N. Laillet
nlaillet.math@gmail.com

Chapitre 20
Probabilités

Table des matières

N. Laillet
nlaillet.math@gmail.com

Le but de ce chapitre est de prolonger les choses vues en Terminale, mais en mettant l'accent sur une approche, qui correspond à l'approche moderne des probabilités : l'univers n'est qu'un ensemble permettant de faire vivre des **variables aléatoires**, lesquelles servent vraiment à modéliser un phénomène aléatoire.
Modéliser, ça n'est pas choisir un univers, mais une variable aléatoire !

1 Notions de probabilités

1.1 Définitions

Définition 1

Un espace probabilisable fini est un couple $(\Omega, \mathcal{P}(\Omega))$ où Ω est un ensemble fini appelé univers. Les éléments de Ω sont appelés *réalisations*, *résultats* ou événements élémentaires, les éléments de $\mathcal{P}(\Omega)$ (i.e. les parties de Ω) sont appelés événements ou évènements.

Afin de ne pas confondre mathématiques et modélisation, on parlera de modélisation après avoir un peu tout défini.

Définition 2

Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable, A et B deux événements.

1. L'événement A et B est défini par l'intersection $A \cap B$.
2. L'événement A ou B est défini par la réunion $A \cup B$.
3. L'événement contraire de A, $\overline{A}$ est défini par le complémentaire de A dans Ω.
4. On dit que deux événements A et B sont incompatibles s'ils sont disjoints, i.e. si $A \cap B = \emptyset$. On note alors $A \cup B = A \sqcup B$.
5. On appelle système complet d'événements (parfois abrégé en sce) toute collection d'événements $(A_1, \ldots, A_n)$, deux à deux disjoints, tels que $A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n = \Omega$.

Proposition 3 (Décomposition sur un système complet d'événements)

Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable, $(A_1, \ldots, A_r)$ un système complet d'événements de Ω. Pour tout B dans Ω,

$$B = \bigsqcup_{k=1}^{r} (A_k \cap B)$$

Exemple 4

1. Imaginons que l'on lance un dé. On **peut** choisir $\Omega = \{1, 2, 3, 4, 5, 6\}$. (j'insiste sur le « on peut » : rien ne nous impose mathématiquement, de choisir cet univers)
   Si l'on considère les événements

N. Laillet
nlaillet.math@gmail.com

- $A =$« le chiffre obtenu est pair » ,
- $B =$« le chiffre obtenu est impair » ,
- $C =$« le chiffre obtenu est premier » ,
- $D = \{4, 6\}$.

Le seul qui est réellement décrit comme un événement est D. Pour les autres, il faudrait dire $A = \{2, 4, 6\}$, $B = \{1, 3, 5\}$ et $C = \{2, 3, 5\}$.
C et D sont incompatibles.
(A, B) forme un système complet d'événements.
L'événement B et C est l'événement $\{3, 5\}$.

2. Quel que soit l'univers Ω, $(\{x\})_{x\in\Omega}$ forme un système complet d'événements de Ω.

Cet exemple montre que les choses ne sont pas tout à fait naturelles si on s'intéresse trop à l'univers. Ce qui nous intéresse, c'est le **résultat** de l'expérience aléatoire. Imaginons que nous lancions un dé, puis tirions une carte. On ne veut pas forcément mettre les mains dans le cambouis pour changer l'univers à nouveau !

Définition 5

Soit Ω un ensemble. Une variable aléatoire X sur Ω à valeurs dans E est une application de Ω dans E.
Si $E \subset \mathbb{R}$, on dit que la variable aléatoire est réelle (on note v.a.r.)

Exemple 6

1. **Remarque.** Comme Ω est fini, $X(\Omega)$ est aussi fini. Par contre, rien ne nous dit que E est fini !
2. Des exemples, si X est le résultat d'un lancer de dé, X est une variable aléatoire réelle à valeurs dans $[\![1, 6]\!]$.
3. Si Y est le résultat d'un autre lancer de dés,
   - $X + Y$ est une variable aléatoire réelle,
   - $\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}$ est une variable aléatoire à valeurs dans $\mathscr{M}_2(\mathbb{R})$,
   - $F \left| \begin{array}{l} \mathbb{R} \to \mathbb{R} \\ t \mapsto X.t + Y \end{array} \right.$ est une variable aléatoire à valeurs dans $\mathbb{R}^{\mathbb{R}}$.

Un exemple de variable aléatoire.

Définition 7

Si $A \subset \Omega$. La fonction indicatrice de A est une variable aléatoire sur Ω, on la note $\mathbb{1}_A$.

N. Laillet
nlaillet.math@gmail.com

Définition 8

Soit Ω un ensemble, X une variable aléatoire à valeurs dans E.

1. Pour toute partie A de E, l'événement $\{X \in A\}$ ou $X \in A$ est défini par
$$\{X \in A\} = X^{-1}(A) = \{\omega \in \Omega,\ X(\omega) \in A\}.$$
2. Si $A = \{a\}$, l'événement $X \in A$ est noté $X = a$.
3. On définit de même, pour une variable aléatoire réelle, et $a \in \mathbb{R}$,
   - $\{X \geqslant a\} = X^{-1}([a, +\infty[)$,
   - $\{X \leqslant a\} = X^{-1}(]-\infty, a])$,
   - $\{X > a\} = X^{-1}(]a, +\infty[)$,
   - $\{X < a\} = X^{-1}(]-\infty, a[)$,

1.2 Probabilité

Définition 9

Soit Ω un ensemble fini. Une probabilité sur Ω est une application $\mathbb{P} : \mathcal{P}(\Omega) \to [0, 1]$, telle que

1. $\mathbb{P}(\Omega) = 1$.
2. $\forall A, B \in \Omega,\ A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Le triplet $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ est appelé espace probabilisé. On le notera souvent, plus simplement, $(\Omega, \mathbb{P})$.

Remarque 10

En théorie, un espace probabilisé est la donné d'un triplet $(\Omega, \mathcal{T}, \mathbb{P})$ où $\mathcal{T}$ est ce qu'on appelle une *tribu* sur Ω : ces subtilités sont du ressort des professeurs de deuxième année, qui travaillent sur des ensembles infinis !

Proposition 11

Soit $(\Omega, \mathbb{P})$ un espace probabilisé. Alors

1. (ensemble vide) $\mathbb{P}(\emptyset) = 0$
2. (complémentaire) $\forall A \subset \Omega, \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
3. (différence) $\forall (A, B) \in \mathcal{P}(\Omega)^2, \mathbb{P}(A \backslash B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.
4. (réunion disjointe) $\forall A_1, \dots, A_n$ deux à deux disjoints, $\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k)$. En particulier, pour tout $A \subset \Omega$, $\mathbb{P}(A) = \sum_{x \in A} \mathbb{P}(\{x\})$.
5. (réunion) $\forall (A, B) \in \mathcal{P}(\Omega)^2,\ \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

N. Laillet
nlaillet.math@gmail.com

6. (croissance) $\forall A \subset B \subset \Omega$, $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.
7. (sous-additivité) $\forall (A_1, \dots, A_n) \in \mathcal{P}(\Omega)^n$, $\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \leqslant \sum_{k=1}^{n} \mathbb{P}(A_k)$, avec égalité si $(A_1, \dots, A_n)$ sont deux à deux disjoints.
8. (décomposition sur un système complet d'événements) Si $(A_1 \dots, A_n)$ est un système complet d'événements et $B \in \mathcal{P}(\Omega)$, alors
$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i).$$
En particulier,
$$\mathbb{P}(B) = \sum_{x \in B} \mathbb{P}(\{x\}).$$

PREUVE

Définition 12

Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini.

1. Un événement A est dit presque sûr s'il est de probabilité 1.
2. Un événement A est dit négligeable s'il est de probabilité nulle.
3. Deux événements A et B sont dits équiprobables si $\mathbb{P}(A) = \mathbb{P}(B)$.
4. On dit que $\mathbb{P}$ est la probabilité uniforme si tous les singletons sont équiprobables, i.e. si pour tout x dans Ω, $\mathbb{P}(\{x\}) = \dfrac{1}{\text{Card}(\Omega)}$.

Remarque 13

1. Par la formule $\mathbb{P}(B) = \sum_{x \in B} \mathbb{P}(\{x\})$, on en déduit qu'une probabilité est définie de manière unique par les valeurs $(\mathbb{P}(\{x\}))_{x \in \Omega}$.
2. Ceci assure l'unicité de la probabilité uniforme : si p est la valeur commune à tous les $\mathbb{P}(\{x\})$, alors
$$1 = \mathbb{P}(\Omega) = \sum_{x \in \Omega} \mathbb{P}(\{x\}) = \sum_{x \in \Omega} p = p.\text{Card}(\Omega),$$
donc $p = \dfrac{1}{\text{Card}(\Omega)}$.
3. **Attention** aux idées reçues. Si on prend un dé truqué, avec comme univers $\Omega = \{1, 2, 3, 4, 5, 6\}$, et comme probabilité sur cet univers $\mathbb{P}$ vérifiant
   - $\mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \dfrac{1}{2}$,
   - $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = 0$,

   alors

N. Laillet
nlaillet.math@gmail.com

- $\mathbb{P}(\{1,2,5,6\}) = 0$ mais $\{1,2,5,6\} \neq \emptyset$,
- $\mathbb{P}(\{3,4\}) = 1$ mais $\{3,4\} \neq \Omega$,
- $\mathbb{P}(\{3\}) = \mathbb{P}(\{4\})$ mais $\{3\} \neq \{4\}$.

4. Exemple de décomposition sur un système complet d'événements.
Si on lance 2 fois de suite un dé, et que l'on prend $\Omega = \{(i,j), (i,j) \in [\![1,6]\!]^2\}$.
On suppose $\mathbb{P}$ uniforme sur Ω.
Si $k \in [\![1,6]\!]$, on note A_k est l'événement « le premier lancer donne k » ($A_k = \{(k,j), j \in [\![1,6]\!]\}$), et B_k est l'événement « le deuxième lancer donne k ».
Si on veut calculer $\mathbb{P}(B_j)$, on peut remarquer que $(A_i)_{i\in[\![1,6]\!]}$ forme un système complet d'événements. AInsi

$$\mathbb{P}(B_j) = \sum_{i=1}^{6} \mathbb{P}(B_j \cap A_i) = \sum_{i=1}^{6} \frac{1}{36} = \frac{1}{6},$$

car $B_j \cap A_i = \{(i,j)\}$: c'est un événement élémentaire.

1.3 Modélisation

Le problème de la modélisation en probabilité est un problème un peu hors des mathématiques. *Pour l'instant*, modéliser un phénomène aléatoire, c'est choisir un univers et une loi de probabilité sur cet univers. Le choix de l'univers n'est pas automatiquement donné par la description du système ! On va voir quelques exemples de cette modélisation.

Exemple 14

Si on lance deux dés, on peut prendre

- $\Omega = [\![1,6]\!]^2$,
- $\Omega = [\![2,12]\!]$,
- $\Omega =$l'ensemble des positions et des vitesses initiales des deux dés.

Et, sur chacun des ces univers, c'est à la personne qui modélise de choisir une loi de probabilité.

Exemple 15

On tire deux boules avec remise dans une urne contenant n boules noires et b boules blanches. Deux possibilités de modélisation

	Premier modèle	Deuxième modèle
Description	On dire soit une boule noire, soit une boule blanche, la proportion des premières est $\frac{n}{n+b}$, celle des secondes est $\frac{b}{n+b}$.	On numérote les boules noires de 1 à n, les boules blanches de $n+1$ à $n+b$. Un tirage correspond à un entier entre 1 et $n+b$.
Univers	L'univers est $\{N,B\}^2$	L'univers est $[\![1,n+b]\!]^2$
Réalisations	$\{N,N\}, \{N,B\}, \{B,B\}, \{B,N\}$	$\{(i,j)\}$ pour $(i,j) \in [\![1,n+b]\!]^2$
Probabilités élémentaires	$\mathbb{P}(\{N,N\}) = \frac{n^2}{(n+b)^2}$, $\mathbb{P}(\{N,B\}) = \frac{nb}{(n+b)^2}$, $\mathbb{P}(\{B,B\}) = \frac{b^2}{(n+b)^2}$, $\mathbb{P}(\{B,N\}) = \frac{bn}{(n+b)^2}$,	$\mathbb{P}(\{(i,j)\}) = \frac{1}{(n+b)^2}$
Événement « deux boules blanches ont été tirées »	$\{B,B\}$	$\{(i,j), (i,j) \in [\![n+1,n+b]\!]^2\}$

N. Laillet
nlaillet.math@gmail.com

Heureusement, ces deux modélisations apportent les mêmes probabilités. La deuxième modélisation est plus lourde mais plus descriptive. Les deux se valent sur ce modèle.

