Chapitre 19 : Probabilités et variables aléatoires finies

Table des matières

1 Espaces probabilisés finis

1.1 Introduction

Étant donnée une expérience (par exemple, un lancé de dé), on qualifie cette expérience d'***aléatoire*** car on ne peut prédire à l'avance le résultat de cette expérience.
L'objet du chapitre est de modéliser les choses afin d'effectuer des calculs de probabilités et prédire non pas ce que va donner l'expérience mais prédire quels sont les événements qui ont plus de chance de se produire que d'autres...

Dans toute la suite, la notation « $\bigsqcup$ »désigne une réunion disjointe d'ensembles. Par exemple, l'ensemble $A \bigsqcup B$ est égal à $A \cup B$ mais signifie de plus que les ensembles A et B sont disjoints. La notation $\bigsqcup\limits_{i \in I} A_i$ correspond à la réunion $\bigcup\limits_{i \in I} A_i$, mais signifie de plus que les ensembles A_i sont deux à deux disjoints pour i décrivant l'ensemble d'indices I.

1.2 Ensemble univers Ω

Définition 1 On considère une expérience aléatoire. Traditionnellement, le résultat de chaque expérience est noté ω et on appelle l'***univers des possibles***, l'ensemble Ω de tous les résultats possibles issus de l'expérience considérée. Cet ensemble Ω est toujours non vide.

Exemple 1 Déterminer l'univers des possibles dans les cas suivants :
- on lance deux dés traditionnels à 6 faces et on fait la somme des deux faces obtenues
- on jette un disque de rayon 1 dans le plan et on compte le nombre de points à coordonnées entières dans ce disque
- on lance une pièce de monnaie et on compte le nombre de lancers nécessaires pour obtenir la séquence « pile », « pile » et « face » pour la première fois
- on pioche [avec remise, sans remise ou simultanément] p boules dans une urne contenant $n \geqslant p$ boules et on regarde le nombre de boules blanches obtenues.

1.3 Événement sur un univers fini

Définition 2 On considère une expérience aléatoire ne pouvant donner qu'un nombre fini de résultats possibles : l'univers Ω des possibles est donc fini. On appelle ***événement***, toute partie de Ω. L'ensemble des événements est alors l'ensemble $\mathscr{P}(\Omega)$ de toutes les parties de Ω.
Si ω est dans Ω, le singleton $\{\omega\}$ est appelé ***événement élémentaire***.
Si A est un événement, la partie $\Omega \setminus A$ est appelée ***événement contraire*** et notée $\overline{A}$.
Si A et B sont deux événements, on dit que A et B sont des ***événements incompatibles*** si $A \cap B = \emptyset$.

1.4 Probabilité $\mathbb{P}$ sur un univers fini

Définition 3 Soit une expérience aléatoire ne pouvant donner qu'un nombre fini de résultats possibles. On appelle ***probabilité sur*** Ω toute fonction $\mathbb{P} : \mathscr{P}(\Omega) \longrightarrow [0, 1]$ telle que les deux conditions suivantes soient réalisées :
- $\mathbb{P}(\Omega) = 1$
- [σ-***additivité***] pour tous événements incompatibles A et B, on a : $\mathbb{P}(A \bigsqcup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Si A est un événement, le nombre $\mathbb{P}(A)$ s'appelle la ***probabilité de l'événement*** A.
On dit qu'un événément A est ***négligeable*** si $\mathbb{P}(A) = 0$ et on dit qu'un événement est ***presque sûr*** ou ***se réalise presque sûrement*** si $\mathbb{P}(A) = 1$.

Définition 4 Si Ω est un univers fini et si $\mathbb{P} : \mathscr{P}(\Omega) \longrightarrow [0,1]$ est une probabilité, le couple $(\Omega, \mathbb{P})$ s'appelle un ***espace probabilisé fini***.

1.5 Premières propriétés

Proposition 1 Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Alors :

- pour tout événement A, on a $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$. En particulier, $\mathbb{P}(\emptyset) = 0$.
- si $A \subset B$, alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$
- pour tous événements $A_1, \cdots, A_r$ deux à deux disjoints, on a :

$$\mathbb{P}\left(\bigsqcup_{i=1}^{r} A_i\right) = \sum_{i=1}^{r} \mathbb{P}(A_i)$$

- pour tous événements A et B, on a : $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- si $\Omega = \{\omega_1, \cdots, \omega_s\}$, construire une probabilité $\mathbb{Q}$ sur Ω revient à se donner un s-uplet de nombres réels $(p_1, \cdots, p_s)$ tous positifs ou nuls et de somme égale à 1. Plus précisément :

▷ si $(p_1, \cdots, p_s)$ est un tel s-uplet, la fonction :

$$\mathbb{Q} : \left| \begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ A & \longmapsto & \displaystyle\sum_{k=1 \text{ tel que } \omega_k \in A}^{s} p_k \end{array} \right.$$

est une probabilité.

▷ si $\mathbb{Q}$ est une probabilité sur Ω, le s-uplet $(p_1, \cdots, p_s)$ avec $p_k = \mathbb{Q}(\{\omega_k\})$ a ses composantes toutes positives ou nulles, de somme égale à 1 et :

$$\forall A \in \mathscr{P}(\Omega),\ \mathbb{Q}(A) = \sum_{k=1 \text{ tel que } \omega_k \in A}^{s} p_k.$$

Exemple 2 • Si $\omega \in \Omega$, la mesure de Dirac δ_ω est la probabilité :

$$\delta_\omega : \left| \begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbf{1}_A(\omega) \end{array} \right. .$$

Cette loi de Dirac est associée au 1-uplet $(1)_\omega$.

