

Chapitre 1

Notions de base pour le calcul de probabilités

1.1 Introduction

Le calcul des probabilités est une branche de mathématiques dont le but est l'étude des phénomènes aléatoires c'est à dire des expériences dont le résultat ne peut être prédit avec certitude : par exemple, si on répète l'expérience plusieurs fois, on peut obtenir des résultats différents. Une forme d'indétermination apparaît dans l'issue de l'expérience. On peut interpréter cette forme de hasard comme notre incompetence à concevoir, expliciter et utiliser les phénomènes physiques considérés ou encore comme un manque d'information sur les conditions de l'expérience. Certains phénomènes sont peut-être par essence même sujets au hasard.

Les exemples de telles expériences sont nombreux : on peut penser au jeu de pile ou face, au lancer d'un dé, au sexe d'un enfant à naître, au temps d'attente d'un client à la poste, à la durée de vie d'une particule radioactive.

Pour étudier ces phénomènes, on doit tout d'abord en faire un modèle mathématique. Le modèle retenu pour décrire les expériences aléatoires est celui d'un triplet communément noté $(\Omega, \mathcal{F}, \mathbb{P})$ et appelé espace probabilisé ou espace de probabilité.

1.2 Axiomes du calcul des probabilités

1.2.1 Notion d'expérience aléatoire

Lorsqu'on étudie un phénomène aléatoire il est possible de l'assimiler à une expérience aléatoire c'est à dire si on répète plusieurs fois de suite la même expérience dans des conditions bien déterminées alors le résultat de cette expérience varie et semble obéir à des considérations stochastiques en un mot au hasard. Dans ce cas on dit qu'on est face à une expérience aléatoire , notée habituellement par la lettre ζ .

Exemples

1.a) Jeu de dés :

On jette sur une surface plane deux dés de couleurs différents et dont les six faces sont numérotées de 1 à 6. A la fin du jet on enregistre les deux chiffres qui apparaissent sur les deux faces supérieures des dés. Ces deux chiffres représentent le résultat de l'expérience.

1.b) Jeu de pile ou face :

On jette sur une surface plane deux pièces de monnaie et on registre ce qu'on voit sur la face supérieure de chacune des deux pièces. Si on pose **F** pour face et **P** pour pile alors le résultat de l'expérience est exprimé par les deux lettres **F, P**.

1.c) Tirage sans remise

On dispose d'une urne contenant N boules dont r sont de couleur rouge et n de couleur noir. L'expérience consiste à retirer de l'urne deux boules successivement et sans remise de la boule retirée. Le résultat de l'expérience est la couleur de la boule obtenue dans le premier choix et celle obtenue dans le second, dans cet ordre.

1.d) Mesure de la taille On enregistre la taille d'un individu choisi au hasard dans une population statistique bien déterminée. Le résultat de l'expérience est un nombre réel positif (bien entendu, après avoir choisi une unité de mesure).

1.2.2 L'ensemble Fondamental Ω .

On associe à l'expérience aléatoire ζ un ensemble fondamental noté Ω et dont les éléments représentent tous les résultats possibles de l'expérience ζ .

Exemples.

- 1.a) $\Omega = \{ (i, j) ; i \text{ et } j \text{ sont deux nombres entiers compris entre 1 et 6} \}$.
- 1.b) $\Omega = \{ PP, FF, PF, FP \}$
- 1.c) $\Omega = \{ (a, b) ; a \text{ représente la couleur de la boule obtenue dans le } 1^{er} \text{ tirage et } b \text{ celle du } 2^{eme} \text{ tirage} \}$
- 1.d) $\Omega =]0, +\infty [$. On peut choisir un intervalle fermé de \mathbb{R}_+ .

Remarque 1.2.1. *Le choix de l'ensemble fondamental Ω se fait selon le cas étudié. Par exemple, dans le jeu de dés on peut supposer que chacun peut se positionner sur chacune de ses arêtes. Dans ce cas on aura des autres issues supplémentaires possibles.*

1.2.3 Notion d'évènements

Considérons une propriété Δ liée à l'issue (résultat de l'expérience aléatoire ζ). A chaque accomplissement de ζ on a deux cas : la propriété Δ est réalisée ou bien elle n'est pas réalisée. Donc, grâce à cette propriété on a partagé l'ensemble fondamental Ω en deux parties disjointes : d'un côté l'ensemble E formé par l'ensemble des points ω , $\omega \in \Omega$, représentant des résultats de ζ avec réalisation de la propriété Δ et de l'autre côté l'ensemble complémentaire \bar{E} dans Ω comportant les points ω de Ω qui correspondent à des résultats de ζ ne réalisant pas la propriété Δ . On dit que l'ensemble E est l'évènement lié à la propriété Δ . On dit aussi que E est l'évènement " E est réalisé". On voit immédiatement que \bar{E} (complémentaire de E dans Ω) est également un évènement auquel on fait référence en écrivant " E est non réalisé".

EXEMPLES.

1.3.a) Dans l'exemple **1.a)** l'évènement E = "la somme des chiffres enregistrés est un chiffre impair" s'écrit comme suit :

$$E = \{(i, j) \in \Omega / i + j = \text{impair}\} = \{(1, 2), (2, 1), (1, 4), (4, 1), (1, 6), (6, 1), (2, 3), (3, 2), \dots\}.$$

1.3.b) Dans le jeu de pile ou face (**exemple 1.b)**) l'évènement E = "On obtient face **F** une fois au moins" s'écrit : $E = \{FF, FP, PF\}$.

1.2.4 Types d'évènements**Evènement certain**

C'est l'évènement lié à une propriété qui est toujours réalisée lorsqu'on effectue l'expérience ζ . Il est représenté par l'ensemble Ω tout entier.

Evènement impossible

C'est l'évènement lié à une propriété qui n'est jamais réalisée; il est représenté par la partie vide de Ω , notée ϕ .

Evènement complémentaire

Si E est un évènement alors la partie \bar{E} complémentaire de la partie E dans Ω est aussi un évènement appelé évènement complémentaire de E , on dit également l'évènement "non E ".

Evènement élémentaire

Les évènements liés à la même expérience ζ et dont les réalisations n'ont lieu que pour une seule issue de ζ s'appellent les évènements élémentaires; ceux sont les parties unitaires de Ω .

1.3 Opérations sur les évènements

On prolonge dans la suite le parallélisme entre les notions ensemblistes et probabilistes. Alors en outre de l'opération de complémentarité définie précédemment il existe aussi d'autres opérations importantes sur l'ensemble fondamental Ω lié à la même expérience ζ .

a) Inclusion

Soient A et B deux évènements de Ω . Si l'expérience qui réalise l'évènement A réalise forcément l'évènement B on dit que A implique B et on écrit $A \subset B$.

b) Intersection.

Soient A et B deux sous ensembles représentant deux évènements de Ω . L'évènement " A et B " est exprimé par l'intersection des sous ensembles A et B de Ω appelé évènement intersection de A et B et on écrit : " A et B " = $A \cap B$.

Définition 1. Si $A \cap B = \phi$ on dit que les évènements A et B sont incompatibles c'est à dire leur réalisation simultanément est impossible.

c) Réunion.

Si A et B sont deux évènements de Ω alors l'évènement " A ou B " est l'évènements représenté par la réunion des deux sous-ensembles A et B de Ω . On écrit " A ou B " = $A \cup B$.

EXEMPLES

1.4.a) Reprenons l'exemple **1.a)** sur le jeu de dés et considérons les évènements suivants :

A = "les deux chiffres enregistrés sont impairs"

B = "les deux chiffres enregistrés sont pairs"

C = "l'un des chiffres enregistrés est pair et l'autre est impair"

Alors on a :

\overline{A} = "au moins l'un des deux chiffres enregistrés est pair"

\overline{B} = "au moins l'un des deux chiffres est impair"

$A \cap B = \phi$; $A \cup B$ = "les deux chiffres enregistrés sont ou bien pairs ou bien impairs"; $C = \overline{A \cup B}$ = "l'un des deux chiffres est pair et l'autre est impair",
 $\overline{A} = B \cup C$, $\overline{B} = A \cup C$, $B \subset \overline{A}$; $A \subset \overline{B}$, $\overline{A} \cap \overline{B} = C$.

1.4.b) Dans l'exemple **1.b)** (Jeu de pile ou face avec 2 pièces) l'évènement élémentaire $\{PP\}$ est l'intersection de l'évènement $A = \{PF, PP\}$ avec l'évènement $B = \{PP, FP\}$. L'évènement $A \cup B$ est : "pile au moins une fois" ,
 $\overline{A \cup B}$ = "face deux fois" = $\{FF\}$.

On va appeler donc évènements certains sous-ensembles de Ω mais pas toujours tous. En revanche, si A et B sont deux "évènements intéressants", alors nous constatons que \overline{A} et $A \cup B$ le sont aussi. On va donc définir des espaces d'évènements stables par passage au complémentaire ou par réunion dénombrable.

Définition 2. *Supposons Ω un ensemble non vide. On appelle tribu sur Ω (ou σ -algèbre) tout ensemble de parties \mathcal{F} de Ω vérifiant les trois conditions suivantes :*

1. $\Omega \in \mathcal{F}$

2. $\forall A, A \in \mathcal{F} \implies \overline{A} \in \mathcal{F}$ (On dit que \mathcal{F} est stable par complémentaire)

3. Pour toute suite $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$ (On dit que \mathcal{F} est stable par union dénombrable).

EXEMPLES

Soit Ω un ensemble non vide. $\mathcal{F}_1 = \{\emptyset, \Omega\}$ est une tribu appelée tribu grossière; $\mathcal{F}_2 = \mathcal{P}(\Omega)$ la collection des parties de Ω , est une tribu; enfin, si $A \subset \Omega$, $\mathcal{F}_3 = \{A, \overline{A}, \emptyset, \Omega\}$ est une tribu sur Ω .

Remarque 1.3.1. *Si \mathcal{F} est une tribu sur Ω , alors $\emptyset \in \mathcal{F}$. De plus, \mathcal{F} est stable par union finie, par intersection finie et dénombrable, par différence . . . Expliquons par exemple pourquoi \mathcal{F} est stable par intersection dénombrable. Soit donc $(A_n)_{n \in \mathbb{N}} \in$*

$\mathcal{F}^{\mathbb{N}}$ on a $\bigcap_{n \in \mathbb{N}} A_n = \overline{\left(\bigcup_{n \in \mathbb{N}} \bar{A}_n \right)}$. Comme, pour tout n , $A_n \in \mathcal{F}$, la condition (2) de la définition implique que $\bar{A}_n \in \mathcal{F}$ pour tout n . La condition (3) entraîne donc que $\bigcup_{n \in \mathbb{N}} \bar{A}_n \in \mathcal{F}$; on peut alors appliquer à nouveau le point (2) pour obtenir que $\overline{\left(\bigcup_{n \in \mathbb{N}} \bar{A}_n \right)} \in \mathcal{F}$.

Définition 3. Soit Ω l'ensemble fondamental lié à l'expérience aléatoire ζ et \mathcal{F} une tribu sur Ω alors le couple (Ω, \mathcal{F}) s'appelle **Espace Probabilisable**. Les éléments de \mathcal{F} sont appelés évènements et ceux de Ω évènements élémentaires.

EXEMPLES

Dans les exemples 1.a), 1.b) et 1.c) on peut prendre $\mathcal{F} = \mathcal{P}(\Omega)$ et on obtient directement l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$.

Remarque 1.3.2. On aimerait toujours travailler avec la tribu $\mathcal{F} = \mathcal{P}(\Omega)$ Mais ceci n'est pas possible sauf si Ω est un ensemble fini ou dénombrable. Si $\Omega = \mathbb{R}$, il n'est pas possible de mesurer toutes les parties de \mathbb{R} sans aboutir à une contradiction et on se limite à une classe de parties appelée tribu borélienne.

