# Table des matières



# I Introduction

**Définition 1 Rappel : Variable aléatoire discrète.** On se donne $(\Omega, \mathcal{T})$.
Une application $X$ définie sur $(\Omega, \mathcal{T})$ est u**ne variable aléatoire discrète** si

- $X$ est définie sur $\Omega$.
- $X(\Omega)$ est fini ou dénombrable,
- l'image réciproque de tout élément de $X(\Omega)$ appartient à $\mathcal{T}$ c'est à dire :

$$\forall x \in X(\Omega),\ X^{-1}(\{x\}) = \{\omega \in \Omega \text{ tels que } X(\omega) = x\} \in \mathcal{T}$$

Si $X(\Omega)$ est fini, on a une variable aléatoire (discrète) finie et si $X(\Omega)$ est dénombrable, on a une variable aléatoire discrète infinie.

N.B. : ici une variable aléatoire n'est pas nécessairement réelle. $X(\omega)$ peut aussi être un vecteur par exemple....

**Définition 2** Soient $X$ et $Y$ deux V.A. discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$, à valeurs dans $E$ et $F$ respectivement. L'application

$$V : \begin{array}{l} \Omega \to E \times F \\ \omega \mapsto (X(\omega), Y(\omega)) \end{array}$$

est un **couple de V.A. discrètes.**

**Remarque 1** On a vu qu'un produit cartésien d'ensembles au plus dénombrables est au plus dénombrable.
Ici, **on pourrait aussi considérer l'application $V$ comme une variable aléatoire discrète définie sur $\Omega$ et à valeurs dans $X(\Omega) \times Y(\Omega)$.** En effet :

- Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, $V^{-1}(\{(x, y)\}) = ((X = x) \cap (Y = y))$ qui est bien un événement (intersection de deux éléments de la tribu $\mathcal{T}$ par hypothèse, donc c'est bien un élément de $\mathcal{T}$).
- $X(\Omega) \times Y(\Omega)$ est au plus dénombrable (étant donné que $X$ et $Y$ sont des V.A. discrètes).

N.B. : Au lieu de noter $((X = x) \cap (Y = y))$, on peut noter $(X = x, Y = y)$.

**Exemple 1** On tire un dé $n$ fois. On note $X$ le plus petit des numéros tirés et $Y$ le plus grand. Puis on pose $V = (X, Y)$.

**Remarque 2 A propos de $V(\Omega)$.**
On a : $\forall \omega \in \Omega,\ (X(\omega), Y(\omega)) \in X(\Omega) \times Y(\Omega)$.
En toute rigueur, $V(\Omega) \subset X(\Omega) \times Y(\Omega)$.
Par exemple, dans l'exemple précédent :

$V(\Omega) =$

$X(\Omega) \times Y(\Omega) =$

Néanmoins, l'usage est de prendre $V(\Omega) = X(\Omega) \times Y(\Omega)$.
Pour certaines valeurs de $x$ et de $y$, on pourra alors avoir $(X = x) \cap (Y = y) = \emptyset$ , donc $P(X = x, Y = y) = 0$. Autrement dit, pour certains $v \in V(\Omega)$, on pourra avoir $P(V = v) = 0$.
**Conséquence** : Le système d'événements $((X = x, Y = y))_{x \in X(\Omega), y \in Y(\Omega)}$ est un système complet d'événements, mais certains d'entre eux peuvent être de probabilité nulle. Aussi lorsque l'on appliquera la formule des probabilités totales, on précisera bien la convention usuelle :

# II Loi conjointe et lois marginales d'un couple de variables aléatoires

**Définition 3** Soit $V = (X, Y)$ un couple de $V.A.$ discrètes.
**La loi conjointe du couple** $V = (X, Y)$ est la loi de $V$ vu comme variable aléatoire (cf remarque 1).
C'est donc la donnée de $V(\Omega) = X(\Omega) \times Y(\Omega)$ et des probabilités $P(X = x, Y = y)$ pour tous les couples $(x, y) \in V(\Omega)$.
Les lois de $X$ et de $Y$ sont appelées **les lois marginales** du couple $(X, Y)$.

Quand les V.A. sont finies, on peut représenter la loi de $V = (X, Y)$ dans un tableau.

**Exemple 2** On lance successivement deux dés.
On note $X$=numéro du premier dé lancé. On note $Y$= somme des deux numéros tirés. Donner la loi de $V = (X, Y)$.

**Exemple 3** On lance deux dés (un vert et un rouge) équilibrés. On note $X$ le plus petit numéro, $Y$ le plus grand. Loi conjointe du couple $(X, Y)$ ? Lois marginales de $X$ et de $Y$ ?

