Probabilité

I. Espace probabilisé

1) Espace probabilisable (théorique). Cas général : Ω peut être infini non dénombrable

Idée : Ω est l'univers, c'est à dire l'ensemble de tous les résultats possibles d'une expérience aléatoire. On veut probabiliser Ω en définissant $P(A) \in [0,1]$ pour certains $A \subset \Omega$. Pour des raisons profondes, il n'est en général pas possible de définir $P(A)$ pour toute partie de Ω. On se restreint à une classe raisonnable d'événements appelée **tribu**.

Le cas le plus fréquent : si Ω est dénombrable, on pourra prendre comme tribu $\mathcal{P}(\Omega)$.

Définition :

Soit Ω un ensemble quelconque (fini ou infini).

Une tribu sur Ω est un ensemble $\mathcal{A}$ de parties de Ω (c'est à dire $\mathcal{A} \subset \mathcal{P}(\Omega)$) telle que :

i) $\Omega \in \mathcal{A}$

ii) Si $A \in \mathcal{A}$, alors on a : $\overline{A} = \Omega \backslash A \in \mathcal{A}$

iii) Si $(A_n)_{n\in\mathbb{N}}$ est une suite de $\mathcal{A}$, alors on a : $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Les éléments de $\mathcal{A}$ sont appelés les évènements.

On dit que $(\Omega, \mathcal{A})$ est un espace probabilisable.

Théorème :

Soit $(\Omega, \mathcal{A})$ un espace probabilisable, alors on a :

1) $\emptyset \in \mathcal{A}$

2) Soit $(A_n)_{n\in\mathbb{N}}$, une suite de $\mathcal{A}$, alors on a : $\displaystyle\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

3) Soient $n \in \mathbb{N}^*$ et $(A_1, A_2, \dots, A_n) \in \mathcal{A}^n$, alors on a : $A_1 \cup A_2 \cup \dots \cup A_n \in \mathcal{A}$ et $A_1 \cap A_2 \cap \dots \cap A_n \in \mathcal{A}$.

4) Soit $(A,B) \in \mathcal{A}^2$, alors $A \backslash B \in \mathcal{A}$.

Exemple :

a) $\mathcal{A} = \mathcal{P}(\Omega)$ est une tribu sur Ω.

b) $\mathcal{A} = \{\emptyset, \Omega\}$ est une tribu sur Ω (on l'appelle tribu triviale).

c) Soit $A \subset \Omega$, déterminer la plus petite tribu contenant A.

d)* Soit $(A_n)_{n\in\mathbb{N}}$ une suites d'évènements d'un espace probabilisable $(\Omega, \mathcal{A})$. Montrer que l'ensemble B des éléments de Ω qui appartiennent à une infinité d'évènements A_n est un évènement de $(\Omega, \mathcal{A})$.

On pourra remarquer que : si $X_\omega = \{k \in \mathbb{N} / \omega \in A_k\}$, alors $[X_\omega$ est infini $\Longleftrightarrow \forall n \in \mathbb{N},\ \exists k \geq n / k \in X_\omega]$

Définition :

Soit $(\Omega, \mathcal{A})$ un espace probabilisable, et soit $I \subset \mathbb{N}$.

On appelle $\boxed{\text{système complet d'évènements}}$, une famille $(A_i)_{i\in I}$ où les $A_i \in \mathcal{A}$ qui vérifie :

• $\Omega = \displaystyle\bigcup_{i\in I} A_i$ et • $\forall (i,j) \in I^2 : i \neq j \Longrightarrow A_i \cap A_j = \emptyset$ (on dit que les $(A_i)_{i\in I}$ sont incompatibles) .

Cas particulier important :

Soit $A \in \mathcal{A}$, alors $(A, \overline{A})$ est un S.C.E. (système complet d'évènements)

2) Espace probabilisé :

Définition : (le plus souvent, la tribu est $\mathcal{P}(\Omega)$)

Une probabilité sur l'espace probabilisable $(\Omega, \mathcal{A})$, est une application $\mathbf{P}$ définie sur $\mathcal{A}$ et à valeurs dans $[0,1]$ telle que :

i) $\mathbf{P}(\Omega) = 1$

ii) propriété de **σ-additivité** : Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements 2 à 2 incompatibles de $\mathcal{A}$, alors on a :

la série $\left(\sum \mathbf{P}(A_n)\right)_{n\in\mathbb{N}}$ converge et $\quad \mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n)$

On dit alors que $(\Omega, \mathcal{A}, \mathbf{P})$ est un espace **probabilisé**.