Définition 4. On appelle tribu borélienne sur \mathbb{R} , notée $\mathbb{B}(\mathbb{R})$, la plus petite tribu, au sens de l'inclusion, contenant tous les intervalles de \mathbb{R} .

Cette définition appelle un commentaire : il n'est pas évident à priori que l'on puisse parler de « la plus petite tribu » contenant les intervalles. Toutefois, la définition a bien un sens car l'intersection d'une famille quelconque de tribus sur Ω est encore une tribu sur Ω . On peut donc considérer l'intersection de toutes les tribus sur \mathbb{R} contenant les intervalles qui devient par construction « la plus petite tribu ».

Définition 5. On appelle système complet d'évènements toute suite E_1, E_2, \dots, E_n d'évènements de \mathcal{F} deux à deux incompatibles et telle que $\bigcup_{i=1}^n E_i = \Omega$. Les ensembles $(E_i)_{1 \leq i \leq n}$ forment une partition de Ω .

1.4 Notion de Probabilité

Définition 6. Soient (Ω, \mathcal{F}) un espace probabilisable et \mathbb{P} une application de \mathcal{F} dans $[0, 1]$. \mathbb{P} est une probabilité sur (Ω, \mathcal{F}) si elle vérifie les conditions suivantes :

1. $\mathbb{P}(\Omega) = 1$
2. Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{F} (événements) deux à deux incompatibles c à d $A_n \cap A_m = \emptyset$ pour $n \neq m$, on a :

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$$

cette propriété s'appelle la σ -additivité.

Définition 7. Soient (Ω, \mathcal{F}) un espace probabilisable et \mathbb{P} une probabilité sur cet espace. Le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ s'appelle un espace probabilisé ou espace de probabilités.

Remarque 1.4.1. Pour toute expérience aléatoire décrite par un espace probabilisable (Ω, \mathcal{F}) il existe un grand nombre de probabilités \mathbb{P} possibles qu'on peut définir sur cet espace. Mais, dans la pratique le choix de la probabilité \mathbb{P} est déterminé soit par des conditions naturelles soit par des considérations expérimentales.

Exemple 1.4.1. On lance un dé sur une surface plane et le résultat auquel on s'intéresse est le nombre indiqué par le dé, l'ensemble fondamental sera typiquement $\Omega = \{1, 2, 3, 4, 5, 6\}$. On utilise comme tribu, l'ensemble $\mathcal{P}(\Omega)$ et, dans le cas d'un dé non pipé, on définit la loi de probabilité \mathbb{P} de telle sorte que chaque singleton ait la même probabilité et il est facile de voir que $\mathbb{P}\{i\} = \frac{1}{6}$ pour tout $i \in \Omega$. Il est aisé de vérifier que cela détermine entièrement \mathbb{P} , et que l'on a alors, pour tout événements $A \in \mathcal{P}(\Omega)$, $\mathbb{P}(A) = \frac{\text{Card}(A)}{6}$.

La définition d'une probabilité conduit aux propriétés suivantes :

Proposition 1.4.1. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Si A, B sont deux événements de \mathcal{F} alors,

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

2. *Propriété d'additivité* : Si $A \cap B = \emptyset$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
3. $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$
4. *Croissance* : Si $A \subset B$, $\mathbb{P}(A) \leq \mathbb{P}(B)$
5. Si $A \subset B$, $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$; en particulier : $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
6. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Preuve. 1. Pour obtenir $\mathbb{P}(\emptyset) = 0$ il suffit de prendre $A_0 = \Omega$ et $A_n = \emptyset$ pour $n \geq 1$ dans la définition et d'appliquer la σ -additivité.

2. Ici il suffit aussi de prendre $A_0 = A$, $A_1 = B$ et $A_n = \emptyset$ pour $n \geq 2$ dans la définition et d'appliquer la σ -additivité.

3. On écrit $B = (B \setminus A) \cup (A \cap B)$ et on utilise l'additivité $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$.

4. et 5. Pour ces deux points on applique le point 3. pour $A \subset B$.

6. On peut écrire $A \cup B = A \cup B \setminus A$ et on utilise l'additivité et le points 3. ■

1.5 Quelques espaces de probabilité importants

1.5.1 L'espace Ω est fini ou dénombrable

Comme on a vu précédemment, dans ce cas, on suppose habituellement que la tribu des évènements \mathcal{F} est $\mathcal{P}(\Omega)$, l'ensemble de toutes les parties de Ω et donc l'espace probabilisable sera $(\Omega, \mathcal{P}(\Omega))$. Les probabilités qu'on peut définir sur cet espace sont alors décrites par le résultat suivant

Proposition 1.5.1. *Soit Ω un ensemble fini ou dénombrable. Soit $\omega \mapsto p_\omega$ une application de Ω dans \mathbb{R}_+ telle que*

$$\sum_{\omega \in \Omega} p_\omega = 1$$

Pour tout $A \subset \Omega$, notons alors :

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$$

Alors $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ est un espace de probabilité (c à d l'application \mathbb{P} définie par la dernière formule est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$). Inversement, toute probabilité \mathbb{P} sur $(\Omega, \mathcal{P}(\Omega))$ est du type précédent, avec $p_\omega = \mathbb{P}(\{\omega\})$.

1.5.2 Le cas équiprobable

Considérons le cas particulier de la proposition 2.5.1 où Ω est fini avec $\text{card}\Omega = n$ et où les événements élémentaires sont équiprobables (c'est à dire qu'ils ont tous la même probabilité). Dans ce cas il est alors facile de déterminer la probabilité \mathbb{P} . En effet comme les $p_\omega = \mathbb{P}(\{\omega\})$ sont tous égaux, alors il est facile de voir que :

$$\forall \omega \in \Omega, p_\omega = \frac{1}{\text{card}\Omega} = \frac{1}{n}$$

Alors, dans ce cas, si $A \subset \Omega$ on a :

$$\boxed{\mathbb{P}(A) = \sum_{\omega \in A} p_\omega = \frac{\text{card}A}{\text{card}\Omega} = \frac{\text{nombre de cas favorable}}{\text{nombre de cas total}}}$$

Ainsi dans ce cas, les calculs de probabilités se ramènent à des problèmes de dénombrement. Cette probabilité \mathbb{P} s'appelle probabilité uniforme sur Ω .

Exemple 1.5.1. Soit l'expérience aléatoire $\zeta =$ "on jette un dé deux fois". Alors $\Omega = \{(i, j); 1 \leq i \leq 6 \text{ et } 1 \leq j \leq 6\}$. Si le dé est bien équilibré, alors il est tout à fait naturel de dire que les événements élémentaires sont équiprobables autrement dit $P(\{(i, j)\}) = \frac{1}{36}$ pour tout $(i, j) \in \Omega$. Si A est l'évènement "la somme des chiffres enregistrés est égale à 6", donc $P(A) = \sum_{(i,j) \in A} P(\{(i, j)\}) = \frac{5}{36}$, car ici $A = \{(1, 5), (5, 1), (2, 4), (4, 2), (3, 3)\}$.

Remarque 1.5.1. Attention : un évènement impossible a la probabilité 0, un évènement certain a la probabilité 1. La réciproque n'est pas vraie.

Pour montrer cela de façon plus claire on considère l'expérience $\zeta =$ "on cherche une pièce de monnaie cachée dans une des mains de plusieurs personnes et on s'arrête à sa découverte". Donc Ω est constitué de toutes les suites du type $(EEE...ER)$ où

E désigne un échec dans la tentative de découvrir la pièce et R une réussite. Dans l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ on pose : Pour tout $\omega \in \Omega$ tel que $\omega = \underbrace{EE\dots ER}_{n+1 \text{ éléments}}$: $\mathbb{P}(\{\omega\}) = q^n p$, $n \in \mathbb{N}$ avec $q = \mathbb{P}(E)$ et $p = \mathbb{P}(R)$; $q = 1 - p$ ($0 < p < 1$ est la probabilité de trouver la pièce dans chaque tentative). Il est facile de voir que : $\sum_{n=0}^{+\infty} q^n p = p \frac{1}{1-q} = 1$. Si A représente l'évènement "on découvre la pièce au troisième essai au moins" alors $\mathbb{P}(A) = \sum_{n=2}^{+\infty} q^n p = q^2$. On remarque que l'évènement $(EE\dots E\dots E\dots)$ a une probabilité nulle par rapport à la probabilité définie pourtant, la possibilité que cet évènement ait lieu existe puisqu'il fait partie de l'ensemble $\mathcal{P}(\Omega)$ (mais cet évènement est irréalisable expérimentalement).

Terminologie concernant les évènements

- Si $\mathbb{P}(A) = 0$, l'évènement A est dit négligeable.
- Si $\mathbb{P}(A) = 1$, l'évènement A est dit presque sûr.

1.5.3 Cas où $\Omega = \mathbb{R}$

Ce cas est naturellement le plus important de tous. La tribu mise sur \mathbb{R} est la tribu de Borel $\mathbb{B}(\mathbb{R})$ définie à la section **2.3** comme la plus petite tribu contenant les intervalles (ouverts, fermés, semi ouverts, demi droites).

Pour décrire les probabilités sur $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$, introduisons une définition importante :

Définition 8. Soit F une fonction de \mathbb{R} dans \mathbb{R} : On dit que F est une fonction de répartition si elle satisfait aux trois propriétés suivantes :

- F est croissante (au sens large) ;
- $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$;
- F est continue à droite en tout point x , c'est-à-dire $\lim_{h \searrow 0} F(x+h) = F(x)$:

On a alors le théorème fondamental suivant :

Théorème 1. Soit \mathbb{P} une probabilité sur $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$. Soit $F_{\mathbb{P}}$ la fonction réelle définie par :

$$\begin{aligned} F_{\mathbb{P}} &: \mathbb{R} \longrightarrow \mathbb{R} \\ F_{\mathbb{P}}(x) &= \mathbb{P}([-\infty, x]) \end{aligned}$$

Alors $F_{\mathbb{P}}$ est une fonction de répartition. Inversement, si F est une fonction de répartition, alors il existe une et une seule probabilité \mathbb{P} sur $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ telle que $F_{\mathbb{P}} = F$.

Commentaires : Ce résultat est assez rassurant : bien qu'on connaisse mal la tribu $\mathbb{B}(\mathbb{R})$, et donc les probabilités définies dessus, il y a en fait bijection entre l'ensemble de toutes les probabilités sur \mathbb{R} et l'ensemble moins abstrait de toutes les fonctions de répartition. Donc pour définir une probabilité sur $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ il nous suffit seulement de définir sa fonction de répartition.

La fonction de répartition permet de calculer les probabilités de tous les intervalles. Pour simplifier, adoptons la notation pour la limite à gauche en x de la fonction croissante F :

$$F(x-0) = \lim_{h \searrow 0} F(x-h)$$

Proposition 1.5.2. Soit F une fonction de répartition d'une probabilité \mathbb{P} sur $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$.

Alors :

$$*\mathbb{P}([-\infty, x]) = F(x-0), \quad \mathbb{P}([x, +\infty]) = 1 - F(x),$$

$$\mathbb{P}([x, +\infty]) = 1 - F(x-0),$$

Et pour $a \leq b$,

$$*\mathbb{P}([a, b]) = F(b) - F(a), \quad \mathbb{P}([a, b]) = F(b-0) - F(a-0).$$

$$*\mathbb{P}([a, b]) = F(b-0) - F(a), \quad \mathbb{P}([a, b]) = F(b) - F(a-0)$$

En particulier : $\mathbb{P}(\{a\}) = F(a) - F(a-0)$.