**Exemple 4** On donne : $\forall (i, j) \in (\mathbb{N}^*)^2, \quad P(X = i, Y = j) = \dfrac{1}{2^{i+j}}$. Loi conjointe du couple $(X, Y)$ ? Lois marginales de $X$ et de $Y$ ?

**Proposition 1** La donnée de la loi conjointe d'un couple de V.A. permet toujours de calculer les lois marginales. En effet : pour $x \in X(\Omega)$ en appliquant la formule des probabilités totales, on a :

$$P(X = x) =$$

et idem pour calculer la loi de $Y$.
**La réciproque est fausse** : la connaissance des lois marginales ne permet pas de calculer la loi conjointe !

**Preuve** Contre-exemple : On possède une urne contenant 3 boule blanches et 4 boules noirs. On fait 2 tirages successifs.
On pose $X = 1$ si la première boule est blanche et $X = 0$ si elle est noire.
On pose $Y = 1$ si la deuxième boule est blanche et $Y = 0$ si elle est noire.
Calculer la loi conjointe et les lois marginales : d'abord lorsque les tirages ont lieu sans remise, puis lorsque les tirages ont lieu avec remise

**Remarque 3** On peut écrire : $\displaystyle\sum_{x\in X(\Omega)} P(X) = \sum_{x\in X(\Omega)} \left(\sum_{y\in Y(\Omega)} P(X=x, Y=y)\right) = 1$

et aussi : $\displaystyle\sum_{y\in Y(\Omega)} P(Y) = \sum_{y\in Y(\Omega)} \left(\sum_{x\in X(\Omega)} P(X=x, Y=y)\right) = 1$

# III Loi conditionnelle

**ATTENTION : La notion d'événement conditionnel n'a AUCUN sens.**
**La notion de variable aléatoire conditionnelle non plus**.
**En revanche** : on peut calculer la probabilité conditionnelle qu'un événement soit réalisé.
Ou la loi d'une V.A. pour une probabilité conditionnelle donnée !

**Définition 4** Soit $A$ un événement de probabilité non nulle. La loi conditionnelle de $X$ sachant $A$ est la loi de $X$ pour la probabilité $P_A$. C'est donc la donnée de $X(\Omega)$ et de $P_A(X=x)$ pour $x \in X(\Omega)$.

C'est la loi de $X$ sachant $A$ qui est conditionnelle, pas la V.A. $X$....

**Définition 5** Soit $(X, Y)$ un couple de V.A. et $x \in X(\Omega)$ tel que $P(X=x) \neq 0$.
**La loi de $Y$ sachant** $(X=x)$ est la loi de $Y$ pour la probabilité $P_{X=x}$.
C'est donc la donnée de $Y(\Omega)$ et de $P_{X=x}(Y=y)$ pour $y \in Y(\Omega)$.

**Exemple 5** Reprendre l' exemple 3 et déterminer la loi conditionnelle de $Y$ sachant que $(X=3)$.

# IV Variables aléatoires indépendantes

## IV.1 Rappels

- La notion d'indépendance est souvent justifiée par les conditions de l'expérience.
  Exemples : lancers successifs d'un dé donné (ou d'une pièce donnée), tirages avec remise dans une urne.
- Deux événements $A$ et $B$ sont indépendants **pour la probabilité** $P$ lorsque :
- La notion d'indépendance dépend de la probabilité considérée. Il est possible que $A$ et $B$ soient indépents pour $P_C$ mais pas pour $P$. Par exemple :
- Si $A$ et $B$ sont indépendants pour la probabilité $P$, alors $A$ et $\bar{B}$ le sont aussi, ainsi que $\bar{A}$ et $B$, ainsi que $\bar{A}$ et $\bar{B}$.

## IV.2 Variables aléatoires discrètes indépendantes

**Définition 6** Deux variables aléatoires discrètes $X$ et $Y$ définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ sont dites indépendantes lorsque : pour tout $A \subset X(\Omega)$ et pour tout $B \subset Y(\Omega)$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants, c'est à dire :

$$\forall A \subset X(\Omega), \forall B \subset Y(\Omega),\ P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B)$$

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

**Remarque 4** Attention : on ne suppose pas que les V.A. sont réelles ; la définition s'applique aussi à des vecteurs aléatoires par exemple.

