Chapitre 31

Couples et vecteurs aléatoires

Compétences attendues.

- ✓ Étudier ou utiliser l'indépendance de variables aléatoires.
- ✓ Obtenir la loi d'un couple ou d'un vecteur, ses lois marginales.
- ✓ Déterminer la loi d'une somme, d'un minimum ou d'un maximum.
- ✓ Calculer une covariance, une variance d'une somme.

Mathieu Mansuy - Professeur en MP2I au Lycée Carnot (Dijon)
Page personnelle : `mathieu-mansuy.fr/`
E-mail : `mathieu.mansuy@ac-dijon.fr`

1 Indépendance de variables aléatoires

Dans tout ce chapitre, (Ω, P) désigne un espace probabilisé fini.

Définition.

Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires définies sur (Ω, P). On dit que X et Y sont *indépendantes*, et on note $X \perp\!\!\!\perp Y$, si pour tous $A \in \mathscr{P}(E)$ et $B \in \mathscr{P}(F)$, les événements $[X \in A]$ et $[Y \in B]$ sont indépendants :

$$P([X \in A] \cap [Y \in B]) = P(X \in A)P(Y \in B).$$

Propriété 1

Deux variables aléatoires X et Y sur (Ω, P) sont indépendantes si, et seulement si, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $[X = x]$ et $[Y = y]$ sont indépendants :

$$P([X = x] \cap [Y = y]) = P(X = x)P(Y = y).$$

Remarque. La plupart du temps l'indépendance de deux variables aléatoires X et Y, si elle n'est pas déjà supposée dans l'énoncé, résulte directement de l'expérience aléatoire. Par exemple dans le cas d'un tirage **avec remise** dans une urne, si X est le numéro de la première boule tirée et Y celui de la seconde, alors les variables X et Y sont indépendantes.

Exercice 1. On tire deux boules sans remise dans une urne qui en contient n, numérotées de 1 à n. On note X le numéro de la première et Y le numéro de la seconde. Les variables X et Y sont-elles indépendantes ?

Propriété 2 (Image de variables aléatoires indépendantes)

Soient X et Y deux variables aléatoires indépendantes, $f : X(\Omega) \to E$ et $g : Y(\Omega) \to F$ deux applications. Alors $f(X)$ et $g(Y)$ sont indépendantes.

Définition.

Soient $X_1, \dots, X_n$ des variables aléatoires définies sur (Ω, P) à valeurs dans $E_1, \dots, E_n$ respectivement. Les variables $X_1, \dots, X_n$ sont dites :

- *deux à deux indépendantes* si pour tout couple (i, j) d'éléments distincts de $[\![1, n]\!]$, X_i et X_j sont indépendantes ;
- *mutuellement indépendantes* si, pour tout $(A_1, \dots, A_n) \in \mathscr{P}(X_1(\Omega)) \times \dots \times \mathscr{P}(X_n(\Omega))$:

$$P([X_1 \in A_1] \cap \dots \cap [X_n \in A_n]) = \prod_{i=1}^{n} P(X_i \in A_i).$$

Remarque. En fixant certains A_i à $X_i(\Omega)$, on constate que si $X_1, \dots, X_n$ sont mutuellement indépendantes, alors toute sous-famille l'est aussi. En particulier, les X_i sont deux à deux indépendantes. Attention, **la réciproque est fausse** : l'indépendance deux à deux n'implique pas l'indépendance mutuelle.

Propriété 3

Soient $X_1, \dots, X_n$ des variables aléatoires sur (Ω, P). Alors $X_1, \dots, X_n$ sont mutuellement indépendantes si, et seulement si :

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

Propriété 4

Soient $X_1, \dots, X_n$ des variables aléatoires sur (Ω, P), et $f_1, \dots, f_n$ des applications telles que pour tout $i \in [\![1, n]\!]$, f_i soit définie sur $X_i(\Omega)$.
Si $X_1, \dots, X_n$ sont mutuellement indépendantes (resp. deux à deux indépendantes), alors les variables $f_1(X_1), \dots, f_n(X_n)$ sont mutuellement indépendantes (resp. deux à deux indépendantes).