Exemple 16

Le jeu de pile ou face. Univers $\Omega = \{P, F\}$. Pièce biaisée p.
Probabilité d'obtenir au moins une fois face en deux lancers ? Deux possibilités " :

1. Soit on prend logiquement $\Omega = \{(P, F), (P, P), (F, P), (F, F)\}$ et l'événement qui nous intéresse est $\{(P, F), (F, P), (F, F)\}$.
2. soit on prend $\Omega = \{(P, F), (P, P), F\}$ (on s'arrête une fois qu'on a eu face !), et sur Ω on peut mettre soit la probabilité uniforme, soit mettre un poids $\frac{1}{2}$ à F.

Il faut garder en tête que mathématiquement, il n'y a aucun problème, il s'agit de deux modélisations différentes.

Il faut garder en tête deux choses :

1. le choix de l'univers n'est pas quelque chose que l'on prendra le temps de détailler, pour deux raisons : soit il est absolument évident, soit on travaille avec des variables aléatoires, qui sont une manière de « lire » l'univers au travers de fonctions définies sur celui-ci.
2. la compétence « bien choisir son univers » n'est pas une compétence mathématique. C'est une compétence utile, de modélisation, de sciences de l'ingénieur, mais pas de mathématiques.

2 Conditionnement

Exemple 17

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, muni d'une probabilité uniforme. A et B deux parties de Ω.
La probabilité de tirer un élément de A est $\mathbb{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$.
Si on sait qu'on est dans A, quelle est la probabilité d'être dans B ? C'est

$$\frac{\text{Card}(A \cap B)}{\text{Card}(A)} = \frac{\frac{\text{Card}(A\cap B)}{\text{Card}(\Omega)}}{\frac{\text{Card}(A)}{\text{Card}(\Omega)}} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

Définition 18

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini et $A \subset \Omega$ de probabilité non nulle. L'application

$$\mathbb{P}_A : \begin{cases} \mathcal{P}(\Omega) \to [0, 1] \\ B \mapsto \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \end{cases}$$

définit une probabilité, appelée probabilité conditionnée à l'événement A.

N. Laillet
nlaillet.math@gmail.com

Si $B \subset \Omega$, on note parfois $\mathbb{P}_A(B)$ sous la forme $\mathbb{P}(B|A)$ et on dit « probabilité de B sachant A »

Démonstration. On vérifie que l'on définit ainsi une probabilité.

- Déjà, si $B \in \mathcal{P}(\Omega)$, on a toujours $A \cap B \subset A$, donc $0 \leqslant \mathbb{P}(A \cap B) \leqslant \mathbb{P}(A)$ donc $0 \leqslant \mathbb{P}_A(B) \leqslant 1$.
- Ensuite, $\mathbb{P}_A(\Omega) = \dfrac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$,
- Enfin, si C et D sont disjoints, alors $C \cap A$ et $D \cap A$ sont disjoints. Donc

$$\begin{aligned}
\mathbb{P}_A(C \cup D) &= \frac{\mathbb{P}(A \cap (C \cup D)}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(A \cap C \cup A \cap D}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(A \cap C) + \mathbb{P}(A \cap D)}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(A)} + \frac{\mathbb{P}(A \cap D)}{\mathbb{P}(A)} \\
&= \mathbb{P}_A(C) + \mathbb{P}_A(D),
\end{aligned}$$

d'où l'additivité

□

Remarque 19

1. **ATTENTION ! ! ! ! ! $B|A$ n'est pas un événement !**
2. $\mathbb{P}_\Omega(A) = A$ et $\mathbb{P}_A(A) = 1$.

Proposition 20

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini et $A \subset \Omega$ de probabilité non nulle, $B \subset \Omega$. Alors

$$\mathbb{P}(A \cap B) = \mathbb{P}_A(B) \times \mathbb{P}(A).$$

Exemple 21

Dans une urne avec cinq boules, deux vertes et trois jaunes, on pioche deux boules sans remise. Quelle est la probabilité d'avoir une verte au second tirage ?
Faisons un raisonnement « lycéen » et voyons les notions utilisées. On fait un arbre. On note V_i l'événement « la i-ème boule tirée est verte » et J_i l'événement « la i-ème boule tirée est jaune » .
On écrit alors, au lycée,

$$\mathbb{P}(V_2) = \frac{2}{5} \times \frac{1}{4} + \frac{3}{5} \times \frac{1}{2}.$$

Qu'y a-t-il derrière ? Deux formules :

N. Laillet
nlaillet.math@gmail.com

- Ce produit $\frac{2}{5} \times \frac{1}{4}$, que l'on généralisera comme « formule des probabilités composées »,
- Cette somme $\frac{2}{5} \times \frac{1}{4} + \frac{3}{5} \times \frac{1}{2}$, que l'on généralisera comme « formule des probabilités totales ».

Proposition 22

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini et $(A_1, \dots, A_n)$ une famille d'événements telle que $\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) > 0$. Alors

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathbb{P}_{\left(\bigcap_{k=1}^{n-1} A_k\right)}(A_n) \times \mathbb{P}_{\left(\bigcap_{k=1}^{n-2} A_k\right)}(A_{n-1}) \times \cdots \times \mathbb{P}_{A_1}(A_2) \times \mathbb{P}(A_1).$$

Démonstration. La preuve se fait par récurrence : on remarque que déjà, si $(A, B) \in \mathbb{P}(\Omega)^2$ avec $\mathbb{P}(A) > 0$,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B),$$

donc

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathbb{P}\left(\bigcap_{k=1}^{n-1} A_k \cap A_n\right) = \mathbb{P}\left(\bigcap_{k=1}^{n-1} A_k\right) \times \mathbb{P}_{\bigcap_{k=1}^{n} A_k} A_n.$$

On conclut par une belle récurrence. □

Exercice 23

Une urne contient t boules turquoise et r roses. On effectue des tirages sans remise. Quelle est la probabilité de tirer pour la première fois une boule rose au k-ième tirage ?

Proposition 24

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini et $(A_1, \dots, A_n)$ un système complet d'événements de probabilités non nulles. Alors pour tout B dans Ω,

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_k) = \sum_{k=1}^{n} \mathbb{P}_{A_k}(B)\mathbb{P}(A_k).$$

Démonstration. La preuve est déjà dans la proposition ! □

N. Laillet
nlaillet.math@gmail.com

Exercice 25

On est toujours dans la situation des boules vertes et jaunes (2 vertes et 3 jaunes).
On tire une boule verte au second tirage. Quelle est la probabilité d'avoir tiré une boule verte au premier tirage ?
On cherche

$$\mathbb{P}_{V_1}(V_2) = \frac{\mathbb{P}(V_1 \cap V_2)}{\mathbb{P}(V_2)} = \frac{\mathbb{P}(V_1 \cap V_2)}{\mathbb{P}(V_1) \times \mathbb{P}_{V_1}(V_2) + \mathbb{P}(J_1)\mathbb{P}_{J_1}(V_2)}.$$

On a utilisé, ici, une manière d'exprimer $\mathbb{P}_{V_1}(V_2)$ à l'aide de $\mathbb{P}_{V_2}(V_1)$: c'est la formule de Bayes.

Proposition 26

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, A et B dans Ω, de probabilités non nulles. Alors

$$\mathbb{P}_B(A) = \frac{\mathbb{P}_A(B)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

Démonstration. La preuve est déja dans la proposition ! □

La formule de Bayes est très efficace pour *remonter le temps*. C'est vraiment l'idée qu'il faut avoir. Nous allons voir deux exemples, un premier « standard » , et un deuxième sous la forme d'un paradoxe (qui n'en est pas un)

Exemple 27 (Remontée dans temps à l'aide de la formule de Bayes)

Une usine fabrique des pièces, avec une proportion de 0,05 de pièces défectueuses. Le contrôle des fabrications est tel que :

- si la pièce est bonne, elle est acceptée avec la probabilité 0,96.
- si la pièce est mauvaise, elle est refusée avec la probabilité 0,98.

On choisit une pièce au hasard et on la contrôle. Quelle est la probabilité qu'il y ait une erreur de contrôle ? qu'une pièce acceptée soit mauvaise ?

N. Laillet
nlaillet.math@gmail.com

Exercice 28 (Paradoxe de Monty Hall)

Vous jouez à un jeu. Trois portes sont fermées : derrière l'une d'elles se trouve une voiture, derrière chacune des deux autres se trouve un porte-clés.
Une candidate choisit l'une des portes. Le présentateur, qui sait où est la voiture, ouvre l'une des deux autres portes, montrant ainsi un porte-clés. Il propose à la candidate de changer de porte. Que doit-elle faire ?
On suppose que la candidate choisit la porte 1, et que le présentateur ouvre la porte 2. On note V_1, V_2, V_3 les événements « la voiture est derrière la porte 1/2/3 ». On note O_2, O_3 les événements « le présentateur ouvre la porte 2/3 ».
Calculer $\mathbb{P}_{O_2}(V_3)$ et conclure.

3 Indépendance

Il y a des événements pour lesquels le conditionnement n'a pas d'influence. Par exemple, si je tire un nombre entre 1 et 8, et que je demande la probabilité d'avoir un nombre pair, elle sera la même si en plus je sais que le nombre est inférieur ou égal à 4. C'est la notion d'indépendance des événements.

Définition 29

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini .

1. Deux événements A et B sont dits indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
   On note parfois $A \perp\!\!\!\perp B$.
2. Des événements (A_k) sont dits indépendants deux à deux si pour tout $i \neq j$, A_i et A_j sont indépendants.

Remarque 30

1. Si $\mathbb{P}(A) \neq 0$, si A et B sont indépendants, $\mathbb{P}_A(B) = \dfrac{\mathbb{P}(A \times B)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(A)} = \mathbb{P}(B)$.
2. Si $\Omega = [\![1, 8]\!]$, et $\mathbb{P}$ est uniforme, si on note $A = \{n \in [\![1, 8]\!],\ n \text{ pair.}\}$ et $B_k = \{n \in [\![1, 8]\!],\ n \leqslant k\}$.
   Les événements A et B_4 sont-ils indépendants ? A et B_3 ?
   - On calcule : $\mathbb{P}(A \cap B_4) = \mathbb{P}(\{1, 2\}) = \dfrac{1}{4}$,
     $\mathbb{P}(A) = \dfrac{1}{2}$ et $\mathbb{P}(B_4) = \dfrac{1}{2}$,
     donc $\mathbb{P}(A) \times \mathbb{P}(B_4) = \dfrac{1}{4} = \mathbb{P}(A \cap B_4)$,
     donc A et B_4 sont indépendants.
   - On calcule : $\mathbb{P}(A \cap B_3) = \mathbb{P}(\{1, 2\}) = \dfrac{1}{4}$,
     $\mathbb{P}(A) = \dfrac{1}{2}$ et $\mathbb{P}(B_3) = \dfrac{3}{8}$,
     donc $\mathbb{P}(A) \times \mathbb{P}(B_3) = \dfrac{3}{16} \neq \mathbb{P}(A \cap B_3)$,
     donc A et B_3 ne sont pas indépendants.

3. **ATTENTION !** Indépendant et incompatibles sont deux termes qui n'ont rien à voir ! Dans l'exemple précédent, A et B_4 sont indépendants mais pas incompatibles.
Question. Quels sont les événements à la fois indépendants et incompatibles ?
Si A et B sont indépendants et disjoints, alors

$$\mathbb{P}(A \cap B) = \begin{cases} 0 \text{ car } A \text{ et } B \text{ sont disjoints.} \\ \mathbb{P}(A) \times \mathbb{P}(B) \text{ car } A \text{ et } B \text{ sont indépendants.} \end{cases}$$

Donc $\mathbb{P}(A) = 0$ ou $\mathbb{P}(B) = 0$, c'est-à-dire que A ou B est négligeable.

