
PROBABILITÉS-STATISTIQUE
COURS ET EXERCICES

Table des matières

1	Analyse combinatoire	1
1.	Cardinal d'un ensemble fini	1
2.	Principes fondamentaux	1
2.1.	Principe multiplicatif	1
2.2.	Principe additif	2
3.	Permutations	3
4.	Arrangements	3
5.	Combinaisons	5
6.	Répartitions d'objets dans des cases discernables	6
6.1.	Cas d'objets discernables	6
6.2.	Cas d'objets indiscernables	7
2	Notions de probabilités	9
1.	Généralités	9
1.1.	Expérience aléatoire	9
1.2.	Événements	9
1.3.	Langage des événements	10
1.4.	Espace probabilisable	11
1.5.	Système complet d'événements	11
2.	Notion de probabilité	12
3.	Probabilités conditionnelles	14
3.1.	Définition	14
3.2.	Formule des probabilités composées	14
3.3.	Formule des probabilités totales	14
3.4.	Formule de Bayes ou probabilités des causes	14
4.	Indépendance d'événements	15
4.1.	Indépendance de deux événements	15
4.2.	Indépendance mutuelle de n événements	15

3	Variables aléatoires réelles	16
1.	Exemples et définitions	16
2.	Loi d'une variable aléatoire réelle	18
2.1.	Fonction de répartition	18
2.2.	Fonction de masse	19
2.3.	Fonction de densité	19
3.	Variables aléatoires réelles discrètes	20
3.1.	Loi d'une v.a.r. discrète	20
3.2.	Loi d'une fonction d'une v.a.r. discrète	21
4.	Variables aléatoires réelles continues	22
4.1.	Loi d'une v.a.r. continue	22
4.2.	Loi d'une fonction continue d'une v.a.r. continue	23
5.	Moments d'une variable aléatoire réelle	24
6.	Fonction caractéristique - Fonction génératrice	26
6.1.	Fonction caractéristique	26
6.2.	Fonction génératrice	27
4	Lois de probabilités usuelles	28
1.	Lois usuelles discrètes	28
1.1.	Loi de Bernoulli $\mathcal{B}(1, p)$	28
1.2.	Loi Binomial $\mathcal{B}(n, p)$	29
1.3.	Loi Hypergéométrique $\mathcal{H}(N, n, p)$ ou $\mathcal{H}(N, Np, n)$	29
1.4.	Loi géométrique $\mathcal{G}(p)$	30
1.5.	Loi Binomial négative $\mathcal{B}_{\mathcal{N}}(r, p)$	31
1.6.	Loi de Poisson $\mathcal{P}(\lambda)$	31
1.7.	Loi uniforme sur $\{1, \dots, n\}; \mathcal{U}_{\{1, \dots, n\}}$	32
1.8.	$X \rightsquigarrow \mathcal{L}$, avec $X(\Omega) = \mathbb{N}$	32
2.	Lois usuelles continues	33
2.1.	Loi uniforme sur $[a, b], a < b \in \mathbb{R}$	33
2.2.	Loi exponentielle $\xi(\theta), \theta > 0$	33
2.3.	Loi Gamma $\gamma(a, \rho), a > 0, \rho > 0$	34
2.4.	Loi normale $\mathcal{N}(m, \sigma^2)$ (ou loi de Gauss)	35
2.5.	Loi de Cauchy	36
5	Vecteurs aléatoires	37
1.	Introduction	37
2.	Loi conjointe d'un vecteur aléatoire	37
2.1.	Fonction de repartition	37

2.2.	Vecteur aléatoire discret	38
2.3.	Vecteur aléatoire absolument continu	38
3.	Lois marginales	39
4.	Espérance mathématique et matrice de covariance	40
5.	Indépendance et conditionnement	42
6.	Vecteurs aléatoires Gaussiens	45
6.1.	Généralités	45
6.2.	Lois dérivées d'un vecteur Gaussien	46
6	Convergences	48
1.	Convergence en loi	48
2.	Convergence en probabilité	48
3.	Convergence presque sûre	49
4.	Théorèmes limites	49
4.1.	Forme multidimensionnelle	49

Chapitre 1

Analyse combinatoire

L'analyse combinatoire a pour but de dénombrer les différentes dispositions que l'on peut former à partir d'un nombre fini d'éléments. Ce chapitre fournit des méthodes de dénombrement particulièrement utiles en probabilités.

1. Cardinal d'un ensemble fini

Définition 1.1. *Un ensemble E non vide est dit fini s'il existe un entier n et une bijection de $\{1, 2, \dots, n\}$ sur E . Lorsqu'il existe, l'entier n est unique et est noté $\text{Card}(E)$. C'est le cardinal ou le nombre d'éléments de E .*

Définition 1.2. *Un ensemble E est dit dénombrable s'il existe une bijection de \mathbb{N} sur E . Un ensemble E est dit infini non dénombrable s'il n'est ni fini, ni dénombrable.*

Soit E un ensemble fini et A, B deux parties de E .

Proposition 1.1.

1. Si \bar{A} est le complémentaire de A dans E alors $\text{Card}(\bar{A}) = \text{Card}(E) - \text{Card}(A)$.
2. $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$.
3. Si $A \cap B = \emptyset$, alors $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B)$
4. $\text{Card}(A \times B) = \text{Card}(A) \times \text{Card}(B)$

2. Principes fondamentaux

2.1. Principe multiplicatif

Ce principe est à la base des techniques de dénombrement. Il permet de compter le nombre de résultats d'expériences qui peuvent se décomposer en une succession de sous-expériences, i.e. en une suite d'expériences partielles, réalisées l'une après l'autre.

Principe : Supposons qu'une expérience est la succession de r sous-expériences. Supposons de plus que la première sous-expérience peut produire n_1 résultats, et pour chaque résultat de la première sous-expérience, il y a n_2 résultats possibles pour la deuxième sous-expérience ; pour chaque résultat des deux premières sous-expériences, il y en a n_3 pour la troisième, etc. Alors, le nombre total de résultats de l'expérience complète est le produit

$$n_1 \times n_2 \times \dots \times n_r.$$

Interprétation avec les ensembles :

Si $E = E_1 \times E_2 \times \dots \times E_r$ (produit cartésien),
alors $Card(E) = Card(E_1) \times Card(E_2) \times \dots \times Card(E_r)$.

Exemple 1 : On considère trois villes : Abidjan, Agboville et Akoupé et On suppose qu'il y a 5 chemins différents pour aller d'Abidjan à Agboville et 3 chemins pour aller d'Agboville à Akoupé. Combien de chemins différents y a-t-il pour aller d'Abidjan à Akoupé en passant par Agboville ?

Exemple 2 : Supposons qu'une plaque d'immatriculation contenant deux lettres distinctes suivies de trois chiffres dont le premier est différent de zéro. Combien de plaques différentes peut-on imprimer ?

Exemple 3 : Nous désirons former un comité de 3 personnes dont 1 président, 1 secrétaire et 1 trésorier à partir d'une classe de 20 élèves. Combien de comités différents est-il possible de faire ?

2.2. Principe additif

Il permet de compter le nombre de résultats d'expériences qui peuvent se décomposer en des sous-expériences mutuellement exclusives (les expériences qui n'ont aucun résultat commun) .

Principe : Supposons qu'une expérience se décompose en r sous-expériences mutuellement exclusives. Si la i -ème sous-expérience a n_i résultats possibles (avec $i = 1, 2, \dots, r$), alors le nombre total de résultats de l'expérience complète est la somme

$$n_1 + n_2 + \dots + n_r.$$

Interprétation avec les ensembles :

Si $E = E_1 \cup E_2 \cup \dots \cup E_r$ avec $E_i \cap E_j = \emptyset$, pour tout $i \neq j$ (partition),
alors $Card(E) = Card(E_1) + Card(E_2) + \dots + Card(E_r)$.

Lorsqu'on veut dénombrer un ensemble fini E , on peut trouver une partition E_1, E_2, \dots, E_n de cet ensemble, où les cardinaux des ensembles E_i sont plus faciles à déterminer. Il ne reste alors qu'à faire la somme des différents cardinaux obtenus

Exemple 4 : On considère trois villes : Abidjan, Agboville et Adzopé. On suppose qu'il y a 5 chemins différents pour aller d'Abidjan à Agboville et 3 chemins pour aller d'Abidjan à Adzopé. Combien de chemins différents y a-t-il pour aller d'Abidjan à l'une des deux autres villes ?

Exemple 5 : En plus des conditions de l'exemple 1 et 4, on suppose qu'il y a 2 chemins pour aller d'Adzopé à Akoupé et 4 chemins directs d'Abidjan à Akoupé. Combien de chemins différents y a-t-il pour aller d'Abidjan à Akoupé ?

3. Permutations

Une permutation de n éléments est un réarrangement ordonné de ces n éléments.

