Chapitre 28 : Probabilités

1 Événements et variables aléatoires

1.1 Généralités

Définition 1.1.

- Un univers (fini) est un ensemble fini non vide Ω
- Un événement est une partie $A \subseteq \Omega$
- Une variable aléatoire (VA) est une application $X : \Omega \to E$ vers un ensemble E

1.2 Opérations

Image d'une VA par une application

Étant donné une VA $X : \Omega \to E$ est une application $f : E \to F$, on définit la VA image de X par f, $f(x)$ comme la composition $f \circ X : \Omega \to F$

Événements définis par une VA

Soit $X : \Omega \to E$

Pour toute partie $S \subseteq E$, on définit l'événement

$$(X \in S) = \{X \in S\} = \{\omega \in \Omega \mid X(\omega) \in S\} = X^{-1}[S]$$

Indicatrice d'un élément

Tout événement $A \subseteq \Omega$ définit une VA

$$\mathbb{1}_A : \begin{cases} \Omega \to \{0,1\} \\ \omega \mapsto \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \notin A \end{cases} \end{cases}$$

1.3 Expériences aléatoires

Considérons un exemple d'expérience aléatoire. On joue à pile ou face n fois de suite.

UNIVERS : On considère $\Omega = \{0,1\}$

Une issue est un résultat possible, càd ici une suite de n lancers.

ÉVÉNEMENTS : L'événement (au sens usuel) "le i-ème lancer donne pile" correspond à l'événement (au sens mathématique) $\pi_i = \{(b_1, \dots, b_n) \in \Omega \mid b_i = 1\}$ (ensemble à 2^{n-1} événements)

"Obtenir que des 'face' " $F = \{(0, 0, \dots, 0)\}$ (est élémentaire = singleton)

"Obtenir un nombre impair de 'pile' " $I = \{(b_1, \dots, b_n) \in \Omega \mid b_1 + \dots + b_n \equiv 1 (\text{mod } 2)\}$

VARIABLES ALÉATOIRES :

- $L_i : \begin{cases} \Omega \to \{0,1\} \\ (b_1, \dots, b_n) \mapsto b_i \end{cases}$ "est le résultat du i-ème lancer"
- $N : \begin{cases} \Omega \to [\![0, n]\!] \\ (b_1, \dots, b_n) \mapsto b_1, \dots, b_n \end{cases}$ est le nombre de "pile" obtenus
- $P : \begin{cases} \Omega \to [\![1, n]\!] \cup \{+\infty\} \\ (b_1, \dots, b_n) \mapsto \min \{i \in [\![1, n]\!] \mid b_i = 1\} \end{cases}$ est le rang de premier "pile"

  (avec la convention $\min \emptyset = +\infty$)

* $R : \begin{cases} \Omega \to P(\llbracket 1, n \rrbracket) \\ (b_1, \dots, b_n) \mapsto \{i \in \llbracket 1, n \rrbracket \mid b_i = 1\} \end{cases}$ est l'ensemble des rangs où l'on a obtenu pile

Lien entre ces objets :

* "Obtenir un nombre impair de 'pile' " et "Obtenir que 'face' " sont incompatibles : $I \cap F = \emptyset$
* $F = \overline{\pi_1 \cup \pi_2 \cup \dots \cup \pi_n} = \overline{\pi_1} \cap \dots \cap \overline{\pi_n}$
* Si $n = 3$, $I = (\pi_1 \cap \overline{\pi_2} \cap \overline{\pi_3}) \cup (\overline{\pi_1} \cap \pi_2 \cap \overline{\pi_3}) \cup (\overline{\pi_1} \cap \overline{\pi_2} \cap \pi_3) \cup (\pi_1 \cap \pi_2 \cap \pi_3)$
* On a $N = L_1 + L_2 + \dots + L_n$
* On a $N = |R|$

  Formellement, $N = \text{Card} \circ R$, où Card : $\begin{cases} P(\llbracket 1, n \rrbracket) \to \llbracket 0, n \rrbracket \\ T \mapsto |T| \end{cases}$

