Chapitre 9

Variables aléatoires discrètes

Table des matières

9.1 Variables aléatoires discrète

9.1.1 Définitions

DÉFINITION 9.1 ★★★ **Variable aléatoire**
On appelle *variable aléatoire discrète* sur l'espace probabilisable $(\Omega, \mathcal{T})$ à valeurs dans un ensemble E une application $X : \Omega \to E$ définie sur Ω

1. dont l'image $X(\Omega)$ est finie ou dénombrable
2. et telle que l'image réciproque de tout élément de $X(\Omega)$ appartient à $\mathcal{T}$:

$$\forall \omega \in X(\Omega), \quad X^{-1}(\{\omega\}) \in \mathcal{T}.$$

Remarque 9.1 L'idée sous-tendant cette définition est qu'on veut que pour tout $x \in E$, l'ensemble $X^{-1}(\{x\})$ soit un événement de $(\Omega, \mathcal{T})$.

Remarque 9.2 Dans le cas où $E = \mathbb{R}$, la variable aléatoire X est dite réelle. L'ensemble $X(\Omega)$ est un sous-ensemble au plus dénombrable de $\mathbb{R}$ et donc il ne contient aucun intervalle réel. En général, $X(\Omega) = \mathbb{N}$ ou $X(\Omega) = \mathbb{Z}$ qui sont des sous-ensembles *discrets* de $\mathbb{R}$, c'est-à-dire des sous-ensembles de $\mathbb{R}$ dont chaque point x admet un voisinage n'intersectant $X(\Omega)$ qu'en ce point x.

PROPOSITION 9.1 ★★★ **Image réciproque d'une réunion, d'une intersection**
ersection
Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $X : \Omega \to E$ une variable aléatoire discrète. Pour tout $A, B \in E$, on a :

1. $X^{-1}(A \cup B) = X^{-1}(A) \cup X^{-1}(B)$;
2. $X^{-1}(A \cap B) = X^{-1}(A) \cap X^{-1}(B)$;
3. $X^{-1}(\overline{A}) = \overline{X^{-1}(A)}$;

et plus généralement, si $(A_i)_{i \in I}$ est une suite de sous-ensembles de E :

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

Démonstration *Prouvons les deux dernières formules, les deux premières en sont des cas particuliers :*

1. *Soit $x \in X^{-1}\left(\bigcup_{i \in I} A_i\right)$. Il existe $y \in \bigcup_{i \in I} A_i$ tel que $X(x) = y$. Alors il existe $i \in I$ tel que $y \in A_i$ et donc $x \in X^{-1}(A_i)$. En conclusion $x \in \bigcup_{i \in I} X^{-1}(A_i)$. Réciproquement, si $x \in \bigcup_{i \in I} X^{-1}(A_i)$ alors il existe $i \in I$ et $y \in A_i$ tel que $X(x) = y$. Mais $y \in \bigcup_{i \in I} A_i$ donc $x \in X^{-1}\left(\bigcup_{i \in I} A_i\right)$. Voilà qui prouve la première égalité.*
2. *Soit $x \in X^{-1}\left(\bigcap_{i \in I} A_i\right)$. Alors pour tout $i \in I$, il existe y_i tel que $X(x) = y_i$. Donc pour $i \in I$, $x \in X^{-1}(A_i)$ et on prouve ainsi que $x \in \bigcap_{i \in I} X^{-1}(A_i)$. Réciproquement, si $x \in \bigcap_{i \in I} X^{-1}(A_i)$ alors pour tout $i \in I$, il existe $y_i \in A_i$ tel que $X(x) = y_i$. Alors $X(x) \in \bigcap_{i \in I} A_i$ et $x \in X^{-1}\left(\bigcap_{i \in I} A_i\right)$ et la seconde égalité est aussi prouvée.*

On prouve encore la troisième formule. On a les équivalences

$$x \in X^{-1}(\overline{A}) \iff X(x) \in \overline{A} \iff X(x) \notin A \iff x \notin X^{-1}(A) \iff x \in \overline{X^{-1}(A)}.$$

PROPOSITION 9.2 ★★★ **L'image réciproque d'une partie de** E **par une variable aléatoire discrète est un événement**
_evenement
Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $X : \Omega \to E$ une variable aléatoire discrète. Pour tout $U \subset E$, on a $X^{-1}(U) \in \mathcal{T}$.

Démonstration *En effet, comme $U \cap X(\Omega)$ est au plus dénombrable, cet ensemble peut s'écrire $U \cap X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ alors*

$$X^{-1}(U) = X^{-1}\left(\bigcup_{n \in \mathbb{N}} \{x_n\}\right) = \bigcup_{n \in \mathbb{N}} X^{-1}(\{x_n\})$$

qui est bien un élément de $\mathcal{T}$ d'après l'axiome **??** *définissant une tribu.*

✎ *Notation 9.1* Pour $x \in E$, on note $(X = x)$ ou $\{X = x\}$ l'événement $X^{-1}(\{x\})$ et pour $A \subset E$, on note $(X \in A)$ ou $\{X \in A\}$ l'événement $X^{-1}(A)$.
Avec ces notations, les trois dernières égalités de la proposition 9.1 s'écrivent :

$$\left\{X \in \overline{A}\right\} = \overline{\{X \in A\}}, \quad \left\{X \in \bigcup_{i \in I} A_i\right\} = \bigcup_{i \in I} \{X \in A_i\} \quad \text{et} \quad \left\{X \in \bigcap_{i \in I} A_i\right\} = \bigcap_{i \in I} \{X \in A_i\}.$$

Remarquons par ailleurs que $\{X \in \varnothing\} = \varnothing$ et que $\{X \in E\} = \Omega$.

evenements

Remarque 9.3 Si non note $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$ alors la famille $(\{X = x_k\})_{k\in\mathbb{N}}$ forme un système complet d'événements de Ω.

DÉFINITION - PROPOSITION 9.2 ★ **Variable aléatoire indicatrice**
Soit $(\Omega, \mathcal{T})$ un espace probabilisable et soit $A \in \mathcal{T}$ alors la variable aléatoire notée 1_A et définie par

$$1_A : \begin{cases} \Omega & \longrightarrow & \{0,1\} \\ \omega & \longmapsto & X(\omega) = \begin{cases} 1 & \text{si} \quad \omega \in A \\ 0 & \text{si} \quad \omega \notin A \end{cases} \end{cases}$$

est une variable aléatoire réelle discrète appelée *variable indicatrice* de A

Démonstration *La variable aléatoire 1_A a comme image $\{0,1\}$ qui est fini et comme $1_A^{-1}(1) = A$, $1_A^{-1} = \overline{A}$, l'image réciproque d'un élément de $1_A(\Omega)$ est bien un élément de $\mathcal{T}$.*

DÉFINITION - PROPOSITION 9.3 ★ **Fonction d'une variable aléatoire**
Soit X une variable aléatoire discrète sur un espace probabilisable $(\Omega, \mathcal{T})$ et à valeurs dans E. Soit f une fonction définie sur E et à valeurs dans un ensemble F. Alors $f \circ X$ est aussi une variable aléatoire discrète appelée *image de la variable aléatoire* X par la fonction f et notée $f(X)$

Démonstration *Comme $X(\Omega)$ est au plus dénombrable, il en est de même de $f(X(\Omega))$. Soit $A \subset F$. On a $(f \circ X)^{-1}(A) = X^{-1}\left(f^{-1}(A)\right)$ mais $f^{-1}(A)$ est une partie de E et d'après la proposition 9.2, $X^{-1}\left(f^{-1}(A)\right)$ est donc un élément de $\mathcal{T}$. Ainsi $f(X)$ est bien une variable aléatoire discrète sur $(\Omega, \mathcal{T})$.*

9.1.2 Loi de probabilité et fonction de répartition d'une variable aléatoire

Rappelons que pour un ensemble E, $\mathscr{P}(E)$ est une tribu sur E.

DÉFINITION - PROPOSITION 9.4 ★★★ **Loi de probabilité d'une variable aléatoire**
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et soit $X : \Omega \to E$ une variable aléatoire discrète. L'application

$$P_X : \begin{cases} \mathscr{P}(X(\Omega)) & \longrightarrow & [0,1] \\ A & \longmapsto & P(X \in A) \end{cases}$$

est une probabilité sur $(X(\Omega), \mathscr{P}(X(\Omega)))$ appelée *loi de probabilité de la variable aléatoire* X.

Démonstration *Vérifions les axiomes définissant une probabilité :*

1. $P_X(X(\Omega)) = P\left(X^{-1}(X(\Omega))\right) = P(\Omega) = 1$.
2. *Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments deux à deux disjoints de $\mathscr{P}(X(\Omega))$. Alors $(X^{-1}(A_n))_{n\in\mathbb{N}}$ est une suite d'éléments deux à deux disjoints de $\mathcal{T}$ et par σ-additivité*

$$P_X\left(\bigcup_{n\in\mathbb{N}} A_n\right) = P\left(X^{-1}\left(\bigcup_{n\in\mathbb{N}} A_n\right)\right) = P\left(\bigcup_{n\in\mathbb{N}} X^{-1}(A_n)\right) = \sum_{n=0}^{+\infty} P\left(X^{-1}(A_n)\right) = \sum_{n=0}^{+\infty} P_X(A_n).$$

Donc P_X définit bien une probabilité sur $(X(\Omega), \mathscr{P}(X(\Omega)))$.

DÉFINITION 9.5 ★★★ **Variables aléatoires de même loi**
On dit que deux variables aléatoires X et Y sont de *même loi* si $P_X = P_Y$. On le note $X \sim Y$.
Si X est une variable aléatoire et que $\mathscr{L}$ est une loi discrète, on note $X \sim \mathscr{L}$ le fait que la loi de X est $\mathscr{L}$.

Remarque 9.4 Dire que deux variables aléatoires sont de même loi revient à dire à la fois que :

1. $X(\Omega) = Y(\Omega)$;
2. $\forall x \in X(\Omega), \quad P(X = x) = P(Y = x)$.

THÉORÈME 9.3 ★★★ $X \sim Y \implies f(X) \sim f(Y)$
Soit X, Y deux variables aléatoire discrètes définies sur un espace probabilisable $(\Omega, \mathcal{T})$ et à valeurs dans E. Soit f une fonction définie sur E et à valeurs dans un ensemble F.
On suppose que X et Y suivent la même loi : $X \sim Y$. Alors $f \circ X$ et $f(Y)$ sont aussi de même loi : $f(X) \sim f(Y)$

Démonstration *Admis.*

✎ *Notation 9.2* Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire discrète. Alors pour $x \in \mathbb{R}$, on note :

— $(X > x) = X^{-1}(]x,+\infty[)$;

— $(X \geqslant x) = X^{-1}([x,+\infty[)$;

— $(X < x) = X^{-1}(]-\infty,x[)$;

— $(X \leqslant x) = X^{-1}(]-\infty,x])$;

DÉFINITION 9.6 ★ **Fonction de répartition d'une variable aléatoire discrète réelle**
On appelle *fonction de répartition* de la variable aléatoire discrète réelle $X : \Omega \to \mathbb{R}$ définie sur un espace probabilisé $(\Omega,\mathcal{T},P)$, la fonction

$$F_X : \left\{ \begin{array}{rcl} \mathbb{R} & \longrightarrow & [0,1] \\ x & \longmapsto & P_X(]-\infty,x]) \end{array} \right.$$

Remarque 9.5

— On a donc

$$F_X(x) = P(X \leqslant x) = P\left(X^{-1}(]-\infty,x])\right).$$

— Si $X(\Omega) = \left\{x_k \mid k \in K\right\}$ avec K au plus dénombrable alors $P(X \leqslant X) = \sum_{k\in L} P(X = x_k)$ où $L = \left\{k \in K \mid x_k \leqslant x\right\}$.

PROPOSITION 9.4 ★ **Propriétés des fonctions de répartition**
Soit X une variable aléatoire discrète réelle définie sur un espace probabilisé $(\Omega,\mathcal{T},P)$ et soit F_X sa fonction de répartition. Alors :

1. F_X est croissante ;
2. F_X admet une limite à droite et à gauche en tout point.
3. $\lim_{-\infty} F_X = 0$ et $\lim_{+\infty} F_X = 1$;
4. Pour tout $x \in \mathbb{R}$,
$$P(X = x) = F_X(x) - F_X(x^-)$$
et la fonction F_X caractérise la loi de X [1].
5. Pour tout $a, b \in \mathbb{R}$, $a \leqslant b$,
$$\boxed{P(a < X \leqslant b) = F_X(b) - F_X(a)}.$$

Démonstration

1. *Soient $x, x' \in \mathbb{R}$ tels que $x \leqslant x'$ alors $]-\infty,x] \subset \left]-\infty,x'\right]$ et $X^{-1}(]-\infty,x]) \subset X^{-1}\left(\left]-\infty,x'\right]\right)$. Il s'ensuit que $P\left(X^{-1}(]-\infty,x])\right) \leqslant P\left(X^{-1}\left(\left]-\infty,x'\right]\right)\right)$ ce qui s'écrit aussi $P_X(X \leqslant x) \leqslant P(X \leqslant x')$. On a démontré que $F_X(x) \leqslant F_X(x')$ donc que F_X est croissante.*
2. *Pour $x_0 \in \mathbb{R}$ fixé, la fonction $F_X : x \mapsto F_X(x)$ pour $x < x_0$ est donc croissante et majorée. Donc par le théorème de la limite monotone, F_X admet une limite à gauche de x_0. On montre de même qu'elle admet une limite à droite de x_0.*
3. *Comme F_X est croissante et majorée par 1 (par définition d'une probabilité), elle admet d'après le théorème de la limite monotone une limite $l \geqslant 1$ en $+\infty$. Mais $]-\infty,+\infty[= \bigcup_{n=0}^{+\infty}]-\infty,n]$ et par continuité croissante de P on a*
$$l = \lim_{n\to\infty} F_X(n) = \lim_{n\to\infty} P(X \in]-\infty,n]) = P(X \in \bigcup_{n=0}^{+\infty}]-\infty,n]) = P_X(\mathbb{R}) = 1$$
donc $\lim_{+\infty} F_X = 1$.
De même, F_X admet une limite $l' \geqslant 0$ en $-\infty$ et pour tout $n \in \mathbb{N}$,
$$l' = \lim_{x\to-\infty} F_X(x) = \lim_{n\to+\infty} P(X \in]-\infty,-n]) = P\left(X \in \bigcap_{n\in\mathbb{N}}]-\infty,-n]\right) = P(\varnothing) = 0$$
en utilisant le théorème de continuité décroissante.
4. *Soit $x \in \mathbb{R}$. Comme $]-\infty,x] = \bigcup_{n=0}^{+\infty} \left]-\infty,x-2^{-n}\right]$, par continuité croissante de P,*
$$P(X < x) = P\left(\bigcup_{n=0}^{+\infty} \left]-\infty,x-2^{-n}\right]\right) = \lim_{n\to+\infty} P(\left]-\infty,x-2^{-n}\right]) = \lim_{n\to+\infty} F_X(x-2^{-n}) = F(x^-).$$
Comme $(X \leqslant x) = (X = x) \setminus (X < x)$, on obtient $P_X(\{x\}) = P(X = x) = F_X(x) - F_X(x^-)$.
5. *Soient $a, b \in \mathbb{R}$, $a \leqslant b$, alors $(a < X \leqslant b) = (X \leqslant b) \setminus (X \leqslant a)$ d'où le résultat.*

