Probabilité finie

La théorie des probabilités fournit des modèles mathématiques permettant l'étude d'expériences dont le résultat ne peut être prévu avec certitude.

Expérience	Univers Ω (ensemble des issues)	cardinal
Lancer d'un dé	Un entier $k \in [\![1,6]\!]$	fini
Prélèvement de n objets en sortie d'une chaîne de production dans l'échantillon	Nombre d'objets défectueux, $k \in [\![0,n]\!]$	fini
Questionnaire à 100 questions binaires	Suite ω de 100 réponses $\omega \in \{0,1\}^{100}$	fini
Lancer d'une pièce jusqu'à la première obtention de pile	Un entier $k \in \mathbb{N}$: le temps d'attente du premier succès	dénombrable
Temps d'attente pour une hotline	un temps $\omega \in [0,+\infty[$	continue
Mouvement d'un grain de pollen dans un liquide	Une fonction continue : la trajectoire $f : \mathbb{R}^+ \to \mathbb{R}^2$	continue
Mouvement d'un cours boursier	Une fonction continue : $f : \mathbb{R}^+ \to \mathbb{R}$	continue

Bien que le résultat précis de chacune de ces expériences soit imprévisible, l'observation et l'intuition nous amènent à penser que ces phénomènes obéissent à certaines lois. Par exemple si on jette 6000 fois le dé, on s'attend à ce que le nombre d'apparitions de la face « 3 » soit voisin de 1000. Si on met en service 100 ampoules, leurs durées de vie observées seront concentrées autour d'une certaine valeur moyenne.
La **théorie des probabilités** modélise l'expérience aléatoire dans un cadre formel en quantifiant le sentiment d'incertitude vis-à-vis d'un événement. La **statistique** permet de confronter les modèles probabilistes avec la réalité observée afin de les valider ou de les invalider. Par exemple si quelqu'un a 60 bonnes réponses sur 100 au questionnaire, est-il légitime de considérer qu'il a « mieux fait » que le hasard ?
Dans le cadre de cours, on limite l'univers à un ensemble **fini**, c'est à dire que l'on peut effectuer une énumération finie des issues de l'expérience aléatoire.

I Modélisation probabilistes pour un univers fini

A Axiomes des probabilité. Premières propriétés

Définition Univers

Un **univers**, noté Ω, est l'ensemble de toutes les **issues** (résultats) qui peuvent être obtenues au cours d'une expérience aléatoire. Une issue est noté ω. Dans toute la suite de cours, on suppose que Ω est fini.

Exemple **Lancer de un dé**

Pour l'expérience aléatoire du lancer d'un dé, l'univers est $\Omega = \{1,2,3,4,5,6\}$.

Définition Événement

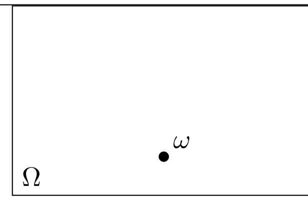
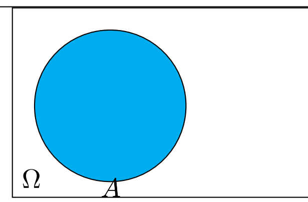
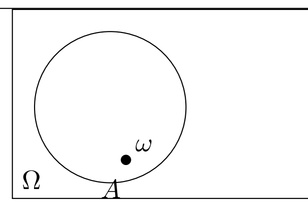
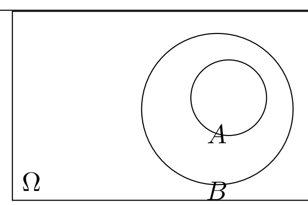
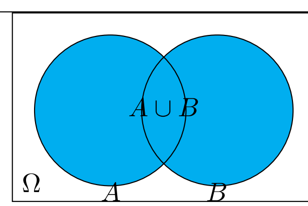
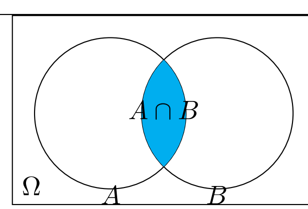
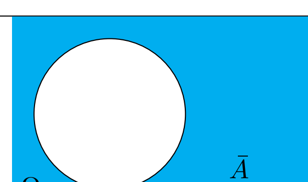
Un **événement** est une partie de l'univers. Un événement est dit **élémentaire** si il contient une unique issue.

Exemple

Pour l'expérience aléatoire du lancer d'un dé, l'événement A="Obtenir un chiffre pair" est l'ensemble $\{2,4,6\}$.
Il n'est pas élémentaire car il est composé de trois issues : 2, 4 et 6.

Définition Notions et opérations

Les notions et opérations que l'on définit sur les événements correspondent aux notions et opérations que l'on définit sur les ensembles.

Notations	Vocabulaire ensembliste	Vocabulaire probabiliste	Diagramme de Venn
$\emptyset$	ensemble vide	événement impossible	
Ω	ensemble plein	événement certain	
ω	élément de Ω	issue	
$\{\omega\}$	singleton	événement élémentaire	
A	sous-ensemble de Ω	événement	
$\omega \in A$	ω appartient à A	Le résultat ω est une des réalisations possibles de A	
$A \subset B$	A est inclus dans B	A implique B	
$A \cup B$	réunion de A et B	A ou B	
$A \cap B$	intersection de A et B	A et B	
			

Définition Tribu

L'ensemble des événements possibles appelé **tribu** est l'ensemble des parties de l'univers. Dans le cadre d'un univers fini, la tribu est l'ensemble des parties de l'univers, soit $\mathcal{P}(\Omega)$.

Exemple

> Pour le lancer d'une pièce, l'univers est $\{P, F\}$ et la tribu est $\{\varnothing, \{P\}, \{F\}, \{P, F\}\}$.

Définition Espace probabilisable

Le couple $(\Omega, \mathcal{P}(\Omega))$ est appelé **espace probabilisable**.

Définition Axiomes des probabilités

Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable fini.
Une **probabilité**, P, sur $(\Omega, \mathcal{P}(\Omega))$ est une application :

$$P \left| \begin{array}{rcl} \mathcal{P}(\Omega) & \longrightarrow & \mathbb{R}^+ \\ A & \longmapsto & P(A) \end{array} \right.$$

vérifiant :

— $P(\Omega) = 1$.

— Pour deux événements disjoints A et B, on a :

$$P(A \cup B) = P(A) + P(B).$$

Le triplet $(\Omega, \mathcal{P}(\Omega), P)$ s'appelle un **espace probabilisé**.

Proposition

1. $P(\emptyset) = 0$.
2. Pour tout $A, B \in \mathcal{P}(\Omega)$ tel que $A \subset B$, on a :

$$P(A) \leq P(B)$$

$$P(B \setminus A) = P(B) - P(A)$$

3. Pour tout $A \in \mathcal{P}(\Omega)$, on a $P(A) \in [0, 1]$.
4. Pour tout $A, B \in \mathcal{P}(\Omega)$, on a

$$P(A \cup B) + P(A \cap B) = P(A) + P(B).$$

5. Pour tout $A, B \in \mathcal{P}(\Omega)$, on a
$$P(A \cap B) \leq P(A) + P(B).$$
6. Si $A_1, \ldots, A_n$ est une famille d'événements deux à deux incompatibles, alors alors

$$P\left(\cup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$

7. Sous additivité : si $A_1, \ldots, A_n$ est une famille d'événements , alors :

$$P\left(\cup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)$$

Démonstration

1. Les deux événements Ω et $\emptyset$ sont disjoints d'où :

$$\begin{aligned} P(\Omega) =& P(\Omega \cup \emptyset) \\ P(\Omega) =& P(\Omega) + P(\emptyset) \\ P(\emptyset) =& 0. \end{aligned}$$

2. Comme $A \subset B$, on a $A \underset{\text{disjoints}}{\bigcup} (B \setminus A) = B$, d'où :

$$P(A) + P(B \setminus A) = P(B)$$

et

$$P(A) \leq P(B)$$

3. Comme $A \subset \Omega$, on a $P(A) \leq P(\Omega) = 1$.
4. Comme $A \cup B = (A \setminus (A \cap B)) \underset{\text{disjoints}}{\bigcup} B$, on a

$$P(A \cup B) = P(A \setminus (A \cap B)) + P(B).$$

Comme $(A \cap B) \subset A$, on a bien :

$$P(A \cup B) = P(A) - P(A \cap B) + P(B).$$

5. Comme $P(A \cup B) + P(A \cap B) = P(A) + P(B)$, on a $P(A \cup B) \leq P(A) + P(B)$.
6. Démonstration par récurrence :
Soit H_n la propriété si $A_1, \ldots, A_n$ est une famille d'événements deux à deux incompatibles, alors

$$P\left(\cup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$

.

— H_1 : on a bien $P(A) = P(A)$.
— $H_n \Rightarrow H_{n+1}$: Soit $A_1, \ldots, A_n, A_{n+1}$ une famille d'événements deux à deux incompatibles. On a :

$$\begin{aligned} P\left(\cup_{i=1}^{n+1} A_i\right) &= P((\cup_{i=1}^n A_i) \cup A_{n+1}) \\ &= P\left(\cup_{i=1}^n A_i\right) + P(A_{n+1}) \text{ car } \cup_{i=1}^n A_i \text{ et } A_{n+1} \text{ sont disjoints} \\ &= \sum_{i=1}^n P(A_i) + P(A_{n+1}) \text{ car } H_n \\ &= \sum_{i=1}^{n+1} P(A_i). \end{aligned}$$

7. Démonstration identique à la précédente en remplaçant les égalités par inégalités.

Définition-Proposition **Probabilité uniforme**

On appelle **probabilité uniforme** sur Ω la probabilité définie par :

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

Démonstration

— $P(\Omega) = \frac{\text{Card}(\Omega)}{\text{Card}(\Omega)} = 1$.

