure ? [...]
C'est ce que nous autres profanes, appelons tout bonnement la moyenne, c'est-à-dire quelque chose dont on ne sait absolument pas ce que c'est. [...]
Quoi qu'il en soit, en effet, la possibilité d'une vie ordonnée repose tout entière sur cette loi des grands nombres ; si cette loi de compensation n'existait pas, il y aurait des années où il ne se produirait rien, et d'autres ou plus rien ne serait sûr ; les famines alterneraient avec l'abondance, les enfants seraient en défaut ou en excès et l'humanité voletterait de côté et d'autre entre ses possibilités célestes et ses possibilités infernales comme les petits oiseaux lorsqu'on s'approche de leur cage. »
L'homme sans qualités, Robert Musil

Notations : Dans ce chapitre, Ω représente un ensemble fini, et $\mathbf{P}$ une probabilité sur Ω.

1 Notion de variable aléatoire réelle ou complexe

A Définition

Exemple *(Exemple suivi...)*

On joue aux dés. Chaque joueur lance deux dés et le gagnant est celui dont la somme des deux dés est la plus grande.
La modélisation naturelle consiste à considérer l'univers constitué de l'ensemble des couples de valeurs possibles et à le munir d'une probabilité uniforme $\mathbf{P}$.

$$\Omega = [\![1,6]\!]^2.$$

Cependant, ce n'est pas le couple (dé 1, dé 2) qui nous intéresse, mais simplement sa somme. L'idée est donc de « transformer » l'univers Ω en un autre plus adapté à la question. Pour cela, on construit l'application $X : \Omega \to [\![2,12]\!]$ qui associe à chaque issue possible, la somme des dés correspondante.

$$X : \left\{ \begin{array}{ccc} \Omega = [\![1,6]\!]^2 & \to & [\![2,12]\!] \\ \omega = (i,j) & \mapsto & i+j. \end{array} \right.$$

On dit que X est une variable aléatoire, et le nouvel univers de travail, beaucoup plus simple, est $[\![2,12]\!]$.
De tous les événements possibles, seuls nous intéressent, pour $k \in [\![2,12]\!]$,

$$A_k : \text{« la somme des dés vaut } k \text{ »}.$$

Ainsi, parmi les 2^{36} événements possibles, nous n'en retenons que 11.
Pour décrire l'événement A_k dans Ω, il faut chercher les antécédents de k par X.
Par exemple, pour A_3, on cherche tous les couples $(i,j) \in \Omega$ tels que $X(i,j) = 3$.
On trouve l'événement $A_3 = \big\{(1,2),(2,1)\big\}$ et on en déduit la probabilité de l'événement A_3 :

$$\mathbf{P}(A_3) = \frac{2}{36} = \frac{1}{18}.$$

On notera cette probabilité $\mathbf{P}(X=3)$.

À présent, formalisons cette notion :

Définition 1.1 *(Variable aléatoire finie)*

Une **variable aléatoire** finie X est une application de Ω dans un ensemble E.
La variable est dite réelle, si elle est à valeurs dans une partie finie $X(\Omega) \subset \mathbf{R}$, elle est à valeurs complexes si $X(\Omega) \subset \mathbf{C}$.
Dans ce cours, $X(\Omega)$ sera appelé indifféremment *univers image* ou *support* de X.

Remarque : L'univers étant supposé fini, on peut écrire $X(\Omega)$ sous la forme

$$X(\Omega) = \{x_i,\ i \in I\}$$

avec I une partie finie de $\mathbf{N}$ ou $\mathbf{Z}$. Par exemple $I = [\![1, n]\!]$.

Explications

La variable aléatoire permet de traduire le résultat d'une expérience dans un ensemble sur lequel on sait travailler : on a plus de facilités à manipuler des nombres que des couples de dés ou des paires de chaussettes de différentes couleurs.

⚠ Le vocabulaire consacré par l'usage est assez troublant : on parle de variable aléatoire, bien qu'il s'agisse en fait d'une *application*. Ces difficultés de notation et de langage sont omniprésentes en probabilités, mais elles se justifient par l'usage.

Notation

Soit X une variable aléatoire dans $X(\Omega)$.
Pour A une partie de $X(\Omega)$, on note $[X \in A]$ l'événement

$$[X \in A] = \{\omega \text{ tel que } X(\omega) \in A\} = X^{-1}(A).$$

Ainsi,

$$\mathbf{P}(X \in A) = \mathbf{P}(\{\omega \text{ tel que } X(\omega) \in A\}) = \mathbf{P}\left(X^{-1}(A)\right).$$

Lorsque $A = \{x\}$ est un singleton, on note plus simplement :

$$[X = x] = \{\omega \text{ tel que } X(\omega) = x\} = X^{-1}(\{x\}).$$

Ainsi

$$\mathbf{P}(X = x) = \mathbf{P}(\{\omega \text{ tel que } X(\omega) = x\}) = \mathbf{P}\left(X^{-1}(\{x\})\right).$$

Explications

Plutôt que d'étudier la réalisation d'un événement $\mathscr{A} \subset \Omega$, on préfère chercher les événements qui réalisent certaines valeurs de X.
Ainsi $[X = A]$ représente tous les événements pour lesquels X prend une valeur dans A.

Exemple *(Exemple suivi)*

Avec l'exemple du lancer de deux dés (équilibrés), on définit la variable aléatoire X qui donne leur somme.
Calculer $\mathbf{P}(X = 2)$, $\mathbf{P}(X \in \{3\})$ et $\mathbf{P}(X \in \{2, 3\})$.

Solution :

$\mathbf{P}(X \in \{2\}) = \mathbf{P}(X = 2) = \mathbf{P}(\{(1,1)\}) = \frac{1}{36}$.

$\mathbf{P}(X \in \{3\}) = \mathbf{P}(X = 3) = \mathbf{P}(\{(1,2),(2,1)\}) = \frac{2}{36} = \frac{1}{18}$.

$\mathbf{P}(X \in \{2,3\}) = \mathbf{P}(\{(1,1),(1,2),(2,1)\}) = \frac{3}{36} = \frac{1}{12}$.

⚠ Si $A = \{x\}$ est un singleton, il n'y a aucune raison pour que l'événement correspondant dans Ω soit lui-même un singleton.
Rappelez vous l'exemple introductif : $[X = 3] = A_3$ n'est pas un singleton.

Propriété 1.2

Soit X une variable aléatoire finie de Ω dans $X(\Omega)$,

$\{[X = x],\ x \in X(\Omega)\}$ forme un système complet d'événements de Ω.

Preuve

- Tout d'abord, tout $\omega \in \Omega$ admet une image par X dans $X(\Omega)$:
  Si on note $x = X(\omega)$, alors $\omega \in [X = x]$, donc $\omega \in \bigcup_{x \in X(\Omega)} [X = x]$.
  Ceci étant vrai pour tous les $\omega \in \Omega$, on a bien $\Omega \subset \bigcup_{x \in X(\Omega)} [X = x]$.
  L'autre inclusion étant évidente, $\Omega = \bigcup_{x \in X(\Omega)} [X = x]$.
- Soient $(x, y) \in (X(\Omega))^2$.
  Si on suppose que $\omega \in [X = x] \cap [X = y]$, alors $X(\omega) = x$ et $X(\omega) = y$, donc $x = y$.
  Ainsi, pour $x \neq y$, $[X = x] \cap [X = y] = \emptyset$.

L'ensemble $\{[X = x],\ x \in X(\Omega)\}$ forme donc bien un système complet d'événements de Ω. ■

Propriété 1.3 *(Loi de probabilité de X)*

Pour X une variable aléatoire $\Omega \to X(\Omega)$,
on définit $\mathbf{P}_X$ par

$$\mathbf{P}_X : \begin{cases} \mathcal{P}(X(\Omega)) & \to [0,1] \\ A & \mapsto \mathbf{P}(X \in A). \end{cases}$$

L'application $\mathbf{P}_X$ définit une probabilité.
On appelle **loi de probabilité** de X, l'application

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

Lorsque deux variables aléatoires X et Y ont même loi, on peut noter $X \sim Y$.

La notation $\mathbf{P}_X$ utilisée pour ce cours n'a rien d'officiel.

Remarque : Il est aussi possible de travailler sur un ensemble plus gros $\mathscr{X}$ qui contient $X(\Omega)$. Auquel cas, tous les événements supplémentaires sont de probabilité nulle.

Explications

La détermination de la probabilité $\mathbf{P}_X$ permet d'oublier complètement l'univers naturel pour ne considérer que $X(\Omega)$.
L'univers image étant fini (de cardinal inférieur à celui de départ), la donnée des probabilités sur les événements élémentaires $[X = x]$ suffit à déterminer $\mathbf{P}_X$.
Remarque : Dans le cas d'un univers infini, ce résultat ne s'applique pas toujours et on pourra être amené à définir différement la loi de probabilité.

Preuve

Il faut prouver que $\mathbf{P}_X$ est bien une probabilité.
Pour cela on peut voir que

$$\forall x \in X(\Omega),\ \mathbf{P}_X(\{x\}) = \mathbf{P}(X = x) = \mathbf{P}(X^{-1}(\{x\})) \in [0, 1].$$

(car $X^{-1}(\{x\})$ est un événement sur Ω et $\mathbf{P}$ est une probabilité sur Ω).
Or $\{[X = x],\ x \in X(\Omega)\}$ forme un système complet d'événements donc

$$\sum_{x \in X(\Omega)} \mathbf{P}_X(\{x\}) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x) = 1.$$

Donc $\mathbf{P}_X$ est bien une probabilité (définie par ses événements élémentaires). ■

Exemple *(Exemple suivi)*

Donner la loi de X pour la somme des deux dés.
Solution :

x	1	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}(X = x)$	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Si on fait la somme des probabilités élémentaires de la deuxième ligne, on doit bien retrouver 1.

Exemple

Une variable aléatoire sur $[\![1, n]\!]$ suit une **probabilité uniforme** si

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

On remarque ici que seul l'ensemble $X(\Omega) = [\![1, n]\!]$ est important, on ne s'occupe plus du tout de l'univers Ω sous-jacent.

L'équiprobabilité sur $X(\Omega)$ n'est pas liée à celle sur Ω : si la variable suit une probabilité uniforme sur $X(\Omega)$, cela ne veut pas dire que l'expérience est équiprobable sur Ω. Et réciproquement, si l'expérience est équiprobable sur Ω, cela ne veut pas dire que la variable ait une probabilité uniforme sur $X(\Omega)$. Par exemple, pour la somme des deux dés, la probabilité est uniforme sur Ω, mais ne l'est pas dans $[\![2, 12]\!]$.

Propriété 1.4

Soit X une variable aléatoire finie de Ω dans $X(\Omega)$, et A une partie de $X(\Omega)$.

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

Preuve

Rappel : on a vu dans le chapitre sur les applications que si $f : E \to F$ est une application, et $(A_i)_{i \in I}$ des parties de F, alors $f^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f^{-1}(A_i)$.
On remarque que c'est également vrai pour une union disjointe. En effet, s'il existe $(i, j) \in I^2$ tel que $x \in f^{-1}(A_i) \cap f^{-1}(A_j)$, alors $f(x) \in A_i$ et $f(x) \in A_j$, donc $f(x) \in A_i \cap A_j$, ce qui prouve que $i = j$ si l'union est initialement disjointe.
Ainsi

$$\begin{aligned}
\mathbf{P}(X \in A) &= \mathbf{P}\left(X^{-1}(A)\right) \quad \text{(définition)} \\
&= \mathbf{P}\left(X^{-1}\left(\bigsqcup_{x \in A} \{x\}\right)\right) \\
&= \mathbf{P}\left(\bigsqcup_{x \in A} X^{-1}(\{x\})\right) \quad \text{(rappel plus haut)} \\
&= \mathbf{P}\left(\bigsqcup_{x \in A} [X = x]\right) \\
&= \sum_{x \in A} \mathbf{P}(X = x) \quad \text{(l'union est disjointe).}
\end{aligned}$$

■

Exemple

Preuve du « rappel » : montrer que pour $(A_i)_{i \in I}$ des parties de $X(\Omega)$

$$X^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} X^{-1}(A_i).$$

Solution :

$$\begin{aligned}
\omega \in X^{-1}\left(\bigcup_{i \in I} A_i\right) &\iff X(\omega) \in \bigcup_{i \in I} A_i \\
&\iff \exists i \in I,\ X(\omega) \in A_i \\
&\iff \exists i \in I,\ \omega \in X^{-1}(A_i) \\
&\iff \omega \in \bigcup_{i \in I} X^{-1}(A_i).
\end{aligned}$$

B Fonction de répartition d'une variable aléatoire réelle

Cette notion est surtout utile pour les variables aléatoires continues qu'on n'étudiera pas cette année, mais elle peut parfois simplifier les raisonnements avec les variables aléatoires discrètes.
La fonction de répartition correspond tout simplement aux « probabilités cumulées croissantes » :

Définition 1.5 *(Fonction de répartition)*

Soit X une variable aléatoire **réelle** sur Ω.
On définit la **fonction de répartition** F_X sur $\mathbf{R}$ par

$$F_X : \left\{ \begin{array}{ccc} \mathbf{R} & \to & \mathbf{R} \\ x & \mapsto & \mathbf{P}(X \leqslant x) = \displaystyle\sum_{x_i \leqslant x} \mathbf{P}(X = x_i). \end{array} \right.$$

⚠ La fonction de répartition est définie à partir d'inégalités sur la variable X, ce qui exclut, de facto, les variables aléatoires à valeurs complexes.