Notation : La fonction "factorielle" est la fonction de \mathbb{N} vers \mathbb{N} qui à tout $n \in \mathbb{N}$ associe $n! = n(n-1) \times \dots \times 2 \times 1$.

On a : $0! = 1$ et $n! = n(n-1)!$.

- Lorsque tous les n éléments sont distincts, on parle de permutations sans répétitions. Dans ce cas le nombre de permutations de n éléments est $n!$.

Remarque : Une permutation de n éléments distincts est une application bijective d'un ensemble E à n éléments dans lui-même.

- Si parmi les n éléments, certains sont semblables, on parle de permutations avec répétitions.

Le nombre de permutations de n éléments dont n_1 sont semblables, n_2 sont semblables, ..., n_r sont semblables est

$$\frac{n!}{n_1! \times n_2! \times \dots \times n_r!}$$

avec $n_1 + n_2 + \dots + n_r = n$.

Exemples :

- Le nombre d'applications d'un ensemble de n éléments distincts dans un ensemble de r éléments distincts $\{a_1, \dots, a_r\}$ qui envoient exactement n_1 éléments dans a_1 , n_2 éléments dans a_2, \dots , n_r éléments dans a_r est égal à

$$\frac{n!}{n_1! \times n_2! \times \dots \times n_r!}$$

4. Arrangements

- Un arrangement sans répétitions de p éléments parmi n éléments distincts est une permutation de p éléments distincts pris parmi les n ($p \leq n$). Cela revient à prendre p éléments distincts parmi les n en tenant compte de l'ordre dans lequel on les choisit.

Remarque :

(i) Les éléments sont pris sans répétition et en tenant compte de l'ordre dans lequel on les choisit.

(ii) Un arrangement sans répétitions de p éléments parmi n est une application injective de $\{1, 2, \dots, p\}$ dans un ensemble F à n éléments.

On note A_n^p le nombre d'arrangements sans répétitions de p éléments parmi n . On a

$$A_n^p = n(n-1) \times \dots \times (n-p+1) = \frac{n!}{(n-p)!}.$$

Propriétés :

(i) $A_n^p = nA_{n-1}^{p-1}$

(ii) $A_n^p = pA_{n-1}^{p-1} + A_{n-1}^p$

Exemples :

- Si E et F sont des ensembles finis tels que $Card(E) = p$ et $Card(F) = n$, alors le nombre d'applications injectives de E dans F est $A_n^p = n(n-1) \times \dots \times (n-p+1)$ si $p \leq n$ et 0 sinon.

- Le nombre de tirages différents sans remise de p boules dans une urne contenant n boules toutes discernables est égal à $A_n^p = n(n-1) \times \dots \times (n-p+1)$ si $p \leq n$ et 0 sinon.

- Le nombre de façons de placer p objets distincts dans n cases discernables, chaque case pouvant contenir au plus 1 objet est $A_n^p = n(n-1) \times \dots \times (n-p+1)$ si $p \leq n$ et 0 sinon.

• Un arrangement avec répétitions (possibles) de p éléments parmi n est une disposition ordonnée de p éléments avec autant de répétitions que l'on souhaite.

Remarque : Un arrangement avec répétitions de p éléments parmi n est une application de $\{1, 2, \dots, p\}$ dans un ensemble F à n éléments.

Le nombre d'arrangements avec répétitions de p éléments parmi n est n^p .

Exemples :

- Si E et F sont des ensembles finis tels que $Card(E) = p$ et $Card(F) = n$, alors le nombre d'applications de E dans F est n^p .

- Le nombre de tirages différents de p boules dans une urne contenant n boules toutes discernables est égal à n^p lorsque les tirages sont non exhaustifs (i.e. avec remise après chaque tirage).

- Le nombre de façons de placer p objets distincts dans n cases discernables, chaque case pouvant contenir éventuellement plusieurs objets est n^p .

5. Combinaisons

• Une combinaison sans répétitions de p éléments pris dans un ensemble à n éléments distincts ($p \leq n$) est un sous-ensemble à p éléments de cet ensemble.

Remarque : Les éléments sont pris sans répétition et ne sont pas ordonnés.

Exemple : Les combinaison sans répétitions de 2 éléments pris dans l'ensemble $\{1, 2, 3, 4\}$ sont : $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, $\{3, 4\}$.

Le nombre de combinaisons (sans répétitions) de p éléments parmi n est noté C_n^p ou $\binom{n}{p}$ qui est appelé coefficient binomial. On a

$$C_n^p = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}.$$

Exemples :

- Si E est un ensemble fini tels que $Card(E) = n$, alors le nombre de parties de E ayant p éléments est C_n^p .

- Le nombre d'applications d'un ensemble de n éléments dans l'ensemble $\{a_1, a_2\}$ ($a_1 \neq a_2$) qui envoient exactement p éléments dans a_1 et les autres dans a_2 est égal C_n^p .

- Le nombre de façons de placer p objets indiscernables dans n cases discernables, chaque case pouvant contenir au plus 1 objet est C_n^p si $p \leq n$ et 0 sinon.

Propriétés :

(i) $C_n^0 = C_n^n = 1$

(ii) $C_n^p = C_n^{n-p}$

(iii) Formule de récurrence (Triangle de Pascal) $C_n^p = C_{n-1}^{p-1} + C_{n-1}^p$

(iv) Formule de Binôme de Newton $(a + b)^n = \sum_{p=0}^n C_n^p a^p b^{n-p}$

En particulier, $2^n = \sum_{p=0}^n C_n^p$ (Le nombre de parties d'un ensemble à n éléments).

• Une combinaison avec répétitions (possibles) de p éléments pris dans un ensemble à n éléments distincts est une disposition non ordonnée de p éléments avec possibilité de répétitions.

Exemple : Les combinaison avec répétitions (possibles) de 2 éléments pris dans l'ensemble $\{1, 2, 3, 4\}$ sont : $\{1, 1\}$, $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 2\}$, $\{2, 3\}$, $\{2, 4\}$, $\{3, 3\}$, $\{3, 4\}$, $\{4, 4\}$.

Le nombre de combinaisons avec répétitions de p éléments parmi n est noté K_n^p ou $\left(\binom{n}{p}\right)$.

On a

$$K_n^p = \left(\binom{n}{p}\right) = C_{n+p-1}^p.$$

Exemples :

Le nombre de façons de placer p objets indiscernables dans n cases discernables, chaque case pouvant contenir éventuellement plusieurs objets est C_{n+p-1}^p .

6. Répartitions d'objets dans des cases discernables

Dans tout ce qui suit, on dispose de n objets et r cases discernables. On désire dénombrer la répartition des objets dans les cases. On suppose que les cases sont discernables.

6.1. Cas d'objets discernables

On suppose que les n objets sont tous discernables.

- **La répartition des objets se fait sans contrainte**

Sous cette hypothèse, une répartition des objets est identifiable à une application de l'ensemble des n objets vers l'ensemble des r cases.

Alors, le nombre de répartitions possibles est r^n .

- **La répartition des objets se fait sans qu'aucune case ne reste vide**

(i) Si $r > n$, alors le nombre de répartitions des objets est zéro.

(ii) Si $r < n$, alors le nombre de répartitions des objets est $A_n^r \times r^{(n-r)}$

En effet, on considère la répartition des objets comme une expérience qu'on subdivise en deux sous-expériences. La première sous-expérience consiste à mettre dans chaque case un objet et la deuxième à répartir le reste des objets sans contrainte. On a A_n^r façons de mettre dans chaque case un objet et un seul. A chacune de ces répartitions, on a $r^{(n-r)}$ façons de répartir le reste des objets. D'après le principe fondamental, on déduit le résultat énoncé.

- **La répartition des objets se fait de telle sorte que la case i contienne n_i objets.**

La répartition des objets est identifiable à une expérience qu'on peut subdiviser en r sous-expériences. La i ème sous-expérience est le choix des objets à mettre dans la case i . A chaque choix des objets des cases $1, \dots, k-1$, on a C_n^{n-k} choix possibles des objets à mettre dans la case k . D'après le principe fondamental, le nombre de répartitions des objets est

$$C_n^{n_1} \times C_{n-n_1}^{n_2} \times \dots \times C_{n-\sum_{i=1}^{r-1} n_i}^{n_r} .$$

En remarquant que $\sum_{i=1}^r n_i = n$, on tire que :

Le nombre de répartitions de n objets discernables dans r cases discernables de telle sorte que la case i contienne n_i objets est

$$\frac{n!}{n_1! \times n_2! \times \dots \times n_r!} .$$

Cette quantité est noté $\binom{n}{n_1, n_2, \dots, n_r}$ et est appelé coefficient multinomial.

6.2. Cas d'objets indiscernables

Ici, on considère que tous les n objets sont indiscernables

- **La répartition des objets se fait sans qu'aucune case ne reste vide**

Puisque les objets sont indiscernables, ce qui différencie les répartitions possibles est la suite n_1, \dots, n_r où n_i désigne le nombre d'objets dans la case i . On doit avoir $n_i \geq 1$.