Propriété 5 (Lemme des coalitions)

Soient $X_1, \dots, X_n$ des variables aléatoires sur (Ω, P). Soient $m \in [\![1, n-1]\!]$, E et F deux ensembles, $f : X_1(\Omega) \times \dots \times X_m(\Omega) \to E$ et $g : X_{m+1}(\Omega) \times \dots \times X_n(\Omega) \to F$ deux applications.
Si $X_1, \dots, X_n$ sont mutuellement indépendantes, alors les coalitions $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.

Remarque. Cet énoncé se généralise sans difficulté au cas d'un nombre fini quelconque de coalitions.

Exemple. Si $(X_1, X_2, X_3, X_4, X_5)$ sont mutuellement indépendantes, alors $\left(X_3, X_1 + X_5, X_2^2 e^{X_4}\right)$ sont encore mutuellement indépendantes. En revanche, on ne peut rien dire de l'indépendance éventuelle de $(X_3, X_1 + X_3, X_1 X_2)$, puisque X_3 figure dans deux des trois variables, de même que X_1 (les « paquets » ne sont pas disjoints).

2 Couples et n-uplets de variables aléatoires

2.1 Définition, loi conjointe

Définition.

Soient E et E' deux ensembles. Si $X : \Omega \to E$ et $Y : \Omega \to E'$ sont deux variables aléatoires sur Ω, alors l'application

$$(X, Y) : \begin{array}{ccc} \Omega & \longrightarrow & E \times E' \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array}$$

est appelée *couple de variables aléatoires sur* Ω. Si X et Y sont des variables aléatoires réelles, on dit que (X, Y) est un *couple de variables aléatoires réelles.*

Remarque. Un couple de variables aléatoires n'est rien d'autre qu'une variable aléatoire à valeurs dans un produit cartésien. Par conséquent, toutes les définitions et résultats du **Chapitre 29. Variables aléatoires finies.** s'appliquent également ici.

Définition.

Soit (X,Y) un couple de variables aléatoires sur un espace probabilisé fini (Ω,P). La loi de (X,Y) est appelée *loi conjointe du couple* (X,Y).

Méthode. Comment déterminer la loi conjointe de (X,Y) ?

Déterminer la loi conjointe de deux variables aléatoires X et Y revient à :

- *déterminer $X(\Omega)=\{x_1,\dots,x_n\}$ et $Y(\Omega)=\{y_1,\dots,y_p\}$;*
- *déterminer pour tout $(i,j)\in[\![1,n]\!]\times[\![1,p]\!]$, la valeur de $p_{i,j}=P([X=x_i]\cap[Y=y_j])$.*

La loi conjointe de (X,Y) peut être représentée par un tableau à double entrée de la forme suivante :

$X \backslash Y$	y_1	y_2	$\dots$	y_m
x_1	$p_{1,1}=P([X=x_1]\cap[Y=y_1])$	$p_{1,2}$	$\dots$	$p_{1,m}$
x_2	$p_{2,1}$	$p_{2,2}$	$\dots$	$p_{2,m}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_n	$p_{n,1}$	$p_{n,2}$	$\dots$	$p_{n,m}$

Notation.

On notera plus simplement $[X=x,Y=y]$ l'événement $[X=x]\cap[Y=y]=[(X,Y)=(x,y)]$.

Exercice 2. Considérons l'expérience aléatoire consistant à jeter deux dés, et notons X la variable aléatoire donnant le plus petit résultat et Y le plus grand. Déterminer la loi conjointe de $Z=(X,Y)$.

Remarques.