4. Quels sont les événements indépendants d'eux-mêmes ?
Si A est indépendant de A, alors $\mathbb{P}(A \cap A) = \mathbb{P}(A) \times \mathbb{P}(A)$, donc $\mathbb{P}(A) = \mathbb{P}(A)^2$, donc $\mathbb{P}(A) = 0$ ou $\mathbb{P}(A) = 1$, c'est-à-dire que A est presque sûr ou négligeable.

5. Soit n un nombre premier, $\mathbb{P}$ la probabilité uniforme sur $\Omega = [\![1, n]\!]$.
Quels sont les événements indépendants dans $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$?
Soient $(A, B) \in \mathcal{P}(\Omega)$, indépendants. Notons $a = \text{Card}(A)$, $b = \text{Card}(B)$ et $c = \text{Card}(A \cap B)$. Alors, comme $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$, on en déduit que

$$\frac{c}{n} = \frac{a}{n} \times \frac{b}{n}, \text{ c'est-à-dire que } nc = ab.$$

Or, n est premier, donc n divise a ou n divise b.
Or, $[\![a, b]\!]$ est dans $[\![0, n]\!]$, donc $a = n$ ou $b = n$, ou $a = 0$, ou $b = 0$.

Proposition 31

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, A et B dans $\mathcal{P}(\Omega)$, $\mathbb{P}(A) > 0$. Alors on a les équivalences

$$A \text{ et } B \text{ sont indépendants} \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B) \Leftrightarrow \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \mathbb{P}(B) \Leftrightarrow \mathbb{P}_A(B) = \mathbb{P}(B).$$

Proposition 32

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, A et B dans $\mathcal{P}(\Omega)$. Les ASSE

1. A et B sont indépendants.
2. $\overline{A}$ et B sont indépendants.
3. A et $\overline{B}$ sont indépendants.
4. $\overline{A}$ et $\overline{B}$ sont indépendants.

Démonstration. On fait suelement la première implication : on suppose A et B indépendants. Mais alors

$$A = (A \cap B) \sqcup (A \cap \overline{B}).$$

N. Laillet
nlaillet.math@gmail.com

Donc $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B})$, donc

$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(\overline{B}).$$

□

Il y a plus fort que l'indépendance, c'est l'indépendance mutuelle.

Définition 33

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini. Des événements $(A_k)_{1 \leqslant k \leqslant n}$ sont dits mutuellement indépendants si pour tout $I \subset [\![1, n]\!]$,

$$\mathbb{P}\left(\bigcap_{k \in I} A_k\right) = \prod_{k \in I} \mathbb{P}(A_k).$$

Exemple 34

On lance deux dés, on prend les événements A, B et C comme

1. A : « le premier dé amène un nombre pair »
2. B : « le deuxième dé amène un nombre pair »
3. C : « la somme des nombres obtenus est paire »

On montre que des événements sont deux à deux indépendants, mais pas mutuellement indépendants.
On remarque que

$$\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2},$$

et que

$$\mathbb{P}(C) = \mathbb{P}((A \cap B) \sqcup (\overline{A} \cap \overline{B})) = \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(\overline{A})\mathbb{P}(\overline{B}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Puis, A et B sont indépendants par hypothèse, et

$$A \cap C = A \cap B = B \cap C,$$

donc

$$\mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = \mathbb{P}(A \cap B) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(C) = \mathbb{P}(B)\mathbb{P}(C) = \mathbb{P}(A)\mathbb{P}(B).$$

Donc les événements sont deux à deux indépendants.
Mais

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \cap B) = \frac{1}{4} \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C),$$

donc on n'a pas d'indépendance mutuelle.
Moralité. L'indépendance 2 à 2 n'implique pas l'indépendance mutuelle. La réciproque est, en revanche, vraie.

N. Laillet
nlaillet.math@gmail.com

4 Variables aléatoires

4.1 Définitions

On définit de cette manière la loi d'une variable.

Proposition 35 (et def)

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, X une v.a. sur Ω à valeurs dans E. L'application

$$\mathbb{P}_X : \begin{array}{l} \mathcal{P}(E) \to [0,1] \\ \quad A \mapsto \mathbb{P}(X \in A) \end{array}$$

définit une probabilité sur E, appelée loi de la variable X.

Démonstration. Démontrons que $\mathbb{P}_X$ est bien une probabilité.

- $\mathbb{P}_X$ est à valeurs dans $[0,1]$ car $\mathbb{P}$ est à valeurs dans $[0,1]$
- X est à valeurs dans E donc, pour tout ω dans Ω, $X(\omega) \in E$, donc $X^{-1}(E) = \Omega$.
  Donc $\mathbb{P}_X(E) = \mathbb{P}(X \in E) = \mathbb{P}(\Omega) = 1$.
- Soient A et B dans $\mathbb{P}(E)^2$ tels que $A \cap B = \emptyset$.
  Alors $X^{-1}(A) \cap X^{-1}(B) = \emptyset$: en effet, si on dispose de ω dans $X^{-1}(A) \cap X^{-1}(B)$, alors $X(\omega) \in A$ et $X(\omega) \in B$, ce qui contredit le fait que $A \cap B = \emptyset$.
  Alors

$$\begin{aligned} \mathbb{P}_X(A \cup B) &= \mathbb{P}(X \in A \cup B) \\ &= \mathbb{P}(\{\omega \in \Omega,\ X(\omega) \in A \text{ ou } X(\omega) \in B\}) \\ &= \mathbb{P}(\{\omega \in \Omega,\ X(\omega) \in A\} \cup \{\omega \in \Omega,\ X(\omega) \in B\}) \\ &= \mathbb{P}(X^{-1}(A) \cup X^{-1}(B)) \\ &= \mathbb{P}(X^{-1}(A)) + \mathbb{P}(X^{-1}(B)) \text{ car } X^{-1}(A) \cap X^{-1}(B) = \emptyset \\ &= \mathbb{P}_X(A) + \mathbb{P}_X(B). \end{aligned}$$

Donc $\mathbb{P}_X$ est une probabilité ! □

Remarque 36

C'est souvent **beaucoup plus parlant** d'écrire $\mathbb{P}(X \in A)$ plutôt que $\mathbb{P}_X(A)$.

Point de méthode 37

En pratique, quand on vous demande de déterminer la loi de X, vous devez faire deux choses : donner les valeurs prises par X et donner la loi de chaque événement élémentaire.

N. Laillet
nlaillet.math@gmail.com

Exemple 38

On lance deux dés équilibrés et on regarde le résultat de la somme des deux, noté X. Quelle est la loi de X ?
X est à valeurs dans $[\![2, 12]\!]$.

a	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = a)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Remarque 39

ATTENTION ! Deux variables aléatoires de même loi ne sont pas nécessairement égales !
Exemple : on tire une fois à pile ou face avec une pièce équilibrée. On note X le nombre de « pile » , Y le nombre de « face » . Alors X et Y sont à valeurs dans $\{0, 1\}$,

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2}, \ \mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = \frac{1}{2},$$

mais $X = 1 - Y$, c'est-à-dire que X et Y sont toujours différentes !

Proposition 40

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un epf.

1. si E et F sont deux ensembles, si X est une variable aléatoire à valeurs dans E, si $f : E \to F$ est une application, alors $f(X)$ est une variable aléatoire à valeurs dans F.
2. si $(E_1, \dots, E_n)$ sont n ensembles, si F est un ensemble, si $g : E_1 \times \cdots \times E_n \to F$ est une application, alors

$$g(X_1, \dots, X_n) \text{ est une variable aléatoire à valeurs dans } F.$$

Exemple 41

La somme, le produit de deux variables aléatoires est une variable aléatoire.

Remarque 42

Abus de notation. Lorsqu'on écrit $f(X)$, on sous-entend $f \circ X : \left| \begin{array}{l} \Omega \to F \\ \omega \mapsto f(X(\omega)) \end{array} \right.$. Cet abus de notation est réservé aux probabilités.
Au contraire, en algèbre linéaire, si $(u, v) \in \mathscr{L}(E)$, on n'écrira jamais $u(v)$ à la place de $u \circ v$!

N. Laillet
nlaillet.math@gmail.com

4.2 Lois usuelles

Définition 43

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un epf. Une variable à valeurs dans un ensemble E fini suit la loi uniforme si pour tout x dans E, $\mathbb{P}(X = x) = \dfrac{1}{\text{Card}(E)}$. On note $X \sim \mathscr{U}(E)$.

Exemple 44

La valeur d'un dé non truqué cubique suit une loi $\mathscr{U}(\llbracket 1, 6 \rrbracket)$.

Définition 45

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un epf. Une variable aléatoire X sur Ω est appelée **variable de Bernoulli** si elle est à valeurs dans $\{0, 1\}$.
Si c'est le cas, on note $p = \mathbb{P}(X = 1)$, $\mathbb{P}(X = 0) = 1 - p$. On note alors $X \sim \mathscr{B}(p)$.

Exemple 46

1. Un tirage à pile ou face : si on lance une fois une pièce et que X correspond au nombre de « pile » , alors X suit une loi de Bernoulli !
2. Si on s'intéresse à une caractéristique dans une population, que l'on tire un individu au hasard, et que X est la variable valant 1 si cet individu a cette caractéristique et 0 sinon, alors X suit une variable de Bernoulli.
3. Un produit de variables de Bernoulli est une variable de Bernoulli (en effet, un produit de fonctions à valeurs dans $\{0, 1\}$ est une fonction à valeurs dans $\{0, 1\}$.

Définition 47

Soit n dans $\mathbb{N}^*$ et $p \in [0, 1]$. Une variable X à valeurs dans $\llbracket 0, n \rrbracket$ suit la loi binomiale de paramètres n et p si pour tout $k \in \llbracket 0, n \rrbracket$,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

On note $X \sim \mathscr{B}(n, p)$.

Remarque 48

N. Laillet
nlaillet.math@gmail.com

1. On définit bien une loi de probabilité !
$$\mathbb{P}(X \in [\![0, n]\!]) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1.$$
2. Pour $n = 1$, on remarque que
$$\mathbb{P}(X = 0) = \binom{1}{0} p^0 (1-p) = 1 - p \text{ et } \mathbb{P}(X = 1) = \binom{1}{1} p \times (1-p)^0 = p,$$
donc si $X \sim \mathscr{B}(1, p)$, alors $X \sim \mathscr{B}(p)$.
3. On démontrera bientôt que c'est la loi suivie par une somme de n variables de Bernoulli indépendantes de paramètre p (mais il faudra définir la notion d'indépendance de variables aléatoires !)

4.3 Couples de variables aléatoires

Définition 49

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un pdf. Soient X et Y deux variables aléatoires à valeurs dans E et F.

1. L'application $\begin{array}{l} \Omega \to E \times F \\ \omega \mapsto (X(\omega), Y(\omega)) \end{array}$ définit une variable aléatoire, notée (X, Y), de loi $\mathbb{P}_{(X,Y)}$, appelée loi jointe (ou conjointe) de X et Y.
2. Les événements élémentaires de (X, Y) sont les $\{X = x\} \cap \{Y = y\}$, notés aussi $\{(X, Y) = (x, y)\}$, ou encore $\{X = x, Y = y\}$.
On note alors indifféremment $\mathbb{P}(X = x, Y = y)$ ou $\mathbb{P}((X, Y) = (x, y))$ ou $\mathbb{P}(\{X = x\} \cap \{Y = y\})$.
3. La loi $\mathbb{P}_X$ est appelée première marginale de (X, Y), la loi $\mathbb{P}_Y$ seconde marginale.

Proposition 50

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un pdf. Soient X et Y deux variables aléatoires respectivement à valeurs dans E et F. Alors

1. pour tout x_0 dans E, $\mathbb{P}(X = x_0) = \displaystyle\sum_{b \in F} \mathbb{P}(X = x_0, Y = b)$,
2. pour tout y_0 dans E, $\mathbb{P}(Y = y_0) = \displaystyle\sum_{a \in E} \mathbb{P}(X = a, Y = y_0)$.