• Si Ω est un ensemble fini de cardinal s, toutes les probabilités sur $\mathscr{P}(\Omega)$ sont des fonctions de la forme $P = \displaystyle\sum_{\omega \in \Omega} \lambda_\omega \cdot \delta_\omega$, où le s-uplet $(\lambda_\omega)_{\omega \in \Omega}$ est composé de réels positifs de somme égale à 1. Les probabilités sont des barycentres à poids positifs de mesures de Dirac.

• Voici une liste de lois usuelles de probabilité :

$\longrightarrow$ si $\Omega = [\![1, n]\!]$, la loi uniforme est $\mathscr{U}_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \delta_i$

$\longrightarrow$ si $\Omega = \{0, 1\}$ et $p \in [0,1]$, la loi de Bernoulli est $\mathscr{B}(p) = p \cdot \delta_1 + (1-p) \cdot \delta_0$

$\longrightarrow$ si $\Omega = [\![0, n]\!]$ et $p \in [0,1]$, la loi binomiale est $\mathscr{B}(n,p) = \displaystyle\sum_{k=0}^{n} \binom{n}{k} \cdot p^k\ (1-p)^{n-k} \cdot \delta_k$

$\longrightarrow$ si $\Omega = \{-1, 1\}$ et $p \in [0,1]$, la loi de Rademacher est $\mathscr{R}(p) = p \cdot \delta_1 + (1-p) \cdot \delta_{-1}$.

2 Probabilités conditionnelles sur un univers fini

Dans tout ce paragraphe, on considère une expérience ne donnant toujours qu'un nombre fini de résultats possibles, formant ainsi un univers Ω des possibles qui est un ensemble fini. On considère également une probabilité $\mathbb{P}$ sur Ω.

2.1 Définition

Définition 5 Soit B un événement de probabilité $\mathbb{P}(B)$ strictement positive. La fonction :

$$\mathbb{P}_B : \left| \begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ A & \longmapsto & \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \end{array} \right.$$

est une probabilité sur Ω. Cette probabilité s'appelle la ***probabilité conditionnelle sachant*** B et on note pour tout $A \in \mathscr{P}(\Omega)$:

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

2.2 Système complet d'événements

Définition 6 Soit une famille finie $(A_1, \cdots, A_r)$ d'événements. On dit que cette famille est un ***système complet d'événements*** (en abrégé ***(S.C.E.)***), si les deux conditions suivantes sont réalisées :

- $\displaystyle\bigcup_{k=1}^{r} A_k = \Omega$ [réunion totale]
- pour tous $i \neq j$ entre 1 et r, on a : $A_i \cap A_j = \emptyset$ [événements deux à deux incompatibles]

Dans ce cas, on note simplement $\displaystyle\bigsqcup_{k=1}^{r} A_k = \Omega$

2.3 Formules

Proposition 2 *formule des probabilités composées*
Soient $A_1, \cdots, A_p$, p événements tels que $\mathbb{P}(A_1 \cap \cdots \cap A_{p-1}) > 0$. Alors :

$$\mathbb{P}(A_1 \cap \cdots \cap A_p) = \mathbb{P}(A_1) \cdot \prod_{k=1}^{p-1} \mathbb{P}(A_{k+1} \mid A_1 \cap \cdots \cap A_k).$$

Proposition 3 *formule des probabilités totales*
Soit $(A_1, \cdots, A_r)$ un S.C.E. Alors, pour tout événement $B \in \mathscr{P}(\Omega)$, on a :

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

Proposition 4 *formule de Bayes*
Soit $(A_1, \cdots, A_r)$ un S.C.E. avec chaque probabilité $\mathbb{P}(A_k)$ strictement positive. Alors, pour tout événement $B \in \mathscr{P}(\Omega)$ tel que $\mathbb{P}(B) > 0$, on a $\mathbb{P}(A_k \mid B) = \dfrac{\mathbb{P}(B \mid A_k) \cdot \mathbb{P}(A_k)}{\mathbb{P}(B)}$ et donc pour tout j entre 1 et r

$$\mathbb{P}(A_j \mid B) = \frac{\mathbb{P}(B \mid A_j) \cdot \mathbb{P}(A_j)}{\displaystyle\sum_{k=1}^{r} \mathbb{P}(B \mid A_k) \cdot \mathbb{P}(A_k)}.$$

Exemple 3 Une compagnie d'assurance automobile a classé ses assurés en trois classes d'âges :
- moins de 25 ans
- de 25 à 50 ans
- plus de 50 ans

Le tableau ci-dessous donne deux informations : la proportion d'assurés appartenant à chaque classe d'âge et la probabilité qu'un assuré d'une classe donnée déclare au moins un accident au cours de l'année :

classe	proportion	probabilité
1	0,2	0,10
2	0,5	0,05
3	0,3	0,08

1. Un assuré est tiré au hasard dans le fichier de la compagnie. Quelle est la probabilité qu'il ait déclaré au moins un accident au cours de l'année ?
2. Quelle est la probabilité qu'un assuré ayant déclaré au moins un accident au cours de l'année ait moins de 25 ans ? ait plus de 50 ans ?
3. Quelle est la probabilité qu'un assuré âgé de 25 ans ou plus ait déclaré au moins un accident au cours de l'année ?
4. Quelle est la probabilité qu'un assuré n'ayant pas déclaré d'accident n'appartienne pas à la classe 2 ? ait moins de 50 ans ?

3 Événements indépendants sur un espace probabilisé fini

Dans tout ce paragraphe, on considère un espace probabilisé fini $(\Omega, \mathbb{P})$.