  On a en fait utilisé la notation $f(x)$ des VA images (càd qu'on a noté $\text{Card}(R)$ plutôt que $\text{Card} \circ R$)
* On a $L_1 + L_2 \leq N$ (en supposant $n \geq 2$)
* $P = \min(R)$
* $(N = 0) = F$
* $(N \equiv 1 (\text{mod } 2)) = (N \text{ impair}) = I$
* $\pi_1 \cap \dots \cap \pi_n = (N = n)$
* $(L_i = 1) = \pi_i$: en fait, $L_i = \mathbb{1}_{\pi_i}$
* $\overline{\pi_1} = (p \geq 2)$
* $(p = n) = \overline{\pi_1} \cap \overline{\pi_2} \cap \dots \cap \overline{\pi_{n-1}} \cap \pi_n = (R = \{n\}) = (N = 1, L_n = 1)$

2 Espaces probabilisés finis

2.1 Généralités

Définition 2.1.

* Une mesure de probabilités sur un univers Ω est une application $P : \mathcal{P}(\Omega) \to [0, 1]$ telle que :
  — $P(\Omega) = 1$
  — Pour tous $A, B \in \mathcal{P}(\Omega)$ disjoints, $P(A \sqcup B) = P(A) + P(B)$
* On appelle espace probabilisé (fini) tout couple (Ω, P), où Ω est un univers et P une mesure de probabilités du Ω

Proposition 2.2. Soit (Ω, P) un espace probabilisé fini (epf).
On a :

* $P(\emptyset) = 0$
* Croissance : pour tous $A, B \in \mathcal{P}(\Omega)$, $A \subseteq B \implies P(A) \leq P(B)$
* $\forall A \in \mathcal{P}(\Omega)$, $P(\overline{A}) = 1 - P(A)$
* Pour tous $A_1, \dots, A_r \in \mathcal{P}(\Omega)$ disjoints,

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

* $\forall A, B \in \mathcal{P}(\Omega)$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2.2 Formule des probabilités globales

Définition 2.3. Soit (Ω, P) un epf. Un système complet d'événements (scé) est une famille $(C_i)_{i=1}^r$ qui forme un recouvrement disjoint de Ω, càd telle que :

* Les C_i sont (deux à deux) disjoints : $\forall i, j \in [\![1, r]\!], i \neq j \implies C_i \cap C_j = \emptyset$
* $\bigcup\limits_{i=1} C_i = \Omega$

Théorème 2.4 (Formule des probabilités totales). Soit (Ω, P) un epf et $(C_i)_{i=1}^r$ un scé.
Alors $\forall A \in \mathcal{P}(\Omega)$, $P(A) = \sum\limits_{i=1}^{r} P(A \cap C_i)$

2.3 Loi d'une VA

Définition 2.5. Soit (Ω, P) un epf et $X : \Omega \to E$ une VA.
La loi de X est la donnée pour tout $S \subseteq E$ de la probabilité $P(X \in S) = P((X \in S))$

Proposition 2.6. Soit (Ω, P) un epf et $X : \Omega \to E$ une VA.
La loi de X est déterminée par les probabilités $P(X = x)$, pour x décrivant $\operatorname{im} X$
Plus précisément, pour tout $S \subseteq E$

$$P(X \in S) = \sum_{x \in S \cap \operatorname{im} X} P(X = x)$$

Définition 2.7. Soit (Ω, P) un epf et E un ensemble fini non vide.
Une VA $X : \Omega \to E$ suit la loi uniforme sur E si $\forall S \in \mathcal{P}(E)$, $P(X \in S) = \frac{|S|}{|E|}$
On note alors $X \sim U(E)$

Définition 2.8. Soit (Ω, P) un epf.
Une VA $X : \Omega \to \{0, 1\}$ suit le loi de Bernoulli de paramètre $p \in [\![0, 1]\!]$ si $P(X = 1) = p$
On note alors $X \sim B(p)$

Remarque importante : Si $A \subseteq \Omega$ est un événement, alors $\mathbb{1}_A \sim B(p)$, où $p = P(A)$

2.4 Couples de VA

Définition 2.9. Soit un epf et $X_1 : \Omega \to E_1$ et $X_2 : \Omega \to E_2$ deux VA.
Le loi conjointe de X_1 et X_2 est la loi de la VA

$$(X_1, X_2) : \begin{cases} \Omega \to E_1 \times E_2 \\ \omega \mapsto (X_1(\omega), X_2(\omega)) \end{cases}$$

Les lois de X_1 et X_2 sont appelées lois marginales de loi conjointe.