Démonstration. Il suffit de remarquer que, comme $Y(\Omega) \subset F$, $\{Y \in F\} = \Omega = \displaystyle\bigsqcup_{b \in F} \{Y = b\}$. Donc $(\{Y = b\})_{b \in F}$ est un système complet d'événements.
Donc, si $x_0 \in E$,
$$\mathbb{P}(X = x_0) = \sum_{b \in F} \mathbb{P}(\{X = x_0\} \cap \{Y = b\}) = \sum_{b \in F} \mathbb{P}(X = x_0, Y = b).$$

N. Laillet
nlaillet.math@gmail.com

On fait de même pour la deuxième proposition. □

Exercice 51

Soit $\alpha \in \mathbb{R}$, $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, X et Y deux variables aléatoires sur Ω à valeurs dans $[\![0, n]\!]^2$, telles que

$$\forall (k, \ell) \in [\![0, n]\!]^2,\ \mathbb{P}(X = k, Y = \ell) = \alpha \binom{n}{k}\binom{n}{\ell}.$$

1. Déterminer α,

2. Déterminer les lois de X et Y.

Remarque 52

La connaissance de la loi jointe permet de retrouver les lois marginales, **MAIS PAS L'INVERSE !** Imaginons que l'on tire des boules dans une urne contenant deux boules numérotées « 0 » et deux boules numérotées « 1 ». Notons X la variable aléatoire du résultat du premier tirage, Y celle du second tirage...

1. dans le cas sans remise,

$$\mathbb{P}(X = 0, Y = 0) = \frac{1}{6}, \mathbb{P}(X = 0, Y = 1) = \frac{1}{3},\ \mathbb{P}(X = 1, Y = 0) = \frac{1}{3},\ \mathbb{P}(X = 1, Y = 1) = \frac{1}{6}.$$

On en déduit que

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2},\ \mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = \frac{1}{2},$$

c'est-à-dire que X et Y suivent une loi de Bernoulli de paramètre $\frac{1}{2}$.

2. dans le cas avec remise,

$$\mathbb{P}(X = 0, Y = 0) = \mathbb{P}(X = 0, Y = 1) = \mathbb{P}(X = 1, Y = 0) = \mathbb{P}(X = 1, Y = 1) = \frac{1}{4},$$

et

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2},\ \mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = \frac{1}{2},$$

c'est-à-dire que X et Y suivent aussi une loi de Bernoulli de paramètre $\frac{1}{2}$.

Ainsi, les lois marginales sont les mêmes alors que les lois jointes sont différentes !

Proposition 53 (Loi conditionnelle)

Soit $(\Omega, \mathbb{P}(\Omega), \mathbb{P})$ un epf, E et F deux ensembles, X et Y deux v.a. sur Ω, à valeurs dans E et F.

1. Si $A \subset E$ et $B \subset F$ sont deux événements tels que $\mathbb{P}(X \in A) \neq 0$, on définit la probabilité de $\{Y \in B\}$ conditionnée à $\{X \in A\}$, notée $\mathbb{P}(Y \in B|X \in A)$ ou $\mathbb{P}_{\{X\in A\}}(Y \in B)$s, par
$$\mathbb{P}_{\{X\in A\}}(Y \in B) = \frac{\mathbb{P}(X \in A, Y \in B)}{\mathbb{P}(X \in A)}.$$
On définit de même, si $\mathbb{P}(Y \in B) \neq 0$, $\mathbb{P}_{\{Y\in B\}}(X \in A)$.
2. Si $A \subset E$ est un événement tel que $\mathbb{P}(X \in A) \neq 0$, on définit alors une loi de probabilité sur F appelée loi de Y conditionnée à $\{X \in A\}$.

Exemple 54

Soient (X, Y) deux variables aléatoires à valeurs dans $[\![0, n]\!]$ telles que

1. $X \sim \mathscr{B}\left(n, \frac{1}{2}\right)$
2. $\forall k \in [\![0, n]\!]$, la loi de Y conditionnée à $\{X = k\}$ est une loi binomiale de paramètres $\left(k, \frac{1}{2}\right)$.

Alors, pour tout k dans $[\![0, n]\!]$, pour tout i dans $[\![0, k]\!]$,
$$\mathbb{P}(Y = i|X = k) = \binom{k}{i}\frac{1}{2^k}.$$

On peut alors retrouver la loi à l'aide de la loi conditionnelle !

Proposition 55 (Probabilités totales)

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, X une variable aléatoire sur Ω à valeurs dans E fini, Y une variable aléatoire sur Ω à valeurs dans F.
Si $(A_1, \dots, A_r)$ est un système complet d'événements de E, si $B \subset F$, si $\forall i,\ \mathbb{P}(X \in A_i) \neq 0$, alors
$$\mathbb{P}(Y \in B) = \sum_{i=1}^{r} \mathbb{P}(X \in A_i, Y \in B) = \sum_{i=1}^{r} \mathbb{P}_{\{X\in A_i\}}(Y \in B)\mathbb{P}(X \in A_i).$$
En particulier, si E est fini et pour tout x dans E, $\mathbb{P}(X = x) \neq 0$, alors
$$\forall y_0 \in F,\ \mathbb{P}(Y = y_0) = \sum_{x\in E} \mathbb{P}_{X=x}(Y = y_0)\mathbb{P}(X = x).$$

N. Laillet
nlaillet.math@gmail.com

Exemple 56

(exemple précédent) Soient (X,Y) deux variables aléatoires à valeurs dans $[\![0,n]\!]$ telles que

1. $X \sim \mathscr{B}\left(n, \frac{1}{2}\right)$
2. $\forall k \in [\![0,n]\!]$, la loi de Y conditionnée à $\{X=k\}$ est une loi binomiale de paramètres $\left(k, \frac{1}{2}\right)$.

Déterminons la loi de Y. Si $i \in [\![0,n]\!]$,

$$\begin{aligned}
\mathbb{P}(Y=i) &= \sum_{k=0}^{n} \mathbb{P}(Y=i|X=k)\mathbb{P}(X=k) \\
&= \sum_{k=0}^{n} \binom{k}{i}\frac{1}{2^k}\binom{n}{k}\frac{1}{2^n} \\
&= \frac{1}{2^n}\sum_{k=i}^{n} \frac{1}{2^k}\frac{k!}{i!(k-i)!}\frac{n!}{k!(n-k)!} \\
&= \frac{1}{2^n}\sum_{k=i}^{n} \frac{1}{2^k}\frac{n!}{i!(n-k)!(k-i)!} \\
&=_{\ell=k-i} \frac{1}{2^{n+i}}\sum_{\ell=0}^{n-i} \frac{1}{2^\ell}\frac{n!}{i!(n-\ell-i)!\ell!} \\
&= \frac{1}{2^{n+i}}\frac{n!}{i!(n-i)!}\sum_{\ell=0}^{n-i} \frac{1}{2^\ell}\frac{(n-i)!}{(n-\ell-i)!\ell!} \\
&= \frac{1}{2^{n+i}}\binom{n}{i} \times \left(1+\frac{1}{2}\right)^{n-i} \\
&= \frac{3^{n-i}}{2^{2n}}\binom{n}{i} \\
&= \left(\frac{3}{4}\right)^{n-i}\left(\frac{1}{4}\right)^{i}\binom{n}{i},
\end{aligned}$$

donc Y suit une loi binomiale de paramètres $\left(n, \frac{1}{4}\right)$.

On peut étendre les définitions précédentes aux p-uplets de variables aléatoires

Définition 57

Soient $X_1, \dots, X_p$ p variables aléatoires à valeurs dans $E_1, \dots, E_p$. Alors

$$X = (X_1, \dots, X_p)$$

définit une variable aléatoire sur $E_1 \times \dots \times E_p$. La loi de X est appelée loi jointe de $X_1, \dots, X_p$.

N. Laillet
nlaillet.math@gmail.com

Pour $k \in [\![1, p]\!]$, la loi de X_k est appelée k-ième loi marginale de X. Elle est donnée par

$$\forall A \in E_k,\ \mathbb{P}_{X_k}(A) = \mathbb{P}_X(E_1 \times \cdots \times E_{k-1} \times A \times E_k \times \cdots \times E_p).$$

Exercice 58

Dans un centre d'appels, un opérateur appelle n personnes qui répondent avec une probabilité p. Quelle est la loi du nombre N_1 de personnes qui décrochent ?
Lors d'une seconde tentative, il rappelle les $n - N_1$ personnes qui n'ont pas répondu. Quelle est la loi du total $N_1 + N_2$ de personnes qui ont répondu ?

4.4 Indépendance

Définition 59

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

1. Deux variables X et Y à valeurs dans E et G sont indépendantes si, pour tout $A \subset E$ et tout $B \subset F$,
$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$
On note alors $X \perp\!\!\!\perp Y$.
2. Des variables aléatoires $(X_1, \ldots, X_p)$ à valeurs dans $E_1, \ldots, E_p$ sont mutuellement indépendantes si pour tous $A_1, \ldots, A_p$ dans $E_1, \ldots, E_p$, on a
$$\mathbb{P}(X_1 \in A_1, \ldots, X_p \in A_p) = \prod_{k=1}^{p} \mathbb{P}(X_k \in A_k).$$

Proposition 60

Deux variables aléatoires X et Y à valeurs dans E et F sont indépendantes si et seulement si pour tout x dans X et y dans Y on a

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

Démonstration. • Si $X \perp\!\!\!\perp Y$, $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X \in \{x\}, Y \in \{y\}) = \mathbb{P}(X \in \{x\})\mathbb{P}(Y \in \{y\}) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$.

N. Laillet
nlaillet.math@gmail.com

- Si $A \subset E$ et $B \subset F$,

$$\begin{aligned}\mathbb{P}(X \in A, Y \in B) &= \sum_{a\in A} \mathbb{P}(X = a, Y \in B) \\ &= \sum_{a\in A}\sum_{b\in B} \mathbb{P}(X = a, Y = b) \\ &= \sum_{a\in A}\sum_{b\in B} \mathbb{P}(X = a)\mathbb{P}(Y = b) \\ &= \sum_{a\in A} \mathbb{P}(X = a) \sum_{b\in B} \mathbb{P}(Y = b) \\ &= \mathbb{P}(A)\mathbb{P}(B),\end{aligned}$$

  d'où l'indépendance !

$\square$

Définition 61

Une variable aléatoire est dite *déterministe* si elle est constante presque sûrement.

Remarque 62

Quelles sont les variables aléatoires indépendantes d'elles-mêmes ?

Soit X une telle variable. Soit a une valeur telle que $\mathbb{P}(X = a) \neq 0$. Alors

$$\mathbb{P}(X = a) = \mathbb{P}(X = a, X = a) = \mathbb{P}(X = a)^2,$$

donc, **comme** $\mathbb{P}(X = a) \neq 0$, $\mathbb{P}(X = a) = 1$, donc X est presque sûrement constante égale à a.

Proposition 63

(lemme des coalitions) Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

1. Si X et Y sont deux variables aléatoires indépendantes sur Ω, à valeurs dans E et F (finis), alors

$$\forall \varphi : E \to G,\ \forall \psi : F \to G, \varphi(X) \text{ et } \psi(Y) \text{ sont indépendantes.}$$

2. Si $(X_1, \ldots, X_n)$ est un n-uplet de variables aléatoires mutuellement indépendantes sur Ω, à valeurs dans $E_1, \ldots, E_n$, alors pour tout k dans $[\![1, n]\!]$, pour toute $f : E_1 \times \cdots \times E_k \to F$, pour toute $g : E_{k+1} \times \cdots \times E_n \to G$, alors $f(X_1, \ldots, X_k)$ et $g(X_{k+1}, \ldots, X_n)$ sont indépendantes.

Démonstration. On ne montre que la première propriété, le reste se démontre de la même manière !
Soient $\varphi : E \to G$ et $\psi : F \to H$. Montrons que $\varphi(X)$ et $\psi(Y)$ sont indépendntes.
Soient $M \subset G$ et $N \subset H$.

N. Laillet
nlaillet.math@gmail.com

Notons $A = \varphi^{-1}(M)$ et $B = \psi^{-1}(N)$, les images réciproques respectives de M par φ et N par ψ. Alors, comme E et F sont finis, A et B sont aussi finis.
Ainsi,

$$\begin{aligned}\mathbb{P}(\varphi(X) \in M, \psi(Y) \in N) &= \mathbb{P}(X \in A, Y \in B)\\ &= \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \text{ par indépendance de } X \text{ et } Y\\ &= \mathbb{P}(\varphi(X) \in M)\mathbb{P}(\psi(Y) \in N),\end{aligned}$$

d'où l'indépendance des deux variables aléatoires ! $\square$

Remarque 64

1. La propriété précédente dit, par exemple, que si (X, Y, Z, T) sont 4 variables aléatoires indépendantes, $X + Y$ et $Z + T$ sont indépendantes.
2. Si $X \perp\!\!\!\perp Y$, alors

$$\mathbb{P}(Y \in B | X \in A) = \mathbb{P}(Y \in B), \ \forall (A, B) \in \mathcal{P}(E) \times \mathcal{P}(F).$$

Définition 65

Des variables aléatoires à valeurs dans le même espace E sont dites i.i.d. si elles sont indépendantes et identiquement distribuées, c'est-à-dire qu'elles suivent la même loi.