Remarque : Par la suite, on dira que 2 événements sont disjoints ou incompatibles.

Propriétés :

Soit $(\Omega, \mathcal{A}, \mathbf{P})$, un espace probabilisé, alors on a :

1) $\mathbf{P}(\emptyset) = 0$.

2) Soit $(A_1, A_2, \ldots, A_n) \in \mathcal{A}^n$ $(n \geqslant 1)$, 2 à 2 disjoints, alors on a :

$\mathbf{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \mathbf{P}(A_1) + \cdots + \mathbf{P}(A_n)$.

3) **Monotonie** : Soit $(A, B) \in \mathcal{A}^2$ tel que $A \subset B$, alors on a : $\mathbf{P}(A) \leqslant \mathbf{P}(B)$ et

$\mathbf{P}(B \backslash A) = \mathbf{P}(B) - \mathbf{P}(A)$. En particulier $\mathbf{P}\left(\overline{A}\right) = 1 - \mathbf{P}(A)$

4) Soient $(A, B) \in \mathcal{A}^2$, quelconques, alors on a : $\quad \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

5) **continuité croissante** :

Soit $(A_n)_{n\in\mathbb{N}}$, une suite croissante d'évènements de $\mathcal{A}$ (i.e. $A_n \subset A_{n+1}$ pour tout n dans $\mathbb{N}$),

alors on a : $\quad \mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbf{P}(A_n)$.

6) **continuité décroissante** :

Soit $(A_n)_{n\in\mathbb{N}}$, une suite décroissante d'évènements de $\mathcal{A}$ (i.e. $A_{n+1} \subset A_n$ pour tout n dans $\mathbb{N}$),

alors on a : $\quad \mathbf{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbf{P}(A_n)$.

7) **sous-additivité (ou inégalité de Boole)** :

Soit $(A_n)_{n\in\mathbb{N}}$, une suite quelconque d'évènements de $\mathcal{A}$, alors on a :

$\mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=0}^{+\infty} \mathbf{P}(A_n)$ (la somme de droite étant éventuellement infinie)

Remarque : *en proba*, si $\sum \mathbf{P}(A_n)$ diverge, on dit que $\sum_{n=0}^{+\infty} \mathbf{P}(A_n) = +\infty$.

Remarque 2 : cas particulier de la propriété 7 : si $(A_1, A_2, \ldots, A_n) \in \mathcal{A}^n$ $(n \geqslant 1)$, on a :

$\mathbf{P}(A_1 \cup A_2 \cup \cdots \cup A_n) \leqslant \mathbf{P}(A_1) + \cdots + \mathbf{P}(A_n)$.

Définition : 2 évènements A et B sont indépendants si $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

Probabilité, suite

3) Conditionnement : $(\Omega, \mathcal{A}, \mathbf{P})$ est un espace **probabilisé**

1 • **Définition d'une proba conditionnelle** :

$\boxed{\text{si } P(B) \neq 0, \text{ alors } P(A/B) = \dfrac{P(A \cap B)}{P(B)}}$ (notée aussi $P_B(A)$).

thm : P_B est une probabilité.

Conséquence : $P(\overline{A}/B) = ...$

Propriété : Si $P(B) \neq 0$ alors :[A et B sont indépendants $\Longleftrightarrow P(A/B) = P(A)$]

2 • **Formule des proba composées** : Si $P(A_1 \cap A_2 \cap ... \cap A_{n-1}) \neq 0$, alors

$\boxed{P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)P(A_2/A_1)P(.... \qquad ...P(A_n/A_1 \cap A_2 \cap ... \cap A_{n-1})}$

3 • **Formule des proba totales** : $I \subset \mathbb{N}$.

Si $(A_i)_{i\in I}$ est un système complet d'évènements, alors pour tout évènement B,

1ère forme : $\boxed{P(B) = \sum_{i\in I} P(A_i \cap B)}$ (pour les lois marginales de couple)

2ème forme : $\boxed{P(B) = \sum_{i\in I} P(A_i)P(B/A_i)}$ avec pour convention, si $P(A_i) = 0$, $P(A_i)P(B/A_i) = 0$.

Application : Soit A un évènement, alors $(A, \overline{A})$ est un système complet d'événements et la formule des

proba totales s'écrit :

4 • **Formule de Bayes** : (la formule à remonter dans le temps)

Si $P(A) \neq 0$ et $P(B) \neq 0$, alors $\boxed{P(A/B) = \dfrac{P(A)}{P(B)}P(...}$

II. Variables aléatoires :

Soit $(\Omega, \mathcal{A}, P)$, un espace probabilisé et E un ensemble quelconque non vide. (en général, $E = \mathbb{R}$).

