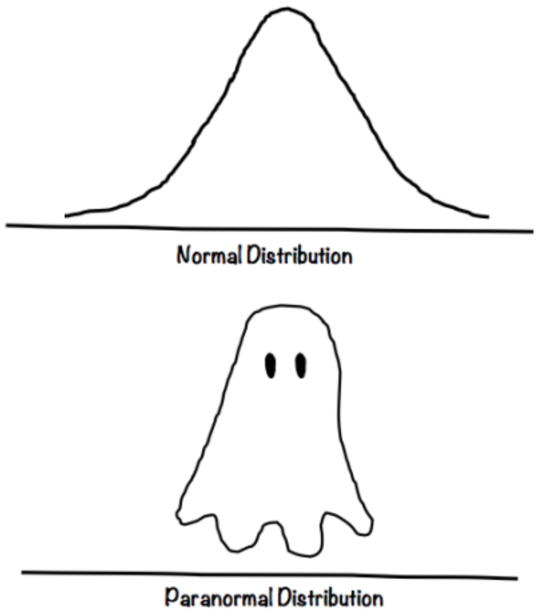
Chapitre 30
Variables Aléatoires

Simon Dauguet
simon.dauguet@gmail.com

10 juin 2025



Ce chapitre prend la suite directe du chapitre sur le dénombrement. Nous allons nous en servir pour calculer les probabilités qu'un événement donné survienne.

En réalité, ce chapitre n'est qu'une introduction au chapitre de probabilités de MP dans lequel vous mêlerez les probabilités, les séries et les séries entières. Les probabilités sont un bon moyens pour mêler un bon nombre de chapitre d'analyse. Les probabilités n'étant qu'un prétexte pour faire de l'analyse.

Toute connaissance dégénère en probabilité.

David Hume (1711-1776)

Le nom seul de calcul des probabilités est un paradoxe : la probabilité, opposée à la certitude, c'est ce qu'on ne sait pas, et comment peut-on calculer ce que l'on ne connaît pas ?

Henri Poincaré

Table des matières

1 Variables aléatoires **3**
- 1.1 Définition . . . 3
- 1.2 Loi d'une variable aléatoire . . . 5

2 Lois usuelles **7**
- 2.1 Loi uniforme . . . 7
- 2.2 Loi de Bernoulli . . . 9
- 2.3 Loi binomiale . . . 10
- 2.4 Autres lois de probabilités (HP) . . . 11

3 Opérations sur les variables aléatoires **11**

4 Couples de variables aléatoires **13**
- 4.1 Loi conjointe . . . 13
- 4.2 Lois marginales . . . 14
- 4.3 Lois conditionnelles . . . 16

5 Variables aléatoires indépendantes **18**

6 Espérance **25**
- 6.1 Définition . . . 25
- 6.2 Propriétés de l'espérance . . . 27
- 6.3 Espérances usuelles . . . 28
- 6.4 Formule de transfert . . . 30
- 6.5 Espérance d'un produit de variables aléatoires . . . 31

7 Variance, Écart type, Covariance **33**
- 7.1 Variance et Écart-type . . . 33
- 7.2 Covariance, Variance d'une somme de variables aléatoires . . . 35
- 7.3 Inégalité de Markov et Bienaymé-Tchebychev . . . 38
- 7.4 Variances usuelles . . . 39

Il y a plusieurs difficultés en probabilités. La première est un problème de vocabulaire. Les probabilités consistent essentiellement à travailler avec des objets mathématiques que nous avons déjà vu (ensembles finis, sous-ensembles etc). Mais tous ces objets auront de nouvelles dénominations. Il s'agit donc de s'habituer à ce nouveau vocabulaire.

Mais la plus grande difficulté des probabilités provient des questions de modélisations. Il s'agit de modéliser mathématiquement une situation concrète décrite. Il faut alors faire le bon choix de modélisation en fonction de ce qui nous intéresse. Et donc, il y a des problèmes d'interprétations qui rentrent en jeu.

1 Variables aléatoires

1.1 Définition

Définition 1.1 (Variable aléatoire) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. On appelle *variable aléatoire* sur $(\Omega, \mathbb{P})$, toute fonction X définie sur Ω et à valeurs dans un ensemble E.

$$X \ : \ \Omega \to E$$

Lorsque $E = \mathbb{R}$, on parle de variable aléatoire réelle.

Exemple 1.1 :
On lance deux dés et donc $\Omega = [\![1, 6]\!] \times [\![1, 6]\!]$. On pose alors $X(\omega_1, \omega_2) = \omega_1 + \omega_2$ qui donne donc la somme des faces des dés tirés. C'est une variable aléatoire réelle.

Remarque :
Si X est une variable aléatoire sur Ω fini, alors $X(\Omega)$ est une sous-ensemble fini de E qui correspond à l'ensemble des valeurs possibles de la variable aléatoire.

Remarque :
La terminologie n'est pas très heureuse. En effet, X est une fonction, ce n'est donc pas une variable. Et elle n'est pas aléatoire. Ce sont les résultats du tirages (et donc ω) qui le sont.

Définition 1.2 (Événements d'une variable aléatoire) :
Soit $X : \Omega \to E$ une variable aléatoire sur un univers fini. Pour toute partie A de E, on note

$\{X \in A\}$ ou $(X \in A)$ l'événement $X^{-1}(A) \subset \Omega$. ie

$$\{X \in A\} = \{\omega \in \Omega,\ X(\omega) \in A\}$$

En particulier, si $a \in E$, on définit l'événement

$$\{X = a\} = X^{-1}(\{a\}) = \{\omega \in \Omega, X(\omega) = a\}$$

Si X est une variable aléatoire réelle et si $a \in \mathbb{R}$, on introduit l'événement

$$\{X \leq a\} = X^{-1}(]-\infty, a]) = \{\omega \in \Omega, X(\omega) \leq a\}$$

et l'on peut définir de la même manière $\{X < a\}$, $\{X \geq a\}$, $\{X \neq a\}$ ...

Exemple 1.2 :
On lance deux dés et on considère la variable aléatoire qui donne la somme des faces obtenue. On a alors les événements

$$\{X = 12\} = \{(6,6)\},\ \{X \leq 3\} = \{(1,1),(1,2),(2,1)\}$$

par exemple

!!! ATTENTION !!!

Ce n'est qu'une notation. A strictement parler, cette notation n'est pas correcte. D'un point de vue ensembliste, c'est même une grosse erreur. Mais c'est une liberté que s'accordent les probabilistes pour de questions de sens (intuitif et ce que représente les notations).

Proposition 1.1 :
Soit $X : \Omega \to E$ une variable aléatoire sur un univers fini et $A, B \subset E$.

La notation choisie est compatible avec les opérations ensemblistes (c'est pour ça qu'on l'a choisie) :

$$\begin{aligned} \{X \in A\} \cap \{X \in B\} &= \{X \in A \cap B\} \\ \{X \in A\} \cup \{X \in B\} &= \{X \in A \cup B\} \\ \{X \in \overline{A}\} &= \overline{\{X \in A\}} \end{aligned}$$

Démonstration :
C'est de la manipulation sur les images réciproques. $\square$

Proposition 1.2 (Décomposition d'un événement d'une variable aléatoire en événements élémentaires) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini, E un ensemble, X une variable aléatoire sur Ω à valeurs dans E et $A \subset E$.

Alors
$$\{X \in A\} = \dot{\bigcup_{x \in A}} \{X = x\}$$
En particulier, avec $A = X(\Omega)$, la famille $(\{X = x\})_{x \in X(\Omega)}$ est uns système complet d'événements appelé *système complet d'événements associé à* X.

Démonstration :
On a
$$\omega \in \{X \in A\} \iff X(\omega) \in A \iff \exists x \in A,\ X(\omega) = x$$
et d'où l'égalité. $\square$

1.2 Loi d'une variable aléatoire

Définition 1.3 (Loi d'une variable aléatoire) :
Soit $X : \Omega \to E$ une variable aléatoire sur un espace probabilisé fini $(\Omega, \mathbb{P})$. On appelle loi de X, l'application
$$\mathbb{P}_X : \begin{array}{ccc} \mathcal{P}(X(\Omega)) & \to & [0,1] \\ A & \mapsto & \mathbb{P}(X \in A) \end{array}$$

Proposition 1.3 :
La loi $\mathbb{P}_X$ définit une loi de probabilité sur l'univers fini $X(\Omega)$

Démonstration :
On a $\mathbb{P}_X(X(\Omega)) = \mathbb{P}(X \in \Omega) = \mathbb{P}(\Omega) = 1$. Et le reste tombe facilement avec les manipulations sur les images directes. $\square$

Remarque :
Il n'est pas nécessaire de connaître complètement l'univers Ω pour raisonner sur des variables aléatoires.