Proposition 2.10 (Calcul des marginales). Avec les notations de la définition :

$$\forall x_1 \in E_1,\ P(X_1 = x_1) = \sum_{x_2 \in \operatorname{im} X_2} P(X_1 = x_1, X_2 = x_2)$$

Remarque important :
La loi conjointe détermine les lois marginales, la réciproque est fausse : il y a plusieurs manières de coupler des lois.
Notamment, pour tout $p \in \left[0, \frac{1}{2}\right]$ on a

$k \setminus l$	0	1
0	p	$\frac{1}{2} - p$
1	$\frac{1}{2} - p$	p

qui constitue un couplage de $B\left(\frac{1}{2}\right)$ avec elle-même.

Trois cas particuliers :

$p = \frac{1}{2}$

$\frac{1}{2}$	0
0	$\frac{1}{2}$

On a $P(X_1 = X_2) = 1$

X_1 et X_2 sont égales (presque sûrement).

$p = 0$

0	$\frac{1}{2}$
$\frac{1}{2}$	0

$X_1 + X_2 = 1$ presque sûrement.

$p = \frac{1}{4}$

$\frac{1}{4}$	$\frac{1}{4}$
$\frac{1}{4}$	$\frac{1}{4}$

$P(X_1 = X_2) = \frac{1}{2}$

Connaître le résultat de X_1 ne donne aucune information sur celui de X_2

On dira que X_1 et X_2 sont indépendantes.

2.5 Construction d'espaces probabilisés finis

Définition 2.11. Une distribution de probabilités sur un univers fini Ω est une famille $(p_\omega)_{\omega\in\Omega}$ de réels ≥ 0 tels que $\sum_{\omega\in\Omega} p_\omega = 1$

Proposition 2.12. Soit $(p_\omega)_{\omega\in\Omega}$ une distribution de probabilités sur un univers fini Ω

Alors il existe une unique mesure de probabilités P sur Ω telle que $\forall\omega \in \Omega$, $P(\{\omega\}) = p_\omega$

2.6 Modélisation(s) d'expériences aléatoires

Imaginons qu'on veuille modéliser l'expérience constituant à tirer aléatoirement une carte dans un jeu de 52 cartes.

Modélisation 1 (normale) : On prend $\Omega = \{2, 3, ..., V, D, R, A\} \times \{P, C, K, T\}$ l'ensemble des 52 cartes, muni de la probabilité uniforme.

L'événement (au sens usuel) "tirer un coeur" correspond à l'événement $\heartsuit = \{(2, C), ..., (A, C)\}$

donc $P(\heartsuit) = \frac{|\heartsuit|}{|\Omega|} = \frac{13}{52} = \frac{1}{4}$

Modélisation 2 (un peu tordue) : On commence par mélanger le jeu. On va poser $J = \{2, ..., R, A\} \times \{P, C, K, T\}$ puis on pose Ω l'ensemble des permutations de J (càd une 52-liste sans répétition de J), muni de la mesure de probabilités uniforme.

L'événement (au sens usuel) "tirer un coeur" correspond à l'événement

$\heartsuit' = \{(C_1, C_2, ..., C_{52}) \in \Omega' \mid C_1 \in \{(2, C), ..., (A, C)\}\}$

On a $|\Omega'| = 52!$ Calculons $|\heartsuit'|$

Pour construire un élément de $\heartsuit'$:

- ∗ On choisit une première carte (un coeur) : 13 possibilités.
- ∗ ―――――― 2è carte (différente de la 1ère) : 51 possibilité.
- ∗ ―――――― 3è carte (différente des précédentes) : 50 possibilités.

Par principe de multiplication, $|\heartsuit'| = 13 \times 51 \times 50 \times ... \times 1 = 13 \times 51!$

Donc $P(\heartsuit') = \frac{13\times 51!}{52!} = \frac{13}{52} = \frac{1}{4}$

En théorie des possibilités, l'epf joue un rôle de seconde plan : dans l'exemple précédent, les détails de Ω n'importent pas. Ce qui compte est qu'il existe une VA donnant le résultat du tirage. Que l'on prenne

$$X : \begin{cases} \Omega \to J \\ C \mapsto C \end{cases} \quad \text{ou} \quad X' : \begin{cases} \Omega' \to J \\ (C_1, ..., C_{52} \mapsto C_1 \end{cases}$$

On a une VA X ou $X' \sim U(J)$ et c'est ce qui compte.