— Soit A et B deux parties disjointes de Ω. On a :

$$P(A\cup B) = \frac{\text{Card}(A\cup B)}{\text{Card}(\Omega)} \overbrace{=}^{\text{A et B sont disjoints}} = \frac{\text{Card}(A)+\text{Card}(B)}{\text{Card}(\Omega)} = \frac{\text{Card}(A)}{\text{Card}(\Omega)}+\frac{\text{Card}(B)}{\text{Card}(\Omega)} = P(A)+P(B).$$

Exemple **Lancer de deux dés**

Pour l'expérience aléatoire du lancer de deux dés, l'univers est $\Omega = \overbrace{\{1,2,3,4,5,6\}}^{\text{issues du premier dé}} \times \overbrace{\{1,2,3,4,5,6\}}^{\text{issues du second dé}}$ et la probabilité est la probabilité uniforme.
L'événement A="Somme des chiffres égale à 2 ou 12" est $\{(1,1),(6,6)\}$.
L'événement B="Somme des chiffres égale à 7" est $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$.
D'où

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)} = \frac{2}{36} = \frac{1}{18},$$

et

$$P(B) = \frac{\text{Card}(B)}{\text{Card}(\Omega)} = \frac{6}{36} = \frac{1}{6}.$$

Proposition **Détermination d'une probabilité par l'image des singletons**

Une probabilité P sur un univers fini est complètement déterminée par les $P(\{\omega\})$ pour tout $\omega\in\Omega$.
$P(\{\omega\})$ est appelé **poids de probabilité**.

Démonstration

En effet, pour $A\subset\Omega$, on a :

$$\begin{aligned} P(A) &= P(\bigcup_{\omega\in A}\{w\}),\\ &= \sum_{\omega\in A} P(\{w\}). \end{aligned}$$

Remarque

— Les poids d'une probabilité P vérifient

$$\sum_{\omega\in\Omega} P(\{\omega\}) = 1.$$

— Il est souvent plus facile de définir une probabilité sur les événements élémentaires ("les issues") que sur l'ensemble des événements.

— Une probabilité sur $\Omega = [\![1,n]\!]$ est la donnée d'une suite finie $(p_1,\dots,p_n)$ de nombres tels que :

1. $p_i = P(\{\omega_i\}), \forall i\in[\![1,n]\!]$
2. $0\leq p_i\leq 1,, \forall i\in[\![1,n]\!]$
3. $p_1+p_2+\ldots+p_n = 1$.

Exemple

Une dé est pipé tel que la chance d'obtenir le chiffre 6 soit 2 fois plus grande que les autres chiffre. Dans ce cas, on a donc :

$$\begin{cases} p_6 = 2p_1 = 2p_2 = 2p_3 = 2p_4 = 2p_5 \\ p_1+p_2+p_3+p_4+p_5+p_6 = 1 \end{cases}$$

Après résolution de ce système d'équations, la probabilité est donc définie par

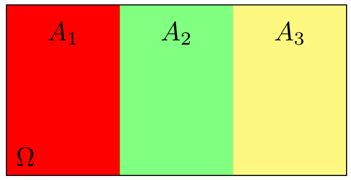
$$p_1 = p_2 = p_3 = p_4 = p_5 = \frac{1}{7} \text{ et } p_6 = \frac{2}{7}.$$

Un système complet d'événements intervient dans la modélisation d'une expérience aléatoire à plusieurs étapes.

Définition **Système complet d'événements**

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé.
Un **système complet d'événements**, appelé aussi **partition**, est une famille $(A_1, \dots, A_n)$ d'événements deux à deux incompatibles tels que

$$\bigcup_{i=1}^{n} A_i = \Omega.$$



(A_1, A_2, A_3) est un système complet d'événements de Ω.

Remarque $A, \bar{A}$ **système complet d'événements**

$(A, \bar{A})$ est un système complet d'événements pour tout événement A.

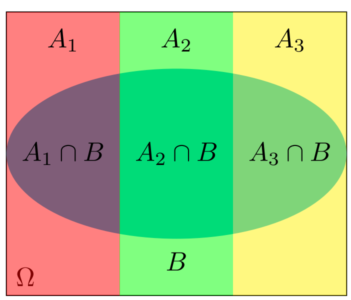
Proposition **Probabilité d'un système complet d'événements**

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé.
Soit $(A_1, \dots, A_n)$ un système complet d'événements.
On a :

$$1 = \sum_{i=1}^{n} P(A_i).$$

Pour tout événement B, on a

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i).$$



Démonstration

On a :

$$P(B) = P(B \cap \Omega)$$

$$P(B) = P(B \cap \left(\bigcup_{i=1}^{n} A_i\right))$$

$$P(B) = P(\bigcup_{i=1}^{n} (B \cap A_i))$$

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i).$$

En particulier si $B = \Omega$, on a :

$$P(\Omega) = \sum_{i=1}^{n} P(\Omega \cap A_i)$$
$$1 = \sum_{i=1}^{n} P(A_i)$$

Remarque

En particulier, pour tout événement A, on a :

$$P(A) + P(\bar{A}) = 1$$

et pour tout événement B, :

$$P(B) = P(B \cap A) + P(B \cap \bar{A})$$

Exemple **3 urnes**

On dispose de 3 urnes U_1, U_2, U_3, chacune contient 10 boules ; parmi elles, U_1 contient 1 blanche, U_2 contient 2 blanches, et U_3 contient 6 blanches. On tire au hasard une urne puis une boule dans cette urne. Quelle est la probabilité de l'événement B="obtenir une blanche" ?
Les événements U_1="tirer l'urne 1, U_2="tirer l'urne 2", U_3="tirer l'urne 3" forment un système complet d'événements. On a donc :

$$P(B) = P(B \cap U_1) + P(B \cap U_2) + P(B \cap U_3).$$

Le but des probabilités conditionnelles est de déterminer les probabilités des termes de cette somme.

B Probabilité conditionnelle

Dans l'expérience aléatoire des trois urnes, la probabilité de tirer une boule blanche est modifiée si l'on dispose de l'information l'urne 1 a été choisie. Le concept de probabilité conditionnelle permet de prendre en compte ce complément d'information.
Par exemple, une classe est constituée de N élèves, dont

1. N_h hommes,
2. N_m majeurs,
3. $N_{h\cap m}$ hommes et majeurs.

Un élève passant au tableau est choisi aléatoirement.
On note H ="élève choisi est un homme" et M ="élève choisi est majeur".
L'équiprobabilité donne :

$$P(M) = \frac{N_m}{N} \text{ et } P(H \cap M) = \frac{N_{h\cap m}}{N}.$$

Quelle est la probabilité que l'élève choisie soit un homme sachant qu'il doit être majeur, noté $P_M(H)$?
Dans cette expérience aléatoire, l'information supplémentaire est l'élève choisi est majeur. L'univers est maintenant l'ensemble des élèves majeurs donc la probabilité est :

$$P_M(H) = \frac{N_{h\cap m}}{N_m}.$$

Mais,

$$P_M(H) = \frac{N_{h\cap m}}{N_m} = \frac{\frac{N_{h\cap m}}{N}}{\frac{N_m}{N}} = \frac{P(H \cap M)}{P(M)}.$$

Par analogie, la définition formelle est :

Définition **Probabilité conditionnelle**

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé. Soit A un événement tel que $P(A) > 0$.
L'application

$$P_A \left| \begin{array}{rcl} \mathcal{P}(\Omega) & \longrightarrow & [0,1] \\ B & \longmapsto & \dfrac{P(A \cap B)}{P(A)} \end{array} \right.$$

est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ appelée **probabilité conditionnellement à A**, ou **probabilité sachant A**.
On note $P(B|A) = P_A(B)$.

Exemple

On lance un dé parfaitement équilibré. La probabilité d'obtenir un 6 est 1/6. On suppose maintenant que ce dé a ses faces impaires peintes en vert, et ses faces paires peintes en bleu. On a aperçu de loin que, sur le dessus du dé, on a obtenu une face bleue. Quelle est la probabilité d'obtenir 6 sachant une face bleue ?
On note $B = \{2, 4, 6\}$ l'événement "face bleue". On a :

$$P_B(\{6\}) = \frac{P(\{6\} \cap B)}{P(B)} = \frac{\frac{\mathrm{Card}(\{6\})}{\mathrm{Card}(\Omega)}}{\frac{\mathrm{Card}(\{2,4,6\})}{\mathrm{Card}(\Omega)}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.$$

Proposition **Formule des probabilités totales**

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé.
Soit $(A_1, \ldots, A_n)$ un système complet d'événements de probabilités non nulles.
Pour tout événement B, on a

$$P(B) = \sum_{i=1}^{n} P_{A_i}(B) P(A_i).$$

Démonstration

Dans l'équation d'un système complet d'événements de la proposition **??**, il suffit de remplacer $P(B \cap A_i)$ pat $P(B|A_i)P(A_i)$.

Exemple **3 urnes**

La probabilité de tirer une boule blanche était de :

$$P(B) = P(B \cap U_1) + P(B \cap U_2) + P(B \cap U_3).$$

Ce qui donne :

$$P(B) = P_{U_1}(B)P(U_1) + P_{U_1}(B)P(U_2) + P_{U_3}(B)P(U_3) = \frac{1}{10}\frac{1}{3} + \frac{2}{10}\frac{1}{3} + \frac{3}{10}\frac{1}{3} = \frac{6}{30} = \frac{1}{5}.$$

Remarque

En particulier, si A et $\bar{A}$ sont deux événements de probabilité non nulle, on a pour tout événement B :

$$P(B) = P_A(B)P(A) + P_{\bar{A}}(B)P(\bar{A}).$$

Proposition **Formule de Bayes**

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé.
Soit A et B deux événements de probabilités non nulles.
Alors

$$P_B(A) = \frac{P_A(B)P(A)}{P(B)}.$$

Cette formule permet d'inverser des conditions.

