

Variables aléatoires discrètes

Table des matières

10 Variables aléatoires discrètes	1
10.1 Variables aléatoires discrète	1
10.1.1 Définitions	1
10.1.2 Loi de probabilité et fonction de répartition d'une variable aléatoire	3
10.1.3 Loi associée à une image dénombrable	4
10.2 Couples de variables aléatoires	5
10.2.1 Loi conjointe, loi marginale	5
10.2.2 lois conditionnelles	6
10.2.3 Indépendance des variables aléatoires discrètes	7
10.2.4 Fonctions de variables aléatoires indépendantes	7
10.2.5 Variables mutuellement indépendantes	8
10.2.6 Le jeu du pile ou face infini	8
10.3 Moments d'une variable aléatoire	9
10.3.1 Espérance	9
10.3.2 Propriétés de l'espérance	11
10.3.3 Variance et covariance	12
10.4 Fonctions génératrices	16
10.5 Lois discrètes usuelles	18
10.5.1 Loi uniforme	18
10.5.2 Loi de Bernoulli	18
10.5.3 Loi binomiale	19
10.5.4 Loi géométrique	21
10.5.5 Loi de Poisson	22
10.6 Résultats asymptotiques	23
10.6.1 Loi binomiale et loi de Poisson	23
10.6.2 Loi faible des grands nombres	23

9.1 Variables aléatoires discrète

9.1.1 Définitions

DÉFINITION 9.1 ♡♡♡ **Variable aléatoire**

On appelle *variable aléatoire discrète* sur l'espace probabilisable (Ω, \mathcal{F}) à valeurs dans un ensemble E une application $X: \Omega \rightarrow E$ définie sur Ω

- ❶ dont l'image $X(\Omega)$ est finie ou dénombrable
- ❷ et telle que l'image réciproque de tout élément de $X(\Omega)$ appartient à \mathcal{F} :

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

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 (Ω, \mathcal{F}) .

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 n'intersectant $X(\Omega)$ qu'en ce point x .

PROPOSITION 9.1 ♡♡♡ **Image réciproque d'une réunion, d'une intersection**

Soit (Ω, \mathcal{F}) un espace probabilisable et $X : \Omega \rightarrow E$ 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 $X^{-1}(\bar{A}) = \overline{X^{-1}(A)}$;

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

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

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

- 1 Soit $x \in X^{-1}(\bigcup_{i \in I} A_i)$. 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 $x \in X^{-1}(\bigcup_{i \in I} A_i)$. Voilà qui prouve la première égalité.
- 2 Soit $x \in X^{-1}(\bigcap_{i \in I} A_i)$. 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 $x \in X^{-1}(\bigcap_{i \in I} A_i)$ et la seconde égalité est aussi prouvée.

On prouve encore la troisième formule. On a les équivalences

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

PROPOSITION 9.2 ♡♡♡ **L'image réciproque d'une partie de E par une variable aléatoire discrète est un événement**

Soit (Ω, \mathcal{F}) un espace probabilisable et $X : \Omega \rightarrow E$ une variable aléatoire discrète. Pour tout $U \subset E$, on a $X^{-1}(U) \in \mathcal{F}$.

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{F} 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 10.1 s'écrivent :

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

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

Remarque 9.3 Si on 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 (Ω, \mathcal{F}) un espace probabilisable et soit $A \in \mathcal{F}$ alors la variable aléatoire notée 1_A et définie par

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

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 $1_A^{-1}(1) = A$, $1_A^{-1} = \bar{A}$, l'image réciproque d'un élément de $1_A(\Omega)$ est bien un élément de \mathcal{F} .

DÉFINITION - PROPOSITION 9.3 ♡ Fonction d'une variable aléatoire

Soit X une variable aléatoire discrète sur un espace probabilisable (Ω, \mathcal{F}) 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 $(f \circ X)^{-1}(A) = X^{-1}(f^{-1}(A))$ mais $f^{-1}(A)$ est une partie de E et d'après la proposition 10.2, $X^{-1}(f^{-1}(A))$ est donc un élément de \mathcal{F} . Ainsi $f(X)$ est bien une variable aléatoire discrète sur (Ω, \mathcal{F}) .

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

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

DÉFINITION - PROPOSITION 9.4 ♡♡♡ Loi de probabilité d'une variable aléatoire

Soit (Ω, \mathcal{F}, P) un espace probabilisé et soit $X : \Omega \rightarrow E$ une variable aléatoire discrète. L'application

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

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

Démonstration Vérifions les axiomes définissant une probabilité :

- $P_X(X(\Omega)) = P(X^{-1}(X(\Omega))) = P(\Omega) = 1$.
- Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'éléments deux à deux disjoints de $\mathcal{P}(X(\Omega))$. Alors $(X^{-1}(A_n))_{n \in \mathbb{N}}$ est une suite d'éléments deux à deux disjoints de \mathcal{F} 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 $(X(\Omega), \mathcal{P}(X(\Omega)))$.

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

- $(X > x) = X^{-1}]x, +\infty[$;
- $(X \geq x) = X^{-1} [x, +\infty[$;
- $(X < x) = X^{-1}]-\infty, x[$;
- $(X \leq x) = X^{-1}]-\infty, x]$;

DÉFINITION 9.5 ♡ Fonction de répartition d'une variable aléatoire discrète réelle

On appelle *fonction de répartition* de la variable aléatoire discrète réelle $X : \Omega \rightarrow \mathbb{R}$ définie sur un espace probabilisé (Ω, \mathcal{F}, P) , la fonction

$$F_X : \begin{cases} \mathbb{R} & \longrightarrow [0, 1] \\ x & \longmapsto P_X]-\infty, x] \end{cases}$$

Remarque 9.4

— On a donc

$$F_X(x) = P(X \leq x) = P(X^{-1}]-\infty, x])$$

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

PROPOSITION 9.3 ♡ Propriétés des fonctions de répartition

Soit X une variable aléatoire discrète réelle définie sur un espace probabilisé (Ω, \mathcal{F}, P) et soit F_X sa fonction de répartition. Alors :

- F_X est croissante ;
- F_X admet une limite à droite et à gauche en tout point.
- $\lim_{-\infty} F_X = 0$ et $\lim_{+\infty} F_X = 1$;
- 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 .