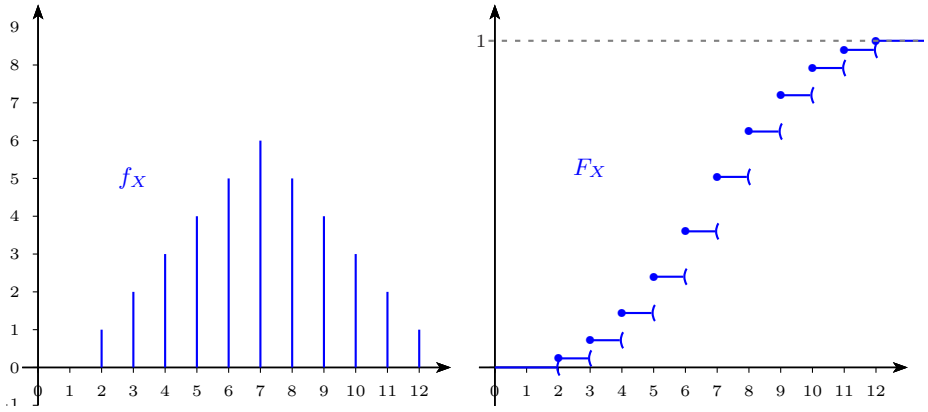
Remarque : Avec des probabilités continues, on prendrait l'intégrale qui est l'équivalent continu de la somme.

Exemple *(Exemple suivi)*

Donner la fonction de répartition pour la variable X correspondant à la somme des deux dés.

Solution :

x	1	2	3	4	5	6	7	8	9	10	11	12
$f_X(x) = \mathbf{P}(X = x)$	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
$F_X(x) = \mathbf{P}(X \leqslant x)$	0	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$



Propriété 1.6

Une fonction de répartition sur $\mathbf{R}$ est positive, croissante et

$$\lim_{x \to -\infty} F_X(x) = 0 \qquad \text{et} \qquad \lim_{x \to +\infty} F_X(x) = 1.$$

Preuve

Soit $x \leqslant y$, alors $[X \leqslant x] \subset [X \leqslant y]$, et par croissance de la probabilité $\mathbf{P}(X \leqslant x) \leqslant \mathbf{P}(X \leqslant y)$, donc $F_X(x) \leqslant F_X(y)$: la fonction F_X est croissante.
Si $X(\Omega)$ est fini, alors il admet un minimum x_m et un maximum x_M.
$\forall x < x_m$, $F_X(x) = 0$, donc $\lim\limits_{x \to -\infty} F_X(x) = 0$. $\forall x \geqslant x_M$, $F_X(x) = 1$, donc $\lim\limits_{x \to +\infty} F_X(x) = 1$. ■

Propriété 1.7

F_X décrit parfaitement la loi de X : la donnée de F_X permet de retrouver la probabilité.
Lorsque les probabilités sont à valeurs entières,

$$\forall k \in \mathbf{Z},\ \mathbf{P}(X = k) = \mathbf{P}(X \leqslant k) - \mathbf{P}(X \leqslant k-1) = F_X(k) - F_X(k-1).$$

Dans le cas général :

$$\forall x \in X(\Omega),\ \mathbf{P}(X = x) = \mathbf{P}(X \leqslant x) - \lim_{x' \to x^-} \mathbf{P}(X \leqslant x') = F_X(x) - F_X(x^-).$$

On note $F_X(x^-)$ la limite à gauche de F_X en x.

C Image d'une variable aléatoire par une fonction

Propriété 1.8

Si X est une variable aléatoire de Ω dans $X(\Omega)$ et $u : X(\Omega) \to F$ une application, alors $u(X)$ définit une nouvelle variable aléatoire sur Ω.

Exemple

Dans le cas d'un double lancer de dés, on peut par exemple associer un gain à chaque résultat.
On note X la variable aléatoire correspondant à somme des valeurs des deux dés, et u l'application qui définit le gain en fonction du résultat.

$$u : \left\{ \begin{array}{ccc} [\![2, 12]\!] & \to & \mathbf{R} \\ x & \mapsto & \left\{ \begin{array}{ll} 10 & \text{si } x = 12 \\ 1 & \text{si } x \in [\![9, 11]\!] \\ -1 & \text{sinon} \end{array} \right. \end{array} \right.$$

On trouve donc $\mathbf{P}(u(X) = 10) = \mathbf{P}(X = 12) = \dfrac{1}{36}$

$$\mathbf{P}(u(X)=1)=\mathbf{P}(X\in[\![9,11]\!])=\frac{4}{36}+\frac{3}{36}+\frac{2}{36}=\frac{9}{36}=\frac{1}{4}$$

$$\mathbf{P}(u(X)=-1)=1-\mathbf{P}(u(X)\neq -1)=1-\frac{1}{36}-\frac{9}{36}=\frac{26}{36}=\frac{13}{18}$$

On peut alors construire le tableau avec la nouvelle loi :

y	-1	1	10
$\mathbf{P}(u(X)=y)$	$\frac{26}{36}$	$\frac{9}{36}$	$\frac{1}{36}$

Explications

Cela consiste à refaire exactement ce que nous avions fait en début de chapitre avec X, mais au lieu de passer de l'espace probabilisé « naturel » $(\Omega,\mathbf{P})$ dans $(X(\Omega),\mathbf{P}_X)$, on passe cette fois-ci de l'univers $(X(\Omega),\mathbf{P}_X)$ dans un nouvel univers $((u\circ X)(\Omega),\mathbf{P}_{u\circ X})$.
Cela revient à considérer la variable aléatoire $u\circ X$ au lieu de X.

2 Espérance et écart type

A Espérance

Définition 2.1 *(Espérance d'une variable aléatoire et théorème de transfert)*

Soit X une variable aléatoire finie de Ω dans $\mathbf{R}$ ou $\mathbf{C}$.
L'espérance de X est définie par

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

Lorsque $\mathbf{E}(X)=0$, on dit que la variable aléatoire est *centrée*.

L'espérance est parfois aussi appelée la *moyenne*, elle représente simplement la valeur moyenne de X sur un très grand nombre d'expériences.
Le terme d'**espérance** fait référence à la théorie des jeux. Il représente le gain moyen que je peux espérer après un grand nombre d'essais.

Dans la pratique, pour calculer l'espérance, on utilise souvent la deuxième expression. En effet, elle a l'avantage de ne pas faire intervenir l'univers initial explicitement, mais uniquement l'univers image.

Preuve *(Égalité des deux formulations)*

$$\begin{aligned}\sum_{\omega\in\Omega}X(\omega)\mathbf{P}(\{\omega\})&=\sum_{x\in X(\Omega)}\ \sum_{\omega\in X^{-1}(\{x\})}X(\omega)\mathbf{P}(\{\omega\})\\ &\text{car les événements } [X=x] \text{ forment un système complet d'événements.}\\ &=\sum_{x\in X(\Omega)}\ \sum_{\omega\in X^{-1}(\{x\})}\mathbf{P}(\{\omega\})\\ &=\sum_{x\in X(\Omega)}x\mathbf{P}(\{\omega\in X^{-1}(\{x\})\})\quad\text{car les événements sont disjoints}\\ &=\sum_{x\in X(\Omega)}x\mathbf{P}(X=x).\end{aligned}$$

■

Exemple

Calculer la valeur moyenne de la somme lorsqu'on lance deux dés équilibrés.
Calculer l'espérance du gain (voir exercice page 4).
Solution :

$$\begin{aligned}\mathbf{E}(X)&=\sum_{k=1}^{12}k\mathbf{P}(X=k)\\ &=2\frac{1}{36}+3\frac{2}{36}+4\frac{3}{36}+5\frac{4}{36}+6\frac{5}{36}+7\frac{6}{36}+8\frac{5}{36}+9\frac{4}{36}+10\frac{3}{36}+11\frac{2}{36}+12\frac{1}{36}.\end{aligned}$$

$\mathbf{E}(X)=7$.
Si on considère l'espérance de gain à partir de la loi $u(X)$ définie précédemment :
$\mathbf{E}(u(X))=-\frac{13}{36}+\frac{9}{36}+10\frac{1}{36}=-\frac{7}{36}\approx -0,19$.
L'espérance de gain est négative, nous n'avons pas intérêt à jouer à ce jeu. Nous perdrions en moyenne $0,19$ à chaque coup.

Exemple *(La roulette)*

La roulette est composée de 37 numéros : $[\![0,36]\!]$.
Faire un pari simple consiste à parier sur la moitié des numéros sauf le 0.
Par exemple, vous pariez sur les rouges.
Si un rouge sort, vous gagnez une fois votre mise, si un noir sort, vous perdez votre mise, et si le zéro sort, vous perdez la moitié de la mise (pour simplifier).
Calculer l'espérance de gain.
Solution :
On définit donc la variable aléatoire qui associe à chaque événement son gain *algébrique* :

- Si ω est un numéro rouge : $X(\omega)=1$,
- Si ω est un numéro noir : $X(\omega)=-1$,
- $X(0)=-\frac{1}{2}$.

L'espérance de gain est donc

$$\mathbf{E}(X)=1\times\mathbf{P}(X=1)-1\times\mathbf{P}(X=-1)-\frac{1}{2}\mathbf{P}(X=-\frac{1}{2})=\frac{18}{37}-\frac{18}{37}-\frac{1}{2}\frac{1}{37}=-\frac{1}{74}.$$

Donc un joueur perd en moyenne $\frac{1}{74}^{\text{ème}}$ de sa mise à chaque coup lorsqu'il fait un pari simple.

Exemple

Donner l'espérance d'une variable uniforme sur $[\![1,n]\!]$.

Solution :

Si X suit une loi uniforme sur $[\![1,n]\!]$, alors

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

Propriété 2.2 *(Propriétés élémentaires)*

X et Y désignent deux variables aléatoires réelles ou complexes.

1. L'espérance est **linéaire** : $\forall (a,b)\in\mathbf{R}^2,\ \mathbf{E}(aX+bY)=a\mathbf{E}(X)+b\mathbf{E}(Y)$.
2. Si $\exists a\in\mathbf{R},\ \forall\omega\in\Omega,\ X(\omega)=a,$ alors $\mathbf{E}(X)=a$.
   L'espérance d'une variable constante est égale à cette constante.

Pour une variable aléatoire X réelle, on a de plus :

3. L'espérance est **positive** : si $\forall\omega\in\Omega,\ X(\omega)\geqslant 0,$ alors $\mathbf{E}(X)\geqslant 0$.
4. L'espérance est **croissante** : si $\forall\omega\in\Omega,\ X(\omega)\leqslant Y(\omega),$ alors $\mathbf{E}(X)\leqslant\mathbf{E}(Y)$.
5. $\left|\mathbf{E}(X)\right|\leqslant\mathbf{E}\left(\left|X\right|\right)$ (inégalité triangulaire).

⚠ La positivité de l'espérance ne veut pas dire que $\mathbf{E}(X)$ est toujours positive. C'est à comprendre dans le même sens que la positivité de l'intégrale par exemple.