9.1.3 Loi associée à une image dénombrable

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Comme $X(\Omega)$ est au plus dénombrable, on peut l'écrire $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$. Remarquons alors que la série de terme général $P(X = x_n)$ est convergente de somme 1.
Ce fait est une conséquence directe de la remarque 9.3. La suite $(\{X = x_n\})_{n\in\mathbb{N}}$ forme un système complet d'événements de Ω. Les ensembles $(X = x_n)$ sont donc deux à deux disjoints et de réunion égale à Ω. Alors d'après l'axiome d'additivité dénombrable, la série de terme général $P(X = x_n)$ est convergente de somme

$$\sum_{n=0}^{+\infty} P(X = x_n) = P\left(\bigcup_{n=0}^{+\infty} (X = x_n)\right) = P(\Omega) = 1.$$

Cette remarque amène la question naturelle suivante. Si (p_n) est une suite de réels positifs telle que la série de terme général p_n est convergente de somme 1 alors pour un espace probabilisable $(\Omega, \mathcal{T})$ et une variable aléatoire discrète $X : \Omega \to \mathbb{R}$, existe-t-il une probabilité P sur $(\Omega, \mathcal{T})$ telle que pour tout $x_n \in X(\Omega)$, $P(X = x_n) = p_n$?
La réponse est positive mais utilise un résultat qui n'est pas au programme en PC. Si on considère une famille $(t_k)_{k\in K}$ de réels *positifs* avec K dénombrable alors la somme $\sum_{k\in K} t_k$, si elle existe, est indépendante de l'ordre choisi pour sommer les t_k.
On a le théorème suivant dont la preuve est hors-programme :

THÉORÈME 9.5 ★★★ Loi associée à une image dénombrable
Soit X une variable aléatoire discrète définie sur un espace probabilisable $(\Omega, \mathcal{T})$ et à valeurs dans un ensemble E. On note $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ et on considère (p_n) une suite réelle telle que :

- H1 (p_n) est une suite de réels positifs ;
- H2 La série de terme générale p_n est convergente de somme 1 :

$$\sum_{n=0}^{\infty} p_n = 1$$

alors il existe une unique probabilité P sur $(\Omega, \mathcal{T})$ telle que

$$\boxed{P(X = x_n) = p_n}.$$

Démonstration *Hors-Programme Procédons par analyse et synthèse :*

Analyse *On suppose qu'une telle probabilité P existe. Alors si $A \in \mathcal{P}(X(\Omega))$ et si on note $k(A) = \{k \in \mathbb{N} \mid x_k \in A\}$, on a :*

$$(X \in A) = \bigcup_{k\in k(A)} (X = x_k)$$

et donc comme les événements $(X = x_k)$ sont deux à deux disjoints, d'après l'axiome d'additivité dénombrable, $P(X \in A) = \sum_{k\in k(A)} P(X = x_k) = \sum_{k\in k(A)} p_k$. Il reste à voir que cette somme existe et est bien définie. L'existence découle du fait que $k(A) \subset \mathbb{N}$, que (p_n) est positive et que $\sum_{n\in\mathbb{N}} p_n = 1$. Le fait qu'elle est bien définie découle des explications précédant le théorème.

Synthèse *On pose pour tout $A \in \mathcal{P}(X(\Omega))$, $P(X \in A) = \sum_{k\in k(A)} p_k$ et on vérifie que P est bien une probabilité sur $(X(\Omega), \mathcal{P}(\Omega))$. Il est tout d'abord clair que $P(X(\Omega)) = \sum_{n\in\mathbb{N}} P(X = x_n) = \sum_{n\in\mathbb{N}} p_n = 1$. Par ailleurs si (A_n) est une suite d'événements de $\mathcal{P}(X(\Omega))$ deux à deux disjoints alors $P(A_n) = \sum_{k\in k(A_n)} p_k$ et comme les A_n sont deux à deux disjoints, les ensembles $k(A_n)$ pour $n \in \mathbb{N}$ forment une partition de $k(A)$ où $A = \bigcup_{n=0}^{+\infty} A_n$ et comme $P(A) = \sum_{k\in k(A)} p_k$ est définie et convergente, on a bien $\sum_{n\in\mathbb{N}} \sum_{k\in k(A_n)} p_k$ qui converge de somme $P(A)$ ce qui permet de prouver l'additivité dénombrable de P.*

Exemple 9.3 On considère la suite de terme général $p_n = 2^{-n-1}$ alors comme $\sum_{n=0}^{+\infty} p_n = 1$, on peut définir la loi de probabilité d'une variable aléatoire discrète X en posant $P(X = x_n) = p_n$.

9.2 Couples de variables aléatoires

9.2.1 Loi conjointe, loi marginale

aleatoires

DÉFINITION - PROPOSITION 9.7 ★ Couple de variables aléatoires discrètes

Soient E et F deux ensembles et $(\Omega, \mathscr{T})$ un espace probabilisable. On considère $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires discrètes. On introduit

$$Z : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & E \times F \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array} \right. .$$

Alors Z est une variable aléatoire discrète sur Ω notée (X,Y) et appelée *couple des variables aléatoires* X et Y. De plus, pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$, on a :

$$\boxed{\big(Z = (x,y)\big) = (X = x) \cap \big(Y = y\big)}.$$

Démonstration *Comme $X(\Omega)$ et $Y(\Omega)$ sont dénombrables, il en est de même de $Z(\Omega) = X(\Omega) \times Y(\Omega)$. De plus, si $(x,y) \in Z(\Omega)$ alors $Z^{-1}(\{(x,y)\}) = X^{-1}(\{x\}) \cap X^{-1}(\{y\}) \in \mathscr{T}$ comme intersection de deux éléments de $\mathscr{T}$. Donc Z est bien une variable aléatoire discrète sur Ω.*

DÉFINITION 9.8 ★ Loi conjointe, lois marginales

Soit $(\Omega, \mathscr{T}, P)$ un espace probabilisé et $X : \Omega \to E$, $Y : \Omega \to F$ deux variables aléatoires discrètes. On appelle :

1. *loi conjointe* du couple (X,Y) la loi de la variable aléatoire $Z = (X,Y)$.
2. *première loi marginale* du couple (X,Y) la loi de X.
3. *seconde loi marginale* du couple (X,Y) la loi de Y.

PROPOSITION 9.6 ★ Expression de la loi conjointe

Soit $(\Omega, \mathscr{T}, P)$ un espace probabilisé et $X : \Omega \to E$, $Y : \Omega \to F$ deux variables aléatoires discrètes. Pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$, on a :

$$\boxed{P_{(X,Y)}((x,y)) = P\big((X = x) \cap (Y = y)\big)}.$$

Démonstration *En utilisant la définition-proposition 9.9,*

$$P_{(X,Y)}((x,y)) = P\big((X,Y) = (x,y)\big) = P\big((X = x) \cap (Y = y)\big).$$

Remarque 9.6 Si $X(\Omega) = \{x_i \mid i \in \mathbb{N}\}$, que $Y(\Omega) = \{y_j \mid j \in \mathbb{N}\}$, que $P(X = x_i) = p_i$ et que $P(Y = y_j) = q_j$ alors

1. La loi conjointe du couple (X,Y) est l'application $X(\Omega) \times Y(\Omega) \to [0,1]$ définie par $(x_i, y_j) \mapsto p_{i,j}$ avec $p_{i,j} = P((X = x_i) \cap (Y = y_i))$.
2. L'application $X(\Omega) \to [0,1]$, $x_i \to p_{i,.}$ où $p_{i,.} = P(X = x_i)$ est la première loi marginale du couple (X,Y), c'est-à-dire la loi de X.
3. L'application $X(\Omega) \to [0,1]$, $y_j \to p_{.,j}$ où $p_{.,j} = P(Y = y_j)$ est la seconde loi marginale du couple (X,Y) ; c'est-à-dire la loi de Y.

_conjointe

Exemple 9.4 On considère une urne contenant 3 boules bleues, 2 boules jaunes et 4 vertes. On extrait 3 boules de cette urne. X désigne le nombre de boules bleues tirées et Y le nombre de boules jaunes. Pour $(k,l) \in [\![1,3]\!] \times [\![1,2]\!]$, on a

$$P_{(X,Y)}(k,l) = P((X = k) \cap (Y = l)) = \frac{\binom{3}{k}\binom{2}{l}\binom{4}{3-k-l}}{\binom{9}{3}}$$

$$P(X = k) = \frac{\binom{3}{k}\binom{6}{3-k}}{\binom{3}{n}} \quad \text{et} \quad P(Y = l) = \frac{\binom{2}{l}\binom{7}{3-l}}{\binom{3}{n}}.$$

Donc :

	Y = 0	Y = 1	Y = 2	loi de X
X = 0	4/84	12/84	4/84	20/84
X = 1	18/84	24/84	3/84	45/84
X = 2	12/84	6/84	0	18/84
X = 3	1/84	0	0	1/84
Loi de Y	35/84	42/84	7/84	1

Remarque 9.7 On voit dans ce tableau que la connaissance de la loi conjointe du couple (X,Y) entraîne celles de X et de Y. La réciproque est fausse, la connaissance des lois de X et de Y n'entraînent en général pas celle de la loi conjointe du couple (X,Y).

THÉORÈME 9.7 ★ **Liens entre les lois marginales et la loi conjointe**

— Pour tout $x \in X(\Omega)$, la série $\sum_{y\in Y(\Omega)} P\big((X,Y)=(x,y)\big)$ converge et sa somme est $P(X=x)$:

$$\boxed{P(X=x) = \sum_{y\in Y(\Omega)} P\big((X,Y)=(x,y)\big) = \sum_{y\in Y(\Omega)} P\left((X=x)\cap(Y=y)\right)};$$

— pour tout $y \in Y(\Omega)$, , la série $\sum_{x\in X(\Omega)} P\big((X,Y)=(x,y)\big)$ converge et sa somme est $P(Y=y)$

$$\boxed{P(Y=y) = \sum_{x\in X(\Omega)} P\big((X,Y)=(x,y)\big) = \sum_{x\in X(\Omega)} P\left((X=x)\cap(Y=y)\right)}.$$

Démonstration *La famille dénombrable $(Y=y)_{y\in Y(\Omega)}$ forme un système complet d'événements de Ω donc la famille $((X=x)\cap(Y=y))_{y\in Y(\Omega)}$ est formée d'événements deux à deux indépendants et par axiome d'additivité dénombrable, la série $\sum_{y\in Y(\Omega)} P\left((X=x)\cap(Y=y)\right) = \sum_{y\in Y(\Omega)} P\big((X,Y)=(x,y)\big)$ converge car converge absolument. De plus,*

$$P(X=x) = P\left((X=x)\cap\Omega\right) = P\big((X=x)\cap\big(\cup_{y\in Y(\Omega)}(Y=y)\big)\big) = \sum_{y\in Y(\Omega)} P\big((X=x)\cap(Y=y)\big).$$

Remarque 9.8 Autrement dit, avec les notations de la remarque précédente, pour tout $(i,j)\in\mathbb{N}^2$,

$$\boxed{P(X=x_i) = \sum_{j=0}^{+\infty} p_{i,j}} \quad \text{et} \quad \boxed{P(Y=y_j) = \sum_{I=0}^{+\infty} p_{i,j}}.$$

Remarque 9.9 On déduit de ces formules que si la loi conjointe de deux variables aléatoires est connue, alors on connaît les lois marginales de ces deux variables aléatoires. La réciproque est fausse.

9.2.2 Extension au cas d'un n-uplets de variables aléatoires

Soit $n \geqslant 1$ un entier.

aleatoires

DÉFINITION - PROPOSITION 9.9 ★ **n-uplets de variables aléatoires discrètes**
Soient $E_1,\dots,E_n$ des ensembles et $(\Omega,\mathscr{T})$ un espace probabilisable. On considère $X_1:\Omega\to E_1,\dots,X_n:\Omega\to E_n$ n variables aléatoires discrètes. On introduit

$$Z:\begin{cases}\Omega & \longrightarrow & E_1\times\dots\times E_n\\ \omega & \longmapsto & (X_1(\omega),\dots,X_n(\omega))\end{cases}.$$

Alors Z est une variable aléatoire discrète sur Ω notée $(X_1,\dots,X_n)$ et appelée *n-uplts des variables aléatoires* $X_1,\dots,X_n$. De plus, pour tout $(x_1,\dots,x_n)\in\in X_1(\Omega)\times\dots\times X_n(\Omega)$, on a :

$$\boxed{(Z=(x_1,\dots,x_n)) = (X_1=x_1)\cap\dots(X_n=x_n)}.$$

Démonstration *Pour tout $i\in[\![1,n]\!]$, l'ensemble $X_i(\Omega)$ est dénombrable. Il en est alors de même de $Z(\Omega)=X_1(\Omega)\times\dots X_n(\Omega)$ qui est un produit fini d'ensemble dénombrable. De plus, si $(x_1,\dots,x_n)\in Z(\Omega)$ alors $Z^{-1}(\{(x_1,\dots,x_n)\}) = X_1^{-1}(\{x_1\})\cap\dots,\cap X_n^{-1}(\{x_n\})\in\mathscr{T}$ comme intersection de n éléments de $\mathscr{T}$. Donc Z est bien une variable aléatoire discrète sur Ω.*

DÉFINITION 9.10 ★ **Loi conjointe, lois marginales**
Soit $(\Omega,\mathscr{T},P)$ un espace probabilisé et $X_1:\Omega\to E_1,\dots,X_n:\Omega\to E_n$ n variables aléatoires discrètes. On appelle :

1. *loi conjointe* du n-uplets $(X_1,\dots,X_n)$ la loi de la variable aléatoire $Z=(X_1,\dots,X_n)$.
2. Les *lois marginales* du n-uplet $Z=(X_1,\dots,X_n)$ sont les lois des variables aléatoires $X_1,\dots,X_n$.