Définition 7 On dit que deux événements A et B sont ***indépendants*** si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.
On dit que des événements $A_1, \cdots, A_p$ sont ***mutuellement indépendants*** si pour toute partie non vide I incluse dans $\{1, \cdots, p\}$, on a :

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

Proposition 5 • Des événements mutuellement indépendants sont indépendants deux à deux. La réciproque est fausse.
• si $A_1, \cdots, A_p$ sont des événements mutuellement indépendants, alors en prenant pour tout $k \in \{1, \cdots, p\}$, $B_k = A_k$ ou $B_k = \overline{A_k}$, les événements de la famille $(B_1, \cdots, B_p)$ restent mutuellement indépendants.

Exemple 4 L'indépendance deux à deux n'implique pas la mutuelle indépendance : par exemple lorsque $\Omega = \{1, 2, 3, 4\}$ $\mathscr{A} = \mathscr{P}(\Omega)$ et $\mathbb{P}$ la probabilité uniforme sur $\mathscr{P}(\Omega)$, alors les événements $A = \{1, 2\}$, $B = \{1, 3\}$ et $C = \{2, 3\}$ sont deux à deux indépendants et non mutuellement indépendants.
• On lance un dé à n faces numérotées de 1 à n, et ceci p fois de suite.
On fixe un vecteur $\vec{u} = (u_1, \cdots, u_p) \in [\![1, n]\!]^p$.
On pose pour tout $i \in [\![1, p]\!]$, l'événement A_i : « au $i^{ème}$ lancer, le dé a fait u_i ».
Alors, les événements $A_1, \cdots, A_p$ sont mutuellement indépendants.

4 Variables aléatoires

Dans tout ce paragraphe, on considère un espace probabilisé fini $(\Omega, \mathbb{P})$.

4.1 Définitions et notations

Définition 8 On appelle ***variable aléatoire réelle*** (en abrégé ***v.a.r.***), toute fonction $X : \Omega \longrightarrow \mathbb{R}$.
L'ensemble des valeurs prises par la v.a.r. X est noté $X(\Omega)$. C'est donc un ensemble fini.
Lorsque A est une partie de $\mathbb{R}$, on adoptera les notations suivantes :

- l'ensemble $X^{[-1]}(A) = \Big\{\omega \in \Omega \mid X(\omega) \in A\Big\}$ sera noté $\{X \in A\}$
- la probabilité $\mathbb{P}(\{X \in A\})$ sera notée $\mathbb{P}(X \in A)$.

Définition 9 Soit $X : \Omega \longrightarrow \mathbb{R}$ une v.a.r. Alors la fonction :

$$\mu : \left| \begin{array}{ccc} \mathscr{P}(X(\Omega)) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbb{P}(X \in A) \end{array} \right.$$

est une probabilité sur l'univers fini $X(\Omega)$. Cette probabilité est appelée la ***loi de la v.a.r.*** X et notée $\mathbb{P}_X$.

Méthode : Comment calculer la loi d'une v.a.r. $X : \Omega \longrightarrow \mathbb{R}$?

- déterminer l'ensemble $X(\Omega)$ des valeurs prises par la v.a.r. X ou un ensemble contenant les valeurs prises par $X(\Omega)$, certaines valeurs pouvant être prises avec une probabilité nulle.
- pour chaque valeur prise $x \in X(\Omega)$, calculer $\mathbb{P}(X = x)$ la probabilité que X vaille x.

Exemple 5 • On considère un entier $n \geqslant 1$ et on note $E = [\![1, n]\!]$. On choisit au hasard une partie $\omega \in \mathscr{P}(E)$ et on note $X(\omega)$ le cardinal de la partie ω. Expliciter la loi de X.
• Soient $1 \leqslant p \leqslant N$ deux entiers. Un sac contient N boules en tout, dont p boules sont blanches. Soit $q \in [\![1, p]\!]$. On pioche q boules du sac. On note X le nombre de boules blanches obtenues.
Déterminer la loi de X lorsque les tirages sont avec remise, sans remise ou simultanés.

Proposition 6 Si $X : \Omega \longrightarrow \mathbb{R}$ est une v.a.r., alors :
• les événements $\{X = x\}$, lorsque x parcourt $X(\Omega)$ forment un S.C.E.
• si $f : \mathbb{R} \longrightarrow \mathbb{R}$ est une fonction, alors la fonction $f(X) = f \circ X : \Omega \longrightarrow \mathbb{R}$ est une autre v.a.r. sur Ω et la loi de cette v.a.r. est donnée par :

$$\forall y \in f(X(\Omega)),\ \mathbb{P}_{f(X)}(\{y\}) = \mathbb{P}(f(X) = y) = \mathbb{P}(X \in \{x \in \mathbb{R} \mid f(x) = y\}).$$

4.2 Espérance d'une v.a.r.

Définition 10 Soit $X : \Omega \longrightarrow \mathbb{R}$ une v.a.r. sur Ω fini. On appelle ***espérance de la v.a.r.*** X et on note $\mathbb{E}(X)$, le nombre :

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

Lorsque $\mathbb{E}(X) = 0$, on dit que la v.a.r. X est ***centrée***.

Remarque : l'espérance d'une variable aléatoire finie – c'est-à-dire qui ne prend qu'un nombre fini de valeurs – est un barycentre des valeurs prises par cette variable aléatoire, les poids étant les probabilités de réalisation de ces valeurs prises.

Méthode : Comment calculer l'espérance d'une v.a.r. $X : \Omega \longrightarrow \mathbb{R}$?

Il y a deux méthodes différentes selon que l'on connaisse facilement la loi de la variable X ou non.
Si on peut connaître la loi de X facilement ...

- déterminer la loi de X ;
- calculer la somme $\displaystyle\sum_{x \in X(\Omega)} x \cdot \mathbb{P}(X = x)$.

Si on ne peut pas connaître la loi de X facilement ...