Preuve

1. **Linéarité :**
   Soient X,Y deux variables aléatoires réelles, et $(a,b)\in\mathbf{R}^2$.
$$\begin{aligned}\mathbf{E}(aX+bY)&=\sum_{\omega\in\Omega}\Big(aX+bY\Big)(\omega)\,\mathbf{P}(\{\omega\})\\&=\sum_{\omega\in\Omega}\Big(aX(\omega)\mathbf{P}(\{\omega\})+bY(\omega)\mathbf{P}(\{\omega\})\Big)\\&=a\sum_{\omega\in\Omega}X(\omega)\mathbf{P}(\{\omega\})+b\sum_{\omega\in\Omega}Y(\omega)\mathbf{P}(\{\omega\})\quad\text{(linéarité de la somme)}\\&=a\mathbf{E}(X)+b\mathbf{E}(Y).\end{aligned}$$
2. **Espérance d'une constante :**
   Si X est une loi certaine : $\forall\omega\in\Omega, X(\omega)=a\in\mathbf{R}$, alors $\mathbf{E}(X):\sum_{x=a}a\mathbf{P}(X=a)=\sum_{x=a}a1=a$.
3. **Positivité :**
   $\forall x\in X(\Omega),\ x\geqslant 0$ par hypothèse, et $\mathbf{P}(X=x)\geqslant 0$ par définition d'une probabilité.
   Donc $\forall x\in X(\Omega),\ x\mathbf{P}(X=x)\geqslant 0$, et par somme de termes positifs, $\mathbf{E}(X)\geqslant 0$.
4. **Croissance :**
   C'est une conséquence de la linéarité et de la positivité avec pour $X\geqslant Y$, $\mathbf{E}(X)-\mathbf{E}(Y)=\mathbf{E}(X-Y)\geqslant 0$.
5. **Inégalité triangulaire :** $-|X|\leqslant X\leqslant|X|$ donc par croissance et linéarité $-\mathbf{E}(|X|)\leqslant\mathbf{E}(X)\leqslant\mathbf{E}(|X|)$ ce qui donne le résultat voulu.

■

Exemple

Si X est une variable aléatoire, alors sa variable aléatoire centrée est $X-\mathbf{E}(X)$. La linéarité de l'espérance montre bien que cette nouvelle variable aléatoire est d'espérance nulle.

Exemple *(Indicatrice)*

Quelle est l'espérance de $\mathbf{1}_A$ pour A une partie de Ω ?

Solution :

$\mathbf{E}(\mathbf{1}_A)=\sum_{\omega\in A}1\times\mathbf{P}(\{\omega\})+\sum_{\omega\notin A}0\times\mathbf{P}(\{\omega\})=\mathbf{P}(A)$.

Exemple

Soit X une variable aléatoire réelle, quelle est l'espérance de $\mathbf{1}_{[a,b]}(X)$?

Solution :

$\mathbf{E}(\mathbf{1}_{[a,b]}(X))=\sum_{X(\omega)\in[a,b]}1\times\mathbf{P}(\{\omega\})+\sum_{X(\omega)\notin A}0\times\mathbf{P}(\{\omega\})=\mathbf{P}(X\in[a,b])$.

Théorème 2.3 *(Théorème de transfert)*

Soit X une variable aléatoire de Ω sur $X(\Omega)$ et $u:X(\Omega)\to\mathbf{K}$,

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

Preuve

C'est la même chose que la preuve donnée pour relier les deux expressions dans la définition de l'espérance.

$$\begin{aligned}\mathbf{E}\left(u(X)\right)&=\sum_{\omega\in\Omega}u\circ X(\omega)\mathbf{P}(\{\omega\})\\&=\sum_{x\in X(\Omega)}\sum_{\omega\in X^{-1}(\{x\})}u\circ X(\omega)\mathbf{P}(\{\omega\})\\&=\sum_{x\in X(\Omega)}u(x)\sum_{\omega\in X^{-1}(\{x\})}\mathbf{P}(\{\omega\})\\&=\sum_{x\in X(\Omega)}u(x)\mathbf{P}\left(X=x\right).\end{aligned}$$

■

Exemple

Nous avions calculé directement l'espérance de gain avec le lancer des deux dés.

On peut aussi faire le calcul avec le théorème de transfert (même si, dans cet exemple c'est plus plus compliqué puisque nous avions déjà calculé la loi de $u(X)$).

$$\begin{aligned}\mathbf{E}(u(X)) &= \sum_{x\in X(\Omega)} u(x)\mathbf{P}(X=x)\\ &= \sum_{x=1}^{12} u(x)\mathbf{P}(X=x)\\ &= u(1)\cdot 0 + u(2)\frac{1}{36} + u(3)\frac{2}{36} + u(4)\frac{3}{36} + u(5)\frac{4}{36} + u(6)\frac{5}{36} + \dots\\ &\quad \dots + u(7)\frac{6}{36} + u(8)\frac{5}{36} + u(9)\frac{4}{36} + u(10)\frac{3}{36} + u(11)\frac{2}{36} + u(12)\frac{1}{36}\\ &= 0 - \frac{1}{36} - \frac{2}{36} - \frac{3}{36} - \frac{4}{36} - \frac{5}{36} - \frac{6}{36} - \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + 10\frac{1}{36}\\ &= -\frac{7}{36} \approx -0,19.\end{aligned}$$

On retrouve le résultat précédent, et on voit bien que dans notre calcul, si on avait factorisé la somme par les valeurs communes de $u(x)$, on aurait retrouvé l'expression calculée sans le théorème de transfert.

⚠ Il en faut pas se tromper sur le théorème du transfert, en général :

$$\mathbf{E}\left(u(X)\right) \neq u\left(\mathbf{E}(X)\right).$$

Par exemple si X suit la loi : $\mathbf{P}(X=1) = \mathbf{P}(X=-1) = \frac{1}{2}$, on pose $Y = \frac{1}{X}$.
$\mathbf{P}(Y=1) = \mathbf{P}\left(\frac{1}{X}=1\right) = \mathbf{P}(X=1) = \frac{1}{2}$ et de même, $\mathbf{P}(Y=-1) = \frac{1}{2}$.
$\mathbf{E}(X) = \mathbf{E}(Y) = 0$ mais $\frac{1}{\mathbf{E}(X)}$ n'existe pas ! (on ne peut pas diviser par 0).

Exemple *(Inégalité de Jensen)*

Si $f : \mathbf{R} \to \mathbf{R}$ est convexe et X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathbf{P})$, alors

$$f\left(\mathbf{E}\left(X\right)\right) \leqslant \mathbf{E}\left(f(X)\right).$$

Définition 2.4 *(Moment d'ordre k)*

Pour X une variable aléatoire réelle finie, on désigne le **moment d'ordre** k de X par

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

Exemple

Le moment d'ordre 1 est l'espérance.

Remarque : Dans cette définition, on a utilisé le théorème de transfert pour exprimer le moment d'ordre k.

B Variance et écart-type

⚠ Ici, la variance est uniquement définie pour une variable aléatoire **réelle.**

Définition 2.5 *(Variance)*

La **variance** d'une variable aléatoire réelle ou complexe X est définie par

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

C'est le moment d'ordre 2 de la variable centrée $X - \mathbf{E}(X)$.

Explications

Le principe de la variance est de mesurer une dispersion, c'est-à-dire un étalement des valeurs.
Elle mesure donc un étalement moyen autour de l'espérance. Comme l'écart moyen entre X est $\mathbf{E}(X)$ est toujours nul par linéarité de la moyenne, on étudie ici « la moyenne du carré de l'écart à la moyenne ».

⚠ La variance n'est pas *homogène* avec X : lorsque X désigne une longueur, $\mathbf{V}(X)$ désigne une longueur au carré. Il s'agit d'un rapport quadratique.

Propriété 2.6 *(Formule de Kœnig-Huygens)*

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

C'est cette formule que l'on utilise en pratique pour le calcul de la variance.

Preuve

Il suffit de développer le carré et d'utiliser la linéarité de $\mathbf{E}$.

$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E}\left(\left(X - \mathbf{E}(X)\right)^2\right) = \mathbf{E}\left(X^2 - 2X\mathbf{E}(X) + \mathbf{E}(X)^2\right)\\ &= \mathbf{E}\left(X^2\right) - 2\mathbf{E}\left(X\mathbf{E}(X)\right) + \mathbf{E}\left(\mathbf{E}(X)^2\right) \qquad \text{(par linéarité)}\\ &= \mathbf{E}\left(X^2\right) - 2\mathbf{E}(X)^2 + \mathbf{E}(X)^2 \qquad \text{(car } \mathbf{E}(X) \text{ et } \mathbf{E}(X)^2 \text{ sont des constantes)}\\ &= \mathbf{E}\left(X^2\right) - \mathbf{E}(X)^2.\end{aligned}$$

■

Propriété 2.7

Pour toutes constantes réelles a et b,

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

Explications

Il est intuitif que $\mathbf{V}$ est insensible à l'ajout d'une constante : la translation des valeurs ne modifie pas leur dispersion.

Par contre, l'homothétie par a influence également la dispersion : c'est un changement d'échelle. Comme la variance est quadratique, elle est multipliée par a^2.

Preuve

$$\begin{aligned}\mathbf{V}(aX+b) &= \mathbf{E}\left((aX+b)^2\right) - (\mathbf{E}(aX+b))^2 \\ &= \mathbf{E}\left(a^2X^2 + 2abX + b^2\right) - (a\mathbf{E}(X)+b)^2 \\ &= a^2\mathbf{E}\left(X^2\right) + 2ab\mathbf{E}(X) + b^2 - a^2\mathbf{E}(X)^2 - 2ab\mathbf{E}(X) - b^2 = a^2\mathbf{V}(X).\end{aligned}$$

■

Propriété 2.8

La variance est une grandeur positive : pour toute variable aléatoire X, $\mathbf{V}(X) \geqslant 0$.

Remarque : ici, on a bien $\mathbf{V}(X) \geqslant 0$ même si X prend des valeurs négatives. Ce n'est pas la même notion que la positivité de l'espérance qui suppose $X \geqslant 0$.

Exemple

Montrer que la variance est nulle si et seulement si la loi est certaine.

$$\begin{aligned}\mathbf{V}(X) = 0 &\iff (\exists\, a \in X(\Omega),\ \mathbf{P}(X=a) = 1) \\ &\iff X = \mathbf{E}(X).\end{aligned}$$

Solution :

Cas de la variance nulle :

On note l'univers image sous la forme $X(\Omega) = \{x_1, x_2, \cdots, x_n\}$, avec $n \in \mathbf{N}^*$.

Ainsi,

$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E}\left((X - \mathbf{E}(X))^2\right) \\ &= \sum_{i=1}^{n} (x_i - \mathbf{E}(X))^2\ \mathbf{P}(X = x_i)\end{aligned}$$

$\mathbf{V}(X)$ s'écrit donc comme la somme de n termes, tous positifs. Donc $\mathbf{V}(X) \geqslant 0$.

Supposons à présent qu'il existe un terme non nul dans la somme pour $i = i_0$.

Alors $\mathbf{V}(X) = (x_i - \mathbf{E}(X))^2\ \mathbf{P}(X = x_i) + \sum_{i \neq i_0} (x_i - \mathbf{E}(X))^2\ \mathbf{P}(X = x_i)$

$$\geqslant (x_i - \mathbf{E}(X))^2\ \mathbf{P}(X = x_i) > 0$$

Donc $\mathbf{V}(X) > 0$.

Et par contraposée, si $\mathbf{V}(X) = 0$, alors $\forall i \in [\![1, n]\!]$, $(x_i - \mathbf{E}(X))^2\ \mathbf{P}(X = x_i) = 0$.

Ainsi, la probabilité ne peut être non nulle que si $x_i = \mathbf{E}(X)$.

Donc X est constante, égale à son espérance.

Explications

Une loi de variance nulle, n'est absolument pas « dispersée » : toutes les valeurs qu'elle peut prendre sont réduite à un seul point. Son espérance est évidemment égale à cette valeur.

Preuve

La variable aléatoire $(X - \mathbf{E}(X))^2$ est à valeurs positives. La positivité de l'espérance permet de conclure. ■

Cette positivité de la variance permet de définir l'écart type :

Définition 2.9 *(Écart type)*

L'**écart type** d'une variable aléatoire réelle X est défini par

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

Une variable est *réduite* lorsque son écart type est égal à 1.

C'est la raison pour laquelle on note souvent la variance $\mathbf{V}(X) = \sigma^2(X) = \sigma_X^2$.

L'intérêt de l'écart type est qu'il est *homogène* avec X : il est dans la même unité que X.

Propriété 2.10

Pour toutes constantes réelles a et b,

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

⚠ Ne pas oublier la valeur absolue !

Preuve

À faire en exercice. ■

Exemple *(Variable aléatoire centrée réduite)*

Si X une variable aléatoire réelle non certaine, alors $\sigma(X) \neq 0$ et on peut poser

$$Y = \frac{X - \mathbf{E}(X)}{\sigma_X}.$$

Montrer que Y est une variable aléatoire **centrée réduite**.

Solution :

Par linéarité : $\mathbf{E}(Y) = \mathbf{E}\left(\dfrac{X - \mathbf{E}(X)}{\sigma_X}\right) = \dfrac{1}{\sigma_X}(\mathbf{E}(X) - \mathbf{E}(X)) = 0$, donc la variable est centrée.