Donnons maintenant des exemples de fonctions de répartition

Définition 9. Fonctions de répartition à densité : Soit f une fonction positive définie sur \mathbb{R} ($\forall x \in \mathbb{R}, f(x) \geq 0$) qui ait des discontinuités au plus en un nombre fini

de points $a_1 < a_2 < \dots < a_N$ et qui soit telle que les intégrales $\int_{a_i}^{a_{i+1}} f(x) dx$ convergent et satisfassent :

$$\sum_{i=0}^N \int_{a_i}^{a_{i+1}} f(x) dx = 1$$

avec la convention $a_0 = -\infty, a_{N+1} = +\infty$ (c à d $\sum_{i=0}^N \int_{a_i}^{a_{i+1}} f(x) dx = \int_{-\infty}^{+\infty} f(x) dx$).

On définit alors la fonction F par $F(x) = \int_{-\infty}^x f(t) dt$. Il est clair que F est une fonction de répartition. Ici, elle est de plus continue et, d'après le théorème fondamental du calcul intégral, elle satisfait $F'(x) = f(x)$ pour tout $x \notin \{a_1, a_2, \dots, a_N\}$. La fonction f s'appelle alors la densité de la fonction de répartition F .

Exemple 1.5.2. Soient f_1 et f_2 les fonctions suivantes :

$$f_1(x) = \frac{1}{2}e^{-|x|}, \forall x \in \mathbb{R}, \quad f_2(x) = \begin{cases} 0, & x < 0 \text{ ou } x > 1 \\ 1, & 0 \leq x \leq 1 \end{cases}$$

on peut montrer facilement que f_1 et f_2 sont des fonctions de densité (pour les points de discontinuités $N = 0$ pour f_1 et $N = 2$ pour f_2). Les fonctions de répartition correspondantes sont respectivement : $F_1(x) = \begin{cases} \frac{1}{2}e^x, & x \leq 0 \\ 1 - \frac{1}{2}e^{-x}, & x \geq 0 \end{cases}$,

$$F_2(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x \geq 1 \end{cases}$$

Définition 10. La probabilité δ_a de Dirac : Si a est un réel, il s'agit de la probabilité définie sur \mathbb{R} par : $\delta_a(A) = \begin{cases} 0, & a \notin A \\ 1, & a \in A \end{cases}$. En appliquant ceci à $A =$

$]-\infty, x]$, on obtient la fonction de répartition : $F_{\delta_a}(x) = \begin{cases} 0, & x < a \\ 1, & x \geq a \end{cases}$.

Si $a = 0$, cette fonction s'appelle l'échelon de Heaviside. Les travaux de 1894 de cet ingénieur électricien sont à la source de la théorie moderne des distributions. Cette théorie permet par exemple de donner un sens à la dérivation de la fonction ci dessus : c'est la probabilité de Dirac δ_a qui jouerait alors le rôle de la dérivée.

1.6 Probabilités Conditionnelles et Indépendance

1.6.1 Probabilités Conditionnelles

La notion de probabilité conditionnelle permet de prendre en compte l'information dont on dispose pour actualiser la probabilité que l'on donne à un événement. Autrement dit, on se place dans un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, et soient deux événements A et B ayant chacun une probabilité non nulle d'être réalisé. On s'intéresse à la nouvelle probabilité de l'un d'eux si l'on est assuré que l'autre s'est produit.

Exemple 1.6.1. Soit l'expérience aléatoire $\zeta =$ "on jette un dé non pipé une seule fois". Alors $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, \mathbb{P} la probabilité uniforme. On pose A : "la face est paire" = $\{2, 4, 6\}$ et B : "La face est inférieure ou égale à 3" = $\{1, 2, 3\}$. On sait bien que $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$. Mais si l'on est assuré de la réalisation de A (i.e. que la face est paire), la probabilité de B semble tomber à $\frac{1}{3}$.

Maintenant, nous définissons formellement ce concept pour le rendre facilement exploitable dans la conduite des calculs de façon sûre et efficace.

Définition 11. Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et B un événement de probabilité non nulle, et soit A un événement. On appelle probabilité conditionnelle de A sachant B , et on note $\mathbb{P}(A/B)$ ou $\mathbb{P}_B(A)$ le nombre

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

Exemple 1.6.2. $M.$ et $Mme H$ ont deux enfants dont une fille. Quelle est la probabilité que l'autre soit un garçon? Même question s'il l'on suppose que la fille est l'aînée? Pour résoudre ce problème rigoureusement, il faut se construire un espace de probabilité adapté. On suppose qu'un nouveau né est un garçon ou une fille avec la même probabilité et indépendamment du sexe de ses aîné(e)s. De là, une famille de deux enfants est dans l'une des quatre configurations de $\Omega = \{FG, GF, FF, GG\}$, chacune avec une probabilité $\frac{1}{4}$. On suppose dans le premier cas que l'événement $A_1 = \{FG, GF, FF\}$ est réalisé et on cherche la probabilité qu'il y ait un garçon

(et une fille) dans la famille soit la probabilité conditionnelle $\mathbb{P}(B_1/A_1)$ avec $B_1 = \{FG, GF\}$. On obtient :

$$\mathbb{P}(B_1/A_1) = \frac{\mathbb{P}(B_1 \cap A_1)}{\mathbb{P}(A_1)} = \frac{\mathbb{P}(\{FG, GF\})}{\mathbb{P}\{FG, GF, FF\}} = \frac{2/4}{3/4} = \frac{2}{3}$$

Par contre, dans le deuxième cas, on cherche $\mathbb{P}(B_2/A_2)$ où $A_2 = \{FG, FF\}$ et $B_2 = \{FG\}$. On obtient $\mathbb{P}(B_2/A_2) = \frac{1}{2}$.

Proposition 1.6.1. Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et B un événement de probabilité non nulle. Alors l'application :

$$\begin{aligned} \mathbb{P}(\cdot/B) &: \mathcal{F} \longrightarrow \mathbb{R}_+ \\ &: A \longmapsto \mathbb{P}(A/B) \end{aligned}$$

est une probabilité sur \mathcal{F} .

Preuve. 1. On a $\mathbb{P}(\Omega/B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$.

2. Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements de \mathcal{F} deux à deux incompatibles c à d $A_n \cap A_m = \emptyset$ pour $n \neq m$, on a :

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n/B\right) = \frac{\mathbb{P}\left(\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap B\right)}{\mathbb{P}(B)} = \frac{\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} (A_n \cap B)\right)}{\mathbb{P}(B)}$$

les événements $(A_n \cap B)_{n \in \mathbb{N}}$ sont eux-mêmes deux à deux incompatibles et par suite

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n/B\right) = \frac{\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} (A_n \cap B)\right)}{\mathbb{P}(B)} = \sum_{n \in \mathbb{N}} \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n/B)$$

Donc $\mathbb{P}(\cdot/B)$ est une probabilité sur \mathcal{F} . ■

Proposition 1.6.2. (Formule des probabilités composées)

Si A et B sont deux événements de probabilités non nulles. Alors

$$\mathbb{P}(A \cap B) = \mathbb{P}(A/B) \mathbb{P}(B) = \mathbb{P}(B/A) \mathbb{P}(A)$$

Généralisation : soient A_1, A_2, \dots, A_n n événements de probabilités strictement positives. Alors :

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

Preuve. facile. ■

Proposition 1.6.3. (*Formule des probabilités totales*).

Soit B un événement tel que $\mathbb{P}(B) > 0$ et $\mathbb{P}(\bar{B}) > 0$. Alors

$$\forall A \in \mathcal{F} : \mathbb{P}(A) = \mathbb{P}(A/B) \mathbb{P}(B) + \mathbb{P}(A/\bar{B}) \mathbb{P}(\bar{B})$$

Généralisation : soient B_1, B_2, \dots, B_n un système complet d'événements pour Ω tel que $\mathbb{P}(B_j) > 0$ pour $j = 1, 2, \dots, n$. Alors :

$$\forall A \in \mathcal{F} : \mathbb{P}(A) = \sum_{j=1}^n \mathbb{P}(A/B_j) \mathbb{P}(B_j)$$

Preuve. On a $\forall A \in \mathcal{F} : A = A \cap \Omega$ et d'autre part $\Omega = \bigcup_{j=1}^n B_j$, alors :

$$\mathbb{P}(A) = \mathbb{P}(A \cap \Omega) = \mathbb{P}\left(A \cap \left(\bigcup_{j=1}^n B_j\right)\right) = \mathbb{P}\left(\bigcup_{j=1}^n (A \cap B_j)\right)$$

les événements $(A \cap B_j)_{j=1, \dots, n}$ sont deux à deux incompatibles, alors

$$\mathbb{P}(A) = \sum_{j=1}^n \mathbb{P}(A \cap B_j) = \sum_{j=1}^n \mathbb{P}(A/B_j) \mathbb{P}(B_j)$$

■

Proposition 1.6.4. (Formule de Bayes). Soient (B_1, B_2, \dots, B_n) un système complet d'événements pour Ω et A un événement de Ω de probabilité strictement positive.

Alors :

$$\forall i \in \{1, 2, \dots, n\} : \mathbb{P}(B_i/A) = \frac{\mathbb{P}(A/B_i) \mathbb{P}(B_i)}{\sum_{j=1}^n \mathbb{P}(A/B_j) \mathbb{P}(B_j)}$$

Preuve. $\forall i \in \{1, 2, \dots, n\} : \mathbb{P}(B_i/A) = \frac{\mathbb{P}(B_i \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A/B_i) \mathbb{P}(B_i)}{\sum_{j=1}^n \mathbb{P}(A/B_j) \mathbb{P}(B_j)}$ ■

Exemple 1.6.3. Pour dépister une maladie, on applique un test sanguin. Si le patient est atteint, le test donne un résultat positif dans 99% des cas. Mais il se peut aussi que le résultat du test soit positif alors que le patient est en bonne santé et cela se produit dans 2% des cas. La proportion de personnes malades dans la population soumise au test est de 10^{-3} . Calculer la probabilité pour qu'un patient soit en bonne santé sachant que le résultat de son test est positif. Pour résoudre ce problème on considère les événements suivants :

M : " Le patient est atteint par la maladie ",

T : " Le résultat du test est positif "

Alors on a : $\mathbb{P}(M) = 0,001$, $\mathbb{P}(T/M) = 0,99$ et $\mathbb{P}(T/\overline{M}) = 0,02$ et on veut calculer $\mathbb{P}(\overline{M}/T)$?

Calculons d'abord $\mathbb{P}(T)$: on a $\{M, \overline{M}\}$ est un système complet d'événements pour Ω alors d'après la **formule des probabilités totales** :

$$\begin{aligned} \mathbb{P}(T) &= \mathbb{P}(T/M) \mathbb{P}(M) + \mathbb{P}(T/\overline{M}) \mathbb{P}(\overline{M}) \\ &= 0,99 \times 0,001 + 0,02 \times (1 - 0,99) = 0,00119 \end{aligned}$$

D'après la **Formule de Bayes**

$$\mathbb{P}(\overline{M}/T) = \frac{\mathbb{P}(T/\overline{M}) \mathbb{P}(\overline{M})}{\mathbb{P}(T/M) \mathbb{P}(M) + \mathbb{P}(T/\overline{M}) \mathbb{P}(\overline{M})} = \frac{0,02 \times 0,999}{0,00119} = 0,168$$

1.6.2 Indépendance stochastique des évènements

Indépendance de deux évènements :

Idée : Si A et B sont deux évènements alors on dit qu'ils sont indépendants si la connaissance de la réalisation de l'un ne change en rien la probabilité de la réalisation de l'autre, soit « $\mathbb{P}(A|B) = \mathbb{P}(A)$ ou bien $\mathbb{P}(B|A) = \mathbb{P}(B)$ ». Pour pouvoir définir cette notion y compris pour les évènements de probabilité nulle, on a la définition suivante :

Définition 12. Deux évènements A et B sont dits indépendants si et seulement si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$$

On obtient souvent des évènements indépendants lorsqu'on reproduit une expérience sans que la première expérience n'interfère avec la seconde. C'est par exemple le cas lorsque l'on joue deux fois à pile ou face. Voici un autre exemple.