1) Généralités :

Définition :

On appelle **variable aléatoire discrète** (**VAD**) sur $(\Omega, \mathcal{A}, P)$ toute application $X : \Omega \longrightarrow E$ vérifiant :

i) L'image $X(\Omega)$ est une partie au plus dénombrable de E (: finie ou dénombrable).

ii) Pour tout $x \in E$, l'ensemble $X^{-1}(\{x\}) \in \mathcal{A}$ (c'est à dire $X^{-1}(\{x\})$ est un événement).

Remarques :

$*$ Si X est une **VAD** et $x \notin X(\Omega)$, alors, $X^{-1}(\{x\}) = \emptyset$ est toujours dans $\mathcal{A}$.

On en déduit que le point ii) de la Définition est équivalent à : Pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathcal{A}$.

$*$ Si $E = \mathbb{R}$, X est une variable aléatoire réelle discrète (**VARD**).

Rappel : $X^{-1}(B) = ...$ et $X^{-1}(\{x\}) = ...$

Notations : En probabilité, on accepte d'alléger les notations ensemblistes habituelles, notamment pour les images réciproques : si $X : \Omega \longrightarrow E$ est une variable aléatoire discrète, et

si $B \subset E$, on notera : $\boxed{X^{-1}(B) = (X \in B)}$ et si $x \in E$, on notera : $\boxed{X^{-1}(\{x\}) = (X = x)}$.

Exemple de ce paragraphe : On lance une pièce 2 fois, alors $\Omega = ...$

Sur $(\Omega, \mathcal{P}(\Omega))$, on définit $X : \Omega \longrightarrow \mathbb{R}$

$\omega \longmapsto$ le nombre de piles de ω.

Dans la pratique, on définit X par : $X =$ le nombre de piles obtenus (ou $X(\omega) =$ le nombre de piles obtenus).

On a alors $X(\Omega) = ...$ et l'événement $(X = 1) = ...$

$(X = 0, X = 1, X = 2)$ est ...

Rem : on note aussi $(P, F) = P_1 \cap F_2$ (ou encore P_1F_2)

Propriété 1 :

Soit $X : \Omega \longrightarrow E$ une **VAD** sur $(\Omega, \mathcal{A})$, on peut donc noter $X(\Omega) = \{x_n \ , n \in I\}$, avec $I \subset \mathbb{N}$.
Alors $\{(X = x_n)_{n\in I}\}$ est un système complet d'évènements.

Autres exemples :

a) On lance un dé jusqu'à l'obtention d'un 4 et on s'intéresse à la variable Y qui compte le nombre de lancers nécessaires pour obtenir le premier 4. Que vaut $Y(\Omega)$?

b) On lance deux dés, alors $\Omega = \{1, 2, 3, 4, 5, 6\}^2$, et Z définie par $Z((i, j)) = i + j$ est une variable aléatoire discrète et $Z(\Omega) =$... , on dira que Z est la somme des résultats des 2 dés, et on note $Z = ...$

Propriété 2 :

Soient X et Y 2 **VARD** sur $(\Omega, \mathcal{A})$, alors $X + Y$, XY , $\max(X, Y)$, $\min(X, Y)$ sont des **VARD**.

Définition de la fonction indicatrice de $A \in \mathcal{P}(\Omega)$, on note $\mathbb{1}_A : \Omega \longrightarrow \mathbb{R}$

$\omega \longmapsto 1$ si $\omega \in A$

$\omega \longmapsto 0$ si $\omega \notin A$.

$\mathbb{1}_A$ est appelée la fonction indicatrice de A : c'est une **VAD**, et on a $\mathbb{1}_A(\Omega) = ...$,

$\forall x \in \mathbb{R}, \quad \mathbb{1}_A^{-1}(\{x\}) = \emptyset$ si ... $\qquad \mathbb{1}_A^{-1}(\{1\}) = ...$ $\qquad$ et $\mathbb{1}_A^{-1}(\{0\}) = ...$

Définition de la loi de probabilité d'une variable aléatoire discrète

Soit X une **VAD** sur $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé.
La loi de probabilité de X est la donnée de :
- $X(\Omega)$, de la forme $\{x_n / n \in I\}$ où $I \subset \mathbb{N}$
- et de $\mathbf{P}(X = x_n)$ pour tout $x_n \in X(\Omega)$

Remarque : $\forall n \in I, \quad \mathbf{P}(X = x_n) \geq 0 \quad$ et $\quad \sum_{n\in I} \mathbf{P}(X = x_n) = 1$.