$\sigma_Y = \sigma\left(\dfrac{X - \mathbf{E}(X)}{\sigma_X}\right) = \dfrac{1}{\sigma_X}(\sigma(X - \mathbf{E}(X))) = \dfrac{1}{\sigma_X}\sigma(X) = 1$, donc la variable est réduite.

3 Lois usuelles

Voici un aperçu rapide des principales lois que l'on rencontre en probabilités finies. Pour l'essentiel, elles ont été vues au lycée.
Lors de l'apprentissage, il est vivement conseillé d'apprendre l'exemple type associé à chaque loi.

A Loi certaine

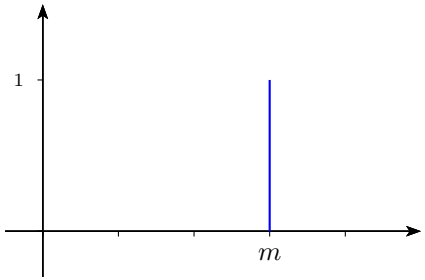
Exemple *(Exemple type : situation déterministe)*
Un processus physique (macroscopique) dont on connait parfaitement les conditions initiales, les conditions de déroulement, et dont on peut donc prévoir l'issue avec certitude. L'événement certain correspond souvent à une constante que l'on ajoute ou soustrait à une autre loi (par exemple une mise initiale dans un jeu de pari : on est totalement sûr de la mettre, quel que soit le tirage qui va suivre).

Définition 3.1 *(Loi certaine)*

Une variable aléatoire X définie sur Ω suit une loi certaine,
s'il existe $m \in \mathbf{R}$ tel que $\mathbf{P}(X = m) = 1$.
On a alors $X(\Omega) = \{m\}$.

Propriété 3.2

Pour une loi certaine égale à m.



Loi de probabilité

Propriété 3.3

Si X est de loi certaine égale à m, alors

$$\mathbf{E}(X) = m \qquad \text{et} \qquad \mathbf{V}(X) = 0.$$

Remarque : On a vu que la réciproque est vraie.

B Loi uniforme : situation équiprobable

Exemple *(Exemple type : lancer d'un dé)*
On lance un dé (équilibré) à 6 faces, on note X la variable aléatoire correspondant au numéro obtenu.

$$\forall k \in [\![1;6]\!],\ \mathbf{P}(X = k) = \frac{1}{6}.$$

Définition 3.4 *(Loi uniforme)*

Soit X une variable aléatoire définie sur Ω à valeurs dans $[\![1,n]\!]$, on dit que X suit une loi uniforme si

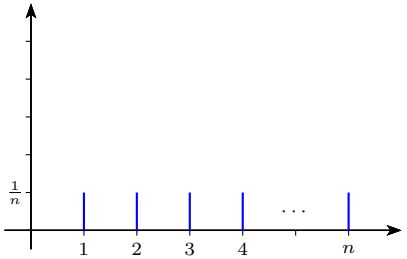
$$\forall k \in [\![1,n]\!],\ \mathbf{P}(X = k) = \frac{1}{n}.$$

On note souvent $X \sim \mathscr{U}([\![1,n]\!])$ ou $X \hookrightarrow \mathscr{U}([\![1,n]\!])$.

Remarque : Le signe $\sim$ n'a rien à voir avec les équivalents en analyse.

Propriété 3.5

Si X suit une loi uniforme dans $[\![1,n]\!]$,



Loi de probabilité

Propriété 3.6

Si X une loi uniforme sur $[\![1,n]\!]$, alors

$$\mathbf{E}(X) = \frac{n+1}{2}.$$

Preuve

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

Exemple *(Variance de la loi uniforme)*
Calculer $\mathbf{V}(X)$.

Solution :

$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \sum_{k=1}^{n} k^2 \mathbf{P}(X=k) - \mathbf{E}(X)^2 \\ &= \frac{n(n+1)(2n+1)}{6}\frac{1}{n} - \left(\frac{n+1}{2}\right)^2 \\ &= \frac{n+1}{2}\left(\frac{n(2n+1)}{3} - \frac{n+1}{2}\right) \\ &= \frac{n+1}{12}(2(2n+1) - 3(n+1)) \\ &= \frac{n+1}{12}(n-1) \\ &= \frac{n^2-1}{12}.\end{aligned}$$

C Loi de Bernoulli : tirage à deux issues

Exemple *(Exemple type : lancer à deux issues)*
Une pièce truquée possède p chances de tomber sur pile (succès) et $q = 1-p$ chances de tomber sur face (échec). On définit la variable aléatoire X qui teste si pile est sorti : X vaut 1 si on obtient pile, et 0 sinon.

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

Définition 3.7 *(Loi de Bernoulli)*

Soit X une variable aléatoire définie sur Ω à valeurs dans $\{0,1\}$.
Pour $p \in [0,1]$, on dit que X suit une loi de Bernoulli de paramètre p si

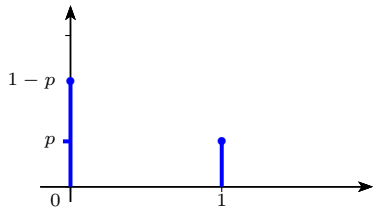
$$\mathbf{P}(X=1) = p \quad \text{et} \quad \mathbf{P}(X=0) = 1-p = q.$$

On note $X \sim \mathscr{B}(p)$ ou $X \hookrightarrow \mathscr{B}(p)$.

Remarque : En général $X(\Omega) = \{0,1\}$, mais si $p=0$, alors $X(\Omega) = \{0\}$ ou si $p=1$, alors $X(\Omega) = 1$. Ce sont des cas critiques qui correspondent à la loi certaine.
Certains ouvrages excluent ces situations de la définition de la loi de Bernoulli.

Propriété 3.8

Pour une loi de Bernoulli de paramètre p,



Loi de probabilité

Propriété 3.9

Soit X une variable aléatoire de loi $\mathscr{B}(p)$,

$$\mathbf{E}(X) = p, \qquad \mathbf{V}(X) = p(1-p) = pq, \qquad \sigma_X = \sqrt{p(1-p)} = \sqrt{pq}.$$

Preuve
$\mathbf{E}(X) = 0 \times \mathbf{P}(X=0) + 1 \times \mathbf{P}(X=1) = p.$
$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = 0^2 \times \mathbf{P}(X=0) + 1^2 \times \mathbf{P}(X=1) - p^2 = p(1-p).$ ■

Exemple
La loi $\mathbf{1}_A$ suit une loi de Bernoulli de paramètre $p = \mathbf{P}(A)$.

D Loi binomiale : tirages avec remise

Exemple *(Exemple type 1 : tirages avec remise)*
On dispose d'une urne avec N boules, dont R sont rouges et $N-R$ sont jaunes.
On effectue n tirages successifs avec remise, la probabilité d'obtenir exactement k boules rouges parmi les n tirages est

$$\mathbf{P}(X=k) = \binom{n}{k}\frac{R^k(N-R)^{n-k}}{N^n}.$$

Exemple *(Exemple type 2 : répétitions indépendantes d'une épreuve de Bernoulli)*
On reprend la pièce de l'épreuve de Bernoulli, et on effectue n lancers indépendants.
On définit la variable aléatoire X qui compte le nombre de fois que l'on a obtenu pile (nombre de succès).

$$\mathbf{P}(X=k) = \binom{n}{k}p^k(1-p)^{n-k}.$$

Définition 3.10 *(Loi binomiale)*

Soit X une variable aléatoire définie sur Ω à valeurs dans $[\![0,n]\!]$,
Pour $p \in [0,1]$, on dit que X suit une loi binomiale de paramètres n et p si

$$\forall k \in [\![0;n]\!],\ \mathbf{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On note $X \sim \mathscr{B}(n,p)$ ou $X \hookrightarrow \mathscr{B}(n,p)$.

Remarque : D'après la formule du binôme de Newton, on retrouve bien

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

La remarque vue pour la loi de Bernoulli à propos du paramètre p est aussi valable pour la loi binomiale.

Propriété 3.11 *(Interprétation de la loi binomiale)*

Si X est la variable aléatoire égale au nombre de succès obtenus par la répétition indépendante de n expériences de Bernoulli de même paramètre $p \in [0,1]$,
alors, X suit une loi binomiale de paramètres n, p : $X \sim \mathscr{B}(n,p)$.

Autrement dit :
Soit $n \in \mathbf{N}^*$, et $\forall i \in [\![1,n]\!]$, $X_i \sim \mathscr{B}(p)$ (indépendantes entre elles). Si on pose $X = \sum_{i=1}^{n} X_i$, alors

$$X \sim \mathscr{B}(n,p).$$

Preuve

- preuve « **intuitive** » : Soient n et k fixés.
  Les tirages successifs s'écrivent comme un n-uplet, composé de succès S et d'échecs $\overline{S}$.
  Un événement élémentaire qui vérifie $X = k$, peut donc s'écrire :

$$\underbrace{\underline{S}\ \underline{S}\ \underline{\overline{S}}\ \underline{S}\ \underline{\overline{S}}\ \underline{\overline{S}}\ \underline{S}\ \cdots\ \underline{\overline{S}}}_{n \text{ tirages dont } k \text{ succès } S \text{ et } n-k \text{ échecs } \overline{S}}$$

  On note S_i : « le $i^{\text{ème}}$ tirage donne un succès ».
  La probabilité de l'événement donné ci-dessus en exemple s'écrit :

$$\mathbf{P}\left(S_1 \cap S_2 \cap \overline{S_3} \cap \cdots \cap \overline{S_n}\right).$$

  Or les événements sont indépendants entre eux, donc

$$\mathbf{P}\left(S_1 \cap S_2 \cap \overline{S_3} \cap \cdots \cap \overline{S_n}\right) = \mathbf{P}(S_1)\ \mathbf{P}(S_2)\ \mathbf{P}\left(\overline{S_3}\right) \cdots \mathbf{P}\left(\overline{S_n}\right).$$

  Et $\forall i \in [\![1,n]\!]$, $\mathbf{P}(S_i) = p$,
  $\mathbf{P}\left(\overline{S_i}\right) = 1-p$.

  Dans le produit, il y a k succès et $n-k$ échecs, la probabilité précédente s'écrit donc

$$\mathbf{P}\left(S_1 \cap S_2 \cap \overline{S_3} \cap \cdots \cap \overline{S_n}\right) = p^k (1-p)^{n-k}.$$

  Or $[X=k]$ correspond à l'union de tous les « événements anagrammes » du précédent. Ces événements sont incompatibles, donc la probabilité de l'union est égale à la somme des probabilités. Ils ont tous la même probabilité $p^k(1-p)^{n-k}$ et il y en a au total $\binom{n}{k}$. Donc

$$\mathbf{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

- preuve « **formelle** » : Avec les notations précédentes.
  On note $\mathcal{P}_k$ l'ensemble des parties à k éléments de $[\![1,n]\!]$.

$$[X=k] = \bigsqcup_{I \in \mathcal{P}_k} \left[\left(\bigcap_{i \in I} S_i \right) \cap \left(\bigcap_{j \in [\![1,n]\!] \setminus I} \overline{S_j} \right) \right].$$

  Cette union est disjointe (événements incompatibles) car les $I \in \mathcal{P}_k$ sont supposés deux à deux différents.

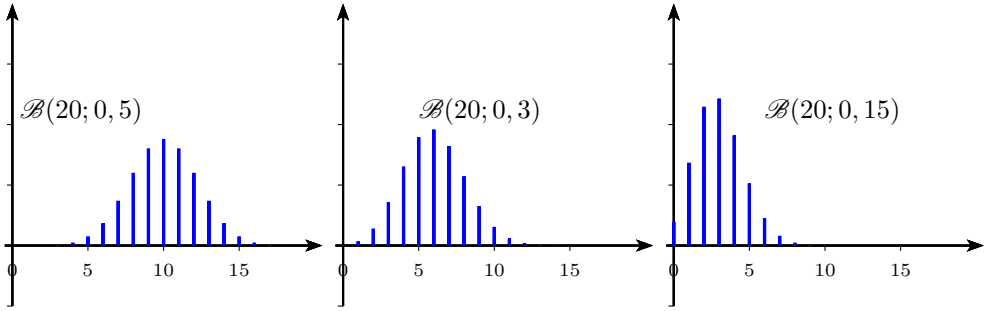
$$\text{Donc } \mathbf{P}([X=k]) = \sum_{I \in \mathcal{P}_k} \mathbf{P}\left[\left(\bigcap_{i \in I} S_i \right) \cap \left(\bigcap_{j \in [\![1,n]\!] \setminus I} \overline{S_j} \right) \right].$$

  Or les événements $(S_i)_{i \in [\![1,n]\!]}$ sont indépendants, donc

$$\begin{aligned}
\mathbf{P}([X=k]) &= \sum_{I \in \mathcal{P}_k} \left(\prod_{i \in I} \mathbf{P}(S_i) \right) \left(\prod_{j \in [\![1,n]\!] \setminus I} \mathbf{P}\left(\overline{S_j}\right) \right) \\
&= \sum_{I \in \mathcal{P}_k} \left(\prod_{i \in I} p \right) \left(\prod_{j \in [\![1,n]\!] \setminus I} (1-p) \right) \\
&= \sum_{I \in \mathcal{P}_k} p^k (1-p)^{n)-k} \qquad (\text{car Card}\, I = k \text{ et Card}\, ([\![1,n]\!] \setminus I) = n-k) \\
&= \binom{n}{k} p^k (1-p)^{n)-k} \qquad (\text{car Card}\, \mathcal{P}_k = \binom{n}{k}).
\end{aligned}$$

■

Lois de probabilités en fonction du paramètre p.



