UNIVERSITÉ BATNA-2

FACULTÉ MATHÉMATIQUES ET INFORMATIQUE

DÉPARTEMENT INFORMATIQUE

COURS

PROBABILITÉS ET STATISTIQUES

L3 INFORMATIQUES

Présenté par

Abderahmene IKASSOULENE

2020 / 2021

Table des matières

Chapitre 1

Espaces probabilisés

1.1 Expériences aléatoires

1.1.1 Vocabulaire

Définition 1.1. *On appelle expérience aléatoire toute expérience dont le résultat ne peut être déterminé a priori, c'est-à-dire qui dépend du hasard.*

Exemple

On lance un dé cubique dont les faces sont numérotées de 1 à 6 et on note le résultat. On e ectue donc une expérience aléatoire.

Définition 1.2. *On appelle univers de l'expérience aléatoire, l'ensemble Ω des issues ou résultats possibles de l'expérience. Les éléments de Ω se notent souvent ω.*

Exemples

1) On lance un dé et on note le résultat, l'univers de l'expérience est alors :

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

2) On lance une pièce et on regarde si elle tombe sur Pile ou Face, l'univers est alors :

$$\Omega = \{Pile, Face\}.$$

3) On choisit simultanément 5 cartes dans un jeu de 32 cartes, l'univers est alors l'ensemble des sous-ensembles de 5 éléments (sans ordre) parmi les 32 cartes.

4) Si on lance trois fois un dé à six faces et que l'on note les trois résultats, on réalise une expérience dont l'univers est

$$\Omega = \{(x, y, z)/x, y, z = 1, .., 6\}$$

. 5) Si on choisit au hasard un réel compris entre 0 et 1, l'univers de l'expérience est [0, 1].

Définition 1.3. *On appelle événement toute partie de l'univers* Ω *de l'expérience aléatoire.*

Remarques

1) Lorsqu'on regarde l'expérience aléatoire, un événement est un fait lié à cette expérience pouvant se produire ou non.

2) L'ensemble des événements est donc l'ensemble des parties de Ω, c'est donc $\mathcal{P}(\Omega)$.

Exemples

1) On lance un dé et on regarde le résultat. Regardons l'événement A : "le chiffre obtenu est un nombre pair". L'événement A est réalisé lorsque le résultat est 2, 4, 6. On écrit alors $A = \{2, 4, 6\}$.

2) Un événement qui est toujours réalisé est appelé un événement certain, il est donc représenté par l'ensemble Ω.

3) Un événement qui n'est jamais réalisé est appelé un événement impossible, il est représenté par l'ensemble vide ϕ.

Remarques

Un événement est une partie de Ω, donc peut être vu comme un sous-ensemble de Ω. On garde donc exactement le même vocabulaire pour les événements que pour les ensembles.

1) L'événement contraire de A est représenté par le complémentaire de A dans Ω que l'on note $\bar{A}$.

2) L'événement "A et B sont réalisés" est représenté par $A \cap B$.

3) L'événement "A ou B est réalisé" est représenté par $A \cup B$.

4) On dit que deux événements A et B sont incompatibles s'ils ne peuvent pas être réalisés simultanément, c'est-à-dire si $A \cap B = \phi$.

5) L'événement "A est réalisé et B n'est pas réalisé" est représenté par $A - B$.

6) On dit que "l'événement A implique l'événement B" si la réalisation de A entraîne la réalisation de B, c'est-à-dire si $A \subset B$.

Définition 1.4. *Les événements qui sont représentés par un singleton $\{\omega\}$ sont appelés des événements élémentaires.*

Exemple

Si on lance un dé et qu'on note le résultat, l'événement "obtenir 2", $A = \{\ 2\}$ est un évènement élémentaire. L'événement "Obtenir un nombre pair", $B = \{2, 4, 6\}$ n'est pas un événement élémentaire.

1.1.2 Fréquence et probabilité

Soit E une expérience aléatoire. Tous les événements liés à (E) n'ont pas tous la même "chance" de se produire. Pour essayer de "mesurer" cette chance, il nous faut répéter plusieurs fois l'expérience. Si on répète n fois l'expérience, on compte le nombre P_n de fois où l'événement A se produit, . Le réel P_n/n est appelé la fréquence d'apparition de A , et on note $f(A) = P_n/n$.

Remarques

1) La fréquence f vérifie les propriétés suivantes :

a) $f(A) \in$[0, 1].

b) Si A et B sont deux événements incompatibles, alors $f(A \cup B) = f(A) + f(B)$.

c) $f(\bar{A}) = 1 - f(A)$.

d)$f(\Omega) = 1$.

e)$f(\phi) = 0$.

2) Si on fait tendre n vers $+\infty$, $f(A)$ admet une limite infinie qui représentera précisément la proportion d'apparition de A sur toutes les chances possibles. C'est cette limite qu'on appellera probabilité de l'événement A, notée $P(A)$.

1.2 Espaces probabilisés

1.2.1 Probabilité sur un ensemble fini

Définition 1.5. *Soit Ω un ensemble fini et soit $\mathcal{P}(\Omega)$ l'ensemble de ses parties.*
Le couple $(\Omega, \mathcal{P}(\Omega))$ s'appelle un espace probabilisable.

Remarque

L'ensemble $\mathcal{P}(\Omega)$ vérifie plusieurs propriétés fondamentales : on dit que c'est une tribu. En effet :

(i) $\Omega \in \mathcal{P}(\Omega)$

(ii) $\forall A \in \mathcal{P}(\Omega), \bar{A} \in \mathcal{P}(\Omega)$

(iii) $\forall A_1, .., A_k \in \mathcal{P}(\Omega)$, $\bigcup_{i=1}^{k} A_i \in \mathcal{P}(\Omega$

($\mathcal{P}(\Omega)$ est stable par passage au complémentaire, et est stable par réunion.)

Définition 1.6. *Soit* $(\Omega, \mathcal{P}(\Omega))$ *un espace probabilisable. On appelle probabilité définie sur cet espace, toute application :* $P : \mathcal{P}(\Omega) \to [0, 1]$

qui vérifie :

i) $P(\Omega) = 1$

ii) Si A *et* B *sont deux événements incompatibles, alors*

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

Le Le triplet $\Omega \in \mathcal{P}(\Omega), P$ *s'appelle un espace probabilisé fini et pour tout événement* A*, le réel* $P(A)$ *s'appelle la probabilité de l'événement* A*.*

1.2.2 Notion de tribu

Il existe des expériences ou l'univers Ω n'est pas fini. Dans ce cas, la construction d'une probabilité sera un peu plus délicate.

Exemple

On considère l'expérience suivante : on lancer une pièce jusqu'à obtenir Pile au moins une fois et on note le nombre de lancers qui ont été nécessaires pour obtenir ce premier Pile. Ici, l'univers de l'expérience est :

$$\Omega = \mathbb{N}^* = \{1, 2, 3, ...\}$$

car le nombre de lancers nécessaires pour obtenir Pile peut être n'importe quel nombre (peut-être très grand).

Remmaque

Contrairement au cas fini, parmi les parties de Ω, il ne faut s'intéresser qu'à celles dont on peut calculer la probabilité

Définition 1.7. *Soit Ω un ensemble et $\mathcal{A}$ une partie de $\mathcal{P}(\Omega)$, donc un ensemble de parties de Ω.*
On dit que $\mathcal{A}$ est une tribu de parties de Ω si :
1) $\Omega \in \mathcal{A}$.
2) Si $A \in \mathcal{A}$, alors $\bar{A} \in \mathcal{A}$.
3) Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'éléments de $\mathcal{A}$, alors $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Exemple

1) L'ensemble $\mathcal{P}(\Omega)$ est une tribu de parties de Ω.
2) Si $A \subset \Omega$, alors $\mathcal{A} = \{\phi, A, \bar{A}, \Omega\}$ est une tribu.

Définition 1.8. *Soient Ω un ensemble et $\mathcal{A}$ une tribu de parties de Ω. Le couple $(\Omega, \mathcal{A})$ s'appelle un espace probabilisable.*
Les événements de $\mathcal{A}$ sont des événements de l'espace probabilisable $(\Omega, \mathcal{A})$.

Définition 1.9. *Soient Ω un ensemble et $\mathcal{A}$ une tribu de parties de Ω. Alors*
1. $\phi \in \mathcal{A}$.
2. Si A et A sont deux événements de $\mathcal{A}$, alors $A \cup B$, $A \cap B$ et $A - B$ sont dans $\mathcal{A}$.
3. Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'éléments de $\mathcal{A}$, alors $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

1.2.3 Espace probabilisé général

Définition 1.10. *Soit $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle probabilité définie sur cet espace, toute application : $P : \mathcal{A} \to [0, 1]$*
qui vérifie :
i) $P(\Omega) = 1$.
ii) Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements deux à deux incompatibles (disjoints), alors

$$P \bigcup_{n\in\mathbb{N}} A_n = \Sigma_{n=0}^{+\infty} P(A_n).$$

Le Le triplet $\Omega \in \mathcal{A}, P$ s'appelle un espace probabilisé.