- interpréter la variable X comme une somme de fonctions simples, par exemple une somme de fonctions indicatrices $X = \displaystyle\sum_{k=1}^{n} \lambda_k \cdot \mathbf{1}_{A_k}$;
- utiliser la linéarité de l'espérance ; par exemple, $\mathbb{E}(X) = \mathbb{E}\left(\displaystyle\sum_{k=1}^{n} \lambda_k \cdot \mathbf{1}_{A_k}\right) = \displaystyle\sum_{k=1}^{n} \lambda_k \cdot \mathbb{P}(A_k)$.

Proposition 7 • L'application $X \longmapsto \mathbb{E}(X)$ est une forme linéaire sur l'espace des v.a.r. sur Ω : $\mathbb{E}(\lambda \cdot X + Y) = \lambda \cdot \mathbb{E}(X) + \mathbb{E}(Y)$, où X et Y sont deux v.a.r. sur Ω.

• Si $X \leqslant Y$ sont deux v.a.r., alors $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

• Si $X : \Omega \longrightarrow \mathbb{R}$ est une v.a.r. et $f : \mathbb{R} \longrightarrow \mathbb{R}$ est une fonction, alors la fonction $f(X) = f \circ X$ est une autre v.a.r. sur Ω et on a la ***formule de transfert*** :

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

• Pour tous a et b réels, pour toute v.a.r. X sur Ω, on a : $\mathbb{E}(a \cdot X + b) = a \cdot \mathbb{E}(X) + b$.

• On a l'***inégalité de Markov*** valable pour tout $a > 0$ et toute v.a.r. $X : \Omega \longrightarrow \mathbb{R}$:

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

Exemple 6 • Si $X : \Omega \longrightarrow \mathbb{R}$ est une variable aléatoire finie, alors pour toute fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$, on a :

$$f(X) = \sum_{x \in X(\Omega)} f(x) \cdot \mathbf{1}_{\{X = x\}},$$

ce qui permet de retrouver la formule de transfert après la prise de l'espérance.

• On pioche q boules avec remise dans un sac contenant N boules dont p sont blanches. Déterminer l'espérance de la variable aléatoire comptant le nombre de boules blanches piochées.

• Calculer l'espérance de la variable aléatoire comptant le cardinal d'une partie prise aléatoirement dans $\mathscr{P}(\llbracket 1, n \rrbracket)$.

Exemple 7 • On dispose d'un tableau de format $2 \times n$, avec $n \geqslant 2$ un entier. On place aléatoirement p croix dans ces $2n$ cases, une croix par case, avec $1 \leqslant p \leqslant 2n$. Calculer le nombre moyen de colonnes entièrement remplies par deux croix dans ce tableau.

• Soit $p \in \mathbb{N}^*$. On lance p dés à six faces parfaitement équilibrés. On note X la somme des faces obtenues. Calculer $\mathbb{E}(X)$.

• On choisit une permutation $\sigma \in \mathfrak{S}_n$ au hasard et de manière équiprobable. Combien la permutation choisie admet-elle en moyenne de points fixes ?

4.3 Variance et écart-type d'une v.a.r.

Définition 11 Soit $X : \Omega \longrightarrow \mathbb{R}$ une v.a.r. sur l'espace probabilisé fini $(\Omega, \mathbb{P})$. On appelle ***variance de*** X et on note $V(X)$, le nombre :

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

On appelle ***écart-type de*** X et on note $\sigma(X)$, le nombre $\sigma(X) = \sqrt{V(X)}$. Si $V(X) = 1$ (ou $\sigma(X) = 1$), on dit que X est une v.a.r. ***réduite***.

Proposition 8 Pour toute v.a.r. $X : \Omega \longrightarrow \mathbb{R}$, on a :

- $V(X) = \mathbb{E}(X^2) - \left(\mathbb{E}(X)\right)^2$ [formule de Huyghens]
- Pour tous réels a et b, $V(a \cdot X + b) = a^2 \cdot V(X)$ et $\sigma(a \cdot X + b) = |a| \cdot \sigma(X)$.
- Si $\sigma(X) > 0$, alors la v.a.r. $\dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est une v.a.r. centrée réduite.
- On a l'***inégalité de Bienaymé-Tchebychev*** valable pour tout $a > 0$ et toute v.a.r. $X : \Omega \longrightarrow \mathbb{R}$:

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

Exemple 8 Montrer que si $X : \Omega \longrightarrow \mathbb{R}$ est une v.a.r. finie de variance nulle, alors X est presque sûrement constante.

Proposition 9 Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. On définit la relation $\mathscr{R}$ sur l'ensemble $E = \mathscr{F}(\Omega, \mathbb{R})$ des variables aléatoires :

$$\forall (X, Y) \in E^2, \ X \mathscr{R} Y \iff X = Y \text{ p.s.}$$

Alors, la relation $\mathscr{R}$ est une relation d'équivalence.
De plus, si X et Y sont deux variables aléatoires égales presque sûrement, alors :

$\longrightarrow$ les variables X et Y ont la même loi ; autrement dit, pour tout réel x, $\mathbb{P}(X = x) = \mathbb{P}(Y = x)$

$\longrightarrow$ pour toute fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$, les variables $f(X)$ et $f(Y)$ restent égales presque sûrement

$\longrightarrow$ les variables X et Y ont la même espérance et en fait, pour toute fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$, $\mathbb{E}(f(X)) = \mathbb{E}(f(Y))$; en particulier, les variables X et Y ont aussi la même variance

$\longrightarrow$ la réciproque est fausse : deux variables aléatoires finies ayant la même espérance et la même variance ne sont pas forcément égales presque sûrement.

Exemple 9 • Soit $X : \Omega \longrightarrow \{0, 1\}$ une variable aléatoire de loi de Bernoulli de paramètre $\dfrac{1}{2}$.
Calculer la loi de $Y = 1 - X$.
• Soient $X : \Omega \longrightarrow \mathbb{R}$ et $Y : \Omega \longrightarrow \mathbb{R}$ deux variables aléatoires finies. Montrer l'équivalence entre les trois points :

$\triangleright$ pour toute fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$, on a :

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

$\triangleright$ pour toute fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$, on a :

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

$\triangleright$ les variables X et Y ont la même loi.