On remarque que lorsque $p = 0,5$ les probabilités sont symétriques par rapport à la moyenne, mais pour $p \neq 0,5$, il existe une asymétrie d'autant plus marquée que p est loin de $0,5$.
On remarque également que la distribution ressemble à une gaussienne (courbe en cloche).

Lorsque n tend vers $+\infty$, la loi binomiale tend vers la loi normale (le théorème central limite nous dit que c'est le cas de toutes les lois construites sur ce modèle : la répétition de n expériences indépendantes qui suivent la même loi).

Propriété 3.12

Soit X une variable aléatoire de loi $\mathscr{B}(n,p)$,

$$\mathbf{E}(X) = np, \qquad \mathbf{V}(X) = np(1-p) = npq, \qquad \sigma_X = \sqrt{np(1-p)} = \sqrt{npq}.$$

Preuve

$$\begin{aligned}
\mathbf{E}(X) &= \sum_{k=0}^{n} k\mathbf{P}(X=k) \\
&= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} n\binom{n-1}{k-1}p^k(1-p)^{n-k} \text{ formule sans nom} \\
&= np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-1-k} \text{ changement d'indice} \\
&= np \text{ binôme de Newton.}
\end{aligned}$$

Pour la variance, d'après Kœnig-Huygens,

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

$$\begin{aligned}
\mathbf{E}\left(X^2\right) &= \sum_{k=0}^{n} k^2\mathbf{P}(X=k) = \sum_{k=0}^{n} k^2\binom{n}{k}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} nk\binom{n-1}{k-1}p^k(1-p)^{n-k} \qquad \text{(formule sans nom)} \\
&= \sum_{k=0}^{n-1} n(k+1)\binom{n-1}{k}p^{k+1}(1-p)^{n-k-1} \qquad \text{(changement d'indice)} \\
&= \sum_{k=0}^{n-1} nk\binom{n-1}{k}p^{k+1}(1-p)^{n-k-1} + n\binom{n-1}{k}p^{k+1}(1-p)^{n-k-1} \\
&= np\sum_{k=0}^{n-1} k\binom{n-1}{k}p^k(1-p)^{n-k-1} + np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-k-1} \\
&= np\sum_{k=1}^{n-1} k\binom{n-1}{k}p^k(1-p)^{n-k-1} + np\,1 \qquad \text{(binôme de Newton)} \\
&= np\sum_{k=1}^{n-1}(n-1)\binom{n-2}{k-1}p^k(1-p)^{n-k-1} + np\,1 \qquad \text{(formule sans nom)} \\
&= n(n-1)p\sum_{k=0}^{n-2}\binom{n-2}{k}p^{k+1}(1-p)^{n-2-k} + np\times 1 \qquad \text{(changement d'indice)}
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}\left(X^2\right) &= n(n-1)p^2\sum_{k=0}^{n-2}\binom{n-2}{k}p^k(1-p)^{n-2-k} + np \\
&= n(n-1)p^2 + np \qquad \text{(binôme de Newton)} \\
&= np(np-p+1).
\end{aligned}$$

Donc

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

Remarque : À la fin du chapitre, on pourra voir que la somme de n variables aléatoires de Bernoulli, indépendantes et de même paramètre p donne une loi binomiale de paramètres n, p.
La linéarité de l'espérance donne alors directement que $\mathbf{E}(X) = np$ *et l'égalité de Bienaymé qui sera aussi vue plus loin, donne également* $\mathbf{V}(X) = npq$. ■

4 Inégalité de Bienaymé-Tchebychev

Remarque : Tchebychev est une translittération, on pourra trouver d'autres orthographes comme Chebyshev.
Cette section est valable uniquement pour les variables réelles.

Propriété 4.1 *(Inégalité de Markov)*

Soient X une variable aléatoire réelle **positive** et $a > 0$,

$$\mathbf{P}(X \geqslant a) \leqslant \frac{\mathbf{E}(X)}{a}.$$

Preuve

$$\begin{aligned}\mathbf{E}(X) &= \sum_{x\in X(\Omega)} x\mathbf{P}(X=x) = \sum_{x<a} x\mathbf{P}(X=x) + \sum_{x\geqslant a} x\mathbf{P}(X=x)\\ &\geqslant \sum_{x\geqslant a} x\mathbf{P}(X=x) \qquad \text{car } X \text{ est à valeurs positives.}\\ &\geqslant \sum_{x\geqslant a} a\mathbf{P}(X=x)\\ &\geqslant a\mathbf{P}(X\geqslant a).\end{aligned}$$

Donc par passage au quotient ($a>0$), on obtient l'inégalité voulue.

Autre preuve (plus formelle, mais plus générale aussi).

La fonction $a\mathbf{1}_{[a,+\infty[}$ vaut : 0 pour $x<a$,
a pour $x\geqslant a$.

On a donc $\forall x\geqslant 0,\ x\geqslant a\mathbf{1}_{[a,+\infty[}(x)$.

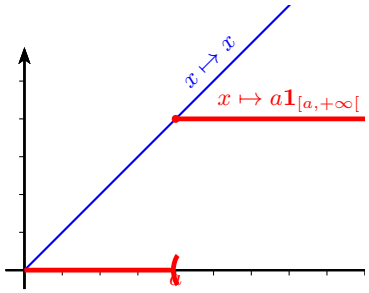
Or X est positive, donc on peut composer *à droite* par X.

$$\forall\omega\in\Omega, X(\omega)\geqslant a\mathbf{1}_{[a;+\infty[}(X(\omega)).$$

C'est-à-dire $X\geqslant a\mathbf{1}_{[a,+\infty[}(X)$.

On passe à l'espérance qui est croissante :

$$\mathbf{E}(X)\geqslant a\mathbf{E}\left(\mathbf{1}_{[a;+\infty[}(X)\right)=a\mathbf{P}(X\geqslant a).$$



■

Théorème 4.2 *(Inégalité de Bienaymé-Tchebychev)*

Soient X une variable aléatoire réelle et $a>0$,

$$\mathbf{P}\left(\left|X-\mathbf{E}(X)\right|\geqslant a\right)\leqslant\frac{\sigma_X^2}{a^2}.$$

Explications

Cette inégalité offre une façon d'interpréter la dispersion d'une variable aléatoire : les résultats sont d'autant plus concentrés autour de l'espérance que l'écart type est faible : la probabilité d'avoir un résultat éloigné de l'espérance décroît avec l'écart type.

La distance à l'espérance est traduite par $|X-\mathbf{E}(X)|\geqslant a$. La probabilité d'avoir la valeur de X éloignée de plus de a de l'espérance est majoré par $\left(\frac{\sigma_X}{a}\right)^2$.

Preuve

$$|X-\mathbf{E}(X)|\geqslant a \iff (X-\mathbf{E}(X))^2\geqslant a^2.$$

Donc $\mathbf{P}(|X-\mathbf{E}(X)|\geqslant a)=\mathbf{P}\left((X-\mathbf{E}(X))^2\geqslant a^2\right)$.

On applique ensuite l'inégalité de Markov à $X-\mathbf{E}(X)$ (on centre la variable). ■

Corollaire 4.3

Soit X une variable aléatoire réelle, soit $a>0$,

$$\mathbf{P}(|X-\mathbf{E}(X)|<a)\geqslant 1-\frac{\sigma_X^2}{a^2}.$$

Exemple

Donner un majorant de la probabilité de s'éloigner de plus de $10\sigma_X$ de la moyenne pour une épreuve donnée.

Solution :

L'inégalité de Bienaymé-Tchebychev donne la majoration :

$$\mathbf{P}(|X-\mathbf{E}(X)|\geqslant 10\sigma_X)\leqslant\frac{\sigma_X^2}{(10\sigma_X)^2}=1\%.$$

Exemple

Un tireur a une probabilité de 70% de toucher sa cible à chaque coup. Chaque coup de fusil est indépendant des précédents.

On note X le nombre de fois qu'il a atteint sa cible.

À partir de combien de séances de tir peut-on affirmer que la probabilité d'avoir taux de succès entre 60% et 80% est supérieur à 90% ?

Solution :

On note X_n la variable aléatoire qui correspond au nombre de succès pour n coups de fusils.

Si on pose $p=0,7$, alors $X_n\sim\mathscr{B}(n,p)$ et $\mathbf{E}(X_n)=np$, $\mathbf{V}(X_n)=np(1-p)$.

On cherche donc $\mathbf{P}\left(0,6\leqslant\frac{X_n}{n}\leqslant 0,8\right)=\mathbf{P}\left(\left|\frac{X_n}{n}-0,7\right|\leqslant 0,1\right)$

$$\begin{aligned}&=\mathbf{P}(|X_n-0,7n|\leqslant 0,1n)\\ &=\mathbf{P}(|X_n-\mathbf{E}(X_n)|\leqslant 0,1n)\\ &\geqslant\mathbf{P}(|X_n-\mathbf{E}(X_n)|<0,1n)\\ &\geqslant 1-\frac{\mathbf{V}(X_n)}{(0,1n)^2} \qquad \text{Bienaymé-Tchebychev}\\ &\geqslant 1-\frac{np(1-p)}{(0,1n)^2}\\ &\geqslant 1-\frac{0,7\times 0,3}{0,1^2n}\\ &\geqslant 1-\frac{21}{n}.\end{aligned}$$

On veut donc que $1-\frac{21}{n}\geqslant 0,9$, c'est-à-dire $n\geqslant 210$.

À partir de 210 coups tirés, la probabilité qu'il ait atteint la cible entre 60% et 80% des fois est supérieure à 90%.

Exercice

Déduire des propriétés précédentes que

1. Pour $a>0$ et pour une variable aléatoire réelle X positive, $\mathbf{P}(X>a)\leqslant\frac{\mathbf{E}(X)}{a}$.

2. Pour $a>0$ et pour une variable aléatoire réelle X, $\mathbf{P}\left(|X-\mathbf{E}\left(X\right)|>a\right)\leqslant\dfrac{\sigma_X^2}{a^2}$.

3. Pour $a>0$ et pour une variable aléatoire réelle X, $\mathbf{P}\left(|X-\mathbf{E}\left(X\right)|\leqslant a\right)\geqslant 1-\dfrac{\sigma_X^2}{a^2}$.

5 Couples de variables aléatoires

A Lois conjointes et marginales

Définition 5.1 *(Loi conjointe)*

Soient $X:\Omega\to E$ et $Y:\Omega\to E'$ deux variables aléatoires.
$Z=(X,Y)$ est une nouvelle variable aléatoire de Ω dans $E\times E'$.
La loi $\mathbf{P}_Z$ du couple $Z=(X,Y)$ est appelé loi conjointe de X et Y.
On définit alors

$$\mathbf{P}\left((X,Y)=(x,y)\right)=\mathbf{P}\left(X=x\text{ et }Y=y\right).$$

Cela revient à chercher la réalisation simultanée (conjointe) de $X=x$ **et** de $Y=y$.

Exemple

On lance deux dés et on note X la valeur donnée par le premier dé et Y la valeur donnée par le second.
Si les dés sont équilibrés alors, pour tout $(x,y)\in[\![1,6]\!]^2$,

$$\mathbf{P}\left((X,Y)=(x,y)\right)=\frac{1}{36}.$$

Exemple

On tire successivement et sans remise deux boules dans une urne qui contient 5 boules rouges et 3 bleues. On note X_1, X_2 les variables aléatoires qui valent 1 si la boule est rouge, et 0 sinon respectivement au premier et au deuxième tirage.
On peut créer le tableau des probabilités conjointes :

		X_2	
		0	1
X_1	0	$\frac{6}{56}$	$\frac{15}{56}$
	1	$\frac{15}{56}$	$\frac{20}{56}$

On vérifie que la somme des probabilités est égale à 1.