Théorème 1.4 :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini, E un ensemble et $X : \Omega \to E$ une variable aléatoire sur Ω. La loi $\mathbb{P}_X$ de X est entièrement déterminée par les valeurs

$$\mathbb{P}_X(x) = \mathbb{P}(X = x) \text{ pour } x \in X(\Omega)$$

En pratique, ce sont par ces valeurs que l'on donne la loi de X.

Autrement dit,
$$\forall A \subset X(\Omega),\ \mathbb{P}_X(A) = \mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x)$$

Ce qui revient encore à dire que la loi $\mathbb{P}_X$ sur $X(\Omega)$ est entièrement déterminé par les probabilités des événements élémentaires de $X(\Omega)$.

Démonstration :
$(\{X = x\})_{x \in X(\Omega)}$ est un SCE de Ω. □

Exemple 1.3 :
On considère toujours un lancé de deux dés. On considère la variable aléatoire qui donne la valeur de la plus grande des deux faces obtenues. Donc $X(\Omega) = [\![1, 6]\!]$ et la loi de X est donnée par

x	1	2	3	4	5	6
$\mathbb{P}_X(x)$	$1/36$	$3/36$	$5/36$	$7/36$	$9/36$	$11/36$

Exemple 1.4 :
On considère une pièce équilibré. On la lance n fois. On note X la variable aléatoire comptant le nombre de pile obtenus. Alors, $\Omega = \{P, F\}^n$, $X(\Omega) = \{0, \ldots, n\}$ et si $k \in \{0, \ldots, n\}$, alors

$$\mathbb{P}(X = k) = \frac{\binom{n}{k}}{2^n}$$

En effet, pour obtenir k pile parmi les n lancers, cela revient à choisir k positions sur les n qui donneront piles et toutes les autres donneront automatiquement un face. Or il ya 2^n tirages possibles.

Remarque :
On peut étendre la loi de X à des événements de E qui ne sont pas nécessairement dans $X(\Omega)$:

$$\forall A \subset E,\ \mathbb{P}(X \in A) = \mathbb{P}(X \in A \cap X(\Omega)) = \sum_{x \in A \cap X(\Omega)} \mathbb{P}_X(x)$$

!!! ATTENTION !!!

La loi de X est entièrement déterminée pas X et la loi $\mathbb{P}$. Mais la loi $\mathbb{P}_X$ ne suffit pas à déterminer la variable X :

$$\mathbb{P}_X = \mathbb{P}_Y \not\Rightarrow X = Y$$

Autrement dit, à une variable aléatoire fixée est associée une unique loi de probabilité. Mais à un loi de donnée peut correspondre plusieurs variables aléatoires.

Contre-exemple :
On lance une pièce et on note $\Omega = \{P, F\}$ le résultat pile ou face du lancé.
On considère la variable aléatoire $X(P) = 1$ et $X(F) = 0$ (le joueur gagne). Alors

$$\mathbb{P}_X(0) = 1/2 \quad \text{et} \quad \mathbb{P}_X(1) = 1/2$$

On considère la variable aléatoire $Y(P) = 0$ et $Y(F) = 1$ (le joueur 2 gagne). Alors

$$\mathbb{P}_Y(0) = 1/2 \quad \text{et} \quad \mathbb{P}_Y(1) = 1/2$$

Donc $\mathbb{P}_X = \mathbb{P}_Y$ mais $X \neq Y$ (ils ne peuvent gagner en même temps, mais ils ont tous les deux mêmes chances de gagner).

2 Lois usuelles

2.1 Loi uniforme

Définition 2.1 (Loi uniforme) :
Soit X une variable aléatoire sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. On dit que X suit une loi uniforme sur un ensemble fini E si

(i) $X(\Omega) = E$

(ii) $\forall x \in E$, $\mathbb{P}(X = x) = 1/n$ avec $n = \mathrm{Card}(E)$.

Exemple 2.1 :
On tire une carte dans un jeu de 32 cartes. La variable aléatoire fournissant la hauteur de la carte suit une loi uniforme sur $\{7, 8, 9, 10, V, D, R, 1\}$.

La variable aléatoire fournissant la couler de la carte suit une loi uniforme sur l'ensemble des quatre couleurs possibles.

Exemple 2.2 :
On lance deux dés discernables équilibrés.

La variable aléatoire fournissant la valeur du premier dé suit une loi uniforme sur $[\![1, 6]\!]$.

Remarque :
Vous savez depuis longtemps que tout ensemble fini de cardinal n est en bijection avec $[\![1, n]\!]$. Donc on peut toujours se ramener à une variable aléatoire prenant ses valeurs dans cet ensemble au moyen d'une bijection.

Définition 2.2 ($X \sim \mathcal{U}(n)$) :
Si la variable aléatoire suit une loi uniforme sur $[\![1, n]\!]$, on note $X \sim \mathcal{U}(n)$.

!!! ATTENTION !!!

C'est la loi de X qui est uniforme, pas la loi $\mathbb{P}$. La loi de X peut être uniforme sans que la loi $\mathbb{P}$ sur Ω le soit.

Contre-exemple :
On lance un dé classique à six faces truqué : on obtient 1 ou 6 avec un probabilité $1/4$ et 2, 3, 4 et 5 sont obtenus avec une probabilité de $1/8$ chacun. On note X la variable aléatoire donnant la parité de la face obtenue. Donc $X(\Omega) = \{0, 1\}$. Alors $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = 1/2$. Donc la loi de X est uniforme alors que $\mathbb{P}$ ne l'est pas.

Remarque (Notations) :
Le programme de sup ne donne pas de notations particulières pour la loi suivie par une variable aléatoire. En MP, il est imposé la notation $\sim$ qui semble celle utilisée par les probabilistes. En revanche, dans les autres filières (PSI, PC et PT), la notation imposée par le programme est $\hookrightarrow$ (ce qui n'est pas en cohérence avec les usages probabilistes).

Pour rester cohérent avec la filière, nous utiliseront donc la notation $\sim$, mais vous pourrez trouver dans la littérature ou dans certains sujet la notation $\hookrightarrow$.

2.2 Loi de Bernoulli

Définition 2.3 (Loi de Bernoulli) :
On dit qu'une variable aléatoire X de $(\Omega, \mathbb{P})$, un espace probabilisé fini, suit une loi de Bernoulli de paramètre $p \in [0, 1]$ si

(i) $X(\Omega) = \{0, 1\}$

(ii) $\mathbb{P}(X = 0) = 1 - p$

(iii) $\mathbb{P}(X = 1) = p$

On note alors $X \sim \mathcal{B}(p)$

En pratique, les seuls cas intéressants seront quand $p \in]0, 1[$.

Remarque :
Les variables aléatoires X suivant une loi de Bernoulli servent à modéliser les situations à deux issues possible seulement :

- succès (valeurs 1, probabilité p)
- échec (valeurs 0, probabilité $1 - p$)

Exemple 2.3 :
Une urne contient des boules blanches en proportion p et des boules noires en proportion $q = 1 - p$. On tire une boule de cette urne. La variable aléatoire valant 1 si la boule tirée est blanche et 0 sinon est une variable aléatoire suivant la loi de Bernoulli de paramètre p.

Remarque :
Les variables aléatoires suivant une loi de Bernoulli sont exactement les fonctions indicatrices des parties F de Ω.

2.3 Loi binomiale

Définition 2.4 (Loi binomiale) :
On dit qu'une variable aléatoire X sur $(\Omega, \mathbb{P})$ un espace probabilisé fini, suit une loi binomiale de paramètre $n \in \mathbb{N}^*$ et $p \in [0, 1]$ si

(i) $X(\Omega) = [\![0, n]\!]$

(ii) $\forall k \in [\![0, n]\!]$, $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

On note alors $X \sim \mathcal{B}(n, p)$

Remarque :
$X \sim \mathcal{B}(p) \iff X \sim \mathcal{B}(1, p)$.

Exemple 2.4 :
On tire n boules avec remises d'une urne contenant des boules blanches en proportion p. Le nombre X de boules blanches tirées suit une loi binomiale de paramètre n et p.

Remarque :
Les lois binomiales sont très utiles pour décrire ce qui s'apparente à un tirage avec remise ou pour mesurer le nombre de succès lorsqu'on répète de façon indépendante une expérience dont la probabilité de réussite est p.

En quelques sortes, la loi binomiale mesure la probabilité de succès d'une répétition de façon indépendante d'expérience suivant une loi de Bernoulli.

Exemple 2.5 :
On tire n fois une pièce équilibrée et l'on compte X le nombre de faces obtenues. Alors $X \sim \mathcal{B}(n, p)$.