2.7 Vers les espaces probabilisés généraux

Définition 2.13. Une tribu (ou une σ-algèbre) sur une ensemble Ω est une partie $a \in \mathcal{P}(\Omega)$ contenant Ω ($\Omega \in a$) stable par passage au complémentaire (si $A \in a, \overline{A} \in a$) et union dénombrable (si $A_0, A_1, ... \in a$, $\bigcup\limits_{n\in\mathbb{N}} A_n \in a$)

Définition 2.14. Un espace probabilisé est un triplet (Ω, a, P) où :
* L'univers Ω est un ensemble non vide.
* a est une tribu sur Ω
* La mesure de probabilités $P : a \to [0,1]$ vérifie :
  — $P(\Omega) = 1$
  — Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de a, deux à deux disjoints, alors $P\left(\bigsqcup\limits_{n\in\mathbb{N}} A_n\right) = \sum\limits_{n\in\mathbb{N}} P(A_n)$ (σ-additivité)

Théorème 2.15 (culturel)**.** Il existe un espace probabilisé $([0,1], B, \lambda)$ où :
* B est une tribu sur $[0,1]$ (tribu des boréliens) contenant les intervalles.
* $\lambda : B \to [0,1]$ est une mesure de probabilités telle que $\forall 0 \leq a < b \leq 1, \lambda(]a,b[) = b - a$ appelée mesure de Lebesgue.

3 Indépendance

Dans cette section, (Ω, P) est un épf.

3.1 Deux événements

Définition 3.1. Soit $A, B \in \mathcal{P}(\Omega)$ deux événements.
On dit que A et B sont indépendants si $P(A \cap B) = P(A)P(B)$

Proposition 3.2. Soit $A, B \in \mathcal{P}(\Omega)$ indépendants.
Alors $\overline{A}$ et B sont indépendants.

3.2 Variables aléatoires

Définition 3.3. Soit $X_1 : \Omega \to E_1$ et $X_2 : \Omega \to E_2$
On dit que X_1 et X_2 sont indépendantes (et on note $X_1 \perp\!\!\!\perp X_2$) si $\forall S_1 \in \mathcal{P}(E_1), \forall S_2 \in \mathcal{P}(E_2)$

$$P(X_1 \in S_1, X_2 \in S_2) = P(X_1 \in S_1)P(X_2 \in S_2)$$

Plus généralement, soit $X_i : \Omega \to E_i$ des VA ($i \in [\![1, n]\!]$).
On dit qu'elles sont indépendantes si $\forall S_1 \in \mathcal{P}(E_1), \dots, \forall S_n \in \mathcal{P}(E_n)$

$$P(X_1 \in S_1, \dots, X_n \in S_n) = P(X_1 \in S_1) ... P(X_n \in S_n)$$

Remarque : Si n VA sont indépendantes (ou, pour insister mutuellement indépendantes), elles sont indépendantes deux à deux, càd $\forall i \neq j \in [\![1, n]\!]$, $X_i \perp\!\!\!\perp X_j$

Théorème 3.4. Soit $X_1 : \Omega \to E_1, \dots, X_n : \Omega \to E_n$ des VA.
Alors $X_1, \dots, X_n$ sont indépendantes ssi

$$\forall x_1 \in E_1, \dots, \forall x_n \in E_n,\ P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) ... P(X_n = x_n)$$

Proposition 3.5. Soit $A, B \in \mathcal{P}(\Omega)$
Alors A et B sont indépendants ssi $\mathbb{1}_A \perp\!\!\!\perp \mathbb{1}_B$

3.3 Plusieurs événements

Définition 3.6. Soit $A_1, \dots, A_n \in \mathcal{P}(\Omega)$
Alors ils sont <u>indépendants</u> si $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ sont indépendants.

Théorème 3.7. Soit $A_1, \dots, A_n \in \mathcal{P}(\Omega)$
Alors $A_1, \dots, A_n$ sont indépendants ssi

$$\forall I \subseteq [\![1, n]\!],\ P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i)$$

3.4 Stabilité de l'indépendance

Théorème 3.8 (Transfert de l'indépendance)**.**
Soit $X_1 : \Omega \to E_1, \dots, X_n : \Omega \to E_n$ des VA indépendantes.
Soit $f_1 : E_1 \to F_1, \dots, f_n : E_n \to F_n$ des applications.
Alors $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes.