Proposition 66

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini. Soit $n \in \mathbb{N}^*$, $(X_1, \ldots, X_n)$ n variables aléatoires indépendantes, suivant toutes une loi de Bernoulli de paramètre $p \in]0, 1[$. Alors

$$X_1 + \cdots + X_n \sim \mathscr{B}(n, p).$$

Démonstration. Démontrons par récurrence sur n la proposition.

Initialisation. Pour $n = 1$, si $X_1 \sim \mathscr{B}(p)$, alors $X_1 \sim \mathscr{B}(1, p)$.
Hérédité. On suppose la propriété vraie pour un certain rang. Soient $X_1, \ldots, X_n, X_{n+1}$ $n+1$ variables aléatoires indépendantes, suivant toutes une loi de Bernoulli de paramètre p. Alors

- $\forall i \in [\![1, n]\!]$, $X_i \in \{0, 1\}$, donc $X_1 + \cdots + X_{n+1} \in [\![0, n+1]\!]$.
- Soit $k \in [\![0, n+1]\!]$. Notons

$$S_n = X_1 + \cdots + X_n, \ S_{n+1} = X_1 + \cdots + X_{n+1} = S_n + X_{n+1}.$$

N. Laillet
nlaillet.math@gmail.com

Alors

$$\begin{aligned}
\mathbb{P}(S_{n+1}=k) &= \mathbb{P}(S_{n+1}=k, X_{n+1}=0)+\mathbb{P}(S_{n+1}=k, X_{k+1}=1)\\
&= \mathbb{P}(S_n+X_{n+1}=k, X_{n+1}=0)+\mathbb{P}(S_n+X_{n+1}=k, X_{k+1}=1)\\
&= \mathbb{P}(S_n=k, X_{n+1}=0)+\mathbb{P}(S_n=k-1, X_{k+1}=1)\\
&= \mathbb{P}(S_n=k)\mathbb{P}(X_{n+1}=0)+\mathbb{P}(S_n=k-1)\mathbb{P}(X_{k+1}=1) \text{ par indépendance.}
\end{aligned}$$

Or,

— si $k=0$, $\mathbb{P}(S_n=k-1)=0$, donc

$$\mathbb{P}(S_{n+1}=0)=\mathbb{P}(S_n=0)\mathbb{P}(X_{n+1}=0)=(1-p)^n\times(1-p)=(1-p)^{n+1}=\binom{n+1}{0}p^0(1-p)^{n+1-0},$$

— si $k=n+1$, $\mathbb{P}(S_n=k)=0$, donc

$$\mathbb{P}(S_{n+1}=n+1)=\mathbb{P}(S_n=n)\mathbb{P}(X_{n+1}=1)=p^n\times p=\binom{n+1}{n+1}p^{n+1}(1-p)^{n+1-(n+1)},$$

— si $k\in[\![1,n]\!]$,

$$\begin{aligned}
\mathbb{P}(S_{n+1}=k) &= \mathbb{P}(S_n=k)\mathbb{P}(X_{n+1}=0)+\mathbb{P}(S_n=k-1)\mathbb{P}(X_{k+1}=1)\\
&= \binom{n}{k}p^k(1-p)^{n-k}\times(1-p)+\binom{n}{k-1}p^{k-1}(1-p)^{n-(k-1)}\times p\\
&= \left[\binom{n}{k}+\binom{n}{k-1}\right]p^k(1-p)^{n+1-k}\\
&= \binom{n}{k+1}p^k(1-p)^{n+1-k},
\end{aligned}$$

par la formule du triangle de Pascal.

□

Remarque 67

1. Le théorème précédent permet d'interpréter la loi binomiale comme celle du nombre de succès après n expériences répétées indépendamment.
2. **Attention !** Le théorème précédent ne dit pa sque si Y suit une loi binomiale, alorsil existe $(X_1,\dots,X_n)$ N variables de Bernoulli telles que $Y=X_1+\cdots+X_n$. Ce n'est, a priori, pas possible ! (c'est encore le problème entre l'égalité en lois et l'égalité de variables aléatoires)
Grossièrement, penser au fait qu'une variable suivant une loi binomiale possède $n+1$ états, alors que n variables de Bernoulli correspondent à n états.

N. Laillet
nlaillet.math@gmail.com

Exercice 68

Quelle est la loi de la somme de deux variables uniformes sur $[\![1, n]\!]$, indépendantes ?

Déjà, pour tout k dans $[\![1, n]\!]$, $\mathbb{P}(X = k) = \mathbb{P}(Y = k) = \dfrac{1}{n}$.
Ensuite, $X + Y$ est à valeurs dans $[\![2, 2n]\!]$.
Soit $k \in [\![2, 2n]\!]$. Alors

- si $k \leqslant n$,

$$
\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{i=1}^{n} \mathbb{P}(X+Y=k,\ Y=i) \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X+Y=k,\ Y=i) \text{ car si } Y \geqslant k,\ X \text{ doit être négatif, impossible.} \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X=k=i,\ Y=i) \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X=k=i,\ Y=i) \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X=k=i)\mathbb{P}(Y=i) \text{ par indépendance.} \\
&= \sum_{i=1}^{k-1} \frac{1}{n^2} =
\end{aligned}
$$

- si $k \leqslant n$,

$$
\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{i=1}^{n} \mathbb{P}(X+Y=k,\ Y=i) \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X+Y=k,\ Y=i) \text{ car si } Y \geqslant k,\ X \text{ doit être négatif, impossible.} \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X=k=i,\ Y=i) \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X=k=i,\ Y=i) \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X=k=i)\mathbb{P}(Y=i) \text{ par indépendance.} \\
&= \sum_{i=1}^{k-1} \frac{1}{n^2} \\
&= \frac{k-1}{n^2}
\end{aligned}
$$

- si $k \geqslant n+1$,

$$\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{i=1}^{n} \mathbb{P}(X+Y=k,\ Y=i) \\
&= \sum_{i=k-n}^{n} \mathbb{P}(X+Y=k,\ Y=i) \text{ car si } Y \leqslant k-n,\ X \text{ doit être } > n, \text{ impossible.} \\
&= \sum_{i=k-n}^{n} \mathbb{P}(X=k=i,\ Y=i) \\
&= \sum_{i=k-n}^{n} \mathbb{P}(X=k=i,\ Y=i) \\
&= \sum_{i=k-n}^{n} \mathbb{P}(X=k=i)\mathbb{P}(Y=i) \text{ par indépendance.} \\
&= \sum_{i=k-n}^{n} \frac{1}{n^2} \\
&= \frac{2n-k+1}{n^2}.
\end{aligned}$$

D'où la loi de $X+Y$.

5 Moments d'une variable aléatoire

Ici toutes nos variables aléatoires sont réelles. Une variable aléatoire réelle est parfois abrégée en v.a.r.

5.1 Espérance

5.1.1 Définition et calculs

Définition 69

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, X une variable aléatoire sur Ω, réelle, à valeurs dans $E \subset \mathbb{R}$, E fini, est

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

Une variable aléatoire d'espérance nulle est dite centrée.

Remarque 70

1. L'espérance doit être vue comme une sorte de moyenne.

N. Laillet
nlaillet.math@gmail.com

2. Si X est une v.a.r. telle que $\mathbb{P}(X=1)=\mathbb{P}(X=-1)=\dfrac{1}{2}$ (variable de Rademacher), alors
$$\mathbb{E}(X)=-1\times\mathbb{P}(X=-1)+1\times\mathbb{P}(X=1)=0.$$
Une variable de Rademacher est donc centrée.
3. En sup, on ne se posera jamais la question de « une variable aléatoire admet-elle une espérance » ? En spé, c'est une autre paire de manches ! En effet, en deuxième année, les variables aléatoires peuvent prendre un ensemble infini, indexé sur $\mathbb{N}$, de valeurs. Si X est à valeurs dans $\mathbb{N}$, on dira que X admet une espérance ssi la série de terme général $k\mathbb{P}(X=k)$ converge.

Proposition 71 (Espérances usuelles)

Soient p un réel de $]0,1[$, a et b deux entiers, X une v.a.r.
1. Si X suit une soi de Bernoulli de paramètre p, alors $\mathbb{E}(X)=p$.
2. Si X suit une loi uniforme sur $[\![a,b]\!]$, alors $\mathbb{E}(X)=\dfrac{a+b}{2}$.
3. Si X suit une loi binomiale de paramètres (n,p), alors $\mathbb{E}(X)=np$.

Démonstration.
1. Si $X\sim\mathscr{B}(p)$,
$$\mathbb{E}(X)=0\times\mathbb{P}(X=0)+1\times\mathbb{P}(X=1)=p.$$
2. Si $X\sim\mathcal{U}([\![a,b]\!])$, alors, pour tout k dans $[\![a,b]\!]$, $\mathbb{P}(X=k)=\dfrac{1}{b-a+1}$, donc
$$\mathbb{E}(X)=\sum_{k=a}^{b}k.\mathbb{P}(X=k)=\sum_{k=a}^{b}k\frac{1}{b-a+1}=\frac{1}{b-a+1}\sum_{k=a}^{b}k=\frac{1}{b-a+1}\frac{(a+b)(b-a+1)}{2}=\frac{a+b}{2}.$$
3. Si $X\sim\mathscr{B}(n,p)$,
$$\begin{aligned}\mathbb{E}(X)&=\sum_{k=0}^{n}k\times\mathbb{P}(X=k)\\&=\sum_{k=0}^{n}k\binom{n}{k}p^k(1-p)^{n-k}\\&=\sum_{k=1}^{n}n\binom{n-1}{k-1}p^k(1-p)^{n-k}\\&=n\sum_{\ell=1}^{n-1}\binom{n-1}{\ell}p^{\ell+1}(1-p)^{n-1-\ell}\\&=np\sum_{\ell=1}^{n-1}\binom{n-1}{\ell}p^{\ell}(1-p)^{n-1-\ell}=np.\end{aligned}$$

□

N. Laillet
nlaillet.math@gmail.com

Remarque 72

L'espérance **ne dépend que de la loi**, c'est-à-dire que si X et Y suivent la même loi, alors $\mathbb{E}(X) = \mathbb{E}(Y)$.

Proposition 73

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé, X une variable aléatoire sur Ω. Alors

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

Démonstration. On sait que $\mathbb{E}(X) = \sum_{x \in E} x \mathbb{P}(\{X = x\})$. Or, $\{X = x\} = X^{-1}(\{x\})$.
Mais, comme X est à valeurs dans E. $\Omega = \sqcup_{x \in E} X^{-1}(\{x\})$.
De plus,

$$\mathbb{P}(X = z) = \sum_{\omega \in X^{-1}(\{x\})} \mathbb{P}(\{\omega\}),$$

donc

$$\mathbb{E}(X) = \sum_{x \in E} x \sum_{\omega \in X^{-1}(\{x\})} \mathbb{P}(\{\omega\}) = \sum_{x \in E} \sum_{\omega \in X^{-1}(\{x\})} X(\omega)\mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}).$$

□

Proposition 74

Soient X et Y deux variables aléatoires réelles, a, b dans $\mathbb{R}$.