Remarque

La série $P(A_n)_{n\in\mathbb{N}}$ est bien convergente.

Définition 1.11. *Un événement* $A \in \mathcal{A}$*est dit :*

1) négligeable si $P(A) = 0$*,*

2) presque sûr si $P(A) = 1$*.*

1.2.4 Propriétés

Proposition 1.1. *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soient* A *et* B *deux événements. Alors*

1) $P(\bar{A}) = 1 - P(A)$*.*

2) En particulier, $P(\phi) = 1 - P(\Omega) = 0$*.*

3) $P(A - B) = P(A) - P(A \cap B)$*.*

4) Si $A \subset B$*, alors* $P(A) \leq P(B)$*.*

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

Remarque

La formule 5) est vraie dans le cas général. On a par exemple pour trois ensembles :

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

1.2.5 Événements élémentaires

Définition 1.12. *Soit* Ω *un ensemble fini :*$\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$*.*

Les événements $\{\omega i\}$ *sont appelés les événements élémentaires.*

Proposition 1.2. *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé fini. Pour tout événement* A*, on a :*

$$P(A) = P(\{\omega/\omega \in A\} = \Sigma_{\omega_i \in A} P(\omega_i)).$$

En particulier, si $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$*, alors*

$$\Sigma_{k=1}^{n} P(\omega_k) = 1.$$

Remarque

On voit que pour calculer $P(A)$ de n'importe quel événement A, il suffit de savoir calculer les probabilités des événements élémentaires.

1.2.6 Le modèle d'équiprobabilité

Définition 1.13. *Soit Ω un ensemble fini.*

On dit que l'espace probabilisé $(\Omega, \mathcal{A}, P)$ est en situation d'équiprobabilité lorsque tous les événements élémentaires ont la même probabilité.

Si $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, l'équiprobabilité se traduit par :

$$\forall k = 1, .., n; P(\omega_k) = 1/n.$$

Théorème 1.1. *(Formule dans le cas d'équiprobabilité)*

Si un espace probabilisé $(\Omega, \mathcal{A}, P)$ est en situation d'équiprobabilité, alors pour tout événement A, on a :

$$P(A) = \frac{Card(A)}{Card(\Omega)} = \frac{Nombre\ de\ cas\ favorables\ a\ A}{Nombre\ de\ cas\ total}$$

Exemples

1) On lance un dé équilibré. On a $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Les événements sont alors tous équiprobables.

2) On lance deux dés équilibrées. On a $\Omega = \{(i, j)/i, j = 1, .., 6\}$.

Les événements sont alors tous équiprobables.

Remarque

La situation d'équiprobabilité est un cas très particulier. Dans la plupart des cas sont en situation de non-équiprobabilité.

Exemple

On lance un dé, et à cause d'un défaut de fabrication, le 6 a deux fois plus de chance de sortir que toutes les autres faces (qui elles sont équiprobables).

On a $\Omega = \{1, 2, 3, 4, 5, 6\}$. Mais on peut considérer aussi $\Omega' = \{1, 2, 3, 4, 5, 6_a, 6_b\}$. Alors

$$P(1) = ... = P(5) = 1/7 \ et \ P(6) = 2/7.$$

les événements élémentaires de Ω ne sont pas équiprobables.

1.2.7 Systèmes complets d'évènements

Définition 1.14. *Soit $(\Omega, \mathcal{A})$ un espace probabilisable.*

On appelle système complet d'événements de Ω toute famille $(A_i)_{i\in I}$ d'éléments de A telle que :

(i) Les événements A_i sont deux à deux incompatibles.

(ii) $\bigcup_{i\in I} A_i = \Omega$.

Example

1) On lance un dé équilibré. On a $\Omega = \{1, 2, 3, 4, 5, 6\}$.

La famille d'événements $(\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\})$ est le système complet d'événements composé des événements élémentaires.

1) On note A l'événement "obtenir un nombre pair" et B l'événement "obtenir un nombre impair". Alors (A, B) forme un système complet d'événements.

Proposition 1.3. *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soit* $(A_i)_{i\in I}$ *un système complet d'événements. Alors*

$$\Sigma_{i\in I} P(A_i) = 1.$$

1.2.8 Théorème de la limite monotone

Définition 1.15. *Soit* $(A_n)_{n\in\mathbb{N}}$ *une suite d'événements d'un espace probabilisé* $(\Omega, \mathcal{A}, P)$ *infni.*

1) On dit que $(A_n)_{n\in\mathbb{N}}$ *est une suite croissante d'événements de* A *lorsque*

$$\forall n \in \mathbb{N}, A_n \subset A_{n+1}.$$

On dit que $(A_n)_{n\in\mathbb{N}}$ *est une suite décroissante d'événements de* A *lorsque*

$$\forall n \in \mathbb{N}, A_{n+1} \subset A_n.$$

Théorème 1.2. *(Théorème de la limite monotone)*

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

1) Si $(A_n)_{n\in\mathbb{N}}$ *est une suite croissante d'événements de* A*, alors la suite* $(P(A_n))_{n\in\mathbb{N}}$ *est convergente et*

$$P(\cup_{n\in\mathbb{N}} A_n) = lim_{n\to+\infty} P(A_n).$$

2) Si $(A_n)_{n\in\mathbb{N}}$ *est une suite décroissante d'événements de* A*, alors la suite* $(P(A_n))_{n\in\mathbb{N}}$ *est convergente et*

$$P(\cap_{n\in\mathbb{N}} A_n) = lim_{n\to+\infty} P(A_n).$$

Exemple

On considère une infinité de lancers de dés et on veut déterminer la probabilité de ne jamais obtenir de 1.

Pour cela, on introduit l'événement B : "ne jamais obtenir de 1",

ainsi que les événements $\forall n \in \mathbb{N}$, B_n : "ne pas obtenir de 1 au cous des n premiers lancers". Alors, on a

$$B = \cap_{n=0}^{n=+\infty} B_n$$

De plus, on sait que

$$\forall n \in \mathbb{N}, P(B_n) = (5/6)^n$$

et "ne pas obtenir de 1 au cours des $n+1$ premiers lancers", implique en particulier que " l'on n'a pas obtenu de 1 au cours des n premiers lancers" : donc $B_{n+1} \subset B_n$ pour tout $n \in \mathbb{N}$, et la suite d'événements $(B_n)_{n\in\mathbb{N}}$ est donc décroissante pour l'inclusion.

D'après le Théorème de Limite Monotone,

$$P(B) = P(\cap_{n\in\mathbb{N}} B_n) = lim_{n\to+\infty} P(B_n) = lim_{n\to+\infty} (5/6)^n = 0.$$

D'où

$$P(\text{"}ne\ jamais\ obtenir\ de\ 1\text{"}) = 0.$$

1.3 Probabilité conditionnelle

La probabilité qu'un événement se réalise peut être influencée si on connaît des informations concernant la réalisation d'un ou plusieurs autres événements. La donnée de la réalisation ou non d'un événement va ainsi réduire l'univers des issues possibles Ω.

Exemple

On lance un dé équilibré et on regarde le résultat.

Notons A l'événement : "on obtient un nombre inférieur ou égal à 5".

Notons B l'événement : "on obtient un nombre supérieur ou égal à 3".

A priori, sans aucune information supplémentaire, on a :

$$P(A) = 5/6, P(B) = 4/6$$

Cependant, supposons que l'on sache que A s'est réalisé, alors les résultats désormais possibles sont $\{1,2,3,4,5\}$, et B est réalisé si ce résultat appartient à $\{3,4,5\}$ qu'est égale à $A \cap B$. Il a donc 3 cas favorables sur 5.

La probabilité que B se réalise sachant que A est réalisé est alors de $\frac{3}{5}$.

On vérifie en réalité que

$$\frac{3}{5} = \frac{Card(A \cap B)}{Card(A)} = \frac{P(A \cap B)}{P(A)}$$

1.3.1 Définition

Théorème 1.3. *(Définition)*

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soit A un événement de probabilité non nulle.

Pour tout évènement B, on définit la probabilité de B sachant A par :

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

Alors la fonction : $P_A : (\Omega, \mathcal{A}) \to [0,1]$ est une probabilité sur $(\Omega, \mathcal{A}, P)$. On note parfois

$$P_A(B) = P(A/B)$$

.

Remarque

Comme P_A est une probabilité, on peut appliquer toutes les propriétés vues précédemment pour les probabilités en général.

Proposition 1.4. *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soient A et B deux événements. Alors*

1) $P_A(\bar{B}) = 1 - P_A(B)$.

2) Si $C \subset B$, alors $P_A(C) \leq P_A(B)$.