Supposons qu'on a n objets indiscernables alignés et qu'on veut les diviser en r groupes non vides. Ces objets peuvent être représentés comme suit :

$$0|0|0|0|\dots|0|0|0$$

où les 0 représentent les n objets indiscernables, les bars | de séparation symbolisant les $n - 1$ espaces entre ces objets. Pour avoir une répartition des objets, il suffit de choisir $r - 1$ des $n - 1$ espaces comme bars de division. Si par exemple, $n = 6, r = 3$ et qu'on choisit les deux bars de séparation comme suit :

$$000|0|00,$$

on obtient la répartition où il y a trois objets dans la première case, un objet dans la deuxième case et deux objets dans la troisième case.

Ainsi,

Le nombre de répartitions de n objets indiscernables dans r cases sans qu'aucune case ne reste vide est C_{n-1}^{r-1} .

Remarque : Notons que le nombre de répartitions de n objets indiscernables dans r cases discernables est identique au nombre de vecteurs (n_1, \dots, n_r) à composantes entières supérieures ou égales à 1 tels que

$$n_1 + \dots + n_r = n.$$

- **La répartition des objets se fait sans aucune contrainte**

Puisque les objets sont indiscernables, ce qui différencie les répartitions possibles est la suite n_1, \dots, n_r où n_i désigne le nombre d'objets dans la case i . Sans contrainte, $n_i \geq 0$.

- Première méthode

Comme précédemment, le nombre de répartitions de n objets indiscernables dans r cases discernables est identique au nombre de vecteurs (n_1, \dots, n_r) à composantes entières supérieures ou égales à 0 tels que

$$n_1 + \dots + n_r = n.$$

Par changement de variable $m_i = n_i + 1, i = 1, \dots, r$, on obtient l'équation suivante :

$$m_1 + \dots + m_r = n + r, m_i \geq 1, i = 1, \dots, r.$$

Le nombre de vecteurs solution de cette équation est $C_{n+r-1}^{r-1} = C_{n+r-1}^n$.

- Deuxième méthode

Désignons par e_i la i ème case, $i = 1, \dots, r$. Ainsi, une répartition de ces n correspond à une combinaison avec répétition de n éléments parmi e_1, e_2, \dots, e_r .

Par exemple $n = 6, r = 5$:

$$(e_1, e_1, e_2, e_4, e_4, e_4)$$

signifie qu'on a la répartition :

- 2 objets sont placés dans la première case,
- 1 objet est placé dans la deuxième case,
- aucun objet dans la troisième case,
- 3 objets dans la quatrième case,
- aucun objet dans la cinquième case.

$$\begin{array}{cccccc} & e_1 & e_2 & e_3 & e_4 & e_5 \\ & | & | & | & | & | \\ 00 & | & 0 & | & | & | \\ & & & \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} \\ & & & & & & \end{array}$$

Par conséquent, il y a autant de façons de répartir les objets dans les cases qu'il y a de combinaisons avec répétition de n éléments choisis parmi r . On a donc $\binom{r}{n} = C_{n+r-1}^n$ répartitions.

Ou bien déplaçons provisoirement les $r - 1$ points intermédiaires et rassemblons les :

$$\begin{array}{cc} \underbrace{||| \dots |||}_{r-1 \text{ bars intermédiaires}} & \underbrace{0000000 \dots 0000000}_n \\ & \text{objets indiscernables} \end{array}$$

Par permutation des $r - 1$ bars (indiscernables) et des n objets indiscernables, on obtient une et une seule répartition des objets. Ainsi

Le nombre de répartitions de n objets indiscernables dans r cases sans contrainte est

$$\frac{(r - 1 + n)!}{(r - 1)! \times n!} = C_{n+r-1}^{r-1} = C_{n+r-1}^n.$$

Chapitre 2

Notions de probabilités

Le but de ce chapitre est d'introduire les bases mathématiques utiles à la modélisation des phénomènes aléatoires. Ainsi, l'objectif visé est de donner un sens mathématique à la notion de "hasard".

1. Généralités

1.1. Expérience aléatoire

◇ Une expérience ou un phénomène est dit aléatoire si on ne peut prévoir d'avance le résultat et qui répétée dans les mêmes conditions peut donner lieu à des résultats différents.

Exemples

- Le lancer d'un dé à six faces.
- Le lancer d'une pièce de monnaie.
- Le tirage d'un numéro parmi n numéros.

◇ L'ensemble des résultats d'une expérience aléatoire est bien déterminé avant sa réalisation, i.e. avant l'épreuve. Cet ensemble est appelé l'espace fondamental ou l'univers de l'expérience. On le note en général Ω .

Exemples

- Pour le lancer d'un dé à six faces numéroté de 1 à 6, $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Pour le lancer d'une pièce de monnaie, $\Omega = \{P, F\}$.

◇ Tout résultat d'une expérience aléatoire s'appelle aussi une éventualité. On note ω le résultat obtenu lors de la réalisation d'une expérience aléatoire.

1.2. Événements

◇ Lorsqu'on effectue une expérience aléatoire, certains faits liés à cette expérience peuvent se produire ou non. On les appelle événements. Ils sont représentés par les sous-ensembles de

l'univers Ω .

Exemple

Pour le lancer d'un dé à six faces numéroté de 1 à 6, "obtenir un nombre pair", "obtenir plus de 4" sont des événements possibles.

Notons A l'événement "obtenir un nombre pair" et B l'événement "obtenir plus de 4".

A est décrit par les éventualités 2,4 et 6. On écrit : $A = \{2, 4, 6\}$.

B est décrit par les éventualités 5 et 6. On écrit : $B = \{5, 6\}$.

◇ Tout événement formé d'une seule éventualité est appelé événement élémentaire.

1.3. Langage des événements

◇ On dit qu'un événement A s'est réalisé si le résultat observé de l'expérience est un élément de A .

◇ Tout événement qui n'est jamais réalisé est dit événement impossible, il correspond à l'ensemble vide \emptyset . Exemple, "obtenir un nombre négatif", dans un lancer de dé est un événement impossible à réaliser.

◇ Tout événement qui est toujours réalisé est dit événement certain, il correspond à l'ensemble Ω . Exemple, dans un lancer de dé "obtenir moins de 7" ou "obtenir plus de 0" sont des phrases équivalentes correspondant toutes deux à un événement toujours réalisé, soit à l'ensemble Ω .

◇ A chaque événement A correspond son événement contraire "non A " noté \bar{A} .

L'événement \bar{A} est réalisé si, et seulement si, A n'est réalisé.

L'événement certain et l'événement impossible sont contraires l'un de l'autre.

On considère deux événements A et B .

◇ On dit que l'événement " A et B ", noté $A \cap B$, s'est réalisé si le résultat observé de l'expérience est un élément de A et de B , i.e. si A et B sont réalisés au cours de la même expérience aléatoire.

◇ Les événements A et B sont dits incompatibles ou disjoints s'ils ne peuvent pas se produire simultanément au cours de la même expérience aléatoire. Autrement dit A et B sont incompatibles si, et seulement si, $A \cap B = \emptyset$.

Deux événements contraires sont incompatibles : $A \cap \bar{A} = \emptyset$.

◇ On dit que l'événement " A ou B ", noté $A \cup B$, s'est réalisé si le résultat observé de l'expérience est un élément de A ou de B , i.e. si l'un au moins des deux événements A ou B est réalisé au cours de la même expérience aléatoire.

Remarque : Pour tout événement A , on a, $\bar{\bar{A}} = C_{\Omega}A$, $A \cup \bar{A} = \Omega$ et $A \cap \bar{A} = \emptyset$.

◊ De même, on définit les événements :

– La différence : $A \setminus B = A \cap \bar{B}$. Remarque : $A \setminus B = A \setminus (A \cap B)$.

L'événement $A \setminus B$ est réalisé si, et seulement si, A est réalisé et B ne l'est pas.

– La différence symétrique : $A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$.

L'événement $A \Delta B$ est réalisé si, et seulement si, l'un et l'un seulement des deux événements A et B est réalisé.

◊ On dit que A implique B si la réalisation de A entraîne la réalisation de B . Autrement dit A implique B si, et seulement si, $A \subset B$.

1.4. Espace probabilisable

Soit Ω l'univers d'une expérience aléatoire. On considère $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω . Un sous-ensemble \mathcal{F} de $\mathcal{P}(\Omega)$ est un ensemble de parties de Ω .

Définition 2.1. *Un sous-ensemble \mathcal{F} de $\mathcal{P}(\Omega)$ est une tribu sur Ω si*

(i) $\Omega \in \mathcal{F}$,

(ii) \mathcal{F} est stable par passage au complémentaire (i.e. $A \in \mathcal{F} \implies \bar{A} \in \mathcal{F}$),

(iii) \mathcal{F} est stable par réunion dénombrable (i.e. $A_i \in \mathcal{F}, i \in \mathbb{N} \implies \bigcup_{i=0}^{\infty} A_i \in \mathcal{F}$).