Exemple 2.6 :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et X une variable aléatoire sur Ω. Soit $n \in \mathbb{N}^*$ et $p \in]0, 1[$.

Montrer que si $X \sim \mathcal{B}(n, p)$, alors $Y = n - X \sim \mathcal{B}(n, 1-p)$.

2.4 Autres lois de probabilités (HP)

Ces lois ne sont pas les seules. Ce sont les seuls qui sont au programme de sup. En deuxième année vous rajouterez :

- La loi de Poisson de paramètre $\lambda > 0$: on dira que $X \sim \mathcal{P}(\lambda)$ si $X(\Omega) = \mathbb{N}$ et $\forall k \in \mathbb{N}$, $\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$
- La loi géométrique de paramètre $p \in]0,1[$: on dira que $X \sim \mathcal{G}(p)$ si $X(\Omega) = \mathbb{N}^*$ et $\forall k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = p(1-q)^{k-1}$.

Ces deux lois ne sont pas au programme de sup car elles étendent les notions de bases vues cette année. En particulier, comme elles sont à valeurs dans un ensemble infini, il y a des problèmes de convergences de séries à prendre en compte (on doit toujours avoir $\mathbb{P}(X \in X(\Omega)) = 1$ et donc $\sum_{k \in \mathbb{N}} \mathbb{P}(X = k)$ doit converger).

Il existe d'autres lois de probabilités, mais qui ne sont pas au programme de spé non plus. Par exemple, la loi hypergéométrique pour laquelle

$$\mathbb{P}(X = k) = \frac{\binom{b}{k}\binom{r}{n-k}}{\binom{b+r}{n}}$$

3 Opérations sur les variables aléatoires

Définition 3.1 (Variable aléatoire composée) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et $X : \Omega \to E$ une variable aléatoire sur E.

Si $f : X(\Omega) \to F$ est une application à valeurs dans un ensemble F, on note $f(X)$ la variable aléatoire $Y = f \circ X$ sur Ω :

$$Y = f(X) \ : \ \begin{array}{ccc} \Omega & \to & F \\ \omega & \mapsto & f(X(\omega)) \end{array}$$

!!! ATTENTION !!!

Les notations sont piégeuse ! La notation $f(X)$, d'un point de vue ensembliste, n'a pas de sens. Ce n'est pas une fonction. La fonction f ne s'applique pas aux fonction de Ω dans E. C'est un liberté sur les notations que s'accordent les probabilistes, au même tire qu'écrire $\{X \in A\}$ n'a pas de sens.

Exemple 3.1 :
On tire simultanément n boules dans une urne contenant p boules blanches et q boules noires avec $n \leq p+q$. On note X la variable aléatoire donnant le nombre de boules blanches du tirage. Déterminer en fonction de X la variable aléatoire Y donnant le nombre de boules rouges du tirage.

Remarque :
Si la fonction f présente une notation usuelle particulière, on adapte celle-ci à la description de la variable aléatoire $f(X)$. On pourra donc noter

$$\sqrt{X},\ X^2,\ |X|,\ aX+b,\ \ln(X),\ \ln(X)$$

Théorème 3.1 (Loi d'une composée de variables aléatoires [✓]) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et X et Y deux variables aléatoire sur Ω. Si $Y = f(X)$, alors la loi de Y est entièrement déterminée par celle de X :

$$\forall B \subset Y(\Omega),\ \mathbb{P}_Y(B) = \mathbb{P}_X(f^{-1}(B))$$

Démonstration :
Par définition,

$$\mathbb{P}_Y(B) = \mathbb{P}(Y \in B) = \mathbb{P}(f(X) \in B) = \mathbb{P}(X \in f^{-1}(B)) = \mathbb{P}_X(f^{-1}(B))$$

□

Remarque :
En pratique, connaître la loi de X suffira pour déterminer les lois des variables aléatoires composées déduites à partir de X.

En particulier, la loi de $f(X)$ est toujours déterminée par les événements élémentaires et donc

$$\forall y \in f(X(\Omega)),\ \mathbb{P}(f(X) = y) = \mathbb{P}(X \in f^{-1}(\{y\})) = \sum_{\substack{x \in X(\Omega) \\ f(x)=y}} \mathbb{P}(X = x)$$

car $\{f(x) = y\} = \bigcup_{\substack{x \in X(\Omega) \\ f(x)=y}} \{X = x\}$.

Exemple 3.2 :
Si X est une variable aléatoire réelle et $f(x) = ax+b$ avec $a \neq 0$, et si on note $X(\Omega) = \{x_1, \ldots, x_n\}$, alors $f(X(\Omega)) = \{y_1, \ldots, y_n\}$ où $\forall i \in \{1, \ldots, n\}$, $y_i = ax_i + b$. Et donc $\forall i \in \{1, \ldots, n\}$, $\mathbb{P}(f(X) = y_i) = \mathbb{P}(aX + b = ax_i + b) = \mathbb{P}(X = x_i)$.

Exemple 3.3 :
Si X est une variable aléatoire réelle telle que $X(\Omega) = [\![-n, n]\!]$ et $f(x) = x^2$, alors, en posant $Y = f(X)$, $Y(\Omega) = \{k^2,\ k \in \{0, \ldots, n\}\}$ et $\mathbb{P}(Y = 0) = \mathbb{P}(X = 0)$ et si $k \in \{1, \ldots, n\}$, $\mathbb{P}(Y = k) = \mathbb{P}(X = k) + \mathbb{P}(X = -k)$.

Exemple 3.4 :
Si $X \sim \mathcal{B}(p)$, alors $Y = 1 - X \sim \mathcal{B}(1-p)$.

Exemple 3.5 :
Si $X \sim \mathcal{B}(n, p)$, alors $Y \sim \mathcal{B}(n, 1-p)$.

4 Couples de variables aléatoires

4.1 Loi conjointe

Définition 4.1 (Couple de variable aléatoire) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini, X et Y deux variables aléatoires sur Ω à valeurs dans E et F respectivement. On appelle couple défini par les variables aléatoires X et Y, la variable aléatoire

$$Z = (X, Y)\ :\ \begin{array}{ccc} \Omega & \to & E \times F \\ \omega & \mapsto & (X(\omega), Y(\omega)) \end{array}$$

Exemple 4.1 :
On choisit une carte dans un jeu de 32 cartes. X représente la valeur de la carte et Y sa couleur. Alors $Z = (X, Y)$ détermine parfaitement la carte tirée.

!!! ATTENTION !!!

⚠ On a $(X,Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$ mais il n'y a pas égalité en général ! Par exemple, si $X = Y = \mathrm{Id}_{\{0,1\}}$, alors $(X,Y)(\{0,1\}) = \{(0,0),(1,1)\} \subsetneq X(\{0,1\}) \times Y(\{0,1\}) = \{(0,0),(0,1),(1,0),(1,1)\}$.

Définition 4.2 (Loi conjointe) :
Soit X et Y deux variables aléatoire sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. On appelle loi conjointe de X et de Y la loi du couple (X,Y).

Remarque :
La loi conjointe de X et de Y est entièrement déterminée par la connaissance des

$$\mathbb{P}(X = x_i, Y = y_j) \quad \text{avec} \quad x_i \in X(\Omega), y_j \in Y(\Omega)$$

En effet, comme (X,Y) est une variable aléatoire à part entière, la famille $(\{(X,Y) = (x,y)\})_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est un SCE de Ω.

Ce qui nous permet de donner la loi conjointe de X et de Y sous forme de tableau par exemple.

Remarque (Notations) :
On notera $\{(X,Y) = (x,y)\}$ ou $\{X = x\} \cap \{Y = y\}$ ou $\{X = x, Y = y\}$ ces événements.

Exemple 4.2 :
On lance deux dés équilibrés et on note X la plus grande des valeurs obtenue et Y la somme. Donner la loi conjointe de X et de Y.

4.2 Lois marginales

Définition 4.3 (Lois marginales) :
Soit Z une variable aléatoire sur $(\Omega, \mathbb{P})$, un espace probabilisé fini et à valeur dans un produit cartésien $E \times F$. On note X la première coordonnée de Z et Y la deuxième, ie on a $\forall \omega \in \Omega$, on note $X(\omega) \in E$ et $Y(\omega) \in F$ tels que $Z(\omega) = (X(\omega), Y(\omega)) \in E \times F$ (ie X est la projection

de Z sur E et Y la projection de Z sur F).
On appelle lois marginales de Z les lois de X et de Y.

Proposition 4.1 (Une loi conjointe détermine les lois marginales) :
Si Z est une variable aléatoire dans un produit cartésien, la loi de Z détermine entièrement ses lois marginales.