Exemple 1.6.4. Si on lance deux fois un dé équilibré, alors $\Omega = E \times E$ avec $E = \{1, 2, 3, 4, 5, 6\}$. Comme le dé est équilibré alors tous les évènements élémentaires ont la même probabilité. Définissons les évènements suivants : $A = \{(i, j), i \text{ pair}\}$, $B = \{(i, j), j \text{ pair}\}$, $C = \{(i, j), i + j \text{ pair}\}$. Il est clair que $P(A) = P(B) = P(C) = \frac{1}{2}$, et $P(A \cap B) = P(B \cap A) = P(A \cap C) = P(C \cap A) = P(B \cap C) = P(C \cap B) = \frac{1}{4}$. Par conséquent les évènements A, B et C sont deux à deux indépendants.

Remarque 1.6.1. (Attention) : Il ne faut pas confondre incompatibilité avec indépendance de deux évènements. Si A et B sont deux évènements incompatibles alors ils ne sont indépendants que si $P(A) = 0$ ou $P(B) = 0$. Lorsque $P(A) \neq 0$ et $P(B) \neq 0$ ils sont toujours dépendants.

Proposition 1.6.5. Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, A et B deux évènements de \mathcal{F} , alors on a :

A et B indépendants $\iff \bar{A}$ et B indépendants $\iff A$ et \bar{B} indépendants $\iff \bar{A}$ et \bar{B} indépendants.

Preuve. On démontre seulement : A et B indépendants $\iff \bar{A}$ et B indépendants (de la même manière pour les autres). On a $B = (A \cap B) \cup (\bar{A} \cap B)$ donc $\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(\bar{A} \cap B)$ (car $(A \cap B)$ et $(\bar{A} \cap B)$ sont incompatibles), alors on a : A et B indépendants $\iff \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \iff \mathbb{P}(\bar{A} \cap B) = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(B)(1 - \mathbb{P}(A)) = \mathbb{P}(B)\mathbb{P}(\bar{A}) \iff \bar{A}$ et B indépendants. ■

Evènements indépendants dans leur ensemble

On se place dans un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et soit une famille finie ou infinie $(A_i)_{i \in I}$ d'évènements de \mathcal{F} . Dire que ces évènements sont indépendants deux à deux équivaut à écrire : $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ pour tout $i, j \in I$. Mais, il y a aussi l'éventualité où $(A_i)_{i \in I}$ sont indépendants (on dit aussi indépendants dans leur ensemble) celle-ci est donnée par la définition suivante :

Définition 13. On dit que les évènements $(A_i)_{i \in I}$ sont indépendants dans leur ensemble si pour tout sous-ensemble $J \subseteq I$, on a

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)$$

Cette condition est plus stricte que l'indépendance des évènements deux à deux, l'indépendance deux à deux correspond à la restriction de cette condition aux sous-ensembles J de cardinal 2. Le piège est le suivant : si A, B, C sont trois évènements indépendants alors (A, B) , (A, C) et (B, C) sont des couples d'évènements indépendants mais la réciproque est fautive. C'est à dire l'indépendance mutuelle de plusieurs évènements **implique** leur indépendance deux à deux mais la réciproque est fautive comme le montre l'exemple suivant :

Exemple 1.6.5. Considérons l'expérience aléatoire $\zeta =$ " On jette deux fois un dé sur une surface plane" et les évènements suivants : $A_1 =$ " on a obtenu le chiffre 1 dans le premier jet", $A_2 =$ " on a obtenu le chiffre 1 dans le deuxième jet" et $A_3 =$ " on a obtenu le même chiffre dans les deux jets". Il est clair que les évènements A_1, A_2 et A_3 sont indépendants deux à deux. En outre $P(A_1) = P(A_2) = P(A_3) = \frac{1}{6}$ et

$A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3 =$ "on a obtenu le chiffre 1 dans les deux jets". Donc $P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2) = \frac{1}{36}$ qui est différent de $P(A_1).P(A_2).P(A_3) = \frac{1}{216}$ et par conséquent les évènements A_1, A_2 et A_3 sont indépendants deux à deux mais ils ne sont pas indépendants dans leur ensemble.

Proposition 1.6.6. — Si $(A_j)_{j \in I}$ est une famille d'évènements indépendants et J une partie finie de I alors pour tout $i \notin J$ on a : $\mathbb{P} \left(A_i / \bigcap_{j \in J} A_j \right) = \mathbb{P}(A_i)$. Autrement dit l'information sur la réalisation des évènements $(A_j)_{j \in J}$ n'a aucune influence sur la probabilité de réalisation de l'évènement A_i pour $i \notin J$. Cette égalité constitue un deuxième critère d'indépendance et donne la signification de celle-ci.

— Dans la suite $(A_j)_{j \in J}$, $J \subset I$ et J finie, d'évènements indépendants on peut remplacer certains évènements A_j par leur complémentaire \bar{A}_j sans perdre la propriété d'indépendance.

Preuve. La démonstration se déduit sans difficulté de la définition d'indépendance des évènements. ■

TD Notions de base pour le calcul des probabilités

Exercice 1. Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et A_1, A_2, \dots, A_n n évènements de \mathcal{F} . Montrer que :

$$(1) . \mathbb{P} \left(\bigcup_{i=1}^n A_i \right) \leq \sum_{i=1}^n \mathbb{P}(A_i) \qquad (2) . \mathbb{P} \left(\bigcap_{i=1}^n A_i \right) \geq 1 - \sum_{i=1}^n \mathbb{P}(\bar{A}_i)$$

Exercice 2. On considère une classe de n élèves. On suppose qu'il n'y a pas d'année bissextile.

(1). Quelle est la probabilité, p_n , pour que deux élèves au moins aient la même date d'anniversaire. Calculer p_{366} .

(2). Quelle est la probabilité, q_n , pour qu'au moins un élève ait la même date d'anniversaire que Ali ?

Exercice 3. On jette trois dés non pipés. Calculer :

(a). La probabilité d'obtenir au moins un as ; (b). La probabilité d'obtenir au moins deux faces portant le même chiffre ; (c) La probabilité que la somme des points marqués sur les trois faces soit paire ; (d). La probabilité que la somme des points soit paire et que deux faces portent le même numéro.

Exercice 4. On cherche un parapluie qui, avec la probabilité $\frac{p}{7}$, se trouve dans l'un quelconque des 7 étages d'un immeuble. On a cherché dans les 6 premiers étages mais on ne l'a pas trouvé. Quelle est la probabilité que le parapluie se trouve au 7^{ème} étage ?.

Exercice 5. On considère trois urnes numérotées de 1 à 3. L'urne n° 1 contient 1 boule blanche et 2 noires, l'urne n° 2 contient 2 boules blanches et 1 noire et L'urne n° 3 contient 3 boules blanches. On choisit une urne au hasard puis on tire une boule dans cette urne. 1. Quelle est la probabilité pour que cette boule soit blanche ? 2. Sachant que la boule tirée est blanche calculer la probabilité qu'elle est tirée de l'urne n° 1 ?.

Exercice 6. Trois machines A , B et C produisent respectivement 50%, 30% et 20% du nombre total de pièces fabriquées par une usine. Le pourcentage de pièces défectueuses pour chaque machine est respectivement 3%, 4% et 5%.

1. Quelle est la probabilité qu'une pièce prise au hasard soit défectueuse ? 2. Quelle est la probabilité qu'une pièce prise au hasard soit bonne ? 3. Si on prend une pièce qui s'avère être défectueuse, quelle est la probabilité qu'elle ait été produite par A , B , C ?

Chapitre 2

Variables Aléatoires

Ce chapitre est consacré à la présentation des variables aléatoires à une dimension. Après un exemple illustrant cette notion, on envisage la définition axiomatique : une variable aléatoire est une application qui associe à chaque élément ω d'un ensemble fondamental un nombre réel. On étudie ensuite les variables aléatoires **discrètes** et **continues**. La première chose à préciser est que dans toute la suite nous utiliserons l'abréviation « v.a. » pour « variable aléatoire » au singulier comme au pluriel.

2.1 Exemple introductif

Considérons l'expérience aléatoire consistant à lancer un dé bien équilibré deux fois de suite. L'ensemble fondamental est $\Omega = \{(i, j), i, j = 1, 2, \dots, 6\}$, les 36 couples sont équiprobables. Associons à chaque couple la somme des deux chiffres c.à.d $(i, j) \longrightarrow i + j$. Nous définissons ainsi une application de l'ensemble fondamental Ω dans l'ensemble des nombres $\{2, 3, 4, \dots, 12\}$. Dans ces conditions, on dit que l'application X : "La somme des deux chiffres obtenus" est une variable aléatoire définie sur l'ensemble fondamental Ω et à valeurs dans l'ensemble $E = \{2, 3, 4, \dots, 12\}$

2.2 Définitions, vocabulaire

Commençons par la définition fondamentale suivante :

Définition 14. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. On appelle **variable aléatoire** définie sur Ω et à valeurs dans (E, \mathcal{B}) toute application X de Ω dans E vérifiant :

$$\forall B \in \mathcal{B}, X^{-1}(B) = \{\omega \in \Omega, X(\omega) \in B\} \in \mathcal{F}$$

Une telle application est dite mesurable. En pratique on ne vérifie que très rarement la mesurabilité. Dans le cas où $E = \mathbb{R}$ on obtient la définition suivante :

Définition 15. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. On appelle **variable aléatoire réelle** définie sur Ω toute application X de Ω dans \mathbb{R} vérifiant :

$$\forall x \in \mathbb{R}, X^{-1}([-\infty, x]) = \{\omega \in \Omega, X(\omega) \leq x\} \in \mathcal{F}$$

* D'après la structure de la tribu borélienne $\mathbb{B}(\mathbb{R})$ (la plus petite tribu qui contient tous les intervalles) la définition a pour conséquence qu'une variable aléatoire X vérifie toujours : $\forall B \in \mathbb{B}(\mathbb{R}), X^{-1}(B) = \{\omega \in \Omega, X(\omega) \in B\} \in \mathcal{F}$.

* On utilise généralement la notation suivante : si $B \in \mathbb{B}(\mathbb{R})$ alors l'ensemble $X^{-1}(B)$ est noté en abrégé $\{X \in B\}$ soit

$$X^{-1}(B) = \{\omega \in \Omega, X(\omega) \in B\} = \{X \in B\}$$

De même l'ensemble $X^{-1}([-\infty, x])$ est noté $= \{X \leq x\}$ soit

$$X^{-1}([-\infty, x]) = \{\omega \in \Omega, X(\omega) \leq x\} = \{X \leq x\}$$

2.2.1 Loi d'une variable aléatoire

Dans l'exemple introductif, si on cherche à évaluer la probabilité attachée à chacun des nombres de l'ensemble E alors il faut évaluer la probabilité attachée à chacun des événements $X^{-1}(\{i\})$, $i = 2, 3, \dots, 12$ qui devrait s'énoncer ainsi : probabilité que

le résultat de la v.a X soit i . Ainsi, à chacun des résultats possibles de l'application X nous associons une probabilité, toutes ces probabilités ensemble forment la loi de probabilité de la v.a X . Introduisons à présent la loi de probabilité d'une variable aléatoire de façon générale.

Théorème 2.2.1. *Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et X une v.a définie sur Ω et à valeurs dans (E, \mathcal{B}) où \mathcal{B} est une tribu sur E . L'application P_X définie par :*

$$\begin{aligned} P_X &: \mathcal{B} \longrightarrow [0, 1] \\ &: B \mapsto P_X(B) = \mathbb{P}(X^{-1}(B)) \end{aligned}$$

est une probabilité sur (E, \mathcal{B}) .

Preuve. * On a tout d'abord $X^{-1}(E) = \Omega$ et par suite : $P_X(E) = \mathbb{P}(X^{-1}(E)) = P_X(\Omega) = \mathbb{P}(\Omega) = 1$.