Proposition 9.8 ★ Expression des lois marginales
Soit $Z = (X_1, \dots, X_n)$ un n-uplets de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{T}, P)$. Pour tout $k \in [\![1, n]\!]$ et $x_k \in X_k(\Omega)$, on a :

$$\boxed{P(X_k = x_k) = \sum_{(x_1,\dots,x_{k-1},x_{k+1},x_n) \in X_1(\Omega)\times\dots\times X_{k-1}(\Omega)\times X_{k+1}(\Omega)\times X_n(\Omega)} P\left((X_1, \dots, X_n) = (x_1, \dots, x_n)\right)}.$$

9.2.3 lois conditionnelles

Définition 9.11 ★ Loi conditionnelle
Soit (X, Y) un couple de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{T}, P)$.

— Pour $x \in X(\Omega)$ tel que $P(X = x) \neq 0$, on appelle *loi de* Y *conditionnée à* $(X = x)$ la probabilité P_x donnée pour tout $y \in Y(\Omega)$ par
$$P_x(y) = P_{(X=x)}(Y = y) = \frac{P\big((X = x) \cap (Y = y)\big)}{P(X = x)}.$$

— Pour $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$, on appelle *loi de* X *conditionnée à* $(Y = y)$ la probabilité P_y donnée pour tout $x \in X(\Omega)$ par
$$P_y(x) = P_{(Y=y)}(X = x) = \frac{P\big((X = x) \cap (Y = y)\big)}{P(Y = y)}.$$

La proposition suivante découle directement de la définition :

Proposition 9.9 ★ Lien entre lois conjointe, marginale et conditionnelle
On a
$$\boxed{P\big((X = x) \cap (Y = y)\big) = P_{(X=x)}(Y = y)P(X = x) = P_x(y)P(x)}.$$

9.2.4 Indépendance des variables aléatoires discrètes

Définition 9.12 ★ Variables aléatoires indépendantes
Soient X, Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{T}, P)$. On dit que (X, Y) est un *couple de variables aléatoires indépendantes* ou que X et Y sont *deux variables aléatoires indépendantes* si et seulement si pour tout $x \in X(\Omega)$ et $y \in Y(\Omega)$ les événements $(X = x)$ et $(Y = y)$ sont indépendants, c'est-à-dire si :
$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad P\big((X = x) \cap (Y = y)\big) = P(X = x)P(Y = y).$$
On le note $X \perp\!\!\!\perp Y$.

Proposition 9.10 ★ Caractérisation de l'indépendance de deux variables aléatoires
Soient X, Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{T}, P)$. On a équivalence entre :
— (X, Y) est un couple de variables aléatoires indépendantes ;
— $\forall A \times B \subset X(\Omega) \times Y(\Omega), \quad P\left((X \in A) \cap (Y \in B)\right) = P(X \in A)P(Y \in B)$.

Démonstration *La preuve est admise en PC.*

$\boxed{\Rightarrow}$ *On suppose que* (X, Y) *est un couple de variables indépendantes. Soient* $A \times B \in X(\Omega) \times Y(\Omega)$. *On effectue dans la démonstration dans le cas où* A *et* B *sont des parties infinies de, respectivement,* $X(\Omega)$ *et* $Y(\Omega)$. *Comme* $X(\Omega)$ *et* $Y(\Omega)$ *sont* *dénombrables, il en est de même de* A *et* B *et on peut écrire* $A = \{a_n \mid n \in \mathbb{N}\}$, $B = \{b_n \mid n \in \mathbb{N}\}$. *Alors :*

$$\begin{aligned}
P((X \in A) \cap (Y \in B)) &= P\left(\left(\cup_{n\in\mathbb{N}}(X = a_n)\right) \cap \left(\cup_{n\in\mathbb{N}}(Y = b_n)\right)\right) \\
&= P\left(\cup_{(m,n)\in\mathbb{N}^2} (X = a_m) \cap (Y = b_n)\right) \\
&= \sum_{(m,n)\in\mathbb{N}^2} P((X = a_m) \cap (Y = b_n)) \text{ par } \sigma\text{-additivité et car les événements } (X = a_m) \cap (Y = b_n) \text{ sont deux à deux disjoints} \\
&= \sum_{(m,n)\in\mathbb{N}^2} P(X = a_m) P(Y = b_n) \text{ car les variables aléatoires X et Y sont indépendantes} \\
&= \sum_{m\in\mathbb{N}} P(X = a_m) \sum_{n\in\mathbb{N}} P(Y = b_n) \text{ par le théorème de Fubini sur les séries} \\
&= P(A)P(B)
\end{aligned}$$

d'où le résultat. Remarquons qu'on a pu appliquer ici l'axiome de σ*-additivité car* $\mathbb{N}^2$ *est dénombrable.*

$\boxed{\Leftarrow}$ *La réciproque est triviale.*

9.2.5 Fonctions de variables aléatoires indépendantes

PROPOSITION 9.11 ★ **Fonctions de deux variables aléatoires indépendantes**

Soient X, Y deux variables aléatoires discrètes indépendantes définies sur un espace probabilisé $(\Omega, \mathscr{T}, P)$. Soient f et g deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$. Alors les variables aléatoires discrètes $f(X)$ et $g(Y)$ sont aussi indépendantes.

$$\boxed{X \perp\!\!\!\perp Y \implies f(X) \perp\!\!\!\perp f(Y)}$$

Démonstration *Soient* $a \in f \circ X(\Omega)$ *et* $b \in g \circ Y(\Omega)$. *Alors d'après la proposition 9.10*

$$P\left((f(X) = a) \cap (g(Y) = b)\right) = P\left((X \in f^{-1}(a)) \cap (Y \in g^{-1}(b))\right) = P(X \in f^{-1}(a)) P(Y \in g^{-1}(b)) = P(f(X) = a) P(f(Y) = b)$$

et donc $f(X)$ *et* $g(Y)$ *sont indépendantes.*

9.2.6 Variables mutuellement indépendantes

DÉFINITION 9.13 ★ **Variables mutuellement indépendantes**

Soit $(X_i)_{i\in[\![1,n]\!]}$ une famille finie de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{T}, P)$. On dit que ces variables aléatoires sont *mutuellement indépendantes* si et seulement si pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$, les événements $(X_1 = x_1), \dots, (X_n = x_n)$ sont mutuellement indépendants, c'est-à-dire si et seulement si :

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega), \quad P\left((X_i)_{i\in[\![1,n]\!]} = (x_i)_{i\in[\![1,n]\!]}\right) = \prod_{i=1}^{n} P(X_i = x_i).$$

Les résultats de première année sur les variables mutuellement indépendantes définies sur des espaces probabilisés finis restent valables sur des espaces dénombrables.

PROPOSITION 9.12

Soient $(X_i)_{i\in[\![1,n]\!]}$ une famille finie de variables aléatoires discrètes mutuellement indépendantes. Soit $(A_1, \dots, A_n) \in \mathscr{P}(X_1(\Omega)) \times \dots \times \mathscr{P}(X_n(\Omega))$. Alors

$$\boxed{P\left((X_i \in A_i)_{i\in[\![1,n]\!]}\right) = \prod_{i\in[\![1,n]\!]} P(X_i \in A_i)}.$$

Démonstration *La preuve est admise car utilise le théorème de Fubini sur les séries*

PROPOSITION 9.13

Soient $(X_i)_{i\in[\![1,n]\!]}$ une famille finie de variables aléatoires discrètes mutuellement indépendantes. Soit $(A_1, \dots, A_n) \in \mathscr{P}(X_1(\Omega)) \times \dots \times \mathscr{P}(X_n(\Omega))$. Alors les événements $(X_i \in A_i)_{i\in[\![1,n]\!]}$ sont mutuellement indépendants.

Démonstration *La preuve est admise*

PROPOSITION 9.14 **Une sous-famille de variables discrètes mutuellement indépendante est mutuellement indépendante**
Soient $(X_i)_{i\in[\![1,n]\!]}$ une famille finie de variables aléatoires discrètes mutuellement indépendantes. Alors toute sous-famille de $(X_i)_{i\in[\![1,n]\!]}$ est composée de variables mutuellement indépendante.

Démonstration *La preuve est admise*
La notion d'indépendance des variables aléatoires s'étend à une suite de variables aléatoires :

DÉFINITION 9.14 ★ **Suite de variables aléatoires indépendantes**
Soit (X_i) une suite de variables aléatoires discrètes. On dit que cette suite est une *suite de variables aléatoires indépendantes* si pour tout sous-ensemble fini non vide $I\subset\mathbb{N}$, la suite de variables aléatoires $(X_i)_{i\in I}$ est une famille finie de variables aléatoires indépendantes.

Terminons par un théorème très important.

THÉORÈME 9.15 ★★★ **Lemme des coalitions**
Soient $X_1,\dots,X_n$ des variables aléatoires discrètes mutuellement indépendantes définies sur Ω à valeurs dans des ensembles $E_1,\dots,E_n$. Soit $p\in[\![1,n-1]\!]$ et soient deux fonctions $f:E_1\times\dots\times E_p\to E$, $g:E_{p+1}\times\dots\times E_n\to F$ alors $f(X_1,\dots,X_p)$ et $g(X_{p+1},\dots,X_n)$ sont indépendantes.

9.2.7 Le jeu du pile ou face infini

ni_de_var

PLAN 9.1 : Lancers infinis, épisode 3/3

Terminons notre série sur les lancers infinis (Les deux premiers épisodes se trouvent page **??** *et* **??***).*

On effectue une infinité de jetés d'une pièce de monnaie et on s'intéresse à chaque fois au fait de savoir si c'est le côté pile ou bien le côté face de la pièce qui est obtenu. On sait que l'univers est modélisable par $\Omega=\{P,F\}^{\mathbb{N}}$ *qu'on muni de la tribu* $\mathcal{T}$ *engendrée par les événements cylindriques. Chaque facteur* $\Omega_k=\{P,F\}$ *de* Ω *peut être pourvu de la probabilité uniforme et les jetés sont supposés mutuellement indépendants.*

Le théorème **??** *page* **??** *permet d'affirmer qu'on peut faire de* $(\Omega,\mathcal{T})$ *un espace probabilisé* $(\Omega,\mathcal{T},P)$ *où pour toute partie cylindrique* $\prod_{k=0}^{+\infty}A_k$ *(où* $A_k\in\mathcal{P}(\Omega_k)$ *pour tout* $k\in\mathbb{N}$ *et où* $A_k=\Omega_k$ *à partir d'un certain rang N (qui dépend de la suite* (A_k) *considérée) sont éléments de* $\mathcal{T}$*), on a*

$$\boxed{P\left(\prod_{k=0}^{+\infty}A_k\right)=\prod_{k=0}^{N}P_k(A_k)}.$$

et où P_k *désigne la probabilité uniforme sur* $\Omega_k=\{P,F\}$.

On peut compléter ce théorème pour décrire l'utilisation de variables aléatoires discrètes sur ce type d'espaces probabilisables.

THÉORÈME 9.16 ★ ***Tribu des parties cylindriques***
On considère une suite dénombrable d'expériences aléatoires mutuellement indépendantes. On note $(\Omega_n,\mathcal{P}(\Omega_n),P_n)$ *les espaces probabilisés associés. On note* $\Omega=\prod_{k=0}^{+\infty}\Omega_n$, $\mathcal{T}$ *la tribu sur* Ω *engendrée par les événements cylindriques et P la probabilité obtenue grâce au théorème* **??***.*

Enfin, on considère une suite $(X_n)_{n\in\mathbb{N}}$ *de variables aléatoires discrètes réelles* X_n *définies sur l'espace probabilisé* $(\Omega_n,\mathcal{P}(\Omega_n),P_n)$*. Alors pour tout* $n>0$ *et pour tous sous-ensembles* $R_1,\dots,R_n$ *de* $\mathbb{R}$*, on a :*

$$\boxed{P\left(\bigcap_{k=0}^{n}(X_i\in R_i)\right)=\prod_{k=0}^{n}P_i(X_i\in R_i)}.$$

Si pour chaque lancer, on introduit une variable aléatoire de Bernoulli $X_k:\Omega_k\to\{0,1\}$ *égale à 1 si on obtient pile et à 0 si on obtient face alors on définit ainsi une suite* (X_n) *de variables aléatoires discrètes réelles et on a, si*

$(x_0,\dots,x_n)\in\{0,1\}^{n+1}$ *toujours d'après le théorème :*

$$\mathrm{P}\left(\bigcap_{k=0}^{n}(\mathrm{X}_i=x_i)\right)=\prod_{k=0}^{n}\mathrm{P}_i(\mathrm{X}_i=x_i)=\frac{1}{2^{n+1}}.$$

9.3 Moments d'une variable aléatoire

9.3.1 Espérance

On rappelle que dans le cas d'un univers $\Omega=\{\omega_1,\dots,\omega_n\}$ fini muni d'une probabilité P, l'espérance d'une variable aléatoire réelle $\mathrm{X}:\Omega\to\mathbb{R}$ est donnée par

$$\mathrm{E}(\mathrm{X})=\sum_{k=1}^{n}\mathrm{X}(\omega_i)\mathrm{P}(\omega_i).$$

Notre objectif maintenant est de généraliser cette notion dans le cas où Ω est dénombrable. Il va se poser le problème de travailler avec des sommes de la forme $\sum_{\omega\in\Omega}\mathrm{X}(\omega)\mathrm{P}(\omega)$ et donc avec des familles sommables.