Démonstration :
Pour $x \in X(\Omega)$,

$$\{X = x\} = \{Z \in \{x\} \times F\}$$

donc $\mathbb{P}_X(x) = \mathbb{P}_Z(\{x\} \times F)$. Et de même, $\forall y \in Y(\Omega)$, $\mathbb{P}_Y(y) = \mathbb{P}_Z(E \times \{y\})$. $\square$

Remarque :
On pourra alors donner les lois marginales sous formes de tableau.

Exemple 4.3 :
On lance deux dés et on considère Z la variable aléatoire donnant le maximum des deux valeurs obtenus et la somme des deux valeurs obtenus. Donner les lois marginales de Z.

!!! ATTENTION !!!

Attention ! Les lois de X et de Y ne suffisent pas à déterminer la loi conjointe. Par exemple

$X\backslash Y$	0	1	$\mathbb{P}_X$
0	$1/2$	0	$1/2$
1	0	$1/2$	$1/2$
$\mathbb{P}_Y$	$1/2$	$1/2$	

et

$X\backslash Y$	0	1	$\mathbb{P}_X$
0	$1/4$	$1/4$	$1/2$
1	$1/4$	$1/4$	$1/2$
$\mathbb{P}_Y$	$1/2$	$1/2$	

Ces deux lois donnent les mêmes lois marginales. Mais donc les lois marginales seuls ne permettent pas de retrouver les lois conjointe. Autrement dit, à une loi conjointe donnée, on peut lui associée des lois marginales unique. Mais l'inverse est faux. Des lois marginales seules ne permettent pas de retrouver la loi conjointe associée. Il faut avoir les $\mathbb{P}(X = x, Y = y)$ qui est plus précis. Il faut avoir chaque élément du tableau et pas seulement les comme par colonnes et pas lignes.

4.3 Lois conditionnelles

Définition 4.4 (Lois conditionnelles) :
Soit X et Y deux variables aléatoires sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. Pour $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) > 0$, on appelle loi conditionnelle de Y sachant $X = x$, la loi de Y pour la probabilité conditionnelle $\mathbb{P}(\,\cdot\,|\{X = x\})$. Autrement dit,

$$\forall B \subset Y(\Omega),\ \mathbb{P}(Y \in B|X = x) = \frac{\mathbb{P}(Y \in B, X = x)}{\mathbb{P}(X = x)}$$

Remarque :
On a

$$\begin{aligned}
\mathbb{P}(Y \in B, X = x) &= \mathbb{P}(\{\omega \in \Omega, (Y(\omega), X(\omega)) \in B \times \{x\}) \\
&= \mathbb{P}(\{\omega \in \Omega, \omega \in Y^{-1}(B) \cap X^{-1}(\{x\})) \\
&= \mathbb{P}(Y^{-1}(B) \cap X^{-1}(\{x\})) \\
&= \mathbb{P}((Y \in B) \cap (X = x))
\end{aligned}$$

d'où la formule.

Proposition 4.2 :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et X et Y deux variables aléatoires sur Ω. La loi conditionnelle de Y sachant $X = x$ est entièrement déterminée par la connaissance des

$$\mathbb{P}(Y = y | X = x) = \frac{\mathbb{P}(Y = y, X = x)}{\mathbb{P}(X = x)} \quad \text{pour tout } y \in Y(\Omega)$$

Exemple 4.4 :
Supposons que X et Y sont des variables aléatoires de loi conjointe

$X \backslash Y$	0	1
0	0.1	0.3
1	0.2	0.4

Donné la loi de Y sachant $X = x$

Proposition 4.3 ([✓]) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Soit deux variables aléatoires X et Y.

La connaissance de la loi de X et de la loi de Y sachant $X = x$ pour tout $x \in X(\Omega)$ suffit à déterminer la loi conjointe (X, Y).

Démonstration :
On pose $Z = (X, Y)$. Soit $(x, y) \in Z(\Omega)$. Alors $x \in X(\Omega)$ et $y \in Y(\Omega)$.

Si $\mathbb{P}(X = x) = 0$, alors $\mathbb{P}(Z = (x, y)) = 0$ car $\{Z = (x, y)\} \subset \{X = x\}$.

Si $\mathbb{P}(X = x) > 0$, alors $\mathbb{P}(Z = (x, y)) = \mathbb{P}(X = x, Y = y) = \mathbb{P}(Y = y | X = x)\mathbb{P}(X = x)$. □

Exemple 4.5 :
Si X et Y sont deux variables aléatoires réelles telles que $X(\Omega) = Y(\Omega) = \{0, \ldots, n\}$, alors

$$\forall k \in \{0, \ldots, 2n\},\ \mathbb{P}(X + Y = k) = \sum_{j=0}^{k} \mathbb{P}(X = j, Y = k - j) = \sum_{j=0}^{k} \mathbb{P}(X = k - j, Y = j)$$

et

$$\forall k \in \{0, \ldots, n^2\},\ \mathbb{P}(XY = k) = \sum_{\substack{x,y \in \{0,\ldots,n\} \\ xy = k}} \mathbb{P}(X = x, Y = y)$$

Définition 4.5 (Extension aux n-uplets de variables aléatoires) :
Soit $X_1, \ldots, X_n$ des variables aléatoires sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. On appelle loi conjointes des variables $X_1, \ldots, X_n$ la loi de la variable $(X_1, \ldots, X_n)$. Elle est déterminée par les $\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)$ où $\forall i \in \{1, \ldots, n\}$, $x_i \in X_i(\Omega)$.

Inversement, si $Z = (X_1, \ldots, X_n)$ est une variable aléatoire à valeur dans un espace produit, on appelle loi marginales de Z les lois des variables $X_1, \ldots, X_n$.

Remarque :
La loi conjointe détermine les lois marginales, mais l'inverse est faux.

Remarque :
On peut définir de la même manière des lois conditionnelles :

$$\mathbb{P}(X_n = x_n | X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = \frac{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)}{\mathbb{P}(X_1 = x_1, \ldots, X_{n-1} = x_{n-1})}.$$

5 Variables aléatoires indépendantes

Définition 5.1 (Variables aléatoires indépendantes) :
Soit X, Y deux variables aléatoires sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. On dit que X et Y sont indépendantes si $\forall A, B \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega))$, $\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$. Autrement dit, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants. On notera $X \perp\!\!\!\perp Y$

Exemple 5.1 :
Une première urne contient 2 boules blanches et 3 boules noires et une seconde contient l'inverse (2 boules noires et 3 boules blanches). On jette une pièce et si on obtient Face, on tire une boule dans la première urne, sinon on tire dans la seconde urne. On note X le résultat du lancé de pièce et Y la couleur de la boule obtenue.

Montrer que X et Y ne sont pas indépendantes.

Théorème 5.1 (CNS de variables aléatoires indépendantes) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et X et Y deux variables aléatoires sur Ω. On a équivalence entre :

(i) $X \perp\!\!\!\perp Y$

(ii) $\forall x \in X(\Omega)$ tel que $\mathbb{P}(X = x) > 0$, la loi de Y sachant $X = x$ est égale à la loi de Y

(iii) $\forall (x, y) \in X(\Omega) \times Y(\Omega)$, $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$

Démonstration :
$\boxed{(i) \implies (ii)}$
Soit $x \in X(\Omega)$ avec $\mathbb{P}(X = x) > 0$. Pour $y \in Y(\Omega)$, on a

$$\mathbb{P}(Y = y | X = x) = \frac{\mathbb{P}(\{Y = y\} \cap \{X = x\})}{\mathbb{P}(X = x)} = \mathbb{P}(Y = y)$$

$\boxed{(ii) \implies (iii)}$
Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. Si $\mathbb{P}(X = x) = 0$, alors $\mathbb{P}(X = x, Y = y) = 0 = \mathbb{P}(X = x)\mathbb{P}(Y = y)$ car $\{X = x\} \cap \{Y = y\} \subset \{X = x\}$.
Si $\mathbb{P}(X = x) > 0$ alors par probabilités composées,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y|X = x) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

$\boxed{(iii) \implies (i)}$
Soit $A \subset X(\Omega)$ et $B \subset Y(\Omega)$. Par la formule des probabilités totales,

$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \sum_{(x,y) \in A \times B} \mathbb{P}(X = x, Y = y)$$

donc

$$\begin{aligned}
\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) &= \sum_{(x,y) \in A \times B} \mathbb{P}(X = x)\mathbb{P}(Y = y) \\
&= \sum_{x \in A} \sum_{y \in B} \mathbb{P}(X = x)\mathbb{P}(Y = y) \\
&= \left(\sum_{x \in A} \mathbb{P}(X = x) \right) \left(\sum_{y \in B} \mathbb{P}(Y = y) \right) \\
&= \mathbb{P}(X \in A)\mathbb{P}(Y \in B)
\end{aligned}$$

□

Exemple 5.2 :

Soit X et Y deux variables aléatoires de loi conjointe donnée par

$X\backslash Y$	0	1
0	$1/12$	$2/12$
1	$3/12$	$6/12$

Montrer que X et Y sont des variables indépendantes.