**Proposition 2** Soient $X$, $Y$ variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Alors $X$ et $Y$ sont indépendantes si et seulement si : pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,

$$P(X = x, Y = y) = P(X = x)\, P(Y = y)$$

Autrement dit : $X \perp\!\!\!\perp Y$ si et seulement si la distribution de probabilités de $(X, Y)$ est donnée par :

$$P(X = x,\ Y = y) = P(X = x)\, P(Y = y).$$

**Exemple 6** Reprendre l'exemple 2. Les V.A. $X$ et $Y$ sont-elles indépendantes ?

**Remarque 5** $X$ et $Y$ sont indépendantes si et seulement si :

$$\forall x \in X(\Omega),\ \text{ tel que } P(X = x) \neq 0,\ P_{X=x}(Y = y) = \frac{P(X = x, Y = y)}{P(X = x)} = P(Y = y).$$

Donc les lois marginales sont égales aux lois conditionnelles.

**Proposition 3** (dém.) Si $X$ et $Y$ sont indépendantes, et si $f$ et $g$ sont deux fonctions définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement, alors $f(X)$ et $g(Y)$ sont indépendantes. Autrement dit :

$$X \perp\!\!\!\perp Y \Longrightarrow f(X) \perp\!\!\!\perp g(Y).$$

**Exemple 7** Si $X$ et $Y$ sont indépendantes, alors $X^2$ et $Y^2$ le sont aussi.
Si $X = (X_1, X_2)$ et $Y = (Y_1, Y_2, Y_3)$ sont indépendantes, alors $X_1 + X_2$ et $Y_1 \cdot Y_2$ sont indépendantes, etc...

## IV.3 Extension de la notion d'indépendance au cas de $n$ variables aléatoires

Rappel : Soit $(A_n)_{n\in[\![ 1\,;\,N ]\!]}$ une famille d'événements. C'est une famille d'événements indépendants lorsque :

**Theorème-Definition 1 Extension de la notion d'indépendance à $n$ variables aléatoires.**

Soient $X_1, \cdots, X_n$ des variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$. Les deux affirmations suivantes sont équivalentes :

(a) $\forall A_1 \subset X_1(\Omega), \cdots, \forall A_n \subset X_n(\Omega),\ P\left(\bigcap_{i=1}^{n}(X_i \in A_i)\right) = \prod_{i=1}^{n} P\left(X_i \in A_i\right).$

(b) $\forall (x_1, \cdots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega),$ les événements $(X_i = x_i)$ sont indépendants.

Lorsque ces affirmations sont vraies, on dit que $X_1, \cdots, X_n$ sont indépendantes.

**Remarque 6 Application** : Modélisation de $n$ expériences aléatoires indépendantes par une suite finie $(X_i)_{1\in[\![ 1\,;\,n ]\!]}$ de variables aléatoires indépendantes.
Exemple : On lance $n$ fois une pièce, on pose $X_i = 1$ si on obtient Pile au $i$-ème lancer et $X_i = 0$ sinon. Les variables aléatoires $X_i$ sont indépendantes. Le nombre de Pile obtenus est alors $X = \sum_{i=}^{n} X_i$.

**Attention** : Ne pas confondre les phrases suivantes, qui ne veulent pas dire la même chose :

- « les V.A. $X_1, \cdots, X_n$ sont deux à deux indépendantes »
- « les V.A. $X_1, \cdots, X_n$ sont indépendantes »

**Proposition 4 Lemme des coalitions.**
Soit $X_1, \cdots, X_n$ des variables aléatoires discrètes définies sur $(\Omega, \mathcal{T}, P)$. Soit $m \in [\![ 1\,;\,n-1 ]\!]$.

Soit $f : X_1(\Omega) \times \cdots X_m(\Omega) \to E$ et soit $g : X_{m+1}(\Omega) \times \cdots X_n(\Omega) \to E$

Si $X_1, \cdots, X_n$ sont indépendantes, alors $f(X_1, \cdots, X_m)$ et $g(X_{m+1}, \cdots, X_n)$ le sont aussi.

**Preuve** On remarque tout d'abord que : si $X_1, \cdots, X_n$ sont indépendantes, alors :
$Y = (X_1, \cdots, X_m)$ et $Z = (X_{m+1}, \cdots, X_n)$ sont des V.A. indépendantes.
On applique ensuite la proposition 3 aux V.A. $Y$ et $Z$.

**Remarque 7** On peut étendre ce lemme au cas de plus de deux coalitions :

### Extension de la notion d'indépendances à des suites infinies de V.A.

**Remarque 8** $X_1, \cdots, X_n$ sont indépendantes si, et seulement si, pour tout $J \subset [\![ 1\,;\,n ]\!]$, la famille de V.A discrètes $(X_j)_{j\in J}$ est une famille de V.A. indépendantes.

**On admet** l'existence d'espaces probabilisés contenant une suite (infinie) de variables aléatoires indépendantes de lois discrètes données.

