Sommaire

1. Espace probabilisé

1.1. Univers et événements

La théorie des probabilités permet l'étude de phénomènes ayant un aspect hasardeux, ou bien, dit d'une façon plus savante, aléatoire.
On appelle cela une expérience aléatoire.

Parmi les cas les plus simples et usuels :

Phénomène	Résultats possibles
lancer d'une pièce	{pile, face}
lancer d'un dé à jouer	$\{1, 2, 3, 4, 5, 6\}$
lancer de deux dés	$\{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$
lancer de deux dés jusqu'à obtenir un double six	$\mathbb{N}^*$

Définition : Les résultats possibles sont appelé les **éventualités**.
L'ensemble des résultats possibles est appelé **univers**, noté souvent Ω.
Un ensemble d'éventualités est un **événement**. *Cet ensemble peut être vide !*
Un événement qui ne contient qu'une seule éventualité est un **événement élémentaire**.
Ainsi : A est un événement $\Leftrightarrow A \subset \Omega \Leftrightarrow A \in \mathscr{P}(\Omega)$.

> Ce chapitre se limite aux univers finis !
> Le dernier exemple n'est pas un univers fini...

- L'ensemble des événements est l'ensemble des parties de Ω, l'univers, noté $\mathscr{P}(\Omega)$;
- L'événement vide est dit **impossible** ;
- L'événement Ω est dit **certain** ;
- Le complémentaire d'un événement A est l'événement **contraire** $\overline{A}$;
- Deux événements disjoints sont dit **incompatibles**.

Exemple : Si on lance un dé :
- faire 6 est une éventualité, assimilée à l'événement élémentaire $\{6\}$;
- faire un résultat pair est un événement, noté aussi $\{2, 4, 6\}$;
- $\{2, 4, 6\}$ et faire un as sont deux événements incompatibles.
- $\{2, 4, 6\}$ et $\{1, 3, 5\}$ sont deux événements contraires.

Les probabilités se font avec des *ensembles*. Cependant, dans la théorie des probabilités, certaines expressions *ensemblistes* sont remplacées par des synonymes *probabilistes*.
On travaille ici dans Ω.

Notation	Vocabulaire ensembliste	Vocabulaire probabiliste
$\emptyset$	Ensemble *vide*	Événement *impossible*
Ω		Événement *certain*
$\omega \in \Omega$	*Élément*	Événement *élémentaire*, *Éventualité*
$A \subset \Omega$	*Sous*-ensemble	*Événement*
$A \cup B$	*Union* de A et B	A *ou* B
$A \cap B$	*Intersection* de A et B	A *et* B
$\complement_\Omega A$	*Complémentaire* de A dans Ω	Événement *contraire*, $\overline{A}$
$A \cap B = \emptyset$	A et B sont *disjoints*	A et B sont *incompatibles*

1.2. Probabilité sur un univers fini

Définition : Une probabilité sur un univers fini Ω est une application $P : \mathscr{P}(\Omega) \to [0,1]$ vérifiant :
- $P(\Omega) = 1$;
- $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$;
- $P\left(\overline{A}\right) = 1 - P(A)$.

La probabilité de la réunion de deux événements incompatibles est la somme des probabilités de ces événements.

Définition : La **probabilité uniforme**, appelée encore **équiprobabilité** est la probabilité définie sur un univers à n éléments par : $\forall \omega \in \Omega,\ P(\omega) = \dfrac{1}{n}$.

L'hypothèse d'équiprobabilité est classique dans les jeux de pile ou face, les jeux de dés, les tirages d'une carte d'un paquet, d'une boule d'une urne...

1.3. Espace probabilisé

Définition : Un **espace probabilisé** est un couple « univers – probabilité », ou encore (Ω, P).

Théorème : L'événement impossible étant incompatibles avec tous les autres, sa probabilité est nulle : $P(\emptyset) = 0$.
$\forall A, B \in \mathscr{P}(\Omega),\ P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
$A \subset B \Rightarrow P(A) \leqslant P(B)$.