3) $P_A(B \cup C) = P_A(B) + P_A(C) - P_A(B \cap C)$.

1.3.2 Grandes formules probabilistes

Remarque

Soient A et B deux événements tels que $P(A) \neq 0$, alors par définition, on a :

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

Cette formule peut être généralisée de la façon suivante :

Proposition 1.5. *(Formule des Probabilités Composées)*
Soient A, B, C *trois événements tels que* $P(A \cap B) \neq 0$*, alors*

$$P(A \cap B \cap C) = P(A)P_A(B)P_{A\cap B}(C)$$

Soient $A_1, ..., A_n$ *une famille d'événements telle que* $P(A_1 \cap A_2 \cap ... \cap A_{n-1}) \neq 0$*, alors on a :*

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A3)...P_{A_1\cap ...\cap A_{n-1}}(A_n)$$

Remarque

Lorsqu'on fait par exemple des tirages SUCCESSIFS et SANS REMISE, alors le deuxième tirage dépend du résultat du premier tirage, autrement dit, pour obtenir des informations sur la deuxième boule tirée, il faut conditionner par le résultat du premier tirage. Puis, le troisième tirage dépend des résultats des deux premiers tirages.

Exemple

Une urne contient 3 boules blanches et 7 noires. On tire successivement trois boules sans remise.
Quelle est la probabilité d'obtenir les 3 boules blanches ?
On note $\forall k = 1, 2, 3; B_k$: "obtenir une boule blanche au k-ième tirage".
On note $A = B1 \cap B2 \cap B3$ l'événement cherché : "obtenir trois boules blanches". Alors d'après la formule :

$$P(A) = P(B) = P(B_1 \cap B_2 \cap B_3) = P(B1)P_{B_1}(B2)P_{B_1\cap B_2}(B3)$$

On a $P(B_1) = 3/10$ (3 blanches sur 10 boules au total).
On a $P_{B_1}(B2) = 2/9$ (il reste alors 2 blanches sur 9 boules au total).
On a $P_{B_1\cap B_2}(B3) = 1/8$ (il reste alors 1 blanche sur 8 boules au total).
Finalement, on a

$$P(A) = 3/10 \cdot 2/9 \cdot 1/8 = 1/120$$

Théorème 1.4. *(Formule des probabilités totales)*
Soit $(A_i)_{i\in I}$ *un système complet d'événements de probabilités non nulles. Alors pour tout événement* B*, on a :*

$$P(B) = \Sigma_{i\in I}P(A_i \cap B) = \Sigma_{i\in I}P(A_i)P_{A_i}(B)$$

Remarque

Cela signifie simplement que si $(A_i)_{i\in I}$ est un système complet d'événements, alors lorsqu'un événement B se réalise, il se réalise soit avec A_1, soit avec A_2, soit avec A_3, . . .

Théorème 1.5. *(Formule de Bayes)*
Soit $(A_i)_{i\in I}$ un système complet d'événements de probabilités non nulles et soit B un événement de probabilité non nulle. Alors pour tout $i_0 \in I$, on a :

$$P_B(A_{i_0}) = \frac{P(A_{i_0})P_{A_{i_0}}(B)}{\Sigma_{i\in I}P(A_i)P_{A_i}(B)}$$

Démonstration. Il suffit d'appliquer la définition de la probabilité conditionnelle et la formule des Probabilités Totales :

$$P_B(A_{i_0}) = \frac{P(A_{i_0}\cap B)}{P(B)} = \frac{P(A_{i_0})P_{A_{i_0}}(B)}{\Sigma_{i\in I}P(A_i)P_{A_i}(B)}$$

□

1.4 Indépendance d'événements

1.4.1 Définition

Définition 1.16. *On dit que deux événements A et B sont indépendants pour la probabilité P si*

$$P(A\cap B) = P(A)\cdot P(B)$$

Remarques

1) Si A et B sont deux événements indépendants (de probabilité non nulle), alors on a

$$P_A(B) = P(B) \text{ et } P_B(A) = P(A)$$

2) La notion d'indépendance dépend de la probabilité P choisie sur les événements élémentaires.

Exemple

Soit $\Omega = \{2, 2, 3, 4, 5, 6\}$ et soit l'espace $(\Omega, \mathcal{P})$, sur lequel on définie deux probabilités P_1 et P_2, telles que

$$P_1(1) = 1/6; P_1(2) = 1/6; P_1(3) = 1/3; P_1(4) = 1/9; P_1(5) = 1/9; P_1(6) = 1/9$$

$$P_2(1) = 1/6; P_2(2) = 1/6; P_2(3) = 1/6; P_2(4) = 1/6; P_2(5) = 1/6; P_2(6) = 1/6$$

Soient les évènements $A = \{1, 2\}$ et $B = \{2, 3\}$, alors $A \cap B = \{2\}$.

$$P_1(A)P_1(B) = (1/6 + 1/6)(1/6 + 1/3) = 1/6 \quad P_1(A \cap B) = 1/6$$

$$P_2(A)P_2(B) = (1/6 + 1/6)(1/6 + 1/6) = 1/9 \quad P_2(A \cap B) = 1/6$$

On remarque que $P_1(A)P_1(B) = P_1(A \cap B)$ et $P_2(A)P_2(B) \neq P_2(A \cap B)$.

Donc A et B sont indépendants pour la probabilité P_1, mais ne sont pas indépendants pour la probabilité P_2.

1.4.2 Propriétés

Proposition 1.6. *Soient A et B deux événements indépendants, alors :*

1) A et $\bar{B}$ sont indépendants,

2) $\bar{A}$ et B sont indépendants,

3) $\bar{A}$ et $\bar{B}$ sont indépendants.

Démonstration. On démontre la première comme suit :

On a $A = A \cap \Omega = A \cap (B \cup \bar{B}) = (A \cap B) \cup (A \cap \bar{B})$ et ces deux événements sont incompatibles, donc

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

alors

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

□

Définition 1.17. *Soient $A_1, ..., A_n$, n événements.*

1) On dit que les événements sont deux à deux indépendants pour la probabilité P si

$$\forall i \neq j, A_i \textit{ et } A_j \textit{ sont independants}$$

2) On dit que les événements sont mutuellement indépendants pour la probabilité P si pour tout $I \subset \{1, .., n\}$

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

Définition 1.18. *Soient* (A_n)*, une suite infinie d'événements.*

On dit que les événements A_n *sont indépendants, si pour toute partie finie* I *de* $\mathbb{N}$*, on a*

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

Chapitre 2

Variables aléatoires

2.1 Définitions générales

2.1.1 Variables aléatoires

On appelle variable aléatoire tout nombre réel aléatoire, c'est-à-dire dont la valeur dépend du résultat d'une expérience probabiliste. Par exemple :

On lance un de. Soit X le resultat obtenu.

Ici X est une variable aléatoire et les valeurs possibles de X sont 1 ; 2 ; 3 ; 4 ; 5 ; 6. Pour chacune de ces valeurs, X a une certaine probabilité de lui être égal. Ici en fait on peut donner directement les probabilités des événements "X = 1" ; "X = 2" ; ... ; "X = 6" : on a P("X = 1") = P("X = 2") = ... = P("X = 6") = 1/6 .

Remarques sur les notations

1) Par convention, les variables aléatoires sont en général notées avec des lettres capitales (X,Y ,T,etc.) pour les différencier des nombres réels non aléatoires.

2) Pour noter les événements relatifs à une variable aléatoire X, comme par exemple "X = 1" ; "X $\leq$ 2" ; "0 $\leq$ X $\leq$ 4", on utilise souvent plutôt les crochets : [X = 1] ; [X $\leq$ 2] ; [0 $\leq$ X $\leq$ 4].

3) De même, plutôt que P("X = 1") ou P([X = 1]) on écrira simplement P(X = 1) ou P[X = 1].

2.1.2 Support d'une variable aléatoire

Le support d'une variable aléatoire est l'ensemble des ses valeurs possibles. C'est la première chose à préciser lorsqu'on considère une variable aléatoire. On notera $\mathcal{S}(X)$ le support d'une variable aléatoire X.

Exemple 1 :

On reprend l'exemple précédent : X est le résultat d'un lancer de dé. Le support de X est alors $\{1; 2; 3; 4; 5; 6\}$.

Exemple 2 :

On lance un dé à 6 faces, et l'on recommence jusqu'à obtenir un 6. On note X le nombre de lancer nécessaires. Alors X$\in \{1; 2; ...\}; \mathcal{S}(X) = \mathbb{N}^{\star}$.

Exemple 3 :

On lance un stylo sur un bureau, et on note X l'angle que fait le stylo avec la surface du bureau. Alors le support de X est $[0\,; \pi]$.

Intéressons-nous dans les 3 cas à la probabilité P(X = x) pour un certain x du support.