1. $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(X)$.
2. si X est à valeurs positives, alors $\mathbb{E}(X) \geqslant 0$.
3. plus généralement, si $\mathbb{P}(X \geqslant 0) = 1$, $\mathbb{E}(X) \geqslant 0$.
4. si X est positive presque sûrement, $\mathbb{E}(X) = 0$ si et seulement si $\mathbb{P}(X = 0) = 1$ (X est presque sûrement nulle).
5. (croissance de l'espérance) si $X \leqslant Y$, alors $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$. Plus généralement, si $\mathbb{P}(X \leqslant Y) = 1$, $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

Démonstration. **1.** On fait le calcul à l'aide de la formule précédente :

$$\begin{aligned}
\mathbb{E}(aX + bY) &= \sum_{\omega \in \Omega} (aX + bY)(\omega)\mathbb{P}(\{\omega\}) \\
&= \sum_{\omega \in \Omega} aX(\omega)\mathbb{P}(\{\omega\}) + bY(\omega)\mathbb{P}(\{\omega\}) \\
&= a\sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) + b\sum_{\omega \in \Omega} Y(\omega)\mathbb{P}(\{\omega\}) \\
&= a\mathbb{E}(X) + b\mathbb{E}(Y).
\end{aligned}$$

N. Laillet
nlaillet.math@gmail.com

2. Fait dans la section suivante.
3. Comme X est à valeurs dans $\mathbb{R}$, on a $\Omega = X^{-1}(\mathbb{R}_-^*) \sqcup X^{-1}(\mathbb{R}_+)$. On écrit donc

$$\begin{aligned}\mathbb{E}(X) &= \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\}) \\ &= \sum_{\omega\in X^{-1}(\mathbb{R}_-^*)} X(\omega)\mathbb{P}(\{\omega\}) + \sum_{\omega\in X^{-1}(\mathbb{R}_+)} X(\omega)\mathbb{P}(\{\omega\}) \\ &= \sum_{\omega\in X^{-1}(\mathbb{R}_+)} X(\omega)\mathbb{P}(\{\omega\}) \text{ par hypothèse} \\ &\geqslant 0,\end{aligned}$$

car $X(\omega) \geqslant 0$ et $\mathbb{P}(\{omega\}) \geqslant 0$ pour tout ω dans $X^{-1}(\mathbb{R}_+)$.
4. On écrit, plus subtilement,

$$\Omega = X^{-1}(\mathbb{R}_-^*) \sqcup X^{-1}(\{0\}) \sqcup X^{-1}(\mathbb{R}_+^*),$$

donc

$$\mathbb{E}(X) = \sum_{\omega\in X^{-1}(\mathbb{R}_+^*)} X(\omega)\mathbb{P}(\{\omega\}) + \sum_{\omega\in X^{-1}(\{0\})} X(\omega)\mathbb{P}(\{\omega\}).$$

Si $\mathbb{P}(X = 0) \neq 1$ alors $\exists\omega \in \Omega$ tel que $X(\omega) > 0$ et $\mathbb{P}(\omega) > 0$, donc $\mathbb{E}(X) > 0$. La contraposée nous assure le résultat.
5. Si $\mathbb{P}(X \leqslant Y) = 1$, alors $\mathbb{P}(Y - X \geqslant 0) = 1$. Donc $\mathbb{E}(Y - X) \geqslant 0$, donc $\mathbb{E}(Y) - \mathbb{E}(X) \geqslant 0$ par linéarité de l'espérance, donc $\mathbb{E}(Y) \geqslant \mathbb{E}(X)$.

□

Remarque 75

1. On peut, grâce à la linéarité de l'espérance, retrouver l'espérance d'une variable suivant une loi binomiale.
   Soit X suivant une loi binomiale. Alors X **suit la même loi que** $X_1 + \cdots + X_n$, où $(X_1, \ldots, X_n)$ sont n vaiid de Bernoulli. (**attention ! Je ne dis pas que** $X = X_1 + \cdots + X_n$)
   Mais alors, comme l'espérance ne dépend que de la loi,
   $$\mathbb{E}(X) = \mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = np.$$
2. Si X est une v.a.r., $X + 2$ est aussi une v.a.r. Quelle est son espérance ?
   C'est $\mathbb{E}(X) + 2$. En effet, par linéarité, $\mathbb{E}(X + 2) = \mathbb{E}(X) + \mathbb{E}(2) = \mathbb{E}(X) + 2$, car 2 doit être interprétée comme une variable aléatoire déterministe, constante égale à 2.

Il faut penser à utiliser les fonctions indicatrices comme des variables aléatoires.

Proposition 76

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini, A un événement.

N. Laillet
nlaillet.math@gmail.com

1. $\mathbb{1}_A : \left| \begin{array}{l} \Omega \to \{0,1\} \\ \omega \mapsto \begin{cases} 0 \text{ si } \omega \notin A \\ 1 \text{ sinon.} \end{cases} \end{array} \right.$ est une variable de Bernoulli de paramètre $\mathbb{P}(A)$.
2. $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.

Proposition 77 (Formule de transfert)

Soit X une variable aléatoire à valeurs dans E et $f : E \to F \subset \mathbb{R}$. Alors

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

Démonstration. On sait que $\mathbb{E}(f(X)) = \sum_{\omega \in \Omega} f(X(\omega))\mathbb{P}(\{x\})$. Mais comme X est à valeurs dans E, on sait que

$$\Omega = \bigsqcup_{x \in E} X^{-1}(\{x\}).$$

Donc

$$\mathbb{E}(f(X)) = \sum_{x \in E} \sum_{\omega \in X^{-1}(\{x\})} f(X(\omega))\mathbb{P}(f(X(\omega))).$$

Mais, si $\omega \in X^{-1}(\{x\})$, $f(X(\omega)) = f(x)$, donc

$$\mathbb{E}(f(X)) = \sum_{x \in E} f(x) \sum_{\omega \in X^{-1}(\{x\})} \mathbb{P}(\{\omega\}) = \sum_{x \in E} f(x)\mathbb{P}(X = x),$$

d'où le résultat désiré. $\square$

Remarque 78

1. Dans la formule de Transfert, la variable X **n'a pas besoin d'être réelle** ! C'est f qui doit être à valeurs réelles.
2. En particulier, X peut être un **couple de variables aléatoires.** Par exemple, si Y est une variable aléatoire à valeurs dans E et Z est une variable aléatoire à valeurs dans F, si $f : (s,t) \mapsto se^t$,
$$\mathbb{E}(f(Y,Z)) = \sum_{y \in E} \sum_{z \in F} ye^z \mathbb{P}(Y = y, Z = z).$$

N. Laillet
nlaillet.math@gmail.com

Exemple 79

Soit $X \sim \mathcal{U}(\llbracket 0, n \rrbracket)$. Calculons $\mathbb{E}(e^X)$. Par la formule de transfert,

$$\mathbb{E}(X) = \sum_{k=0}^{n} e^k \mathbb{P}(X = k) = \sum_{k=0}^{n} \frac{1}{n+1} e^k = \frac{1}{n+1} \frac{1 - e^{n+1}}{1 - e}.$$

5.1.2 Espérance et indépendance

Proposition 80

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini. Soient X et Y deux variables aléatoires réelles, à valeurs respectives dans E et dans F, parties finies de $\mathbb{R}$.
Si X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Démonstration. ⊛ On aura besoin ici de la formule de transfert. On sait que

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{x \in E} \sum_{y \in F} xy \mathbb{P}(X = x, Y = y) \\
&= \sum_{x \in E} \sum_{y \in F} xy \mathbb{P}(X = x)\mathbb{P}(Y = y) \text{ par indépendance de } X \text{ et } Y \\
&= \sum_{x \in E} x\mathbb{P}(X = x) \sum_{y \in F} y\mathbb{P}(Y = y) \\
&= \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}$$

□

Remarque 81

ATTENTION !!!!! La réciproque est très fausse !
En effet, soit X uniforme sur $-1, 0, 1$ et Y la fonction indicatrice de $X = 0$, c'est-à-dire que Y vaut 1 si $X = 0$ et 0 sinon.
Alors

- $XY = 0$ car si $X \neq 0$, $Y = 0$, donc $\mathbb{E}(XY) = 0$,
- $\mathbb{E}(X) = 0$, donc $\mathbb{E}(X)\mathbb{E}(Y) = 0$.

Mais X et Y ne sont pas indépendantes : en effet, $\mathbb{P}(X = 0, Y = 1) = \mathbb{P}(X = 0) \neq \mathbb{P}(X = 0) \times \mathbb{P}(Y = 1)$.

Exercice 82

Démontrer que deux variables aléatoires à valeur dans un même ensemble E sont indépendantes si et seulement si pour toutes fonctions f et g définies sur E et à valeurs dans $\mathbb{R}$, $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$.
Réponse.

N. Laillet
nlaillet.math@gmail.com

Le sens direct est immédiat car si $X \perp\!\!\!\perp Y$, alors $f(X) \perp\!\!\!\perp g(Y)$.
Pour le sens réciproque, on suppose que pour toutes fonctions f et g définies sur E, $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$. Soit $x \in E$ et $y \in E$. Alors si $f : t \mapsto \begin{cases} 1 \text{ si } t = x \\ 0 \text{ sinon.} \end{cases}$ et $g : t \mapsto \begin{cases} 1 \text{ si } t = y \\ 0 \text{ sinon.} \end{cases}$, on a

$$f(X) = \mathbb{1}_{\{X=x\}},\ g(Y) = \mathbb{1}_{\{Y=y\}}, \text{ donc } f(X)g(Y) = \mathbb{1}_{\{X=x,Y=y\}},$$

d'où, comme, par hypothèse, $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$,

$$\mathbb{E}(\mathbb{1}_{\{X=x,Y=y\}}) = \mathbb{E}(\mathbb{1}_{\{X=x\}})\mathbb{E}(\mathbb{1}_{\{Y=y\}}),$$

on en déduit

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

d'où l'indépendance de X et Y.

5.1.3 Inégalités

Proposition 83 (Inégalité de Markov)

Soit X une v.a.r. positive, $a > 0$. Alors pour tout $a > 0$, $\mathbb{P}(X \geqslant a) \leqslant \dfrac{\mathbb{E}(X)}{a}$.

Démonstration. ⊛ On sait que pour tout ω dans Ω,

$$X(\omega) \geqslant X(\omega)\mathbb{1}_{\{X\geqslant a\}}(\omega) \geqslant a\mathbb{1}_{X\geqslant a}$$

En effet,

- pour la première inégalité, on sait que pour tout ω, $0 \leqslant X(\omega) \leqslant 1$ donc, comme $0 \leqslant \mathbb{1}_{\{X\geqslant a\}}(\omega) \leqslant 1$, on en déduit la première inégalité.
- pour la seconde, si $X(\omega) \geqslant a$, l'inégalité est immédiate, et, si $X(\omega) < a$, $\mathbb{1}_{\{X\geqslant a\}}(\omega) = 0$.

Donc, par croissance de l'espérance,

$$\mathbb{E}(X) \geqslant \mathbb{E}(a\mathbb{1}_{\{X\geqslant a\}}) = a\mathbb{E}(\mathbb{1}_{\{X\geqslant a\}} = a\mathbb{P}(X \geqslant a),$$

d'où le résultat, et l'inégalité de Markov en divisant par a. □

Remarque 84

Il est tout à fait autorisé, si l'on comprend bien que l'on revient aux éléments de Ω,

$$X \geqslant X\mathbb{1}_{\{X\geqslant a\}} \geqslant a\mathbb{1}_{\{X\geqslant a\}}$$

N. Laillet
nlaillet.math@gmail.com

Proposition 85 (Inégalité de Cauchy-Schwarz)

Soient X et Y deux variables aléatoires. Alors $\mathbb{E}(XY)^2 \leqslant \mathbb{E}(X^2)\mathbb{E}(Y^2)$, avec égalité si, et seulement si X et Y sont presque sûrement colinéaires, c'est-à-dire qu'il existe λ dans $\mathbb{R}$ tel que $X = \lambda Y$ presque sûrement.

Démonstration. ⊛ La méthode est importante ! Définissons, pour tout t dans $\mathbb{R}$,

$$Q(t) = \mathbb{E}((X + tY)^2) = \mathbb{E}(X^2 + 2tXY + t^2X^2) = \mathbb{E}(X^2) + 2t\mathbb{E}(XY) + t^2\mathbb{E}(Y^2).$$

Mais, par positivité de l'espérance, pour tout t dans $\mathbb{R}$, $Q(t) \geqslant 0$. Donc, comme Q est un polynôme du second degré et de signe constant, donc **son discriminant est négatif.** Ainsi,

$$4\mathbb{E}(XY)^2 - 4\mathbb{E}(X^2)\mathbb{E}(Y^2) \leqslant 0,$$

c'est-à-dire

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

On a égalité si et seulement si le discriminant de Q est nul, c'est-à-dire que Q admet une racine réelle. Donc il y a égalité si et seulement s'il existe $t_0 \in \mathbb{R}$ tel que

$$\mathbb{E}((X + t_0Y)^2) = 0,$$

i.e. $X + t_0Y = 0$ presque sûrement, i.e. $X = -t_0Y$ presque sûrement, c'est-à-dire que X et Y sont presque sûrement colinéaires. □

Exercice 86

Soit X une var et $f : \mathbb{R} \to \mathbb{R}_+$ croissante. Alors pour tout a dans $\mathbb{R}$ tel que $f(a) > 0$,

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(f(X))}{f(a)}.$$

5.2 Variance et covariance

5.2.1 Variance

Définition 87

Soit X une v.a.r.