1.4. Indépendance et conditionnement

Définition : Deux événements A et B sont **indépendants** si et seulement si : $P(A \cap B) = P(A) \,.\, P(B)$.

On ne confondra pas événements indépendants et événements incompatibles !
Avec un dé à jouer, les événements *nombre pair* et *nombre impair* sont incompatibles mais pas indépendants !

Exemple : Si on lance deux pièces discernables et non truquées, les événements « la première est pile », et, « la seconde est pile » sont indépendants...
La démonstration est laissée au lecteur !

Définition : Si $P(B) > 0$, on appelle **probabilité conditionnelle** de A sachant B,
le réel : $P_B(A) = P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.

Théorème : A et B sont indépendants si et seulement si $P_B(A) = P(A|B) = P(A)$.

1.5. Formules des probabilités composées et des probabilités totales

On a d'abord la formule des probabilités composées, qu'on écrit arbitrairement pour 3 événements.

Théorème : Si on a des événements tels que $P(A_1 \cap A_2 \cap A_3) > 0$, alors :

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2)$$

Elle se généralise pour n événements

Théorème : Si on a $A_1, A_2, \dots A_n$ une famille d'événements de conjonction non impossible, c'est à dire telle que : $P\left(\bigcap_{1 \leqslant k \leqslant n} A_k\right) \neq 0$, alors :

$$P\left(\bigcap_{1 \leqslant k \leqslant n} A_k\right) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \cdots \cdot P(A_n|A_1 \cap A_2 \cdots \cap A_{n-1})$$

Et on a aussi la formule des probabilités totales, toujours écrites pour 3, puis pour n, événements.
Il nous faut d'abord définir :

Définition : (A_1, A_2, A_3) forme un **système complet d'événements** si et seulement si :

$$\begin{cases} A_1 \cup A_2 \cup A_3 = \Omega \\ A_1 \cap A_2 = \emptyset,\ A_2 \cap A_3 = \emptyset,\ A_1 \cap A_3 = \emptyset \end{cases}$$

Définition : $(A_1, A_2, \dots, A_n)$ forme un **système complet d'événements** si et seulement si :

$$\begin{cases} A_1 \cup A_2 \cup \cdots \cup A_n = \Omega \\ i \neq j \Rightarrow A_i \cap A_j = \emptyset \end{cases}$$

Théorème :
Si on a (A_1, A_2, A_3) un système complet d'événements, alors, pour tout événement B, on a :

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3)$$

Théorème :
Si on a $(A_1, \dots, A_n)$ un système complet d'événements, alors, pour tout événement B, on a :

$$P(B) = \sum_{k=1}^{n} P(B \cap A_k) = \sum_{k=1}^{n} P(B|A_k) \cdot P(A_k)$$

> La formule des probabilités totales s'utilise souvent avec un système complet composé de 2 événements A et son contraire $\overline{A}$.
> Cela donne, pour tout événement B : $P(B) = P(B \cap A) + P(B \cap \overline{A}) = P(B|A) \cdot P(A) + P\left(B|\overline{A}\right) \cdot P\left(\overline{A}\right)$.

1.6. Formule de Bayes

Ce qui nous donne facilement la formule de Bayes :

Théorème : Si $P(A) > 0$ et $P(B) > 0$, alors : $P_B(A) = P(A|B) = \dfrac{P_A(B) \cdot P(A)}{P(B)} = \dfrac{P(B|A) \cdot P(A)}{P(B)}$

Démonstration : Il suffit d'écrire : $P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$. ∎

Cette formule se généralise avec un système complet d'événements $(A_1, A_2, \ldots, A_n)$, c'est toujours la formule de Bayes :

Théorème : Soit un système complet d'événements $(A_1, A_2, \ldots, A_n)$, alors : $P(A_j|B) = \dfrac{P(B|A_j) \cdot P(A_j)}{\sum\limits_{i=1}^{n} P(B|A_i) \cdot P(A_i)}$.