- Dans cet exemple, $Z(\Omega)=\{(i,j)\in[\![1,6]\!]\mid i\leq j\}$ et seule la partie triangulaire supérieure du tableau suffit à caractériser la loi conjointe. En fait, on a toujours $(X,Y)(\Omega)\subset X(\Omega)\times Y(\Omega)$ et la loi de (X,Y) est bien caractérisée par la donnée de $P(X=x,Y=y)$ pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$.
- La famille $([X=x]\cap[Y=y])_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est un système complet d'événements. En particulier :

$$\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} P(X=x,Y=y)=1.$$

  Ainsi, lorsqu'on représente la loi conjointe de (X,Y) par un tableau, la somme de toutes les probabilités qui y figurent vaut 1.

2.2 Lois marginales

Définition.

Soit (X,Y) un couple de variables aléatoires sur (Ω,P).
La loi de X est appelée *première loi marginale du couple* (X,Y), celle de Y *deuxième loi marginale*.

Propriété 6

Soit (X,Y) un couple de variables aléatoires. Notons $X(\Omega)=\{x_1,\dots,x_n\}$ et $Y(\Omega)=\{y_1,\dots,y_p\}$. Alors :

$$\forall i\in[\![1,n]\!],\quad p_i=P(X=x_i)=\sum_{j=1}^{p}P([X=x_i]\cap[Y=y_j])=\sum_{j=1}^{p}p_{i,j},$$

$$\forall j\in[\![1,p]\!],\quad q_j=P(Y=y_j)=\sum_{i=1}^{n}P([X=x_i]\cap[Y=y_j])=\sum_{i=1}^{n}p_{i,j}.$$

Remarque. Lorsque la loi de (X,Y) est représentée par un tableau, on obtient $P(X=x_i)$ en sommant les termes de la i-ème ligne et $P(Y=y_j)$ en sommant les termes de la j-ème colonne :

X \ Y	y_1	$\dots$	y_m	Somme par lignes 1ère loi marginale $\downarrow$
x_1	$P([X=x_1]\cap[Y=y_1])$	$\dots$	$P([X=x_1]\cap[Y=y_m])$	$P([X=x_1])$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_n	$P([X=x_n]\cap[Y=y_1])$	$\dots$	$P([X=x_n]\cap[Y=y_m])$	$P([X=x_n])$
Somme par colonnes 2ème loi marginale $\rightarrow$	$P([Y=y_1])$	$\dots$	$P([Y=y_m])$	Somme totale $=1$

Exercice 3. Déterminer les lois marginales du couple (X,Y) de l'exercice précédent.

X \ Y	$Y=1$	$Y=2$	$Y=3$	$Y=4$	$Y=5$	$Y=6$	
$X=1$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X=2$	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X=3$	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X=4$	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X=5$	0	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	
$X=6$	0	0	0	0	0	$\frac{1}{36}$	

✍ **À retenir. Lien entre loi conjointe et lois marginales.**

- À partir de la loi conjointe de (X,Y), on peut obtenir les lois marginales de X et de Y. En revanche, il n'est en général **pas possible de retrouver la loi conjointe à partir des lois marginales** (on ne peut retrouver l'ensemble des valeurs figurant dans le tableau ci-dessus par la seule connaissance de sa dernière ligne et de sa dernière colonne) ;
- Il y a cependant un cas où c'est possible : lorsque les variables X et Y sont **indépendantes**,

$$p_{i,j}=P([X=x_i]\cap[Y=y_j])=P([X=x_i])P(Y=y_j])=p_i\times q_j.$$

Dans le cas **indépendant**, la loi conjointe se retrouve à partir des lois marginales par produit.

$$\begin{tikzcd} \boxed{\text{Loi conjointe de } (X,Y)} \arrow[r, bend left, "\text{Formule de probabilités totales}"] & \boxed{\text{Lois marginales de } X \text{ et de } Y} \arrow[l, dashed, bend left, "\text{Si } X \text{ et } Y \text{ sont indépendantes}"] \end{tikzcd}$$

2.3 Lois conditionnelles

Définition.

Soit (X,Y) un couple de variables aléatoires, et soit $x \in X(\Omega)$ tel que $P(X=x) \neq 0$.