* Si $(B_n)_{n \in \mathbb{N}}$ une suite d'éléments de \mathcal{B} deux à deux disjoints, alors $(X^{-1}(B_n))_{n \in \mathbb{N}}$ est une suite d'éléments de \mathcal{F} deux à deux disjoints. Comme $X^{-1}\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \bigcup_{n \in \mathbb{N}} X^{-1}(B_n)$, on a, \mathbb{P} étant une probabilité, $P_X\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \mathbb{P}\left(X^{-1}\left(\bigcup_{n \in \mathbb{N}} B_n\right)\right) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} X^{-1}(B_n)\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(X^{-1}(B_n)) = \sum_{n \in \mathbb{N}} P_X(B_n)$.
Ceci montre que P_X est une de probabilité sur (E, \mathcal{B}) . ■

Définition 16. *La probabilité P_X donnée par le théorème ci-dessus s'appelle la loi de probabilité de la v.a X .*

Maintenant, la question qui se pose est de savoir comment décrire la loi d'une variable aléatoire. Mais cela dépend directement de l'espace (E, \mathcal{B}) et dans ce cas on a deux espaces importants : le cas où $E = \mathbb{R}$ dans ce cas la v.a X est dite **continue**, et le cas où E est au plus dénombrable (fini ou dénombrable) et dans ce cas on a la définition suivante :

Définition 17. Une variable aléatoire réelle X est dite **discrète** si l'ensemble $X(\Omega)$ est au plus dénombrable (fini ou dénombrable) : il existe une suite de réels $(x_i)_{i \in \mathbb{N}}$ distincts tels que $X(\Omega) = \{x_i, i \in \mathbb{N}\}$.

Loi de probabilité d'une v.a continue

Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et X une v.a définie sur Ω et à valeurs dans $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ et soit P_X sa loi de probabilité. Comme nous avons vu que la détermination de la loi P_X ce n'est rien d'autre que une détermination d'une loi de probabilité sur $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ et d'après ce qu'on a vu dans le précédent chapitre que chaque probabilité sur $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ est caractérisée par sa fonction de répartition F . Ceci justifie la définition suivante :

Définition 18. Soit X une variable aléatoire réelle. On appelle fonction de répartition de X , notée F (ou F_X si besoin), la fonction suivante :

$$\begin{aligned} F &: \mathbb{R} \longrightarrow [0, 1] \\ &: x \longmapsto F(x) = P_X(\text{]}-\infty, x]) = \mathbb{P}(X \leq x) \end{aligned}$$

et dans ce cas la loi de probabilité P_X de la v.a X est entièrement déterminée par sa fonction de répartition F .

La terminologie « fonction de répartition de X » est un peu impropre car en fait F dépend de la probabilité P_X et pas directement de X . Un des intérêts de la fonction de répartition est qu'elle permet d'exprimer $\mathbb{P}(X \in I)$ pour tout intervalle réel $I \subseteq \mathbb{R}$. C'est l'objet du résultat qui suit :

Proposition 2.2.1. Soit X une variable aléatoire réelle et F sa fonction de répartition. Alors :

1. F est croissante
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$;
3. F est continue à droite en tout point x , c'est-à-dire $\lim_{h \searrow 0} F(x+h) = F(x)$
4. si a et b sont deux réels tels que $a < b$, on a (notant $F(x-0) = \lim_{h \nearrow 0} F(x-h)$)

- (a) . $\mathbb{P}(X = b) = P_X(\{b\}) = F(b) - F(b - 0)$
- (b) . $\mathbb{P}(a < X \leq b) = P_X(]a, b]) = F(b) - F(a)$
- (c) . $\mathbb{P}(a \leq X \leq b) = P_X([a, b]) = F(b) - F(a - 0)$
- (d) . $\mathbb{P}(a < X < b) = P_X(]a, b[) = F(b - 0) - F(a)$
- (e) . $\mathbb{P}(a \leq X < b) = P_X([a, b[) = F(b - 0) - F(a - 0)$
- (f) $\mathbb{P}(X < b) = P_X(]-\infty, b[) = F(b - 0)$
- (g) $\mathbb{P}(X \geq b) = P_X([b, +\infty[) = 1 - F(b - 0)$
- (h) $\mathbb{P}(X > b) = P_X(]b, +\infty[) = 1 - F(b)$

Une conséquence de ce résultat est que, F est continue en un point x si et seulement si $P_X(\{x\}) = 0$. Notons également que comme F est croissante, ses points de discontinuité forment un ensemble au plus dénombrable. Aussi si F est continue on obtient (b) = (c) = (d) = (e) = $F(b) - F(a)$, et (g) = (h) = $1 - F(b)$.

Exemple 2.2.1. Soit X une variable aléatoire réelle de fonction de répartition F donnée par, $[x]$ désignant la partie entière de x ,

$$F(x) = \begin{cases} \frac{1}{5}e^{x-1}, & \text{si } x < 1 \\ \frac{1}{4}[x], & \text{si } 1 \leq x < 3 \\ 1 - \frac{1}{2x}, & \text{sinon} \end{cases}$$

On obtient facilement les probabilités suivantes, $P(X = 0) = 0$, $P(X = 2) = F(2) - F(2 - 0) = 1/2 - 1/4 = 1/4$,

de même que, notant que $F(3) = 5/6$, $F(1) = 1/4$ et $F(1 - 0) = 1/5$: $P(1 < X \leq 3) = F(3) - F(1) = 7/12$, $P(1 \leq X \leq 3) = F(3) - F(1 - 0) = 19/30$

Nous en venons à un résultat très profond. Il montre en fait que la loi d'une v.a.r. est entièrement caractérisée par sa fonction de répartition.

Théorème 2. La loi P_X d'une variable aléatoire réelle X est entièrement déterminée par sa fonction de répartition F .

Plus précisément, si $F : \mathbb{R} \longrightarrow [0, 1]$ est une fonction croissante, continue à droite vérifiant $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$, il existe une unique probabilité P sur

$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que :

$$\forall x \in \mathbb{R}, F(x) = P([-\infty, x])$$

Nous allons à présent nous intéresser à deux types particuliers de v.a.r. : les variables aléatoires **discrètes** d'une part et d'autre part les variables aléatoires **absolument continues**.

2.3 Variables aléatoires discrètes

Comme nous avons vu dans la définition une variable aléatoire réelle X est dite **discrète** si l'ensemble $X(\Omega)$ est au plus dénombrable c'est à dire il existe une suite de réels $(x_i)_{i \in \mathbb{N}}$ distincts tels que $X(\Omega) = \{x_i, i \in \mathbb{N}\}$. Si on note pour $i \in \mathbb{N}$, $p_i = \mathbb{P}(X = x_i)$ on doit avoir $\sum_{i \in \mathbb{N}} p_i = 1$. Certains des p_i peuvent être nuls.

Exemple 2.3.1. Imaginons qu'on lance un dé plusieurs fois de suite :

*Si X représente le résultat du premier lancer, alors : $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$ et $p_1 = p_2 = \dots = p_6 = \frac{1}{6}$.

Si X représente à présent le premier lancer pour lequel on obtient 1, alors : $X(\Omega) = \mathbb{N}^$ et pour tout $k \in \mathbb{N}^*$: $p_k = \mathbb{P}(X = k) = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}$.

Nous allons voir que la loi d'une v.a.r. discrète est relativement simple à caractériser. Tout d'abord, pour x réel déterminons l'ensemble $\{X \leq x\}$. Comme $X(\Omega) = \{x_i, i \in \mathbb{N}\}$ on a $\{X \leq x\} = \bigcup_{i \in \mathbb{N}} \{X \leq x\} \cap \{X = x_i\} = \bigcup_{i \in \mathbb{N}, x_i \leq x} \{X = x_i\}$

On déduit de cette égalité ensembliste deux choses : tout d'abord, X est mesurable si et seulement si les ensembles $\{X = x_i\}$ appartiennent à \mathcal{F} pour tout $i \in \mathbb{N}$; et d'autre part, en utilisant la σ -additivité de la probabilité \mathbb{P} , on obtient :

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(\bigcup_{i \in \mathbb{N}, x_i \leq x} \{X = x_i\}\right) = \sum_{i \in \mathbb{N}, x_i \leq x} \mathbb{P}(\{X = x_i\}) = \sum_{i \in \mathbb{N}, x_i \leq x} p_i$$

On voit alors que la fonction de répartition de X – et par suite sa loi – est entièrement déterminée par les réels $p_i = \mathbb{P}(\{X = x_i\})$. On obtient donc le résultat suivant :

Théorème 3. Soit X une v.a.r. discrète prenant les valeurs $(x_i)_{i \in \mathbb{N}}$. La loi de X est complètement déterminée par les valeurs

$$p_i = \mathbb{P}(\{X = x_i\}) = P_X(\{x_i\})$$

Plus précisément, on a pour tout borélien B ,

$$\mathbb{P}(\{X \in B\}) = \sum_{i \in \mathbb{N}, x_i \in B} p_i$$

Pour une variable discrète, quelle allure a le graphe de la fonction F ? Nous avons déjà dit, avec les notations du théorème précédent, que $F(x) = \sum_{i \in \mathbb{N}, x_i \leq x} p_i$. Si on suppose que les x_i sont classés par ordre croissant, on a pour $x < x_0$, $F(x) = 0$, puis pour $x \in [x_0, x_1[$, $F(x) = p_0$, puis sur $[x_1, x_2[$, $F(x) = p_0 + p_1 \dots$. F est donc une fonction constante par morceaux.

Exemple 2.3.2. Soit X une variable aléatoire prenant les valeurs $-2, 1, 2$ avec probabilité respectives $\mathbb{P}(\{X = -2\}) = \frac{1}{3}$, $\mathbb{P}(\{X = 1\}) = \frac{1}{2}$, $\mathbb{P}(\{X = 2\}) = \frac{1}{6}$. La fonction de répartition de X est donnée par :

$$F(x) = \begin{cases} 0, & \text{si } x < -2 \\ \frac{1}{3}, & \text{si } x \in [-2, 1[\\ \frac{5}{6}, & \text{si } x \in [1, 2[\\ 1, & \text{sinon} \end{cases}$$

2.4 Variables aléatoires absolument continues

Dans ce paragraphe, on considère le cas des variables aléatoires réelles absolument continues ; commençons par une définition :

Définition 19. Soit la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$. f est une densité de probabilité si :

1. f est positive. c.à.d $\forall x \in \mathbb{R} : f(x) \geq 0$.
2. $\int_{\mathbb{R}} f(x) dx = 1$.

Exemple 2.4.1. La fonction $f(x) = \frac{1}{4}e^{-\frac{|x|}{2}}$ est une densité de probabilité .

Définition 20. Soit X une variable aléatoire réelle de fonction de répartition F . On dit que X est absolument continue s'il existe une densité de probabilité f telle que :

$$\forall x \in \mathbb{R}, F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t) dt$$

Dans ce cas on dit que X a pour densité f ou encore f est la densité de probabilité de la variable aléatoire X .

Si X est absolument continue de densité f , on a pour tout intervalle, toute réunion d'intervalles (et même tout borélien) B ,

$$\mathbb{P}(X \in B) = \int_B f(t) dt$$

La question qui se pose à présent est la suivante : comment reconnaître qu'une v.a.r. X possède une densité ? Voici un résultat qui précise les liens entre fonction de répartition et densité de probabilité.

Théorème 4. Soit X une variable aléatoire de fonction de répartition F .

1. Si X possède une densité, alors F est continue et donc $\mathbb{P}(X = x) = 0$ pour tout $x \in \mathbb{R}$.

2. Si F est dérivable (de classe C^1) alors X possède une densité f définie par $f(x) = F'(x)$.

2.5 Les caractéristiques d'une variable aléatoire

Une variable aléatoire est totalement déterminée soit par sa fonction de répartition, soit par sa densité ou encore par sa fonction de probabilité. En statistique, on se contente généralement de certaines valeurs caractéristiques qui ne décrivent pas entièrement la variable aléatoire mais qui sont en général très importantes. Les valeurs caractéristiques les plus importantes sont l'espérance mathématique et la variance.