DÉFINITION 9.15 ★★★ **Espérance d'une variable aléatoire**
Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega,\mathcal{T},\mathrm{P})$ à valeurs dans $[0,+\infty]$. L'*espérance* de X, notée E(X), est la somme dans $[0,+\infty]$ de la famille de réels positifs $(\mathrm{P}(\mathrm{X}=x))_{x\in\mathrm{X}(\Omega)}$:

$$\boxed{\mathrm{E}(\mathrm{X})=\sum_{x\in\mathrm{X}(\Omega)}x\mathrm{P}(\mathrm{X}=x)}$$

avec la convention :
- $x\mathrm{P}(\mathrm{X}=x)=0$ si $x=+\infty$;
- $\mathrm{P}(\mathrm{X}=+\infty)=0$.

Remarque 9.10 Si X est à valeurs dans $\mathbb{N}$, la sommabilité de la famille $(\mathrm{P}(\mathrm{X}=n)=$ équivaut à la convergence de la série $\sum_{n\geqslant0}n\mathrm{P}(\mathrm{X}=n)$ et l'on a (dans $[0,+\infty]$) : $\mathrm{E}(\mathrm{X})=\sum_{n=0}^{+\infty}\mathrm{P}(\mathrm{X}=n)n$.

DÉFINITION 9.16 ★★★ **Variable aléatoire d'espérance finie**
Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega,\mathcal{T},\mathrm{P})$ et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. On dit que X est d'*espérance finie* si la famille $(x\mathrm{P}(\mathrm{X}=x))_{x\in\mathrm{X}(\Omega)}$ est sommable. Sa somme est notée E(X) et est appelée l'*espérance de* X :

$$\boxed{\mathrm{E}(\mathrm{X})=\sum_{x\in\mathrm{X}(\Omega)}^{+\infty}x\mathrm{P}(\mathrm{X}=x)}.$$

Remarque 9.11
- Si X est positive, elle admet toujours une espérance, finie ou infinie.
- Si X n'est pas positive et que $(x\mathrm{P}(\mathrm{X}=x))_{x\in\mathrm{X}(\Omega)}$ n'est pas sommable, alors X n'admet pas espérance.
- Dans le cas d'un univers Ω fini, la définition de l'espérance l'espérance coincide avec celle de première année.
- Toute variable aléatoire à support fini admet une espérance finie.

Les deux propositions suivantes portent sur l'espérance de deux variables aléatoires fondamentales : les constantes et les indicatrices.

PROPOSITION 9.17 ★★★ **Espérance d'une variable aléatoire discrète constante**
_constante
Soit $(\Omega,\mathcal{T},\mathrm{P})$ un espace probabilisé et $\mathrm{X}:\Omega\to\mathbb{R}$ une variable aléatoire discrète réelle. On suppose que :

H1 il existe $m\in\mathbb{R}$ tel que : $\forall\omega\in\Omega,\quad \mathrm{X}(\omega)=m$

alors X admet une espérance donnée par $\boxed{\mathrm{E}(\mathrm{X})=m}$.

Démonstration *Comme* $\mathrm{X}(\Omega)=\{m\}$, *la famille définissant l'espérance se réduit à* $(m\mathrm{P}(\mathrm{X}=m))$ *qui vaut* m.

PROPOSITION 9.18 ★★★ **Espérande d'une indicatrice**
Soit $(\Omega,\mathscr{T},\mathrm{P})$ un espace propbabilisé. Soit $\mathrm{A}\in\mathscr{T}$ un événement et soit $\mathbb{1}_\mathrm{A}$ sa fonction indicatrice :

$$\mathbb{1}_\mathrm{A}(x)=\begin{cases}1 & \text{si} \quad x\in\mathrm{A}\\ 0 & \text{sinon}\end{cases}.$$

Alors $\mathbb{1}_\mathrm{A}$ admet une espérance finie et

$$\boxed{\mathrm{E}(\mathbb{1}_\mathrm{A})=\mathrm{P}(\mathrm{A})}.$$

Démonstration *En effet, la famille* $(x\mathrm{P}(\mathbb{1}_\mathrm{A}=x))_{x\in\mathrm{X}(\Omega)}$ *se réduit à* $(0,\mathrm{P}(\mathbb{1}_\mathrm{A})=1)$ *c'est à dire* $(0,\mathrm{P}(\mathrm{A}))$ *et est bien sommable de somme* $\mathrm{P}(\mathrm{A})$.

DÉFINITION 9.17 ★ **Variable centrée, variable centrée associée à une variable aléatoire**
- Une variable aléatoire réelle discrète admettant une espérance nulle est dite *centrée*.
- À une variables aléatoire réelle discrète X qui admet une espérance finie, on peut associer la variable aléatoire $\mathrm{X}-\mathrm{E}(\mathrm{X})$ qui est appelée *variable centrée associée à* X.

Pour calculer l'espérance d'une variable aléatoire à valeurs entières, le théorème suivant s'avère parfois pratique :

PROPOSITION 9.19 ★ **Calcul pratique de l'espérance d'une variable aléatoire à valeurs entières**
Soit X une variable aléatoire à valeurs dans $\mathbb{N}\cup\{+\infty\}$ et définie sur un espace probabilisé $(\Omega,\mathscr{T},\mathrm{P})$. Alors

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

Démonstration *D'après les conventions introduites en début de section, on peut supposer que* X *est à valeurs dans* $\mathbb{N}$.
Soit $k\in\mathbb{N}$. *On sait que*

$$\mathrm{P}(\mathrm{X}=k)=\mathrm{P}(\mathrm{X}\geqslant k)-\mathrm{P}(\mathrm{X}\geqslant k+1)$$

donc on a pour tout $n\in\mathbb{N}$,

$$\sum_{k=0}^{n}k\mathrm{P}(\mathrm{X}=k)=\sum_{k=0}^{n}k(\mathrm{P}(\mathrm{X}\geqslant k)-\mathrm{P}(\mathrm{X}\geqslant k+1))=\sum_{k=1}^{n}\mathrm{P}(\mathrm{X}\geqslant k)-n\mathrm{P}(\mathrm{X}\geqslant n+1)\quad(\star)$$

Procédons maintenant à la démonstration de la proposition.
- *Supposons que* X *ne soit pas d'espérance finie. Alors la série de terme général* $k\mathrm{P}(\mathrm{X}=k)$ *diverge. Il en est alors de même de la série de terme général* $\mathrm{P}(\mathrm{X}\geqslant k)$ *car :*
$$\sum_{k=1}^{n}\mathrm{P}(\mathrm{X}\geqslant k)\geqslant\sum_{k=0}^{n}k\mathrm{P}(\mathrm{X}=k).$$
Ainsi, $\mathrm{E}(\mathrm{X})=\sum_{n=1}^{+\infty}\mathrm{P}(\mathrm{X}>n)=+\infty$.
- *On suppose maintenant que* X *est d'espérance finie. Alors la série* $\sum_{k\in\mathbb{N}}k\mathrm{P}(\mathrm{X}=k)$ *converge et on peut considérer son reste d'ordre* n *qu'on note* $\tilde{\mathrm{R}}_n$. *Par définition d'une probabilité, comme la famille* $(\mathrm{X}=k)_{k\in\mathbb{N}}$ *est formée d'événements deux à deux incompatibles, la série* $\sum_{k\in\mathbb{N}}\mathrm{P}(\mathrm{X}=k)$ *est elle aussi convergente. Il vient alors*
$$\tilde{\mathrm{R}}_n=\sum_{k=n+1}^{+\infty}k\mathrm{P}(\mathrm{X}=k)\geqslant\sum_{k=n+1}^{+\infty}(n+1)\mathrm{P}(\mathrm{X}=k)=(n+1)\sum_{k=n+1}^{+\infty}\mathrm{P}(\mathrm{X}=k)=(n+1)\mathrm{P}(\mathrm{X}\geqslant n+1)=n\mathrm{R}_n\geqslant 0.$$
Mais $\tilde{\mathrm{R}}_n\xrightarrow[n\to+\infty]{}0$ *donc, par le théorème d'encadrement,* $(n+1)\mathrm{P}(\mathrm{X}\geqslant n+1)=n\mathrm{R}_n\xrightarrow[n\to+\infty]{}0$. *On en déduit que* $\sum_{k=0}^{n}\mathrm{P}(\mathrm{X}\geqslant k)$ *converge vers* $\mathrm{E}(\mathrm{X})$. *Donc* $\mathrm{E}(\mathrm{X})=\sum_{n=0}^{+\infty}\mathrm{P}(\mathrm{X}\geqslant n)$

Remarque 9.12 Avec les mêmes hypothèses, on a aussi

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

9.3.2 Propriétés de l'espérance

Le théorème suivant est très pratique car il permet de calculer l'espérance d'une variable aléatoire $f(X)$ sans connaître sa loi de probabilité, seule celle de X suffit. Sa démonstration est hors programme et on l'admettra.

THÉORÈME 9.20 ★ **Théorème de transfert**
Soit X une variable aléatoire discrète et soit f une application à valeurs réelles définie sur $X(\Omega)$. Alors la variable aléatoire $f(X)$ est d'espérance finie si et seulement si la la famille $\big(P(X=x_n)f(x_n)\big)$ est sommable. Dans ce cas, on a :

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

Démonstration *Ce théorème est admis.*

Remarque 9.13 Cette formule s'applique aux couples et aux n-uplets de variables aléatoires.

PROPOSITION 9.21 ★ **Linéarité de l'espérance**

1. L'ensemble $\mathscr{E}$ des variables aléatoires discrètes de Ω dans $\mathbb{R}$ dont l'espérance est finie est un espace vectoriel.
2. L'application espérance

$$E : \left\{ \begin{array}{ccc} \mathscr{E} & \longrightarrow & \mathbb{R} \\ X & \longmapsto & E(X) \end{array} \right.$$

est une forme linéaire sur $\mathscr{E}$

Démonstration *La preuve est admise*

PROPOSITION 9.22 ★ **Positivité, croissance de l'espérance**

1. L'application espérance E est positive dans le sens où si $X \in \mathscr{E}$ est telle que $X(\Omega) \subset \mathbb{R}_+$ alors $E(X) \geqslant 0$.
2. L'application espérance est croissante dans le sens où si $X, Y \in \mathscr{E}$ sont telles que $\forall \omega \in \Omega, \quad X(\omega) \leqslant Y(\omega)$ alors $E(X) \leqslant E(Y)$.

Démonstration

1. *Si X est à valeurs positives alors $E(X)$ est la somme d'une série à termes positifs. Alors $E(X)$ est positive.*
2. *Si X et Y vérifient la propriété indiquée, alors $Y-X$ est à valeurs dans $\mathbb{R}_+$. Par linéarité de l'application espérance, $E(X-Y)$ est finie, égale à $E(X)-E(Y)$ et positive d'après le point précédent. On en déduit que $E(X) \leqslant E(Y)$.*

On a aussi le théorème suivant :

_esperance

THÉORÈME 9.23 ★ **Critère de majoration**
Si X et Y sont deux variables aléatoires réelles discrètes telles que :

- H1 $|X| \leqslant Y$;
- H2 Y admet une espérance finie

alors $E(Y) \in \mathbb{R}_+$ et X a une espérance finie. De plus,

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

Démonstration *La preuve est admise*

THÉORÈME 9.24 ★★★ **Variable aléatoire d'espérance nulle**
Soit X une variable aléatoire $\boxed{\text{positive}}$. Alors X est d'espérance nulle si et seulement si l'événement $(X=0)$ est presque sûr.

$$\boxed{X \geqslant 0 \implies \big[E(X)=0 \iff X=0 \text{ ps}\big]}.$$

Démonstration

$\boxed{\Rightarrow}$ *Supposons que* X *est d'espérance nulle. Comme pour tout* $x \in X(\Omega)$, *on a* $0 \leqslant x P(X=x) \leqslant E(x)$, *il vient que* $x P(X=x) = 0$ *et donc* $P(X=x) = 0$ *si* $x \neq 0$. *On en déduit que*

$$P(X=0) = 1 - P(X \neq 0) = 1 - \sum_{x \in X(\Omega) \setminus \{0\}} P(X=x) x = 1$$

et $X = 0$ *presque sûrement.*

$\boxed{\Leftarrow}$ *Réciproquement, si* $(X=0)$ *est presque sûr alors pour tout* $x \in X(\Omega) \setminus \{0\}$, *on a* $P(X=x) = 0$ *et donc* $x P(X=x) = 0$. *On en déduit que* $E(X) = 0$.

La proposition suivante, qu'on admettra, est très utile en pratique.