**Définition 7** Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires discrètes.

- C'est une suite de variables aléatoires indépendantes lorsque, pour toute partie **finie** $I$ de $\mathbb{N}$, la famille $(X_i)_{i\in I}$ est une famille de variables aléatoires indépendantes.
- C'est une suite de variables aléatoires identiquement distribuées lorsqu'elles ont toute la même loi.

Lorsque que $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires **indépendantes Et identiquement distribuées**, on peut écrire que c'est « une suite de variables aléatoires *i.i.d.* ».

**Exemple 8** Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d., suivant toutes une loi de Bernoulli : $X_i \sim \mathcal{B}(p)$. Cette suite modélise le "jeu de Pile ou Face " (infinité de lancers d'une pièce, et $X_i = 1$ si Pile au $i$-ème lancer).

# V A propos d'espérance....

## V.1 Retour sur le théorème de transfert

**Théorème 1 Théorème de transfert.** Soit $X$ une V.A. définie sur $(\Omega, \mathcal{T}, P)$ et soit $f$ une fonction définie sur $X(\Omega)$ et à valeurs dans $\mathbb{R}$.

1. Si $X(\Omega) = \{x_1, \cdots, x_n\}$ est fini, alors $f(X)$ a une espérance et
$$E\left(f\left(X\right)\right) = \sum_{i=1}^{n} f(x_i)P(X = x_i)$$
2. Si $X(\Omega) = \{x_i,\ i \in \mathbb{N}\}$, alors :

   $f(X)$ est d'espérance finie *si et seulement si* $\displaystyle\sum_{i=0}^{+\infty} f(x_i)P(X = x_i)$ est $\boxed{\textbf{absolument}}$ convergente.

   Dans ce cas, on a :
$$E\left(f\left(X\right)\right) = \sum_{i=0}^{+\infty} f(x_i)P(X = x_i).$$

**Remarque 9** Dans ce théorème, $X$ est une variable aléatoire qui n'est pas nécessairement réelle. $X$ peut être un vecteur aléatoire. En revanche, $f$ étant à valeurs réelles, $f(X)$ est bien une variable aléatoire réelle.
Si $f$ est une fonction de deux variables, et $(X, Y)$ un couple de V.A., on peut donc appliquer le théorème de transfert et écrire (sous réserve d'absolue convergence) :
$$E(f(X,Y)) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} f(x,y)P(X = x, Y = y)$$
Cette formule peut aussi s'appliquer à des $n$-uplets de variables aléatoires....

## V.2 Calcul de $E(XY)$ lorsque $X$ et $Y$ sont des variables aléatoires réelles

Si $X$ et $Y$ sont deux V.A. réelles, alors le théorème de transfert permet d'écrire, **sous réserve d'absolue convergence** :
$$E(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x, Y = y)$$

**Théorème 2** (dém.) : Si $X$ et $Y$ sont deux V.A. réelles **indépendantes** et d'espérances finies, alors $XY$ est d'espérance finie et de plus $E(XY) = E(X)\ E(Y)$.

**Remarque 10** On peut étendre ce théorème au cas de $n$ variables aléatoires réelles :

La réciproque est fausse ! $E(XY) = E(X)\ E(Y) \not\Longrightarrow X$ et $Y$ indépendantes...

Contre-exemple : soit $X$ telle que $X(\Omega) = \{-1; 1; 0\}$ muni de la loi uniforme. Soit $Y = X^2$.
$X$ et $Y$ sont-elle indépendantes ? Calculer $E(XY)$, $E(X)$ et $E(Y)$. Conclure...

## V.3 Inégalité de Cauchy-Schwarz

**Théorème 3 Inégalité de Cauchy-Schwarz** : si $X^2$ et $Y^2$ sont d'espérance finie, alors $XY$ l'est aussi et de plus

$$E(XY)^2 \leqslant E(X^2)E(Y^2).$$

On a égalité si et seulement si : $\exists(\alpha,\beta)\in\mathbb{R}^2,$ tel que $(\alpha,\beta)\neq(0,0)$ et $\alpha X+\beta Y=0$ presque sûrement.

**Preuve** On suppose que $X^2$ et $Y^2$ sont d'espérance finie.
Montrons que $XY$ est aussi d'espérance finie.
On sait montre sans peine que : $\forall(x,y)\in\mathbb{R}^2,\ |xy|\leq\dfrac{x^2+y^2}{2}$.
D'après les hypothèses, on peut dire que $\left(\dfrac{X^2+Y^2}{2}\right)$ est d'espérance finie.
L'inégalité $|XY|\leq\dfrac{X^2+Y^2}{2}$ permet alors d'affirmer que $XY$ est d'espérance finie.
Posons $H(u)=E\left((uY+X)^2\right)$. Alors