Démonstration

On a $P(A \cap B) = P(B)P_B(A)$ et $P(A \cap B) = P(B \cap A) = P(A)P_A(B)$. Comme $P(B) \neq 0$, on a le résultat voulu.

Proposition **Formule de Bayes usuelle**

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé.
Soit $(A_1, \dots, A_n)$ un système complet d'événements de probabilités non nulles.
Pour tout événement B de probabilité non nulle, et on a

$$\forall k \in [\![1, n]\!] : \quad P_B(A_k) = \frac{P_{A_k}(B)P(A_k)}{\sum_{i=1}^n P_{A_i}(B)P(A_i)}.$$

Remarque

Le système $(A_1, \dots, A_n)$ représente souvent une liste de causes pouvant amener l'événement B lors d'une étape suivante de l'expérience par exemple dans le problème des 3 urnes. Il est alors généralement facile de déterminer la probabilité qu'une certaine conséquence B ait lieu, sachant que la cause A_i a eu lieu, c'est-à-dire la probabilité conditionnelle $P_{A_i}(B)$ en respectant l'ordre temporel. Ces données permettent, grâce à la formule de Bayes, de remonter le temps, en déterminant la probabilité qu'une certaine cause A_i ait eu lieu sachant la conséquence B. Pour cette raison, cette formule est aussi souvent appelée formule de probabilité des causes.

Exemple **3 urnes**

On cherche à connaître la probabilité que l'urne 1 est été choisie sachant que l'on a tiré une boule blanche, c'est à dire $P_B(U_1)$.
On a donc

$$P_B(U_1) = \frac{P_{U_1}(B)P(U_1)}{P_{U_1}(B)P(U_1) + P_{U_2}(B)P(U_2) + P_{U_3}(B)P(U_3)} = \frac{\frac{1}{3}\frac{1}{10}}{\frac{1}{3}\frac{1}{10} + \frac{1}{3}\frac{2}{10} + \frac{1}{3}\frac{6}{10}} = \frac{1}{9}.$$

Exemple **Test de dépistage**

Vous êtes directeur de cabinet du ministre de la santé. Une maladie est présente dans la population, dans la proportion d'une personne malade sur 10000. Un responsable d'un grand laboratoire pharmaceutique vient vous vanter son nouveau test de dépistage : si une personne est malade, le test est positif à 99%. Si une personne n'est pas malade, le test est positif à 0,1%. Ces chiffres ont l'air excellent, vous ne pouvez qu'en convenir. Toutefois, avant d'autoriser la commercialisation de ce test, vous faites appel au statisticien du ministère : ce qui vous intéresse, ce n'est pas vraiment les résultats présentés par le laboratoire, c'est la probabilité qu'une personne soit malade si le test est positif. La formule de Bayes permet de calculer cette probabilité.
On note M l'événement : "La personne est malade", et T l'événement : "Le test est positif". Le but est de calculer $P_T(M)$. Les données que vous avez en main sont $P(M) = 0,0001$ (et donc $P(\bar{M}) = 0,9999$), $P_M(T) = 0,99$ et $P_{\bar{M}}(T) = 0,001$. La formule de Bayes donne :

$$P_T(\bar{M}) = \frac{P_{\bar{M}}(T)P(\bar{M})}{P_M(T)P(M) + P_{\bar{M}}(T)P(\bar{M})} = \frac{0,999910^{-3}}{10^{-4}0,99 + 0,999910^{-3}} \approx 0,91.$$

C'est catastrophique ! Il y a que 91% de chances qu'une personne positive au test ne soit pas malade ! C'est tout le problème des tests de dépistage pour des maladies rares : ils doivent être excessivement performants, sous peine de donner beaucoup trop de "faux-positifs".

Définition **Arbre de probabilité**

Un **arbre de probabilité** est un schéma permettant de résumer une expérience aléatoire connaissant des probabilités conditionnelles, soit un graphe orienté et pondéré obéissant aux règles suivantes :

1. la somme des probabilités des branches issues d'un même sommet donne 1,
2. la probabilité d'un chemin est le produit des probabilités des branches qui le composent,
3. la probabilités de la branche allant du sommet A vers le sommet B est la probabilité conditionnelle de B sachant que A est déjà réalisé $P_A(B)$.

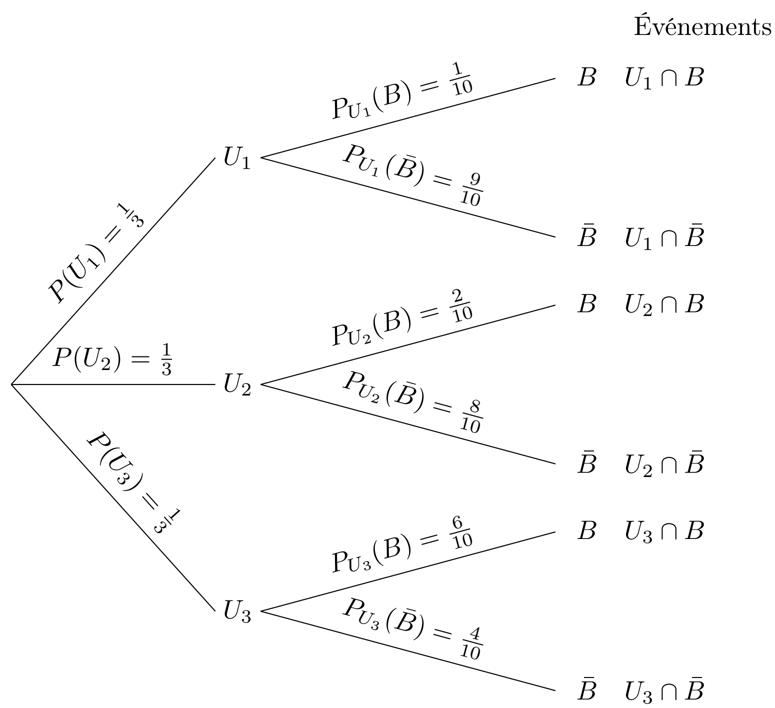
On retrouve alors la propriété de la probabilité conditionnelle :

$$P(A \cap B) = P_A(B)P(A).$$

Ainsi que la formule des probabilités totales, si $A_1, A_2, \ldots, A_n$ définit un système complet d'évènements de probabilité non nulles, on a :

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P_{A_i}(B)P(A_i).$$

Exemple **3 urnes**



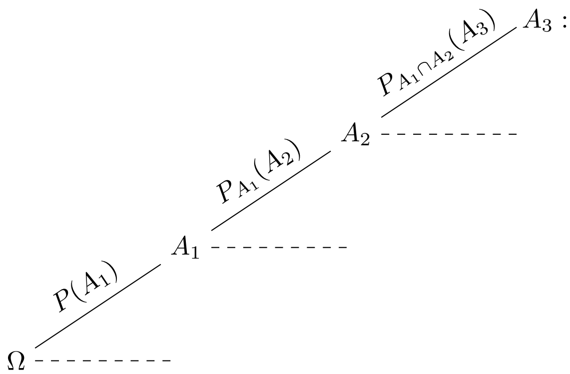
La probabilité de l'événement B est la somme des probabilités des parcours qui mènent à B, d'où

$$P(B) = P_{U_1}(B)P(U_1) + P_{U_1}(B)P(U_2) + P_{U_1}(B)P(U_3).$$

Proposition Formule des probabilités composées

Soit $A_1, A_2, \dots, A_n$ n événements vérifiant $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \neq 0$.
Alors

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)\dots P_{A_1\cap A_2 \cap \dots \cap A_{n-1}}(A_n).$$



$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)$$

Cas pour n=3

Démonstration

Démontrons par récurrence la propriété (H_n) pour tout $n \in \mathbb{N}^*$:

Si $A_1, A_2, \dots, A_n$ n événements vérifiant $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \neq 0$,
Alors $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)\dots P_{A_1\cap A_2 \cap \dots \cap A_{n-1}}(A_n)$.

— *Initialisation :* Soit A_1 un événement. On a bien $P(A_1) = P(A_1)$.

— *Hérédité :* Soit $n \in \mathbb{N}^*$. Supposons la propriété (H_n) vraie au rang n et vérifions au rang $n+1$.
Soit $A_1, A_2, \dots, A_{n+1}$ $n+1$ événements vérifiant $P(A_1 \cap A_2 \cap \dots \cap A_n) \neq 0$.
Comme $P(A_1 \cap A_2 \cap \dots \cap A_n) \neq 0$, on a :

$$P(A_1 \cap A_2 \cap \dots \cap A_n \cap A_{n+1}) = P(A_1 \cap A_2 \cap \dots \cap A_n)P_{A_1\cap A_2 \cap \dots \cap A_n}(A_n + 1).$$

Comme $A_1 \cap A_2 \cap \dots \cap A_n \subset A_1 \cap A_2 \cap \dots \cap A_{n-1}$, on a $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \neq 0$. On applique l'hypothèse de récurrence à $P(A_1 \cap A_2 \cap \dots \cap A_n)$, d'où

$$P(A_1 \cap A_2 \cap \dots \cap A_n \cap A_{n+1}) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)\dots P_{A_1\cap A_2 \cap \dots \cap A_{n-1}}(A_n)P_{A_1\cap A_2 \cap \dots \cap A_n}(A_n+1).$$

La propriété (H_n) est donc vraie pour tout $n \in \mathbb{N}^*$.