Exemple

On lance deux dés équilibrés et on note X le minimum et Y le maximum.

		Y					
		1	2	3	4	5	6
X	1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	2		$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	3			$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	4				$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	5					$\frac{1}{36}$	$\frac{2}{36}$
	6						$\frac{1}{36}$

On vérifie que la somme des probabilités est égale à 1.

Dans cet exemple, on voit que l'univers image $(X,Y)(\Omega)$ n'est pas à confondre avec l'ensemble $X(\Omega)\times Y(\Omega)$ dont il n'est qu'une partie. Comme $X(\Omega)\times Y(\Omega)$ est plus simple à calculer, on travaillera souvent sur cet univers par abus quitte à donner la probabilité 0 aux événements « *en trop* ». Mais il faut garder en tête cette distinction. Ici, l'univers image $(X,Y)(\Omega)$ est donné par l'ensemble des couples $\left\{(x,y)\in[\![1,6]\!]^2,\ x\leqslant y\right\}$, qui est représenté par un triangle dans le tableau. Il est différent de l'univers $X(\Omega)\times Y(\Omega)=\left\{(x,y)\in[\![1,6]\!]^2\right\}$ représenté par le tableau entier.

Définition 5.2 *(Loi marginale)*

Soient $X:\Omega\to E$ et $Y:\Omega\to E'$ deux variables aléatoires.
La loi $\mathbf{P}_X$ de la variable X est la première loi marginale de (X,Y),
et $\mathbf{P}_Y$ est la seconde loi marginale.

Explications

La loi marginale suivant X revient à étudier la probabilité de réalisation de X quelle que soit la valeur prise par Y.
Dans le cas des probabilités discrètes on peut écrire les lois sous la forme de tableaux

à deux entrées :

		$\boldsymbol{Y}$						
		y_1	y_2	$\cdots$	y_j	$\cdots$	y_q	Σ
$\boldsymbol{X}$	x_1	$p_{1,1}$	$p_{1,2}$	$\cdots$	$p_{1,j}$	$\cdots$	$p_{1,q}$	$P_X(x_1)$
	x_2	$p_{2,1}$	$p_{2,2}$	$\cdots$	$p_{2,j}$	$\cdots$	$p_{2,q}$	$P_X(x_2)$
	$\vdots$				$\vdots$			$\vdots$
	x_i	$p_{i,1}$	$p_{i,2}$	$\cdots$	$p_{i,j}$	$\cdots$	$p_{i,q}$	$P_X(x_i)$
	$\vdots$				$\vdots$			$\vdots$
	x_p	$p_{p,1}$	$p_{p,2}$	$\cdots$	$p_{p,j}$	$\cdots$	$p_{p,q}$	$P_X(x_p)$
	Σ	$P_Y(y_1)$	$P_Y(y_2)$	$\cdots$	$P_Y(y_j)$	$\cdots$	$P_Y(y_q)$	$\to 1$
		$\downarrow$						
		1						

Pour obtenir $p_i = \mathbf{P}(X = x_i) = \mathbf{P}_X(x_i)$, on fait la somme de la ligne x_i.
Pour connaître $p'_j = \mathbf{P}(Y = y_i) = \mathbf{P}_Y(y_j)$, on somme suivant la colonne y_j.
Cela se résume par la formule suivante :

Propriété 5.3

La loi conjointe détermine les lois marginales :

$$\mathbf{P}_X(x) = \sum_{y \in Y(\Omega)} \mathbf{P}\left(X = x \text{ et } Y = y\right).$$

$$\mathbf{P}_Y(y) = \sum_{x \in X(\Omega)} \mathbf{P}\left(X = x \text{ et } Y = y\right).$$

Preuve

C'est la formule des probabilités totales, simplement parce que

$$\begin{aligned}\mathbf{P}_X(x) &= \mathbf{P}\left([X = x] \bigcap \left(\bigsqcup_{y \in Y(\Omega)} [Y = y]\right)\right)\\ &= \mathbf{P}\left(\bigsqcup_{y \in Y(\Omega)} ([X = x] \cap [Y = y])\right)\\ &= \sum_{y \in Y(\Omega)} \mathbf{P}\left([X = x] \cap [Y = y]\right) \quad \text{événements incompatibles.}\end{aligned}$$

■

Remarque : $\left\{\{X = x \text{ et } Y = y\},\ (x, y) \in (X, Y)(\Omega)\right\}$ forme un système complet d'événements.

Exemple

Pour le lancer de deux dés équilibrés : X traduit le résultat du premier lancer, et Y le résultat du second.

$$\mathbf{P}_X(4) = \mathbf{P}(X = 4) = \sum_{k=1}^{6} \mathbf{P}(X = 4 \cap Y = k) = 6 \times \frac{1}{36} = \frac{1}{6}.$$

Exemple

Pour l'exemple du lancer de deux dés avec X le minimum et Y le maximum.

		Y						
		1	2	3	4	5	6	Σ
X	1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{11}{36}$
	2		$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
	3			$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{7}{36}$
	4				$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{5}{36}$
	5					$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$
	6						$\frac{1}{36}$	$\frac{1}{36}$
	Σ	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	

$\rightarrow 1$

$\downarrow$

1

⚠ En général, les lois marginales ne permettent pas de définir la loi conjointe : il ne suffit pas de connaître la dernière ligne et la dernière colonne du tableau pour pouvoir le remplir tout entier !

Exemple

On tire successivement deux boules dans une urne avec R rouges et $N-R$ noires. X vaut 1 si la première boule tirée est rouge, Y vaut 1 si la seconde boule tirée est rouge, 0 sinon.
Donner la loi de Y.
Solution :

$$\begin{aligned}\mathbf{P}_Y(1) = \mathbf{P}(Y=1) &= \mathbf{P}(X=1 \cap Y=1) + \mathbf{P}(X=0 \cap Y=1)\\ &= \mathbf{P}_{X=1}(Y=1)\mathbf{P}(X=1) + \mathbf{P}_{X=0}(Y=1)\mathbf{P}(X=0)\\ &= \frac{R}{N}\times\frac{R-1}{N-1} + \frac{N-R}{N}\times\frac{R}{N-1} = \frac{R(R-1)+R(N-R)}{N(N-1)}\\ &= \frac{R}{N}.\end{aligned}$$

Propriété 5.4 *(Loi image d'un couple)*

Soient $X : \Omega \to E$ et $Y : \Omega \to E'$ deux variables aléatoires.
Soit $u : (X,Y)(\Omega) \to F$. La loi $\mathbf{P}_u$ est définie par

$$\mathbf{P}_u\left(u(X,Y)=t\right) = \sum_{\substack{(x,y)\in(X,Y)(\Omega)\\ u(x,y)=t}} \mathbf{P}\left(X=x \text{ et } Y=y\right).$$

On remarque que $\mathbf{P}$ est une probabilité sur $(X,Y)(\Omega)$ (fini) si elle est à valeurs positive et si $\sum_{(x,y)\in(X,Y)(\Omega)} \mathbf{P}(X=x, Y=y) = 1$.

Exemple *(Loi image d'un couple)*

Soit le tirage de deux dés. X est la variable aléatoire correspondant au résultat du premier dé, et Y au résultat du second.
La loi de la somme est donnée pour $u(X,Y) = X+Y$.

$$\mathbf{P}(X+Y=k) = \sum_{i=1}^{k-1} \mathbf{P}\big([X=i]\cap[Y=k-i]\big).$$

B Espérance

Théorème 5.5 *(Théorème de transfert)*

Soient X et Y deux variables aléatoires finies sur Ω, et u une application définie sur $(X,Y)(\Omega)$ dans $\mathbf{R}$.

$$\mathbf{E}\left(u(X,Y)\right) = \sum_{(x,y)\in(X,Y)(\Omega)} u(x,y)\mathbf{P}(X=x, Y=y).$$

Exemple

Calculer l'espérance de la somme de deux variables aléatoires X et Y en utilisant le théorème de transfert pour retrouver la linéarité.
Solution :

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

On retrouve la formule de la linéarité de l'espérance.

C Covariance

Définition 5.6 *(Covariance)*

Soient X et Y deux variables aléatoires finies sur Ω, on définit la covariance de X, Y par

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

Si on prend $Y = X$, on retrouve la formule de la variance de X : $\mathbf{Cov}(X,X) = \mathbf{V}(X)$. Comme pour la variance, on utilisera plutôt la formule de Kœnig-Huygens pour calculer la covariance.

Propriété 5.7 *(Formule de Kœnig-Huygens)*

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

Propriété 5.8

Pour X, Y, X_1 et X_2 des variables aléatoires sur Ω, et $(a,b) \in \mathbf{R}^2$,

1. $\mathbf{Cov}(X,Y) = \mathbf{Cov}(Y,X)$. (symétrie)
2. $\mathbf{Cov}(aX+b,Y) = a\,\mathbf{Cov}(X,Y)$.
3. $\mathbf{Cov}(aX_1+X_2,Y) = a\,\mathbf{Cov}(X_1,Y) + \mathbf{Cov}(X_2,Y)$. (linéarité)
4. $\mathbf{Cov}(X,X) = \mathbf{V}(X) \geqslant 0$. (positive)

Propriété 5.9 *(Variance d'une somme)*

Soient X et Y deux variables aléatoires finies sur Ω,

$$\mathbf{V}(X+Y) = \mathbf{V}(X) + \mathbf{V}(Y) + 2\,\mathbf{Cov}(X,Y).$$

Preuve

$$\begin{aligned}\mathbf{V}(X+Y) &= \mathbf{E}\left((X+Y)^2\right) - \mathbf{E}\left(X+Y\right)^2\\ &= \mathbf{E}\left(X^2+Y^2+2XY\right) - \left(\mathbf{E}\left(X\right)+\mathbf{E}\left(Y\right)\right)^2\\ &= \mathbf{E}\left(X^2\right)+\mathbf{E}\left(Y^2\right)+2\mathbf{E}\left(XY\right) - \mathbf{E}\left(X\right)^2 - \mathbf{E}\left(Y\right)^2 - 2\mathbf{E}(X)\mathbf{E}(Y)\\ &= \mathbf{V}(X)+\mathbf{V}(Y)+2\mathbf{Cov}(X,Y).\end{aligned}$$

■

6 Indépendance des variables aléatoires

A Loi conditionnelle

Définition 6.1 *(Loi conditionnelle)*

Pour (X,Y) deux variables aléatoires finies sur (Ω), et on définit la loi conditionnelle de X sachant $(Y=y)$ par

$$f_{X|Y=y} : \begin{cases} X(\Omega) & \to [0,1] \\ x & \mapsto \mathbf{P}_{[Y=y]}(X=x) = \mathbf{P}\left(X=x|Y=y\right) = \dfrac{\mathbf{P}\left(X=x,Y=y\right)}{\mathbf{P}(Y=y)}. \end{cases}$$

On parle aussi de probabilité conditionnelle de X sachant $[Y=y]$.

Remarque : Cette définition suppose que $\mathbf{P}(Y=y) \neq 0$.

Travailler sur une loi conditionnelle ne modifie pas les variables aléatoires, mais seulement la loi de probabilité. Cela illustre que les mêmes variables aléatoires peuvent être étudiées avec des lois de probabilités différentes.

Lorsque nous ne connaissons que les lois marginales, nous ne pouvons pas retrouver la loi conjointe. Par contre, si nous connaissons à la fois les lois marginales et les lois conditionnelles par rapport à une variable, alors il est possible de reconstruire la loi conjointe. En effet

$$\mathbf{P}\left((X,Y)=(x,y)\right) = \mathbf{P}\left([X=x]\cap[Y=y]\right) = \mathbf{P}_{[Y=y]}\left(X=x\right)\mathbf{P}\left(Y=y\right)$$

Exemple

Pour le lancer de deux dés, on peut considérer la probabilité que le plus grand dé soit y sachant que le plus petit est x. Si les dés sont équilibrés, on obtient la loi de probabilité conditionnelle.
Par exemple pour $x=2$:

Y	1	2	3	4	5	6
$\mathbf{P}_{[X=2]}(Y=y)$	0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$

Cette probabilité conditionnelle se construit facilement à partir du tableau des probabilités conjointes et marginales. Il suffit de sélectionner la bonne ligne $X=2$ et de diviser chaque probabilité conjointe par la probabilité marginale (somme des probabilités conjointes de la ligne).

B Variables aléatoires indépendantes

Définition 6.2 *(Variables aléatoires indépendantes)*

Deux variables aléatoires X et Y définies sur Ω sont dites **indépendantes**, si

$$\forall (x,y) \in X(\Omega) \times Y(\Omega), \quad \mathbf{P}\left([X=x] \cap [Y=y]\right) = \mathbf{P}(X=x)\mathbf{P}(Y=y).$$

On note $X \perp\!\!\!\perp Y$.