$$H(u)=E(u^2Y^2+2XYu+X^2)=u^2E(Y^2)+2E(XY)u+E(X^2).$$

**1er cas** : **Supposons que** $E(Y^2)\neq 0$**.** Alors $H$ est une fonction polynomiale du second degré. Comme elle est de signe constant, son discriminant est négatif ou nul.
Or $\Delta=4E(XY)^2-4E(Y^2)\times E(X^2)$ donc

$$\Delta\leq 0 \iff E(XY)^2\leq E(X^2)E(Y^2)$$

On a égalité si et seulement $\Delta=0$ ce qui est équivalent à dire que $H$ admet une unique solution réelle :

$$\exists u_0\in\mathbb{R},\ H(u_0)=0 \text{ c'est à dire : } E\left((X+u_0Y)^2\right)=0.$$

*Rappel : si $Z$ est une variable aléatoire réelle discrète* ***positive et d'espérance nulle****, alors $(Z=0)$ est presque sûr (Prop 4 page 4 du 2e poly de probas).*
Ainsi en appliquant ce résultat à $Z=(uY+X)^2$, on obtient :

$$\exists u_0\in\mathbb{R},\ \text{tel que } P(u_0Y+X=0)=1.$$

Ou encore : il existe $(a,b)\neq(0,0)$ tel que $aX+bY=0$ presque sûrement.
**2e cas** : **Supposons que** $E(Y^2)=0$**.** Alors $Y=0$ presque sûrement.
D'une part, l'inégalité de Caucht-Scwarz est une égalité.
Par ailleurs, en prenant $a=1$ et $b=0$ on a : $aX+bY=0$ presque sûrement.
**Conclusion** : l'inégalité de Cauchy-Schwarz est toujours vraie. Et elle est une égalité si et seulement si $\exists(\alpha,\beta)\in\mathbb{R}^2$ tel que $(\alpha,\beta)\neq(0,0)$ et $\alpha X+\beta Y=0$ **presque sûrement.**

## V.4 Séries génératrices et variables aléatoires indépendantes

**Définition 8** (Rappel) Soit $X$ une V.A. à valeurs dans $\mathbb{N}$.On appelle **série génératrice de** $X$ la série

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

**Théorème 4** (Rappel) En reprenant les notations de la définition 8, on a :

- $[-1,1] \subset \mathcal{D}_{G_X}$ (en notant $\mathcal{D}_{G_X}$ le domaine de définition de $G_X$).
  Le rayon de convergence $R$ de la série génératrice vérifie $R \geqslant 1$.
- $G_X$ est de classe $\mathcal{C}^\infty$ au moins sur l'intervalle $]-1,1[$.
- $G_X(1) = 1$.
- La loi d'une variable aléatoire à valeurs dans $\mathbb{N}$ est caractérisée par sa série génératrice $G_X$.
  Plus précisément : $P(X=n) = \dfrac{G_X^{(n)}(0)}{n!}$.
  Par conséquent, si $X$ et $Y$ sont deux V.A. à valeurs dans $\mathbb{N}$, alors : $(X \sim Y) \iff G_X = G_Y$.
- $X$ admet une espérance $E(X)$ si et seulement si $G_X$ est dérivable (à gauche) en 1 et, **si tel est le cas** : $E(X) = G'_X(1)$.
- $X$ admet une variance si et seulement si $G_X$ est deux fois dérivable (à gauche) en 1. Dans le cas où $G''_X(1)$ existe, on a : $V(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$.

**Théorème 5 Série génératrice de la somme de deux V.A. indépendantes**
Soient $X, Y$ deux V.A. à valeurs dans $\mathbb{N}$, indépendantes.
Soit $R_X$ (respectivement $R_Y$) le rayon de convergence de $G_X$ (resp. $G_Y$) et soit $\mathcal{D}_X$ (respectivement $\mathcal{D}_Y$) le domaine de définition de $G_X$ (resp. $Y$).
Alors le rayon de convergence $G_{X+Y}$ de $X+Y$ vérifie : $R_{X+Y} \geqslant \min(R_X, R_Y)$. De plus :

$$\forall t \in \mathcal{D}_X \cap \mathcal{D}_Y,\ G_{X+Y}(t) = G_X(t)\, G_Y(t).$$