Exemple

On considère une urne contenant 4 boules blanches et 3 boules noires. On tire une à une et sans remise 3 boules de l'urne. Quelle est la probabilité pour que la première boule tirée soit blanche, la seconde blanche et la troisième noire ?
On note B_i l'événement : "La i-ème boule tirée est blanche (resp. noire)". On cherche à calculer $P(B_1 \cap B_2 \cap \bar{B}_3)$, ce que l'on va faire en utilisant la formule des probabilités composées :

$$P(B_1 \cap B_2 \cap \bar{B}_3) = P(B_1)P(B_2|B_1)P(\bar{B}_3|B_1 \cap B_2).$$

Chacune des probabilités qui apparaît est facile à calculer, car $P(B_1) = 4/7$, $P(B_2|B_1) = 3/6$ (il reste 6 boules dont 3 blanches) et $P(\bar{B}_3|B_1 \cap B_2) = 3/5$. Finalement, on obtient $P(B_1 \cap B_2 \cap \bar{B}_3) = 6/35$.

C Indépendance

On considère l'expérience aléatoire de lancer successivement deux fois un dé.
Soit A un événement relié au premier lancer, par exemple obtenir un chiffre pair sur le premier dé et soit B un événement relié au second lancer, par exemple obtenir un chiffre impair sur le second dé. L'information de savoir que l'événement A est réalisé, ne modifie pas la probabilité de l'événement B, autrement dit $P_A(B) = P(B)$. Mais,

d'après la définition de la probabilité conditionnelle, on a : $P_A(B) = \frac{P(A\cap B)}{P(A)}$. De ces deux égalités, on obtient $P(A \cap B) = P(A)P(B)$ et on dit que les deux événements sont indépendants.
Par analogie, la définition formelle est :

Définition Indépendance

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

- Deux événements A et B sont **indépendants** si $P(A \cap B) = P(A)P(B)$.
- Une famille $(A_1, \ldots, A_n)$ d'événements sont **(mutuellement) indépendants** si pour toute partie $I \subset [\![1, n]\!]$, on a

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

Exemple **Lancer de deux dés**

A="Obtenir 5 ou 6 sur le premier dé"
$A = \{(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$.
B="Obtenir 6 sur le second dé"
$B = \{(1,6),(2,6),(3,6),(4,6),(5,6),(6,6)\}$
$A \cap B$ ="Obtenir 5 ou 6 sur le premier dé et obtenir 6 sur le second dé".
$A \cap B = \{(5,6),(5,6),(6,6)\}$. On a $P(A) = \frac{12}{36} = \frac{1}{3}, P(B) = \frac{6}{36} = \frac{1}{6}$ et $P(A \cap B) = \frac{3}{36} = \frac{1}{12} = P(A)P(B)$.
Donc les événements A et B sont indépendants.

Remarque

Trois événements peuvent être indépendants deux à deux, sans pour autant être mutuellement indépendants.
Par exemple, on considère un dé équilibré à 4 faces. On a $\Omega = \{1,2,3,4\}$ avec $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$.
Soit les événements $A = \{1,2\}, B = \{1,3\}$ et $C = \{1,4\}$.
On a :

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

et

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

Les événements A, B et C sont deux à deux indépendants mais :

$$P(A \cap B \cap C) = P(\{1\}) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C).$$

Donc les événements ne sont pas mutuellement indépendants.

Proposition Lien avec la probabilité conditionnelle

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé.
Soit A un événement de probabilité non nulle et B un événement quelconque.
Alors les événements A et B sont indépendants si et seulement si $P_A(B) = P(B)$.

Démonstration

Soit A et B deux événements tel que $P(A) \geq 0$.

$$\begin{aligned} A, B \text{ sont indépendants } &\Leftrightarrow P(B \cap A) = P(B)P(A) \\ &\Leftrightarrow \frac{P(B \cap A)}{P(A)} = P(B) \text{ car } P(A) \neq 0 \\ &\Leftrightarrow P_A(B) = P(B). \end{aligned}$$

II Variable aléatoire

Exemple **Jeu**

Pour attirer les clients, un casino propose un nouveau jeu : le croupier lance simultanément 2 dés et calcule leur somme,

— si la somme est égale à 2 ou 12, le joueur gagne 2 euros,

— si la somme est égale à 7, le casino gagne 1 euro,

— dans les autres cas, c'est nul (le joueur gagne 0 euro).

A votre avis, ce jeu est-il favorable au joueur ou au casino ? Quel est le gain moyen du joueur ?
Dans ce jeu, on fait intervenir le hasard en observant la somme des points marqués par deux dés. Considérons le jet d'un dé bleu et d'un dé rouge et notons S la somme des points obtenus. On modélise cette expérience en prenant l'équiprobabilité sur l'univers

$$\Omega = \overbrace{\{1,2,3,4,5,6\}}^{\text{issues du dé bleu}} \times \overbrace{\{1,2,3,4,5,6\}}^{\text{issues du dé rouge}}.$$

Une issue, ω, est un couple (b, r) où b désigne le chiffre du dé bleu et r celui du rouge. La somme S est l'application :

$$S \left| \begin{array}{ccc} \Omega & \longrightarrow & \{2,3,\dots,12\} \\ (b,r) & \longmapsto & b+r \end{array} \right.$$

Cette application est représentée par ce tableau :

$b\backslash r$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
2	5	6	7	8	9	10
1	6	7	8	9	10	11
2	7	8	9	10	11	12

On dit que S est une variable aléatoire sur $\{2, 3, \dots, 12\}$.
En fait, l'observation qui nous intéresse dans cette expérience, ce n'est pas ω, mais seulement $S(\omega)$. On aimerait connaître la probabilité que la somme des points prenne une valeur donnée, soit $P(S = k)$ pour k entier fixé entre 2 et 12.
La méthode est de déterminer l'ensemble des issues donnant $(S = k)$ en lisant le tableau ci-dessous puis de calculer la probabilité de cet événement. Par exemple, on a $(S = 3) = \{(1, 2), (2, 1)\}$, puis

$$P(S=3) = P(\{(1,2),(2,1)\}) \overset{\text{équiprobabilité}}{\overbrace{=}} \frac{\mathrm{Card}(\{(1,2),(2,1)\})}{\mathrm{Card}(\Omega)} = \frac{2}{36}.$$

On obtient ainsi :

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

Cela revient à considérer un nouvel univers :

$$\Omega' = \{2,3,4,5,6,7,8,9,10,11,12\}$$

et à munir cet ensemble de la probabilité P_S définie par le tableau des $P(S = k)$. Cette nouvelle probabilité s'appelle loi de la variable aléatoire S.

A Définition

Définition Variable aléatoire

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé fini.
Une **variable aléatoire réelle** est une application :

$$X \left| \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & X(\omega) \end{array} \right. .$$

On note $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ l'ensemble des images de la variable aléatoire.
Comme Ω est fini, $X(\Omega)$ est aussi finie. On s'autorise désormais à écrire l'ensemble des images sous la forme $X(\Omega) = \{x_1, x_2, \dots, x_n\}$.

Remarque

- Malgré son nom, une variable aléatoire n'est pas une variable (c'est une fonction) et elle n'est pas aléatoire.
- Il est toujours possible de quantifier les valeurs prises par une variable aléatoire. Par exemple si X représente le sexe de l'enfant à la naissance, on posera le label 1 pour le sexe féminin et le label 0 pour le sexe masculin et dans ce cas $X(\Omega) = \{0, 1\}$.
- Dans ce cours l'espace de probabilité est finie et la variable aléatoire réelle, on écrira "Soit X une variable aléatoire" au lieu de "Soit X une variable aléatoire réelle sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), P)$."

Exemple **Variable aléatoire indicatrice de A**

Soit A un événement de Ω.
La variable aléatoire indicatrice de A, noté 1_A, est une fonction explicitant l'appartenance ou non à un événement A de toute issue de Ω :

$$1_A \left| \begin{array}{ccc} \Omega & \longrightarrow & \{0, 1\} \\ \omega & \longmapsto & \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases} \end{array} \right. .$$

Définition Événements associés à une variable aléatoire

Soit X une variable aléatoire.
Pour tout $A \subset \mathbb{R}$, on définit l'événement $\{X \in A\}$ comme étant

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}.$$

Dans le cadre de ce chapitre, nous aurons recours aux notations suivantes, pour $x \in \mathbb{R}$ et $A \subset \mathbb{R}$:

$$\begin{aligned}
(X \in A) &= X^{-1}(A) = \{\omega \in \Omega | X(\omega) \in A\} \\
(X = x) &= X^{-1}(\{x\}) = \{\omega \in \Omega | X(\omega) = x\} \\
(X \leq x) &= X^{-1}(]-\infty, x]) = \{\omega \in \Omega | X(\omega) \leq x\} \\
(X < x) &= X^{-1}(]-\infty, x[) = \{\omega \in \Omega | X(\omega) < x\} \\
(X \geq x) &= X^{-1}([x, +\infty[) = \{\omega \in \Omega | X(\omega) \geq x\} \\
(X > x) &= X^{-1}(]x, +\infty[) = \{\omega \in \Omega | X(\omega) > x\}
\end{aligned}$$

Remarque

Dans les démonstrations, sans perte de généralité, on considérera uniquement les parties de $X(\Omega)$ et non de $\mathbb{R}$ pour les rendre plus lisibles.

Exemple **Jeu**

Par exemple, l'ensemble des issues donnant une somme à 7 est :

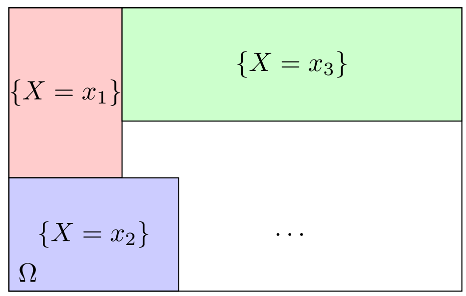
$$\{S = 7\} = S^{-1}(\{7\}) = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}.$$

et l'ensemble des issues donnant une somme à 2 ou 12 est :

$$\{S \in \{2,12\}\} = S^{-1}(\{2,12\}) = \{(1,1),(6,6)\}.$$

Proposition

Soit X une variable aléatoire.
L'ensemble $(X = x_i)_{1 \leq i \leq n}$ forme un système complet d'événements de Ω.