Remarque : On peut voir que pour des variables aléatoires indépendantes on a nécessairement $(X,Y)(\Omega) = X(\Omega) \times Y(\Omega)$.

Propriété 6.3

Si les événements $[X=x]$ et $[Y=y]$ sont tous de probabilité non nulle pour $(x,y) \in X(\Omega) \times Y(\Omega)$, alors, l'indépendance de X et Y est aussi caractérisée par :

$$\forall (x,y) \in X(\Omega) \times Y(\Omega), \quad \mathbf{P}_{[Y=y]}(X=x) = \mathbf{P}(X=x).$$

ou

$$\forall (x,y) \in X(\Omega) \times Y(\Omega), \quad \mathbf{P}_{[X=x]}(Y=y) = \mathbf{P}(Y=y).$$

Explications

Il est sous-entendu que les variables sont indépendantes pour la probabilité $\mathbf{P}$. L'indépendance dépend du choix de probabilité même si ce n'est pas dit explicitement.

Exemple

Si X et Y désignent respectivement le minimum et le maximum lors du lancer de deux dés équilibrés, alors ces variables aléatoires ne sont pas indépendantes. Par exemple $\mathbf{P}(X=6, Y=5) = 0$, mais $\mathbf{P}(X=6)\mathbf{P}(Y=5) = \frac{1}{36}\frac{9}{39} \neq 0$.

Exemple

Si X et Y désignent respectivement la valeur du premier dé et la valeur du second dé lors du lancé de deux dés équilibrés, alors ces variables aléatoires sont indépendantes.

Exemple *(Variables de Bernoulli indépendantes)*

Montrer que deux variables de Bernoulli X_1 et X_2 sont indépendantes si et seulement si les événements $[X_1=1]$ et $[X_2=1]$ sont indépendants.

Solution :

Si A et B sont indépendants, alors $\overline{A}$ et B aussi, A et $\overline{B}$ aussi et $\overline{A}$ et $\overline{B}$ aussi (en exercice).

Dans le cas des variables de Bernoulli, comme l'univers image se réduit à $[X=1]$ et $[X=0] = \overline{[X=1]}$, la vérification proposée est donc suffisante.

Théorème 6.4

Soient deux variables aléatoires **indépendantes** X et Y définies sur Ω,

$$\forall (A,B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega)), \quad \mathbf{P}(X \in A\,,\, Y \in B) = \mathbf{P}(X \in A)\mathbf{P}(Y \in B).$$

Preuve

On écrit A et B comme union d'événements élémentaires.

$$A = \{x_i\}_{i\in I} \text{ et } B = \{y_j\}_{j\in J}\,.$$

$$\begin{aligned}
\mathbf{P}(X \in A\,,\, Y \in B) &= \mathbf{P}\left(\bigcup_{(i,j)\in I\times J} ([X=x_i] \cap [Y=y_j])\right) \\
&= \sum_{(i,j)\in I\times J} \mathbf{P}\left([X=x_i] \cap [Y=y_j]\right) \\
&= \sum_{(i,j)\in I\times J} \mathbf{P}\left([X=x_i]\right)\mathbf{P}\left([Y=y_j]\right) \quad \text{indépendance} \\
&= \sum_{i\in I} \mathbf{P}\left([X=x_i]\right) \sum_{j\in J} \mathbf{P}\left([Y=y_j]\right) \quad \text{indices séparables} \\
&= \mathbf{P}(X \in A)\mathbf{P}(Y \in B).
\end{aligned}$$

■

Il existe aussi une réciproque :
Si pour *toutes* parties A de $X(\Omega)$ et B de $Y(\Omega)$, on a

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\mathbf{P}(Y \in B)$$

alors les variables aléatoires sont indépendantes pour $\mathbf{P}$.
Pour le montrer, il suffit de particulariser avec des événements singleton.

Théorème 6.5 *(L'indépendance passe à l'image)*

Soient deux variables aléatoires indépendantes $X : \Omega \to E$ et $Y : \Omega \to E'$.
Pour toutes les applications $f : E \to F$ et $g : E' \to F'$,

$$f(X) \text{ et } g(Y) \text{ sont indépendantes.}$$

Explications

Supposons que l'on associe un gain à chacune des variables aléatoires (selon la valeur qu'elles prennent), si les variables sont indépendantes, alors les gains le sont aussi.

Preuve

$\mathbf{P}(f(X)=t) = \mathbf{P}(X \in f^{-1}(\{t\}))$ et $\mathbf{P}(g(Y)=u) = \mathbf{P}(X \in g^{-1}(\{u\}))$.
Donc en appliquant le théorème précédent :

$$\begin{aligned}
\mathbf{P}([f(X)=t] \cap [g(Y)=u]) &= \mathbf{P}\left(X \in f^{-1}(\{t\})\right)\mathbf{P}\left(Y \in g^{-1}(\{u\})\right) \\
&= \mathbf{P}\left(f(X)=t\right)\mathbf{P}\left(g(Y)=u\right).
\end{aligned}$$

■

Exemple

Si X et Y sont indépendantes, alors X^2 est indépendante avec Y.

Théorème 6.6 *(Espérance du produit)*

Soient deux variables aléatoires X et Y.

$$X \perp\!\!\!\perp_{\mathbf{P}} Y \Rightarrow \mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y).$$

⚠ La réciproque est **fausse**, il ne suffit pas d'avoir $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ pour que les variables aléatoires soient indépendantes. C'est une erreur fréquente.

Preuve

On note $X(\Omega) = \{x_i\}_{i\in[\![1,p]\!]}$ et $Y(\Omega) = \{y_j\}_{j\in[\![1,q]\!]}$.

$$\begin{aligned}\mathbf{E}(XY) &= \sum_{i=1}^{p}\sum_{j=1}^{q} x_i y_j \mathbf{P}\left((X=x_i)\cap(Y=y_j)\right)\\ &= \sum_{i=1}^{p}\sum_{j=1}^{q} x_i y_j \mathbf{P}\left(X=x_i\right)\mathbf{P}\left(Y=y_j\right) \quad \text{(Variables indépendantes).}\\ &= \sum_{i=1}^{p} x_i \mathbf{P}\left(X=x_i\right)\sum_{j=1}^{q} y_j \mathbf{P}\left(Y=y_j\right) = \mathbf{E}(X)\mathbf{E}(Y).\end{aligned}$$

■

En particulier, deux variables aléatoires indépendantes ont une covariance nulle. Par contre, ce n'est pas parce que la covariance est nulle que les variables sont indépendantes.

Propriété 6.7

Si X, Y sont deux variables aléatoires indépendantes, alors $\mathbf{Cov}(X,Y) = 0$.

Preuve

Découle de la formule de Kœnig-Huygens. ■

Définition 6.8 *(Corrélation)*

Deux variables aléatoires X, Y sont dites **corrélées** si $\mathbf{Cov}\left(X,\,Y\right) \neq 0$.
Dans la cas contraire, elles sont dites **décorrélées.**

Si $\mathbf{Cov}\left(X,\,Y\right) > 0$, alors elles sont positivement corrélées.
Si $\mathbf{Cov}\left(X,\,Y\right) < 0$, alors elles sont négativement corrélées.

Propriété 6.9 *(Égalité de Bienaymé)*

Si X, Y sont deux variables aléatoires indépendantes, alors

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

Preuve

Car la covariance est nulle. ■

Exemple

Cette égalité permet de retrouver la variance d'une variable suivant une loi binomiale très simplement.

⚠ Comme précisé plus haut, l'indépendance implique la décorrélation, mais la réciproque est fausse.
En d'autres termes, deux variables corrélées ne peuvent pas être indépendantes, mais ce n'est pas parce que les variables sont corrélées qu'elles sont nécessairement indépendantes.

Exemple

On considère deux variables X et Y définies par la loi conjointe :

		Y		
		-1	0	1
X	-1	0	$\frac{1}{8}$	0
	0	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
	1	0	$\frac{1}{8}$	0

$XY = 0$ donc $\mathbf{E}\left(XY\right) = 0 = \mathbf{E}\left(X\right)\mathbf{E}\left(Y\right)$ et pourtant les variables aléatoires ne sont pas indépendantes car $\mathbf{P}\left((X,\,Y)=(1,\,1)\right) = 0 \neq \mathbf{P}\left(X=1\right)\mathbf{P}\left(Y=1\right)$.

7 Généralisation à n variables aléatoires

A Indépendance mutuelle

Définition 7.1 *(Variables mutuellement indépendantes)*

Soient $X_1, X_2, \cdots, X_n$ des variables sur un même univers Ω.
$(X_1, X_2, \cdots, X_n)$ sont **mutuellement indépendantes** si

$$\forall (x_i)_{i\in[\![1,n]\!]} \in \prod_{i=1}^{n} X_i\left(\Omega\right), \qquad \mathbf{P}\left(\bigcap_{i=1}^{n}[X_i = x_i]\right) = \prod_{i\in I}\mathbf{P}\left(X_i = x_i\right).$$

Propriété 7.2 *(Lien avec l'indépendance des événements)*

Soient n variables aléatoires mutuellement indépendantes $(X_i)_{i\in[\![1,n]\!]}$,
Pour $A_1, A_2, \cdots, A_n$ parties de $X_1(\Omega), X_2(\Omega), \cdots, X_n(\Omega)$,
les événements $(X_i \in A_i)$ sont mutuellement indépendants.

Remarque : La réciproque est évidemment vraie.

Preuve

On reproduit la même preuve que pour deux événements.
Soit I une sous-famille de $[\![1;n]\!]$.
Pour tout $i \in I$, on note $B_i = A_i$, et pour $i \notin I$, on pose $B_i = X_i(\Omega)$.
Ainsi $\bigcap_{i\in I}[X_i \in A_i] = \bigcap_{i=1}^n [X_i \in B_i]$.
Pour $i \in [\![1, n]\!]$ on écrit B_i comme réunion d'événements élémentaires.

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{i\in I}[X\in A_i]\right) &= \mathbf{P}\left(\bigcap_{i=1}^n [X\in B_i]\right)\\
&= \mathbf{P}\left(\bigcap_{i=1}^n \left(\bigsqcup_{x_i\in B_i}[X_i = x_i]\right)\right)\\
&= \mathbf{P}\left(\bigsqcup_{(x_i)_{i\in[1,n]}\in\prod_{i=1}^n B_i}\left(\bigcap_{i=1}^n [X_i=x_i]\right)\right) \quad \text{(distributivité)}\\
&= \sum_{(x_i)_{i\in[1,n]}\in\prod_{i=1}^n B_i} \mathbf{P}\left(\bigcap_{i=1}^n [X_i=x_i]\right) \quad \text{(union disjointe)}\\
&= \sum_{(x_i)_{i\in[1,n]}\in\prod_{i=1}^n B_i} \prod_{i=1}^n \mathbf{P}\left([X_i=x_i]\right) \quad \text{(indépendance)}
\end{aligned}$$

On refait à présent la distributivité dans l'autre sens, mais avec la somme et le produit au lieu de l'union et de l'intersection et on trouve

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{i\in I}[X\in A_i]\right) &= \prod_{i=1}^n \sum_{x_i\in B_i} \mathbf{P}\left([X_i=x_i]\right)\\
&= \prod_{i=1}^n \mathbf{P}\left(\bigsqcup_{x_i\in B_i}[X_i=x_i]\right) \quad \text{(union disjointe)}\\
&= \prod_{i=1}^n \mathbf{P}\left([X_i\in B_i]\right) = \prod_{i\in I}\mathbf{P}\left([X_i\in A_i]\right).
\end{aligned}$$

■

Propriété 7.3 *(Indépendance des sous-familles)*

Si $(X_1, X_2, \cdots, X_n)$ sont des variables mutuellement indépendantes, alors toute sous-famille l'est aussi.

Preuve

D'après la propriété précédente. ■

Théorème 7.4 *(L'indépendance passe à l'image)*

Si $(X_1, X_2, \cdots, X_n)$ sont des variables mutuellement indépendantes, et $(u_1, \cdots, u_n) \in \mathbf{R}^{\mathbf{R}}$, alors $(u_1(X_1), u_2(X_2), \cdots, u_n(X_n))$ sont aussi mutuellement indépendantes.

Exemple

Si X, Y, Z sont indépendantes, alors $X+Y$ et Z sont indépendantes.