PROPOSITION 9.25 ★ **Espérance du produit de deux variables aléatoires indépendantes**

Si X et Y sont deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ admettant des espérances finies alors

$$\boxed{X \text{ et } Y \text{ indépendantes} \implies [XY \text{ admet une espérance finie et } E(XY) = E(X)E(Y)]}.$$

Démonstration *La preuve est hors programme.*

De manière plus générale, on a la proposition suivante :

PROPOSITION 9.26 ★ **Espérance d'un produit de variables aléatoires mutuellement indépendantes**

Si $X_1, \ldots, X_n$ sont n variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ admettant des espérances finies alors

$$\boxed{X_1, \ldots, X_n \text{ mutuellement indépendantes} \implies [X_1 \ldots X_n \text{ admet une espérance finie et } E(X_1 \ldots X_n) = E(X_1) \ldots E(X_n)]}.$$

Démonstration *Hors programme On effectue une récurrence sur* $n \geqslant 2$. *La propriété est vraie au rang 2 d'après la proposition précédente. On la suppose vraie au rang* n *et on la montre au rang* $n+1$. *Soient* $X_1, \ldots, X_n, X_{n+1}$ *des variables aléatoires réelles discrètes admettant des espérances finies. Comme ces* $n+1$ *variables aléatoires sont mutuellement indépendantes, les variables aléatoires* $Y_n = X_1 \ldots X_n$ *et* X_{n+1} *sont indépendantes. D'après l'hypothèse de récurrence,* Y_n *admet une espérance finie donnée par* $E(Y_n) = E(X_1) \ldots E(X_n)$. *Alors d'après la proposition précédente,* $Y_n X_{n+1}$ *admet une espérance finie donnée par* $E(Y_n X_{n+1} = E(Y_n) E(X_{n+1})$. *Alors* $X_1 \ldots X_{n+1}$ *admet une espérance finie donnée par* $E(X_1 \ldots X_{n+1}) = E(X_1) \ldots E(X_{n+1})$ *ce qui prouve la proposition par récurrence.*

THÉORÈME 9.27 ★ **Inégalité de Markov**

Soit X une variable aléatoire réelle discrète *positive* admettant une espérance finie. Alors on a *l'inégalité de Markov* :

$$\forall a > 0, \quad \boxed{P(X \geqslant a) \leqslant \frac{E(x)}{a}}.$$

Démonstration *On note* $X(\Omega) = \{x_k \mid k \in \mathbb{N}\} \subset \mathbb{R}$. *Notons* $K_{<a} = \{k \in \mathbb{N} \mid x_k < a\}$ *et* $K_{\geqslant a} = \{k \in \mathbb{N} \mid x_k \geqslant a\}$. *Remarquons que pour tout* $n \in \mathbb{N}$,

$$\sum_{k \in K_{<a} \cap [\![0,n]\!]} x_k P(X = x_k) \leqslant \sum_{k \in [\![0,n]\!]} x_k P(X = x_k).$$

De plus, $\sum_{k \in \mathbb{N}} x_k P(X = x_k)$ *est convergente et à terme général positif. Donc* $\sum_{k \in [\![0,n]\!]} x_k P(X = x_k)$ *est bornée. Mais alors* $\sum_{k \in K_{<a} \cap [\![0,n]\!]} x_k P(X = x_k)$ *est aussi bornée et comme son terme général est positif, la série* $\sum_{k \in K_{<a}} x_k P(X = x_k)$ *est convergente. On montre de même que* $\sum_{k \in K_{\geqslant a}} x_k P(X = x_k)$ *est convergente et en passant aux sommes partielles que*

$$E(X) = \sum_{k=0}^{+\infty} x_k P(X = x_k) = \sum_{k \in K_{<a}} x_k P(X = x_k) + \sum_{k \in K_{\geqslant a}} x_k P(X = x_k)$$

On montre aussi que, comme X *est positive,* $\sum_{k \in K_{<a}} x_k P(X = x_k) \geqslant 0$ *et donc que*

$$E(X) \geqslant \sum_{k \in K_{\geqslant a}} x_k P(X = x_k).$$

Il vient alors

$$E(X) \geqslant \sum_{k \in K_{\geqslant a}} x_k P(X = x_k) \geqslant a \sum_{k \in K_{\geqslant a}} P(X = x_k) = a P(X \geqslant a)$$

ce qui prouve l'inégalité.

9.3.3 Variance et covariance

DÉFINITION 9.18 ★ **Moment d'ordre m**
Soit X une variable aléatoire réelle discrète. Si X^m admet une espérance finie alors on dit que X admet un *moment d'ordre m*. Celui-ci est donnée par $E(X^m)$.

Remarque 9.14 Dire que X admet un moment d'ordre 1 revient à dire que X a une espérance finie.

aleatoires

PROPOSITION 9.28 ★ **Inégalité de Cauchy-Schwarz**
Soient X et Y deux variables aléatoires réelles discrètes admettant des moments d'ordre 2 alors

1. XY est d'espérance finie
2. On a l'inégalité de Cauchy-Schwarz :

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

Démonstration *On sait que pour tout $(x,y) \in \mathbb{R}^2$, $|xy| \leqslant \frac{1}{2}(x^2+y^2)$. Mais alors si X^2 et Y^2 ont des espérances finies, il en est de même de la combinaison linéaire $\frac{1}{2}(X^2+Y^2)$ et le théorème 9.23 permet alors d'affirmer que XY a aussi une espérance finie.*
Il s'ensuit alors que, si X^2 et Y^2 ont des espérances finies, pour tout $\lambda \in \mathbb{R}$, $(\lambda X+Y)^2 = \lambda^2 X^2 + 2\lambda XY + Y^2$ a une espérance finie comme combinaison linéaire de variables aléatoires admettant des espérances finies.
Alors si $Q(\lambda) = E\left((\lambda X+Y)^2\right)$, on a pour tout $\lambda \in \mathbb{R}$, $Q(\lambda) \geqslant 0$.
Mais, si $E(X^2) \neq 0$, $Q(\lambda) = \lambda^2 E(X^2) + 2\lambda E(XY) + E(Y^2)$ est un trinôme du second degré en λ et on doit avoir pour son discriminant réduit $\delta \leqslant 0$. Comme $\delta = (E(XY))^2 - E(X^2)E(Y^2)$, on obtient que $|E(XY)| \leqslant \sqrt{E(X^2)}\sqrt{E(Y^2)}$.
Par ailleurs, si $E(X^2) = 0$ alors $Q(\lambda) = 2\lambda E(XY) + E(Y^2) \geqslant 0$ ce qui n'est possible que si $E(XY) = 0$ et l'inégalité est encore vérifiée.

COROLLAIRE 9.29 ★ **Si X a un moment d'ordre 2 alors X a un moment d'ordre 1**
Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2 alors X admet un moment d'ordre 1, c'est-à-dire une espérance finie.

Démonstration *On applique le théorème précédent à $Y = 1$ qui admet aussi un moment d'ordre 2. On obtient que $X = XY$ admet un moment d'ordre 1 et que $|E(X)| \leqslant \sqrt{E(X^2)}$.*

rdre_2_sev

COROLLAIRE 9.30 ★ **L'ensemble des variables aléatoires admettant un moment d'ordre 2 est un espace vectoriel**
L'ensemble des variables aléatoires réelles discrètes admettant un moment d'ordre 2 est un sous-espace vectoriel de l'espace vectoriel des variables aléatoires réelles discrètes admettant un moment d'ordre 1.

Démonstration *Le corollaire précédent nous permet d'affirmer que l'ensemble des variables aléatoires admettant un moment d'ordre 2 est un sous-ensemble du $\mathbb{R}$-espace vectoriel des variables aléatoires admettant un moment d'ordre 1.*
Ensuite, l'ensemble des variables aléatoires admettant des moments d'ordre 2 n'est pas vide, il contient par exemple la variable aléatoire nulle. De plus, si X et Y admettent des moments d'ordre 2 alors $(X+Y)^2 = X^2 + 2XY + Y^2$. D'après le théorème précédent, XY admet un moment d'ordre 1 donc, par combinaison linéaire, il en est de même de la somme et X + Y admet bien un moment d'ordre 2. Enfin, si $\alpha \in \mathbb{R}$ et si X admet un moment d'ordre 2, αX admet aussi, par linéarité de l'espérance, un moment d'ordre 2.

DÉFINITION 9.19 ★★★ **Variance et écart-type**
Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On définit alors la *variance* de X, notée V(X), par :

$$V(X) = E\left((X - E(X))^2\right)$$

et l'*écart type* de X, noté $\sigma(X)$, par :

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

Remarque 9.15

— La définition de la variance a bien un sens car comme X admet un moment d'ordre 2, elle admet un moment d'ordre 1 et E(X) est bien définie. De plus, la variable aléatoire réelle discrète $(X - E(X))^2 = X^2 - 2XE(X) + (E(X))^2$ est une combinaison linéaire de variables aléatoires réelles discrètes admettant des moments d'ordre 1 donc elle admet un moment d'ordre 1.

— Comme $(X - E(X))^2 \geqslant 0$, il en est de même de son espérance et l'écart type $\sigma(X)$ est lui aussi bien définit.

— Si $E(X) = m$, $X(\Omega) = \left\{x_k \mid k \in \mathbb{N}\right\}$ et que X admet un moment d'ordre 2 alors $V(X) = \sum_{k=0}^{+\infty} (x_k - m)^2 P(X = x_k)$.

THÉORÈME 9.31 ★★★ **Formule de König-Huygens**
Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors :

1. On a la formule de König-Huygens :
$$\boxed{\mathrm{V}(\mathrm{X}) = \mathrm{E}(\mathrm{X}^2) - (\mathrm{E}(\mathrm{X}))^2};$$
2. Pour tout $(a,b) \in \mathbb{R}^2$,
$$\boxed{\mathrm{V}(a\mathrm{X}+b) = a^2\mathrm{V}(\mathrm{X})}.$$

Démonstration *Comme X admet un moment d'ordre 2, elle admet un moment d'ordre 1 qu'on note m.*

1. *Il vient alors, par linéarité de l'espérance :*
$$\mathrm{V}(\mathrm{X}) = \mathrm{E}\left((\mathrm{X}-m)^2\right) = \mathrm{E}\left(\mathrm{X}^2 - 2m\mathrm{X} + m^2\right) = \mathrm{E}\left(\mathrm{X}^2\right) - 2m\mathrm{E}(\mathrm{X}) + m^2 = \mathrm{E}\left(\mathrm{X}^2\right) - m^2$$
d'où le résultat.
2. *De même, en utilisant la formule de König-Huygens et le fait que $\mathrm{E}(a\mathrm{X}+b) = a\mathrm{E}(\mathrm{X})+b$ pour $a,b \in \mathbb{R}$, on a que :*
$$\begin{aligned}\mathrm{V}(a\mathrm{X}+b) &= \mathrm{E}\left((a\mathrm{X}+b)^2\right) - (\mathrm{E}(a\mathrm{X}+b))^2 = \mathrm{E}\left(a^2\mathrm{X}^2 + 2ab\mathrm{X} + b^2\right) - (a\mathrm{E}(\mathrm{X})+b)^2 \\ &= a^2\mathrm{E}(\mathrm{X}^2) + 2abm + b^2 - a^2(\mathrm{E}(\mathrm{X}))^2 - 2abm + b^2 = a^2\left(\mathrm{E}(\mathrm{X}^2) - (\mathrm{E}(\mathrm{X}))^2\right) = a^2\mathrm{V}(\mathrm{X})\end{aligned}$$

DÉFINITION 9.20 ★ **Variable réduite, variable centrée réduite**
Soit X une variable aléatoire discrète admettant un moment d'ordre 2 et telle que $\sigma(\mathrm{X}) > 0$.

1. La variable aléatoire $\frac{\mathrm{X}}{\sigma(\mathrm{X})}$ a un écart-type égal à 1 est appelée *variable réduite associée à* X.
2. La variable aléatoire $\frac{\mathrm{X}-\mathrm{E}(\mathrm{X})}{\sigma(\mathrm{X})}$ a une espérance nulle et un écart-type égal à 1 est est appelée *variable centrée réduite associée à* X.

THÉORÈME 9.32 ★ **Inégalité de Bienaymé-Tchebychev**
Soit X une variable aléatoire discrète admettant un moment d'ordre 2 alors on a l'inégalité de Bienaymé-Tchebychev :
$$\forall \varepsilon > 0, \quad \boxed{\mathrm{P}\left(|\mathrm{X}-\mathrm{E}(\mathrm{X})| \geqslant \varepsilon\right) \leqslant \frac{\mathrm{V}(\mathrm{X})}{\varepsilon^2}}.$$

Démonstration *Comme X admet un moment d'ordre 2, elle admet un moment d'ordre 1 qu'on note m. On applique alors l'inégalité de Markov à la variable aléatoire discrète positive $\mathrm{Y} = |\mathrm{X}-m|$. Remarquons que $\mathrm{E}(\mathrm{Y}^2) = \mathrm{V}(\mathrm{X})$. Il vient alors pour tout $\varepsilon > 0$,*
$$\mathrm{P}\left(|\mathrm{X}-\mathrm{E}(\mathrm{X})| \geqslant \varepsilon\right) = \mathrm{P}\left(\mathrm{Y}^2 \geqslant \varepsilon^2\right) \leqslant \frac{\mathrm{E}(\mathrm{Y}^2)}{\varepsilon^2} = \frac{\mathrm{V}(\mathrm{X})}{\varepsilon^2}.$$

DÉFINITION 9.21 ★ **Covariance de deux variables aléatoires**
Soient X et Y deux variables aléatoires réelles discrètes admettant un moment d'ordre 1 (c'est-à-dire d'espérance finie). On appelle alors *covariance de* X *et* Y le nombre réel noté $\mathrm{cov}(\mathrm{X},\mathrm{Y})$ et donné, si elle existe, par :
$$\mathrm{cov}(\mathrm{X},\mathrm{Y}) = \mathrm{E}\left([\mathrm{X}-\mathrm{E}(\mathrm{X})]\,[\mathrm{Y}-\mathrm{E}(\mathrm{Y})]\right).$$

Remarque 9.16 La covariance de deux variables aléatoires réelles discrètes admettant un moment d'ordre 2 est bien définie par le théorème 9.28.

PROPOSITION 9.33 ★ **Existence de la covariance**
covariance
Soient X et Y deux variables aléatoires réelles discrètes admettant un moment d'ordre 2 alors la covariance de X et Y existe et on a :
$$\boxed{\mathrm{cov}(\mathrm{X},\mathrm{Y}) = \mathrm{E}(\mathrm{XY}) - \mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{Y})}.$$

Démonstration *Comme* X *et* Y *admettent des moments d'ordre* 2*, elles admettent des moments d'ordre* 1 *et on a*

$$(X - E(X))(Y - E(Y)) = XY - XE(Y) - YE(X) + E(X)E(Y)$$

et d'après le théorème 9.28, XY *admet un moment d'ordre* 1*. Comme l'ensemble des variables aléatoires d'espérance finie est un* $\mathbb{R}$*-espace vectoriel,* $(X - E(X))(Y - E(Y))$ *admet un moment d'ordre* 1 *et* $\operatorname{cov}(X, Y)$ *est bien définie. La formule mentionnée est alors une conséquence directe de la linéarité de l'espérance.*

Il découle alors facilement des résultats précédents la proposition suivante :

PROPOSITION 9.34 ★ **Règles de calcul pour la covariance**
Soient X, X', Y, Y' des variables aléatoires réelles discrètes admettant un moment d'ordre 2 et soient $a, b, c, d \in \mathbb{R}$. Alors :

1. $\operatorname{cov}(Y, X) = \operatorname{cov}(X, Y)$;
2. $\operatorname{cov}(X, X) = V(X)$;
3. $\operatorname{cov}(aX + b, cY + d) = ac\operatorname{cov}(X, Y)$;
4. $\operatorname{cov}\left(aX + bX', cY + dY'\right) = ac\operatorname{cov}(X, Y) + bc\operatorname{cov}\left(X', Y\right) + ad\operatorname{cov}\left(X, Y'\right) + bd\operatorname{cov}\left(X', Y'\right)$.