• Soient $X : \Omega \longrightarrow \mathbb{R}$ et $Y : \Omega \longrightarrow \mathbb{R}$ deux variables aléatoires finies. Montrer l'équivalence entre les deux points :

▷ pour tout événement $A \subset \Omega$,
$$\mathbb{E}(X \cdot \mathbf{1}_A) = \mathbb{E}(Y \cdot \mathbf{1}_A)$$

▷ les variables X et Y sont égales presque sûrement.

5 Couple de v.a.r.

Dans ce paragraphe on se donne un espace probabilisé fini $(\Omega, \mathbb{P})$.

5.1 Définitions, lois de couple

Définition 12 On appelle ***couple de v.a.r. sur*** Ω, la donnée d'une application $(X, Y) : \Omega \longrightarrow \mathbb{R}^2$, avec $X : \Omega \longrightarrow \mathbb{R}$ et $Y : \Omega \longrightarrow \mathbb{R}$ deux v.a.r.

Définition 13 Soit $(X, Y) : \Omega \longrightarrow \mathbb{R}$ un couple de v.a.r. sur Ω. La ***loi conjointe de*** (X, Y) est la loi du couple (X, Y), c'est-à-dire la donnée de l'ensemble (fini) des valeurs prises puis pour chaque valeur prise (x, y) la probabilité $\mathbb{P}((X, Y) = (x, y)) = \mathbb{P}(X = x \text{ et } Y = y) = \mathbb{P}(X = x, Y = y)$.
Les lois de X et de Y sont alors appelées ***lois marginales***.
Si $x \in X(\Omega)$ vérifie $\mathbb{P}(X = x) > 0$, l'application $\mu : A \longmapsto \mathbb{P}(Y \in A \mid X = x)$ est une probabilité appelée la ***loi conditionnelle de*** Y ***sachant*** $\{X = x\}$ et notée $\mathbb{P}_{Y \mid X=x}$.

Remarque 1 On peut restituer sous forme d'un tableau à deux entrées la loi conjointe de (X, Y), chaque case renfermant $\mathbb{P}(X = x, Y = y)$. Les marginales de X et de Y se lisent alors sur les marges horizontales ou verticales de ce tableau.

5.2 Covariance

Définition 14 Soit $(X, Y) : \Omega \longrightarrow \mathbb{R}^2$ un couple de v.a.r. sur Ω. On appelle ***covariance du couple*** (X, Y), le nombre :
$$\mathrm{Cov}(X, Y) = \mathbb{E}\Big((X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))\Big).$$

Proposition 10 [formule de Huyghens]
Soit $(X, Y) : \Omega \longrightarrow \mathbb{R}^2$ un couple de v.a.r. sur Ω. Alors,
$$\mathrm{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - E(X) \cdot \mathbb{E}(Y).$$

5.3 V.a.r. indépendantes

Définition 15 Soit $(X, Y) : \Omega \longrightarrow \mathbb{R}$ un couple de v.a.r. sur Ω. On dit que les v.a.r. X et Y sont ***indépendantes*** si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a :
$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y).$$
On dit que p v.a.r. $X_1, \cdots, X_p$ sur Ω sont ***mutuellement indépendantes*** si pour tout $(x_1, \cdots, x_p)$ avec chaque x_k dans $X_k(\Omega)$, on a :
$$\mathbb{P}\Big((X_1, \cdots, X_p) = (x_1, \cdots, x_p)\Big) = \prod_{k=1}^{p} \mathbb{P}(X_k = x_k).$$

5.4 Propriétés

Proposition 11 Soient X et Y deux v.a.r. indépendantes sur Ω. Alors :

- pour tout $A \subset X(\Omega)$ et tout $B \subset Y(\Omega)$, on a :

$$\mathbb{P}((X,Y) \in A \times B) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B).$$

- On a encore la $\boxed{\textit{formule de transfert}}$: pour toute fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ et pour tout couple (Z_1, Z_2) de v.a.r. sur Ω, alors :

$$\mathbb{E}(f(Z_1, Z_2)) = \sum_{(a,b))\in Z_1(\Omega)\times Z_2(\Omega)} f(a,b) \cdot \mathbb{P}(Z_1 = a, Z_2 = b).$$

Plus généralement, si $\vec{Z} = (Z_1, \cdots, Z_d) : \Omega \longrightarrow \mathbb{R}^d$ est un vecteur aléatoire, si $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ est une fonction, alors :

$$\mathbb{E}\Big(f(Z_1, \cdots, Z_d)\Big) = \sum_{(z_1,\cdots,z_d)\in\vec{Z}(\Omega)} f(z_1, \cdots, z_d) \cdot \mathbb{P}\Big(Z_1 = z_1, \cdots, Z_d = z_d\Big).$$