Exemples

- $\{\emptyset, \Omega\}$ est une tribu appelée tribu grossière. C'est la plus petite tribu sur Ω .

- $\mathcal{P}(\Omega)$ est une tribu appelée tribu triviale. C'est la plus grande tribu sur Ω .

- Pour tout $A \subset \Omega$, $\{\emptyset, A, \bar{A}, \Omega\}$ est une tribu. C'est la tribu engendrée par l'événement A , c-à-d la plus petite tribu sur Ω contenant A . On la note $\sigma(A)$.

Le couple (Ω, \mathcal{F}) formé d'un ensemble Ω et d'une tribu \mathcal{F} est appelé un espace probabilisable.

Définition 2.2. *On appelle tribu borélienne la tribu définie sur \mathbb{R} et qui contient tous les intervalles de \mathbb{R} . On la note $\mathcal{B}_{\mathbb{R}}$.*

1.5. Système complet d'événements

Dans cette partie, I désigne une partie de \mathbb{N} ou de \mathbb{Z}

Définition 2.3. *Soit (Ω, \mathcal{F}) un espace probabilisable. On appelle système complet d'événements toute famille $\{A_i, i \in I\}$ d'événements de \mathcal{F} deux à deux incompatibles tels que*

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

Exemple

Pour tout $A \subset \Omega$, $\{A, \bar{A}\}$ est un système complet d'événements.

2. Notion de probabilité

On considère une expérience aléatoire d'univers Ω . Soit $A \subset \Omega$. On répète l'expérience n fois. Soit n_A le nombre de réalisation de A . Soit $f_n(A)$ la fréquence de réalisation de A . Alors

$$f_n(A) = \frac{n_A}{n} \text{ et } 0 \leq f_n(A) \leq 1.$$

On a

(i) $f_n(\Omega) = 1$

(ii) Si A et B sont deux événement incompatibles alors $f_n(A \cup B) = f_n(A) + f_n(B)$

On montre que lorsque n tend vers l'infini, $f_n(A)$ converge vers une quantité $P(A)$ qui sera appelée probabilité de l'événement A .

Définition 2.4. Soit Ω l'univers d'une expérience aléatoire. Soit \mathcal{F} une tribu définie sur Ω . On appelle probabilité sur l'espace probabilisable (Ω, \mathcal{F}) , toute application $P : \mathcal{F} \rightarrow [0, 1]$ telle que

(i) $P(\Omega) = 1$

(ii) Pour toute suite $(A_n)_{n \geq 0}$ d'éléments de \mathcal{F} deux à deux incompatibles on a

$$P\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} P(A_n).$$

La quantité $P(A)$ s'appelle la probabilité de l'événement A et le triplet (Ω, \mathcal{F}, P) espace probabilisé.

Exemple : Probabilité uniforme

Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable fini c'est à dire tel que Ω est fini. On suppose que tous les événements élémentaires ont la même probabilité ce qui s'énonce en disant qu'on a l'hypothèse d'équiprobabilité. Sous cette hypothèse, la probabilité d'un événement A est alors donnée par

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{\text{Nb de cas favorables à } A}{\text{Nb de cas possibles}}.$$

Cette probabilité est dite uniforme sur Ω . Notons qu'il est impossible d'avoir l'hypothèse d'équiprobabilité sur un espace probabilisable non fini.

Propriété 2.1. Toute probabilité sur un espace probabilisable (Ω, \mathcal{F}) vérifie les propriétés suivantes :

(P1) $P(\emptyset) = 0$

(P2) $P(\bar{A}) = 1 - P(A)$

(P3) $P(A \setminus B) = P(A) - P(A \cap B)$

(P4) Si $A \subset B$ alors $P(A) \leq P(B)$

(P5) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(P6) *Formule de Poincaré :*

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{k=3}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k} P(A_{i_1} \cap \dots \cap A_{i_k}),$$

soit

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset [1, n], |I|=k} P\left(\bigcap_{i \in I} A_i\right).$$

(P7) *Pour toute suite croissante $(A_n)_{n \geq 0}$ d'éléments de \mathcal{F} , on a*

$$P\left(\bigcup_{n=0}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

(P8) *Pour toute suite décroissante $(A_n)_{n \geq 0}$ d'éléments de \mathcal{F} , on a*

$$P\left(\bigcap_{n=0}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

Définition 2.5. *Soit (Ω, \mathcal{F}, P) un espace probabilisé.*

- *Si A est un événement de probabilité nulle, on dit que A est négligeable.*
- *Soit \mathcal{P} une propriété. Si $A = \{\omega \in \Omega / \omega \text{ vérifie } \mathcal{P}\}$ et $P(A) = 1$, on dit que \mathcal{P} est vraie presque sûrement (p.s.).*

Remarque : Ces notion dependent de P .

★ Détermination d'une probabilité sur un espace probabilisable dénombrable.

Soit I une partie de \mathbb{N} ou \mathbb{Z} , $\Omega = \{\omega_i, i \in I\}$ un ensemble et $(p_i)_{i \in I}$ une famille de nombre réels. Soit P une application sur $(\Omega, \mathcal{P}(\Omega))$ telle que pour tout $i \in I$, $P(\{\omega_i\}) = p_i$.

Une condition nécessaire et suffisante pour que P soit une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ est que :

$$\begin{cases} \forall i \in I, p_i \geq 0 \\ \sum_{i \in I} p_i = 1. \end{cases}$$

P est alors unique et on a pour tout événement A , $P(A) = \sum_{i \in I / \omega_i \in A} p_i$.

Remarque : La probabilité P sur un espace probabilisable dénombrable $(\Omega, \mathcal{P}(\Omega))$ est entièrement déterminée par la connaissance des probabilités des événements élémentaires.

3. Probabilités conditionnelles

Dans cette section, notre objectif est quantifier les chances de réalisation d'un événement lorsqu'on dispose d'informations sur le résultat de l'expérience sans le connaître.

Exemple. On lance succesivement deux dés non pipés. On suppose qu'on a l'information suivante : " la somme des chiffres obtenus est 8". On cherche sous cette information à évaluer la chance que le premier chiffre du résultat obtenu soit pair.

3.1. Définition

Définition 2.6. Soit (Ω, \mathcal{F}, P) un espace probabilisé. Soient A et B deux événements tels que $P(B) > 0$. On appelle probabilité conditionnelle de A sachant B , la quantité $\frac{P(A \cap B)}{P(B)}$. Pour tout événement B tel que $P(B) > 0$, on appelle probabilité conditionnelle sachant B l'application notée $P(.|B)$ ou P_B définie sur \mathcal{F} et qui à tout élément A de \mathcal{F} associe $\frac{P(A \cap B)}{P(B)}$ qu'on note $P(A/B)$ ou $P_B(A)$.

3.2. Formule des probabilités composées

Soit A_1, \dots, A_n une famille d'événements tels que $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$. Alors, on a $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \times P(A_2/A_1) \times P(A_3/A_1 \cap A_2) \times \dots \times P(A_n/A_1 \cap \dots \cap A_{n-1})$.

Exemple :

3.3. Formule des probabilités totales

Soit $(A_i)_{i \in I}$ un système complet d'événements de probabilités non nulles. Pour tout événement B , on a

$$P(B) = \sum_{i \in I} P(A_i \cap B) = \sum_{i \in I} P(A_i)P(B/A_i).$$

Exemple :

3.4. Formule de Bayes ou probabilités des causes

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

$$P(A_{i_0}/B) = \frac{P(A_{i_0})P(B/A_{i_0})}{\sum_{i \in I} P(A_i)P(B/A_i)}.$$

Exemple :

4. Indépendance d'événements

4.1. Indépendance de deux événements

Définition 2.7. Soit (Ω, \mathcal{F}, P) un espace probabilisé. Soit A et B deux événements. On dit que A et B sont indépendants si $P(A \cap B) = P(A)P(B)$.

Remarque : Si A et B sont tels que $P(A) \neq 0$ et $P(B) \neq 0$, alors A et B sont indépendants si $P(A/B) = P(A)$ et $P(B/A) = P(B)$. i.e. la réalisation de l'un n'a pas d'influence sur la chance de réalisation de l'autre.

Exemple :

Propriétés

- (i) Les événements A et B sont indépendants si et seulement si les événements A et \bar{B} sont indépendants.
- (ii) Les événements A et B sont indépendants si et seulement si les événements \bar{A} et \bar{B} sont indépendants.
- (iii) L'événement certain est indépendant de tout événement.
- (iv) L'événement impossible est indépendant de tout événement.
- (v) Deux événements incompatibles en dehors de l'univers et de l'événement impossible ne sont pas indépendants.