1. Pour le premier exemple, si je prends x = 4, j'ai P(X = 1) = 1=6.

2. Pour le deuxième exemple, si je prends x = 2, j'ai P(X = 2) = 5/6 · 1/6 =5/36.

3. Pour le troisième exemple, si je prends x = $\frac{\pi}{4}$, quelle est la chance que le stylo fasse exactement un angle de $\frac{\pi}{4}$? La réponse est 0. On pourra lancer le style autant de fois qu'on veut, il ne tombera jamais exactement sur $\frac{\pi}{4}$, même s'il se rapprochera beaucoup. On aboutit donc au paradoxe suivante :

$$\forall x \in [0; \pi]; P(X = x) = 0 \textit{ mais pourtant } P(X \in [0; \pi]) = 1.$$

Cela vient du fait qu'on ne peut pas écrire

$$P(X \in [0; \pi]) = \Sigma_{x\in[0;\pi]} P(X = x)$$

alors qu'on pouvait écrire dans le deuxième cas

$$P(X \in \mathbb{N}) = \Sigma_{n\in\mathbb{N}^{\star}} P(X = n).$$

Cela est du à la nature de l'ensemble $[0\,;\pi]$, qui est un intervalle continu de $\mathbb{R}$, alors que $\mathbb{N}$, est l'ensemble des entiers. On dit que $[0\,;\pi]$ est infini indénombrable, alors que $\mathbb{N}$ est infini dénombrable. En pratique, les variables discrètes auront comme support un ensemble fini (ex : $\{1; 2; 3; 4; 5; 6\}$), ou un ensemble infini dénombrable comme $\mathbb{N}$, ou $\mathbb{Z}$.

Remarque

Il existe également des variables qui ne sont ni discrètes ni continues.

Exemple 4 : Je joue au golf, et j'appelle X la distance entre ma balle et le trou après un coup. J'ai $\mathcal{S} = [0\,;+\infty[$. J'ai P(X = 0) > 0 car il est possible que je mette la balle dans le trou, et P(X = x) = 0 pour x > 0 : Donc X n'est ni discrète ni continue.

2.1.3 Fonction de répartition d'une variable aléatoire

Définition 2.1. *La fonction de répartition d'une variable aléatoire X est la fonction définie pour tout $t \in \mathbb{R}$ par*

$$F_X(t) = P(X \leq t).$$

Autrement dit, $F_X(t)$ est la probabilité de l'événement "la valeur de X est inférieure ou égale à t".

Exemple 1 :

X est le résultat du lancer d'un dé. On a vu que P(X = 1) = P(X =2) = ... = P(X = 6) = 1/6 : c'est la loi de X. On peut la présenter sous forme de tableau :

x	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

On peut aussi présenter ce résultat sous la forme d'un graphique, ou diagramme en bâtons (figure 2.1 à gauche). Enfin on peut aussi tracer le graphe de la fonction de répartition F_X (figure 2.1 à droite). On voit ici que la fonction de répartition est constante par morceaux : elle présente des sauts pour les valeurs du support de X, mais reste constante entre deux de ces valeurs. Ce sera toujours le cas pour des variables discrètes.

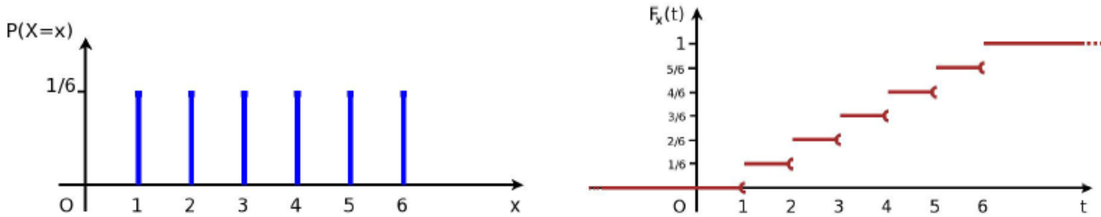

Figure 2.1 $-$ *Loi et fonction de repartition de lancer d'un de*

Proposition 2.1. *On a $F_X(t) \in [0;1]$ pour tout $t \in \mathbb{R}$; et F_X est une fonction croissante.*

Démonstration. $F_X(t) \in [0;1]$ car c'est une probabilité. De plus, si $t \leq u$, on a $[X \leq t] \subset [X \leq u]$ et donc $P(x \leq t) \leq P(x \leq u)$, c'est-à-dire $F_X(t) \leq F_X(u)$. Donc F_X est croissante. □

Grâce à la fonction de répartition, on peut calculer la probabilité de tomber dans un intervalle quelconque : Pour $t, s \in \mathbb{R}$ tel que $t \leq s$,

$$P(X \in]t;s]) = P((X \leq s) \backslash (X \leq t)) = P(X \leq s) - P(X \leq t) = F_X(s) - F_X(t).$$

2.2 Variables aléatoires discrètes

Définition 2.2. *(Loi d'une variable aléatoires discrète)*
Donner la loi d'une variable aléatoire discrète X, c'est calculer les probabilités P(X = x) pour toutes les valeurs x possibles prises par X (autrement dit pour tous les x appartenant au support de X).

Lorsque toutes les probabilités formant la loi de X sont égales, comme dans l'exemple du dé, on parle de loi uniforme. C'est l'exemple le plus simple de variable aléatoire.

Définition 2.3. *Soit un ensemble fini $\Omega = \{x_1; x_2; ...; x_n\}$. On dit que la variable X suit la loi uniforme sur Ω lorsque $\mathcal{S}(X) = \Omega$ et $P(X = x_i) = 1/n$ pour tout $1 \leq i \leq n$. On note $X \sim \mathcal{U}_\Omega$.*

Exemple

Si X est le résultat d'un lancer de dé, $X \sim \mathcal{U}_\Omega$ avec n = 6 et $\Omega = \{1; 2; 3; 4; 5; 6\}$.

Si X est le résultat d'un pile-ou-face, $X \sim \mathcal{U}_\Omega$ avec n = 2 et $\Omega = \{"pile"; "face"\}$.

Dans ce qui suit, on donne quelques fameuses variables aléatoires discrètes :

2.2.1 La loi binomiale

Exemple (Loi binomiale pour n = 3)

On lance trois pièces de monnaie. Soit X "le nombre de Faces obtenues". Quelle est la loi de X ?
Le support de X est ici $\mathcal{S}(X) = \{0; 1; 2; 3\}$. On doit donc calculer P(X = 0) ; P(X = 1) ; P(X = 2) et P(X = 3) pour décrire complètement la loi de X.
Définissons L_1 ="la première pièce tombe sur Face" ; et de même L_2 et L_3. On peut clairement supposer que les trois lancers de pièce sont indépendants ici ; et donc que $L_1; L_2; L_3$ sont indépendants.
•$[X = 0] = L_1^c \cap L_2^c \cap L_3^c$ donc $P(X = 0) = P(L_1^c)P(L_2^c)P(L_3^c)$ grâce à l'indépendance. Ainsi

$$P(X = 0) = 1/2 \cdot 1/2 \cdot 1/2 = 1/8.$$

•$[X = 1] = (L_1 \cap L_2^c \cap L_3^c) \cup (L_1^c \cap L_2 \cap L_3^c) \cup (L_1^c \cap L_2^c \cap L_3)$. Cette union est disjointe, donc on peut additionner les probabilités :

$$P(X = 1) = P(L_1 \cap L_2^c \cap L_3^c)P(L_1 \cap L_2 \cap L_3^c)P(L_1 \cap L_2^c \cap L_3),$$

$$P(X = 1) = P(L_1)P(L_2^c)P(L_3^c) + P(L_1^c)P(L_2)P(L_3^c) + P(L_1^c)P(L_2^c)P(L_3),$$

$$(par\ independance)$$

$$P(X = 1) = 3 \cdot 1/2 \cdot 1/2 \cdot 1/2 = 3/8.$$

• Pour calculer intelligemment $P(X = 2)$, on introduit Y "le nombre de piles obtenus". Il est clair que le nombre de "piles" suit la même loi que le nombre de "faces", donc $Y \sim X$. On a de plus Y = X-3. Donc $P(X = 2) = P(Y = 1) = P(X = 1) = 3/8$.
• Avec le même raisonnement, P(X = 3) = P(Y = 0) = P(X = 0) = 1/8.