2.5.1 L'espérance mathématique

Définition 21. L'espérance mathématique, notée $E(X)$, d'une variable aléatoire X , appelée également moyenne, si elle existe.

1. Dans le cas d'une variable aléatoire discrète X , dont les valeurs possibles sont $x_i, i \in \mathbb{N}$, l'espérance de X est définie par l'expression

$$E(X) = \sum_{i \in \mathbb{N}} x_i \mathbb{P}(X = x_i) = \sum_{i \in \mathbb{N}} x_i p_i, \text{ si elle existe.}$$

En termes concrets, l'espérance de X est la moyenne pondérée des valeurs que X peut prendre, les poids étant les probabilités que ces valeurs soient prises.

2. Dans le cas d'une variable aléatoire absolument continue X , dont la densité de probabilité est f , l'espérance de X est définie par :

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx, \text{ si elle existe.}$$

Exemple 2.5.1. On lance un dé équilibré une seule fois et soit X la v.a qui représente le chiffre obtenu. Alors X prend les valeurs 1, 2, 3, 4, 5, 6 avec la même probabilité $p_i = \frac{1}{6}$, pour $i = 1, 2, 3, 4, 5, 6$. Donc $E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3,5$.

Exemple 2.5.2. Soit X la v.a.r, dont la densité de probabilité est f définie par $f(x) = \begin{cases} e^{-x}, & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases}$. Alors :

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^0 0 dx + \int_0^{+\infty} x e^{-x} dx = 1$$

Remarque 2.5.1. – L'espérance n'est pas toujours définie. Considérons le cas où X est une variable aléatoire pouvant prendre une infinité de valeurs distinctes, $E(X) = \sum_{i=1}^{+\infty} x_i p_i$. Lorsque la somme a un sens et ne dépend pas de l'ordre des termes, et dans ce cas seulement on considère que X a une espérance mathématique.

– L'espérance d'une variable aléatoire n'appartient pas obligatoirement aux valeurs possibles de X . Ainsi l'espérance mathématique des valeurs prises par un dé idéal est 3,5 alors que le dé ne peut faire apparaître que des valeurs entières. $E(X)$ est un nombre autour duquel se répartissent les valeurs possibles de la v. a X .

Définition 22. Une variable aléatoire X est dite centrée si $E(X) = 0$.

2.5.2 Espérance d'une fonction d'une v.a.r

D'une façon générale on peut définir l'espérance d'une fonction d'une v.a.r X , alors on a la définition suivante :

Définition 23. Soient X une v.a.r définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ et $\Phi : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction. On définit la v.a $Y = \Phi(X)$ par :

$$\begin{aligned} Y &: \Omega \longrightarrow \mathbb{R} \\ &: \omega \mapsto \Phi(X(\omega)) \end{aligned}$$

Et si Φ est intégrable on a :

1. Dans le cas d'une variable aléatoire discrète X , dont les valeurs possibles sont $x_i, i \in \mathbb{N}$, l'espérance de Y est définie par l'expression

$$E(Y) = E(\Phi(X)) = \sum_{i \in \mathbb{N}} \Phi(x_i) \mathbb{P}(X = x_i) = \sum_{i \in \mathbb{N}} \Phi(x_i) p_i, \text{ si elle existe.}$$

2. Dans le cas d'une variable aléatoire absolument continue X , dont la densité de probabilité est f , l'espérance de Y est définie par :

$$E(Y) = E(\Phi(X)) = \int_{-\infty}^{+\infty} \Phi(x) f(x) dx, \text{ si elle existe.}$$

2.5.3 Propriétés de l'espérance mathématique

1. L'espérance mathématique d'une constante est la constante elle même. c'est à dire Si : $X = c, (c \in \mathbb{R})$ alors $E(X) = c$

2. Si X est une variable aléatoire et $\alpha \in \mathbb{R}$, alors : $E(\alpha X) = \alpha E(X)$.
3. Si X et Y sont deux variables aléatoires, alors : $E(X + Y) = E(X) + E(Y)$
4. En utilisant les points 2. et 3. on obtient : Si X et Y sont deux variables aléatoires, et $\alpha, \beta \in \mathbb{R}$, alors $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$. C'est à dire l'espérance mathématique est un opérateur linéaire.
5. D'une façon plus générale si $(X_i)_{i=1,2,\dots,n}$ est une suite de variables aléatoires et $(\alpha_i)_{i=1,2,\dots,n}$ une suite de nombres réels, alors $E\left(\sum_{i=1}^n \alpha_i X_i\right) = \sum_{i=1}^n \alpha_i E(X_i)$.
6. Si X est une v.a d'espérance $E(X)$ alors $Y = X - E(X)$ est une v.a centrée ($E(Y) = 0$) c'est à dire toute v.a on peut la centrer.
7. Si X est une v.a positive c'est à dire $P(X \geq 0) = 1$ alors $E(X) \geq 0$.

2.5.4 Moments d'une variable aléatoire

Définition 24. On appelle moment d'ordre r de la variable aléatoire X l'espérance $E(X^r)$ et il est défini par :

1. Dans le cas d'une variable aléatoire discrète X , dont les valeurs possibles sont $x_i, i \in \mathbb{N}$:

$$E(X^r) = \sum_{i \in \mathbb{N}} x_i^r \mathbb{P}(X = x_i) = \sum_{i \in \mathbb{N}} x_i^r p_i, \text{ si elle existe.}$$

2. Dans le cas d'une variable aléatoire absolument continue X , dont la densité de probabilité est f :

$$E(X^r) = \int_{-\infty}^{+\infty} x^r f(x) dx, \text{ si elle existe.}$$

On remarque que :

- * L'espérance $E(X)$ d'une v a X ce n'est rien d'autre que son moment d'ordre 1.
- * Pour $r = 0$ on obtient $E(X^0) = E(1) = 1$.

Définition 25. On appelle moment d'ordre r centré autour de la valeur $a \in \mathbb{R}$ de la variable aléatoire X l'espérance $E((X - a)^r)$ et il est défini par :

1. Dans le cas d'une variable aléatoire discrète X , dont les valeurs possibles sont $x_i, i \in \mathbb{N}$:

$$E((X - a)^r) = \sum_{i \in \mathbb{N}} (x_i - a)^r \mathbb{P}(X = x_i) = \sum_{i \in \mathbb{N}} (x_i - a)^r p_i, \text{ si elle existe.}$$

2. Dans le cas d'une variable aléatoire absolument continue X , dont la densité de probabilité est f :

$$E((X - a)^r) = \int_{-\infty}^{+\infty} (x - a)^r f(x) dx, \text{ si elle existe.}$$

2.5.5 Variance et Ecart-type d'une variable aléatoire

Si X est une variable aléatoire réelle, la première information que l'on cherche est la valeur moyenne, $E(X)$. Ensuite, on s'intéresse à la dispersion de X autour de cette valeur moyenne : c'est la notion de variance.

Définition 26. Soit X une variable aléatoire réelle telle que $E(X^2)$ existe. Alors on appelle variance de X , notée $V(X)$ ou $\text{Var}(X)$ le moment de X d'ordre 2 centré autour de $E(X)$. C'est à dire

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

Définition 27. On appelle écart-type de la v.a X , noté σ ou σ_X la racine carrée (positive) de sa variance. C'est à dire :

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

Exemple 2.5.3. On lance un dé équilibré une seule fois et soit X la v.a qui représente le chiffre obtenu. Alors X prend les valeurs 1, 2, 3, 4, 5, 6 avec la même probabilité $p_i = \frac{1}{6}$, pour $i = 1, 2, 3, 4, 5, 6$. Donc on a $E(X) = 3,5$. et pour calculer la variance on doit d'abord calculer $E(X^2)$ et on a :

$$E(X^2) = \sum_{i=1}^6 x_i^2 p_i = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + \dots + 6^2 \cdot \frac{1}{6} = \frac{1}{6} (1 + 4 + 9 + 16 + 25 + 36) = \frac{91}{6} = 15,17$$

$$V(X) = E(X^2) - (E(X))^2 = 15,17 - (3,5)^2 = 15,17 - 12,25 = 2,92 \text{ et } \sigma_X = \sqrt{V(X)} = \sqrt{2,92} = 1,7$$

2.5.6 Propriétés de la variance d'une variable aléatoire

1. La variance et l'écart-type sont des valeurs positives c'est à dire : $V(X) \geq 0$ et $\sigma_X \geq 0$.
2. La variance est nulle si et seulement si la v.a est constante. c'est à dire $X = c$, $(c \in \mathbb{R}) \Leftrightarrow V(X) = 0$
3. Si X est une variable aléatoire et $\alpha, \beta \in \mathbb{R}$, alors : $V(\alpha X + \beta) = \alpha^2 V(X)$., et $\sigma_{\alpha X + \beta} = |\alpha| \sigma_X$.

Définition 28. Soit X une variable aléatoire réelle, X est dite **réduite** si $V(X) = 1$.

Définition 29. Soit X une variable aléatoire réelle, d'espérance $E(X)$ et d'écart-type σ_X . On appelle variable aléatoire **centrée réduite** associée à X , notée X^* la v.a.r définie par :

$$X^* = \frac{X - E(X)}{\sigma_X}$$

on peut vérifier sans difficulté que $E(X^*) = 0$ et $V(X^*) = 1$.

2.6 Lois Usuelles

Dans cette section nous passons en revue certaines des lois connues, discrètes ou continues, nous explicitons certaines de leurs caractéristiques telles que la moyenne et la variance.

2.6.1 Lois discrètes usuelles

2.6.2 Loi uniforme :

Une variable aléatoire X suit une loi uniforme discrète si elle prend ses valeurs dans l'ensemble fini $E = \{x_1, x_2, \dots, x_n\}$ avec la même probabilité c'est à dire : $p_i = P(X = x_i) = \frac{1}{\text{card}E} = \frac{1}{n}$. Dans ce cas, on dit que les événements $\{X = x_i\}$ sont équiprobables. On dit dans ce cas que X suit une loi uniforme sur E et on note par :

$$\begin{aligned}
 X &\sim \mathcal{U}_{\{x_1, x_2, \dots, x_n\}} \\
 * E(X) &= \sum_{i=1}^n x_i P(X = x_i) = \sum_{i=1}^n x_i p_i = \frac{1}{n} \sum_{i=1}^n x_i \\
 * E(X^2) &= \sum_{i=1}^n x_i^2 P(X = x_i) = \sum_{i=1}^n x_i^2 p_i = \frac{1}{n} \sum_{i=1}^n x_i^2 \\
 * V(X) &= E(X^2) - (E(X))^2
 \end{aligned}$$

Cas particulier

Si $x_i = i$ c à d $E = \{1, 2, \dots, n\}$ alors : $E(X) = \frac{1}{n} \sum_{i=1}^n i = \frac{n+1}{2}$, $E(X^2) = \frac{1}{n} \sum_{i=1}^n i^2 = \frac{(n+1)(2n+1)}{6}$, $V(X) = E(X^2) - (E(X))^2 = \frac{n^2-1}{12}$

2.6.3 Loi de Bernoulli $\mathcal{B}(p)$

Définition 30. Une épreuve (expérience) de Bernoulli est une épreuve à deux issues, souvent appelées : **succès** et **échec**.

On considère une expérience aléatoire et un événement A lié à cette expérience tel que $P(A) = p$. On effectue une fois cette expérience et on désigne par X l'application qui prend la valeur 1 c'est A se réalise et la valeur 0 sinon (X est le nombre de

réalisation de A) c'est à dire nous avons ici une épreuve de Bernoulli tel que le **succès** est la réalisation de A et l'**échec** est la réalisation de \bar{A} .