4.2. Indépendance mutuelle de n événements

Définition 2.8. Soient A_1, \dots, A_n n événements d'un espace probabilisé (Ω, \mathcal{F}, P) . Les événements A_1, \dots, A_n sont dits mutuellement indépendants si pour tout $k \in \{2, \dots, n\}$ et pour tout $i_1, \dots, i_k \in \{1, \dots, n\}$ avec $i_1 < i_2 < \dots < i_k$, on a

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j}).$$

Remarque

L'indépendance mutuelle de n événements est décrite à l'aide de $\sum_{k=2}^n C_n^k = 2^n - n - 1$ relations.

Exemple A, B, C sont mutuellement indépendants si

- (1) $P(A \cap B) = P(A)P(B)$
- (2) $P(A \cap C) = P(A)P(C)$
- (3) $P(B \cap C) = P(B)P(C)$
- (4) $P(A \cap B \cap C) = P(A)P(B)P(C)$

On a $\sum_{k=2}^3 C_3^k = 2^3 - 3 - 1 = 4$ relations.

Chapitre 3

Variables aléatoires réelles

1. Exemples et définitions

Exemple 3.1. *On lance une pièce de monnaie trois fois et on note le résultat.*

L'univers de cette expérience aléatoire est

$$\Omega = \{PPP, PPF, PFP, FPP, PFF, FPF, FFP, FFF\},$$

où P désigne "pile" et F "face".

Soit X : le nombre de piles obtenus.

X associe un nombre réel $X(\omega)$ à chaque résultat $\omega \in \Omega$.

L'ensemble des valeurs possibles pour X est $X(\Omega) = \{0, 1, 2, 3\}$.

Exemple 3.2. *Soient A et B les points de coordonnées respectives $(1, 0)$ et $(1, 1)$ dans un repère orthonormé d'origine O . On tire un point au hasard dans le triangle OAB .*

L'univers de cette expérience aléatoire est

$$\Omega = \text{l'intérieur du triangle } OAB.$$

On note X l'abscisse du point tiré.

De même, X associe un nombre réel $X(\omega)$ à chaque résultat $\omega \in \Omega$.

L'ensemble des valeurs possibles pour X est $X(\Omega) =]0, 1[$.

Définition 3.1. *Une variable aléatoire réelle (v.a.r) X sur un univers Ω est une application qui associe à chaque résultat $\omega \in \Omega$ un nombre réel $x = X(\omega)$.*

L'ensemble des valeurs possibles pour X est $X(\Omega)$, il est appelé support de X et dénoté S_X .

Si le support $X(\Omega)$ est fini ou dénombrable, on dit que X est une variable aléatoire discrète (v.a.r.d). Dans le cas contraire, X est dit continue (v.a.r.c).

Note : L'exemple 3.1 illustre une variable aléatoire discrète, l'exemple 3.2 une variable aléatoire continue. Il existe aussi des v.a. mixtes, i.e. discrète pour certains éléments de l'univers Ω et continue pour d'autres.

Remarque.

La définition précédente de variable aléatoire est relative à l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$. Mais elle peut être généralisée à tout espace probabilisable (Ω, \mathcal{F}) , où $\mathcal{F} \subset \mathcal{P}(\Omega)$ est une tribu sur Ω .

Définition 3.2. Soit (Ω, \mathcal{F}) un espace probabilisable et \mathcal{B} la tribu borélienne sur \mathbb{R} . Une variable aléatoire réelle sur (Ω, \mathcal{F}) est une application

$$X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B})$$

telle que

$$\forall B \in \mathcal{B}, \quad X^{-1}(B) \in \mathcal{F} \tag{3.1}$$

Dans la pratique, pour vérifier la relation 3.1, il suffit de la vérifier pour tout B de \mathcal{B}_{S_X} , la tribu induite sur S_X par \mathcal{B} .

Notation : $X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}$ est notée $\{X \in B\}$.

L'exemple suivant montre l'importance de la relation 3.1.

Exemple 3.3. Considérons l'expérience aléatoire qui consiste à lancer un dé.

Alors, $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{F}_1 = \mathcal{P}(\Omega)$.

Soit l'application X :

$$X(\omega) = \begin{cases} 1, & \text{si } \omega = 6 \\ 0, & \text{sinon.} \end{cases}$$

On a $S_X = \{0, 1\}$, $\mathcal{B}_{S_X} = \mathcal{P}(S_X) = \{\emptyset, \{0\}, \{1\}, S_X\}$.

On a $X^{-1}(\emptyset) = \emptyset \in \mathcal{F}_1$; $X^{-1}(\{0\}) = \{1, 2, 3, 4, 5\} \in \mathcal{F}_1$; $X^{-1}(\{1\}) = \{6\} \in \mathcal{F}_1$; $X^{-1}(S_X) = \Omega \in \mathcal{F}_1$;

Ainsi, X est une variable aléatoire sur (Ω, \mathcal{F}_1) .

Par contre, X n'est pas une variable aléatoire sur (Ω, \mathcal{F}_2) avec $\mathcal{F}_2 = \{\emptyset, \{5, 6\}, \{1, 2, 3, 4\}, \Omega\}$, car $X^{-1}(\{0\}), X^{-1}(\{1\}) \notin \mathcal{F}_2$ (la condition 3.1 n'est pas vérifiée).

Proposition 3.1. Une application $X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B})$ est une variable aléatoire réelle ssi

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

Dans la suite, on considère un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

2. Loi d'une variable aléatoire réelle

Soit X une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $(\mathbb{R}, \mathcal{B})$.

On définit une probabilité sur \mathcal{B} :

$$\forall B \in \mathcal{B}, P_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{X \in B\}).$$

Cette probabilité est appelée loi de X .

En pratique, il est difficile de pouvoir déterminer l'application P_X de façon explicite. Donc pour déterminer la loi d'une variable aléatoire, on considère des fonctions qui la caractérisent.

2.1. Fonction de répartition

Définition 3.3. Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{F}, \mathbb{P})$. On appelle fonction de répartition de X la fonction $F_X : \mathbb{R} \rightarrow \mathbb{R}$ telle que

$$F_X(x) = \mathbb{P}(X^{-1}(]-\infty, x])) = \mathbb{P}(X \leq x).$$

La fonction de répartition vérifie les importantes propriétés suivantes :

Propriété 3.1.

- (1) $0 \leq F_X(x) \leq 1$, pour tout $x \in \mathbb{R}$
- (2) $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$
- (3) F_X est croissante (au sens large)
- (4) F_X est continue à droite en tout point de $\mathbb{R} : \forall x \in \mathbb{R}, \lim_{u \rightarrow x} F_X(u) = F_X(x)$.

Réciproquement, on prouve qu'une fonction vérifiant les propriétés précédentes est la fonction de répartition d'une certaine variable aléatoire réelle (discrète ou continue).

On a aussi les propriétés suivantes : $\forall a, b \in \mathbb{R}$,

- $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$
- $\mathbb{P}(a \leq X \leq b) = F_X(b) - \lim_{u \nearrow a} F_X(u)$
- $\mathbb{P}(a < X < b) = \lim_{u \nearrow b} F_X(u) - F_X(a)$
- $\mathbb{P}(a \leq X < b) = \lim_{u \nearrow b} F_X(u) - \lim_{u \nearrow a} F_X(u)$.
- Si X est une v.a.r. discrète, alors $F_X(x)$ est une fonction en escalier
- Si X est continue alors $F_X(x)$ est une fonction continue.

2.2. Fonction de masse

Définition 3.4. Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{F}, \mathbb{P})$. On appelle fonction de masse de X la fonction $p_X : \mathbb{R} \rightarrow \mathbb{R}$ telle que

$$p_X(x) = \mathbb{P}(X^{-1}(\{x\})) = \mathbb{P}(X = x).$$

La fonction de masse en x_0 représente donc le saut de la fonction de répartition en ce point :

$$p_X(x_0) = F_X(x_0) - \lim_{u \rightarrow x_0^-} F_X(u).$$

On prouve qu'il existe un nombre fini ou dénombrable de points où la fonction de masse est non nulle. Cela signifie encore que F_X est continue sauf en un nombre fini ou dénombrable de points.

Remarque :

- La fonction de masse p_X d'une variable aléatoire réelle continue X est nulle sur tout \mathbb{R} ; $\forall x \in \mathbb{R}, p_X(x) = \mathbb{P}(X = x) = 0$.

- La fonction de masse p_X d'une variable aléatoire réelle discrète X est nulle en dehors de son support S_X ; $\forall x \in \mathbb{R} \setminus S_X, p_X(x) = 0$ et $\sum_{x \in S_X} p_X(x) = 1$.

Réciproquement, on prouve qu'une fonction p vérifiant

$$(1) \quad \forall x \in \mathbb{R}, p(x) \geq 0$$

$$(2) \quad \sum_{x \in S} p(x) = 1.$$

est la fonction de masse d'une certaine variable aléatoire réelle discrète de support S .