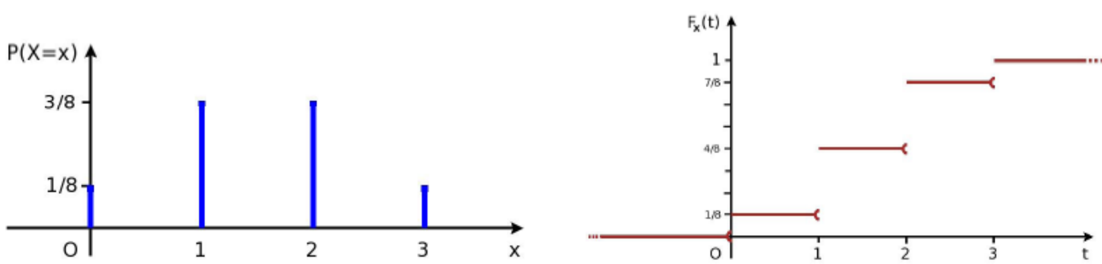

Figure 2.2 – Loi et fonction de repartition de X (loi binomiale, n = 3)

La loi de X est en fait un exemple de loi binomiale. La figure 2.2 montre le graphique de cette loi ainsi que la fonction de répartition.

Supposons désormais que la pièce soit biaisée(non homogène), c'est-à-dire qu'elle a une probabilité p de tomber sur pile, où $p \in [0;1]$ est un paramètre inconnu. Alors on a avec le même raisonnement

$$P(X=0) = (1-p)^3$$

$$P(X=1) = 3p^2(1-p)$$

$$P(X=2) = 3p(1-p)^2$$

$$P(X=3) = 3p$$

Finalement, supposons qu'on ne lance plus 3 pièces mais n pièces, où $n \in \mathbb{N}^\star$. Cette fois $\mathcal{S}(X) = \{0;1;2;...;n\}$. Pour donner la loi de X, il faut calculer pour tout $k \in \{0;1;2;...;n\}$, $P(X=k)$. L'évènement $X = k$ survient si exactement k pièces tombent sur Pile, et n-k tombent sur Face. On numérote les pièces de 1 à n. L'évènement X = k survient donc si il y a un sous-ensemble E à k éléments tel que L_i est vrai pour $i \in E$, et L_i^c est vrai pour $i \notin E$. En sommant sur tous les sous-ensembles E possibles on a

$$P(X=k) = \sum_{les\ sous-ensembles\ E\ a\ k\ elements} P(L_i\ est\ vrai\ pour\ i \in E,\ L_i^c\ est\ vrai\ pour\ i \notin E)$$

$$P(X=k) = \sum_{les\ sous-ensembles\ E\ a\ k\ elements} \prod_{i \in E} P(L_i) \prod_{i \notin E} P(L_i^c)$$

(car les évènements sont indépendants)

$$P(X=k) = \sum_{les\ sous-ensembles\ E\ a\ k\ elements} p^k(1-p)^{1-k}$$

$$P(X=k) = \{le\ nombre\ de\ sous-ensembles\ E\ a\ k\ elements\} p^k(1-p)^{1-k}$$

$$P(X=k) = C_n^p\, p^k(1-p)^{1-k}.$$

On remarque que la somme de toutes les probabilités est bien égale à 1, car par la formule du binôme

$$\sum_{k=0}^{n} P(X=k) = \sum_{k=0}^{n} C_n^p\, p^k(1-p)^{1-k} = (p+(1-p))^n = 1^n = 1.$$

Définition 2.4. *Soit $n \in \mathbb{N}^\star$ et $p \in [0;1]$. On appelle loi binomiale de paramètre n et p, notée $\mathcal{B}(n;p)$ la loi d'une variable X telle que $S(X) = \{0;1;2;...;n\}$ et pour $0 \leq k \leq n$,*

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

Une variable binomiale représente le nombre de succès lors de n répétitions indépendantes d'expériences aléatoires qui ont chacune une probabilité p de réussir.

Exemple

• Je choisis au hasard 6 étudiants dans la promo, et j'appelle X le nombre de ces étudiants qui auront plus de 16 en note finale. Comme j'ai choisi les étudiants au hasard, les succès des uns et des autres sont des variables aléatoires indépendantes les unes des autres. En effet, savoir qu'un des étudiants a eu 16 ne me donne aucun renseignement sur les résultats des autres. Donc, en supposant que 2/3 des étudiants de la promo auront plus de 16, $X \sim B(6;2/3)$.

• Je choisis cette fois un étudiant au hasard dans la promo, et je considère les notes qu'il aura aux 6 matières du semestre. Je note X le nombre de ces notes qui seront supérieures à 16. Cette fois, X ne suit pas une loi binomiale car les résultats de cet étudiant aux différentes matières dépendent de certains facteurs : Le niveau de l'étudiant, la quantité de travail qu'il va fournir, etc...Par exemple, si je sais que l'étudiant a eu 16 à la 1 ère matière, c'est probablement un bon étudiant, et il aura probablement de bonnes notes dans les autres matières...

La figure 2.3 suivante montre des exemples de lois binomiales pour différents paramètres :
lois binomiales (à gauche) et fonctions de répartition F_X correspondantes (à droite) pour différents paramètres : 1 ère ligne : $n = 5; p = 0.4$; 2 ème ligne : $n = 15; p = 0.6$; 3 ème ligne : $n = 10; p = 0.05$.

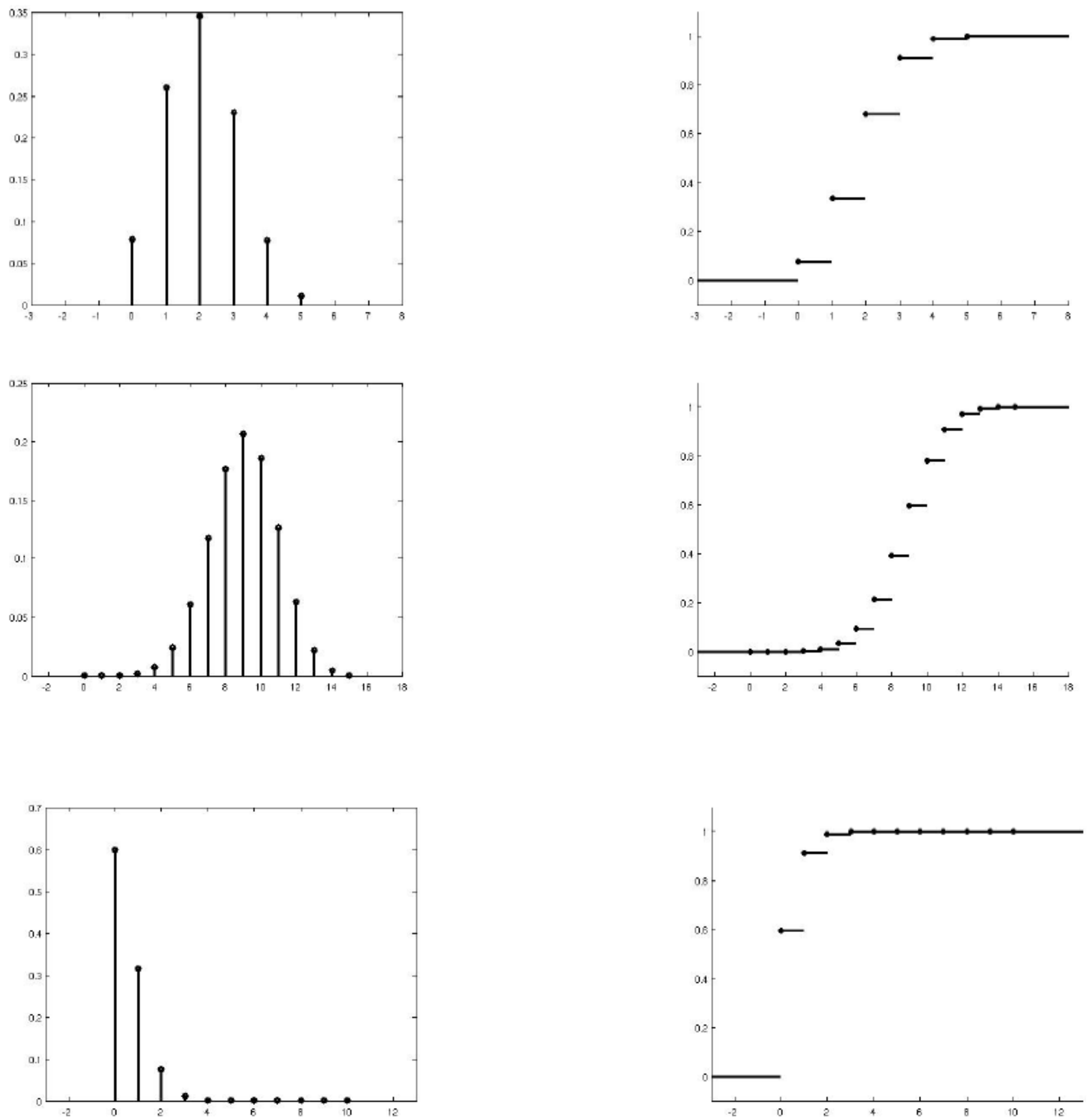

Figure 2.3 – *lois binomiales* (*a gauche*) *et fonctions de repartition* (*a droite*)

2.2.2 Loi de Bernoulli

On appelle expérience de Bernoulli une expérience aléatoire qui n'a que deux issues possibles ("Pile" ou "Face", "Noir" ou "Blanc", "succès" ou "échec", 0 ou 1, etc...).