Démonstration : Dans la formule précédente, il suffit, d'une part de remplacer A par A_j, et d'autre part, d'écrire : $P(B) = \sum\limits_{i=1}^{n} P(B \cap A_i) = \sum\limits_{i=1}^{n} P(B|A_i) \cdot P(A_i)$. ∎

> Cette dernière formule s'utilise souvent aussi avec un système complet composé de 2 événements A et son contraire $\overline{A}$.
> On obtient : $P(A|B) = \dfrac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P\left(B|\overline{A}\right) \cdot P\left(\overline{A}\right)}$.

Exemple : Cette formule permet, en quelques sorte, de remonter dans le passé. Considérons une urne contenant 2 boules blanches et 3 boules noires, indiscernables au toucher.
On tire une boule sans la regarder, on la garde.
Puis on tire une seconde boule.
On appelle A l'événement : la première boule est blanche, et B, l'événement : la deuxième boule est blanche.
La question est la suivante : Si la seconde boule est blanche, que peut-on dire sur la première ?
Il est clair que $P(A) = \dfrac{2}{5}$ et $P\left(\overline{A}\right) = \dfrac{3}{5}$, que $P(B \cap A) = \dfrac{1}{4}$ et que $P\left(B \cap \overline{A}\right) = \dfrac{2}{4}$.
On en déduit : $P(B) = \dfrac{1}{4} \cdot \dfrac{2}{5} + \dfrac{2}{4} \cdot \dfrac{3}{5} = \dfrac{2}{5}$.
Et enfin : $P(A|B) = \dfrac{1}{4}$, en vous laissant le calcul !
Ainsi, en voyant la deuxième boule, on a pu affiner la probabilité de la première !

2. Variables aléatoires sur un univers fini

2.1. Variable aléatoire

Définition : Une **variable aléatoire** est une application : $\Omega \to E$.
Quand $E = \mathbb{R}$, on parle de variable aléatoire réelle, ce qui est le cas le plus courant pour nous.

Exemple : Cet exemple nous servira de fil rouge.
On joue à pile ou face, contre une banque, avec une pièce non truquée, donc une équiprobabilité.
On mise 1 €. Si c'est pile, on perd notre €, si c'est face, la banque nous donne 4 €.
On a donc : $\Omega = \{\text{pile,face}\}$, ainsi que : P(pile)=P(face)= 1/2.
En appelant X la variable aléatoire de nos gains, on a : $X(\text{pile}) = 0 - 1 = -1$ et $X(\text{face}) = 4 - 1 = 3$.

Définition : La **loi** de la variable aléatoire X est, pour tous les $x \in X(\Omega)$, la donnée de $P(X = x)$

Exemple : En continuant notre exemple, on a $X(\Omega) = \{-1, 3\}$, et $P(X = -1) = P(X = 3) = 1/2$.

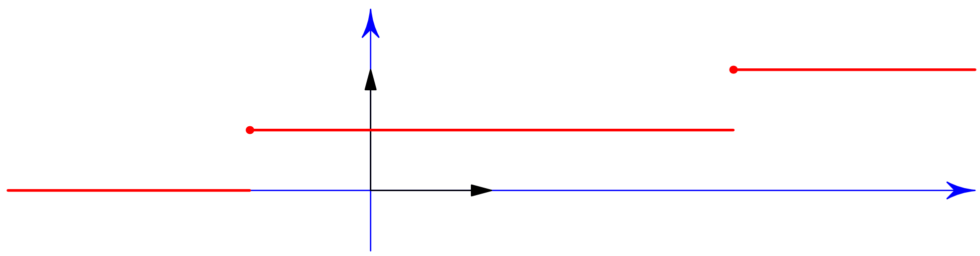
Théorème : La loi d'une variable aléatoire X, sur espace probabilisé (Ω, P), permet de définir une probabilité sur $X(\Omega)$, en posant : $P_X(x) = P(X = x)$.