- Pour tout $a, b \in \mathbb{R}$, $a \leq b$,

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

Démonstration

1. Soient $x, x' \in \mathbb{R}$ tels que $x \leq x'$ alors $]-\infty, x] \subset]-\infty, x']$ et $X^{-1}(]-\infty, x]) \subset X^{-1}(]-\infty, x'])$. Il s'ensuit que $P(X^{-1}(]-\infty, x])) \leq P(X^{-1}(]-\infty, x']))$ ce qui s'écrit aussi $P_X(X \leq x) \leq P(X \leq x')$. On a démontré que $F_X(x) \leq F_X(x')$ 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 $l \geq 1$ en $+\infty$. Mais $]-\infty, +\infty[= \bigcup_{n=0}^{+\infty}]-\infty, n]$ et par continuité croissante de P on a

$$l = \lim_{n \rightarrow +\infty} F_X(n) = \lim_{n \rightarrow +\infty} P(X \in]-\infty, n]) = P(X \in \bigcup_{n=0}^{+\infty}]-\infty, n]) = P_X(\mathbb{R}) = 1$$

donc $\lim_{+\infty} F_X = 1$.

De même, F_X admet une limite $l' \geq 0$ en $-\infty$ et pour tout $n \in \mathbb{N}$,

$$l' = \lim_{x \rightarrow -\infty} F_X(x) = \lim_{n \rightarrow +\infty} P(X \in]-\infty, -n]) = P\left(X \in \bigcap_{n \in \mathbb{N}}]-\infty, -n]\right) = P(\emptyset) = 0$$

en utilisant le théorème de continuité décroissante.

4. Soit $x \in \mathbb{R}$. Comme $]-\infty, x] = \bigcup_{n=0}^{+\infty}]-\infty, x - 2^{-n}]$, par continuité croissante de P ,

$$P(X < x) = P\left(\bigcup_{n=0}^{+\infty}]-\infty, x - 2^{-n}]\right) = \lim_{n \rightarrow +\infty} P(]-\infty, x - 2^{-n}]) = \lim_{n \rightarrow +\infty} F_X(x - 2^{-n}) = F(x^-).$$

Comme $(X \leq x) = (X = x) \cup (X < x)$, on obtient $P_X(\{x\}) = P(X = x) = F_X(x) - F_X(x^-)$.

5. Soient $a, b \in \mathbb{R}$, $a \leq b$, alors $(a < X \leq b) = (X \leq b) \setminus (X \leq a)$ d'où le résultat.

9.1.3 Loi associée à une image dénombrable

Soit X une variable aléatoire discrète sur un espace probabilisé (Ω, \mathcal{F}, 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 10.3. 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 (Ω, \mathcal{F}) et une variable aléatoire discrète $X : \Omega \rightarrow \mathbb{R}$, existe-t-il une probabilité P sur (Ω, \mathcal{F}) 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 9.4 ♡♡♡ Loi associée à une image dénombrable

Soit X une variable aléatoire discrète définie sur un espace probabilisable (Ω, \mathcal{F}) 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 (Ω, \mathcal{F}) telle que

$P(X = x_n) = p_n$

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

DÉFINITION - PROPOSITION 9.6 ♡ Couple de variables aléatoires discrètes

Soient E et F deux ensembles et (Ω, \mathcal{F}) un espace probablisable. On considère $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ deux variables aléatoires discrètes. On introduit

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

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 :

$$(Z = (x, y)) = (X = x) \cap (Y = y) .$$

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 $Z^{-1}(\{(x, y)\}) = X^{-1}(\{x\}) \cap Y^{-1}(\{y\}) \in \mathcal{F}$ comme intersection de deux éléments de \mathcal{F} . Donc Z est bien une variable aléatoire discrète sur Ω .

DÉFINITION 9.7 ♡ Loi conjointe, lois marginales

Soit (Ω, \mathcal{F}, P) un espace probablisé et $X : \Omega \rightarrow E$, $Y : \Omega \rightarrow F$ 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 9.5 ♡ Expression de la loi conjointe

Soit (Ω, \mathcal{F}, P) un espace probablisé et $X : \Omega \rightarrow E$, $Y : \Omega \rightarrow F$ deux variables aléatoires discrètes. Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a :

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

Démonstration En utilisant la définition-proposition 10.6,

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

Remarque 9.5 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 $X(\Omega) \times Y(\Omega) \rightarrow [0, 1]$ définie par $(x_i, y_j) \mapsto p_{i,j}$ avec $p_{i,j} = P((X = x_i) \cap (Y = y_j))$.
2. L'application $X(\Omega) \rightarrow [0, 1]$, $x_i \mapsto p_{i,\cdot}$, où $p_{i,\cdot} = P(X = x_i)$ est la première loi marginale du couple (X, Y) , c'est-à-dire la loi de X .
3. L'application $Y(\Omega) \rightarrow [0, 1]$, $y_j \mapsto p_{\cdot,j}$ où $p_{\cdot,j} = P(Y = y_j)$ est la seconde loi marginale du couple (X, Y) ; c'est-à-dire la loi de Y .