Définition 2.5. *Une variable aléatoire* X *suit la loi de Bernoulli de paramètre* $p \in [0;1]$ *si* $X \in \{0;1\}$ *et* $P(X=1)=p, P(X=0)=1-p$. *On note* $X \sim \mathcal{B}(p)$.

Exemple

- On jette une pièce en l'air. Soit X la variable

$X = 1$ si la pièce tombe sur pile, $X = 0$ sinon. Alors $X \sim \mathcal{B}(1/2)$.

- On jette un dé. Soit la variable

$X = 1$ si le dé fait 6, $X = 0$ sinon. Alors $X \sim B(1/6.)$

Exemple

Soit $n > 1$ et $p \in [0;1]$. Soit $X_i; i = 1;...;n$ des variables de Bernoulli indépendantes de paramètre p. Soit $X = \sum_{i=1}^{i=n} X_i$. Alors $X \sim B(n;p)$. En effet, en réfléchissant un peu, on voit que X est le nombre d'expériences de Bernoulli qui réussissent, où chaque expérience a une probabilité p de réussir, indépendamment des autres.

2.2.3 La Loi géométrique

Exemple (Loi géométrique)

On lance un dé et on recommence jusqu'à l'apparition du premier 6 (par exemple [X = 5] ="on obtient 6 pour la première fois au 5 ème lancer"). Déterminer la loi de X.

Le support de X est ici $\{1;2;3;....\} = \mathbb{N}^\star$. En toute rigueur il faudrait ici remarquer que X n'est pas nécessairement défini en envisageant le cas où le 6 n'apparaît jamais ; ou bien rajouter la valeur $+\infty$ au support en définissant l'événement $[X = +\infty]$ ="le 6 n'apparaît jamais". Cependant on peut montrer que la probabilité de cet événement est nulle, et donc qu'il n'y a pas lieu de le prendre en compte.

Pour tout $n \geq 1$, notons A_n l'événement "le n ème lancer vaut 6". A_n ne dépend que du résultat du n ème lancer, et il est clair que les lancers sont des expériences indépendantes. Par conséquent les A_n sont des événements indépendants.

• [X = 1] = A 1, donc P(X = 1) = P(A 1) = 1/6.

•$[X = 2] = A_1^c \cap A_2$, donc $P(X = 2) = P(A_1^c)P(A_2)$ par indépendance. Donc

$$P(X = 2) = 5/6 \cdot 1/6.$$

•$[X = 3] = A_1^c \cap A_2^c \cap A_3$, donc $P(X = 3) = P(A_1^c)P(A_2^c)P(A_3)$ grâce à l'indépendance. Donc

$$P(X = 3) = (5/6)^2 \cdot 1/6.$$

• Plus généralement, on aura $[X = n] = A_1^c \cap A_2^c \cap A_3 \cdots A_{n-1}^c \cap A_n$, donc $P(X = n) = P(A_1^c)P(A_2^c) \cdots P(A_{n-1}^c)P(A_n)$, toujours grâce à l'indépendance. Ainsi $P(X = n) = (5/6)^{n-1} \cdot 1/6$. Finalement cette dernière formule est valable pour tout $n \geq 1$; on a donc ainsi bien déterminé la loi de X. Cette loi s'appelle la loi géométrique de paramètre 1/6 .

Définition 2.6. *On dit qu'une variable aléatoire suit la loi géométrique de paramètre p, où $p \in [0; 1]$ fixé, si le support de X est égal à $\mathbb{N}^\star$ et que*

$$P(X = n) = p(1 - p)^{n-1}.$$

On note $\mathcal{G}(p)$ la loi géométrique de paramètre p, et on note $X \sim \mathcal{G}(p)$ pour dire que X suit une loi géométrique de paramètre p.

Une autre manière de définir la loi géométrique est via sa fonction de répartition :

$$P(X \leq n) = 1 - P(X > n) = 1 - P(A_1^c \cap \cdots \cap \ A_n^c)$$

$$= 1 - P(A_1^c)P(A_2^c) \cdots P(A_n^c)$$
$$= 1 - (1 - p)^n.$$

Pour une loi géométrique il est souvent plus simple de considérer la fonction "quantile" :

$$P(X > n) = P(rater\ n\ fois) = (1 - p)^n.$$

2.2.4 La loi de Poisson

On considère un ensemble d'évènements qui peuvent intervenir dans le temps à n'importe quel moment, indépendamment les uns des autres, et tel qu'aucune période ne soit privilégiée par rapport à une autre par rapport à l'arrivée de ces évènements.

Exemples

• Clients qui rentrent dans un bureau de poste.

• Poissons qui mordent à un hameçon (en supposant qu'on peut les décrocher et remettre la ligne instantanément).

• Naissances dans une population donnée.

• Avions qui passent dans le ciel.

• Tremblements de terre dans le monde.

Alors le nombre d'évènements qui surviennent dans une période donnée (1 heure, 1 journée,...) suit une loi de Poisson. On voit bien que le nombre d'avions qui passent dans le ciel et le nombre de tremblements de terre ne suivent pas la même loi. Il s'agit en fait d'une famille de lois qui dépendent d'un paramètre $\lambda > 0$ (comme la famille des géométriques, des exponentielles, ...)

Définition 2.7.

Soit $\lambda \geq 0$. Une variable X suit la loi de Poisson de paramètre λ si pour tout entier $k \geq 0$,

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Le support de X est ici $\mathbb{N}$. La loi de Poisson de paramètre λ est notée $\mathcal{P}(\lambda)$ (et donc $X \sim \mathcal{P}(\lambda)$).
λ nous donne (ou est défini par) le nombre moyen d'évènements qui surviennent sur la période donnée.

Exercice

Expliquer pourquoi les variables aléatoires suivantes ne suivent pas une loi de Poisson :

• Nombre de fois que le record du 100 m sera battu dans le prochain siècle.

• Nombre d'étudiants qui vont valider la matière dans un groupe de 30 étudiants.

• Nombre de gagnants au loto dans une population de 10 millions d'habitants.

2.2.5 Le lien Poisson - binomiale

La loi de Poisson est la version en temps continu de la loi binomiale (mais ça reste une variable discrète). En effet, considérons l'exemple du pêcheur, et du nombre X de poissons (sans majuscule) attrapés en une heure. On note $\lambda = \mathbb{E}(X)$ le nombre moyen de poissons attrapés en une heure : $X \sim \mathcal{P}(\lambda)$. On découpe l'heure en 60 minutes et on introduit Y le nombre de minutes où il a attrapé 1 ou plusieurs poissons. Comme ce qui se passe lors d'une minute donnée (disons la 10 ème minute) est indépendant de ce qui se passe à une autre (la 34 ème minute par exemple), alors Y est une loi binomiale de paramètre $n = 60$, et $p = P$(il attrape un poisson ou plus sur une durée d'une minute). Donc Y s'obtient à partir de X en "découpant" le temps. Pour résumer :

• Le nombres de poissons attrapés suit une loi de Poisson.

• Le nombres de minutes où il a attrapé un ou des poisson(s) suit une loi binomiale.

2.2.6 Les événements élémentaires [X = x]

1) Les événements [X = x], considérés pour tous les x appartenant au support, forment une partition (ou un système complet) de Ω . En effet, X prend nécessairement une et une seule de ces valeurs, ce qui prouve bien que la réunion des [X = x] est égale à Ω, et que $[X = x] \cap [X = y] = \Phi$ si $x \neq y$. Par conséquent la somme totale des $P(X = x)$ doit toujours être égale à 1, ce qui s'écrit, en notant $\mathcal{S}$ le support de X,

$$\sum_{x \in \mathcal{S}} P(X = x) = 1.$$

Il faut toujours penser à le vérifier lorsqu'on calcule une loi. Pour les exemples précédents, on a :

• pour le dé : $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$,

• pour la loi géométrique, on a une série géométrique

$$\sum_{n=1}^{n=+\infty} p(1-p)^{n-1} = p \sum_{n=1}^{n=+\infty} (1-p)^{n-1} = p \frac{1}{1-(1-p)} = 1.$$

• Pour la loi binomiale : si $X \sim \mathcal{B}(n; p)$,

$$\sum_{x \in \mathcal{S}(\mathcal{X})} P(X = x) = \sum_{k=0}^{k=n} C_n^k p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1.$$

2) De plus ces événements [X = x] sont les événements élémentaires pour la variable X, au sens où tout événement relatif à X s'exprime comme une union de ces événements, et sa probabilité est la somme des $P(X = x)$ correspondants.