Remarque 9.17 On tire des règles de calcul précédentes que la covariance est une forme bilinéaire symétrique positive sur l'espace des variables aléatoires réelles discrètes admettant un moment d'ordre 2.

PROPOSITION 9.35 ★ **La covariance de deux variables aléatoires réelles discrètes indépendantes est nulle**
Soient X et Y deux variables aléatoires réelles discrètes *indépendantes* admettant un moment d'ordre 2 alors $\operatorname{cov}(X, Y) = 0$.

Démonstration *Comme* X *et* Y *sont indépendantes,* $E(XY) = E(X)E(Y)$ *et la formule de la proposition 9.33 permet de conclure en la nullité de leur covariance.*

DÉFINITION 9.22 ★ **Variables aléatoires non corrélées**
Deux variables aléatoires réelles discrètes admettant une covariance nulle sont dites *non corrélées*

PROPOSITION 9.36 ★ **Variance de la somme de deux variables aléatoires**
Soient X et Y deux variables aléatoires réelles discrètes admettant un moment d'ordre 2. Alors :

1. $(X + Y)^2$ admet une espérance finie et
$$\boxed{V(X + Y) = V(X) + V(Y) + 2\operatorname{cov}(X, Y)} ;$$
2. Si X et Y sont indépendantes,
$$\boxed{V(X + Y) = V(X) + V(Y)}.$$

Démonstration

1. *Comme* $(X + Y)^2 = X^2 + 2XY + Y^2$ *et que* X *et* Y *admettent des moments d'ordre deux alors, d'après le théorème 9.28, il vient que* XY *admet un moment d'ordre* 1 *et, par structure d'espace vectoriel, que* $(X + Y)^2$ *admet une espérance finie. On peut alors considérer la variance de* $X + Y$ *et, en utilisant la formule de König-Huygens et en notant* $m = E(X)$*,* $m' = E(Y)$*, on obtient que :*
$$V(X + Y) = E\left((X + Y)^2\right) - (E(X + Y))^2 = \left[E(X^2) - m^2\right] + \left[E(Y^2) - m'^2\right] + 2\left[E(XY) - mm'\right] = V(X) + V(Y) + 2\operatorname{cov}(X, Y)$$
2. *Si* X *et* Y *sont indépendantes, alors* $\operatorname{cov}(X, Y) = 0$ *d'où la formule.*

Remarque 9.18 Ce résultat se généralise facilement à une somme de n variables aléatoires $X_1, \dots, X_n$ admettant des moments d'ordre 2 :

$$V(X_1, \dots, X_n) = \sum_{i=1}^{n} V(X_i) + 2\sum_{1 \leqslant i < j \leqslant n} \operatorname{cov}(X_i, X_j)$$

et si les n variables aléatoires $X_1, \dots, X_n$ sont deux à deux indépendantes (ou plus généralement si celles-ci ne sont pas deux à deux corrélées), cette formule se réduit à :

$$V(X_1, \dots, X_n) = \sum_{i=1}^{n} V(X_i).$$

Définition 9.23 ★ **Coefficient de corrélation linéaire**
Soient X et Y deux variables aléatoires réelles discrètes admettant un écart-type non nul. Le *coefficient de corrélation linéaire de* X *et* Y est le réel, noté $\rho(X,Y)$ et donné par

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma(X)\sigma(Y)}.$$

Proposition 9.37
Soient X et Y deux variables aléatoires réelles discrètes admettant un moment d'ordre 2. Alors :

1. $|\operatorname{cov}(X,Y)| \leqslant \sigma(X)\sigma(Y)$;
2. Si de plus les écarts-types de X et Y sont non nuls alors $\rho(X,Y) \in [-1,1]$.

Démonstration *Comme* X *et* Y *admettent des moments d'ordre* 2*, il en est de même, d'après le 9.30, de* $tX+Y$ *où* $t \in \mathbb{R}$*. Alors on considère* $Q(t) = V(tX+Y) = t^2V(X) + 2t\operatorname{cov}(X,Y) + V(Y) \geqslant 0$.

- *Si* $V(X) \neq 0$ *alors* $Q(t)$ *est un trinôme positif du second degré en* t*, son discriminant est donc négatif ou nul d'où l'inégalité.*
- *Si* $V(X) = 0$*, alors* $Q(t)$ *est une fonction affine et positive ce qui n'est possible que si* $\operatorname{cov}(X,Y) = 0$ *et l'inégalité est encore vraie dans ce cas.*

Le fait que $\rho(X,Y) \in [-1,1]$ *quand les écarts-types de* X *et* Y *sont non nuls découle directement de cette inégalité.*

Remarque 9.19 On a $\rho(X,Y) = \pm 1$ si et seulement si il existe $t_0 \in \mathbb{R}$ tel que $Q(t_0) = V(t_0X+Y) = 0$. Mais d'après l'exercice **??** page **??**, ceci est équivalent à l'existence de $t_1 \in \mathbb{R}$ tel que $P(t_0X+Y = t_1) = 1$. Alors $\rho(X,Y) = \pm 1$ si et seulement si l'événement il existe deux réels a et b tels que l'événement $Y = aX+b$ est presque certain (c'est-à-dire de probabilité 1).

9.4 Fonctions génératrices

Définition 9.24 ★★★ **Fonction génératrice d'une variable aléatoire à valeurs dans $\mathbb{N}$**
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On définit sa *fonction génératrice* G_X par :

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

pour les $t \in \mathbb{R}$ tels que la variable aléatoire t^X admet une espérance finie.

Remarque 9.20 Comme $X(\Omega) = \mathbb{N}$, si $Y = t^X$ alors $Y(\Omega) = \left\{t^{X(\omega)} \mid \omega \in \Omega\right\} \subset \left\{t^n \mid n \in \mathbb{N}\right\}$ donc $E(Y) = \sum_{\omega\in\Omega} Y(\omega)P(\{\omega\}) = \sum_{n\in\mathbb{N}} t^n P(X=n)$ car si $n \notin X(\Omega)$ alors $P(X=n) = 0$.
Si t est fixé, t^X est une variable aléatoire réelle dont l'espérance est obtenue par le théorème de transfert.

Théorème 9.38 ★ **Propriétés des fonctions génératrices**
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

1. La série entière définissant G_X a un rayon de convergence supérieur ou égal à 1.
2. La série entière définissant G_X converge normalement sur tout segment inclus dans $]-R,R[$ et converge aussi au point $t=1$. De plus, $\lim_{t\to 1^-} G_X(t) = 1$.
3. La fonction G_X est continue de classe $\mathscr{C}^\infty$ sur $]-R,R[$.
4. $\forall n \in \mathbb{N}$, $\boxed{P(X=n) = \dfrac{G_X^{(n)}(0)}{n!}}$. Autrement dit, la fonction G_X caractérise la loi de X.

Démonstration *La famille* $(X=k)_{k\in\mathbb{N}}$ *forme un système complet d'événements donc par axiome d'additivité dénombrable, la série* $\sum_{k\in\mathbb{N}} P(X=k)$ *converge et a pour somme* 1*. On en déduit que la série entière définissant* G_X *a un rayon de convergence* $\geqslant 1$.
On pose pour $n \in \mathbb{N}$, $f_n : \left\{\begin{array}{ccc} [0,1] & \longrightarrow & \mathbb{R} \\ t & \longmapsto & P(X=n)\,t^n \end{array}\right.$*. On a* $\|f_n\|_{\infty,[0,1]} = P(X=n)$ *et* $\sum_{n\in\mathbb{N}} P(X=n)$ *qui est convergente. Donc* $\sum_{n\in\mathbb{N}} f_n$ *est normalement convergente et donc uniformément convergente sur* $[0,1]$*. De plus, pour tout* $n \in \mathbb{N}$, $f_n \in \mathscr{C}^0([0,1])$*. On en déduit par critère de contuité sur séries de fonctions que* G_X *est continue sur* $[0,1]$*. En particulier, on a* $\lim_{t\to 1^-} G_X(t) = 1$.
Le reste du théorème est une conséquence des résultats du chapitre sur les séries entières.

Lemme 9.39 ★ Limite au bord du disque de convergence d'une série entière à terme général positif

Soit $\sum a_n t^n$ une série entière de rayon de convergence $R = 1$ et de somme S. On suppose que :

(H1) $\forall n \in \mathbb{N}, \quad a_n \geqslant 0$

alors on a équivalence entre :

1. La série $\sum a_n$ converge.
2. La somme S admet une limite finie à gauche en 1.

Dans ce cas, $\boxed{\lim_{t\to 1^-} S(t) = \sum_{n=0}^{+\infty} a_n}$.

Démonstration *Remarquons que S est croissante sur $[0,1[$. Elle admet alors, d'après le théorème de la limite monotone, une limite l quand $t \to 1^-$ finie ou infinie .*
Si $\sum a_n$ converge alors comme pour tout $n \in \mathbb{N}$ et $t \in [0,1]$, $a_n t^n \leqslant a_n$, $\sum a_n t^n$ converge normalement sur $[0,1]$ et sa somme est donc continue sur $[0,1]$ ce qui permet d'écrire que $l = \sum_{n=0}^{+\infty} a_n$.
Sinon, comme $\sum_{n=0}^{N} a_n t^n \leqslant S(t) \leqslant l$ pour tous $t \in [0,1[$ et $N \in \mathbb{N}$ alors en passant à la limite quand $t \to 1^-$, il vient $\sum_{n=0}^{N} a_n \leqslant l$ et en faisant $N \to +\infty$, on a alors $l = +\infty$

Proposition 9.40 ★ Espérance et variance de X

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Alors :

1. La variable aléatoire X admet une espérance finie si et seulement si G_X est dérivable à gauche en 1^-. Dans ce cas,
$$\boxed{E(X) = \lim_{t\to 1^-} G'_X(t)}.$$
2. La variable aléatoire X^2 admet une espérance finie (et donc X admet une variance finie) si et seulement si G_X est deux fois dérivable à gauche en 1^-. Dans ce cas,
$$\boxed{E(X(X-1)) = \lim_{t\to 1^-} G''_X(t)} \quad \text{et} \quad \boxed{V(X) = \lim_{t\to 1^-} G''_X(t) + \lim_{t\to 1^-} G'_X(t) - \left(\lim_{t\to 1^-} G'_X(t)\right)^2}.$$
3. Si le rayon de convergence R de la série entière définissant G_X est > 1 alors G_X est dérivable en 1 et ces formules deviennent :
$$\boxed{E(X) = G'(1)}, \quad \boxed{E(X(X-1)) = G''_X(1)} \quad \text{et} \quad \boxed{V(X) = G''_X(1) + G'_X(1) - \left(G'_X(1)\right)^2}$$

Démonstration *Cette démonstration n'est pas exigible.*

- *Si la série entière définissant G_X a un rayon de convergence $R > 1$ alors comme G_X est de classe $\mathscr{C}^\infty$ sur $]-R,R[$ et est en particulier dérivable deux fois en 1. Par théorème de dérivation terme à terme pour les séries entières, il vient que, pour tout $t \in]-R,R[$, on a $G'(t) = \sum_{n=0}^{+\infty} n t^{n-1} P(X=n)$ et $G''_X(t) = \sum_{n=0}^{+\infty} n(n-1) t^{n-2} P(X=n)$. Les séries $\sum nP(X=n)$ et $\sum n(n-1)P(X=n)$ sont absolument convergentes sur $]-R,R[$ et les variables aléatoires X et $X(X-1)$ admettent donc une espérance finie. L'existence d'un espérance pour $X(X-1)$ étant une conséquence du théorème de transfert. Il s'ensuit alors, par linéarité de l'espérance, que X^2 admet aussi une espérance finie. De plus, $G'(1) = \sum_{n\geqslant 0}^{+\infty} nP(X=n) = E(X)$ et $G''(1) = \sum_{n=0}^{+\infty} n(n-1)P(X=n) = E(X(X-1))$ d'où la formule pour V(X).*
- *Si la série entière définissant G_X a un rayon de convergence $R = 1$ alors nous allons avoir besoin du lemme précédent qu'on applique tout d'abord à la série entière $\sum nP(X=n)t^{n-1}$ pour le premier résultat et à la série entière $\sum n(n-1)P(X=n)t^{n-2}$ pour le second.*

Proposition 9.41 ★ Fonction génératrice et somme

Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que :

(H1) X et Y sont indépendantes.

Alors :

$$\forall t \in]-1,1[, \quad \boxed{G_{X+Y}(t) = G_X(t)G_Y(t)}.$$

De manière plus générale, si $X_1, \dots, X_n$ sont n variables aléatoires à valeurs dans $\mathbb{N}$ mutuellement indépendantes alors :

$$\forall t \in]-1,1[, \quad \boxed{G_{X_1+\dots+X_n}(t) = G_{X_1}(t)\dots G_{X_n}(t)}.$$

Démonstration *On a*

$$G_{X+Y}(t) = E\left(t^{X+Y}\right) = E\left(t^X t^Y\right) = E\left(t^X\right)E\left(t^Y\right) = G_X(t)G_Y(t)$$

l'avant dernière égalité provenant du fait que X *et* Y *sont indépendantes et donc que pour toute fonction* $f : \mathbb{N} \to \mathbb{R}$, $f \circ X$ *et* $f \circ Y$ *sont aussi indépendantes.*
La formule générale est une conséquence du lemme des coalitions.

9.5 Lois discrètes usuelles

Nous présentons dans cette section différentes lois de probabilité usuelles.

9.5.1 Loi uniforme

La loi uniforme servira quand on choisit au hasard un objet parmi n. On numérote en effet ces objets de 1 à n et X désigne le numéro de l'objet choisit.

DÉFINITION 9.25 ★ **Loi uniforme**
Soit $n \in \mathbb{N}^*$. On dit que la variable aléatoire X suit la *loi uniforme sur* $[\![1,n]\!]$ si et seulement si X peut prendre les valeurs 1, 2, …, n avec la probabilité $\frac{1}{n}$ ou autrement dit, si et seulement si

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

On écrit alors $X \sim \mathcal{U}([\![1,n]\!])$.