Démonstration

- *disjoints :* Soit $i \neq j$. $\{i\}$ et $\{j\}$ sont deux ensembles disjoints. Donc $X^{-1}(\{i\})$ et $X^{-1}(\{j\})$ sont aussi disjoints.
- $\cup_{i=1}^{n}(X = x_i) = \cup_{i=1}^{n} X^{-1}(\{x_i\}) \overbrace{=}^{\text{Formule de Hausdorff}} X^{-1}(\cup_{i=1}^{n}\{x_i\}) = X^{-1}(X(\Omega)) = \Omega.$

B Loi de probabilité

Définition-Proposition **Loi d'une variable aléatoire**

Soit X une variable aléatoire.
L'application

$$P_X \left| \begin{array}{rcl} \mathcal{P}(X(\Omega)) & \longrightarrow & \mathbb{R}^+ \\ A & \longmapsto & P(X^{-1}(A)) \end{array} \right.$$

est une probabilité sur $X(\Omega)$, appelée **loi de la variable X**.

Démonstration

- $P_X(X(\Omega)) = P(X^{-1}(X(\Omega))) = P(\Omega) = 1$
- Soit A et B deux événements incompatibles de $X(\Omega)$. Ainsi $X^{-1}(A)$ et $X^{-1}(B)$ sont deux événements incompatibles de Ω. On a

$$P_X(A \cup B) = P(X^{-1}(A \cup B)) = P(X^{-1}(A) \cup X^{-1}(B))$$

Comme $X^{-1}(A)$ et $X^{-1}(B)$ incompatibles, on a :

$$P_X(A \cup B) = P(X^{-1}(A)) + P(X^{-1}(B)) = P_X(A) + P_X(B).$$

Proposition **Déterminer une loi**

Déterminer **la loi de probabilité P_X** de la variable aléatoire X, c'est donner

1. l'ensemble $X(\Omega) = \{x_1, \dots, x_n\}$ des valeurs prises par X,
2. pour chaque x_i de $X(\Omega)$, la probabilité $p_i = P(X = x_i)$.

:

Démonstration

Soit $A \in \mathcal{P}(X(\Omega))$. Il existe $I \subset [\![1,n]\!]$ tel que $A = \{x_i : i \in I\}$.
On a

$$P(X \in A) = P(\bigcup_{i \in I}(X = x_i)) \overbrace{=}^{(\{X=x_i\})_{i\in I} \text{ disjoints}} \sum_{i \in I} P(X = x_i).$$

Remarque

- $\forall i \in [\![1,n]\!]$: $\quad p_i = \sum_{\omega \in \Omega \text{ tel que } X(\omega)=x_i} P(\{\omega\})$,
- Comme les événements $X = x_1, X = x_2, \dots, X = x_n$ forme un système complet d'événements de Ω, on a :

$$\sum_{x_i \in X(\Omega)} P(X = x_i) = \sum_{i=1}^{n} p_i = 1.$$

Exemple **Jeu**

La variable aléatoire somme $S \left| \begin{array}{rcl} \Omega & \longrightarrow & \{2,3,\dots,12\} \\ (b,r) & \longmapsto & b+r \end{array} \right.$ est déterminée par ce tableau :

$b\backslash r$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
2	5	6	7	8	9	10
1	6	7	8	9	10	11
2	7	8	9	10	11	12

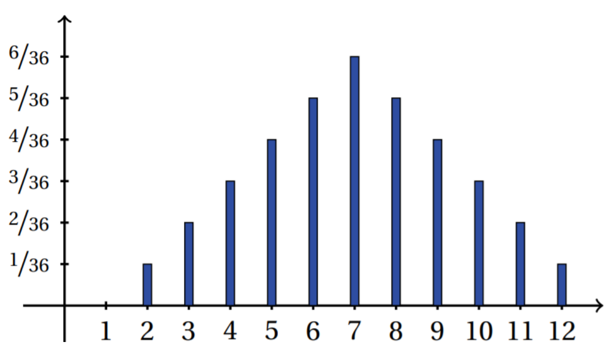
Du fait de l'équiprobabilité, on détermine la loi de probabilité S en calculant le nombre de cases. Par exemple,

$$P(S = 4) = P(\{(1,3),(2,2),(3,1)\}) = \frac{\text{Card}(\{(1,3),(2,2),(3,1)\})}{\text{Card}(\{1,\dots,6\}^2)} = \frac{3}{36}.$$

On obtient :

k	2	3	4	5	6	7	8	9	10	11	12
$p_k = P(S = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Sa représentation graphique est :



Théorème **Construire un espace de probabilité à partir d'une loi**

Soit $(x_1,\dots,x_n)$ une famille finie de réels et $(p_1,\dots,p_n)$ une famille de réels positifs telle que $\sum_{i=1} p_i = 1$. Alors il existe un espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ et une variable aléatoire X sur cette espace et à valeurs dans $\{x_1,\dots,x_n\}$ telle que :

$$\forall i \in [\![1,n]\!] : \quad P(X = x_i) = p_i.$$

Démonstration

Soit $\Omega = \{x_1, \dots, x_n\}$ et P la probabilité définie sur $(\Omega, \mathcal{P}(\Omega))$ par :

$$\forall i \in [\![1, n]\!] : \quad P(\{x_i\}) = p_i.$$

Soit X l'application identité. On vérifie que $P(X = x_i) = P(\{x_i\}) = p_i$, pour tout $i \in [\![1, n]\!]$.

Remarque

Ce théorème permet de définir une variable aléatoire par sa loi de probabilité sans avoir à étudier l'expérience aléatoire sous-jacente, c'est à dire définir l'espace de probabilité $(\Omega, \mathcal{P}(\Omega), P)$ et la fonction de la variable aléatoire. Par exemple, on modélise l'expérience aléatoire d'un lancer de dé en posant : la variable aléatoire X représente le chiffre du dé avec $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$.

Définition-Proposition **Fonction d'une variable aléatoire**

Soit X une variable aléatoire.
Soit $f : \mathbb{R} \to \mathbb{R}$ une application quelconque.
L'application $f \circ X$

$$f \circ X \left| \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & f(X(\omega)) \end{array} \right.$$

est une variable aléatoire réelle.
L'usage veut qu'on la note abusivement $f(X)$ au lieu de $f \circ X$.

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[rr, bend right, "f \circ X"'] & \mathbb{R} \arrow[r, "f"] & \mathbb{R} \end{tikzcd}$$

On a

$$\forall y \in f(X)(\Omega) : \quad P(f(X) = y) = \sum_{x \in X(\Omega) \text{ tel que } f(x) = y} P(X = x) = \sum_{i \in [\![1,n]\!] \text{ tel que } f(x_i) = y} P(X = x_i).$$

Exemple **Jeu**

Le gain obtenu est fonction de la somme obtenue avec les deux dés. La modélisation est d'appliquer une fonction à la somme, s, des dés :

$$G \left| \begin{array}{ccc} [\![2, 12]\!] & \longrightarrow & \{-1, 0, 2\} \\ s & \longmapsto & \begin{cases} 2 & \text{si } s = 2 \text{ ou } 12 \\ -1 & \text{si } s = 7 \\ 0 & \text{sinon} \end{cases} \end{array} \right. .$$

La variable aléatoire Gain est $G \circ X$.
Vérifions sur un exemple la modélisation. Si le jet des dés donne $\omega = (1, 6)$, le gain de -1. On a bien $G \circ S(\omega = (1, 6)) = G(S(1, 6)) = G(1 + 6) = G(7) = -1$.
La loi de probabilité S est déterminé par :

k	-1	2	0
$P(G(S) = k)$	$\frac{6}{36} = P(S = 7)$	$\frac{2}{36} = P(S = 2) + P(S = 12)$	$\frac{28}{36} = 1 - P(G(S) = -1) + P(G(S) = 2)$

C Vecteurs aléatoires

Définition-Proposition Loi conjointe

Soit X, Y deux variables aléatoires.
On note (X, Y) le **couple de variables aléatoires** prenant ses valeurs dans $\mathbb{R}^2$.
La **loi conjointe** du couple (X, Y) est déterminé par :

1. $X(\Omega) \times Y(\Omega) = \{x_1, \dots, x_n\} \times \{y_1, \dots, y_m\}$, les valeurs prises par le couple
2. $\forall (i, j) \in [\![1, n]\!] \times [\![1, m]\!] : \quad P(X = x_i \cap Y = y_j)$.

On note $p_{i,j} = P(X = x_i \cap Y = y_j)$.
Les événements $((X = x_i) \cap (Y = y_j))_{\forall (i,j) \in [\![1,n]\!] \times [\![1,m]\!]}$ forment un système complet d'événements de Ω. En particulier, on a :

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P(X = x_i \cap Y = y_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{i,j} = 1.$$

Démonstration

Démontrons que la loi conjointe est déterminée par

$$\forall (i, j) \in [\![1, n]\!] \times [\![1, m]\!] : \quad P(X = x_i \cap Y = y_j).$$

Soit $(A, B) \subset X(\Omega) \times Y(\Omega)$. Il existe $I \subset [\![1, n]\!]$ et $J \subset [\![1, m]\!]$ tel que $A = \{x_i : i \in I\}$ et $B = \{y_j : j \in J\}$.
On a :

$$P((X, Y) \in (A, B)) = P(X \in A \cap Y \in B) = P\left(\left(\bigcup_{i \in I} (X = x_i)\right) \cap \left(\bigcup_{j \in J} (Y = y_j)\right)\right) = P(\bigcup_{i \in I, j \in J} (X = x_i) \cap (Y = y_j))$$

. Comme $(\{(X = x_i) \cap (Y = y_j)\})_{i \in I, j \in J}$ est une famille d'ensembles disjoints. On a

$$P((X, Y) \in (A, B)) = P(X \in A \cap Y \in B) = \sum_{i \in I, j \in J} P(X = x_i \cap Y = y_j).$$

Définition Lois marginales

Soit (X, Y) un couple de variables aléatoires.
La loi de X appelé **première loi marginale** et la loi de Y **second loi marginale** du couple.