Définition : La **fonction de répartition** de la variable aléatoire *réelle* X sur l'espace probabilisé (Ω, P), notée F_X, est une application : $\mathbb{R} \to [0, 1]$ définie par : $F_X(x) = P(X \leqslant x)$.

Exemple : Toujours avec notre exemple, la fonction de répartition F_X est :

- nulle sur $]-\infty, -1[$;
- vaut $1/2$ sur $[-1, 3[$;
- puis vaut 1 sur $[3, +\infty[$.

On a représenté le graphe de cette fonction de répartition.



2.2. Espérance

Définition : Soit X, une variable aléatoire *réelle*, sur espace probabilisé fini (Ω, P).
Alors **l'espérance** de X est donnée par : $E(X) = \sum_{\omega \in \Omega} P(\{\omega\})\, X(\omega)$.

> L'espérance est la moyenne des valeurs de X pondérée par leur probabilité.

On a maintenant le petit théorème de transfert :

Théorème : Soit X, une variable aléatoire *réelle*, sur espace probabilisé fini (Ω, P).
Ω étant fini, il en est de même pour $X(\Omega)$.
Alors l'espérance de X vaut : $E(X) = \sum_{x \in X(\Omega)} x \,.\, P(X = x)$.

Démonstration : La probabilité que $X = x$ est égale à la somme des probabilités des ω tels que $X(\omega) = x$.
En regroupant les ω selon leur image, on a le résultat annoncé. ∎

Puis le théorème de transfert :

Théorème : Soit X, une variable aléatoire *réelle*, sur espace probabilisé fini (Ω, P).
Ω étant fini, il en est de même pour $X(\Omega)$.
On considère aussi l'application φ de $\mathbb{R} \to \mathbb{R}$.
Alors l'espérance de $\varphi(X)$ vaut : $E(\varphi(X)) = \sum_{x \in X(\Omega)} \varphi(x)\, P(X = x)$.

Théorème : L'espérance est une application linéaire ; en particulier, $E(aX + b) = a\, E(x) + b$.

Exemple : Si Y est la variable aléatoire qui vaut le résultat dans un lancer d'un dé non truqué, alors, $E(Y) = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = \frac{7}{2}$.

Exemple : Toujours avec notre fil rouge,...
En utilisant le petit théorème de transfert : $E(X) = -1 \,.\, 1/2 + 3 \,.\, 1/2 = 1$.
Notre espérance de gain est strictement positive, on a intérêt à jouer !

Définition : On dit qu'une variable aléatoire X est **centrée** si et seulement si $E(X) = 0$.

2.3. Variance et écart-type

Définition : Soit X, une variable aléatoire *réelle*, sur espace probabilisé fini (Ω, P).
Alors, la **variance** de X est : $V(X) = E\left((X - E(X))^2\right)$.
Et **l'écart-type** de X est : $\sigma(X) = \sqrt{V(X)}$.

> La variance est une moyenne des écarts à la moyenne de la variable aléatoire : elle traduit la *dispersion* de ses valeurs.

On obtient facilement, par linéarité, le théorème de Kœning-Huygens :

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

Démonstration :
$V(X) = E\left((X - E(X))^2\right) = E\left(X^2 - 2\,X\,E(X) + E(X)^2\right) = E\left(X^2\right) - 2\,E(X)\,E(X) + E(X)^2 = E\left(X^2\right) - E(X)^2$ ■

Théorème : Si X est une variable aléatoire réelle, a et b deux réels, alors : $V(aX + b) = a^2 V(X)$.