Corollaire 3.9. Si $X_1 : \Omega \to E_1, \dots, X_n : \Omega \to E_n$ sont indépendantes et que, pour tout $i \in [\![1, n]\!]$, A_i est un événement défini par X_i, alors $A_1, \dots, A_n$ sont indépendants.

Théorème 3.10 (Lemme de coalitions)**.** Soit $X_1 : \Omega \to E_1, \dots, X_n : \Omega \to E_n$ indépendantes et $r \in [\![0, n]\!]$
Alors les VA

$$Y : \begin{cases} \Omega \to E_1 \times \dots \times E_r \\ \omega \mapsto (X_1(\omega), \dots, X_r(\omega)) \end{cases} \qquad Z : \begin{cases} \Omega \to E_{r+1} \times \dots \times E_n \\ \omega \mapsto (X_{r+1}(\omega), \dots, X_n(\omega)) \end{cases}$$

sont indépendantes.

3.5 Indépendance et mesure uniforme

Théorème 3.11. Soit $E_1, \dots, E_n$ des ensembles finis non vides et on munit $\Omega = E_1 \times \dots \times E_n$ de la probabilité uniforme.
Alors

$$X_1 : \begin{cases} \Omega \to E_1 \\ (x_1, \dots, x_n) \mapsto x_1 \end{cases} \quad , \dots , \quad X_n : \begin{cases} \Omega \to E_n \\ (x_1, \dots, x_n) \mapsto x_n \end{cases}$$

sont indépendantes.

3.6 Loi binomiale

Définition 3.12. Une VA $X : \Omega \to [\![0, n]\!]$ <u>suit la loi binomiale</u> de paramètre n et $p \in [0, 1]$ ($X \sim B(n, p)$) si $\forall k \in [\![0, n]\!]$

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

Théorème 3.13. Soit $X_1, \dots, X_n : \Omega \to \{0, 1\}$ indépendantes suivant la loi de Bernoulli $B(p)$ (pour $p \in [0, 1]$)
Alors $X_1 + \dots + X_n \sim B(n, p)$

4 Probabilités conditionnelles

On se place dans un espace probabilisé fini (Ω, P)

4.1 Définition

Définition 4.1. Soit $B \in \mathcal{P}(\Omega)$ un événement non négligeable (càd $P(B) > 0$). Soit $A \in \mathcal{P}(\Omega)$
On définit la probabilité conditionnelle de A sachant B

$$P_B(A) = P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Proposition 4.2. Soit $B \in \mathcal{P}(\Omega)$ tel que $P(B) > 0$
Alors P_B est une mesure de probabilités sur Ω pour laquelle B est presque sûr.

Attention ! La notation $(A \mid B)$ seule n'a aucun sens.
Par exemple, la phrase française " Sachant que $X = Y$, $'X = 0'$ équivaut à $'Y = 0'$ " ne se traduit pas en $(X = 0 \mid X = Y) = (Y = 0)$ qui n'a aucun sens.
Elle ne se traduit pas non plus en $P(X = 0 \mid X = Y) = P(Y = 0)$ (qui a un sens mais qui est faux en général).
La bonne traduction est $(X = Y, X = 0) = (X = Y, Y = 0)$

Proposition 4.3. Soit $A, B \in \mathcal{P}(\Omega)$ tels que $P(B) > 0$
Alors A et B sont indépendants ssi $P(A \mid B) = P(A)$

Définition 4.4. Soit $X : \Omega \to E$ une VA et $A \in \mathcal{P}(\Omega)$ un événement non négligeable.
La loi conditionnelle de X sachant A est la donnée, pour tout $S \subseteq E$ de $P(X \in S \mid A)$

4.2 Probabilités composées et probabilités totales

Proposition 4.5 (Formule des probabilités composées)**.**
Soit $A_1, \dots, A_n \in \mathcal{P}(\Omega)$ tels que $\bigcap_{i=1}^{n-1} A_i$ soit non négligeable.
On a

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_2 \cap A_1)...P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right)$$

Remarque : Si $\bigcap_{i=1}^{n-1} A_i$ est négligeable, la formule reste correcte si on pose la convention $0\times$ inepte $= 0$