Exemple 5.3 :
Montrer que si $X, Y \sim \mathcal{B}(p)$, alors X et Y sont indépendantes si, et seulement si, $\{X = 1\}$ et $\{Y = 1\}$ sont indépendants.

Remarque :
Si X et Y sont deux variables aléatoires indépendantes, alors les lois de X et Y détermine la loi conjointe de (X, Y).

Théorème 5.2 (Composées de variables aléatoires) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et $X : \Omega \to E$ et $Y : \Omega \to F$ des variables aléatoires et $f : E \to E'$ et $g : F \to F'$ des applications, où E, F, E', F' sont des ensembles.

$$X \perp\!\!\!\perp Y \implies f(X) \perp\!\!\!\perp f(Y)$$

Démonstration :
Soit $x' \in f(X(\Omega))$ et $y' \in g(Y(\Omega))$. On a

$$\mathbb{P}(f(X) = x', g(Y) = y') = \mathbb{P}(\{X \in f^{-1}(\{x'\})\} \cap \{Y \in g^{-1}(\{y'\})\})$$

Les variables X et Y sont indépendantes, donc

$$\begin{aligned}\mathbb{P}(f(X) = x', g(Y) = y') &= \mathbb{P}(X \in f^{-1}(\{x'\}))\mathbb{P}(Y \in g^{-1}(\{y'\})) \\ &= \mathbb{P}(f(X) = x')\mathbb{P}(g(Y) = y')\end{aligned}$$

□

Exemple 5.4 :

Si $X, Y \sim \mathcal{U}(n)$ indépendantes, déterminer les lois de $X+Y$ et $X-Y$.

Définition 5.2 (Variables aléatoires mutuellement indépendantes) :
Soit $X_1, \dots, X_n$ des variables aléatoires sur Ω, un ensemble fini. On dit que $X_1, \dots, X_n$ sont mutuellement indépendantes si

$$\forall m \in \{1, \dots, n\},\ \forall i_1, \dots, i_m \in \{1, \dots, n\},\ i_1 < \dots < i_m,\ \forall (A_{i_k})_{1 \leq k \leq m} \in \prod_{k=1}^{m} \mathcal{P}(X_{i_k}(\Omega)),$$
$$\mathbb{P}\left(\bigcap_{k=1}^{m} \{X_{i_k} \in A_{i_k}\}\right) = \prod_{k=1}^{m} \mathbb{P}(X_{i_k} = A_{i_k})$$

autrement dit si pour tout $i \in \{1, \dots, n\}$ et pour tout $A_i \subset X_i(\Omega)$, les événements $\{X_1 \in A_1\}, \dots, \{X_n \in A_n\}$ sont mutuellement indépendants.

Ce qui revient à dire que $\forall (A_1, \dots, A_n) \in \mathcal{P}(X_1(\Omega)) \times \dots \times \mathcal{P}(X_n(\Omega))$, $\forall I \subset \{1, \dots, n\}$ non vide,

$$\mathcal{P}\left(\bigcap_{i \in I} \{X_i \in A_i\}\right) = \prod_{i \in I} \mathcal{P}(X_i \in A_i)$$

Remarque :
Lorsqu'on répète n fois la même expérience aléatoire et l'on note $X_1, \dots, X_n$ les résultats successifs, si le résultat d'une expérience n'influe pas sur le résultat de l'expérience suivante, les variables aléatoires $X_1, \dots, X_n$ seront supposées mutuellement indépendants.

Exemple 5.5 :
On tire des boules dans une urne contenant des boules blanches et noires. On note X_i la couleur obtenue lors du i-ème tirage.

Si le tirage a lieu avec remises, les variables aléatoires seront mutuellement indépendantes. Si l'on tire sans remise, les variables aléatoires ne sont plus indépendantes.

Théorème 5.3 (Caractérisation de variables aléatoires mutuellement indépendantes) :

Soit $X_1, \dots, X_n$ des variables aléatoires sur Ω, un ensemble fini.

Les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes si et seulement si $\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$, les événements $\{X_1 = x_1\}, \dots, \{X_n = x_n\}$ sont mutuellement indépendants.

Démonstration :
Le sens direct est évident par définition de variables aléatoires mutuellement indépendantes (prendre $A_i = \{x_i\}$).

Supposons que $\forall (x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$, les événements $\{X_1 = x_1\}, \ldots, \{X_n = x_n\}$ sont mutuellement indépendants.

Soit $(A_1, \ldots, A_n) \in \mathcal{P}(X_1(\Omega)) \times \cdots \times \mathcal{P}(X_n(\Omega))$ et $I \subset \{1, \ldots, n\}$ non vide. Alors

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i\in I}\{X_i \in A_i\}\right) &= \mathbb{P}\left(\bigcup_{x_1\in A_1,\ldots,x_n\in A_n}^{\cdot} \left(\bigcap_{i\in I}\{X_i = x_i\}\right)\right) \\
&= \sum_{x_1\in A_1,\ldots,x_n\in A_n} \mathbb{P}\left(\bigcap_{i\in I}\{X_i = x_i\}\right) && \text{loi de proba} \\
&= \sum_{x_1\in A_1}\left(\cdots\left(\sum_{x_n\in A_n}\left(\prod_{i\in I}\mathbb{P}(X_i = x_i)\right)\right)\cdots\right) && \text{hyp} \\
&= \prod_{i\in I}\left(\sum_{x_i\in A_i}\mathbb{P}(X_i = x_i)\right) && \text{factorisation} \\
&= \prod_{i\in I}\mathcal{P}(X_i \in A_i)
\end{aligned}$$

□

Proposition 5.4 :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et $X_1, \ldots, X_n$ des variables aléatoires sur Ω.

Si $X_1, \ldots, X_n$ sont mutuellement indépendantes, alors $X_1, \ldots, X_n$ sont deux à deux indépendantes.

Démonstration :
Ça provient du cas des événements. □

!!! ATTENTION !!!

Ne pas confondre l'indépendance mutuelle et l'indépendance deux à deux. Comme dans le cas des événements, ce n'est qu'une implication. La réciproque est fausse en général, dès que $n > 2$.

Contre-exemple :
Si on lance deux dés discernables et que l'on note X et Y les parités de chaque dés et Z la parité de la somme, alors les variables X, Y, Z sont deux à deux indépendantes mais pas mutuellement indépendantes.

Proposition 5.5 (Composées de variables aléatoires mutuellement indépendantes) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Soit $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes sur Ω. Soit $f_1, \ldots, f_n$ des fonctions définies sur $X_1(\Omega), \ldots, X_n(\Omega)$ respectivement.

Alors $f_1(X_1), \ldots, f_n(X_n)$ sont des variables aléatoires mutuellement indépendantes.

Démonstration :
Il suffit d'adapter la démo pour un couple de variables aléatoires. □

Définition 5.3 (Variables aléatoires indépendantes identiquement distribuées (VAIID)) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et $X_1, \ldots, X_n$ des variables aléatoires sur Ω.

Si $X_1, \ldots, X_n$ sont mutuellement indépendantes et qu'elles suivent toutes la même loi (*i.e.* $\forall i \in \{1, \ldots, n\}$, $X_i \sim X_1$), on dit que les variables aléatoires $X_1, \ldots, X_n$ sont des *variables aléatoires indépendantes identiquement distribuées* (VAIID).

Définition 5.4 (Schéma de Bernoulli) :
On appelle schéma de Bernoulli de paramètre $p \in]0,1[$ toute suite $X_1, \ldots, X_n$ VAIID de loi $\mathcal{B}(p)$.

Exemple 5.6 :
On lance n fois une pièce de on pose $X_i = 1$ ou 0 selon que le i-ème tirage a donné Face ou non. Les variables $X_1, \ldots, X_n$ définissent un schéma de Bernoulli.

Exemple 5.7 :
On tire avec remise n boules dans une urne contenant des boules blanches et noires. On pose $X_i = 1$ ou 0 selon que le i-ème tirage est une boule blanche ou noire. On a encore un schéma de Bernoulli.

Théorème 5.6 (Loi d'un schéma de Bernoulli) :
Si $X_1, \ldots, X_n$ est un schéma de Bernoulli de paramètre $p \in]0,1[$, on note

$$S = X_1 + \cdots + X_n$$

alors $S \sim \mathcal{B}(n,p)$.