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 $(k, l) \in \llbracket 1, 3 \rrbracket \times \llbracket 1, 2 \rrbracket$, on a

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

$$P(X = k) = \frac{\binom{3}{k} \binom{6}{3-k}}{\binom{9}{3}} \quad \text{et} \quad P(Y = l) = \frac{\binom{2}{l} \binom{7}{3-l}}{\binom{9}{3}}.$$

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 9.6 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 fautive, la connaissance des lois de X et de Y n'entraînent en général pas celle de la loi conjointe du couple (X, Y) .

THÉORÈME 9.6 ♡ Liens entre les lois marginales et la loi conjointe

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

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

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

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

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 $\sum_{y \in Y(\Omega)} P((X = x) \cap (Y = y)) = \sum_{y \in Y(\Omega)} P((X, Y) = (x, y))$ converge car converge absolument. De plus,

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

Remarque 9.7 Autrement dit, avec les notations de la remarque précédente, pour tout $(i, j) \in \mathbb{N}^2$,

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

Remarque 9.8 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 fautive.

9.2.2 lois conditionnelles

DÉFINITION 9.8 ♡ Loi conditionnelle

Soit (X, Y) un couple de variables aléatoires discrètes définies sur un espace probabilisé (Ω, \mathcal{F}, 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

$$P_x(y) = P_{(X=x)}(Y = y) = \frac{P((X = x) \cap (Y = y))}{P(X = x)} .$$

— 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

$$P_y(x) = P_{(Y=y)}(X = x) = \frac{P((X = x) \cap (Y = y))}{P(Y = y)}.$$

La proposition suivante découle directement de la définition :

PROPOSITION 9.7 ♡ Lien entre lois conjointe, marginale et conditionnelle

On a

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

9.2.3 Indépendance des variables aléatoires discrètes

DÉFINITION 9.9 ♡ Variables aléatoires indépendantes

Soient X, Y deux variables aléatoires définies sur un espace probabilisé (Ω, \mathcal{F}, 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 :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad P((X = x) \cap (Y = y)) = P(X = x)P(Y = y).$$

PROPOSITION 9.8 ♡ Caractérisation de l'indépendance de deux variables aléatoires

Soient X, Y deux variables aléatoires définies sur un espace probabilisé (Ω, \mathcal{F}, P) . On a équivalence entre :

- (X, Y) est un couple de variables aléatoires indépendantes ;
- $\forall A \times B \subset X(\Omega) \times Y(\Omega), \quad P((X \in A) \cap (Y \in B)) = P(X \in A)P(Y \in B)$.

Démonstration *La preuve est admise en PC.*

\Rightarrow On suppose que (X, Y) est un couple de variables indépendantes. Soient $A \times B \subset X(\Omega) \times Y(\Omega)$. 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 :

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

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

\Leftarrow La réciproque est triviale.

9.2.4 Fonctions de variables aléatoires indépendantes

PROPOSITION 9.9 ♡ Fonctions de deux variables aléatoires indépendantes

Soient X, Y deux variables aléatoires discrètes indépendantes définies sur un espace probabilisé (Ω, \mathcal{F}, 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.

Démonstration Soient $a \in f \circ X(\Omega)$ et $b \in g \circ Y(\Omega)$. Alors d'après la proposition 10.8

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

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

9.2.5 Variables mutuellement indépendantes

DÉFINITION 9.10 ♡ Variables mutuellement indépendantes

Soit $(X_i)_{i \in [1, n]}$ une famille finie de variables aléatoires discrètes définies sur un espace probabilisé (Ω, \mathcal{F}, 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 :

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

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

PROPOSITION 9.10

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

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

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

PROPOSITION 9.11

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

Démonstration La preuve est admise

PROPOSITION 9.12 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 9.11 ♡ 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 hors programme mais malgré tout bien utile.

THÉORÈME 9.13 ♡♡♡ Lemme des coalitions HORS PROGRAMME

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 $f : E_1 \times \dots \times E_p \rightarrow E$, $g : E_{p+1} \times \dots \times E_n \rightarrow F$ alors $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont indépendantes.

9.2.6 Le jeu du pile ou face infini

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{F} 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 (Ω, \mathcal{F}) un espace probabilisé (Ω, \mathcal{F}, 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{F}), on a

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

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 9.14 ♡ 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 $\Omega = \prod_{k=0}^{+\infty} \Omega_k$, \mathcal{F} 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 :

$$P\left(\bigcap_{k=0}^n (X_k \in R_k)\right) = \prod_{k=0}^n P_k(X_k \in R_k).$$

Si pour chaque lancer, on introduit une variable aléatoire de Bernoulli $X_k : \Omega_k \rightarrow \{0, 1\}$ 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 :

$$P\left(\bigcap_{k=0}^n (X_k = x_k)\right) = \prod_{k=0}^n P_k(X_k = x_k) = \frac{1}{2^{n+1}}.$$

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 $X : \Omega \rightarrow \mathbb{R}$ est donnée par

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

Notre objectif maintenant est de généraliser cette notion dans le cas où Ω est dénombrable. Il va se poser deux problèmes :

- 1 Tout d'abord celui de la convergence de la somme $\sum_{\omega \in \Omega} X(\omega)P(\omega)$
- 2 Et celui ensuite posé par le fait que la valeur d'une telle somme, quand elle existe, dépend de l'ordre dans lequel cette sommation a été effectuée.

Une fois encore, le problème posé par l'ordre de sommation va être résolu en considérant des séries absolument convergentes.

On aboutit à la définition suivante :

DÉFINITION 9.12 ♡♡♡ Espérance d'une variable aléatoire

Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, \mathcal{F}, P) et à valeurs dans un ensemble dénombrable $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$. On dit que X est d'espérance finie si la série $\sum x_k P(X = x_k)$ est absolument convergente. Sa somme est notée $E(X)$ et est appelée l'espérance de X :

$$E(X) = \sum_{k=0}^{+\infty} x_k P(X = x_k).$$

Remarque 9.9 Comme indiqué plus haut, la série $\sum x_k P(X = x_k)$ étant absolument convergente, sa somme $\sum_{k=0}^{+\infty} x_k P(x_k)$ est indépendante de l'ordre d'énumération.