2.3. Fonction de densité

Définition 3.5. Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{F}, \mathbb{P})$. On appelle fonction de densité de X la fonction $f_X : \mathbb{R} \rightarrow \mathbb{R}$ telle que

$$\forall B \in \mathcal{B}, P_X(B) = \mathbb{P}(X \in B) = \int_B f_X(t) dt.$$

En particulier,

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(t) dt.$$

La fonction de densité vérifie les propriétés suivantes :

Propriété 3.2.

$$(1) \quad f_X(x) \geq 0, \text{ pour tout } x \in \mathbb{R}$$

$$(2) \quad f \text{ est continue par morceaux}$$

$$(3) \int_{-\infty}^{+\infty} f(t)dt = 1.$$

Réciproquement, on prouve qu'une fonction vérifiant les propriétés précédentes est la fonction de densité d'une certaine variable aléatoire réelle continue.

3. Variables aléatoires réelles discrètes

3.1. Loi d'une v.a.r. discrète

La loi de probabilité P_X d'une v.a.r. discrète X est entièrement déterminée par sa fonction de masse p_X .

Soit X une variable aléatoire réelle discrète (v.a.r.d) définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ de support $X(\Omega) = \{x_i, i \in I\}$, avec I une partie de \mathbb{N} .

Remarque : $\{X = x_i\}_{i \in I}$ est un système complet d'événements appelé système complet associé à la v.a.r. X .

En pratique, pour déterminer la loi de X , on détermine les valeurs x_i susceptibles d'être prises par X , puis les probabilités $p_i = p_X(x_i) = \mathbb{P}(X = x_i)$.

Si $I = [[1, n]]$ avec n petit, on représente souvent les résultats dans un tableau.

Exemple 3.4. *Une urne contient 6 boules numérotées de 1 à 6. On tire sans remise 3 boules. On note X le plus petit nombre tiré et Y le plus grand nombre tiré. Donner les lois de X et Y .*

Exemple 3.5. *On jette une pièce de monnaie. On désigne par X le nombre de jets successifs nécessaires pour obtenir le côté pile pour la première fois. Déterminer la loi de X .*

Soit I une partie de \mathbb{N} ou \mathbb{Z} et $\{(x_i, p_i), i \in I\}$.

Une condition nécessaire et suffisante pour qu'une partie de \mathbb{R}^2 , $\{(x_i, p_i), i \in I\}$ telle que $x_i \neq x_j, \forall i \neq j$ soit la loi de probabilité d'une v.a.r. discrète est que :

$$\begin{cases} \forall i \in I, p_i \geq 0 \\ \sum_{i \in I} p_i = 1. \end{cases}$$

Soit X une v.a.r.d de support $X(\Omega) = \{x_i, i \in I\}$. Alors sa fonction de répartition F_X est constante par morceaux et les points de discontinuité sont les x_i .

Plus précisément, si $X(\Omega) = \{x_1, x_2, \dots, x_i, \dots\}$, où les valeurs sont rangées dans l'ordre croissant et $p_i = \mathbb{P}(X = x_i)$ la loi de X , alors, pour tout $x \in \mathbb{R}$, $F_X(x)$, est définie par :

$$\begin{aligned} \forall x \in]-\infty, x_1[, & F_X(x) = 0. \\ \forall k \geq 1, \forall x \in [x_k, x_{k+1}[, & F_X(x) = p_1 + p_2 + \dots + p_k = \sum_{i=1}^k p_i. \end{aligned}$$

- Si $X(\Omega)$ est fini et si x_n est la plus grande valeur, on a

$$\forall x \in [x_n, +\infty[, \quad F_X(x) = p_1 + p_2 + \dots + p_n = \sum_{i=1}^n p_i = 1.$$

- Si $X(\Omega)$ est infini

$$\sum_{i=1}^{+\infty} p_i = 1.$$

Exemple 3.6. On considère une variable aléatoire discrète X telle que $X(\Omega) = \{1, 2, 3, 4\}$ et $\mathbb{P}(X = 1) = \frac{1}{3}$, $\mathbb{P}(X = 2) = \frac{1}{4}$, $\mathbb{P}(X = 3) = \frac{1}{6}$. Déterminer sa fonction de répartition.

Connaissant, la fonction de répartition F_X d'une variable aléatoire réelle discrète X , on peut reconstituer sa loi :

$$\forall x \in X(\Omega), \quad \mathbb{P}(X = x) = F_X(x) - \lim_{u \underset{<}{\rightarrow} x} F_X(u).$$

En particulier, si $X(\Omega) = \{x_1, x_2, \dots, x_i, \dots\}$, où les valeurs sont rangées dans l'ordre croissant, alors

$$\mathbb{P}(X = x_1) = F_X(x_1),$$

$$\forall i \geq 2, \quad \mathbb{P}(X = x_i) = F_X(x_i) - F_X(x_{i-1}).$$

Remarque : Il est parfois plus simple de déterminer la loi d'une variable aléatoire réelle à partir de sa fonction de répartition.

Exemple 3.7. On jette successivement et indépendamment sur une table deux dés tétraédriques dont les faces sont numérotées de 1 à 4. On note X_1 le premier nombre obtenu, X_2 le deuxième nombre obtenu et X le plus grand des deux nombres obtenus.

Donner la loi de X .

3.2. Loi d'une fonction d'une v.a.r. discrète

Soit X une variable aléatoire réelle discrète (v.a.r.d) définie sur $(\Omega, \mathcal{F}, \mathbb{P})$, de support S_X et de loi p_X . Soit φ une fonction numérique définie sur S_X . Soit $Y = \varphi(X)$. Alors Y est une v.a.r.d de support $S_Y = \varphi(S_X) = \{y \in \mathbb{R} | \exists x \in S_X, y = \varphi(x)\}$ et de loi p_Y définie par :

$$y \in S_Y, \quad p_Y(y) = \mathbb{P}(Y = y) = \sum_{x \in S_X | \varphi(x)=y} p_X(x) = \sum_{x \in S_X | \varphi(x)=y} \mathbb{P}(X = x).$$

Exemple 3.8. On joue à un jeu de pile ou face. On perd 1 point si on obtient face, on gagne 1 point sinon. Pour tout entier non nul n , on note X_n la somme des points obtenus au cours des n premiers jeux.

1) Déterminer la loi X_3 .

2) On pose $Y = |X_3|$ et $Z = 2X_3 + 3$. Déterminer la loi de Y et de Z .

Exemple 3.9. Soit X v.a.r.d de support $S_X = \{-2, -1, 0, 1, 2\}$ et de masse de probabilité p_X définie par :

x	-2	-1	0	1	2
$p_X(x)$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	p

- 1) Trouver la valeur de p .
- 2) On pose $Y = X^2$. Déterminer la loi de Y .

4. Variables aléatoires réelles continues

On rappelle qu'une variable aléatoire réelle X est dite continue lorsqu'elle n'est pas discrète, autrement dit lorsque son support S_X est un intervalle de \mathbb{R} éventuellement \mathbb{R} lui-même.

4.1. Loi d'une v.a.r. continue

La fonction de répartition F_X d'une variable aléatoire réelle continue X est continue et caractérise la loi de X .

$$\checkmark \mathbb{P}(a < X \leq b) = \mathbb{P}(a < X < b) = \mathbb{P}(a \leq X \leq b) = F_X(b) - F_X(a).$$

La v.a.r.c X est dite absolument continue, si sa fonction de répartition F_X est absolument continue c'est à dire s'il existe une fonction densité f_X , telle que

$$\forall x \in \mathbb{R}, F_X(x) = \int_{-\infty}^x f_X(t)dt.$$

Ce qui revient à dire que F_X est dérivable en tout point x de \mathbb{R} où f_X est continue et on a $F'_X(x) = f_X(x)$.

Dans ce cas, la loi de X est déterminée par la densité f_X .

$$\checkmark \mathbb{P}(a \leq X \leq b) = \int_a^b f_X(t)dt = F_X(b) - F_X(a).$$

Remarque : Une variable aléatoire réelle peut être ni discrète, ni absolument continue.

Exercice 3.1. On considère la fonction F définie par :

$$F(x) = \begin{cases} 0, & \text{si } x < 0 \\ \frac{3}{4}x^2 - \frac{1}{4}x^3, & \text{si } 0 \leq x < 2 \\ 1, & \text{si } x \geq 2 \end{cases}$$

- 1) Montrer que F est la fonction de répartition d'une variable aléatoire réelle continue X dont on déterminera l'ensemble des valeurs possibles.
- 2) Montrer que la v.a. X est absolument continue et déterminer sa densité.
- 3) Calculer $\mathbb{P}(X < 1,5)$, $\mathbb{P}(0,5 < X \leq 1,5)$.

Exercice 3.2. On considère la fonction f définie par :

$$f(x) = \begin{cases} ke^{-2x}, & \text{si } x \geq 0 \\ 0, & \text{si non} \end{cases}$$

1) Déterminer la valeur de k pour que f soit une densité de probabilité d'une variable aléatoire réelle continue X .