Proposition 4.6 (Formule des probabilités totales)**.** Soit $A \in \mathcal{P}(\Omega)$ et $C_1, \dots, C_n \in \mathcal{P}(\Omega)$ des événements (non négligeables) formant un système complet d'événements.
Alors

$$P(A) = \sum_{i=1}^{n} P(A \cap C_i) = \sum_{i=1}^{n} P(A \mid C_i)P(C_i)$$

Corollaire 4.7. Soit $A, B \in \mathcal{P}(\Omega)$ (tels que $P(B) \in]0, 1[$)
Alors $P(A) = P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B})$

4.3 Formules de Bayes

Proposition 4.8. Soit A, B deux événements non négligeables.
On a

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

Proposition 4.9. Soit A un événement et $(C_i)_{i=1}^n$ un scé, tous non négligeables.
Alors, pour tout $i \in [\![1, n]\!]$

$$P(C_i \mid A) = \frac{P(A \mid C_i)P(C_i)}{P(A)} = \frac{P(A \mid C_i)P(C_i)}{\sum\limits_{j=1}^{n} P(A \mid C_j)P(C_j)}$$

5 Espérance

Dans toute cette section, X sera une variable aléatoire à valeurs réelle ou complexe, voire à valeurs dans un espace vectoriel réel.

Définition 5.1. Soit X une VA complexe.
On définit son espérance :

$$E(X) = \sum_{x \in \mathrm{im}(X)} P(X = x)x$$

Remarque : $E(X)$ ne dépend que de la loi de X :
Si X' est une VA de même loi que X (ce que l'on note $X \sim X'$), alors $E(X) = E(X')$

Théorème 5.2. L'espérance possède les propriétés suivantes :
Linéarité : L'espérance est une forme linéaire $\mathbb{C}^\Omega \to \mathbb{C}$ (ou $\mathbb{R}^\Omega \to \mathbb{R}$)
Concrètement, si X, Y sont deux VA complexes et $\lambda \in \mathbb{C}$, alors $E(X + \lambda Y) = E(X) + \lambda E(Y)$
Positivité : Si $X : \Omega \to \mathbb{R}_+$ est une VA positive, on a $E(X) \geq 0$
Croissante : Si $X, Y : \Omega \to \mathbb{R}$ sont deux VA telles que $X \leq Y$ (càd $\forall \omega \in \Omega$, $X(\omega) \leq Y(\omega)$), alors $E(X) \leq E(Y)$
Inégalité triangulaire : Si $X : \Omega \to \mathbb{C}$ est une VA complexe, $|E(X)| \leq E(|X|)$

5.1 Définition

Lemme 5.3. Si X est une VA complexe, $E(X) = \sum\limits_{\omega \in \Omega} P(\{\omega\})X(\omega)$

Proposition 5.4. Soit $X \sim B(n, p)$
Alors $E(X) = np$

Théorème 5.5 (Formule de transfert)**.** Soit $X : \Omega \to E$ et $f : E \to \mathbb{C}$ une application.
Alors

$$E(f(x)) = \sum_{x \in \mathrm{im}(X)} P(X = x)f(x)$$

Corollaire 5.6 (Formule de transfert bivariée)**.** Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux VA et $f : E \times F \to \mathbb{C}$
Alors

$$E(f(X, Y)) = \sum_{\substack{x \in \mathrm{im}\, X \\ y \in \mathrm{im}\, Y}} P(X = x, Y = y)f(x, y)$$

Définition 5.7. Une VA $X : \Omega \to \mathbb{C}$ est centrée si $E(X) = 0$

5.2 Applications

Théorème 5.8 (Formule du crible ou d'inclusion-exclusion)**.** Soit $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$
On a

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} P\left(A_{i_1} \cap \ldots \cap A_{i_k}\right)$$

Corollaire 5.9. Soit Ω un ensemble non vide et $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$
Alors

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} |A_{i_1} \cap \ldots \cap A_{i_k}|$$

5.3 Indépendance

Théorème 5.10. Soit $X, Y : \Omega \to \mathbb{C}$ deux VA indépendantes.
Alors $E(XY) = E(X)E(Y)$
Évidemment, cela se généralise à n VA.