Proposition Relations

Soit (X, Y) un couple de variables aléatoires.
On a :

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

et

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

$X \backslash Y$	y_1	$\dots$	y_j	$\dots$	y_m	$P(X = x_i)$	
x_1	$p_{1,1}$	$\dots$	$p_{1,j}$	$\dots$	$p_{1,m}$	$P(X = x_1)$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	
x_i	$p_{i,1}$	$\dots$	$p_{i,j}$	$\dots$	$p_{i,m}$	$P(X = x_i)$	$\leftarrow \sum_{k=1}^{m} p_{i,k}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	
x_n	$p_{n,1}$	$\dots$	$p_{n,j}$	$\dots$	$p_{n,m}$	$P(X = x_n)$	
$P(Y = y_j)$	$P(Y = y_1)$	$\dots$	$P(Y = y_j)$	$\dots$	$P(Y = y_m)$	1	
			$\uparrow$ $\sum_{k=1}^{n} p_{k,j}$				

Démonstration

Comme les événements $((Y = y_j))_{\forall j \in [\![1,m]\!]}$ forment un système complet d'événements de Ω, d'après la proposition **??** on a

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

.

Exemple

On tire deux nombres au hasard dans $\{-1, 1\}$. On note X leur somme, et Y leur produit. On cherche à déterminer la loi conjointe de (X, Y).
L'espace de probabilité est est $(\{-1; 1\}^2, \mathcal{P}(\{-1; 1\}^2), P)$ avec P la probabilité uniforme. Pour tout $\omega = (r, b) \in \Omega$, on a $X(r, b) = r + b$ et $Y(r, b) = r.b$. Sous forme de tableau, on a :

$r\backslash b$	-1	1
-1	-2	0
1	0	2

Définition de X

$r\backslash b$	-1	1
-1	1	-1
1	-1	1

Définition de Y

X prend ses valeurs dans $\{-2, 0, 2\}$ et Y dans $\{-1, 1\}$.
Les variables aléatoires étant finies, il suffit de déterminer toutes les probabilités $P(X = x \text{ et } Y = y)$ pour tout couple $(x, y) \in \{-2, 0, 2\} \times \{-1, 1\}$. On a :

- $P(X = 2, Y = 1) = P(\{(1, 1)\}) \overbrace{=}^{\text{equiprobabilité}} \frac{1}{4}$.
- $P(X = 2, Y = -1) = P(\emptyset) = 0$.
- $P(X = 0, Y = 1) = P(\emptyset) = 0$.
- $P(X = 0, Y = -1) = P(\{(1, -1), (-1, 1)\}) = \frac{2}{4} = \frac{1}{2}$.
- $P(X = -2, Y = 1) = P(\{(-1, -1)\}) = \frac{1}{4}$.
- $P(X = -2, Y = -1) = P(\emptyset) = 0$.

La loi de probabilité conjointe (X, Y) est donc déterminée par :

$X\backslash Y$	-1	1	$P(X = x)$
-2	0	1/4	1/4
0	1/2	0	1/2
2	0	1/4	1/4
$P(Y = y)$	1/2	1/2	1

Exemple

Dans une classe, la répartition en fonction de l'age et du genre est :

Age\ Genre	Fille	Garçon	Total
18	5	10	15
19	2	6	8
20	0	1	1
Total	7	17	24

On tire au hasard un élève dans la classe.
X représente l'age de l'élève.
Y représente le genre de l'élève avec la label 0 pour une fille et le label 1 pour un garçon.
La loi de probabilité conjointe (X, Y) est déterminée par :

$X\backslash Y$	0	1	$P(X=x)$
18	$5/24$	$10/24$	$15/24$
19	$2/24$	$6/24$	$8/24$
20	$0/24$	$1/24$	$1/24$
$P(Y=y)$	$7/24$	$17/24$	1

D Indépendance et lois conditionnelles

Définition-Proposition **Lois conditionnelles**

Soit (X,Y) un couple de variables aléatoires.

1. On appelle **loi conditionnelle** de X sachant $(Y=y_j)$ pour $P(Y=y_j)\neq 0$, la probabilité définie par :
$$\forall i\in [\![1,n]\!]:\quad P_{Y=y_j}(X=x_i)=\frac{P(X=x_i\cap Y=y_j)}{P(Y=y_i)}$$
2. On appelle **loi conditionnelle** de Y sachant $(X=x_i)$ pour $P(X=x_i)\neq 0$, la probabilité définie par :
$$\forall j\in [\![1,m]\!]:\quad P_{X=x_i}(Y=y_j)=\frac{P(X=x_i\cap Y=y_j)}{P(X=x_i)}$$

Remarque

Les lois marginales et les lois conditionnelles déterminent la loi conjointe :
$$\forall (i,j)\in [\![1,n]\!]\times[\![1,m]\!]:\quad P(X=x_i\cap Y=y_j)=P_{X=x_i}(Y=y_j)P(X=x_i)$$
$$\forall (i,j)\in [\![1,n]\!]\times[\![1,m]\!]:\quad P(X=x_i\cap Y=y_j)=P_{Y=y_j}(X=x_i)P(Y=y_j)$$

Définition **Indépendance**

Soit (X,Y) un couple de variables aléatoires.
On dit que X et Y sont **indépendantes** si
$$\forall (i,j)\in [\![1,n]\!]\times[\![1,m]\!]:\quad P((X=x_i)\cap(Y=y_j))=P(X=x_i)P(Y=y_j).$$

Remarque

Dans la modélisation d'une expérience aléatoire, l'hypothèse d'indépendance des variables aléatoires est souvent une donnée de l'expérience et non pas une propriété à vérifier. Par exemple, la modélisation de l'expérience du jeter de deux dés serait :

— X_1, la variable aléatoire représentant le chiffre du premier dé,
— X_2, la variable aléatoire représentant le chiffre du second dé,
— X_1 et X_2, supposées indépendantes,
— $S=X_1+X_2$, la somme des deux dés.

Théorème

Soit (X,Y) un couple de variables aléatoires.
Si X et Y sont indépendantes, il en va de même pour les variables aléatoires $f(X)$ et $g(Y)$ où $f:\mathbb{R}\to\mathbb{R}$ et $g:\mathbb{R}\to\mathbb{R}$ sont deux applications quelconques.

E Espérance

L'espérance mathématique d'une variable aléatoire réelle est la moyenne pondérée par les probabilités d'apparition de chaque valeur. Le théorème de la loi forte des grands nombres démontrera que l'espérance est la valeur que l'on s'attend à trouver, en moyenne, si l'on répète un grand nombre de fois la même expérience aléatoire.

Définition Espérance d'une variable aléatoire

Soit X une variable aléatoire.
L'**espérance** de X est

$$E(X) = \sum_{\forall i \in [\![1,n]\!]} x_i P(X = x_i)$$

Exemple **Somme de deux dés**

Reprenons l'exemple du lancer de deux dés en notant S la somme des chiffres obtenus. L'espérance de S vaut alors :

$$E(S) = \sum_{s \in S(\Omega)} sP(S = s) = 2.\frac{1}{36}+3.\frac{2}{36}+4.\frac{3}{36}+5.\frac{4}{36}+6.\frac{5}{36}+7.\frac{6}{36}+8.\frac{5}{36}+9.\frac{4}{36}+10.\frac{3}{36}+11.\frac{2}{36}+12.\frac{1}{36} = 7.$$

Théorème Formule de transfert

Soit X une variable aléatoire et $f : \mathbb{R} \to \mathbb{R}$.
On a

$$E(f(X)) = \sum_{\forall i \in [\![1,n]\!]} f(x_i) P(X = x_i).$$

Démonstration

$$E(f(X)) = \sum_{\forall j \in [\![1,m]\!]} y_j P(f(X) = y_i) = \sum_{\forall j \in [\![1,m]\!]} y_j \sum_{\forall i \in [\![1,n]\!] : f(x_i) = y_j} P(X = x_i) = \sum_{\forall j \in [\![1,m]\!]} \sum_{\forall i \in [\![1,n]\!] : f(x_i) = y_j} y_j P(X = x_i).$$

D'où

$$E(f(X)) = \sum_{\forall j \in [\![1,m]\!]} \sum_{\forall i \in [\![1,n]\!] : f(x_i) = y_j} f(x_i) P(X = x_i) = \sum_{\forall i \in [\![1,n]\!]} f(x_i) P(X = x_i).$$

Remarque

La formule de transfert permet de calculer $E(f(X))$ sans avoir à déterminer la loi de $f(X)$.
Dans l'exemple du jeu, l'espérance de la variable aléatoire $G(X)$ représente le gain moyen du joueur. On applique la formule du transfert :

$$E(G(S)) = \sum_{s \in \{2,3,\ldots,12\}} G(s)P(S = s) = 2P(S = 2) - 1P(S = 7) + 2P(S = 12) = 2\frac{1}{36} - 1\frac{6}{36} + 2\frac{2}{36} = -\frac{1}{18}.$$

En conclusion, ce jeu n'est pas favorable au joueur.