Remarque 9.10 Dans le cas d'un univers Ω fini, l'espérance d'une variable aléatoire est toujours définie.

PROPOSITION 9.15 ♡ **Espérance d'une variable aléatoire discrète constante**

Soit (Ω, \mathcal{F}, P) un espace probabilisé et $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire discrète réelle. On suppose que :

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

alors X admet une espérance donnée par $E(X) = m$.

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

DÉFINITION 9.13 ♡ **Variable centrée, variable centrée associée à une variable aléatoire**

- Une variable aléatoire réelle discrète admettant une espérance nulle est dite *centrée*.
- À une variable aléatoire réelle discrète X qui admet une espérance finie, on peut associer la variable aléatoire $X - E(X)$ 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 9.16 ♡ **Calcul pratique de l'espérance d'une variable aléatoire à valeurs entières**

Soit X une variable aléatoire à valeurs dans \mathbb{N} et définie sur un espace probabilisé (Ω, \mathcal{F}, P) . La variable aléatoire X admet une espérance finie si et seulement si la série de terme général $P(X \geq n)$ est convergente. Dans ce cas,

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

Démonstration Soit $k \in \mathbb{N}$. On sait que

$$P(X = k) = P(X \geq k) - P(X \geq k + 1)$$

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

$$\sum_{k=0}^n kP(X = k) = \sum_{k=0}^n k(P(X \geq k) - P(X \geq k + 1)) = \sum_{k=1}^n P(X \geq k) - nP(X \geq n + 1) \quad (*)$$

Procédons maintenant à la démonstration de la proposition :

⇒ Si $E(X)$ est finie alors la série $\sum kP(X = k)$ converge donc on peut considérer son reste d'ordre n qu'on note \tilde{R}_n .

Par définition d'une probabilité, comme la famille $(X = k)_{k \in \mathbb{N}}$ est formée d'événements deux à deux incompatibles, la série $\sum_{k \in \mathbb{N}} P(X = k)$ est elle aussi convergente. On considère alors aussi R_n son reste d'ordre n qu'on note R_n . On a

$$R_n = \sum_{k=n+1}^{+\infty} P(X = k) = P(X \geq n + 1).$$

Il vient alors

$$\tilde{R}_n = \sum_{k=n+1}^{+\infty} kP(X = k) \geq \sum_{k=n+1}^{+\infty} (n+1)P(X = k) = (n+1) \sum_{k=n+1}^{+\infty} P(X = k) = (n+1)P(X \geq n + 1) = nR_n \geq 0.$$

Mais $\tilde{R}_n \xrightarrow{n \rightarrow +\infty} 0$ donc, par le théorème d'encadrement, $(n+1)P(X \geq n + 1) = nR_n \xrightarrow{n \rightarrow +\infty} 0$. On en déduit que $\sum_{k=0}^n P(X \geq k)$

converge vers $E(X)$. Donc $E(X) = \sum_{n=0}^{+\infty} P(X \geq n)$

⇐ On suppose maintenant que la série de terme général $P(X \geq n)$ est convergente. De l'égalité (*), on tire pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^n kP(X = k) \leq \sum_{k=1}^n P(X \geq k).$$

Les séries $\sum kP(X = k)$ et $\sum P(X = k)$ sont toutes deux à termes positifs. La seconde converge, donc ses sommes partielles sont majorées. On en déduit que les sommes partielles de la première sont aussi majorées et qu'elle est aussi convergente. Donc X admet une espérance finie.

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 9.17 ♡ Théorème de transfert

Soit X une variable aléatoire discrète et soit f une application à valeurs réelles définie sur l'image $\{x_n \mid n \in \mathbb{N}\}$ de X . Alors la variable aléatoire $f(X)$ est d'espérance finie si et seulement si la série $\sum P(X = x_n) f(x_n)$ converge absolument. Dans ce cas, on a :

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

Démonstration Ce théorème est admis.

PROPOSITION 9.18 ♡ Linéarité de l'espérance

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

$$E: \begin{cases} \mathcal{E} & \longrightarrow \mathbb{R} \\ X & \longmapsto E(X) \end{cases}$$

est une forme linéaire sur \mathcal{E}

Démonstration La preuve est admise

PROPOSITION 9.19 ♡ Positivité, croissance de l'espérance

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

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 $E(X) \leq E(Y)$.

On a aussi le théorème suivant :

THÉORÈME 9.20 ♡ Critère de majoration

Si X et Y sont deux variables aléatoires réelles discrètes telles que :

(H1) $|X| \leq Y;$

(H2) Y admet une espérance finie

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

$$|E(X)| \leq E(Y).$$

Démonstration La preuve est admise

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

PROPOSITION 9.21 ♡ Espérance du produit de deux variables aléatoires indépendantes

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

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

Démonstration La preuve est hors programme.

De manière plus générale, on a la proposition suivante :

PROPOSITION 9.22 ♡ Espérance d'un produit de variables aléatoires mutuellement indépendantes

Si X_1, \dots, X_n sont n variables aléatoires discrètes définies sur un espace probabilisé (Ω, \mathcal{F}, P) admettant des espérances finies alors

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

Démonstration *Hors programme* On effectue une récurrence sur $n \geq 2$. 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, \dots, 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 $Y_n = X_1 \dots X_n$ et X_{n+1} sont indépendantes. D'après l'hypothèse de récurrence, Y_n admet une espérance finie donnée par $E(Y_n) = E(X_1) \dots E(X_n)$. 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 $X_1 \dots X_{n+1}$ admet une espérance finie donnée par $E(X_1 \dots X_{n+1}) = E(X_1) \dots E(X_{n+1})$ ce qui prouve la proposition par récurrence.