Exemple de ce paragraphe :

On lance une pièce 2 fois, et $X =$ le nombre de piles obtenus. Déterminer la loi de X :

Propriété : (admise)

Soit $X : \Omega \longrightarrow E$ une **VAD** sur l'espace probabilisable $(\Omega, \mathcal{A})$. On pose $X(\Omega) = \{x_n \ , n \in I\}$, avec $I \subset \mathbb{N}$.
Soit $(p_n)_{n\in I}$ une suite de réels positifs vérifiant $\sum_{n\in I} p_n = 1$.
Alors il existe une probabilité $\mathbf{P}$ sur $(\Omega, \mathcal{A})$ telle que, pour tout $n \in I$, $\quad \mathbf{P}(X = x_n) = p_n$.

Définition

2 **VAD** X et Y sont indépendantes si
$\forall x \in X(\Omega), \forall y \in Y(\Omega), \quad (X = x)$ et $(Y = y)$ sont indépendantes (c'est à dire : ...

Probabilité, fin

Propriétés (admises) :

Si X et Y sont indépendantes alors :

• $\forall A \subset X(\Omega),\ \forall B \subset Y(\Omega), \quad \mathbf{P}(X \in A \text{ et } Y \in B) = \mathbf{P}(X \in A)\mathbf{P}(Y \in B)$

• Toute fonction de X est indépendantes de toute fonction de Y.

Exemples :

1. Si X et Y sont indépendantes, on obtient $\mathbf{P}(X \leq x \text{ et } Y \leq y) =$...
2. Si X et Y sont indépendantes alors X^2 et e^Y sont indépendantes.

Quelques méthodes à connaitre :

• max $(X, Y) \leq x \Longleftrightarrow X \leq x$ **et** $Y \leq x$ (inégalité large ou stricte)

• min $(X, Y) \geq x \Longleftrightarrow X \geq x$ **et** $Y \geq x$

• Si $X(\Omega) = Y(\Omega) \subset \mathbb{N}$, alors $(X = Y) = (X = 0 \text{ et } Y = 0)$ ou $(X = 1 \text{ et } Y = 1)$ ou ....

$$\boxed{(X = Y) = \bigcup_{k \in \mathbb{N}} ((X = k) \cap (Y = k))}$$

Propriété : Si $X(\Omega) \subset \mathbb{Z}$, alors $\mathbf{P}(X = k) = \mathbf{P}(X \leqslant k) - \mathbf{P}(X \leqslant k - 1)$.

2) Lois classiques :

1. **Loi uniforme sur $[\![1, n]\!]$** :

$X \sim U([\![1, n]\!])$ si $X(\Omega) =$ .... et $P(X = k) =$ ...

$E(X) =$ ... et $V(X) =$

2. **Loi de Bernoulli de paramètre $p \in]0, 1[$** : une épreuve de Bernoulli a 2 issus possibles : succès et échec (le succès correspondant à $(X = 1)$)

$X \sim B(p)$ si $X(\Omega) =$ .... et $\begin{cases} P(X = 0) = \text{...} \\ P(X = 1) = \text{...} \end{cases}$

$E(X) =$ ... et $V(X) =$

3. **Loi binomiale de paramètres $n \geqslant 1$ et $p \in]0, 1[$** :

Modèle : nombre de succès lors de n épreuves de Bernoulli indépendantes de paramètre p.

$X \sim B(n, p)$ si $X(\Omega) =$ .... et $P(X = k) =$ ...

$E(X) =$ ... et $V(X) =$

4. **Loi géométrique de paramètre $p \in]0, 1[$** :

Modèle : rang du premier succès lors d'une suite d'épreuves de Bernoulli indépendantes de paramètre p.

$X \sim \mathcal{G}(p)$ si $X(\Omega) =$ .... et $P(X = k) =$ ... .

$E(X) =$ ... et $V(X) =$ ...

5. **Loi de Poisson de paramètre $\lambda > 0$** :

$X \sim \mathcal{P}(\lambda)$ si $X(\Omega) =$ .... et $P(X = k) =$ ...

$E(X) = V(X) =$ ...

3) Espérance : en classe

4) Paramètres de dispersion : en classe