Démonstration : $E((aX + b)^2) = E(a^2X^2 + 2abX + b^2) = a^2E(X^2) + 2abE(X) + b^2$,
tandis que $\left(E(aX + b)\right)^2 = (aE(X) + b)^2 = a^2E(X)^2 + 2abE(X) + b^2$.
En soustrayant, et après simplification, on a bien : $V(aX + b) = a^2V(X)$. ■

Exemple : On reprend notre variable aléatoire Y de notre lancer d'un dé.
Alors : $E\left(Y^2\right) = 1\frac{1}{6} + 4\frac{1}{6} + 9\frac{1}{6} + 16\frac{1}{6} + 25\frac{1}{6} + 36\frac{1}{6} = \frac{91}{6}$.
Et donc : $V(Y) = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$ et aussi : $\sigma(Y) = \frac{1}{2}\sqrt{\frac{35}{3}}$.

Enfin, on obtient aussi l'inégalité de Bienaymé-Chebychev :

Théorème : Soit X, une variable aléatoire *réelle*, sur espace probabilisé fini (Ω, P).
Alors, $\forall t \in \mathbb{R}^*$, $P\left(|X - E(X)| \geqslant t\right) \leqslant \frac{V(X)}{t^2}$.

2.4. Lois usuelles

- Loi certaine : $X(\Omega)$ ne contient qu'une seule valeur.
  Son espérance est cette valeur et sa variance est nulle.
- Loi uniforme : Toutes les valeurs x de $X(\Omega)$ ont la même probabilité.
  En particulier, quand on a équiprobabilité sur Ω et la variable aléatoire X injective, alors la loi de X est uniforme.
  En effet, dans ce cas : $P(x = X(\omega)) = P(\{\omega\})$.
- Loi de Bernoulli de paramètre $p \in [0, 1]$, notée $\mathscr{B}(p)$.
  X ne prend que deux valeurs, 0 et 1, et : $P(X = 1) = p$.
  Son espérance est p et sa variance $p(1 - p) = pq$, en notant $q = 1 - p$.
- Loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0, 1]$ notée $\mathscr{B}(n, p)$
  X ne prend que $n + 1$ valeurs, $0, 1, 2, \dots, n$ et : $P(X = k) = p^k(1 - p)^{n-k}\binom{n}{k}$.
  Son espérance est np et sa variance est $np(1 - p) = npq$, en notant $q = 1 - p$.

3. Couple de variables aléatoires

3.1. Lois...

Définition : Soit X, Y, deux variables aléatoires sur espace probabilisé fini (Ω, P).
Le **couple de variables aléatoires** (X, Y) est l'application : $\Omega \to E$ qui, à ω, associe le couple $(X(\omega), Y(\omega))$.

Définition : La **loi du couple**, ou **loi conjointe**, de (X, Y) est, pour tous les $x \in X(\Omega)$ et pour tous les $y \in Y(\Omega)$, la donnée de $P(X = x \cap Y = y)$
Les lois de X et de Y sont les **lois marginales** du couple (X, Y).

Théorème : Les lois marginales ne déterminent pas la loi conjointe

Démonstration : Il suffit bien sûr d'un contre exemple.
On choisit un lancer simultané de deux dés, chacun des deux dés étant parfait, ce qui donne une équiprobabilité.
On choisit les variables aléatoires X, la somme des deux dés, et Y leur produit.
On a $P(X = 2) = \dfrac{1}{36}$, puisqu'il faut un double as. On a aussi $P(Y = 36) = \dfrac{1}{36}$, puisqu'il faut un double six.
Par contre, on a $P(X = 2 \cap Y = 36) = 0$, ce qui ne peut, en aucun cas, se déduire des lois marginales ! ■

> Ce qu'on vient de faire pour les couples de variables aléatoires peut se généraliser aux n-uplets : $(X_1, X_2, \ldots, X_n)$

Définition :
La **loi conditionnelle** de Y sachant $(X = x)$ est la données des $P(Y = y|X = x) = P_{X=x}(Y = y)$, pour tous les $y \in Y(\Omega)$.