On appelle *loi conditionnelle de Y sachant que* $[X=x]$ la loi de Y pour la probabilité $P_{[X=x]}$. Autrement dit, il s'agit de la donnée de $Y([X=x])$ et des probabilités $P_{[X=x]}(Y=y)$ pour tout $y \in Y([X=x])$.

On définit de même la *loi conditionnelle de X sachant que* $[Y=y]$.

Propriété 7

Soit (X,Y) un couple de variables aléatoires. On suppose que $P(X=x) \neq 0$ et $P(Y=y) \neq 0$ pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$. Alors, pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$:

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

$$P(X=x) = \sum_{y \in Y(\Omega)} P(Y=y)P_{[Y=y]}(X=x),$$

$$P(Y=y) = \sum_{x \in X(\Omega)} P(X=x)P_{[X=x]}(Y=y).$$

Exercice 4. Une urne contient 4 boules numérotées de 1 à 4. On tire une boule de l'urne, on note X le numéro qu'elle porte, et on ne la remet pas dans l'urne. On effectue alors n tirages avec remise dans cette urne, et on note Y le nombre de boules portant des numéros impairs ainsi obtenues.

Déterminer les lois de X et Y, ainsi que la loi du couple (X,Y).

2.4 Généralisation aux n-uplets de variables aléatoires

Définition.

Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires sur Ω à valeurs dans $E_1, E_2, \ldots, E_n$ respectivement. On dit que l'application :

$$(X_1, \ldots, X_n) : \begin{array}{ccc} \Omega & \longrightarrow & E_1 \times \cdots \times E_n \\ \omega & \longmapsto & (X_1(\omega), \ldots, X_n(\omega)) \end{array}$$

est un *n-uplet de variables aléatoires sur* Ω, ou encore un *vecteur aléatoire*.

La loi de la variable aléatoire $(X_1, \ldots, X_n)$ est appelée *loi conjointe* de $(X_1, \ldots, X_n)$, et les lois des X_i sont appelées *lois marginales du n-uplet* $(X_1, \ldots, X_n)$.

Comme pour les couples, on montre que :

- la loi conjointe est caractérisée par la donnée de :

$$P\left([X_1 = x_1] \cap \cdots \cap [X_n = x_n]\right) \quad \text{où} \quad (x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega).$$

- La formule des probabilités totales permet d'obtenir la i-ème loi marginale à partir de la loi conjointe :

$$\forall x \in X_i(\Omega), \quad P\left(X_i = x\right) = \sum_{\substack{x_1 \in X_1(\Omega), \ldots, x_{i-1} \in X_{i-1}(\Omega) \\ x_{i+1} \in X_{i+1}(\Omega), \ldots, x_n \in X_n(\Omega)}} P\left([X_1 = x_1] \cap \cdots \cap [X_i = x] \cap \cdots \cap [X_n = x_n]\right).$$

2.5 Variable aléatoire fonction d'un nombre fini de variables aléatoires

Loi d'une somme de variables aléatoires réelles

Propriété 8 (*Produit de convolution discret*)

Soient X et Y des variables aléatoires **indépendantes** à valeurs dans $\mathbb{N}$. Alors :

$$\forall n \in \mathbb{N}, \quad P(X + Y = n) = \sum_{k=0}^{n} P(X = k)P(Y = n - k).$$

Propriété 9 (Stabilité de la loi binomiale par somme)

Soient $n_1, n_2 \in \mathbb{N}^*$ et $p \in [0, 1]$. Considérons X_1, X_2 deux variables aléatoires **indépendantes** sur un même espace probabilisé (Ω, P) telles que $X_1 \sim \mathscr{B}(n_1, p)$ et $X_2 \sim \mathscr{B}(n_2, p)$.
Alors $X_1 + X_2 \sim \mathscr{B}(n_1 + n_2, p)$.