Autrement dit, si $X_1, \ldots, X_n \sim \mathcal{B}(p)$ sont mutuellement indépendantes, alors $\sum_{i=1}^n X_i \sim \mathcal{B}(n,p)$.

Démonstration :
On raisonne par récurrence sur n.

La propriété est vraie quand $n = 0$ car $\mathbb{P}(S = 0) = 1$. Supposons la propriété vraie à un certain rang $n \geq 0$.

Soit $X_1, \ldots, X_{n+1}$ un schéma de Bernoulli de paramètre p. Posons $S = X_1 + \cdots + X_{n+1}$ et $T = X_1 + \cdots + X_n$. Par hypothèse $T \sim \mathcal{B}(n,p)$.

De plus, comme $\forall i \in \{1, \ldots, n\}$, $X_i(\Omega) = \{0,1\}$, on a $S(\Omega) = \{0, \ldots, n+1\}$. On note que $\{S = 0\} = \cap_{i=1}^{n+1} \{X_i = 0\}$ et par indépendance mutuelle, $\mathbb{P}(S = 0) = \prod_{i=1}^{n+1} \mathbb{P}(X_i = 0) = (1-p)^{n+1}$. Soit $k \in \{1, \ldots, n+1\}$. Les événements $\{X_{n+1} = 0\}$ et $\{X_{n+1} = 1\}$ constituent un système complet d'événements.

$$\mathbb{P}(S = k) = \mathbb{P}(\{S = k\} \cap \{X_{n+1} = 0\}) + \mathbb{P}(\{S = k\} \cap \{X_{n+1} = 1\})$$

Or

$$\{S = k\} \cap \{X_{n+1} = 0\} = \{T = k\} \cap \{X_{n+1} = 0\}$$

De plus, les événements $\{T = k\}$ et $\{X_{n+1} = 0\}$ sont indépendants.

$$\mathbb{P}(\{S = k\} \cap \{X_{n+1} = 0\}) = \mathbb{P}(T = k)\mathbb{P}(X_{n+1} = 0) = \binom{n}{k} p^k (1-p)^{n-k} \times (1-p)$$

et de même

$$\mathbb{P}(\{S = k\} \cap \{X_{n+1} = 1\}) = \mathbb{P}(T = k-1)\mathbb{P}(X_{n+1} = 1) = \binom{n}{k-1} p^{k-1} (1-p)^{n-k-1} \times p$$

et donc

$$\mathbb{P}(S = k) = \left(\binom{n}{k} + \binom{n}{k-1} \right) p^k (1-p)^{n+1-k}$$

récurrence établie. □

6 Espérance

6.1 Définition

Définition 6.1 (Espérance d'une variable aléatoire) :
On appelle espérance d'une variable aléatoire réelle X sur $(\Omega, \mathbb{P})$ un espace probabilisé fini, le réel

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

Cette quantité ne dépend que de la loi de X.

Remarque :
Avec $\mathbb{P}(X = x) = 0$ si $x \notin X(\Omega)$, on peut également écrire $\mathbb{E}(X) = \sum_{x \in \mathbb{R}} x\mathbb{P}(X = x)$.

Remarque :
L'espérance ne dépend que de la loi de X. Donc deux variables aléatoires différentes de même lois auront la même espérance. Et en particulier, on ne peut pas retrouver la loi de X (ni X) à partir de la seule information de l'espérance.

L'espérance est la moyenne des valeur prise par X. Autrement dit, en considérant X comme le gain obtenu en jouant à un jeu de hasard, l'espérance est le gain moyen que l'on peut espérer recevoir. ATTENTION ! Ça ne veut pas dire que c'est le gain qu'on aura en jouant. C'est le gain moyen. Il y aura des parties où on perdra, d'autres où on gagnera. C'est la différence entre probabilités et statistiques.

Théorème 6.1 (Calcul de l'espérance) :
Soit X un variable aléatoire réelle sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. Alors

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\})$$

Démonstration :
On a $\{X = x\} = \bigcup_{\omega \in \{X=x\}} \{\omega\}$. Donc

$$\mathbb{P}(X = x) = \sum_{\omega \in \{X=x\}} \mathbb{P}(\{\omega\})$$

puisque cette réunion est disjointe. Puis

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \sum_{\omega \in \{X=x\}} \mathbb{P}(\{\omega\}) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} X(\omega)\mathbb{P}(\{\omega\})$$

Les événements $\{X = x\}$ avec x parcourant $X(\Omega)$ constituent une partition de Ω. Autrement dit $\Omega = \bigcup_{x \in X(\Omega)} \bigcup_{\omega \in \{X=x\}} \{\omega\}$. Donc

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\})$$

□

Remarque :
L'espérance peut donc être vu comme une moyenne pondérée. C'est la moyenne des valeurs prises par la variables pondéré par la probabilité que la variable X prennent ces valeurs.

Définition 6.2 (Variable aléatoire centrée) :
Une variable aléatoire réelle X est dite centrée si son espérance est nulle.

Autrement dit, une variable aléatoire correspond à un jeu équilibré, pour lequel, on gagne autant que l'on perd *en moyenne*.

6.2 Propriétés de l'espérance

Théorème 6.2 (Linéarité) :
Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Pour toutes variables aléatoires réelles X et Y sur Ω et pour tout $\lambda, \mu \in \mathbb{R}$,

$$\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$$

Autrement dit, l'espérance est une application linéaire sur l'ensemble des variables aléatoires réelles sur Ω.

Démonstration :
On a

$$\mathbb{E}(\lambda X + \mu Y) = \sum_{\omega \in \Omega} (\lambda X + \mu Y)(\omega) \mathbb{P}(\{\omega\}) = \lambda \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) + \mu \sum_{\omega \in \Omega} Y(\omega) \mathbb{P}(\{\omega\})$$

et donc le résultat. □

En particulier,

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$$

Proposition 6.3 ("Centrage" d'une variable aléatoire) :
Si X est une variable aléatoire réelle sur Ω un univers fini, alors $Y = X - \mathbb{E}(X)$ est une variable aléatoire centrée.

Démonstration :
Par linéarité,

$$\mathbb{E}(Y) = \mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(X) = 0$$

□

Remarque :
Par un schéma de Bernoulli, on peut retrouver la l'espérance d'une loi binomiale.

Théorème 6.4 (Positivité) :
Soit Ω un univers fini. Si X est une variable aléatoire sur Ω à valeurs dans $\mathbb{R}_+$, alors

$$\mathbb{E}(X) \geq 0$$

Démonstration :

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\})$$

Mais $\forall \omega \in \Omega$, $X(\omega) \geq 0$ et $\mathbb{P}$ étant une probabilité sur Ω, elle est à valeur positive. □

Corollaire 6.5 (Croissance de l'espérance) :
Soit Ω un univers fini. Soit X et Y deux variables aléatoires réelles sur Ω telles que $X \leq Y$, alors

$$\mathbb{E}(X) \leq \mathbb{E}(Y)$$

Démonstration :
Comme pour l'intégration, on considère $Z = Y - X$. Elle est positive. Et la linéarité de l'espérance finit le travail. □

Proposition 6.6 (Inégalité triangulaire) :
Soit Ω un univers fini et X une variable aléatoire réelle sur Ω. Alors

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|).$$

Démonstration :
On a $-|X| \leq X \leq |X|$. Donc, par croissance de l'espérance et par linéarité, $-\mathbb{E}(|X|) \leq \mathbb{E}(X) \leq \mathbb{E}(|X|)$ et donc $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$. □

6.3 Espérances usuelles

Proposition 6.7 (Espérance d'une variable aléatoire constante) :
Si X est une variable aléatoire réelle constante C, alors

$$\mathbb{E}(X) = C$$

Démonstration :
On a $\mathbb{E}(X) = C\mathbb{P}(X = C) = C$. □

Proposition 6.8 (Espérance d'une fonction indicatrice) :
Soit Ω un univers fini et A un événement de Ω. Alors $\mathbb{1}_A$ est une variable aléatoire réelle sur Ω et
$$\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A).$$

Démonstration :
$\mathbb{E}(\mathbb{1}_A) = \sum_{x \in \mathbb{1}_A(\Omega)} x\mathbb{P}(\mathbb{1}_A = x) = 1 \times \mathbb{P}(\mathbb{1}_A^{-1}(\{1\})) = \mathbb{P}(A)$. □

Proposition 6.9 (Espérance d'une loi uniforme) :
Soit X une variable aléatoire réelle suivant une loi uniforme sur $[\![p, q]\!]$. Alors
$$\mathbb{E}(X) = \frac{p+q}{2}$$