THÉORÈME 9.23 ♡ Inégalité de Markov

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

$$\forall a > 0, \quad P(X \geq a) \leq \frac{E(X)}{a}$$

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 $K_{\geq a} = \{k \in \mathbb{N} \mid x_k \geq a\}$. Remarquons que pour tout $n \in \mathbb{N}$,

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

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 $\sum_{k \in K_{\geq a}} x_k P(X = x_k)$ est convergente et en passant aux sommes partielles que

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

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

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

Il vient alors

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

ce qui prouve l'inégalité.

9.3.3 Variance et covariance

DÉFINITION 9.14 ♡ 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 9.11 Dire que X admet un moment d'ordre 1 revient à dire que X a une espérance finie.

PROPOSITION 9.24 ♡ 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 :

$$|E(XY)| \leq \sqrt{E(X^2)} \sqrt{E(Y^2)}$$

Démonstration On sait que pour tout $(x, y) \in \mathbb{R}^2$, $|xy| \leq \frac{1}{2}(x^2 + y^2)$. 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 10.20 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 $Q(\lambda) = E((\lambda X + Y)^2)$, on a pour tout $\lambda \in \mathbb{R}$, $Q(\lambda) \geq 0$. 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 $\delta \leq 0$. Comme $\delta = (E(XY))^2 - E(X^2)E(Y^2)$, on obtient que $|E(XY)| \leq \sqrt{E(X^2)}\sqrt{E(Y^2)}$. Par ailleurs, si $E(X^2) = 0$ alors $Q(\lambda) = 2\lambda E(XY) + E(Y^2) \geq 0$ ce qui n'est possible que si $E(XY) = 0$ et l'inégalité est encore vérifiée.

COROLLAIRE 9.25 ♡ 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 $|E(X)| \leq \sqrt{E(X^2)}$.

COROLLAIRE 9.26 ♡ 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 9.15 ♡♡♡ 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 :

$$V(X) = E((X - E(X))^2)$$

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

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

Remarque 9.12

- 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 $(X - E(X))^2 \geq 0$, il en est de même de son espérance et l'écart type $\sigma(X)$ est lui aussi bien défini.
- Si $E(X) = m$, $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$ 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 9.27 ♡♡♡ 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 :

$$V(X) = E(X^2) - (E(X))^2;$$

- 2 Pour tout $(a, b) \in \mathbb{R}^2$,

$$V(aX + b) = a^2 V(X).$$

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 :

$$V(X) = E((X - m)^2) = E(X^2 - 2mX + m^2) = E(X^2) - 2mE(X) + m^2 = E(X^2) - m^2$$

d'où le résultat.

- 2 De même, en utilisant la formule de König-Huygens et le fait que $E(aX + b) = aE(X) + b$ pour $a, b \in \mathbb{R}$, on a que :

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

DÉFINITION 9.16 ♡ **Variable réduite, variable centrée réduite**

Soit X une variable aléatoire discrète admettant un moment d'ordre 2 et telle que $\sigma(X) > 0$.

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

THÉORÈME 9.28 ♡ **Inégalité de Bienaymé-Tchebychev**

Soit X une variable aléatoire discrète admettant un moment d'ordre 2 alors on a l'inégalité de Bienaymé-Tchebychev :

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

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 $Y = |X - m|$. Remarquons que $E(Y^2) = V(X)$. Il vient alors pour tout $\varepsilon > 0$,

$$P(|X - E(X)| \geq \varepsilon) = P(Y^2 \geq \varepsilon^2) \leq \frac{E(Y^2)}{\varepsilon^2} = \frac{V(X)}{\varepsilon^2}.$$

DÉFINITION 9.17 ♡ **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 $\text{cov}(X, Y)$ et donné, si elle existe, par :

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Remarque 9.13 La covariance de deux variables aléatoires réelles discrètes admettant un moment d'ordre 2 est bien définie par le théorème 10.24.

PROPOSITION 9.29 ♡ **Existence de la 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 :

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

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 10.24, 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 $\text{cov}(X, Y)$ 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 9.30 ♡ **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 $\text{cov}(Y, X) = \text{cov}(X, Y)$;
- 2 $\text{cov}(X, X) = V(X)$;
- 3 $\text{cov}(aX + b, cY + d) = ac \text{cov}(X, Y)$;
- 4 $\text{cov}(aX + bX', cY + dY') = ac \text{cov}(X, Y) + bc \text{cov}(X', Y) + ad \text{cov}(X, Y') + bd \text{cov}(X', Y')$.

Remarque 9.14 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 9.31 ♡ **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 $\text{cov}(X, Y) = 0$.

Démonstration Comme X et Y sont indépendantes, $E(XY) = E(X)E(Y)$ et la formule de la proposition 10.29 permet de conclure en la nullité de leur covariance.

DÉFINITION 9.18 ♡ **Variabes aléatoires non corrélées**
Deux variables aléatoires réelles discrètes admettant une covariance nulle sont dites *non corrélées*

PROPOSITION 9.32 ♡ **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

$$\boxed{V(X + Y) = V(X) + V(Y) + 2 \text{cov}(X, Y)} ;$$

- 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 10.24, 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 :

$$V(X + Y) = E\left((X + Y)^2\right) - (E(X + Y))^2 = \left[E(X^2) - m^2\right] + \left[E(Y^2) - m'^2\right] + 2[E(XY) - mm'] = V(X) + V(Y) + 2 \text{cov}(X, Y)$$

- 2 Si X et Y sont indépendantes, alors $\text{cov}(X, Y) = 0$ d'où la formule.