Exemple

L'événement "le dé tombe sur un nombre pair supérieur à 3" est égal à $[X = 4][\cup[X = 6]$ si X est le résultat du dé. Sa probabilité est donc de $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

Exemple

Si X est une variable discrète quelconque, l'événement $[X \leq t]$ est l'union des évènements [X = x] pour tous les x appartenant au support de X tels que $x \leq t$. On peut ainsi donner une formule pour la fonction de répartition d'une variable discrète :

$$F_X(t) = \sum_{x \in \mathcal{S}, x \leq t} P(X = x).$$

2.3 Variables aléatoires continues, ou à densité

2.3.1 Variable aléatoire uniforme

On jette un stylo sur une table, et on note X l'angle entre 0 et π qu'il forme avec le bord de la table. Quelle est la loi de X ?

L'ensemble des valeurs possibles pour cette variable aléatoire est $[0; \pi]$. En suivant l'idée vue auparavant, on voudrait donc chercher à calculer tous les $P(X = \alpha)$ pour $\alpha \in [0; \pi]$. En fait on verra que $P(X = \alpha)$ sera toujours égal à 0, ce qui signifie que quelle que soit la valeur de α, le stylo n'a aucune chance de former exactement (et c'est le mot important ici) un certain angle α avec la table, de la même manière qu'aucun être humain sur terre ne mesure exactement 1m70, au milliardième de nanomètre près. On est obligé pour donner un sens aux probabilités ici, de considérer des intervalles et non des valeurs uniques, et de regarder les probabilités que X soit dans ces intervalles. Par exemple on peut raisonnablement penser ici que $P(0 \leq X \leq \frac{\pi}{2}) = \frac{1}{2}$, ou que $P(\frac{\pi}{4} \leq X \leq \frac{\pi}{2}) = \frac{1}{4}$. Plus généralement, on peut supposer que la probabilité que X appartienne à un intervalle $[\alpha; \beta] \subset [0; \pi]$ correspond à la proportion d'angles compris dans cet intervalle, c'est-à-dire que pour $0 \leq \alpha \leq \beta \leq \pi$;

$$P(\alpha \leq X \leq \beta) = c(\beta - \alpha)$$

où c est une certaine constante de proportionnalité que l'on va déterminer.
Ceci permet de caractériser entièrement la variable X car la probabilité de tout événement lié à X peut se calculer à partir de cette formule.

Définition 2.8. *Pour une variable aléatoire continue X, il existe une fonction $f_X : \mathbb{R} \to \mathbb{R}^+$ telle que*

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx \quad \textit{pour tout } a, b \textit{ dans } \mathbb{R}$$

Cette fonction f_X est appelée densité de X.

Reprenons l'exemple du stylo qui tombe sur le bureau. C'est en fait une variable de densité

$$f(x) = \begin{cases} c & \text{si } x \in [0, \pi], \\ 0 & \text{sinon.} \end{cases} \tag{2.1}$$

En effet, pour $[\alpha; \beta] \subset [0; \pi]$, la probabilité que $X \in [a; b]$ est proportionnel à la longueur de l'intervalle $[a; b]$:

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx = c(b-a)$$

Pour déterminer la valeur de c, il suffit de remarquer que

$$1 = P(0 \leq X \leq \pi) = \int_0^\pi f_X(x)dx = \int_0^\pi c\ dx = c(\pi - 0) = c\pi;$$

d'où $c = 1/\pi$.

Si par exemple $a = -1; b \in [0, \pi]$,

$$P(a \leq X \leq b) = P(-1 \leq X \leq b) = \int_{-1}^b f_X(x)dx = \int_0^b f_X(x)dx = \int_0^b \frac{1}{\pi}dx = \frac{b}{\pi}.$$

La formule est bien vérifiée dans ce cas de figure. Elle marche dans tous les cas de figure.

On aura donc par exemple :

$$P(0 \leq X \leq \pi/2) = \int_0^{\pi/2} \frac{1}{\pi}dx = \frac{1}{2},$$

ou encore :

$$P(\pi/4 \leq X \leq \pi/2) = \int_{\pi/4}^{\pi/2} \frac{1}{\pi}dx = \frac{1}{4}.$$

C'est en fait un cas particulier de la loi uniforme.

Définition 2.9. *La loi uniforme sur un intervalle $[\alpha, \beta]$ est la loi de densité*

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{si } x \in [\alpha, \beta], \\ 0 & \text{sinon.} \end{cases} \tag{2.2}$$

On note $X \sim U([\alpha, \beta])$ ("X suit la loi uniforme sur $[\alpha, \beta]$").

Remarque 2.1. *Ne pas confondre loi uniforme pour une variable discrète et pour une variable continue !*

Remarque 2.2. *on peut prendre $a = -\infty$ ou $b = +\infty$dans cette formule. On a par exemple*

$$P(-\infty \leq X \leq +\infty) = \int_{-\infty}^{+\infty} f_X(x)dx$$

et bien évidemment $P(-\infty \leq X \leq +\infty) = 1$, donc toute fonction de densité doit avoir une intégrale entre $-\infty$ et $+\infty$ égale à 1. Il faut bien penser à le vérifier.

Interprétation graphique : La probabilité $P(a \leq X \leq b)$ correspond à l'aire du domaine situé sous le graphe de f_X entre les abscisses a et b (voir figure 2.4).

Remarque 2.3. *on a bien $P(X = a) = P(a \leq X \leq a) = \int_a^a f_X(x)dx = 0$. Du coup,*

$$P(a \leq X \leq b) = P("a < X \leq b" \cup "X = a") = P(a < X \leq b) + P(X = a) = P(a < X \leq b).$$

Pour des raisons similaires,

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b).$$

Il n'est pas important de préciser si les frontières de l'intervalle sont incluses ou non dans ce genre de calcul, pour des variables a densité.

Remarque 2.4. *Calculer la loi d'une variable à densité, c'est calculer sa densité (de la même manière que pour calculer la loi d'une variable discrète, il fallait calculer les $P(X = x)$, pour $x \in \mathcal{S}(X)$).*

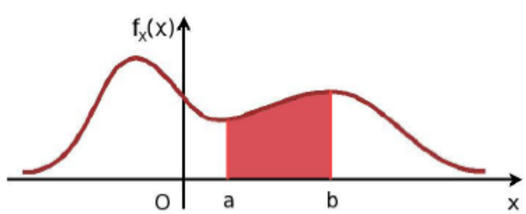

Figure 2.4 - Graphe d'une densité de probabilité f_X. La partie colorée correspond à $P(a \leq X \leq b)$.

2.3.2 Variable gaussienne

Je relève chaque jour la différence X entre la température extérieure et la moyenne saisonnière.

L'ensemble des valeurs possibles pour cette variable aléatoire est $\mathbb{R}$. En suivant le même raisonnement qu'auparavant, $P(X = t) = 0$ pour $t \in \mathbb{R}$, et il faut ici aussi considérer des intervalles, et une fonction de densité pour étudier cette variable. La distribution gaussienne est souvent utilisée pour ce type de variable qui a en général des petites variations, positives et négatives, et très rarement de grandes variations.

On appelle variable gaussienne de déviation standard σ^2 la variable continue X avec la densité

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}.$$

$\star$ Cette densité est souvent utilisée par les mathématiciens car elle a des bonnes propriétés mathématiques et correspond à ce que l'on observe pour de nombreux phénomènes naturels.
$\star$ On divise par $\sqrt{2\pi\sigma^2}$ car $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} = \sqrt{2\pi\sigma^2} \neq 1$, donc il faut diviser pour avoir une vraie densité de probabilité (i.e. une fonction positive dont l'intégrale vaut 1)
On note cette loi $\mathcal{N}(0;\sigma^2)$. Voici la courbe de la fonction f_X pour $\sigma = \sqrt{2}$:

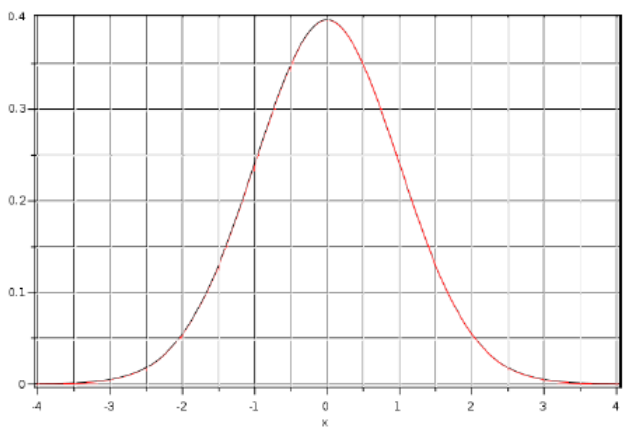

*Figure*2.5 *– la courbe de Gauss pour* $\sigma = \sqrt{2}$

On l'appelle courbe de Gauss, ou courbe en cloche. Elle présente la particularité d'avoir une grosse bosse au milieu, et de décroître ensuite très rapidement vers 0. Cela implique que la variable aléatoire qui a cette densité à de plus forte chances de prendre des valeurs proches de 0, ou plus précisément dans l'intervalle $[-\sigma, \sigma]$.