1. Soit $k \in \mathbb{N}$. Le moment d'ordre k de X est le réel $\mathbb{E}(X^k)$, son moment d'ordre k centré est $\mathbb{E}((X - \mathbb{E}(X))^k)$.
2. La variance de X est son moment d'ordre 2 centré, $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$.
3. L'écart-type de X est $\sigma(X) = \sqrt{\mathbb{V}(X)}$.
4. Une variable centrée est dite réduite si sa variance est égale à 1.

N. Laillet
nlaillet.math@gmail.com

Remarque 88

1. La variance, comme tous les moments (centrés) d'une variable aléatoire, ne dépend que de la loi de la variable aléatoire.
2. $X - \mathbb{E}(X)$ est bien centrée : en effet,
$$\mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(X) = 0.$$
3. La variance permet de mesurer la dispersion, i.e. les écarts à la moyene : la variance, c'est

   « la moyenne des carrés des écarts à la moyenne. »
4. Que signifie, pour une variable aléatoire, d'être de variance nulle ?
$$\begin{aligned} \mathbb{V}(X) = 0 &\Leftrightarrow \mathbb{E}((X - \mathbb{E}(X))^2) = 0 \\ &\Leftrightarrow X - \mathbb{E}(X) = 0 \text{ p.s.} \\ &\Leftrightarrow X = \mathbb{E}(X) \text{ p.s.} \\ &\Leftrightarrow X \text{ est constante p.s.} \end{aligned}$$
5. De même qu'au 3. de la remarque 70, il n'est pas évident qu'une variable aléatoire possède un moment d'ordre k. En sup, oui c'est évident car les variables aléatoires sont à valeurs dans un ensemble fini. En spé, ça l'est moins, et il faut utiliser de la théorie des séries.

Proposition 89 (Relation de Koenig-Huygens)

Soit X une v.a.r. Alors $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

Démonstration. La preuve est immédiate :
$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2. \end{aligned}$$
□

Proposition 90 (Variances usuelles)

1. Si $X \sim \mathbb{U}([\![0, n]\!])$, $\mathbb{V}(X) = \dfrac{n(n+2)}{12}$ (ne pas la connaître par coeur, savoir la retrouver)
2. Si $X \sim \mathscr{B}(p)$, $\mathbb{V}(X) = p(1-p)$ (à connaître par coeur).
3. Si $X \sim \mathscr{B}(n, p)$, $\mathbb{V}(X) = np(1-p)$ (à connaître par coeur).

N. Laillet
nlaillet.math@gmail.com

Démonstration. **1.** Par la formule de Koenig-Huygens, $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Or, on a déjà vu que $\mathbb{E}(X) = \dfrac{n}{2}$.

Par la formule de transfert,

$$\mathbb{E}(X^2) = \sum_{k=0}^{n} k^2 \mathbb{P}(X = k) = \frac{1}{n+1} \sum_{k=0}^{n} k^2 = \frac{1}{n+1} \frac{n(n+1)(2n+1)}{6} = \frac{n(2n+1)}{6},$$

donc

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

2. Si $X \sim \mathscr{B}(p)$, $\mathbb{E}(X) = p$ et $\mathbb{E}(X^2) = p$ car $X^2 = X$, donc

$$\mathbb{V}(X) = p - p^2 = p(1-p).$$

3. (en cours, on ne va pas faire le calcul, mais vu qu'on peut le faire, faisons-le ici !) Toujours par la formule de transfert,

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} kn \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&=_{\ell=k-1} \sum_{\ell=0}^{n-1} (\ell+1) n \binom{n-1}{\ell} p^{\ell+1} (1-p)^{n-1-\ell} \\
&= \sum_{\ell=0}^{n-1} \ell n \binom{n-1}{\ell} p^{\ell+1} (1-p)^{n-1-\ell} + n \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^{\ell+1} (1-p)^{n-1-\ell} \\
&= A + B.
\end{aligned}$$

Déjà,

$$B = np \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^{\ell} (1-p)^{n-1-\ell} = np.$$

N. Laillet
nlaillet.math@gmail.com

Ensuite,

$$\begin{aligned}
A &= n\sum_{\ell=0}^{n-1}\ell\binom{n-1}{\ell}p^{\ell+1}(1-p)^{n-1-\ell}\\
&= n\sum_{\ell=1}^{n-1}\ell\binom{n-1}{\ell}p^{\ell+1}(1-p)^{n-1-\ell}\\
&= n\sum_{\ell=1}^{n-1}(n-1)\binom{n-2}{\ell-1}p^{\ell+1}(1-p)^{n-1-\ell}\\
&=_{j=\ell-1} n(n-1)\sum_{j=0}^{n-2}\binom{n-2}{j}p^{j+2}(1-p)^{n-2-j}\\
&= n^2p^2\sum_{j=0}^{n-2}\binom{n-2}{j}p^{j}(1-p)^{n-2-j} = n(n-1)p^2,
\end{aligned}$$

donc

$$\mathbb{E}(X^2) = n^2p^2 + np - np^2 = n^2p^2 + np(1-p),$$

donc

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = n^2p^2 + np(1-p) - n^2p^2 = np(1-p),$$

d'où le résultat désiré !

□

Proposition 91

1. Soit X une v.a.r., a et b des réels. Alors
$$V(aX+b) = a^2V(X).$$
2. Soit X une v.a.r., d'espérance μ, de variance σ^2. Alors $\dfrac{X-\mu}{\sigma}$ est centrée réduite.

Démonstration. **1.** On calcule

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

par linéarité de l'espérance. Donc

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

donc

$$\mathbb{V}(aX+b) = \mathbb{E}\left(\left[aX+b-\mathbb{E}(aX+b)\right]^2\right) = \mathbb{E}(a^2(X-\mathbb{E}(X))^2) = a^2\mathbb{E}((X-\mathbb{E}(X))^2) = a^2\mathbb{V}(X).$$

2. Déjà,

$$\mathbb{E}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}(\mathbb{E}(X)-\mu) = \frac{1}{\sigma}(\mu-\mu) = 0,$$

N. Laillet
nlaillet.math@gmail.com

donc $\dfrac{X-\mu}{\sigma^2}$ est bien centrée. Ensuite,

$$\mathbb{V}\left(\frac{X-\mu}{\sigma}\right) = \mathbb{V}\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}\mathbb{V}(X) = \frac{\sigma^2}{\sigma^2} = 1,$$

donc $\dfrac{X-\mu}{\sigma^2}$ est centrée réduite.

□

Proposition 92

1. Soient X et Y deux v.a.r. indépendantes. Alors
$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$
2. Soient $(X_1, \dots, X_n)$ n v.a.r. indépendantes. Alors
$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathbb{V}(X_i).$$

Démonstration. **1.** Par la formule de Koenig-Huygens,

$$\begin{aligned}
\mathbb{V}(X+Y) &= \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2 \\
&= \mathbb{E}(X^2 + 2XY + Y^2) - \Big(\mathbb{E}(X) + \mathbb{E}(Y)\Big)^2 \\
&= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)^2 \\
&= \mathbb{E}(X^2) + 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y^2) - \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)^2 \text{ par indépendance de } X \text{ et } Y \\
&= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\
&= \mathbb{V}(X) + \mathbb{V}(Y),
\end{aligned}$$

d'où le résultat !

□

Remarque 93

1. Comme cette proposition est une conséquence de $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ lorsque $X \perp\!\!\!\perp Y$, on peut aussi préciser que « $V(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$ » **n'entraîne pas** que $X \perp\!\!\!\perp Y$. (prendre le même contre-exemple que précédemment)
2. Que se passe-t-il lorsque les variables ne sont pas indépendantes ? Y a-t-il une formule ?

5.2.2 Covariance

Définition 94

Soient X et Y deux v.a.r. sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

1. On appelle *covariance* de X et Y le réel $\mathrm{Cov}(X,Y) = \mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))$.
2. Deux v.a.r. de covariance nulle sont dites non corrélées.

Proposition 95

Soient X et Y deux v.a.r. sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

1. $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
2. $V(X) = \mathrm{Cov}(X,X)$
3. Cov est symétrique, i.e. $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
4. Cov est bilinéaire, i.e.

$$\begin{aligned}
&\forall (A,B,C) \text{ v.a.r}, \forall (\lambda,\mu) \in \mathbb{R}^2,\\
&\mathrm{Cov}(\lambda A + \mu B, C) = \lambda \mathrm{Cov}(A,C) + \mu \mathrm{Cov}(B,C),\\
&\mathrm{Cov}(A, \lambda B + \mu C) = \lambda \mathrm{Cov}(A,B) + \mu \mathrm{Cov}(A,C).
\end{aligned}$$

Démonstration.

1. Effectuons simplement le calcul :

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))\\
&= \mathbb{E}(XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y))\\
&= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y) \text{ (on n'oublie pas que } \mathbb{E}(X) \text{ et } \mathbb{E}(Y) \text{ sont des } \textbf{constant}\\
&= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}$$

2. On calcule
$$\mathrm{Cov}(X,X) = \mathbb{E}((X-\mathbb{E}(X))(X-\mathbb{E}(X))) = \mathbb{V}(X).$$

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

4. Par symétrie, il ne suffit de démontrer la linéarité que par rapport à la première variable. Soient A, B, C trois v.a.r., λ et μ deux réels. Alors

$$\begin{aligned}
\mathrm{Cov}(\lambda A + \mu B, C) &= \mathbb{E}\Big((\lambda A + \mu B)C\Big) - \mathbb{E}(\lambda A + \mu B)\mathbb{E}(C)\\
&= \mathbb{E}(\lambda AC + \mu BC) - (\lambda\mathbb{E}(A) + \mu\mathbb{E}(B))\mathbb{E}(C)\\
&= \lambda\mathbb{E}(AC) + \mu\mathbb{E}(BC) - \lambda\mathbb{E}(A)\mathbb{E}(C) - \mu\mathbb{E}(B)\mathbb{E}(C)\\
&= \lambda\Big(\mathbb{E}(AC) - \mathbb{E}(A)\mathbb{E}(C)\Big) + \mu\Big(\mathbb{E}(BC) - \mathbb{E}(B)\mathbb{E}(C)\Big)\\
&= \lambda\mathrm{Cov}(A,C) + \mu\mathrm{Cov}(B,C),
\end{aligned}$$

N. Laillet
nlaillet.math@gmail.com

d'où la linéarité par rapport à la première variable et, par symétrie, la linéarité par rapport à la seconde.

□

Proposition 96

Soient X et Y deux v.a.r. sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

1. $\mathbb{V}(X+Y) = \mathbb{V}(X) + 2\mathrm{Cov}(X,Y) + \mathbb{V}(Y)$.
2. Plus généralement si $X_1, \ldots, X_n$ sont des v.a.r., $V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbb{V}(X_k) + 2\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i, X_j)$

Démonstration.

1. On utilise la proposition précédente
$$\begin{aligned}\mathbb{V}(X+Y) &= \mathrm{Cov}(X+Y, X+Y)\\ &= \mathrm{Cov}(X, X+Y) + \mathrm{Cov}(Y, X+Y) \text{ par linéarité par rapport à la première variable}\\ &= \mathrm{Cov}(X,X) + \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X) + \mathrm{Cov}(Y,Y) \text{ par linéarité par rapport à la seconde variable}\\ &= \mathbb{V}(X) + 2\mathrm{Cov}(X,Y) + \mathbb{V}(Y) \text{ par la formule reliant variance et covariance, ainsi que par symétrie de la}\end{aligned}$$
D'où le résultat désiré.
2. On fait cette preuve par récurrence.

□

Proposition 97 (Variables non corrélées, indépendantes.)