(On utilisera le fait que $\int_0^{+\infty} f(t)dt = \lim_{x \rightarrow +\infty} \int_0^x f(t)dt$)

2) Déterminer sa fonction de répartition.

3) Calculer $\mathbb{P}(X > 1)$.

4.2. Loi d'une fonction continue d'une v.a.r. continue

Soit X une variable aléatoire réelle discrète (v.a.r.d) définie sur $(\Omega, \mathcal{F}, \mathbb{P})$, de support S_X et de loi p_X . Soit φ une fonction continue définie sur S_X . Soit $Y = \varphi(X)$. Alors Y est une v.a.r.c de support $S_Y = \varphi(S_X) = \{y \in \mathbb{R} | \exists x \in S_X, y = \varphi(x)\}$. Sa loi s'obtient en déterminant sa fonction de répartition F_Y à l'aide de celle de X , F_X selon la régularité de φ .

Cas d'une transformation monotone

Dans le cas où φ est dérivable, strictement croissante, on a :

$$F_Y(y) = P[Y \leq y] = P[\varphi(X) \leq y] = P[X \leq \varphi^{-1}(y)] = F_X(\varphi^{-1}(y)),$$

La densité correspondante est :

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{\varphi'(\varphi^{-1}(y))} f_X(\varphi^{-1}(y)).$$

Dans le cas où φ est dérivable, strictement décroissante, on a :

$$F_Y(y) = P[Y \leq y] = P[\varphi(X) \leq y] = P[X \geq \varphi^{-1}(y)] = 1 - F_X(\varphi^{-1}(y)),$$

La densité correspondante est :

$$f_Y(y) = \frac{d}{dy} F_Y(y) = -\frac{1}{\varphi'(\varphi^{-1}(y))} f_X(\varphi^{-1}(y)).$$

Donc de façon générale, si φ est dérivable, strictement monotone, la densité f_Y de $Y = \varphi(X)$ est donnée par :

$$f_Y(y) = \left| \frac{1}{\varphi'(\varphi^{-1}(y))} \right| f_X(\varphi^{-1}(y))$$

ou

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|,$$

avec $x = \varphi^{-1}(y)$ exprimé en terme de y .

Exemple 3.10. Soit X une v.a.r.c de densité f_X :

$$f_X(x) = \begin{cases} 1, & \text{si } 0 \leq x \leq 1 \\ 0, & \text{sinon.} \end{cases}$$

Déterminer la densité de $Y = X^2$.

Cas d'une transformation non monotone

Soit X une v.a.r.c de densité f_X :

$$f_X(x) = \begin{cases} \frac{1}{2}, & \text{si } -1 \leq x \leq 1 \\ 0, & \text{sinon.} \end{cases}$$

Déterminer la densité de $Y = X^2$.

5. Moments d'une variable aléatoire réelle

Définition 3.6. Soit X une variable aléatoire réelle. On appelle espérance mathématique ou moyenne de X la quantité si elle existe :

$$\mathbb{E}(X) = \begin{cases} \sum_{x \in X(\Omega)} xp_X(x) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x), & \text{si } X \text{ est discrète,} \\ \int_{-\infty}^{+\infty} xf_X(x)dx, & \text{si } X \text{ est absolument continue.} \end{cases}$$

Exemple 3.11. Soit X v.a.r.d de support $S_X = \{-2, -1, 0, 1, 2\}$ et de masse de probabilité p_X définie par :

x	-2	-1	0	1	2
$p_X(x)$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	p

- 1) Trouver la valeur de p .
- 2) Calculer l'espérance mathématique $\mathbb{E}(X)$ de X .

Exemple 3.12. On effectue une succession de lancers de dé cubique bien équilibré, jusqu'à ce que l'on obtienne 6. Soit X le nombre de lancers effectués.

Calculer l'espérance mathématique $\mathbb{E}(X)$ de X .

Exercice 3.3. On suppose que la durée de vie d'un individu dans une population donnée est modélisée par une v.a.c. X dont la fonction densité est donnée par :

$$f(x) = \begin{cases} kx^2(100 - x)^2, & \text{si } 0 \leq x \leq 100 \\ 0, & \text{sinon.} \end{cases}$$

où k est une contante positive.

- 1) Déterminer la valeur de k .
- 2) Calculer la probabilité qu'un individu meure entre 60 ans et 70 ans.
- 3) Quelle est l'espérance de vie d'un individu dans cette population ?

Théorème 3.1. Soit X une variable aléatoire réelle.

- Si X est discrète de loi p , et g une fonction définie de \mathbb{R} dans \mathbb{R} . Alors

$$\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x)p(x) = \sum_{x \in X(\Omega)} g(x)\mathbb{P}(X = x),$$

si cette somme est finie. Réciproquement, si cette égalité tient pour toute fonction g continue positive ou bornée, alors p est la loi de la v.a.r. discrète X .

- Si X est continue de densité f , et g une fonction définie de \mathbb{R} dans \mathbb{R} . Alors

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx,$$

si cette intégrale existe. Réciproquement, si cette égalité tient pour toute fonction g continue positive ou bornée, alors f est la densité de la v.a.r continue X .

Exemple 3.13. Soit X une v.a. de densité

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

Déterminer la loi de X^2 .

Posons $Y = X^2 = \varphi(X)$, où $\varphi(x) = x^2$. Pour toute fonction g continue positive, on a

$$\begin{aligned} \mathbb{E}(g(Y)) &= \mathbb{E}(g(\varphi(X))) = \int_{-\infty}^{+\infty} g(\varphi(x))f(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(x^2)e^{-\frac{1}{2}x^2} dx, \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} g(x^2)e^{-\frac{1}{2}x^2} dx \end{aligned}$$

Pour $x \in [0, +\infty[$, on pose $y = x^2$ et donc $x = \sqrt{y}$. D'où

$$\mathbb{E}(g(Y)) = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} g(y)e^{-\frac{1}{2}y} \frac{1}{2\sqrt{y}} dy = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}y} \mathbf{1}_{[0, +\infty[} dy,$$

Par conséquent, Y a pour densité $f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}y} \mathbf{1}_{[0, +\infty[}$.

Propriété 3.3. - $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.

- Si $X \leq Y$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

- Si X et Y sont indépendantes, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Définition 3.7. Soit X une variable aléatoire réelle. On appelle variance de X la quantité si elle existe :

$$V(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

Propriété 3.4. - $V(X) \geq 0$

- $V(aX + b) = a^2V(X)$.

- Si X et Y sont indépendantes, alors $V(X + Y) = V(X) + V(Y)$.

Définition 3.8. Soit X une variable aléatoire réelle.

- On appelle moment d'ordre $k \geq 1$ de X la quantité : $m_k = \mathbb{E}(X^k)$.
- On appelle moment centré d'ordre $k \geq 1$ de X la quantité : $\mu_k = \mathbb{E}[(X - \mathbb{E}(X))^k]$.

Remarque 3.1. $\mathbb{E}(X) = m_1$ et $V(X) = \mu_2$.

Définition 3.9. Soit X une variable aléatoire réelle. On appelle écart type de X la racine carrée de la variance $V(X)$: $\sigma_X = \sqrt{V(X)}$.

Proposition 3.2 (Inégalité de Markov). Soit X une variable aléatoire réelle positive. Alors

$$\mathbb{P}(X \geq \alpha) \leq \frac{1}{\alpha} \mathbb{E}(X), \quad \forall \alpha > 0.$$

Proposition 3.3 (Inégalité de Bienaymé- Tchebychev). Soit X une variable aléatoire réelle. Alors

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \alpha) \leq \frac{1}{\alpha^2} V(X), \quad \forall \alpha > 0.$$

6. Fonction caractéristique - Fonction génératrice

6.1. Fonction caractéristique

Définition 3.10. Soit X une variable aléatoire réelle. On appelle fonction caractéristique de X la fonction $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}$ définie par :

$$\varphi_X(t) = \mathbb{E}(e^{itX})$$

Dans le cas discret :

$$\varphi_X(t) = \sum_{i \in I} e^{itx_i} \mathbb{P}(X = x_i)$$

Dans le cas continu :

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx = \hat{f}(t).$$

✓ Formule d'inversion de Fourier.

Si $|\varphi_X(t)|$ est intégrable, on a :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \hat{f}(t) dt.$$

Théorème 3.2. Soit X et Y deux variables aléatoires réelles définies sur $(\Omega, \mathcal{F}, \mathbb{P})$. Si $\varphi_X(t) = \varphi_Y(t)$ alors X et Y ont la même loi.

Proposition 3.4. Soit X une variable aléatoire réelle.