**Extension au cas de $n$ variables aléatoires indépendantes.**

**Exemple 9** Montrer les propositions suivantes :

- Soit $X \sim \mathcal{B}(n,p)$ et $Y \sim \mathcal{B}(m,p)$, **indépendantes.** Alors $X+Y \sim \mathcal{B}(n+m,p)$
- Soit $k \in \mathbb{N}$. Si $X_1, \cdots, X_k$ sont des V.A. **indépendantes** telles que : $\forall i \in [\![ 1\,;\,k ]\!],\ X_i \sim \mathcal{B}(n_i,p)$, alors $X_1 + \cdots + X_k \sim \mathcal{B}(n_1 + \cdots + n_k, p)$
- Soit $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$, **indépendantes.** Alors $X+Y \sim \mathcal{P}(\lambda+\mu)$
- Soit $k \in \mathbb{N}$. Si $X_1, \cdots, X_k$ sont des V.A. **indépendantes** telles que : $\forall i \in [\![ 1\,;\,k ]\!],\ X_i \sim \mathcal{P}(\lambda_i)$ Alors $X_1 + \cdots + X_k \sim \mathcal{P}(\lambda_1 + \cdots + \lambda_k)$

# VI Covariance de deux V.A.R. et variance d'une somme

**Theorème-Definition 2** Soient $X$ et $Y$ deux V.A.R. discrètes d'espérance finies, telles que $E(XY)$ existe.
Alors $E(\,(X-E(X))\,(Y-E(Y))\,)$ existe est appelé covariance de $X$ et de $Y$ et est noté $\text{cov}(X,Y)$.

**Remarque 11**
1. La covariance de $X$ et $Y$ peut être négative.
2. $V(X)=\text{cov}(X,X)$.
3. Rappelons des conditions **suffisantes** permettant d'assurer l'existence de $\text{cov}(X,Y)$
   - Si $X$ et $Y$ sont des V.A.R. admettant un moment d'ordre 2, alors $E(XY)$ existe.
   - Si $X$ et $Y$ sont indépendantes et si leurs espérances existent, alors $E(XY)$ existe et vaut :

**Proposition 5** Si $E(X), E(Y)$ et $E(XY)$ existent alors

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

**Preuve** C'est une conséquence de la linéarité de l'espérance. Remarquons tout d'abord que $(X-E(X))\,(Y-E(Y))=XY-E(X)Y-XE(Y)+E(X)E(Y)$.
Ainsi $(X-E(X))\,(Y-E(Y))$ estune somme de variables aléatoires d'espérances finies, donc c'est aussi une variable aléatoire d'espérance finie.

$$\begin{aligned}\text{cov}(X,Y) &= E\big[\,XY-E(X)Y-XE(Y)+E(X)E(Y)\,\big]\\ &= E[XY]\,-E[E(X)Y]\,-E[E(Y)X]\,+E[E(X)E(Y)]\\ &= E(XY)\,-E(X)\,E(Y)\,-E(X)\,E(Y)\,+E(X)\,E(Y)\,=\,E(XY)-E(X)\times E(Y)\end{aligned}$$

**Proposition 6** Si $X$ et $Y$ sont indépendantes et si $E(X)$ et $E(Y)$ existent, alors $\text{cov}(X,Y)=0$.
⚠ Attention : la réciproque est fausse.

**Preuve** La première affirmation est immédiate d'après le théorème 2.
Nous avons déjà montré que la réciproque est fausse.

**Définition 9** Deux variables aléatoires dont la covariance est nulle sont dites décorrélées.

**Proposition 7** Si $X$ et $Y$ sont deux V.A.R. admettant des moments d'ordre 2, alors V(X+Y) existe et

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

**Proposition 8** Si $X_1,\cdots,X_n$ sont $n$ VAR admettant des moment d'ordre 2, alors $V\left(\sum_{i=1}^{n}X_i\right)$ existe et vaut :

$$V\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}V(X_i)+2\times\sum_{1\leq i<j\leq n}\text{cov}(X_i,X_j)$$

**Preuve** On suppose vérifiées les hypothèses du théorème. Justifions que $V\left(\sum_{i=1}^{n} X_i\right)$ existe.

Posons $Y = \sum_{i=1}^{n} X_i$. Il faut donc justifier que $Y^2$ est d'espérance finie. Or $Y^2 = \sum_{i=1}^{n} X_i^2 + 2 \sum_{1\leq i<j\leq n} X_iX_j$.
Par hypothèse : $\forall i \in [\![ 1 \,;\, n ]\!]$, $E(X_i^2)$ existe.
Par ailleurs, on sait que l'existence de moments d'ordre 2 pour $X_i$ et $X_j$ assure l'existence de $E(X_iX_j)$ pour tout $i \neq j$ (ainsi que celle de $E(X_i)$ pour tout $i$).
Donc $E(Y^2)$ existe, donc $V(Y)$ existe.