Pour savoir quelle est la probabilité que ma température reste sous un seuil donné, on peut utiliser la fonction de répartition :

$$P(X \leq t) = \int_{-\infty}^{t} f_X(x)dx = \int_{-\infty}^{t} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx.$$

En supposant qu'on mesure la température en degrés celsius, on choisit $\sigma = 5$ (on verra plus tard comment choisir une valeur pertinente pour σ), ce qui veut dire que la température varie en moyenne de $\pm 5°$ de la moyenne. Remarquons que si on définit $Y = \frac{X}{5}$, alors Y suit des variations d'amplitude égale à 1. On choisit donc comme loi $Y \sim \mathcal{N}(0;1)$. On a renormalisé X pour obtenir une loi gaussienne de "variabilité" 1 (on appellera en fait ce nombre la "variance").

On appelle cette loi la loi Gaussienne standard, ou loi normale, ou loi gaussienne centrée réduite.

Si maintenant on appelle Z la température "tout court", on a $Z = m + X$, où $m = 17$ est la moyenne saisonnière (par exemple...). La valeur moyenne de Z est m et on note la loi de Z par $N(m;\sigma^2)$.

En conclusion :

$\star$ $Y \sim \mathcal{N}(0;1)$ signifie que la densité de Y est $f_Y(y) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$.

$\star$ Pour $m \in \mathbb{R}, \sigma > 0$, on note $\mathcal{N}(m;\sigma^2)$ la loi de la variable $Z = m + \sigma Y$. Elle a pour densité

$$f_Z(z) = \frac{e^{-\frac{(z-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

Proposition 2.2. *Soit $X \sim \mathcal{N}(0,1)$, et $\sigma > 0$. Alors $\sigma X \sim \mathcal{N}(0,\sigma^2)$. De plus pour $m \in \mathbb{R}$, $m + \sigma X \sim \mathcal{N}(m;\sigma^2)$*

Exemple Un joueur joue aux fléchettes. Soit X la distance entre sa fléchette et le centre de la cible. Alors il est naturel de modéliser X par une loi $\mathcal{N}(0,\sigma^2)$, cela signifie que la fléchette va être typiquement à une distance plus ou moins σ du centre de la cible, et plus rarement à de plus larges distances. Si le joueur est mauvais, on peut mettre $\sigma = 50$ (on compte en centimètres), et s'il est très bon, on peut mettre $\sigma = 2$ (toujours en cm).

2.3.3 Loi exponentielle

Un pêcheur met son hameçon à l'eau. Il appelle X le temps (aléatoire) qu'il devra attendre avant qu'un poisson s'attaque à son appât. On voit bien que X est continue car la probabilité $P(X = t) = 0$ pour un temps donné $t > 0$. On peut aussi dire que X est une variable aléatoire sans mémoire car elle satisfait la condition suivante : sachant qu'aucun poisson n'a mordu au temps t, la probabilité qu'il morde dans les prochaines minutes ne dépend pas de t. Autrement dit, ça n'est pas parce que vous n'avez attrapé aucun poisson en une heure dans un lac rempli de poissons que vous avez plus de chances d'y arriver que vous n'en aviez au départ. Cette propriété ressemble à celle des variables discrètes de loi géométriques : si on lance le dé 100 fois sans faire de 6, la probabilité qu'on fasse 6 au 101 ème coup n'a pas changé, elle est toujours de 1/6.

Cette loi est en fait l'analogue continu de la loi géométrique. C'est la loi du temps que mettra un évènement à survenir sachant que la probabilité qu'il survienne ne change pas avec le temps.

Soit $a > 0$ un réel. On dit que X suit la loi exponentielle de paramètre a si elle admet la densité

$$f_X(x) = ae^{-ax}1_{\mathbb{R}^+} = \begin{cases} ae^{-ax} & \text{si } x \in \mathbb{R}^+, \\ 0 & \text{sinon.} \end{cases}$$

On note $X \sim \mathcal{E}(a)$. On vérifie que

$$\int_{-\infty}^{+\infty} f_X(x) = a\int_0^{+\infty} e^{-ax}dx = a\big[\frac{e^{-ax}}{-a}\big]_0^{+\infty} = a(0 - (-1/a)) = 1.$$

Si on divisait le temps en minutes au lieu de considérer tous les temps possibles, la loi Y de la première minute où un poisson mord est pour $k \in \mathbb{N}$

$$P(Y = k) = p(1-p)^{k-1}$$

où p est la probabilité qu'un poisson morde dans une minute donnée.

Autrement dit :

⋆ Le premier temps où l'on va attraper un poisson suit une loi exponentielle.

⋆ Le nombre de minutes (entières) à attendre avant d'attraper un poisson suit une loi géométrique.

Pour $X \sim \mathcal{E}(a)$, sa fonction de répartition est :

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)dt = \int_{-\infty}^{x} ae^{-ax}1_{\mathbb{R}^+} = \begin{cases} 1 - e^{-ax} & \text{si } x \in \mathbb{R}^+, \\ 0 & \text{sinon.} \end{cases}$$

Il est plus simple de retenir que pour une variable exponentielle,

$$P(X > x) = 1 - P(x \le x) = 1 - F_X(x) = \begin{cases} e^{-ax} & \text{si } x \in \mathbb{R}+, \\ 1 & \text{sinon.} \end{cases}$$

On peut déduire facilement que la densité $f_X(x) = F_X^{'}(x)$:

$$F_X^{'}(x) = \big[\int_{-\infty}^{x} f_X(t)dt\big]^{'} = f_X(x).$$

La figure 2.6 suivante montre la densités et la fonctions de répartition de la loi exponentielle pour deux valeurs différentes du paramètre a.

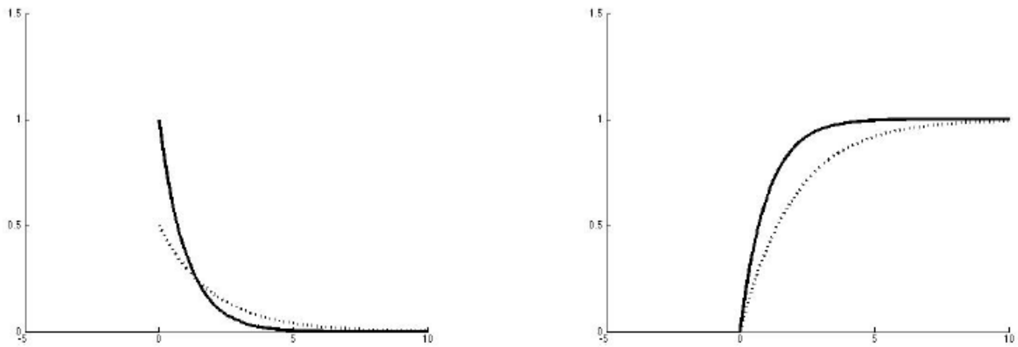

Figure 2.6 – Graphes de la densite et de la fonction derepartition de la loi exponentielle pour $a = 1$ (traits pleins) et $a = 0,5$ (pointilles).

La densite est la derivee de la fonction de repartition.

2.3.4 Fonction de répartition d'une variable aléatoire à densité

La fonction de répartition de la variable aléatoire X est par définition donnée comme suit :

$$F_X(x) = P(X \leq x) = P(-\infty < X \leq x) = \int_{-\infty}^{x} f_X(t)dt.$$

Ainsi F_X est une primitive de la fonction densité f_X, et f_X est donc la dérivée de F_X.
On répète que vouloir donner une valeur (autres que 0) à la quantité $P(X = x)$ n'a pas de sens, et en particulier il ne faut surtout pas écrire $P(X = x) = f_X(x)$.

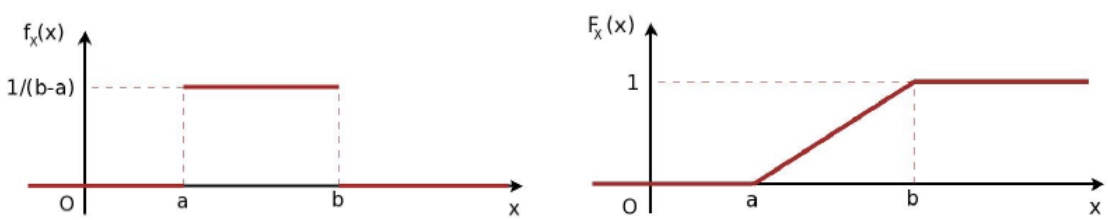

Figure 2.7 – Graphes de la densite et de la fonction de repartition de la loi uniforme sur un intervalle $[a, b]$.

Chapitre 3

Espérance et variance

3.1 Espérance