Remarque 9.21 La loi uniforme $\mathcal{U}$ est bien une loi de probabilité car $\sum_{k\in[\![1,n]\!]} \mathcal{U}(k) = \sum_{k\in[\![1,n]\!]} \frac{1}{n} = 1$.

PROPOSITION 9.42 **Espérance et variance de la loi uniforme**
Soit X une variable aléatoire suivant la loi uniforme. Alors

$$\boxed{E(X) = \frac{n+1}{2}} \quad \text{et} \quad \boxed{V(X) = \frac{n^2-1}{12}}.$$

Démonstration *On a*

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

De même

$$E(X^2) = \sum_{k=1}^{n} k^2 P(X=k) = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

donc

$$V(X) = E(X^2) - E^2(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n+1}{12}\left(2(2n+1) - 3(n+1)\right) = \frac{n^2-1}{12}.$$

Remarque 9.22 La fonction génératrice de la loi uniforme est définie via « la série » :

$$G_X(t) = \sum_{k=1}^{n} t^k P(X=k) = \frac{1}{n}\sum_{k=1}^{n} t^k = \frac{t(1-t^n)}{n(1-t)}$$

sur le disque $]-1,1[$. On peut alors re-calculer l'espérance et la variance de la loi uniforme en utilisant sa fonction génératrice.

9.5.2 Loi de Bernoulli

La loi de Bernoulli code une expérience aléatoire ne présentant que deux issues comme par exemple le jeu de pile ou face.

DÉFINITION 9.26 ★ **Loi de Bernoulli**
Soit $0 < p < 1$. On dit que la variable aléatoire X suit la *loi de Bernoulli de paramètre* $p \in [0,1]$ si et seulement si X peut prendre les valeurs 0 où 1 avec, respectivement, les probabilités $q = 1 - p$ et p :

$$P_X : \left\{ \begin{array}{ccc} \{0,1\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \begin{cases} q & \text{si } x = 0 \\ p & \text{si } x = 1 \end{cases} \end{array} \right. .$$

On écrit $X \sim \mathscr{B}(p)$

Remarque 9.23 La loi de Bernoulli est bien une loi de probabilité. En effet, $P(\{0\}) + P(\{1\}) = p + (1-p) = 1$.

PROPOSITION 9.43 **Espérance et variance de la loi de Bernoulli**
Soit X une variable aléatoire suivant la loi de Bernoulli $\mathscr{B}(p)$. Alors

$$\boxed{E(X) = p} \qquad \text{et} \qquad \boxed{V(X) = pq}$$

où $q = 1 - p$.

Démonstration *On a*

$$E(X) = 0.q + 1.p = p$$

et comme $X^2 = X$, *il vient* $E(X^2) = E(X)$ *donc*

$$V(X) = E(X^2) - E^2(X) = E(X) - E^2(X) = p - p^2 = pq.$$

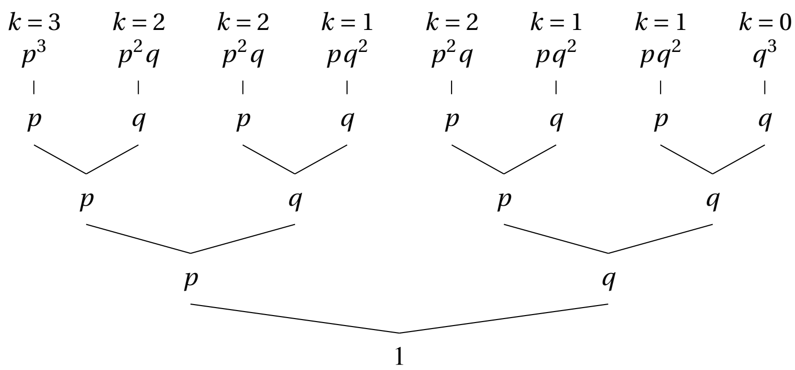
Remarque 9.24 La série génératrice de la loi de Bernoulli de paramètre p est donnée, pour tout $t \in \mathbb{R}$ par :

$$G_X(t) = t^0 q + t^1 p = q + tp.$$

On retrouve alors facilement les valeurs de l'espérance et de la variances calculées précédemment.

9.5.3 Loi binomiale

On effectue n épreuves de Bernoulli indépendantes (comme n tirages à pile ou face) avec comme probabilités de réussite $p \in [0,1]$ et d'échec $q = 1 - p$. On compte alors le nombre k de réussites à l'issue des n épreuves et on nomme X la variable aléatoire comptant ce nombre de succès. Cette variable aléatoire suit la loi binomiale que nous allons définir.
Cette loi peut être représentée par un arbre binaire.



On a $P(X = 0) = q^3$, $P(X = 1) = \binom{3}{1} pq^2$, $P(X = 2) = \binom{3}{2} p^2 q$ et $P(X = 3) = p^3$.
Elle permet par exemple de modéliser l'expérience consistant à effectuer n tirages avec remise d'une boule dans une urne contenant m boules blanches et n boules rouges et à compter le nombre de fois qu'une couleur fixée à l'avance sorte.

Définition 9.27 ★ Loi binomiale ou loi des tirages avec remise
Soient $p \in [0,1]$ et $n \in \mathbb{N}^*$. On dit que la variable aléatoire X suit la *loi de binomiale de paramètres p* si et seulement si X peut prendre les valeurs 0, 1, ..., n avec les probabilités $\mathrm{P}(\mathrm{X}=k) = \binom{n}{k} p^k q^{n-k}$ où $q = 1-p$. Sa loi est donc donnée par :

$$\mathrm{P_X} : \left\{ \begin{array}{ccc} [\![0,n]\!] & \longrightarrow & \mathbb{R} \\ k & \longmapsto & \dbinom{n}{k} p^k q^{n-k} \end{array} \right. .$$

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

Remarque 9.25 On a $\sum_{k=0}^n \mathrm{P_X}(k) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = \left(p+q\right)^n = 1$ d'après la formule du binôme donc on a bien défini une loi de probabilité.

Prouvons le résultat annoncé au début de cette section :

Proposition 9.44 ★ Une somme de variables de Bernoulli indépendantes et de même paramètre suit la loi binomiale
Soient $\mathrm{X}_1,\dots,\mathrm{X}_n$ n variables aléatoires indépendantes suivant la même loi de Bernoulli $\mathrm{X}_i \sim \mathscr{B}(p)$ alors leur somme suit la loi binomiale $\mathscr{B}(n,p)$: $\mathrm{X}_1+\ldots+\mathrm{X}_n \sim \mathscr{B}(n,p)$.

Démonstration *On effectue la preuve par une récurrence sur $n \geqslant 2$. Si $\mathrm{X}_1 \sim \mathscr{B}(p)$ et $\mathrm{X}_2 \sim \mathscr{B}(p)$ sont deux variables aléatoires de Bernoulli indépendantes alors, si $\mathrm{Y} = \mathrm{X}_1+\mathrm{X}_2$:*

- $\mathrm{P}(\mathrm{Y}=2) = \mathrm{P}\left((\mathrm{X}_1=1)\cap(\mathrm{X}_2=1)\right) = \mathrm{P}(\mathrm{X}_1=1)\mathrm{P}(\mathrm{X}_2=1) = p^2$;
- $\mathrm{P}(\mathrm{Y}=1) = \mathrm{P}\left([(\mathrm{X}_1=1)\cap(\mathrm{X}_2=0)]\cup[(\mathrm{X}_1=0)\cap(\mathrm{X}_2=1)]\right) = \mathrm{P}(\mathrm{X}_1=1)\mathrm{P}(\mathrm{X}_2=0)+\mathrm{P}(\mathrm{X}_1=0)\mathrm{P}(\mathrm{X}_2=1) = 2pq$;
- $\mathrm{P}(\mathrm{Y}=0) = \mathrm{P}\left((\mathrm{X}_1=0)\cap(\mathrm{X}_2=0)\right) = \mathrm{P}(\mathrm{X}_1=0)\mathrm{P}(\mathrm{X}_2=0) = q^2$.

Dans tous les cas, on a $\mathrm{P}(\mathrm{X}_1+\mathrm{X}_2=k) = \binom{2}{k}p^k q^{2-k}$ et la propriété est vraie au rang 2
On suppose la propriété vraie au rang $n-1$ pour $n \geqslant 3$ et on la prouve au rang n. On considère n variables aléatoires indépendantes suivant une même loi de Bernoulli $\mathscr{B}(p)$. Soit $k \in [\![0,n]\!]$. Il est clair que $\mathrm{P}(\mathrm{X}_1+\ldots+\mathrm{X}_n=n) = \mathrm{P}\left((\mathrm{X}_1=1)\cap\ldots\cap(\mathrm{X}_n=1)\right) = \prod_{k=1}^n \mathrm{P}(\mathrm{X}_k=1) = p^n$. De même, $\mathrm{P}(\mathrm{X}_1+\ldots+\mathrm{X}_n=0) = q^n$. Soit $k \in]1,n-1[$. On a alors, avec $\mathrm{Y} = \mathrm{X}_1+\ldots+\mathrm{X}_n$) :

$$\begin{aligned}
\mathrm{P}(\mathrm{Y}=k) &= \mathrm{P}\left([(\mathrm{X}_1+\ldots+\mathrm{X}_{n-1}=k)\cap(\mathrm{X}_n=0)]\cup[(\mathrm{X}_1+\ldots+\mathrm{X}_{n-1}=k-1)\cap(\mathrm{X}_n=1)]\right) \\
&= \mathrm{P}\left([(\mathrm{X}_1+\ldots+\mathrm{X}_{n-1}=k)\cap(\mathrm{X}_n=0)]\right)+\mathrm{P}\left([(\mathrm{X}_1+\ldots+\mathrm{X}_{n-1}=k-1)\cap(\mathrm{X}_n=1)]\right) \\
&= \mathrm{P}\left(\mathrm{X}_1+\ldots+\mathrm{X}_{n-1}=k\right)\mathrm{P}\left(\mathrm{X}_n=0\right)+\mathrm{P}\left(\mathrm{X}_1+\ldots+\mathrm{X}_{n-1}=k-1\right)\mathrm{P}\left(\mathrm{X}_n=1\right) \\
&= \binom{n-1}{k}p^k q^{n-1-k}q+\binom{n-1}{k-1}p^{k-1}q^{n-1-k+1} \\
&= \left(\binom{n-1}{k}+\binom{n-1}{k-1}\right)p^k q^{n-k} \\
&= \binom{n}{k}p^k q^{n-k}
\end{aligned}$$

et la propriété est prouvée par récurrence.

Proposition 9.45 ★ Espérance et variance de la loi binomiale
Soit X une variable aléatoire suivant la loi binomiale $\mathscr{B}(p)$. Alors

$$\boxed{\mathrm{E}(\mathrm{X}) = np}, \qquad \boxed{\mathrm{V}(\mathrm{X}) = npq} \qquad \text{et} \qquad \boxed{\mathrm{G_X}(t) = (q+pt)^n}$$

où $q = 1-p$.

Démonstration *Pour tout $t \in \mathbb{R}$,*

$$\mathrm{G_X}(t) = \sum_{k=0}^n t^k \mathrm{P}(\mathrm{X}=k) = \sum_{k=0}^n \binom{n}{k}(pt)^k q^{n-k} = \left(q+pt\right)^n$$

par la formule du binôme. Nous proposons maintenant trois méthodes de calculs différentes de l'espérance et de la variance de X.

1. *Comme X est la somme de n variables aléatoires deux à deux indépendantes suivant une même loi de Bernoulli de paramètre p, on a :*

$$E(X) = E(X_1 + \ldots + X_n) = E(X_1) + \ldots + E(X_n) = np$$
$$V(X) = V(X_1 + \ldots + X_n) = V(X_1) + \ldots + V(X_n) = npq$$

2. *On utilise dans la seconde méthode la série génératrice de la loi binomiale et en particulier que $E(X) = G_X'(1) = np$. Par ailleurs $E(X(X-1)) = G_X''(1) = n(n-1)p^2$ et $E(X^2) = n(n-1)p^2 + E(X) = n(n-1)p^2 + np$. Grâce à la formule de Koenig-Huygens, il vient alors que $V(X) = E(X^2) - (E(X))^2 = npq$.*

3. *Pour la dernière démonstration, effectuons un calcul direct. Grâce à la formule du binôme, il vient en utilisant la petite formule de Pascal $k\binom{n}{k} = n\binom{n-1}{k-1}$:*

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

P ar ailleurs, pour la variance, calculons au préalable $E(X(X-1))$:

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

Donc $E(X^2) = E(X) + n(n-1)p^2 = np + n(n-1)p^2$. Alors

$$V(X) = E(X^2) - E^2(X) = np + n(n-1)p^2 - n^2p^2 = np - np^2 = npq.$$

9.5.4 Loi géométrique

□ On considère comme précédemment une suite d'épreuves de Bernoulli indépendantes de même paramètres p mais maintenant la suite est infini et on s'intéresse au rang X d'apparition du premier succès.
On a, avec $\Omega = \{S, E\}^{\mathbb{N}}$, $X(\Omega) = \mathbb{N}^*$ et si A est l'événement « les $k-1$ premiers résultats sont des échecs » et si B est l'événement « le k-ième résultat est un succès alors $P(X = k) = P(A \cap B)$. Mais A et B sont des événements indépendants donc

$$P(X = k) = P(A)P(B) = (1-p)^{k-1}p.$$

La loi de X est appelée loi géométrique.

DÉFINITION 9.28 ★ **Loi géométrique**
Soit $0 < p < 1$. On dit que la variable aléatoire X suit la *loi géométrique de paramètre* $p \in [0,1]$ si et seulement si $X(\Omega) = \mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$, $P(X = k) = (1-p)^{k-1}p$. On écrit $X \sim \mathscr{G}(p)$.

Remarque 9.26 La loi géométrique est bien une loi de probabilité. En effet, avec $q = 1-p$, $\sum_{k\geqslant 1} q^{k-1}p = \frac{p}{q}\sum_{k\geqslant 1} q^k = \frac{p}{q}$ qui est convergente car $0 < q < 1$ de somme $\sum_{k=1}^{+\infty} q^{k-1}p = \frac{p}{q}\left(\frac{1}{1-q} - 1\right) = 1$ et donc $P(X(\Omega)) = 1$.