$$\begin{aligned}
V\left(\sum_{i=1}^{n} X_i\right) &= E\left(\left(\sum_{i=1}^{n} X_i\right)^2\right) - \left(E\left(\sum_{i=1}^{n} X_i\right)\right)^2 \\
&= E\left(\sum_{i=1}^{n} X_i^2 + 2 \sum_{1\leq i<j\leq n} X_iX_j\right) - \left(\sum_{i=1}^{n} E\left(X_i\right)\right)^2 \\
&= \sum_{i=1}^{n} E(X_i^2) + 2 \sum_{1\leq i<j\leq n} E(X_iX_j) - \left(\sum_{i=1}^{n} \left(E\left(X_i\right)\right)^2 + 2 \sum_{1\leq i<j\leq n} E(X_i) \times E(X_j)\right) \\
&= \sum_{i=1}^{n} \left(E(X_i^2) - (E(X_i))^2\right) + 2 \sum_{1\leq i<j\leq n} (E(X_iX_j) - E(X_i)E(X_j)) \\
&= \sum_{i=1}^{n} V(X_i) + 2 \sum_{1\leq i<j\leq n} \operatorname{cov}(X_i, X_j)
\end{aligned}$$

**Corollaire 1** Si $X_1, \cdots, X_n$ sont $n$ VAR admettant des moment d'ordre 2 et **deux à deux décorrélées**, alors $V\left(\sum_{i=1}^{n} X_i\right)$ existe et vaut : $V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$

**Preuve** Immédiat d'après la propriété précédente. Remarquons que l'on peut appliquer ce théorème au cas où les V.A. sont indépendantes car : si les V.A. sont indépendantes, elles sont *a fortiori* indépendantes 2 à 2 et donc 2 à 2 décorrélées.

# VII Inégalités probabilistes

**Théorème 6 Inégalité de Markov.**
Soit $X$ une V.A.R. discrète **à valeurs positives** et d'espérance finie. Alors :

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

**Preuve** Soit $a > 0$ et soit $Y$ la fonction définie sur $\Omega$ par

$$Y(\omega) = a \text{ si } a \leqslant X(\omega) \text{ et } Y(\omega) = 0 \text{ sinon.}$$

Alors $\forall \omega \in \Omega,\ Y(\omega) \leqslant X(\omega)$. Par ailleurs $Y$ est une variable aléatoire réelle discrète. En effet :

- elle est bien définie sur $\Omega$
- $Y(\Omega)$ est fini
- $Y^{-1}(\{a\}) = (X \geqslant a)$ est bien un événement, ainsi que $Y^{-1}(\{0\}) = (X < a)$

Enfin, $E(Y)$ existe car $Y$ est une V.A.R. finie. De plus $E(Y) = 0 \cdot P(Y = 0) + a \cdot P(Y = a) = a \cdot P(X \geqslant a)$.
Par croissance de l'espérance : $0 \leqslant E(Y) \leqslant E(X)$ c'est à dire : $a \cdot P(X \geqslant a) \leqslant E(X)$.
Donc $P(X \geqslant a) \leqslant \dfrac{E(X)}{a}$.

**Théorème 7** Inégalité de Bienaymé-Tchébychev.
Soit $X$ une VAR discrète admettant une espérance $E(X)$ et un variance $V(X)$. Alors :

$$\forall \varepsilon > 0,\ P(|X - E(X)| \geqslant \varepsilon) \leqslant \frac{V(X)}{\varepsilon^2} = \frac{(\sigma_X)^2}{\varepsilon^2}.$$

**Preuve** Il suffit d'appliquer l'inégalité de Markov à la variable aléatoire positive $(X - E(X))^2$. On obtient alors, pour $a = \epsilon^2 > 0$ :

$$P\left( (X - E(X))^2 \geqslant \varepsilon^2 \right) \leqslant \frac{V(X)}{\varepsilon^2}$$

On remarque ensuite que : $(X - E(X))^2 \geqslant \varepsilon^2 \iff |X - E(X)| \geqslant \varepsilon$

d'où : $P\left( |X - E(X)| \geqslant \varepsilon \right) \leqslant \dfrac{V(X)}{\varepsilon^2}$

**Remarque 12** C'est l'inégalité de Bienaymé-Tchébychev qui permet de comprendre ce que mesure la variance : pour $\varepsilon$ fixé, la probabilité que l'écart entre $X$ et $E(X)$ soit supérieur à $\varepsilon$ est d'autant plus petite que $V(X)$ est faible. La variance donne donc une indication de la dispersion de $X$ autour de son espérance, c'est à dire sa tendance à s'écarter de sa moyenne.