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

1. deux variables aléatoires réelles X et Y sur Ω son non corrélées si, et seulement si $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.
2. si $X_1, \ldots, X_n$ sont des v.a.r. deux à deux non corrélées, $\mathbb{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbb{V}(X_k)$.
3. si deux variables aléatoires réelles X et Y sur Ω sont indépendantes, alors elles sont non corrélées.

Démonstration.

1. On remarque que $\mathbb{V}(X+Y) - \mathbb{V}(X) - \mathbb{V}(Y) = 2\mathrm{Cov}(X,Y)$. Donc X et Y sont non corrélées ssi $2\mathrm{Cov}(X,Y) = 0$, i.e. ssi $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.
2. On utilise la formule 2. de la proposition 96.
3. si $X \perp\!\!\!\perp Y$, alors $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y) = 0$ par indépendance.

□

N. Laillet
nlaillet.math@gmail.com

Remarque 98

La réciproque du 2. de la proposition 97 est fausse ! La formule peut tout à fait être vérifiée sans qu'on ait non-corrélation. Par exemple, si X et Y sont deux variables de Radema indépendantes, de paramètre $\frac{1}{2}$, et si Z est l'indicatrice de $X = -Y = 1$ ($Z = 1$ si $X = 1$ et $Y = -1$, $Z = 0$ sinon), alors $\mathrm{Cov}(X, Y) = 0$ par indépendance de X et Y, mais

$$\mathrm{Cov}(X, Z) = \mathbb{E}(XZ) - \mathbb{E}(X)\mathbb{E}(Z) = \mathbb{E}(XZ) \text{ car } X \text{ est centrée.}$$

Or,

$$\mathbb{E}(XZ) = \mathbb{P}(X = 1, Z = 1) = \mathbb{P}(X = 1, Y = -1) = \frac{1}{4} \neq 0,$$

donc X et Z sont corrélées. De même pour Y et Z, on a $\mathrm{Cov}(Y, Z) = -\frac{1}{4}$. Ainsi,

$$\begin{aligned}\mathbb{V}(X + Y + Z) &= \mathbb{V}(X) + \mathbb{V}(Y) + \mathbb{V}(Z) + 2\mathrm{Cov}(X, Y) + 2\mathrm{Cov}(X, Z) + 2\mathrm{Cov}(Y, Z)\\ &= \mathbb{V}(X) + \mathbb{V}(Y) + \mathbb{V}(Z) + 0 + 2 \times \frac{1}{4} + 2 \times \frac{-1}{4} = 0.\end{aligned}$$

5.2.3 Inégalité pour la covariance

Proposition 99 (Inégalité de Cauchy-Schwarz pour la covariance)

Soient X et Y deux v.a.r. Alors

$$\mathrm{Cov}(X, Y)^2 \leqslant \mathbb{V}(X)\mathbb{V}(Y),$$

avec égalité si et seulement si X est presque sûrement égale à une fonction affine de Y, i.e. ssi il existe α et β tels que $X = \alpha Y + \beta$ presque sûrement.

Démonstration. ⊛ Soit $Q : t \mapsto \mathbb{V}(X + tY)$. Pour tout t dans $\mathbb{R}$,

$$Q(t) = \mathbb{V}(X) + 2\mathrm{Cov}(X, tY) + \mathbb{V}(tY) = \mathbb{V}(X) + 2t\mathrm{Cov}(X, Y) + t^2\mathbb{V}(Q),$$

donc Q est un polynôme du second degré en t.
Or, une variance est toujours positive, donc le discriminant de Q est négatif ou nul. Donc

$$4\mathrm{Cov}(X, Y)^2 - 4\mathbb{V}(X)\mathbb{V}(Y) \leqslant 0, \text{ i.e. } \mathrm{Cov}(X, Y)^2 \leqslant \mathbb{V}(X)\mathbb{V}(Y).$$

Ensuite,

$$\begin{aligned}\mathrm{Cov}(X, Y)^2 = \mathbb{V}(X)\mathbb{V}(Y) &\Leftrightarrow \text{ le discriminant de } Q \text{ est nul}\\ &\Leftrightarrow \exists t_0 \in \mathbb{R},\ \mathbb{V}(X + t_0Y) = 0\\ &\Leftrightarrow \exists t_0 \in \mathbb{R}, \exists K \in \mathbb{R},\ X + t_0Y = K \text{ p.s.}\\ &\Leftrightarrow \exists t_0 \in \mathbb{R}, \exists K \in \mathbb{R},\ X = -t_0Y + K \text{ p.s.,}\end{aligned}$$

d'où le cas d'égalité ! □

N. Laillet
nlaillet.math@gmail.com

Définition 100

Soient X et Y deux v.a.r. On définit le coefficient de corrélation de X et de Y comme $\rho = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}$.

Remarque 101

1. Par l'inégalité de Cauchy-Schwarz, $\rho \leqslant 1$.
2. Si $\rho = 0$, alors X et Y sont non corrélées.
3. Si ρ est proche de 1, alors X et Y sont fortement corrélées.

5.2.4 Vers la loi des grands nombres

De même qu'on a l'inégalité de Markov pour les variables aléatoires, une fois qu'on a introduit la notion de variance, il existe une autre inégalité, appelée inégalité de Bienaymé-Tchebycheff.

Proposition 102 (Inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire réelle. Alors pour tout $\varepsilon > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) \leqslant \frac{V(X)}{\varepsilon^2}.$$

Démonstration. Soit $\varepsilon > 0$. Alors

$$\begin{aligned}\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) &= \mathbb{P}((X - \mathbb{E}(X))^2 \geqslant \varepsilon^2) \text{ car } \varepsilon > 0\\ &\leqslant \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{\varepsilon^2} \text{ par l'inégalité de Markov}\\ &= \frac{\mathbb{V}(X)}{\varepsilon^2}.\end{aligned}$$

□

Remarque 103

1. Cette inégalité montre qu'une variable aléatoire ne peut pas « trop » s'éloigner de son espérance.
2. Il y a des conditions pour que les inégalités de Markov et de Bienaymée-Tchebychev puissent s'appliquer : pour l'inégalité de Markov, il faut que X admette une espérance. Pour celle de Bienaymé-Tchebycheff, il faut que X admette une variance (non trivial dans le cas d'un univers infini !)

N. Laillet
nlaillet.math@gmail.com

Cette inégalité sert à comprendre les bases des statistiques : vous vous êtes peut-être déjà demandés pourquoi, pour savoir la proportion d'un certain caractère dans une population, on faisait un test sur un « grand » échantillon de personnes, et que l'on jugeait ensuite que cela suffisait ?

Proposition 104 (Résultat admis, ultra-HP, dû à Kolmogorov)

Soit X une v.a. Il existe un univers Ω (infini) sur lequel on peut construire une suite (X_k) de v.a. i.i.d.

Proposition 105 (Loi faible des grands nombres)

Soient $(X_k)_{k\in\mathbb{N}}$ une suite de variables aléatoires i.i.d. d'espérance m. Soit pour tout n dans $\mathbb{N}$ $S_n = X_1 + \cdots + X_n$. Alors pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \underset{+\infty}{\longrightarrow} 0.$$

Démonstration. Soit $n \in \mathbb{N}^*$.
Par linéarité de l'espérance,

$$\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\mathbb{E}(S_n) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = m$$

Ainsi,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) &= \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| \geqslant \varepsilon\right)\\ &\leqslant \frac{\mathbb{V}\left(\frac{S_n}{n}\right)}{\varepsilon^2} \text{ par l'inégalité de Bienaymé-Tchebycheff.}\end{aligned}$$

Mais

$$\begin{aligned}\mathbb{V}\left(\frac{S_n}{n}\right) &= \frac{1}{n^2}\mathbb{V}(S_n)\\ &= \frac{1}{n^2}\mathbb{V}\left(\sum_{k=1}^{n}X_k\right)\\ &= \frac{1}{n^2} \times n\sum_{k=1}^{n}\mathbb{V}(X_k) \text{ par indépendance}\\ &= \frac{1}{n}\mathbb{V}(X_1) \text{ car les } X_k \text{ suivent la même loi,}\end{aligned}$$

donc

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{V}(X_1)}{n\varepsilon^2} \underset{n\to+\infty}{\longrightarrow} 0,$$

d'où le résultat ! □

N. Laillet
nlaillet.math@gmail.com

Remarque 106

Cette loi des grands nombres est à la base des statistiques. Pour estimer l'espérance d'un phénomène aléatoire, il suffit de répéter un grand nombre de fois ce phénomène et d'en faire la moyenne.
Cette proposition pose la question de la convergence des variables aléatoires :

- convergence dite « en probabilité » : on dit que $(Y_n)_{n\in\mathbb{N}}$ converge en probabilité vers Y si $\forall \varepsilon > 0,\ \mathbb{P}(|Y_n - Y| \geqslant \varepsilon) \underset{n\to+\infty}{\longrightarrow} 0$.
- convergence dite « L^1 » : on dit que $(Y_n)_{n\in\mathbb{N}}$ converge « L^1 » vers Y si $\mathbb{E}(|Y_n - Y|) \underset{n\to+\infty}{\longrightarrow} 0$.

La loi forte des grands nombres existe : $\mathbb{E}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \underset{n\to+\infty}{\longrightarrow} 0$.

5.3 Processus composé – HP

On fait dans cette section un exercice important, qui est formateur.

Proposition 107 (Formule de Wald)

Soit $n \in \mathbb{N}^*$, N une variable aléatoire à valeurs dans $[\![0, n]\!]$, $(X_i)_{i\in[\![1,n]\!]}$ n v.a.i.i.d. et indépendantes de N, toutes définies sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

Soit $Y = \sum_{i=1}^{N} X_i$, c'est-à-dire que pour tout ω dans Ω,

$$Y(\omega) = \sum_{i=1}^{N(\omega)} X_i(\omega).$$

Alors

$$\mathbb{E}(Y) = \mathbb{E}(N)\mathbb{E}(X_1).$$

Remarque 108

Attention ! On ne peut pas écrire

$$\mathbb{E}(Y) = \sum_{i=1}^{N} \mathbb{E}(X_i).$$

En effet, ce qui est à gauche est un réel, ce qui est à droite est une variable aléatoire (à cause du N).

N. Laillet
nlaillet.math@gmail.com

Démonstration. Soit $A = Y(\Omega)$ (l'ensemble des valeurs prises par Y). Alors

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{y\in A} y\mathbb{P}(Y=y) \\
&= \sum_{y\in A} y \sum_{k=0}^{n} \mathbb{P}(Y=y, N=k) \text{ car } (\{N=k\})_{0\leqslant k\leqslant n} \text{ est un système complet d'événements} \\
&= \sum_{y\in A} y \sum_{k=0}^{n} \mathbb{P}\left(\sum_{i=1}^{k} X_i = y, N=k\right) \\
&= \sum_{y\in A} y \sum_{k=0}^{n} \mathbb{P}\left(\sum_{i=1}^{k} X_i = y\right) \mathbb{P}(N=k) \text{ par indépendance.} \\
&= \sum_{k=0}^{n} \sum_{y\in A} y\mathbb{P}\left(\sum_{i=1}^{k} X_i = y\right) \mathbb{P}(N=k) \\
&= \sum_{k=0}^{n} \mathbb{P}(N=k) \sum_{y\in A} y\mathbb{P}\left(\sum_{i=1}^{k} X_i = y\right).
\end{aligned}$$

Or, si $k \in [\![0, n]\!]$, $\sum_{i=1}^{k} X_i$ est à valeurs dans A, donc

$$\begin{aligned}
\sum_{y\in A} y\mathbb{P}\left(\sum_{i=1}^{k} X_i = y\right) &= \mathbb{E}\left(\sum_{i=1}^{k} X_i\right) \\
&= \sum_{i=1}^{k} \mathbb{E}(X_i) \text{ par linéarité} \\
&= k\mathbb{E}(X_1) \text{ car } (X_1, \dots, X_n) \text{ sont iid}
\end{aligned}$$

Donc

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{k=0}^{n} \mathbb{P}(N=k)\, k\mathbb{E}(X_1) \\
&= \mathbb{E}(X_1) \sum_{k=0}^{n} k\mathbb{P}(N=k) \\
&= \mathbb{E}(X_1)\mathbb{E}(N),
\end{aligned}$$

d'où le résultat ! □