- On a $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$ puis $\mathrm{Cov}(X,Y) = 0$ et la réciproque est fausse.
- Pour toutes fonctions $f : X(\Omega) \longrightarrow \mathbb{R}$ et $g : Y(\Omega) \longrightarrow \mathbb{R}$, les v.a.r. $f(X)$ et $g(Y)$ restent indépendantes.
- $\boxed{[\textit{lemme des coalitions}]}$ Si $X_1, \cdots, X_p$ sont des variables aléatoires mutuellement indépendantes, si $(U_1, \cdots, U_s)$ est une partition de $[\![1,p]\!]$, en prenant pour tout $j \in [\![1,s]\!]$, une variable aléatoire Y_j de la forme $Y_j = f_j\Big((X_k)_{k\in U_j}\Big)$, alors les v.a.r. $Y_1, \cdots, Y_r$ restent mutuellement indépendantes. En particulier, les variables $\varphi(X_1, \cdots, X_r)$ et $\psi(X_{r+1}, \cdots, X_p)$ sont indépendantes, pour toutes fonctions $\varphi : \mathbb{R}^r \longrightarrow \mathbb{R}$ et $\psi : \mathbb{R}^{p-r} \longrightarrow \mathbb{R}$.
- Si $X_1, \cdots, X_p$ sont des variables mutuellement indépendantes, si pour tout $i \in [\![1,p]\!]$, A_i est une partie de $\mathbb{R}$, alors :

$$\mathbb{P}\left(\bigcap_{i=1}^{p} \{X_i \in A_i\}\right) = \prod_{i=1}^{p} \mathbb{P}(X_i \in A_i).$$

De plus,

$$\mathbb{E}\Big(X_1 \times \cdots \times X_p\Big) = \mathbb{E}(X_1) \times \cdots \times \mathbb{E}(X_p).$$

Exemple 10 Soient $A_1, \cdots, A_p$ des événements. A-t-on l'équivalence entre les assertions suivantes :

- les événements $A_1, \cdots, A_p$ sont mutuellement indépendants
- les variables aléatoires $\mathbf{1}_{A_1}, \cdots, \mathbf{1}_{A_p}$ sont mutuellement indépendantes ?

Exemple 11 Soit $\mathscr{A}$ un ensemble fini, $\mu : \mathscr{P}(\mathscr{A}) \longrightarrow [0,1]$ une probabilité et enfin un entier $s \geqslant 1$.
Alors, il existe s variables aléatoires $X_1, \cdots, X_s$ définies sur un même univers Ω et à valeurs dans $\mathscr{A}$ telles que les variables aléatoires X_i soient mutuellement indépendantes et chaque variable X_i suive la loi μ.

Proposition 12 • L'application $(X,Y) \longmapsto \mathbb{E}(X,Y)$ définit un produit scalaire sur l'espace des v.a.r. finies définies de Ω vers $\mathbb{R}$ – en convenant que $X = 0 \Longleftrightarrow \mathbb{P}(X = 0) = 1$.
L'inégalité de Cauchy-Schwarz donne alors :

$$|\mathrm{Cov}(X,Y)| \leqslant \sigma(X) \cdot \sigma(Y).$$

- Si $X_1, \cdots, X_n$ sont n variables aléatoires finies définies de Ω vers $\mathbb{R}$, alors la linéarité de l'espérance permet d'écrire :

$$\mathbb{E}\Big(X_1 + \cdots + X_n\Big) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n).$$

De plus,

$$V\Big(X_1+\cdots+X_n\Big)=\sum_{k=1}^{n}V(X_k)+2\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(X_i,X_j).$$

Lorsque les variables $X_1,\cdots,X_n$ sont indépendantes deux à deux, alors :

$$V\Big(X_1+\cdots+X_n\Big)=\sum_{k=1}^{n}V(X_k).$$

Exemple 12 Le couple (X,Y) dont la loi est :

$X\setminus Y$	0	1
0	0	$\frac{1}{4}$
1	$\frac{1}{2}$	0
2	0	$\frac{1}{4}$

est de covariance nulle mais les variables X et Y ne sont pas indépendantes.

Exemple 13 • Soit $(X_n:\Omega\longrightarrow\mathbb{R})_{n\in\mathbb{N}}$ une suite de variables aléatoires mutuellement indépendantes, finies et identiquement distribuées.

1. Montrer que les nombres $\mathbb{E}(X_n)$ et $\sigma(X_n)$ ne dépendent pas de la variable n. On pose dans la suite $m=\mathbb{E}(X_n)$ et $\sigma=\sigma(X_n)$.
2. Montrer que si $\sigma=0$, alors les X_n sont presque sûrement constantes.
3. On suppose $\sigma>0$. Montrer que :

$$\forall\varepsilon>0,\ \forall n\in\mathbb{N}^*,\ \mathbb{P}\left(\left|\frac{X_1+\cdots+X_n}{n}-m\right|\geqslant\varepsilon\right)\leqslant\frac{\sigma^2}{n\ \varepsilon^2}.$$

4. En déduire la loi faible des grands nombres :

$$\forall\varepsilon>0,\ \lim_{n\longrightarrow+\infty}\mathbb{P}\left(\left|\frac{X_1+\cdots+X_n}{n}-m\right|\geqslant\varepsilon\right)=0.$$

• En fait, on a la $\boxed{\text{loi forte des grands nombres}}$: si $(X_n:\Omega\longrightarrow\mathbb{R})_{n\in\mathbb{N}}$ est une suite de variables aléatoires mutuellement indépendantes, finies et identiquement distribuées, en posant $m=\mathbb{E}(X_n)$, alors :

$$\lim_{n\longrightarrow+\infty}\frac{X_1+\cdots+X_n}{n}=m\ \text{, presque sûrement.}$$

Si A est un événement, $\mathbb{P}(A)$ est la limite presque sûre de la proportion du nombre de réalisations de l'événement A sur n expériences réalisées dans des conditions indépendantes.

6 Lois usuelles finies

6.1 La loi uniforme sur un ensemble fini

Définition 16 Soit E un ensemble fini. On rappelle que le ***cardinal*** noté $\mathrm{Card}(E)$ de l'ensemble E est le nombre fini de ses éléments.