Démonstration :
Par sommation arithmétique, $\mathbb{E}(X) = \sum_{k=p}^{q} k\frac{1}{q-p+1} = \frac{p+q}{2}$ □

Proposition 6.10 (Espérance d'une loi de Bernoulli) :
Si $X \sim \mathcal{B}(p)$, alors
$$\mathbb{E}(X) = p$$

Démonstration :
$\mathbb{E}(X) = 0(1-p) + 1 \times p = p$ □

Proposition 6.11 (Espérance d'une loi binomiale) :
Si $X \sim \mathcal{B}(n, p)$, alors
$$\mathbb{E}(X) = np$$

Démonstration :
On a $\mathbb{E}(X) = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}$. Mais il est bien connu que $k\binom{n}{k} = n\binom{n-1}{k-1}$. Donc
$$\mathbb{E}(X) = \sum_{k=1}^{n} n\binom{n-1}{k-1}p^k(1-p)^{n-k} = np(p+(1-p))^{n-1} = np$$
□

6.4 Formule de transfert

Théorème 6.12 (Formule de transfert [✓]) :
Soit Ω un univers fini. Si X est une variable aléatoire réelle sur Ω et $f : X(\Omega) \to \mathbb{R}$, alors
$$\mathbb{E}(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbb{P}(X = x)$$
L'espérance de $f(X)$ ne dépend donc que de la loi de X.

Démonstration :
Par définition,
$$\mathbb{E}(f(X)) = \sum_{\omega \in \Omega} f(X(\omega))\mathbb{P}(\{\omega\})$$
Comme on est sur un univers fini, $X(\Omega)$ est fini. On note $X(\Omega) = \{x_1, \dots, x_n\}$. Les événements $\{X = x_1\}$, ..., $\{X = x_n\}$ constituent un système complet d'événements. Et par conséquent,
$$\mathbb{E}(f(X)) = \sum_{k=1}^{n} \sum_{\omega \in \{X = x_k\}} f(X(\omega))\mathbb{P}(\{\omega\}) = \sum_{k=1}^{n} f(x_k) \sum_{\omega \in \{X = x_k\}} \mathbb{P}(\{\omega\}) = \sum_{k=1}^{n} f(x_k)\mathbb{P}(X = x_k)$$
□

Remarque :
Si $f = \mathrm{Id}_{\mathbb{R}}$, on retrouve la définition de l'espérance de X (ce qui est heureux) :
$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x)$$

Remarque :
On notera qu'il n'est pas nécessaire de connaître la loi de $f(X)$ pour connaître l'espérance de $f(X)$.

Exemple 6.1 :
$\mathbb{E}(X^2) = \sum_{x\in X(\Omega)} x^2\mathbb{P}(X = x)$. On a aussi $\mathbb{E}(e^X) = \sum_{x\in X(\Omega)} e^x\mathbb{P}(X = x)$. Et on peut en faire plein.

6.5 Espérance d'un produit de variables aléatoires

Théorème 6.13 (Espérance d'un produit) :
Soit Ω un univers fini. Soit X et Y deux variables aléatoires réelles sur Ω indépendantes. Alors

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Démonstration :
Par définition,

$$\mathbb{E}(XY) = \sum_{\omega\in\Omega} X(\omega)Y(\omega)\mathbb{P}(\{\omega\}) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy\mathbb{P}(X = x, Y = y)$$

Par indépendance,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

et donc

$$\mathbb{E}(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy\mathbb{P}(X = x)\mathbb{P}(Y = y)$$

En remettant tout ça dans l'ordre, ce qui est possible puisque les variables sont séparées,

$$\mathbb{E}(XY) = \left(\sum_{x\in X(\Omega)} x\mathbb{P}(X = x)\right)\left(\sum_{y\in Y(\Omega)} y\mathbb{P}(Y = y)\right) = \mathbb{E}(X)\mathbb{E}(Y)$$

□

!!! ATTENTION !!!

La réciproque est fausse en général ! On peut avoir $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ même avec des variables dépendantes.

Contre-exemple :
Soit X_1 et X_2 telles que $X_1(\Omega) = \{-1, 0, 1\}$, $\mathbb{P}(X_1 = 0) = 1/2$, $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/4$ et $X_2 = \mathbb{1}_{\{X_1=0\}}$. Alors $X_1X_2 = 0$. Comme $\mathbb{E}(X_1) = 0$, on a bien $\mathbb{E}(X_1X_2) = 0 = \mathbb{E}(X_1)\mathbb{E}(X_2)$.
Mais $\mathbb{P}(X_1 = 1, X_2 = 1) = 0 \neq 1/8 = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 1)$. Donc les deux variables ne sont pas indépendantes.

Contre-exemple :
Soit X et Y deux variables aléatoires déterminées par leur loi conjointe suivante :

$X\backslash Y$	-1	0	1
-1	$1/6$	$1/6$	$1/6$
1	$1/4$	0	$1/4$

Alors X et Y ne sont pas indépendantes mais $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Corollaire 6.14 :
Si $X_1, \dots, X_n$ sont des variables aléatoires réelles mutuellement indépendantes,

$$\mathbb{E}(X_1 \times \cdots \times x_n) = \mathbb{E}(X_1) \times \cdots \times \mathbb{E}(X_n)$$

Démonstration :
On fait simplement une récurrence sur le nombre de variables aléatoires. $\square$

Définition 6.3 (Moments d'une variable aléatoire réelle) :
Soit X une variable aléatoire réelle sur Ω, un univers fini. Soit $p \geq 1$.
Le *moment d'ordre p de X* est l'espérance de X^p, *i.e.* c'est

$$\mathbb{E}(X^p) = \sum_{x \in X(\Omega)} x^p \mathbb{P}(X = x).$$

Remarque :
En particulier, le moment d'ordre 1 est l'espérance de la variable aléatoire.

7 Variance, Écart type, Covariance

7.1 Variance et Écart-type

Définition 7.1 (Variance, Écart-type) :
On appelle variance d'une variable aléatoire réelle sur un espace probabilisé fini $(\Omega, \mathbb{P})$, le réel

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X = x)$$

On appelle écart-type de de la variable X, le réel

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

Remarque :
Variance et écart-type ne dépendent que de la loi de X.

Remarque :
La variance correspond donc au moment d'ordre 2 de la variable centrée associée à la variable aléatoire X.

Remarque :
Variance et écart-type sont des indicateurs de dispersion. Ces deux quantités permettent de mesurer la dispersion des valeurs prises par la variable X par rapport à sa valeur moyenne (ici l'espérance). Précisément, l'écart-type est la distance euclidienne dans $\mathbb{R}^n$ pour le produit scalaire canonique entre le vecteurs dont les coordonnées sont les valeurs prises par X et le vecteur dont toutes les coordonnées sont $\mathbb{E}(X)$.

Plus l'écart-type (et donc la variance) sont petits et plus X est concentrée autour de son espérance. Éventuellement, pour une variable aléatoire constante, on aura (voir plus bas) $\mathbb{V}(X) = 0$. Et inversement, si $\mathbb{V}(X) = 0$, alors X est presque sûrement constante.

Physiquement parlant, l'espérance et écart-type s'exprime dans la même unité que X si elle en a une. Ces quantités sont même homogènes à X.

Théorème 7.1 (Expression de la variance [✓]) :
Si X est une variable aléatoire réelle sur $(\Omega, \mathbb{P})$ un espace probabilisé fini, alors

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Démonstration :
En développant,

$$(X - \mathbb{E}(X))^2 = X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2$$

et par linéarité de l'espérance,

$$\mathbb{V}(X) = \mathbb{E}(X^2) - 2\mathbb{E}(X)E(X) + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

□

Remarque :
Ce théorème s'appelle le théorème de Kœnig-Huygens.

Exemple 7.1 :
Calculer $\mathbb{V}(X)$ si $X \sim \mathcal{U}(n)$.

Théorème 7.2 (Propriété algébrique de la variance) :
Soit Ω un univers fini. Si X est une variable aléatoire réelle sur Ω, alors $\forall a, b \in \mathbb{R}$,

$$\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$$

et donc $\sigma(aX + b) = |a|\sigma(X)$.

Démonstration :

On calcul :

$$\begin{aligned}\mathbb{V}(aX+b) &= \mathbb{E}((aX+b)^2) - \mathbb{E}(aX+b)^2 \\ &= \mathbb{E}(a^2X^2 + 2abX + b^2) - (a\mathbb{E}(X)+b)^2 \\ &= a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - a^2\mathbb{E}(X)^2 - 2ab\mathbb{E}(X) - b^2 \\ &= a^2(\mathbb{E}(X^2) - \mathbb{E}(X)^2) \\ &= a^2\mathbb{V}(X)\end{aligned}$$

□

Remarque :
C'est normal qu'ajouter b à X ne change pas la variance. C'est une translation des toutes la valeurs. Donc ça change la valeur moyenne, mais ça ne change pas la dispersion des valeurs autour de la moyenne. Et la variance mesure cette dispersion. Donc c'est normal.