Corollaire 10

- Soit $p \in [0, 1]$, et soient $X_1, \ldots, X_k$ des variables aléatoires sur (Ω, P) mutuellement indépendantes telles que $X_i \sim \mathscr{B}\left(n_i, p\right)$ pour tout $i \in [\![1, k]\!]$. Alors :

$$\sum_{i=1}^{k} X_i \sim \mathscr{B}\left(\sum_{i=1}^{k} n_i, p\right).$$

- En particulier, si $X_1, \ldots, X_n$ sont n variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre p, alors $\sum_{i=1}^{n} X_i$ suit la loi binomiale $\mathscr{B}(n, p)$.

Remarque. Ceci justifie qu'une loi binomiale $\mathscr{B}(n, p)$ s'interprète comme le nombre de succès lors d'une répétition de n épreuves de Bernoulli indépendantes de même paramètre p.

Loi d'un maximum ou d'un minimum de variables aléatoires réelles

Propriété 11

Soient $X_1, \dots, X_n$ des variables aléatoires à valeurs dans $\mathbb{N}$, et $Z = \max(X_1, \dots, X_n)$.

- Pour tout $k \in \mathbb{N}$, $[Z \leq k] = \bigcap_{i=1}^{n} [X_i \leq k]$.
- Si de plus les variables $X_1, \dots, X_n$ sont mutuellement indépendantes, alors :

$$\forall k \in \mathbb{N}, \quad P(Z \leq k) = P(X_1 \leq k) \times \cdots \times P(X_n \leq k).$$

- Pour tout $k \in \mathbb{N}$, $P(Z = k) = P(Z \leq k) - P(Z \leq k-1)$.

Exercice 5. Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes suivant toutes une loi $\mathscr{U}(\llbracket 1, N \rrbracket)$. Déterminer la loi de $Z = \max(X_1, \dots, X_n)$.

Remarque. On procède de même pour la loi d'un minimum $Z = \min(X_1, \dots, X_n)$ en considérant pour tout $k \in \mathbb{N}$, les événements $[Z \geq k]$ et en notant que $P(Z = k) = P(Z \geq k) - P(Z \geq k+1)$.

3 Covariance

3.1 Espérance d'un produit

Propriété 12

Soient X et Y deux variables aléatoires réelles sur un espace probabilisé (Ω, P). Alors :

$$E(XY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xyP([X = x] \cap [Y = y]).$$

Propriété 13

Si X et Y sont deux variables aléatoires **indépendantes** sur (Ω, P), alors :

$$E(XY) = E(X)E(Y).$$

☠ **Danger.**

Sans l'hypothèse d'indépendance, cette propriété est fausse : par exemple, on n'a pas en général $E(X \times X) = E(X) \times E(X)$ (sauf si X suit une loi certaine).

Corollaire 14

Si $X_1, \dots, X_n$ sont des variables aléatoires réelles **indépendantes** sur (Ω, P), alors :

$$E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E\left(X_i\right).$$

3.2 Covariance

Définition.

Soient X et Y deux variables aléatoires réelles sur (Ω, P). On appelle *covariance de X et Y* le réel :

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

On dit que X et Y sont *décorrélées* ou *non corrélées* si $\mathrm{Cov}(X, Y) = 0$.

Exercice 6. Soient X et Y des variables aléatoires réelles sur (Ω, P). On suppose que Y est une variable aléatoire certaine. Montrer que X et Y sont décorrélées.

Propriété 15 (*Formule de Huygens* (1629 - 1695))

Si X et Y sont deux variables aléatoires réelles sur (Ω, P), alors :

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

Propriété 16

La covariance est une forme bilinéaire symétrique et positive sur l'ensemble des variables aléatoires réelles sur (Ω, P) : pour toutes variables X, Y, Z, T,

- *Symétrie :* $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.
- *Bilinéarité* (c'est-à-dire linéaire par rapport à chacune des variables) :
  - *linéarité à gauche* : $\forall \lambda \in \mathbb{R}$, $\mathrm{Cov}(\lambda X + Y, Z) = \lambda \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$;
  - *linéarité à droite* : $\forall \lambda \in \mathbb{R}$, $\mathrm{Cov}(X, \lambda Z + T) = \lambda \mathrm{Cov}(X, Z) + \mathrm{Cov}(X, T)$;
- *Positivité :* $\mathrm{Cov}(X, X) = V(X) \geq 0$.