Théorème Espérance du produit de variables aléatoires

Soit X et Y deux variables aléatoires indépendantes.
Alors

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

Démonstration

On applique le théorème de transfert à la variable aléatoire $f(X,Y) = XY$ d'où

$$\begin{aligned} E(XY) &= \sum_{\forall (i,j) \in [\![1,n]\!] \times [\![1,m]\!]} x_i y_j P(X = x_i \cap Y = y_j) \\ &\overbrace{=}^{\text{indépendance}} \sum_{\forall (i,j) \in [\![1,n]\!] \times [\![1,m]\!]} x_i y_j P(X = x_i) P(Y = y_j) \\ &= \sum_{i=0}^{n} \sum_{j=0}^{m} x_i y_j P(X = x_i) P(Y = y_j) \\ &= \sum_{i=0}^{n} x_i P(X = x_i) \sum_{j=0}^{m} y_j P(Y = y_j) \\ &= E(X)E(Y). \end{aligned}$$

Proposition **Propriétés de l'espérance**

Soit X et Y deux variables aléatoires réelles et $(a,b) \in \mathbb{R}^2$.

— $E(a) = a$.

— *linéarité* : $E(aX + bY) = aE(X) + bE(Y)$.

— Si $P(X \geq 0) = 1$, alors $E(X) \geq 0$.

— Si $P(X \leq Y) = 1$, alors $E(X) \leq E(Y)$.

Définition **Centrée**

Une variable aléatoire est dite **centrée** si son espérance est nulle.

Exemple **Centrer une variable aléatoire**

Soit X une variable aléatoire.
La variable aléatoire $Y = X - E(X)$ est centrée car $E(Y) = E(X - E(X)) = E(X) - E(E(X)) = E(X) - E(X) = 0$

Démonstration

— Dans l'expression $E(a)$, a est la variable aléatoire définie par $\omega \mapsto a$ d'où $E(a) = \overbrace{a}^{\in \mathbb{R}} P(a = \overbrace{a}^{\in \mathbb{R}}) = \overbrace{a}^{\in \mathbb{R}}$.

— On applique le théorème de transfert à la variable aléatoire $f(X,Y) = aX + bY$ d'où

$$\begin{aligned} E(aX + bY) &= \sum_{\forall i \in [\![1,n]\!]} \sum_{\forall j \in [\![1,m]\!]} (ax_i + by_j) P(X = x_i \cap Y = y_j) \\ &= a \sum_{\forall i \in [\![1,n]\!]} x_i \sum_{\forall j \in [\![1,m]\!]} P(X = x_i \cap Y = y_j) + b \sum_{\forall j \in [\![1,m]\!]} y_j \sum_{\forall i \in [\![1,n]\!]} P(X = x_i \cap Y = y_j) \\ &= a \sum_{\forall i \in [\![1,n]\!]} x_i P(X = x_i) + b \sum_{\forall j \in [\![1,m]\!]} y_j P(Y = y_j) \\ &= aE(X) + bE(Y). \end{aligned}$$

— Soit X une variable aléatoire telle que $P(X \geq 0) = 1$ d'où $P(X < 0) = 0$.
De plus, pour tout $x \in \mathbb{R}$, $P(X = x) = 0$ car $\{X = x\} \subset \{X < 0\}$.

$$E(X) = \sum_{\forall i \in [\![1,n]\!]} x_i P(X = x_i) = \sum_{\forall i \in [\![1,n]\!] : x_i \geq 0} x_i P(X = x_i) \geq 0.$$

— Soit X une variable aléatoire telle que $P(Y - X \leq 0) = 1$. Ainsi $E(Y - X) \geq 0$ et $E(X) \leq E(Y)$.

F Variance

La variance est une mesure de la dispersion des valeurs d'une loi de probabilité.

Définition **variance d'une variable aléatoire**

Soit X une variable aléatoire.
La **variance** de X est

$$E\left((X - E(X))^2\right)$$

et on la note $V(X)$.
L'**écart-type** est $\sigma(X) = \sqrt{V(X)}$.

Proposition **Formule de Kœnig-Hugens**

Soit X une variable aléatoire.
Pour tout $a > 0$, on a :

$$V(X) = E(X^2) - (E(X))^2$$

Démonstration

$$V(X) = E\left((X - E(X))^2\right)$$
$$V(X) = E\left(X^2 - 2E(X)X + E(X)^2\right)$$
$$V(X) = E(X^2) - 2E(X)E(X) + E(X)^2 \text{ car } E \text{ linéaire et } E(a) = a$$
$$V(X) = E(X^2) - (E(X))^2$$

Exemple **Somme de deux dés**

Reprenons l'exemple du lancer de deux dés en notant S la somme des chiffres obtenus. On a :

$$E(S^2) \overbrace{=}^{\text{th transfert}} \sum_{s \in S(\Omega)} s^2 P(S = s) = 2^2.\frac{1}{36} + 3^2.\frac{2}{36} + 4^2.\frac{3}{36} + \cdots + 12^2.\frac{1}{36} = \frac{329}{6}.$$

La variance de S vaut alors :

$$V(S) = E(S^2) - (E(S))^2 = \frac{329}{6} - 7^2 = \frac{35}{6}.$$

et son écart-type

$$\sigma(S) = \sqrt{V(S)} \approx 2,42.$$

Définition **Réduite**

Une variable aléatoire est dite **réduite** si sa variance est égale à .

Exemple **Centrer et réduire une variable aléatoire**

Soit X une variable aléatoire.
La variable aléatoire $Y = \frac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Définition **Covariance de deux variables aléatoires**

Soit X et Y deux variables aléatoires.
La **covariance** de X et Y est

$$E((X - E(X))(Y - E(Y)))$$

et on la note $\operatorname{Cov}(X, Y)$.

Proposition **Propriétés**

Soit X et Y deux variables aléatoires.

— $V(aX+b) = E\left((X-E(X))^2\right)$.

— $Var(X+Y) = Var(X) + Var(Y) + 2\text{Cov}(X,Y)$.

— Si X et Y sont indépendantes, alors $V(X+Y) = V(X) + V(Y)$.

Démonstration

— $V(aX+b) = E((aX+b-E(aX+b))^2) = E(a^2(X-E(X))^2) = a^2V(X)$ car E linéaire..

— $V(X+Y) = E((X+Y-E(X+Y))^2) = E((X-E(X))^2) + E((Y-E(Y))^2) + 2E((X-E(X)(Y-E(Y))) = Var(X) + Var(Y) + 2\text{Cov}(X,Y)$.

— Si X et Y sont indépendantes, alors $X-E(X)$ et $Y-E(Y)$ sont indépendantes, d'où

$$E((X-E(X)(Y-E(Y))) = E(X-E(X))E(Y-E(Y)) = 0.$$

Définition **Corrélation**

Soit X et Y deux variables aléatoires de variance non nulles.
Le **coefficient de corrélation** de X et Y est

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Proposition **Propriétés**

Soit X et Y deux variables aléatoires de variance non nulles.

— $\rho(X,Y) \in [-1,1]$

— $\rho(X,Y) = 1$ si et seulement si il existe deux réels a et b tels que l'événement $\{Y = aX+b\}$ soit certain.

— Si X et Y sont indépendantes, alors $\rho(X,Y) = 0$. On dit qu'elles sont décorrélées. La réciproque est fausse.

Démonstration

On utilise l'inégalité de Cauchy-Schwarz $|\text{Cov}(X,Y)| \leq \sigma(X)\sigma(Y)$ (voir un cours d'algèbre bilinéaire pour la preuve de cette inégalité).

G Lois usuelles

Loi uniforme

Lorsque on se trouve dans une situation d'équiprobabilité par exemples "jeter un dé équilibré", " tirer une boule dans une urne de boules distinctes", "choisir un élève au hasard pour passer au tableau", on modélise l'expérience aléatoire avec une variable aléatoire suivant une loi uniforme.

Définition-Proposition **Loi uniforme**

Une variable aléatoire X suit la **loi uniforme** sur $[\![1,n]\!]$ si

1. $X(\Omega) = [\![1,n]\!]$,
2. $\forall i \in [\![1,n]\!] : \quad P(X=i) = \frac{1}{n}$.

On note $X \hookrightarrow \mathcal{U}(E)$.
On a : $E(X) = \frac{n+1}{2}$ et $V(X) = \frac{n^2-1}{2}$.

Démonstration

L'espérance est :

$$E(X) = \sum_{i=1}^{n} iP(X=i) = \sum_{i=1}^{n} i\frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Pour la variance, on a :

$$E(X^2) = \sum_{i=1}^{n} i^2 P(X=i) = \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{1}{n}\frac{(2n+1)n(n+1)}{6} = \frac{(2n+1)(n+1)}{6}.$$

Ainsi, on a :

$$V(X) = E(X^2) - (E(X))^2 = \frac{(2n+1)(n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n+1}{2}\left(\frac{2(2n+1)-3(n+1)}{3}\right),$$

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

Loi de Bernoulli

Lorsque l'issue est binaire du type "succès ou échec", "vrai ou faux", "marche ou arrêt", pile ou face", etc, on modélise l'expérience aléatoire avec une variable aléatoire suivant une loi de Bernoulli.

Définition-Proposition **Loi de Bernoulli**

Une variable aléatoire X suit la **loi de Bernoulli** de paramètre p si

1. $X(\Omega) = \{0, 1\}$,
2. $P(X = 1) = p \quad \text{et} \quad P(X = 0) = 1 - p$.

On note $X \hookrightarrow \mathcal{B}(p)$.
On a : $E(X) = p$ et $V(X) = p(1-p)$.

Démonstration

L'espérance est :

$$E(X) = 0P(X=0) + 1P(X=1) = p.$$

Pour la variance, on a :

$$E(X^2) = 0^2P(X=0) + 1^2P(X=1) = p^2.$$

Ainsi, on a :

$$V(X) = p - p^2 = p(1-p).$$

Définition-Proposition **Loi indicatrice**

Soit A est un événement de Ω.
La variable aléatoire réelle 1_A définie par

$$1_A(\omega) = \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \in \bar{A} \end{cases}$$

est une variable de Bernoulli de paramètre $p = P(A)$. On l'appelle **loi indicatrice** de l'événement A.