PROPOSITION 9.46 **Fonction génératrice, espérance et variance de la loi géométrique**
Soit $X \sim \mathscr{G}(p)$ avec $p \in]0,1[$. Alors la fonction génératrice G_X de cette loi a un rayon de convergence égal à $R = \frac{1}{q} > 1$ où $q = 1-p$ et pour tout $t \in]-R, R[$, on a :

$$\boxed{G_X(t) = \sum_{k=1}^{+\infty} (1-p)^{k-1} p t^k = \frac{pt}{1-tq}}.$$

De plus, X admet une espérance et une variance finie données par

$$\boxed{\mathrm{E}(\mathrm{X}) = \frac{1}{p}} \qquad \text{et} \qquad \boxed{\mathrm{V}(\mathrm{X}) = \frac{q}{p^2}}.$$

Démonstration *La série entière* $\sum_{k=1}^{+\infty} q^{k-1} p t^k$ *admet, d'après le critère de d'Alembert pour les séries entières, comme rayon de convergence* $\mathrm{R} = 1/q$. *La fonction* $\mathrm{G_X}$ *donnée comme la somme de cette série entière est donc définie sur* $]-\mathrm{R}, \mathrm{R}[$ *et pour tout* t *dans cet intervalle :*

$$\mathrm{G_X}(t) = \sum_{k=1}^{+\infty} q^{k-1} p t^k = \frac{p}{q} \sum_{k=1}^{+\infty} (qt)^k = \frac{pt}{1-tq}.$$

Comme $\mathrm{R} > 1$, X *admet une espérance et une variance respectivement données par*

$$\mathrm{E}(\mathrm{X}) = \mathrm{G}'_x(1) = \frac{1}{p} \quad \text{et} \quad \mathrm{V}(\mathrm{X}) = \mathrm{G}''_{\mathrm{X}}(1) + \mathrm{G}'_{\mathrm{X}}(1) - \left(\mathrm{G}'_{\mathrm{X}}(1)\right)^2 = \frac{q}{p^2}$$

PROPOSITION 9.47 ★ $\mathrm{P}(\mathrm{X} > n) = q^n$
Soit $(\Omega, \mathscr{T}, \mathrm{P})$ un espace probabilisé et soit $\mathrm{X} : \Omega \to \mathbb{N}^*$ une variable aléatoire discrète. On a l'équivalence :

$$\boxed{\mathrm{X} \sim \mathscr{G}(p) \Longleftrightarrow \forall n \in \mathbb{N}^*, \quad \mathrm{P}(\mathrm{X} > n) = q^n}$$

avec $q = 1 - p$.

Démonstration

$\boxed{\Rightarrow}$ *En effet, par axiome d'additivité dénombrable, il vient :*

$$\mathrm{P}(\mathrm{X} > n) = \mathrm{P}\left(\bigcup_{k=n+1}^{+\infty} (\mathrm{X} = k)\right) = \sum_{k=n+1}^{+\infty} q^{k-1} p = q^n.$$

$\boxed{\Leftarrow}$ *Réciproquement, on a bien* $\mathrm{X}(\omega) = \mathbb{N}^*$ *et pour tout* $n \in \mathbb{N}^*$, $\mathrm{P}(\mathrm{X} = n) = \mathrm{P}(\mathrm{X} > n-1) - \mathrm{P}(\mathrm{X} > n) = q^{n-1} - q^n = q^{n-1}(1-q)$ *donc* $\mathrm{X} \sim \mathscr{G}(p)$.

THÉORÈME 9.48 ★ **Caractérisation de la loi géométrique comme loi sans mémoire**
La loi géométrique $\mathrm{X} \sim \mathscr{G}(p)$ est la seule loi de probabilité discrète à valeurs dans $\mathbb{N}^*$ telle que

$$\mathrm{P}_{\mathrm{X}>n}(\mathrm{X} > n + m) = \mathrm{P}(\mathrm{X} > m)$$

pour tout $m, n \in \mathbb{N}$.

Démonstration

— *Soient* $m, n \in \mathbb{N}$. *On suppose que* $\mathrm{X} \sim \mathscr{G}(p)$. *Comme*

$$\mathrm{P}(\mathrm{X} > n) = q^n \neq 0$$

, on peut calculer

$$\mathrm{P}_{\mathrm{X}>n}(\mathrm{X} > n+m) = \frac{\mathrm{P}\left((\mathrm{X} > n) \cap (\mathrm{X} > n+m)\right)}{\mathrm{P}(\mathrm{X} > n)} = \frac{\mathrm{P}(\mathrm{X} > n+m)}{\mathrm{P}(\mathrm{X} > n)} = \frac{q^{n+m} p}{q^n p} = q^m p = \mathrm{P}(\mathrm{X} > m)$$

et la loi géométrique est bien une loi sans mémoire.

— *Soit maintenant* Y *une variable aléatoire à valeurs dans* $\mathbb{N}^*$ *sans mémoire. Alors pour tout* $m, n \in \mathbb{N}$, $\mathrm{P}_{\mathrm{Y}>n}(\mathrm{Y} > n+m) = \mathrm{P}(\mathrm{Y} > m)$. *Alors* $\mathrm{P}\left((\mathrm{Y} > n) \cap (\mathrm{Y} > n+m)\right) = \mathrm{P}(\mathrm{Y} > n)\mathrm{P}(\mathrm{Y} > m)$ *ce qui amène* $\mathrm{P}(\mathrm{Y} > m+n) = \mathrm{P}(\mathrm{Y} > n)\mathrm{P}(\mathrm{Y} > m)$.
Mais alors si $q = \mathrm{P}(\mathrm{Y} > 1)$, *on obtient par une récurrence simple* $\mathrm{P}(\mathrm{Y} > n) = q^n$ *et par le théorème précédent, on sait que* Y *suit donc la loi géométrique* $\mathscr{G}(p)$ *avec* $p = 1 - q$.

9.5.5 Loi de Poisson

Comme on l'expliquera dans la section suivante, la loi de Poisson peut être vue comme limite de lois binomiales. On peut l'interpréter comme la loi du nombre de succès lorsque on compte un nombre très élevé d'épreuves ayant chacune une probabilité p très faible de succès.

DÉFINITION 9.29 ★ **Loi de Poisson**
Soit $\lambda \in \mathbb{R}_+^*$. On dit que la variable aléatoire X suit la *loi de Poisson de paramètre* λ si et seulement si $X(\Omega) = \mathbb{N}$ et pour tout $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$. On écrit $X \sim \mathscr{P}(\lambda)$.

Remarque 9.27 La loi géométrique est bien une loi de probabilité. En effet, $e^{\lambda} = \sum_{k\geqslant 0}^{+\infty} \dfrac{\lambda^k}{k!}$ donc $P(X(\Omega)) = 1$.

PROPOSITION 9.49 **Fonction génératrice, espérance et variance de la loi de poisson**
Soit $X \sim \mathscr{P}(\lambda)$ avec $\lambda \in \mathbb{R}_+^*$. Alors la fonction génératrice G_X de cette loi a un rayon de convergence égal à $R = +\infty$ et pour tout $t \in$
R, on a :

$$\boxed{G_X(t) = \sum_{k=0}^{+\infty} e^{-\lambda}\frac{t^k\lambda^k}{k!} = e^{-\lambda}e^{\lambda t}}.$$

De plus, X admet une espérance et une variance finie données par

$$\boxed{E(X) = \lambda} \qquad \text{et} \qquad \boxed{V(X) = \lambda}.$$

Démonstration *La série entière* $\sum_{k\geqslant 0} e^{-\lambda}\dfrac{t^k\lambda^k}{k!}$ *admet, d'après le critère de d'Alembert pour les séries entières, un rayon de convergence* $R = +\infty$. *La fonction* G_X *donnée comme la somme de cette série entière est donc définie sur* $\mathbb{R}$ *et pour tout* $t \in \mathbb{R}$ *:*

$$G_X(t) = \sum_{k=0}^{+\infty} e^{-\lambda}\frac{t^k\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{+\infty}\frac{t^k\lambda^k}{k!} = e^{-\lambda}e^{\lambda t}.$$

Comme $R > 1$, X *admet une espérance et une variance respectivement données par*

$$E(X) = G_x'(1) = \lambda \quad \text{et} \quad V(X) = G_X''(1) + G_X'(1) - \left(G_X'(1)\right)^2 = \lambda.$$

PROPOSITION 9.50 ★ **La somme deux variables aléatoires indépendantes suivant une loi de Poisson est une variable aléatoire suivant aussi la loi de Poisson**
Soient $X \sim \mathscr{P}(\lambda)$ et $X \sim \mathscr{P}(\lambda')$ où $\lambda, \lambda \in \mathbb{R}_+^*$ deux variables aléatoires indépendantes alors $X + Y \sim \mathscr{P}(\lambda + \lambda')$.

Démonstration *La somme de deux variables aléatoires discrète est encore une variable aléatoire discrète donc* $Z = X + Y$ *est une variable aléatoire discrète. De plus, comme* X *et* Y *sont à valeurs dans* $\mathbb{N}$, *il en est de même de* Z.
Pour $n \in \mathbb{N}$, *par indépendance de* X *et* Y *et par incompatibilité des événements de la famille* $((X = i) \cap (Y = n - i))_{i\in[\![0,n]\!]}$, *on a :*

$$\begin{aligned}
P(Z = n) &= \sum_{i=0}^{n} P\left((X = i) \cap (Y = n - i)\right)\\
&= \sum_{i=0}^{n} P(X = i)P(Y = n - i)\\
&= \sum_{i=0}^{n} e^{-\lambda}\frac{\lambda^i}{i!}e^{-\lambda'}\frac{\lambda'^{n-i}}{(n-i)!}\\
&= \frac{e^{-(\lambda+\lambda')}}{n!}\sum_{i=0}^{n}\binom{n}{i}\lambda^i\lambda'^{n-i}\\
&= \frac{e^{-(\lambda+\lambda')}}{n!}\left(\lambda+\lambda'\right)^n
\end{aligned}$$

d'où le résultat.
Proposons une autre preuve de ce résultat utilisant les fonctions génératrices. Comme X *et* Y *sont indépendantes, on a pour* $t \in \mathbb{R}$ *:*

$$G_Z(t) = G_{X+Y}(t) = G_X(t)G_Y(t) = e^{-\lambda}e^{\lambda t}e^{-\lambda'}e^{\lambda' t} = e^{-(\lambda+\lambda')}e^{(\lambda+\lambda')t} = \sum_{n=0}^{+\infty}\frac{\left(\lambda+\lambda'\right)^n}{n!}\mathrm{e}^{-(\lambda+\lambda')}t^n$$

et par unicité du DSE, il vient pour tout $n \in \mathbb{N}$, $P(Z = n) = \dfrac{\left(\lambda+\lambda'\right)^n}{n!}\mathrm{e}^{-(\lambda+\lambda')}$

9.6 Résultats asymptotiques

9.6.1 Loi binomiale et loi de Poisson

THÉORÈME 9.51 ★ **Approximation de la loi binomiale par la loi de Poisson**
Pour tout $n \in \mathbb{N}$, on considère $X_n \sim \mathscr{B}(n, p_n)$ où (p_n) est une suite de paramètre tels que $np_n \to \lambda \in \mathbb{R}_+^*$. Alors

$$\forall k \in \mathbb{N}, \quad \boxed{\lim_{n\to+\infty} P(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}}$$

Démonstration *Pour tout $n \in \mathbb{N}$, on note $q_n = 1 - p_n$. Pour tout $k \in \mathbb{N}$,*

$$P(X_n = k) = \begin{cases} \binom{n}{k} p_n^k q_n^{n-k} & \text{si} \quad 0 \leqslant k \leqslant n \\ 0 & \text{si} \quad k > n \end{cases}.$$

Par ailleurs, si $k \in [\![0, n]\!]$,

$$P(X_n = k) = \frac{1}{k!}\frac{n!}{(n-k)!}p_n^k q_n^{n-k} = \frac{1}{k!}n(n-1)\dots(n-k+1)p_n^k q_n^{n-k} = \frac{1}{k!}(1-\frac{1}{n})\dots(1-\frac{k-1}{n})(np_n)^k q_n^{n-k}$$

mais pour k fixé, $\prod_{i=1}^{k-1}\left(1-\frac{i}{n}\right) \xrightarrow[n\to+\infty]{} 1$. De plus,

$$q_n^{n-k} = (1-p_n)^{n-k} = e^{(n-k)\ln(1-p_n)} = e^{n\ln(1-p_n)}e^{-k\ln(1-p_n)}$$

mais $\ln(1-p_n) \underset{n\to+\infty}{\sim} -p_n$ car $p_n \xrightarrow[n\to+\infty]{} 0$ (en effet, comme $\lambda \neq 0$, $p_n \underset{n\to+\infty}{\sim} \lambda/n$) et donc $-k\ln(1-p_n) \xrightarrow[n\to+\infty]{} 0$ et $e^{n\ln(1-p_n)} \xrightarrow[n\to+\infty]{} e^{-\lambda}$.
On en tire que $q_n^{n-k} \xrightarrow[n\to+\infty]{} e^{-\lambda}$ puis finalement que $P(X_n = k) \xrightarrow[n\to+\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}$.

9.6.2 Loi faible des grands nombres

THÉORÈME 9.52 ★★★ **Loi faible des grands nombres**
Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles. On suppose que :

H1 les variables aléatoires X_n sont deux à deux indépendantes;

H2 qu'elles suivent toutes la même loi;

H3 qu'elles admettent un moment d'ordre 2

alors si m est l'espérance commune à toutes ces variables et si $S_n = \sum_{i=1}^n X_i$ alors :

$$\forall \varepsilon > 0, \quad \boxed{\lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = 0}.$$

Démonstration *Par linéarité de l'espérance,*

$$E\left(\frac{S_n}{n}\right) = \frac{1}{n}\sum_{k=1}^n E(X_k) = m$$

et comme les variables sont deux à deux indépendantes,

$$V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^n V(X_i) = \frac{\sigma^2}{n}$$

où σ est l'écart-type commun à toutes ces variables aléatoires.
On utilise alors l'inégalité de Bienaymé-Tchebychev pour $\varepsilon > 0$ fixé :

$$0 \leqslant P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

et par le théorème d'encadrement, il vient que $\lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = 0$.