Propriété 17

Soient X et Y des variables aléatoires réelles sur (Ω, P).
Si X et Y sont **indépendantes**, **alors** $\mathrm{Cov}(X, Y) = 0$.

Danger.

Si $\mathrm{Cov}(X,Y) \neq 0$, **alors** X et Y ne sont pas indépendantes. Attention, **la réciproque est fausse** : on peut avoir $\mathrm{Cov}(X,Y) = 0$ avec X et Y non indépendantes.

Exercice 7. Soient X et Y deux variables aléatoires indépendantes, suivant une même loi de Bernoulli $\mathscr{B}(p)$ de paramètre $p \in]0,1[$. On pose $S = X + Y$ et $U = X - Y$. Calculer $\mathrm{Cov}(S,U)$. Les variables S et U sont-elles indépendantes ?

Le saviez-vous ?

Le statisticien et chimiste anglais William Gosset (1876 - 1937) travaillait aux brasseries Guiness à Dublin. Dans le but d'améliorer la qualité du houblon, il l'étudia de manière statistique. Cela l'amena à introduire la notion de variance sous le nom de *fluctuation*. Son compatriote Ronald Fisher (1890 - 1962) préféra la nommer *variance*, mot qui a pour sens désaccord en anglais. Le préfixe *co* vient de la préposition latine *cum* signifiant *avec*. Ainsi *covariance* correspond à « désaccord avec ».

L'entreprise Guiness avait pour règle que ses chimistes ne publient pas leurs découvertes. Gosset argua que ses travaux ne seraient d'aucune utilité pour les concurrents et obtint l'autorisation de publier mais sous un pseudonyme, Student, pour éviter les difficultés avec les autres membres de son équipe. C'est ce nom qu'on retrouve aujourd'hui encore associé à de nombreux outils statistiques.

William Gosset, alias Student

3.3 Variance d'une somme

Propriété 18 (Variance d'une somme)

Soient X et Y des variables aléatoires réelles sur (Ω, P). Alors :

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

Si de plus X et Y sont **indépendantes**, alors :

$$V(X+Y) = V(X) + V(Y).$$

Astuce.

On retiendra en particulier de la première égalité que $\mathrm{Cov}(X,Y) = \dfrac{1}{2}(V(X+Y) - V(X) - V(Y))$, ce qui permet de calculer la covariance à partir des variances.

Propriété 19

Plus généralement, si $X_1, \ldots, X_n$ sont n variables aléatoires réelles sur (Ω, P), alors :

$$V\left(X_1 + X_2 + \cdots + X_n\right) = \sum_{i=1}^{n} V\left(X_i\right) + 2 \sum_{1 \leq i < j \leq n} \operatorname{Cov}\left(X_i, X_j\right).$$

En particulier, si $X_1, X_2, \ldots, X_n$ sont deux à deux **indépendantes** :

$$V\left(X_1 + \cdots + X_n\right) = V\left(X_1\right) + V\left(X_2\right) + \cdots + V\left(X_n\right).$$

Astuce.

Si on connait l'espérance et la variance d'une loi de Bernoulli (p et $p(1-p)$ resp.), alors on retrouve sans calcul celles d'une loi binomiale. En effet, si $X_i \sim \mathscr{B}(p)$ pour tout $i = 1, \ldots, n$, alors $X = \sum_{i=1}^{n} X_i \sim \mathscr{B}(n, p)$, et :

$$E(X) \overset{\text{lin. de } E}{=} E(X_1) + \cdots + E(X_n) = np,$$

$$V(X) \overset{\text{indep. des } X_i}{=} V(X_1) + \cdots + V(X_n) = np(1-p).$$

3.4 Inégalités de concentration

Propriété 20 (*Inégalité de Markov* (1856 - 1922))