Définition 7.2 (Variable réduite) :
Si une variable aléatoire réelle X a une variance de 1, on dit que X est réduite.

Exemple 7.2 :
Si X est une variable aléatoire réelle sur $(\Omega, \mathbb{P})$ d'espérance m et d'écart-type $\sigma > 0$, alors

$$Y = \frac{X-m}{\sigma}$$

est une variable centrée réduite.

7.2 Covariance, Variance d'une somme de variables aléatoires

Définition 7.3 (Covariance, Variables aléatoires non corrélées) :
Soit (X, Y) deux variables aléatoires réelles sur un univers Ω fini.

On appelle *covariance* du couple (X, Y) le nombre

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

Si $\mathrm{Cov}(X,Y)=0$, on dit que X et Y sont *non corrélées*.

Remarque :
La covariance mesure l'écartement simultané dans le même sens en moyenne des deux variables aléatoires. Autrement dit, $\mathrm{Cov}(X,Y)\geq 0$ si X et Y s'écartent de leur espérance dans le même sens (toutes les deux dessous ou toutes les deux au-dessus) et $\mathrm{Cov}(X,Y)\leq 0$ si elles s'écartent de leur espérance respective dans des sens contraire (l'une au-dessus et l'autre dessous).

Proposition 7.3 (Propriété de la covariance) :
Soit Ω un univers fini et X et Y deux variables aléatoires réelles sur Ω.
Alors

(i) Cov est une forme bilinéaire symétrique positive sur l'ensemble des variables aléatoires réelles sur Ω.
(ii) $\mathrm{Cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$
(iii) $\mathbb{V}(X+Y)=\mathbb{V}(X)+2\,\mathrm{Cov}(X,Y)+\mathbb{V}(Y)$
(iv) Si X et Y sont indépendantes, alors $\mathrm{Cov}(X,Y)=0$. La réciproque est fausse.

Démonstration :

(i) C'est une conséquence de la linéarité de l'espérance, de la positivité de l'espérance et de la commutativité du produit dans $\mathbb{R}$.
(ii) $\mathrm{Cov}(X,Y)=\mathbb{E}(EX-\mathbb{E}(X)Y-\mathbb{E}(Y)X+\mathbb{E}(X)\mathbb{E}(Y))=\mathbb{E}(XY)-2\mathbb{E}(X)\mathbb{E}(Y)+\mathbb{E}(X)\mathbb{E}(Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$ par linéarité de l'espérance.
(iii) On a

$$\mathbb{V}(X+Y)=\mathbb{E}((X-\mathbb{E}(X)+Y-\mathbb{E}(Y))^2)=\mathbb{E}((X-\mathbb{E}(X))^2)+2\mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))+\mathbb{E}((Y-\mathbb{E}(Y))^2)$$

(iv) Évident avec le point (ii). Et des contre-exemples ont déjà été donnés.

□

Remarque :
La covariance est donc presque un produit un scalaire. On peut donc presque appliqué l'inégalité de Cauchy-Schwarz :

Proposition 7.4 (Inégalité de Cauchy-Schwarz) :
Soit X et Y deux variables aléatoires réelles sur Ω un univers fini.
Alors

$$\operatorname{Cov}(X,Y)^2 \leq \mathbb{V}(X)\mathbb{V}(Y)$$

Autrement dit, $|\operatorname{Cov}(X,Y)| \leq \sigma(X)\sigma(Y)$.

Démonstration :
On adapte la démonstration de l'inégalité de Cauchy-Schwartz : On a $\forall t \in \mathbb{R}$, $0 \leq \mathbb{V}(X+tY) = \mathbb{V}(X) + 2t\operatorname{Cov}(X,Y) + t^2\mathbb{V}(Y)$. Donc $\operatorname{Cov}(X,Y)^2 - \mathbb{V}(X)\mathbb{V}(Y) \leq 0$. Et donc le résultat. □

Proposition 7.5 (Variance d'une somme de variables aléatoires réelles) :
Soit Ω un univers fini et $X_1,\ldots,X_n$ des variables aléatoires réelles sur Ω.
Alors

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

De plus, si $(X_1,\ldots,X_n)$ sont deux à deux indépendantes, alors

$$\mathbb{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}(X_i)$$

Remarque :
De plus, si elles sont VAIID, alors $\mathbb{V}\left(\sum_{i=1}^n X_i\right) = n\mathbb{V}(X_1)$.

Démonstration :
Les cas $n=1$ et $n=2$ sont triviaux. Supposons que ce soit vrai pour $n-1$ variables aléatoires réelles. Soit $X_1,\ldots,X_n$ des variables aléatoires réelles, alors

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

□

7.3 Inégalité de Markov et Bienaymé-Tchebychev

Théorème 7.6 (Inégalité de Markov [✓]) :
Soit X une variable aléatoire à valeur dans $\mathbb{R}_+$ sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. Alors,

$$\forall a \geq 0,\ a\mathbb{P}(X \geq a) \leq \mathbb{E}(X)$$

Démonstration :
On a

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\})$$

et comme $X(\omega)$ est positif,

$$\mathbb{E}(X) \geq \sum_{\substack{\omega \in \Omega \\ X(\omega) \geq a}} X(\omega)\mathbb{P}(\{\omega\}) \geq \sum_{\substack{\omega \in \Omega \\ X(\omega) \geq a}} a\mathbb{P}(\{\omega\}) = a\mathbb{P}(X \geq a)$$

$\square$

Théorème 7.7 (Inégalité de Bienaymé-Tchebychev [✓]) :
Soit X une variable aléatoire réelle sur $(\Omega, \mathbb{P})$, un espace probabilisé fini. Alors,

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

Démonstration :
On a $\{|X - \mathbb{E}(X)| \geq \varepsilon\} = \{(X - \mathbb{E}(X))^2 \geq \varepsilon^2\}$ et par l'inégalité de Markov appliquée à la variable positive $Y = (X - \mathbb{E}(X))^2$:

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

$\square$

Remarque :
En passant aux événements contraire, on a égalament,

$$\forall a > 0,\ \mathbb{P}(|X - \mathbb{E}(X)| < a) \geq 1 - \frac{\mathbb{V}(X)}{a}.$$

Exemple 7.3 :
On veut estimer l'équilibre d'une pièce. On note p la probabilité inconnue que la pièce donne Face lors d'un lancer.

On lance n fois la pièce et on pose X la variable aléatoire correspondant au nombre de lancé donnant Face. On sait que $X \sim \mathcal{B}(n, p)$. Sachant que $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = np(1-p) \leq n/4$, l'inégalité de Bienaymé-Tchebychev, donne

$$\mathbb{P}(|X/n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$$

pour tout $\varepsilon > 0$.

En particulier, pour $\varepsilon = 0,01$, on obtient X/n est une valeur approchée de p à ε près avec une probabilité supérieure à 5% sous réserve de prendre $n \geq 50000$.

7.4 Variances usuelles

Proposition 7.8 (Variance d'une variable aléatoire constante) :
Soit X une variable aléatoire constante sur $(\Omega, \mathbb{P})$ un espace probabilisé fini. Alors

$$\mathbb{V}(X) = 0$$

Démonstration :
Si $X = C$,

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = C^2 - C^2 = 0$$

□

Proposition 7.9 (Variance d'une loi de Bernoulli) :
Si $X \sim \mathcal{B}(p)$, alors

$$\mathbb{V}(X) = p(1-p)$$

Démonstration :
$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1-p)$ □

Proposition 7.10 (Variance d'une loi binomiale) :
Si $X \sim \mathcal{B}(n,p)$, alors
$$\mathbb{V}(X) = np(1-p)$$

Démonstration :
On a $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2$. Or
$$\mathbb{E}(X(X-1)) = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k (1-p)^{n-k}$$
et avec tous ce qu'on sait que les coefficient binomiaux,
$$k(k-1)\binom{n}{k} = n(n-1)\binom{n-2}{k-2}$$
donc
$$\mathbb{E}(X(X-1)) = n(n-1)\sum_{k=2}^{n}\binom{n-2}{k-2}p^k(1-p)^k = n(n-1)p^2(p+(1-p))^{n-2}$$
et donc
$$\mathbb{V}(X) = n(n-1)p^2 + np - (np)^2 = np(1-p)$$

□