3.2. Variables aléatoires indépendantes

Définition : On dit que (X, Y) est un couple de variables aléatoires **indépendantes** si et seulement si pour tout les $x \in X(\Omega)$ et $y \in Y(\Omega)$, on a : $P(X = x \cap Y = y) = P(X = x) \cdot P(Y = y)$.

Théorème : Si (X, Y) est un couple de variables aléatoires indépendantes, alors,
pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, on a : $P((X, Y) \in A \times B) = P(X \in A) \cdot P(Y \in B)$.

Théorème : Si (X, Y) est un couple de variables aléatoires indépendantes, alors, $(f(X), g(Y))$ est aussi un couple de variables aléatoires indépendantes.

3.3. Variables aléatoire mutuellement indépendantes

Définition : On dit que $(X_1, X_2, \ldots, X_n)$ sont des variables aléatoires **mutuellement indépendantes** si et seulement si : $\forall (x_1, x_2, \ldots, x_n) \in X_1(\Omega) \times X_2(\Omega) \times \cdots \times X_n(\Omega)$: $P\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) = \prod_{i=1}^{n} P(X_I = x_i)$

Théorème : Si $(X_1, X_2, \ldots, X_n)$ sont des variables aléatoires mutuellement indépendantes,
alors : $\forall (A_1, A_2, \ldots, A_n) \subset X_1(\Omega) \times X_2(\Omega) \times \cdots \times X_n(\Omega)$, les événements $(X_i \in A_i)$ sont mutuellement indépendants.

Théorème : Si $(X_1, X_2, \ldots, X_n)$ sont des variables aléatoires mutuellement indépendantes, alors, les variables aléatoires $f\left(X_1, \ldots, X_p\right)$ et $g\left(X_{p+1}, \ldots, X_n\right)$ sont aussi indépendantes.

Théorème : Si $(X_1, X_2, \dots, X_n)$ sont des variables aléatoires de BERNOULLI mutuellement indépendantes, de même loi $\mathscr{B}(p)$, alors, la variable aléatoire $S = X_1 + X_2 + \cdots + X_n$ suit la loi binomiale $\mathscr{B}(n, p)$.

Démonstration : La somme vaut k si et seulement si k variables valent 1 et les $n-k$ autres valent 0.
Il y a $\binom{n}{k}$ façons de choisir les variables qui valent 1, chacune de ces façons a une probabilité de $p^k(1-p)^{n-k}$.
Donc : $P(S = k) = \binom{n}{k}p^k(1-p)^{n-k}$, c'est la loi binomiale annoncée. ■

3.4. Covariance et coefficient de corrélation linéaire

Définition :
La **covariance** du couple de variables aléatoires (X, Y) est : $\mathrm{Cov}(X, Y) = E\Big(\big(X - E(X)\big) . \big(Y - E(Y)\big)\Big)$.

> La covariance de (X, X) est la variance de X !

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

Démonstration : Il suffit d'utiliser la linéarité de l'espérance.
$\mathrm{Cov}(X, Y) = E\Big(\big(X - E(X)\big) . \big(Y - E(Y)\big)\Big) = E(XY - XE(Y) - E(X)Y + E(X)E(Y))$
$= E(XY) - E(X)E(Y)$ ■

Théorème :
Si X et Y sont deux variables aléatoires, alors : $V(aX + bY) = a^2V(X) + 2ab\,\mathrm{Cov}(X, Y) + b^2V(Y)$.

Démonstration : $E((aX + bY)^2) = E(a^2X^2 + 2abXY + b^2Y^2) = a^2E(X^2) + 2abE(XY) + b^2E(Y^2)$.
On a aussi : $E(aX + bY)^2 = (aE(X) + bE(Y))^2 = a^2E(X)^2 + 2abE(X)E(Y) + b^2E(Y)^2$.
En faisant la différence entre ces deux égalités et en utilisant les propriétés de la variance et de la covariance, on a immédiatement le résultat demandé. ■

Définition :
Le **coefficient de corrélation linéaire** du couple de variables aléatoires (X, Y) est : $\rho(X, Y) = \dfrac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$.
C'est le quotient de la covariance par le produit des écart-types.