Remarque 9.15 Ce résultat se généralise facilement à une somme de n variables aléatoires X_1, \dots, X_n admettant des moments d'ordre 2 :

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

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 9.19 ♡ **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

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

PROPOSITION 9.33
Soient X et Y deux variables aléatoires réelles discrètes admettant un moment d'ordre 2. Alors :

- 1 $|\text{cov}(X, Y)| \leq \sigma(X)\sigma(Y)$;
- 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 10.26, de $tX + Y$ où $t \in \mathbb{R}$. Alors on considère $Q(t) = V(tX + Y) = t^2 V(X) + 2t \operatorname{cov}(X, Y) + V(Y) \geq 0$.

- 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 9.16 On a $\rho(X, Y) = \pm 1$ si et seulement si il existe $t_0 \in \mathbb{R}$ tel que $Q(t_0) = V(t_0 X + Y) = 0$. Mais d'après l'exercice ?? page ??, ceci est équivalent à l'existence de $t_1 \in \mathbb{R}$ tel que $P(t_0 X + Y = t_1) = 1$. 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 9.20 ♥♥♥ **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 :

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

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

Remarque 9.17 Comme $X(\Omega) = \mathbb{N}$, si $Y = t^X$ alors $Y(\Omega) = \{t^{X(\omega)} \mid \omega \in \Omega\} \subset \{t^n \mid n \in \mathbb{N}\}$ 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 9.34 ♥ **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, $\lim_{t \rightarrow 1^-} G_X(t) = 1$.
- 3 La fonction G_X est continue de classe \mathcal{C}^∞ sur $] -R, R[$.

- 4 $\forall n \in \mathbb{N}$, $P(X = n) = \frac{G_X^{(n)}(0)}{n!}$. 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 ≥ 1 .

On pose pour $n \in \mathbb{N}$, $f_n : \begin{cases} [0, 1] & \rightarrow \mathbb{R} \\ t & \mapsto P(X = n) t^n \end{cases}$. 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 $n \in \mathbb{N}$, $f_n \in \mathcal{C}^0([0, 1])$. On en déduit par critère de continuité sur séries de fonctions que G_X est continue sur $[0, 1]$. En particulier, on a $\lim_{t \rightarrow 1^-} G_X(t) = 1$.

Le reste du théorème est une conséquence des résultats du chapitre sur les séries entières.

LEMME 9.35 ♥ **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) $\forall n \in \mathbb{N}, a_n \geq 0$

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, $\lim_{t \rightarrow 1^-} S(t) = \sum_{n=0}^{+\infty} a_n$.

Démonstration Remarquons que S est croissante sur $[0, 1]$. Elle admet alors, d'après le théorème de la limite monotone, une limite l quand $t \rightarrow 1^-$ finie ou infinie.

Si $\sum a_n$ converge alors comme pour tout $n \in \mathbb{N}$ et $t \in [0, 1]$, $a_n t^n \leq a_n$, $\sum a_n t^n$ 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 $\sum_{n=0}^N a_n t^n \leq S(t) \leq l$ pour tous $t \in [0, 1]$ et $N \in \mathbb{N}$ alors en passant à la limite quand $t \rightarrow 1^-$, il vient $\sum_{n=0}^N a_n \leq l$ et en faisant $N \rightarrow +\infty$, on a alors $l = +\infty$

PROPOSITION 9.36 ♡ **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,

$$E(X) = \lim_{t \rightarrow 1^-} G'_X(t).$$

- 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,

$$E(X(X-1)) = \lim_{t \rightarrow 1^-} G''_X(t) \quad \text{et} \quad V(X) = \lim_{t \rightarrow 1^-} G''_X(t) + \lim_{t \rightarrow 1^-} G'_X(t) - \left(\lim_{t \rightarrow 1^-} G'_X(t) \right)^2.$$

- 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 :

$$E(X) = G'(1), \quad E(X(X-1)) = G''(1) \quad \text{et} \quad V(X) = G''(1) + G'(1) - (G'(1))^2$$

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 \mathcal{C}^∞ 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 $G''(t) = \sum_{n=0}^{+\infty} n(n-1) t^{n-2} P(X = n)$. Les séries $\sum n P(X = n)$ 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'une 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, $G'(1) = \sum_{n \geq 0} n P(X = n) = E(X)$ et

$$G''(1) = \sum_{n=0}^{+\infty} n(n-1) P(X = n) = E(X(X-1)) \text{ d'où la formule pour } V(X).$$

— 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 $\sum n P(X = n) t^{n-1}$ pour le premier résultat et à la série entière $\sum n(n-1) P(X = n) t^{n-2}$ pour le second.

PROPOSITION 9.37 ♡ **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 :

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

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

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

Démonstration On a

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

l'avant dernière égalité provenant du fait que X et Y sont indépendantes et donc que pour toute fonction $f : \mathbb{N} \rightarrow \mathbb{R}$, $f \circ X$ 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 9.21 ♡ Loi uniforme

Soit $n \in \mathbb{N}^*$. On dit que la variable aléatoire X suit la *loi uniforme sur* $\llbracket 1, n \rrbracket$ si et seulement si X peut prendre les valeurs $1, 2, \dots, n$ avec la probabilité $\frac{1}{n}$ ou autrement dit, si et seulement si

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

On écrit alors $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$.

Remarque 9.18 La loi uniforme \mathcal{U} est bien une loi de probabilité car $\sum_{k \in \llbracket 1, n \rrbracket} \mathcal{U}(k) = \sum_{k \in \llbracket 1, n \rrbracket} \frac{1}{n} = 1$.

PROPOSITION 9.38 Espérance et variance de la loi uniforme

Soit X une variable aléatoire suivant la loi uniforme. Alors

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

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

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

donc

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

Remarque 9.19 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 9.22 ♡ 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 ou 1 avec, respectivement, les probabilités $q = 1 - p$ et p :

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

On écrit $X \hookrightarrow \mathcal{B}(p)$

Remarque 9.20 La loi de Bernoulli est bien une loi de probabilité. En effet, $P(\{0\}) + P(\{1\}) = p + (1 - p) = 1$.