Définition 17 Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. On dit que la v.a.r. $X : \Omega \longrightarrow \mathbb{R}$ suit ***une loi uniforme*** si :

$$\forall k \in X(\Omega),\ \mathbb{P}(X = k) = \frac{1}{\text{Card}(X(\Omega))}.$$

Autrement dit, toutes les valeurs prises par X sont ***équiprobables***.
La loi uniforme de la v.a.r. X est alors :

$$\forall A \in \mathscr{P}(X(\Omega)),\ \mathbb{P}(X \in A) = \frac{\text{Card}(A)}{\text{Card}(X(\Omega))} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Proposition 13 Si X suit la loi uniforme sur $X(\Omega) = \{1, \cdots, n\}$, alors :

$$\mathbb{E}(X) = \frac{n+1}{2} \text{ et } V(X) = \frac{n^2 - 1}{12}.$$

Exemple 14 • La loi uniforme apparaît en cas d'équiprobabilité, lorsque l'on tire un objet parmi n objets au total.

• Dans une urne contenant N boules dont p boules sont blanches, on pioche s boules simultanément. Déterminer la loi de la variable X comptant le nombre de boules blanches piochées.

• Au cours de n années consécutives, on a mesuré différentes hauteurs de crues du Rhône. On suppose que les hauteurs observées sont toutes différentes. Soit j un entier entre 1 et n. Quelle est la probabilité que la $j^{ème}$ année soit une année record ?

6.2 La loi de Bernoulli

Définition 18 Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Soit $p \in [0, 1]$. On dit que la v.a.r. $X : \Omega \longrightarrow \mathbb{R}$ suit ***la loi de Bernoulli de paramètre*** p si :

$$X(\Omega) = \{0, 1\} \text{ et } \left\{ \begin{array}{l} \mathbb{P}(X = 1) = p \\ \mathbb{P}(X = 0) = 1 - p \end{array} \right. .$$

Dans ce cas, on note $X \sim \mathscr{B}(p)$ ou encore $X \hookrightarrow \mathscr{B}(p)$, pour signifier que X suit la loi de Bernoulli de paramètre p.

Proposition 14 Si $X \sim \mathscr{B}(p)$, alors :

$$\mathbb{E}(X) = p \text{ et } V(X) = p(1 - p).$$

Exemple 15 • Cette loi apparaît à chaque situation où le résultat de l'expérience n'a que deux éventualités (tests, tirages pile ou face,...)

• Si $A \in \mathscr{P}(\Omega)$, alors $\mathbf{1}_A$ l'indicatrice de A suit la loi $\mathscr{B}(\mathbb{P}(A))$.

• Soit $n \geqslant 2$ un entier. On considère n^2 variables aléatoires mutuellement indépendantes $X_{i,j} : \Omega \longrightarrow \mathbb{R}$, lorsque le couple (i, j) parcourt $[\![1, n]\!]^2$, chaque variable $X_{i,j}$ suivant la loi de Bernoulli de paramètre $\frac{1}{2}$.

1. Quelle est la probabilité que la matrice aléatoire $M = \left(X_{i,j}\right)_{1 \leqslant i,j \leqslant n}$ vue comme une variable aléatoire $M : \Omega \longrightarrow \mathscr{M}_n(\mathbb{Z}/2\mathbb{Z})$ soit inversible ?

2. On se place dans le cas $n = 100$. On vous propose le jeu suivant : « vous donnez 100 euros pour jouer à ce jeu. Vous piochez ensuite une matrice dans $\mathscr{M}_{100}(\mathbb{Z}/2\mathbb{Z})$. Si la matrice que vous avez piochée est inversible dans $\mathscr{M}_{100}(\mathbb{Z}/2\mathbb{Z})$, alors vous gagnez 500 euros. ».
   Jouez-vous à ce jeu ?

6.3 La loi binomiale

Définition 19 Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Soient $n \in \mathbb{N}^*$ et $p \in [0,1]$. On dit que la v.a.r. $X : \Omega \longrightarrow \mathbb{R}$ suit ***la loi binomiale de paramètres*** n ***et*** p si :

$$X(\Omega) = [\![0,n]\!] \text{ et } \forall k \in [\![0,n]\!],\ \mathbb{P}(X = k) = \left(\begin{array}{c} n \\ k \end{array}\right) \cdot p^k\ (1-p)^{n-k}.$$

Dans ce cas, on note $X \sim \mathscr{B}(n,p)$ ou $X \hookrightarrow \mathscr{B}(n,p)$, pour signifier que X suit la loi binomiale de paramètres n et p.

Proposition 15 • Si $X_1, \cdots, X_n$ sont n v.a.r. mutuellement indépendantes, suivant chacune la loi de Bernoulli $\mathscr{B}(p)$, alors la v.a. $Y = X_1 + \cdots + X_n$ suit la loi binomiale $\mathscr{B}(n,p)$.
• Si $X \sim \mathscr{B}(n,p)$, alors :

$$\mathbb{E}(X) = np \text{ et } V(X) = n \cdot p\ (1-p).$$

• Si $X \sim \mathscr{B}(n,p)$ et $Y \sim \mathscr{B}(m,p)$ et que X et Y sont indépendantes, alors $X + Y \sim \mathscr{B}(n+m,p)$.

Exemple 16 • Cette loi apparaît lors de tirages avec remise.
• Dans un sac contenant une proportion de boules blanches égale à p, on effectue n tirages successifs avec remise. La v.a. X comptant le nombre de boules blanches obtenues suit la loi $\mathscr{B}(n,p)$.

Exemple 17 Soient n et p deux entiers strictement positifs. On considère p variables aléatoires $X_1, \cdots, X_p$ définies d'un univers Ω vers l'ensemble $\mathscr{P}([\![1,n]\!])$ des parties de $[\![1,n]\!]$. On suppose que les variables X_k sont mutuellement indépendantes et suivent toutes la loi uniforme sur $\mathscr{P}([\![1,n]\!])$.

1. Calculer l'espérance de la variable $Y = \#\left(\bigcap_{k=1}^{p} X_k\right)$.
2. Expliciter la loi de la variable Y.