Loi Binomiale

Lorsque on répète n fois une expérience de Bernoulli indépendamment des unes et des autres et que l'on compte le nombre de succès, par exemples "nombre de faces obtenus après n lancers", on modélise l'expérience aléatoire avec une variable aléatoire suivant une loi binomiale.
En effet, soit $X_1, X_2, \ldots, X_n$ n variables aléatoires indépendantes de loi de Bernoulli de même paramètre p.
On pose $X = X_1 + X_2 + \cdots + X_n$ la variable aléatoire représentant le nombre de succès.
On a $X(\Omega) = [\![0, n]\!]$ et

$$(X = k) = (X_1 + X_2 + \cdots + X_n = k) = \bigcup_{I \subset [\![1,n]\!] : \mathrm{Card}(I) = k} (\cap_{i\in I}(X_i = 1)) \cap \left(\cap_{i\in [\![1,n]\!]\setminus I}(X_i = 0)\right)$$

Comme l'union est disjointe, on a :

$$P(X=k)=\sum_{I\subset[\![1,n]\!]:\mathrm{Card}(I)=k} P((\cap_{i\in I}(X_i=1))\cap\left(\cap_{i\in[\![1,n]\!]\setminus I}(X_i=0)\right))$$

Comme les variables aléatoires sont indépendantes, on a :

$$P(X=k)=\sum_{I\subset[\![1,n]\!]:\mathrm{Card}(I)=k} p^k(1-p)^{n-k}$$

Le nombre de parties à k éléments de l'ensemble $[\![1,n]\!]$ est le coefficient binomial correspondant à la position des k succès, d'où

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

où p^k est la probabilité d'obtention de ces k succès et $(1-p)^{n-k}$ est la probabilité d'obtention des échecs.
On définit donc :

Définition Loi binomiale

Une variable aléatoire X suit la **loi binomiale** de paramètres n et p si

1. $X(\Omega)=[\![1,n]\!]$,
2. $\forall i\in[\![1,n]\!]:\quad P(X=i)=\binom{n}{i}p^i(1-p)^{n-i}$.

On note $X\hookrightarrow\mathcal{B}(n,p)$.
On a : $E(X)=np$ et $V(X)=np(1-p)$.

Proposition Loi des tirages avec remise

Soit $X_1,X_2,\dots,X_n$ n variables aléatoires indépendantes de loi de Bernoulli de même paramètre p.
Alors $X=X_1+X_2+\cdots+X_n$ suit une loi binomiale de paramètres n,p.
En d'autres termes, si l'expérience aléatoire est une répétition de n épreuves identiques et indépendantes tel que chaque épreuve est une expérience Bernouilli de paramètre p et si S est égale au nombre de succès des n épreuves de l'expérience, alors S suit une loi binomiale de paramètres n,p.
Inversement, si X suit une loi binomiale de paramètres n,p, alors il existe $X_1,X_2,\dots,X_n$ n variables aléatoires indépendantes de loi de Bernoulli de même paramètre p tel que $X=X_1+X_2+\cdots+X_n$.

Démonstration

La preuve a été faite dans l'introduction de la loi binomiale.

Remarque

Pour $n=1$, on retrouve la loi de Bernoulli de paramètre p.

Proposition Espérance et Variance

Soit $X\hookrightarrow\mathcal{B}(n,p)$.
On a : $E(X)=np$ et $V(X)=np(1-p)$.

Démonstration

Au lieu d'un calcul directe de l'espérance $E(X)=\sum_{i=0}^n i\binom{n}{i}p^i(1-p)^{n-i}p^i(1-p)^{n-i}$, nous utilisons la proposition sur la loi des tirages avec remise.
Il existe $X_1,X_2,\dots,X_n$ n variables aléatoires indépendantes de loi de Bernoulli de même paramètre p tel que $X=X_1+X_2+\cdots+X_n$. Ainsi :

$$E(X)=E(X_1+X_2+\cdots+X_n)\overbrace{=}^{\text{linéarité}}E(X_1)+E(X_2)+\cdots+E(X_n)=np.$$

Aussi, la variance est égale à :

$$V(X)=V(X_1+X_2+\cdots+X_n)\overbrace{=}^{\text{indépendantes}}V(X_1)+V(X_2)+\cdots+V(X_n)=np(1-p).$$

H Convergence et approximations

Proposition Inégalité de Markov

Soit X une variable aléatoire réelle positive.
Pour tout $a > 0$, on a :

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

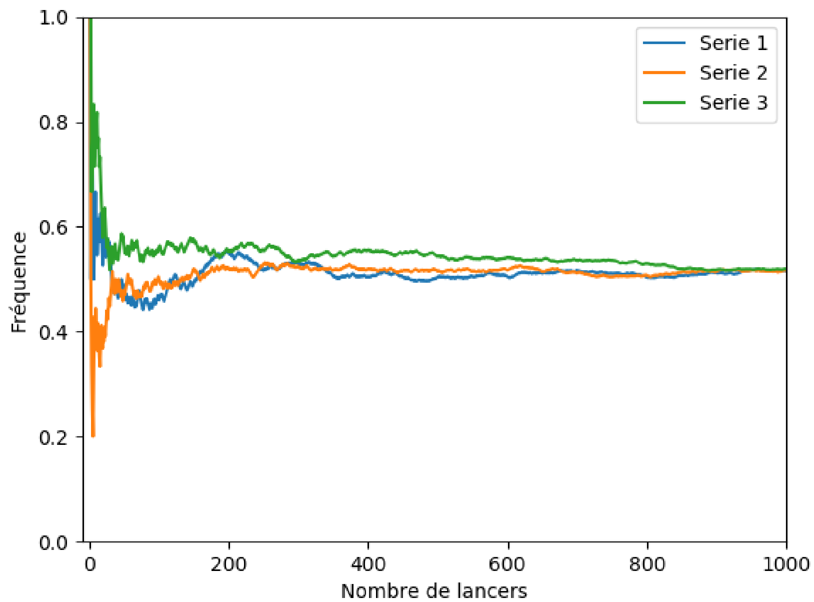
Proposition Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle. Pour tout $a > 0$, on a

$$P(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}$$

Cette inégalité présente un intérêt théorique en majorant la probabilité qu'une variable aléatoire s'écarte de sa moyenne. Nous allons l'utiliser pour prouver la loi faible des grands nombres.

Lors d'un lancer d'une pièce de monnaie équilibrée, les deux côtés « pile » et « face » apparaissent de façon équiprobable pour des raisons de symétrie : on ne s'attend pas plus à l'un ou à l'autre côté. Cette mesure de l'attente s'appuie souvent sur une considération statistique : on observe que la fréquence des occurrences de chaque côté se rapproche de $1/2$.



Le théorème de la loi des grands nombres permet de justifier ce résultat en interprétant la probabilité comme une fréquence de réalisation.
La modélisation de cette expérience aléatoire est :

1. **n$^{\text{ième}}$ lancer** : pour tout $i \in \mathbb{N}^*$, la variable aléatoire X_i, représente le résultat du n$^{\text{ième}}$ lancer. Elle suit une loi Bernoulli de paramètre p (l'issue 0 représente le pile et l'issue 1 le face). Elles sont supposées indépendantes. On a $E(X_i) = (1-p).0 + p.1 = p$.
2. **fréquence** : pour tout $n \in \mathbb{N}^*$, la variable aléatoire S_n, représente la moyenne des résultats obtenus au cours des n premiers lancers, soit :
$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

L'issue, ω, correspondant à une succession de faces, $X_n(\omega) = 1$ pour tout $n \in \mathbb{N}^*$, existe. Donc 1 est une valeur possible de la variable aléatoire S_n. Cependant, le théorème de la loi faible des grands nombres prouve que la **probabilité** que S_n s'écarte de l'espérance $E(X_n)$ tend vers 0 quand n tend vers l'infini.

Théorème Loi faible des grands nombres

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et de même loi, admettant une espérance m et un écart type σ.
On pose $S_n = \frac{X_1+X_2+\cdots+X_n}{n}$.
Alors

$$\forall \epsilon > 0 : \quad P\left(|S_n - m| \geq \epsilon\right) \xrightarrow[n\to+\infty]{} 0.$$

On dit que la suite des variables aléatoires (s_n) **converge en probabilité** vers m.

Démonstration

D'après l'inégalité de Bienaymé-Tchebychev et par indépendance des variables aléatoires,

$$\forall \epsilon > 0 : \quad P\left(|S_n - m| \geq \epsilon\right) \leq \frac{V(S_n)}{\epsilon^2} = \frac{\frac{V(X_1+X_2+\cdots+X_n)}{n^2}}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \xrightarrow[n\to+\infty]{} 0.$$

Théorème Théorème de Bernoulli

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements indépendants de même probabilité p.
La suite des variables aléatoires $\frac{1}{n}\sum_{i=0}^{N} 1_{A_i}$ converge en probabilité vers p.

Démonstration

Les variables aléatoires 1_{A_n} sont indépendantes de même loi de Bernoulli. D'après la loi faible des grands nombres, la suite des variables aléatoires $\frac{1}{n}\sum_{i=0}^{N} 1_{A_i}$ converge en probabilité vers $E(X_1) = p$.

Remarque

Ce théorème assure qu'en répétant une expérience aléatoire, la fréquence des occurrences d'un événement converge en probabilité vers la probabilité de l'événement. Ce théorème apporte la cohérence entre l'approche de la modélisation probabiliste et l'approche fréquentiste.
Le TP Python https://github.com/VincentTariel/cours/blob/master/probabilite/simulation_variable_aleatoire_avtivite_python.pdf permet de vous familiariser sur les applications de ce théorème.