X est donc la variable aléatoire définie par :

$$X = \begin{cases} 1, & \text{si } A \text{ est réalisé} \\ 0, & \text{si } A \text{ n'est pas réalisé} \end{cases}$$

la loi de X est donnée par : $P(X = 1) = p$, $P(X = 0) = 1 - p = q$

On dit que X suit une loi Bernoulli de paramètre p et on note par : $X \sim \mathcal{B}(p)$
(on utilise aussi la notation $X \sim \mathcal{B}(1, p)$)

Dans ce cas :

$$* E(X) = \sum_{i=0}^1 iP(X=i) = 0 \times P(X=0) + 1 \times P(X=1) = p$$

$$* E(X^2) = \sum_{i=0}^1 i^2 P(X=i) = 0^2 \times P(X=0) + 1^2 \times P(X=1) = p$$

$$* V(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1-p) = pq$$

2.6.4 Loi Binômiale $\mathcal{B}(n, p)$

On considère une épreuve de Bernoulli qui nous donne A ou \bar{A} . On effectue, dans les mêmes conditions, n fois de suite cette épreuve avec :

* La probabilité que A se réalise est p dans chaque épreuve.

* Les n répétitions de l'épreuve sont indépendantes les unes des autres.

Soit X le nombre d'obtention de l'évènement A au cours de n épreuves. Alors X est la variable aléatoire qui prend les valeurs $0, 1, \dots, n$ sa loi de probabilité est donnée par :

$$\forall k \in \{0, 1, \dots, n\} : P(X = k) = C_n^k p^k (1-p)^{n-k}$$

On dit que X suit la loi Binomiale de paramètres n et p et on note : $X \sim \mathcal{B}(n, p)$

Le premier paramètre de la loi n , est le nombre de répétitions de l'épreuve de Bernoulli; le second paramètre p , est la probabilité que A se réalise dans chaque épreuve.

Montrons que $\forall k \in \{0, 1, \dots, n\} : P(X = k) = C_n^k p^k (1-p)^{n-k}$:

On a la probabilité que A est réalisé k fois exactement dans la suite des répétitions de l'épreuve, et \bar{A} est réalisé $n - k$ fois est, puisque les répétitions sont indépendantes les unes des autres, $p^k (1 - p)^{n-k}$. Et comme il y a autant de n -uplets où A apparaît exactement k fois que de sous-ensembles à k éléments dans un ensemble à n éléments, on la formule.

Remarks 2.6.1. 1. On a : $\sum_{k=0}^n P(X = k) = \sum_{k=0}^n C_n^k p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1$ (binôme de Newton), donc il s'agit vraiment d'une loi de probabilité

2. On remarque que $P(X = k) = C_n^k p^k (1 - p)^{n-k}$ est $(k + 1)^{\text{ième}}$ terme dans la formule de Newton de $(p + (1 - p))^n$ et pour cela on l'appelle la loi Binomiale.

3. il est clair que X peut être regardé comme la somme de n variables indépendantes de Bernoulli X_i de même paramètre p , soit : $X = \sum_{k=1}^n X_i$ où :

$$X_i = \begin{cases} 1, & \text{si } A \text{ est réalisé dans l'épreuve } i \\ 0, & \text{si } A \text{ n'est pas réalisé dans l'épreuve } i \end{cases}$$

Proposition 2.6.1. Soit X une variable aléatoire tel que $X \sim \mathcal{B}(n, p)$. Alors

$$* E(X) = np \quad * V(X) = np(1 - p)$$

Preuve. La démonstration est basée toujours sur la formule de Newton. On a :

$$* E(X) = \sum_{k=0}^n k P(X = k) = \sum_{k=0}^n k C_n^k p^k (1 - p)^{n-k} = np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1 - p)^{n-k} = np (p + (1 - p))^{n-1} = np$$

$$* E(X^2) = \sum_{k=0}^n k^2 P(X = k) = \sum_{k=0}^n k^2 C_n^k p^k (1 - p)^{n-k} = np \sum_{k=1}^n k C_{n-1}^{k-1} p^{k-1} (1 - p)^{n-k} \\ = np \sum_{k=1}^n (k - 1) C_{n-1}^{k-1} p^{k-1} (1 - p)^{n-k} + np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1 - p)^{n-k}$$

$$= n(n - 1) p^2 \sum_{k=2}^n C_{n-2}^{k-2} p^{k-2} (1 - p)^{n-k} + np = n(n - 1) p^2 + np.$$

$$* V(X) = E(X^2) - (E(X))^2 = n(n - 1) p^2 + np - (np)^2 = np(1 - p) \quad \blacksquare$$

Remarque 2.6.1. On peut aussi utiliser que $X = \sum_{k=1}^n X_i$ avec $X_i \sim \mathcal{B}(p)$ pour tout $i = 1, 2, \dots, n$ et donc on a :

$$* E(X) = E\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n E(X_k) = \sum_{k=1}^n p = np$$

* Comme les variables aléatoires $(X_i)_{i=1,2,\dots,n}$ sont indépendantes alors :

$$V(X) = V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k) = \sum_{k=1}^n p(1-p) = np(1-p).$$

2.6.5 Loi Hypergéométrique $\mathcal{H}(N, n, m)$

Soit une population de N individus parmi lesquels m ($m \leq N$) individus possèdent un caractère particulier. Il s'agit par exemple du nombre d'individus qui souffrent d'une certaine maladie, ou le nombre des pièces défectueuse dans un grand lot de fabrication. On prélève un échantillon de n ($n \leq N$) individus parmi cette population (le tirage pouvant s'effectuer d'un seul coup ou au fur et à mesure mais sans remise). On note X la variable aléatoire d'individus de l'échantillon possédant le caractère envisagé, alors X prend les valeurs de l'ensemble $X(\Omega) = \{0, 1, \dots, \min(n, m)\}$. Dans ce cas X suit une loi dite **hypergéométrique** et on a :

$$\forall k \in X(\Omega) : P(X = k) = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n}$$

On note $X \sim \mathcal{H}(N, n, m)$

Si on pose $p = \frac{m}{N}$ (la proportion des individus qui possèdent le caractère envisagé) on peut ici utiliser la notation : $X \sim \mathcal{H}(N, n, p)$

$$\text{Montrons que } \forall k \in X(\Omega) : P(X = k) = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} :$$

On peut voir facilement que : $P(X = k) = \frac{\text{Card}(X^{-1}(\{k\}))}{\text{Card}(\Omega)}$ et on a : $\text{Card}(\Omega) = C_N^n$ (Car on fait des tirages sans remise de n éléments parmi N éléments), et $\text{Card}(X^{-1}(\{k\})) = C_m^k C_{N-m}^{n-k}$ = nombre de cas où l'échantillon de n éléments contient k qui possède le caractère et $n - k$ qui ne le possèdent pas.

Proposition 2.6.2. Soit X une variable aléatoire tel que $X \sim \mathcal{H}(N, n, m)$ et $p = \frac{m}{N}$.

Alors

$$* E(X) = np \qquad * V(X) = np(1-p) \frac{N-n}{N-1}$$

Preuve. On a :

$$\begin{aligned}
 * E(X) &= \sum_{k=0}^n k P(X = k) = \sum_{k=0}^n k \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} = \sum_{k=1}^n k \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} = \frac{nm}{N} \sum_{k=1}^n \frac{C_{m-1}^{k-1} C_{N-m}^{n-k}}{C_{N-1}^{n-1}} = \\
 \frac{nm}{N} &= np \quad (\text{Car } \sum_{k=1}^n \frac{C_{m-1}^{k-1} C_{N-m}^{n-k}}{C_{N-1}^{n-1}} = 1 \text{ c'est la somme des probabilités de la loi } \mathcal{H}(N-1, n-1, m-1))
 \end{aligned}$$

De la même manière on peut calculer $E(X^2)$ et ainsi $V(X)$. ■

Remarque 2.6.2. * On remarque que si on pose $p = \frac{m}{N}$, alors, p est la probabilité de tomber sur un individu présentant le caractère envisagé lorsque l'on tire au hasard un individu dans la population. Avec cette notation, on a : $E(X) = np$. On retrouve, bien que la probabilité de tirer un individu présentant le caractère varie lors de la constitution de l'échantillon de taille n , le résultat de l'espérance d'une variable aléatoire suivant la loi $\mathcal{B}(n, p)$.

* On remarque aussi que la variance de X ne diffère de la variance d'une variable aléatoire qui suit la loi $\mathcal{B}(n, p)$ que d'un facteur $\frac{N-n}{N-1}$ qui s'appelle facteur d'exhaustivité.

* Si dans la même population on fait les tirages des n individus avec remise alors la variable aléatoire X suit la loi $\mathcal{B}(n, p)$ et pour cette raison on dit que la loi **binomiale** est un schéma de tirage avec remise et la loi **hypergéométrique** est le schéma de tirage sans remise.

2.6.6 Loi géométrique $\mathcal{G}(p)$

C'est la loi du nombre d'essais(ou d'épreuves) nécessaires pour faire apparaître pour la première fois un événement de probabilité p . Alors, on considère une expérience aléatoire et soit A un évènement lié à cette expérience avec $P(A) = p$. On effectue, dans les mêmes conditions, cette expérience plusieurs fois jusqu'à l'obtention de A pour la première fois. Soit donc X le nombre de fois nécessaires pour obtenir A pour la première fois. Alors X est la variable aléatoire qui prend ses valeurs dans l'ensemble \mathbb{N}^* (c'est à dire ici $X(\Omega) = \mathbb{N}^*$), avec :

$$\forall k \in \mathbb{N}^* : P(X = k) = (1 - p)^{k-1} p = q^{k-1} p$$

et on dit que X suit une loi géométrique de paramètre p et on note : $X \sim \mathcal{G}(p)$

Soient les évènements A_i : " A est réalisé dans l'expérience i " pour $i \geq 1$ alors on a $(A_i)_{i \in \mathbb{N}^*}$ est une suite d'évènements indépendants avec $P(A_i) = p$ pour tout $i \in \mathbb{N}^*$. On peut écrire : $\forall k \in \mathbb{N}^* : P(X = k) = P(\overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_{k-1}} \cap A_k) = P(\overline{A_1}) P(\overline{A_2}) \dots P(\overline{A_{k-1}}) P(A_k) = (1-p)^{k-1} p = q^{k-1} p$

Proposition 2.6.3. Soit X une variable aléatoire tel que $X \sim \mathcal{G}(p)$. Alors

$$* E(X) = \frac{1}{p} \quad * V(X) = \frac{q}{p^2}$$

Preuve. On utilise dans la démonstration la somme de la série géométrique et ses dérivés

$$\begin{aligned} * E(X) &= \sum_{k=0}^{\infty} k P(X = k) = \sum_{k=0}^{\infty} k q^{k-1} p = p \sum_{k=1}^{\infty} k q^{k-1} = p \sum_{k=1}^{\infty} \frac{dq^k}{dq} = p \frac{d\left(\sum_{k=0}^{\infty} q^k\right)}{dq} = \\ &= p \frac{d\left(\frac{1}{1-q}\right)}{dq} = p \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} \\ * E(X^2) &= \sum_{k=0}^{\infty} k^2 P(X = k) = \sum_{k=0}^{\infty} k^2 q^{k-1} p = p \sum_{k=1}^{\infty} k^2 q^{k-1} = p \sum_{k=1}^{\infty} (k^2 - k + k) q^{k-1} \\ &= p \sum_{k=1}^{\infty} k(k-1) q^{k-1} + p \sum_{k=1}^{\infty} k q^{k-1} = pq \sum_{k=2}^{\infty} k(k-1) q^{k-2} + E(X) \\ &= pq \sum_{k=2}^{\infty} \frac{d^2 q^k}{dq^2} + \frac{1}{p} = pq \frac{d^2\left(\sum_{k=0}^{\infty} q^k\right)}{dq^2} + \frac{1}{p} \\ &= pq \frac{d^2\left(\frac{1}{1-q}\right)}{dq^2} + \frac{1}{p} = pq \frac{d\left(\frac{1}{(1-q)^2}\right)}{dq} + \frac{1}{p} = pq \frac{2}{(1-q)^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p} \\ * V(X) &= E(X^2) - (E(X))^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q+p-1}{p^2} = \frac{q}{p^2}. \quad \blacksquare \end{aligned}$$

Exemple 2.6.1. On jette une pièce de monnaie équilibrée jusqu'à l'obtention de pile. Soit X le nombre de jets nécessaires. Alors ici $X \sim \mathcal{G}\left(\frac{1}{2}\right)$ c'est à dire ($q = p = \frac{1}{2}$) :

$$\forall k \in \mathbb{N}^* : P(X = k) = \left(\frac{1}{2}\right)^k$$

et : $* E(X) = 2 \quad * V(X) = 2$

2.6.7 Loi de poisson $\mathcal{P}(\lambda)$

Soit $\lambda \in \mathbb{R}$. On dit que la variable aléatoire discrète X , à valeurs dans \mathbb{N} , suit la loi de Poisson de paramètre λ , notée $X \sim \mathcal{P}(\lambda)$, si sa loi est donnée par la relation suivante :

$$\forall k \in \mathbb{N} : P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Rappel : $\sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x \quad \forall x \in \mathbb{R}$.