Théorème : On a : $|\rho(X, Y)| \leqslant 1$, ou encore : $-1 \leqslant \rho(X, Y) \leqslant 1$

Démonstration : On considère la variable aléatoire $\lambda X + Y$.
Sa variance vaut $V(\lambda X + Y) = E\left((\lambda X + Y)^2\right) - E(\lambda X + Y)^2$
$= \lambda^2E(X^2) + 2\lambda\,E(XY) + E(Y^2) - \lambda^2E(X)^2 - 2\lambda\,E(X)E(Y) - E(Y)^2$
$= \lambda^2V(X) + 2\lambda\,(E(XY) - E(X)E(Y)) + V(Y) = \lambda^2V(X) + 2\lambda\,\mathrm{Cov}(X, Y) + V(Y) \geqslant 0$
Ceci étant vrai pour tout λ, le discriminant de cette expression est négatif !
On a donc : $\mathrm{Cov}(X, Y)^2 - V(X)V(Y) \leqslant 0$, ou encore : $|\mathrm{Cov}(X, Y)| \leqslant \sigma(X)\sigma(Y)$.
On en déduit immédiatement le résultat sur le coefficient de corrélation. ■

> Le coefficient de corrélation est le cosinus de l'angle entre les deux variables aléatoires !

Théorème : Si $\rho(X, Y) = \pm 1$, alors : $\exists \mu \in \mathbb{R}$ tel que $Y - E(Y) = \mu(X - E(X))$.

Démonstration : Dans la démonstration du théorème précédent, le discriminant est nul !
D'où l'existence d'un λ tel que $V(\lambda X + Y) = 0\dots$
C'est à dire en fait $\lambda X + Y = Cte = E(\lambda X + Y)$, ce qui donne le résultat demandé. ■

Théorème :
Si X et Y sont indépendantes, alors $\rho(X, Y) = 0$.
La réciproque est fausse.

Démonstration : Comme les variables aléatoire sont indépendantes :
$E(XY) = \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_i P(X = x_i \cap Y = y_j) = \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_i P(X = x_i) P(Y = y_j)$.
D'où : $E(XY) = \sum_{x_i \in X(\Omega)} \sum_{y_j \in Y(\Omega)} x_i y_i P(X = x_i) P(Y = y_j) = \sum_{x_i \in X(\Omega)} x_i P(X = x_i) \sum_{y_j \in Y(\Omega)} y_j P(Y = y_j)$
$= E(X)E(Y)$
La covariance et le coefficient de corrélation sont donc nuls quand les variables aléatoires sont indépendantes.
Montrons maintenant que la réciproque est fausse.
On considère un lancer de dés non truqués et les variables aléatoires suivantes :

- X qui vaut 1 si on tire un as ou un 6, 0 si on tire un 2 ou un 5, et -1 si on tire un 3 ou un 4 ;
- Y qui vaut 1 si on tire un 2 ou un 5, et 0 dans les autres cas.

On a facilement $E(X) = 0$, XY qui est la variable aléatoire nulle, donc $E(XY) = 0$.
Ce qui donne $\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$, alors même que X et Y ne sont pas indépendantes :
$P(X = 1 \cap Y = 1) = 0$, alors que $p(X = 1) = P(Y = 1) = \frac{1}{3}$! ■

Théorème : En résumé, si X et Y sont indépendantes, alors :

- $E(XY) = E(X)E(Y)$
- $\mathrm{Cov}(X, Y) = 0$
- $\rho(X, Y) = 0$
- $V(X + Y) = V(X) + V(Y)$