PROPOSITION 9.39 Espérance et variance de la loi de Bernoulli
Soit X une variable aléatoire suivant la loi de Bernoulli $\mathcal{B}(p)$. Alors

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

où $q = 1 - p$.

Démonstration On a

$$E(X) = 0 \cdot q + 1 \cdot p = p$$

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

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

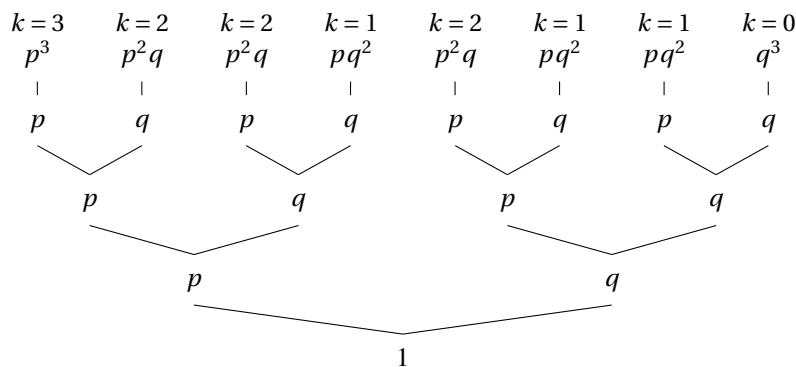
Remarque 9.21 La série génératrice de la loi de Bernoulli de paramètre p est donnée, pour tout $t \in \mathbb{R}$ par :

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

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 $P(X = 0) = q^3$, $P(X = 1) = \binom{3}{1} p q^2$, $P(X = 2) = \binom{3}{2} p^2 q$ 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 9.23 ♡ **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 $0, 1, \dots, n$ avec les probabilités $P(X = k) = \binom{n}{k} p^k q^{n-k}$ où $q = 1 - p$. Sa loi est donc donnée par :

$$P_X : \begin{cases} [0, n] & \rightarrow \mathbb{R} \\ k & \mapsto \binom{n}{k} p^k q^{n-k} \end{cases} .$$

On écrit $X \hookrightarrow \mathcal{B}(n, p)$.

Remarque 9.22 On a $\sum_{k=0}^n P_X(k) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1$ 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 9.40 ♡ Une somme de variables de Bernoulli indépendantes et de même paramètre suit la loi binomiale

Soient X_1, \dots, X_n n variables aléatoires indépendantes suivant la même loi de Bernoulli $X_i \hookrightarrow \mathcal{B}(p)$ alors leur somme suit la loi binomiale $\mathcal{B}(n, p) : X_1 + \dots + X_n \hookrightarrow \mathcal{B}(n, p)$.

Démonstration On effectue la preuve par une récurrence sur $n \geq 2$. Si $X_1 \hookrightarrow \mathcal{B}(p)$ et $X_2 \hookrightarrow \mathcal{B}(p)$ sont deux variables aléatoires de Bernoulli indépendantes alors, si $Y = X_1 + X_2$:

- $P(Y = 2) = P((X_1 = 1) \cap (X_2 = 1)) = P(X_1 = 1)P(X_2 = 1) = p^2$;
- $P(Y = 1) = P([(X_1 = 1) \cap (X_2 = 0)] \cup [(X_1 = 0) \cap (X_2 = 1)]) = P(X_1 = 1)P(X_2 = 0) + P(X_1 = 0)P(X_2 = 1) = 2pq$;
- $P(Y = 0) = P((X_1 = 0) \cap (X_2 = 0)) = P(X_1 = 0)P(X_2 = 0) = q^2$.

Dans tous les cas, on a $P(X_1 + X_2 = k) = \binom{2}{k} p^k q^{2-k}$ et la propriété est vraie au rang 2

On suppose la propriété vraie au rang $n-1$ pour $n \geq 3$ et on la prouve au rang n . On considère n variables aléatoires indépendantes suivant une même loi de Bernoulli $\mathcal{B}(p)$. Soit $k \in [0, n]$. Il est clair que $P(X_1 + \dots + X_n = n) = P((X_1 = 1) \cap \dots \cap (X_n = 1)) = \prod_{k=1}^n P(X_k = 1) = p^n$. De même, $P(X_1 + \dots + X_n = 0) = q^n$. Soit $k \in]1, n-1[$. On a alors, avec $Y = X_1 + \dots + X_n$:

$$\begin{aligned} P(Y = k) &= P([(X_1 + \dots + X_{n-1} = k) \cap (X_n = 0)] \cup [(X_1 + \dots + X_{n-1} = k-1) \cap (X_n = 1)]) \\ &= P([(X_1 + \dots + X_{n-1} = k) \cap (X_n = 0)]) + P([(X_1 + \dots + X_{n-1} = k-1) \cap (X_n = 1)]) \\ &= P(X_1 + \dots + X_{n-1} = k) P(X_n = 0) + P(X_1 + \dots + X_{n-1} = k-1) P(X_n = 1) \\ &= \binom{n-1}{k} p^k q^{n-1-k} q + \binom{n-1}{k-1} p^{k-1} q^{n-1-k+1} p \\ &= \left(\binom{n-1}{k} + \binom{n-1}{k-1} \right) p^k q^{n-k} \\ &= \binom{n}{k} p^k q^{n-k} \end{aligned}$$

et la propriété est prouvée par récurrence.

PROPOSITION 9.41 ♡ Espérance et variance de la loi binomiale

Soit X une variable aléatoire suivant la loi binomiale $\mathcal{B}(n, p)$. Alors

$$E(X) = np$$

$$V(X) = npq$$

et

$$G_X(t) = (q + pt)^n$$

où $q = 1 - p$.