5.4 Espérance conditionnelle

Définition 5.11. Soit $X : \Omega \to \mathbb{C}$ une VA et $B \in \mathcal{P}(\Omega)$ non négligeable.
On définit l'espérance conditionnelle de X sachant B :

$$E(X \mid B) = \sum_{i \in \operatorname{im} X} P(X = x \mid B)x$$

Théorème 5.12 (Formule des espérances totales)**.** Soit $X : \Omega \to \mathbb{C}$ une VA et $(C_i)_{i=1}^n$ un scé non négligeables.
Alors

$$E(X) = \sum_{i=1}^{n} E(X \mid C_i)P(C_i)$$

6 Moments d'ordre deux

Remarque : Pour $k \in \mathbb{N}$, l'espérance $E(X^k)$ s'appelle le k-ième moment de X

6.1 Variance et écart-type

Définition 6.1. Soit $X : \Omega \to \mathbb{R}$ une VA réelle.
On définit sa variance :

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

On définit l'écart-type :

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

Proposition 6.2 (Formule de König-Huygens)**.** Soit $X : \Omega \to \mathbb{R}$ Alors

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

Exemple important : Si $X \sim B(p)$, on a $E(X) = p$
Donc $V(X) = p(1 - p)$

Proposition 6.3. Soit $X : \Omega \to \mathbb{R}$ et $a, b \in \mathbb{R}$
On a $V(aX + b) = a^2V(X)$

Proposition 6.4. Soit $X : \Omega \to \mathbb{R}$
On a $V(X) = 0 \implies X$ est constante presque sûrement.
(et signifie qu'il existe $c \in \mathbb{R}$ tel que $P(X = c) = 1$

6.2 Covariance

Définition 6.5. Soit $X, Y : \Omega \to \mathbb{R}$
On définit la covariance de X et Y :

$$\operatorname{cov}(X, Y) = E\left((X - EX)(Y - EY)\right)$$

Proposition 6.6. On a

$$\operatorname{cov}(X, Y) = E(XE) - EX \cdot EY$$

Remarque : Si $X \perp\!\!\!\perp Y$, on a donc $\operatorname{cov}(X,Y) = 0$. On dit que X et Y sont décorrélées.
Si $\operatorname{cov}(X,Y) > 0$ (resp. < 0), on dit que X et Y sont positivement (resp. négativement) corrélées.

Remarque : La covariance est presque un produit scalaire (il manque le caractère défini)
On garde donc toutes les propriétés liées à la bilinéarité et au caractère positif, notamment :

* L'identité remarquable $V(X+Y) = V(X) + 2\operatorname{cov}(X,Y) + V(Y)$
* L'inégalité de Cauchy-Schwarz (sans le cas d'égalité)

Théorème 6.7. Soit $X, Y : \Omega \to \mathbb{R}$
On a
$$\operatorname{cov}(X,Y) \leq |\operatorname{cov}(X,Y)| \leq \sqrt{V(X)}\sqrt{V(Y)}$$

Théorème 6.8. Soit $X_1, \dots, X_n : \Omega \to \mathbb{R}$ deux à deux décorrélées.
Alors
$$V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n)$$

Corollaire 6.9. Soit $X \sim B(n,p)$
Alors $V(X) = np(1-p)$

7 Inégalités de concentration

7.1 Inégalité de Markov

Théorème 7.1. Soit $X : \Omega \to \mathbb{R}_+$ (à valeurs positives)
Alors $\forall a > 0$
$$P(X \geq a) \leq \frac{E(X)}{a}$$

7.2 Inégalité de Bienaymé-Tchebychev

Théorème 7.2. Soit $X : \Omega \to \mathbb{R}$
Alors $\forall a > 0$
$$P(|X - EX| \geq a) \leq \frac{V(X)}{a^2}$$

7.3 Loi faible des grands nombres

Théorème 7.3. Soit $X_1, \dots, X_n : \Omega_n \to \mathbb{R}$ indépendantes et de même loi.
Notons μ leur espérance.
On définit
$$S_n = \frac{X_1 + \dots + X_n}{n}$$
Alors $\forall \varepsilon > 0$
$$P(|S_n - \mu| \geq \varepsilon) \xrightarrow[n \to +\infty]{} 0$$

7.4 Théorème d'approximation de Weierstrass

Théorème 7.4. Soit $f \in C^0([0,1])$
Alors il existe une suite de fonctions polynomiales $(P_n)_{n \in \mathbb{N}^*}$ qui converge uniformément vers f