Alors on a : $\sum_{k=0}^{+\infty} P(X = k) = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^\lambda = 1$.

La variable aléatoire de Poisson est utilisée pour modéliser les phénomènes comme, le nombre d'accident pendant un jour, nombre d'appel téléphoniques pendant un intervalle de temps, nombre de pièces défectueuses dans une livraison importante, nombre de clients qui arrivent à la poste

Proposition 2.6.4. *Soit X une variable aléatoire tel que $X \sim \mathcal{P}(\lambda)$. Alors*

$$* E(X) = \lambda \qquad * V(X) = \lambda$$

Preuve. On a

$$* E(X) = \sum_{k=0}^{\infty} k P(X = k) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda e^\lambda e^{-\lambda} = \lambda.$$

$$* E(X^2) = \sum_{k=0}^{\infty} k^2 P(X = k) = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} (k^2 - k + k) \frac{\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=1}^{\infty} k(k-1) \frac{\lambda^k}{k!} + e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} + E(X)$$

$$= \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda = \lambda^2 e^{-\lambda} e^\lambda + \lambda = \lambda^2 + \lambda$$

$$* V(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \quad \blacksquare$$

2.6.8 Approximation de la loi binômiale $\mathcal{B}(n, p)$ par une loi de Poisson $\mathcal{P}(\lambda)$

Soit X une variable aléatoire telle que $X \sim \mathcal{B}(n, p)$ alors on a le résultat suivant :

Théorème 5. $\lim_{\substack{n \rightarrow +\infty \\ np \rightarrow \lambda}} \mathcal{B}(n, p) = \mathcal{P}(\lambda).$

C'est à dire : quand $n \rightarrow +\infty$ tel que $np \rightarrow \lambda$ alors $X \rightsquigarrow \mathcal{P}(\lambda).$

*Ce théorème veut dire que si n est suffisamment **grand** et p est suffisamment **petit** tel que $np = \lambda$ alors on peut faire l'approximation $P(X = k) = C_n^k p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$. Et donc au lieu d'utiliser la loi binomiale on utilise la loi de poisson pour les calculs de $P(X = k)$. On pratique on applique cette approximation quand $n \geq 30$ et $np \leq 5$.*

Preuve. Comme $X \rightsquigarrow \mathcal{B}(n, p)$ alors $P(X = k) = C_n^k p^k (1-p)^{n-k}$, pour $k = 0, 1, \dots, n$ et donc on doit montrer que $\forall k \in \mathbb{N} : \lim_{\substack{n \rightarrow +\infty \\ np \rightarrow \lambda}} C_n^k p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$

Mais, comme $np \rightarrow \lambda$ on peut écrire : $p \simeq \frac{\lambda}{n}$ (quand $n \rightarrow +\infty$) et donc on a :

$$\begin{aligned} \lim_{\substack{n \rightarrow +\infty \\ np \rightarrow \lambda}} C_n^k p^k (1-p)^{n-k} &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \lim_{n \rightarrow +\infty} \frac{n(n-1)\dots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow +\infty} \frac{n^k \left(1 - \frac{1}{n}\right)\dots\left(1 - \frac{k-1}{n}\right)}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

Car : $\lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n}\right) = \lim_{n \rightarrow +\infty} \left(1 - \frac{2}{n}\right) = \lim_{n \rightarrow +\infty} \left(1 - \frac{k-1}{n}\right) = 1$

et : $\lim_{n \rightarrow +\infty} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \lim_{n \rightarrow +\infty} e^{\ln\left(1 - \frac{\lambda}{n}\right)^{n-k}} = \lim_{n \rightarrow +\infty} e^{(n-k)\ln\left(1 - \frac{\lambda}{n}\right)} = \lim_{n \rightarrow +\infty} e^{(n-k)\left(-\frac{\lambda}{n} + o\left(\frac{\lambda}{n}\right)\right)} = e^{-\lambda}$ ■

Exemple 2.6.2. Si par exemple $X \rightsquigarrow \mathcal{B}(100, 0,03)$ alors on peut utiliser l'approximation par la loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda = 100 \times 0,03 = 3$.

2.6.9 Lois absolument continues usuelles

2.6.10 Loi Uniforme Continue $\mathcal{U}_{[a,b]}$

Soient $a, b \in \mathbb{R}$ vérifiant $a < b$. Soit X une variable aléatoire réelle de fonction de densité f . On dit que X suit la loi uniforme sur $[a, b]$ et on écrit : $X \rightsquigarrow \mathcal{U}_{[a,b]}$ si sa

fonction de densité f est donnée par la formule suivante :

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a, b] \\ 0, & \text{si } x \notin [a, b] \end{cases}$$

Le fait que f soit constante sur $[a, b]$ correspond au fait que selon cette loi on choisit les points dans l'intervalle $[a, b]$ avec les mêmes chances, ce qui explique le nom **uniforme**.

La fonction de répartition de la loi $\mathcal{U}_{[a,b]}$ est donnée par :

$$F(x) = \begin{cases} 0, & \text{si } x \leq a \\ \frac{x-a}{b-a}, & \text{si } x \in [a, b] \\ 1, & \text{si } x \geq b \end{cases}$$

$$\text{En effet : } \forall x \in \mathbb{R} : F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$

$$\text{Si : } x \leq a : F(x) = \int_{-\infty}^x 0 dt = 0$$

$$\text{Si : } a \leq x \leq b : F(x) = \int_{-\infty}^a 0 dt + \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

$$\text{Si : } x \geq b : F(x) = \int_{-\infty}^a 0 dt + \int_a^b \frac{1}{b-a} dt + \int_b^x 0 dt = 1$$

Les caractéristiques de la loi $\mathcal{U}_{[a,b]}$

$$* E(X) = \frac{a+b}{2}, \quad * E(X^2) = \frac{a^2+ab+b^2}{3} \quad * V(X) = \frac{(b-a)^2}{12}$$

En effet :

$$* E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^a x \cdot 0 dx + \int_a^b \frac{x}{b-a} dx + \int_b^{+\infty} x \cdot 0 dx = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{b-a}$$

$$* E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_{-\infty}^a x^2 \cdot 0 dx + \int_a^b \frac{x^2}{b-a} dx + \int_b^{+\infty} x^2 \cdot 0 dx = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

$$* V(X) = E(X^2) - E(X)^2 = \frac{(b-a)^2}{12}$$

2.6.11 Loi Normale (loi de Laplace-Gauss) $\mathcal{N}(m, \sigma^2)$

La démonstration classique du résultat suivant utilise un calcul d'intégrale double. Ce résultat sera ici admis.

Théorème 6. Soient $m \in \mathbb{R}$, et $\sigma \in \mathbb{R}_+^*$. Alors la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par :

$$\forall x \in \mathbb{R} : f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

est une densité de probabilité sur \mathbb{R} .

C'est à dire : $\forall x \in \mathbb{R} : f(x) \geq 0$ (très clair),

$$\text{et } \int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = 1.$$

On remarque d'après le théorème que : $\int_{-\infty}^{+\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \sigma\sqrt{2\pi}$.

Et avec un changement de variable ($y = \frac{x-m}{\sigma}$) dans l'intégral on obtient aussi

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$$

Définition 31. (Loi normale $\mathcal{N}(m, \sigma^2)$) Soient $m \in \mathbb{R}$, et $\sigma \in \mathbb{R}_+^*$. On dit que la variable aléatoire continue X suit la loi normale $\mathcal{N}(m, \sigma^2)$, lorsque la densité de probabilité de X est la fonction f définie sur \mathbb{R} par :

$$\forall x \in \mathbb{R} : f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Proposition 2.6.5. Soit X une variable aléatoire suivant la loi $\mathcal{N}(m, \sigma^2)$, alors :

$$*E(X) = m \qquad *V(X) = \sigma^2.$$

Remarque 2.6.3. On remarque que les paramètres m et σ^2 de la loi normale $\mathcal{N}(m, \sigma^2)$ correspondent respectivement à l'espérance et à la variance de toute variable aléatoire suivant cette loi.

Définition 32. (Loi normale centrée réduite $\mathcal{N}(0, 1)$) On dit que la variable aléatoire continue X suit la loi normale centrée réduite si $m = 0$ et $\sigma = 1$, c'est à

dire $X \sim \mathcal{N}(0, 1)$, sa densité de probabilité f est définie par :

$$\forall x \in \mathbb{R} : f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

alors on a : $*E(X) = 0$ $*V(X) = 1$

Notation : La fonction de répartition d'une variable aléatoire X qui suit la loi normale centrée réduite $\mathcal{N}(0, 1)$ est notée $\Pi : .$ On a donc :

$$\forall t \in \mathbb{R} : \Pi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx$$

nous n'avons pas une formule explicite pour la fonction de répartition de la loi normale mais ses valeurs sont données dans des tableaux numériques.

On remarque que : $* \forall t \in \mathbb{R} : \Pi(t) = 1 - \Pi(-t)$

* Si $X \sim \mathcal{N}(m, \sigma^2)$ alors $X^* = \frac{X-m}{\sigma} \sim \mathcal{N}(0, 1)$, donc on peut toujours utiliser le tableau de la fonction de répartition de la loi normale centrée réduite pour calculer les valeurs de celle de la loi $\mathcal{N}(m, \sigma^2)$ et on a : si F est la fonction de répartition de $\mathcal{N}(m, \sigma^2)$ alors : $\forall t \in \mathbb{R} : F(t) = \Pi\left(\frac{t-m}{\sigma}\right)$.

2.6.12 Approximation de la loi binômiale $\mathcal{B}(n, p)$ et la loi de Poisson $\mathcal{P}(\lambda)$ par la loi normale

Approximation de la loi binomiale $\mathcal{B}(n, p)$ par une loi normale

Soit la variable aléatoire X qui suit une loi binomiale $\mathcal{B}(n, p)$ de si n tend vers l'infini, la variable X tend vers la loi normale $\mathcal{N}(m, \sigma^2)$ avec $m = np$ et $\sigma^2 = np(1-p)$. C'est à dire en pratique, on fait l'approximation de la binomiale par une loi normale quand n est suffisamment grand ($n \geq 50$) et p plus voisin de 0, 5.

Approximation de la loi de Poisson $\mathcal{P}(\lambda)$

Si $\lambda > 18$, la loi de Poisson $\mathcal{P}(\lambda)$ peut être assimilée à une loi normale $\mathcal{N}(m, \sigma^2)$ où $m = \lambda$ et $\sigma^2 = \lambda$.