Exemple 18 On se donne un entier $n \geqslant 2$, puis p un réel dans $]0,1[$. On considère n pièces de monnaie donnant « *pile* » avec une probabilité égale à p et donc « *face* » avec une probabilité égale à $q = 1 - p$.
On lance les n pièces de monnaie et on ne garde que les pièces qui ont donné « *pile* ». On note X_1 le nombre de pièces restantes après cette première étape.
À la deuxième étape, on lance les X_1 pièces qui ont été conservées et on les relance. On ne garde que les pièces qui ont redonné « *pile* ». On note X_2 le nombre de pièces restantes après cette deuxième étape.
On poursuit le processus en relançant à la $k^{ème}$ étape les X_{k-1} pièces restantes et en conservant que les X_k pièces ayant donné « *pile* » en vue de l'étape suivante, jusqu'à ne plus avoir de pièces de monnaie à lancer.
Tous les lancers sont indépendants.

1. Déterminer la loi de la variable X_2.
2. Déterminer la loi de la variable X_k, pour $k \in \mathbb{N}$.
3. Déterminer la loi de la variable Y indiquant le dernier numéro de l'étape avant que le processus s'arrête.
   En particulier, quelle est la probabilité que le jeu ne finisse jamais ?

7 Annexe : espaces probabilisés discrets dénombrables

7.1 Univers dénombrable

Définition 20 On dit qu'un ensemble est ***dénombrable*** s'il est en bijection avec $\mathbb{N}$. On parle alors d'***univers dénombrable discret*** pour signifier que l'univers Ω est en bijection avec $\mathbb{N}$. Les parties de Ω sont encore appelées ***événements***.

7.2 Probabilité sur un univers dénombrable

Définition 21 Soit Ω un univers dénombrable. On appelle ***probabilité sur*** Ω toute fonction $\mathbb{P} : \mathscr{P}(\Omega) \longrightarrow [0,1]$ telle que :

- $\mathbb{P}(\Omega) = 1$
- pour toute famille $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles, la série $\sum_{n=0}^{+\infty} \mathbb{P}(A_n)$ est convergente

et :

$$\mathbb{P}\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

Proposition 16 Une probabilité $\mathbb{P}$ sur un univers $\Omega = \left\{\omega_n \ ; \ n \in \mathbb{N}\right\}$ dénombrable est uniquement caractérisée par une suite $(p_n)_{n\in\mathbb{N}}$ vérifiant :

$$\forall n \in \mathbb{N},\ p_n \geqslant 0 \text{ et } \sum_{n=0}^{+\infty} p_n = 1.$$

En effet, si $(p_n)_{n\in\mathbb{N}}$ est une telle suite, l'application :

$$\mathbb{P} : A \longmapsto \sum_{k=0 \ ; \ \omega_k \in A}^{+\infty} p_k$$

définit une probabilité sur Ω et si $\mathbb{P}$ est une probabilité sur Ω, alors la suite $(p_n = \mathbb{P}(\{\omega_n\}))_{n\in\mathbb{N}}$ vérifie les deux conditions ci-dessus.

Définition 22 On appelle ***espace probabilisé discret***, la donnée d'un univers Ω dénombrable et d'une probabilité sur Ω pour former $(\Omega, \mathbb{P})$.

Exemple 19 • En posant $\Omega = \mathbb{N}$ et :

$$\mathbb{P} : \left| \begin{array}{ccc} \mathscr{P}(\mathbb{N}) & \longrightarrow & [0,1] \\ A & \longmapsto & \displaystyle\sum_{k\in A} \frac{1}{2^{k+1}} \end{array} \right.$$

on a un espace probabilisé discret $(\Omega, \mathbb{P})$.

- Soit $s > 1$. On considère l'espace probabilisé $\Omega = \mathbb{N}^*$, $\mathscr{A} = \mathscr{P}(\mathbb{N}^*)$ et la probabilité définie par :

$$\forall k \in \mathbb{N}^*,\ \mathbb{P}(\{k\}) = \frac{1}{\zeta(s)} \cdot \frac{1}{k^s}.$$

Pour tout $m \in \mathbb{N}^*$, on note $A_m = m \cdot \mathbb{N}^*$. On note $\mathscr{Q}$ l'ensemble des nombres premiers.

1. Montrer que les événements $(A_m)_{m\in\mathscr{Q}}$ sont mutuellement indépendants.
2. En déduire que : $\zeta(s) = \displaystyle\prod_{p\in\mathscr{Q}} \left(1 - \frac{1}{p^s}\right)^{-1}$, en calculant $\mathbb{P}(\{1\})$ de deux manières différentes.

7.3 Variables aléatoires discrètes

Définition 23 Soit $(\Omega, \mathbb{P})$ un espace probabilisé discret. On appelle ***variable aléatoire discrète***, toute fonction $X : \Omega \longrightarrow \mathbb{R}$ telle que $X(\Omega)$ est au plus dénombrable et pour toute partie A de $\mathbb{R}$, l'ensemble $X^{[-1]}(A)$ appartient au domaine de définition de la probabilité $\mathbb{P}$.
On parle alors de loi de la variable pour désigner la probabilité μ définie par :

$$\forall A \in \mathscr{P}(X(\Omega)),\ \mu(A) = \mathbb{P}(X \in A).$$

7.4 Espérance d'une v.a. discrète

Définition 24 Soit $X : \Omega \longrightarrow \mathbb{R}$ une v.a. discrète. On pose $X(\Omega) = \{x_n\ ;\ n \in \mathbb{N}\}$, avec les x_n tous différents. On dit que la v.a. X ***admet une espérance*** si la série $\displaystyle\sum_n |x_n| \cdot \mathbb{P}(X = x_n)$ est convergente. On note alors dans ce cas :

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} x_n \cdot \mathbb{P}(X = x_n).$$