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

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

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 :

$$E(X) = E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n) = np$$

$$V(X) = V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n) = npq$$

- 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 $k \binom{n}{k} = n \binom{n-1}{k-1}$:

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

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

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

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

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

9.5.4 Loi géométrique

On considère comme précédemment une suite d'épreuves de Bernoulli indépendantes de même paramètre p mais maintenant la suite est infinie 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 9.24 ♡ 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 $X \hookrightarrow \mathcal{G}(p)$.

Remarque 9.23 La loi géométrique est bien une loi de probabilité. En effet, avec $q = 1-p$, $\sum_{k \geq 1} q^{k-1} p = \frac{p}{q} \sum_{k \geq 1} q^k = \frac{p}{q}$

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 9.42 Fonction génératrice, espérance et variance de la loi géométrique

Soit $X \hookrightarrow \mathcal{G}(p)$ 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 :

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

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

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

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 $R = 1/q$. La fonction G_X donnée comme la somme de cette série entière est donc définie sur $]-R, R[$ et pour tout t dans cet intervalle :

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

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

$$E(X) = G'_X(1) = \frac{1}{p} \quad \text{et} \quad V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 = \frac{q}{p^2}$$

THÉORÈME 9.43 ♡ Caractérisation de la loi géométrique comme loi sans mémoire

La loi géométrique $X \hookrightarrow \mathcal{G}(p)$ est la seule loi de probabilité discrète à valeurs dans \mathbb{N}^* telle que

$$P_{X > n}(X > n+m) = P(X > m)$$

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

Démonstration

— Soient $m, n \in \mathbb{N}$. On suppose que $X \hookrightarrow \mathcal{G}(p)$. On calcule

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

et comme cette probabilité est non nulle, on peut calculer

$$P_{X>n}(X > n+m) = \frac{P((X > n) \cap (X > n+m))}{P(X > n)} = \frac{P(X > n+m)}{P(X > n)} = \frac{q^{n+m} p}{q^n p} = q^m p = P(X > m)$$

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 $m, n \in \mathbb{N}$, $P_{Y>n}(Y > n+m) = P(Y > m)$. Alors $P((Y > n) \cap (Y > n+m)) = P(Y > n)P(Y > m)$ ce qui amène $P(Y > m+n) = P(Y > n)P(Y > m)$.

Mais alors si $q = P(Y > 1)$, on obtient par une récurrence simple $P(Y > n) = q^n$ et $P(Y = n) = P(Y > n-1) - P(Y > n) = q^{n-1} - q^n = q^{n-1}(1-q)$ et Y suit donc la loi géométrique $\mathcal{G}(p)$ 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 9.25 ♡ 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 $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$. On écrit $X \hookrightarrow \mathcal{P}(\lambda)$.

Remarque 9.24 La loi géométrique est bien une loi de probabilité. En effet, $e^\lambda = \sum_{k \geq 0} \frac{\lambda^k}{k!}$ donc $P(X(\Omega)) = 1$.

PROPOSITION 9.44 Fonction génératrice, espérance et variance de la loi de poisson

Soit $X \hookrightarrow \mathcal{P}(\lambda)$ 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 $t \in \mathbb{R}$, on a :

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

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

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

Démonstration La série entière $\sum_{k \geq 0} e^{-\lambda} \frac{t^k \lambda^k}{k!}$ 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

$$E(X) = G'_X(1) = \lambda \quad \text{et} \quad V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 = \lambda$$

PROPOSITION 9.45 ♡ La somme deux variables aléatoires indépendantes suivant une loi de Poisson est une variable aléatoire suivant aussi la loi de Poisson

Soient $X \hookrightarrow \mathcal{P}(\lambda)$ et $X \hookrightarrow \mathcal{P}(\lambda')$ où $\lambda, \lambda' \in \mathbb{R}_+^*$ deux variables aléatoires indépendantes alors $X + Y \hookrightarrow \mathcal{P}(\lambda + \lambda')$.

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 :

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

d'où le résultat.

Proposons une autre preuve de ce résultat utilisant les fonctions génératrices. Comme X et Y sont indépendantes, on a pour $t \in \mathbb{R}$:

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

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

9.6 Résultats asymptotiques

9.6.1 Loi binomiale et loi de Poisson

THÉORÈME 9.46 ♡ Approximation de la loi binomiale par la loi de Poisson

Pour tout $n \in \mathbb{N}$, on considère $X_n \rightarrow \mathcal{B}(n, p_n)$ où (p_n) est une suite de paramètres tels que $np_n \rightarrow \lambda \in \mathbb{R}_+^*$. Alors

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

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

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

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

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

mais pour k fixé, $\prod_{i=1}^{k-1} \left(1 - \frac{i}{n}\right) \xrightarrow{n \rightarrow +\infty} 1$. 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 $\ln(1-p_n) \underset{n \rightarrow +\infty}{\sim} -p_n$ car $p_n \xrightarrow{n \rightarrow +\infty} 0$ (en effet, comme $\lambda \neq 0$, $p_n \underset{n \rightarrow +\infty}{\sim} \lambda/n$) et donc $-k \ln(1-p_n) \xrightarrow{n \rightarrow +\infty} 0$ et $e^{n \ln(1-p_n)} \xrightarrow{n \rightarrow +\infty} e^{-\lambda}$.

On en tire que $q_n^{n-k} \xrightarrow{n \rightarrow +\infty} e^{-\lambda}$ puis finalement que $P(X_n = k) \xrightarrow{n \rightarrow +\infty} e^{-\lambda} \frac{\lambda^k}{k!}$.

9.6.2 Loi faible des grands nombres

THÉORÈME 9.47 ♡♡♡ 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 :

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

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 :

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

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