Théorème 7.5 *(Lemme des coalitions)*

Si $(X_1, X_2, \cdots, X_n)$ sont des variables mutuellement indépendantes, alors pour tout $m \in [\![1, n]\!]$, $f : \mathbf{R}^m \to \mathbf{R}$ et $g : \mathbf{R}^{n-m} \to \mathbf{R}$, on a

$$f(X_1, \cdots, X_m) \text{ et } g(X_{m+1}, \cdots, X_n) \text{ indépendantes.}$$

Ce résultat se généralise à plus de deux coalitions.

Preuve

On montre le résultat pour 2 coalitions. On remarque ensuite qu'il se généralise sans problème à davantage de coalitions par récurrence. Le théorème précédent (l'indépendance passe à l'image) peut alors être considéré comme un cas particulier avec exactement n coalitions.
Pour la preuve, on montre que les variables aléatoires $X = (X_1, \cdots, X_m)$ et $Y = (X_{m+1}, \cdots, X_n)$ sont indépendants (on étend naturellement la notion de variable aléatoire aux fonctions à valeurs dans $\mathbf{R}^k$).
Pour cela on constate simplement que

$$\begin{aligned}
&\mathbf{P}\left(X=(x_1,\cdots,x_m)\ , Y=(x_{m+1},\cdots,x_n)\right)\\
&\qquad = \mathbf{P}\left(X_1=x_1,\cdots,X_n=x_n\right)\\
&\qquad = \prod_{i=1}^n \mathbf{P}\left(X_i=x_i\right)\\
&\qquad = \prod_{i=1}^m \mathbf{P}\left(X_i=x_i\right)\prod_{i=m+1}^n \mathbf{P}\left(X_i=x_i\right)\\
&\qquad = \mathbf{P}\left(X=(x_1,\cdots,x_m)\right)\mathbf{P}\left(Y=(x_{m+1},\cdots,x_n)\right).
\end{aligned}$$

Ensuite, on utilise le passage à l'image (non détaillé). ■

Propriété 7.6 *(Variance d'une somme)*

Si $(X_1, X_2, \cdots, X_n)$ sont n variables aléatoires mutuellement indépendantes, alors

$$\mathbf{E}(X_1\, X_2 \cdots X_n) = \mathbf{E}(X_1)\, \mathbf{E}(X_2) \cdots \mathbf{E}(X_n).$$

$$\mathbf{V}(X_1 + X_2 + \cdots + X_n) = \mathbf{V}(X_1) + \mathbf{V}(X_2) + \cdots + \mathbf{V}(X_n).$$

Preuve

Par récurrence. ■

B Application aux lois de Bernoulli et binomiale

Propriété 7.7 *(Somme de lois de Bernoulli)*

Soient $(X_i)_{i\in[\![1,n]\!]}$ des variables aléatoires mutuellement indépendantes sur Ω,
Soit $p \in [0,1]$,
Si $\forall i \in [\![1,n]\!], X_i \hookrightarrow \mathscr{B}(p)$, alors

$$\sum_{i=1}^{n} X_i \hookrightarrow \mathscr{B}(n,p).$$

Lemme 7.8 *(Rappel : Formule de Vandermonde)*

Soient $n, m, k \in N^*$, avec $k \leqslant n+m$.

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

Preuve

L'idée (comme souvent pour ce genre de relations) est de repasser aux polynômes.

$$\begin{aligned}
&(1+X)^{m+n} = (1+X)^n (1+X)^m \\
\iff &\sum_{k=0}^{m+n} \binom{m+n}{k} X^k = \sum_{i=0}^{m} \binom{n}{i} X^i \sum_{j=0}^{m} \binom{m}{j} X^j \\
\iff &\sum_{k=0}^{m+n} \binom{m+n}{k} X^k = \sum_{i=0}^{n}\sum_{j=0}^{m} \binom{n}{i}\binom{m}{j} X^{j+i}.
\end{aligned}$$

On identifie le coefficient de X^k dans les deux expressions :
Dans la deuxième expression, pour $i \geqslant 0$ fixé, $j = k - i \geqslant 0$. Donc i varie dans $[\![0,k]\!]$

$$\binom{n}{k} = \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i}.$$

On peut aussi la retrouver facilement par un raisonnement de dénombrement. ■

Propriété 7.9 *(Stabilité de la loi binomiale)*

Soient X, Y deux variables aléatoires **indépendantes** sur Ω,
Si $X \hookrightarrow \mathscr{B}(n,p)$ et $Y \hookrightarrow \mathscr{B}(m,p)$ avec le **même paramètre** p,
alors $X + Y \hookrightarrow \mathscr{B}(n+m,p)$.

Preuve

$$\begin{aligned}
\mathbf{P}(X+Y=k) &= \sum_{i=0}^{k} \mathbf{P}(X=i \cap Y=k-i) \\
&= \sum_{i=0}^{k} \mathbf{P}(X=i)\mathbf{P}(Y=k-i) \quad \text{variables indépendantes} \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} \\
&= \binom{n+m}{k} p^k (1-p)^{n+m-k} \quad \text{Formule de Vandermonde.}
\end{aligned}$$

■

Exemple

Un institut de recherche effectue des tests sur le comportement des souris dans un labyrinthe. Deux chercheurs se partagent la tâche : sur les 300 essais, l'un en effectue 100, et le second 200.
Le labyrinthe comporte un embranchement, et les souris ont la probabilité p d'aller à gauche et la probabilité $1-p$ d'aller à droite. Les épreuves sont indépendantes les unes des autres.
On compte pour chaque chercheur le nombre de souris qui sont allées à droite.
On note ce nombre X_1 pour le premier chercheur et X_2 pour le second.
Alors $X_1 \hookrightarrow \mathscr{B}(100,p)$ et $X_2 \hookrightarrow \mathscr{B}(200,p)$.
On comprend bien que si on met ensemble les deux expériences, alors $X_1 + X_2 \hookrightarrow \mathscr{B}(300,p)$.

Corollaire 7.10

Soient $(X_i)_{i\in[\![1,m]\!]}$ des variables aléatoires mutuellement indépendantes sur Ω,
Soit $p \in [0,1]$ et $(n_1, n_2, \cdots, n_m) \in \mathbf{N}^m$,
Si $\forall i \in [\![1,m]\!], X_i \hookrightarrow \mathscr{B}(n_i,p)$, alors

$$\sum_{i=1}^{m} X_i \hookrightarrow \mathscr{B}(n_1 + n_2 + \cdots + n_m, p).$$

C Un beau résultat pour finir : la loi faible des grands nombres

Ce dernier résultat est hors programme en première année.

Théorème 7.11 *(Loi faible des grands nombres (⋆))*

Si $(X_1, X_2, \cdots, X_n, \cdots)$ est une suite de variables aléatoires

1. mutuellement indépendantes,
2. de même loi $\mathbf{P}_X$.

Alors pour tout $\varepsilon > 0$,

$$\lim_{n\to+\infty} \mathbf{P}\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mathbf{E}(X)\right| \geqslant \varepsilon\right) = 0.$$

Explications

Ce résultat permet de justifier que, quelque soit notre exigence de précision, lorsque l'on répète un grand nombre de fois une même expérience aléatoire, de façon indépendante, et que l'on calcule la moyenne, on est presque sûr d'obtenir l'espérance (à l'incertitude près fixée au préalable et arbitrairement petite).

Preuve

On note $Y_n = \frac{X_1+X_2+\cdots+X_n}{n}$.
Alors $\mathbf{E}(Y_n) = \mathbf{E}(X)$ par linéarité de l'espérance.
Or

$$\mathbf{V}(Y_n) = \frac{1}{n^2}\mathbf{V}(X_1+X_2+\cdots+X_n) = \frac{\mathbf{V}(X_1)+\mathbf{V}(X_2)+\cdots+\mathbf{V}(X_n)}{n^2} = \frac{\mathbf{V}(X)}{n}.$$

Donc d'après l'inégalité de Bienaymé-Tchebychev, pour tout $\varepsilon > 0$

$$\mathbf{P}\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mathbf{E}(X)\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbf{V}(X)}{n\varepsilon^2}.$$

Donc pour $n \to +\infty$, on obtient la limite voulue. ■

Exemple

Donner un contre-exemple à la loi faible des grands nombres lorsque les variables ne sont pas supposées mutuellement indépendantes.

Solution :

On peut poser $X_1 \hookrightarrow \mathscr{B}\left(\frac{1}{2}\right)$, et $\forall n \geqslant 2,\ X_n = X_1$.
Alors $X_n \hookrightarrow \mathscr{B}\left(\frac{1}{2}\right)$. et $\frac{1}{n}\sum_{k=1}^{n} X_k = X_1$. C'est donc une variable aléatoire à valeurs dans $\{0,1\}$.
En particulier pour $\varepsilon = \frac{1}{4}$,

$$\mathbf{P}\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mathbf{E}(X)\right| \geqslant \varepsilon\right) = \mathbf{P}\left(\left|X_1 - \frac{1}{2}\right| \geqslant \frac{1}{4}\right) = 1.$$

Ce qui ne tend pas vers 0 quand $n \to +\infty$.

8 Complément hors programme : fonction génératrice

Cette notion est hors programme, nous en esquissons simplement le contour car elle est simple, elle facilite les calculs... et vous en parlerez l'an prochain.

Définition 8.1 *(Fonction génératrice)*

Soit $\mathbf{P}$ une probabilité sur un univers Ω et X une variable aléatoire définie sur Ω et à valeurs dans $[\![0,n]\!]$.
On définit la fonction génératrice de X par

$$\forall s \in \mathbf{R},\ G(s) = \mathbf{E}\left(s^X\right) = \sum_{k=0}^{n} s^k \mathbf{P}(X=k).$$

La fonction génératrice est une application polynomiale. Nous pouvons donc utiliser tous les outils d'analyse et sur les polynômes pour l'étudier. Ses dérivées successives correspondent aux *moments* de la variable. En particulier, on retrouve facilement l'espérance et la variance (qui correspond au moment centré d'ordre 2).

Propriété 8.2

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

Preuve

$G(1) = \sum_{k=0}^{n} \mathbf{P}(X=k)1^k = \sum_{k=0}^{n} \mathbf{P}(X=k) = 1$
car $\{[X=k],\ k \in [\![0,n]\!]\}$ forme un système complet d'événements.
$G'(1) = \sum_{k=0}^{n} k\mathbf{P}(X=k)1^{k-1} = \sum_{k=0}^{n} k\mathbf{P}(X=k) = \mathbf{E}(X)$.
$G''(1) = \sum_{k=0}^{n} k(k-1)\mathbf{P}(X=k)1^{k-2} = \sum_{k=0}^{n} k(k-1)\mathbf{P}(X=k) = \mathbf{E}(X(X-1))$
(d'après la formule de transfert).

$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 \\ &= \mathbf{E}\left(X(X-1)+X\right) - \left(G'(1)\right)^2 \\ &= \mathbf{E}\left(X(X-1)\right) + \mathbf{E}(X) - \left(G'(1)\right)^2 = G''(1) + G'(1) - \left(G'(1)\right)^2.\end{aligned}$$

■

Exemple *(Loi binomiale)*

Exemple de l'utilisation de la fonction génératrice pour trouver l'espérance et la variance de la loi binomiale. La fonction génératrice de la loi binomiale peut s'écrire :

$$G_X(x) = \sum_{k=0}^{n} x^k P(X=k) = \sum_{k=0}^{n} x^k \binom{n}{k} p^k (1-p)^{n-k} = (xp+1-p)^n.$$

Or $\mathbf{E}(X) = G'_X(1)$, et $\mathbf{E}\left(X(X-1)\right) = G''_X(1)$.

Donc $\mathbf{V}(X) = \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 = \mathbf{E}\left(X(X-1)+X\right) - (\mathbf{E}(X))^2$
$= \mathbf{E}\left(X(X-1)\right) + \mathbf{E}(X) - (\mathbf{E}(X))^2$
$= G''_X(1) + G'_X(1)(X) - (G'_X(1))^2$.
$\forall x \in \mathbf{R},\ G'_X(x) = np(xp+1-p)^{n-1}$. Donc $\mathbf{E}(X) = G'_X(1) = np$.
$\forall x \in \mathbf{R},\ G''_X(x) = n(n-1)p^2(xp+1-p)^{n-2}$. Donc $G''_X(1) = n(n-1)p^2$.
Ainsi

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

La fonction génératrice contient toute l'information sur la loi de X.

Propriété 8.3

$\forall k \in \mathbf{N}$,

$$\mathbf{P}(X=k) = \frac{G^{(k)}(0)}{k!}.$$

Preuve

D'après l'égalité de Taylor sur les polynômes, $G(s) = \sum_{k=0}^{+\infty} \frac{1}{k!} G^{(k)}(0) s^k$ et par identification, on a donc l'égalité voulue. ■