**Théorème 8** Autre version de l'inégalité de Bienaymé-Tchébychev. En reprenant les mêmes hypothèses que précédemment, on a : $P(|X - E(X)| < \varepsilon) \geqslant 1 - \dfrac{V(X)}{\varepsilon^2}$

**Preuve** On reprend l'inégalité $P(|X - E(X)| \geqslant \varepsilon) \leq \dfrac{V(X)}{\varepsilon^2}$. Par passage au complémentaire, on en déduit :

$$1 - P(|X - E(X)| < \varepsilon) \leq \frac{V(X)}{\varepsilon^2}, \quad \text{ce qui s'écrit : } P(|X - E(X)| < \varepsilon) \geqslant 1 - \frac{V(X)}{\varepsilon^2}$$

**Exemple 10** On réalise 400 fois la même expérience, dont la probabilité de succès est 0,8=$\frac{4}{5}$. On suppose que les 400 expériences sont indépendantes. Soit $X$ le nombre de succès obtenus. Calculer $E(X)$ puis donner un minorant de $P(300 < X < 340)$.

**Exemple 11** On utilise un dé équilibré. Cherchons le nombre de lancers qu'il faut effectuer pour pouvoir affirmer avec un risque d'erreur inférieur à 5%, que la fréquence d'apparition du numéro 1 au cours de ces $n$ lancers sera dans l'intervalle $]\frac{1}{6} - \frac{1}{100}, \frac{1}{6} + \frac{1}{100}[$ ?

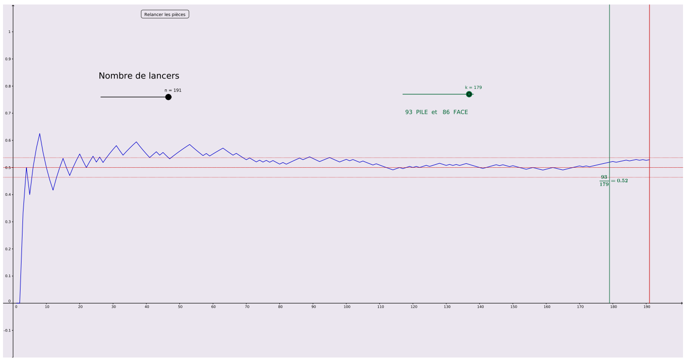
# VIII Loi faible des grands nombres

**Théorème 9 Loi faible des grands nombres.**
Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires i.i.d. admettant un moment d'ordre 2.
Alors si $S_n = \sum_{k=1}^{n} X_k$, $m = \mathrm{E}(X_1)$ et $\sigma = \sigma(X_1)$, on a :

$$\text{pour tout } \varepsilon > 0 \ : \ P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \xrightarrow[n\to\infty]{} 0.$$

$$\text{Plus précisément } : \forall \varepsilon > 0,\ P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$$

**Avant de le montrer, interprétons ce résultat**. Considérons des lancers successifs d'une pièce équilibrée.




*Intuitivement*, la fréquence d'apparition de « Pile » lors d'un très grand nombre de lancers devrait être proche de $\frac{1}{2}$. Rien ne nous garantit qu'au cours d'une série particulière de 10 000 lancers d'une pièce équilibrée, la fréquence d'apparition soit proche de $\frac{1}{2}$. **En revanche**, la *probabilité* que la fréquence soit proche de $1/2$ tend vers 1 quand le nombre de tirage tend vers $+\infty$.

La V.A.R. $\frac{S_n}{n}$ représente **la fréquence** de succès lors de $n$ épreuves de Bernoulli indépendantes. Plus généralement, réalisons une expérience aléatoire dont le résultat est une variable aléatoire $X$. La loi de $X$ est inconnue, mais on a bon espoir de la connaître de mieux en mieux en répétant un grand nombre de fois l'expérience. A chacune des expériences, correspond une variable aléatoire $X_1, X_2, \cdots, X_n$. Raisonnablement, on peut supposer que les $n$ variables aléatoires réelles sont indépendantes et suivant une même loi (mais c'est une hypothèse...). Tout expérimentateur pensera à calculer la moyenne des valeurs observées

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

Il est important de noter que $M_n$ est elle-même une variable aléatoire, somme de $n$ variables aléatoires indépendantes. Il n'y a aucune raison de supposer que $M_n$ suit la même loi que $X$ (et ce n'est généralement pas le cas).
L'idée fondamentale est que $M_n$ représente une approximation « acceptable »de l'espérance de $X$ et qu'elle permet de s'en approcher de mieux en mieux au fur et à mesure de la répétition des expériences.

**Preuve** D'après l'inégalité de Bienaymé-Tchebychev : $P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{V\left(\frac{S_n}{n}\right)}{\varepsilon^2}$.
Or $V\left(\frac{S_n}{n}\right) = \frac{1}{n^2} V(S_n)$.
Comme les $X_i$ sont indépendantes : $V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i) = nV(X_1)$ et donc

$$V\left(\frac{S_n}{n}\right) = \frac{nV(X_1)}{n^2} = \frac{\sigma^2}{n}$$