Soit X est une variable aléatoire **positive**. Alors :

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

Propriété 21 (*Inégalité de Bienaymé* (1796 - 1878) *-Tchebychev* (1821 - 1894))

Soit X une variable aléatoire sur (Ω, P). Alors :

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

Remarque. Les inégalités de Markov et de Bienaymé-Tchebychev permettent d'estimer la « concentration » des valeurs prises par X autour de son espérance. Et ceci sans connaitre la loi de X : il suffit simplement de connaître son espérance et sa variance. Ces majorations sont cependant assez grossières, souvent bien supérieures aux probabilités considérées.

Exercice 8. Combien de fois doit-on lancer un dé équilibré pour garantir avec moins de 5% d'erreur que la fréquence d'apparition du 2 sera $\frac{1}{6} \pm 0,01$?

Remarque. Cette inégalité est loin d'être optimale : un calcul direct en utilisant la loi binomiale prouve qu'en fait $n \geq 5395$ suffit.

3.5 Vers la loi faible des grands nombres

Considérons $X_1, \ldots, X_n$ des variables aléatoires réelles indépendantes, qui suivent toute la même loi, et notons $S_n = \sum_{i=1}^{n} X_i$. Notons également μ l'espérance des X_i et σ^2 leur variance. Alors :

$$E\left(\frac{S_n}{n}\right) \overset{\text{lin. de } E}{=} \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n} n\mu = \mu \quad \text{et} \quad V\left(\frac{S_n}{n}\right) \overset{\text{indép. des } X_i}{=} \frac{1}{n^2}\sum_{i=1}^{n} V\left(X_i\right) = \frac{1}{n^2}\left(n\sigma^2\right) = \frac{\sigma^2}{n}.$$

Par l'inégalité de Bienaymé-Tchebychev appliquée à $\frac{S_n}{n}$, il vient alors, pour tout $\varepsilon > 0$:

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

Si on s'autorise à passer à la limite lorsque $n \to +\infty$ (ce qui en toute rigueur nécessiterait de travailler sur un univers infini, hors programme en première année), on obtient par encadrement :

$$\lim_{n \to +\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0.$$

Ce résultat est appelé la *loi faible des grands nombres*. Et il est particulièrement remarquable car il justifie l'interprétation « fréquentiste » que nous avons des probabilités. En effet, si on répète un grand nombre n de fois une même expérience, qu'on note $X_1, \ldots, X_n$ les résultats de l'expérience considérée, et qu'on calcule la moyenne $\frac{S_n}{n}$ de ces résultats, alors la probabilité que cette moyenne soit à distance plus de $\varepsilon > 0$ de $\mu = E\left(X_1\right)$ tend vers 0 lorsque n tend vers $+\infty$. En résumé :

> *si on répète un grand nombre de fois l'expérience, la moyenne des résultats va être proche de l'espérance.*

En particulier, considérons une succession d'expériences identiques et indépendantes, au cours de chacune desquelles un événement A est susceptible de se produire avec probabilité $p = P(A)$. Pour tout $i \geq 1$, on définit la variable X_i par :

$$X_i = \begin{cases} 1 & \text{si l'événement } A \text{ se réalise au cours de la } i\text{-ème expérience,} \\ 0 & \text{sinon.} \end{cases}$$

Les variables aléatoires X_i sont indépendantes et suivent toutes la loi de Bernoulli de paramètre $E(X_1) = p$. La loi faible des grands nombres affirme que la fréquence $\frac{S_n}{n}$ de réalisation de l'événement A au cours des n premières expériences est « proche » de $p = P(A)$. Ainsi,

> *la fréquence statistique de réalisation d'un événement « tend » vers la probabilité de cet événement.*

Ce résultat est fondamental : c'est lui qui justifie que le formalisme que nous employons correspond bien à l'intuition que l'on se fait d'une probabilité.