- Si $\mathbb{E}(|X|) < +\infty$, alors φ_X est dérivable en 0 et $\mathbb{E}(X) = -i\varphi'_X(0)$.
- Si $\mathbb{E}(X^2) < +\infty$, alors φ_X est deux fois dérivable en 0 et $\mathbb{E}(X^2) = -\varphi''_X(0)$. Ainsi

$$V(X) = -\varphi''_X(0) + [\varphi'_X(0)]^2.$$

-De manière générale, si $k \geq 1$, $\mathbb{E}(|X|^k) < +\infty$, alors φ_X est k -fois dérivable en 0 et $\mathbb{E}(X^k) = (-i)^k \varphi_X^k(0)$.

Remarque 3.2. - $|\varphi_X(t)| \leq 1$

- Si $Y = aX + b$, alors $\varphi_Y(t) = e^{itb} \varphi_X(at)$.

- Si X et Y sont indépendantes, alors $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.

6.2. Fonction génératrice

Définition 3.11. Soit X une variable aléatoire réelle à valeurs dans \mathbb{N} . On appelle fonction génératrice de X la fonction $G_X : [-1, 1] \rightarrow \mathbb{R}$ définie par :

$$G_X(s) = \sum_{k=1}^{+\infty} s^k \mathbb{P}(X = k) = \mathbb{E}(s^X)$$

Théorème 3.3. Soit X et Y deux variables aléatoires réelles définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans \mathbb{N} . Si $G_X(s) = G_Y(s)$, $\forall s \in [-1, 1]$, alors X et Y ont la même loi.

Proposition 3.5. Soit X une variable aléatoire réelle à valeurs dans \mathbb{N} .

- G_X est continue sur $[-1, 1]$ et indéfiniment dérivable sur $] - 1, 1[$.

- Si $\mathbb{E}(X) < +\infty$, alors $\mathbb{E}(X) = \lim_{s \rightarrow 1^-} G_X'(s)$.

- Si $\mathbb{E}(X^2) < +\infty$, alors $\mathbb{E}(X(X-1)) = \lim_{s \rightarrow 1^-} G_X''(s)$.

- Si X et Y sont indépendantes, alors $G_{X+Y}(s) = G_X(s)G_Y(s)$.

Remarque 3.3. Si on connaît la fonction génératrice d'une variable aléatoire réelle X à valeurs dans \mathbb{N} , il suffit de faire un développement en série entière pour avoir la loi de X .

Chapitre 4

Lois de probabilités usuelles

1. Lois usuelles discretées

1.1. Loi de Bernoulli $\mathcal{B}(1, p)$

On appelle épreuve de Bernoulli, toute expérience aléatoire à deux issues, l'une est appelée succès et l'autre l'échec.

Exemples :

- Le jet d'une pièce de monnaie
- Le jet d'un dé cubique et l'observation d'une face marquée.
- Se présenter à un examen
- Dans une population subdivisée en deux parties, le tirage d'un individu.

Le résultat de cette expérience est représenté par une v.a. X , appelée v.a de Bernoulli, définie par :

$$X(\omega) = \begin{cases} 1, & \text{si le choix de } \omega \text{ donne un succès} \\ 0, & \text{sinon.} \end{cases}$$

Ainsi, $X(\Omega) = \{0, 1\}$. Si la probabilité du succès est p ($p = \mathbb{P}(\text{succs})$), alors $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$. On dit que X suit une loi de Bernoulli de paramètre p , ce qu'on écrit $X \rightsquigarrow \mathcal{B}(1, p)$.

En pratique : Ces variables aléatoires servent à modéliser des expériences aléatoires n'ayant que deux issues possibles, que lon code avec 0 ou 1 : lancer d'un pile/face, expérience avec succès/échec, état d'un composant fonctionne/en panne...

- $F_X(x) = \begin{cases} 0, & \text{si } x < 0 \\ 1 - p, & \text{si } 0 \leq x < 1 \\ 1, & \text{si } x \geq 1. \end{cases}$
- $\mathbb{E}(X) = p, \quad V(X) = p(1 - p)$
- $\varphi_X(t) = (1 - p) + pe^{it}, \quad G_X(s) = (1 - p) + ps$

1.2. Loi Binomial $\mathcal{B}(n, p)$

On effectue n répétitions indépendantes d'une épreuve de Bernoulli et on note X le nombre de fois où on a obtenu le succès. On définit ainsi une variable aléatoire X qui suit une loi Binomial de paramètres n et $p = \text{Prob}(\text{Succès})$, caractérisé par

$$X(\Omega) = \{0, 1, \dots, n\} \text{ et pour tout } k \in X(\Omega), \quad \mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k}.$$

On écrit $X \rightsquigarrow \mathcal{B}(n, p)$.

$$\bullet F_X(x) = \begin{cases} 0, & \text{si } x < 0 \\ (1-p)^n, & \text{si } 0 \leq x < 1 \\ \sum_{i=1}^k C_n^i p^i (1-p)^{n-i}, & \text{si } k \leq x < k+1 \\ 1, & \text{si } x \geq n \end{cases}$$

- $\mathbb{E}(X) = np, \quad V(X) = np(1-p)$
- $\varphi_X(t) = (1-p + pe^{it})^n, \quad G_X(s) = (1-p + ps)^n.$

Remarque : $X \rightsquigarrow \mathcal{B}(n, p) \iff X = \sum_{i=1}^n X_i$, avec $X_i \rightsquigarrow \mathcal{B}(1, p), \forall i = 1, \dots, n$ indépendantes.

Cette loi apparaît à chaque fois que l'on somme n variables aléatoires indépendantes et identiquement distribuées de loi de Bernoulli de paramètre p .

Exemple 4.1. *On lance dix fois dans les mêmes conditions une pièce de monnaie, quelle est la probabilité d'obtenir 5 faces.*

Le lancer d'une pièce de monnaie est une épreuve de Bernoulli. Le fait de lancer dix fois dans les mêmes conditions la pièce constitue une répétition indépendante de la même épreuve. X = nombre de fois où l'on obtient face. Alors $X \rightsquigarrow \mathcal{B}(n, p)$, avec $n = 10$ et $p = \frac{1}{2}$.

1.3. Loi Hypergéométrique $\mathcal{H}(N, n, p)$ ou $\mathcal{H}(N, Np, n)$

On effectue n tirage sans remise dans une urne contenant N objets dont Np objets A (p étant la proportion des objets A) et on note X le nombre d'objets A tirés. Alors X suit une loi Hypergéométrique de paramètres N, Np et n ou N, p et n , caractérisé par

$$X(\Omega) = \{\max(0, n - Nq), \dots, \min(n, Np)\} \text{ avec } Np \in \mathbb{N}, q = 1 - p$$

et pour tout

$$k \in X(\Omega), \quad \mathbb{P}(X = k) = \frac{C_{Np}^k C_{Nq}^{n-k}}{C_N^n}.$$

On écrit $X \rightsquigarrow \mathcal{H}(N, Np, n)$.

- $\mathbb{E}(X) = np, \quad V(X) = npq \frac{N-n}{N-1}$

Exemple : Une urne contient 10 pièces dont 4 sont défectueuses. On y prend au hasard un lot de 5 pièces. Soit X la variable aléatoire donnant le nombre de pièces non défectueuses contenues dans le lot.

1. Quel est l'ensemble des valeurs possibles de X ?
2. Donner la loi de probabilité de X
3. Calculer l'espérance, la variance et l'écart type de X .
4. Quelle est la probabilité que le lot ne contienne aucune pièce défectueuse ?

1.4. Loi géométrique $\mathcal{G}(p)$

On effectue des répétitions indépendantes de la même épreuve de Bernoulli jusqu'à l'obtention du premier succès et on note X le nombre de répétition effectuées ou nécessaires. Alors X suit une loi géométrique de paramètre p , avec $p = \text{Prob}(\text{Succes})$, caractérisé par $X(\Omega) = \mathbb{N}^*$ et pour tout $k \in X(\Omega)$, $\mathbb{P}(X = k) = p(1 - p)^{k-1}$.

On écrit $X \rightsquigarrow \mathcal{G}(p)$.

- $F_X(x) = \begin{cases} 0, & \text{si } x < 1 \\ 1 - (1 - p)^n, & \text{si } n \leq x < n + 1. \end{cases}$
- $\mathbb{E}(X) = \frac{1}{p}, \quad V(X) = \frac{1 - p}{p^2}$
- $\varphi_X(t) = \frac{pe^{it}}{1 - (1 - p)e^{it}}, \quad G_X(s) = \frac{ps}{1 - (1 - p)s}$

Remarque : Si X_i est la variable de Bernoulli $\mathcal{B}(1, p)$ associée à la i -ème épreuve, alors $X = \inf\{k \geq 1, X_k = 1\} \rightsquigarrow \mathcal{G}(p)$.

De même si $(X_n)_{n \in \mathbb{N}}$ une suite de variable aléatoire indépendante de même loi.

Si $X = \inf\{k \geq 1, X_k \in B\}$ où $B \in \mathcal{B}_{\mathbb{R}}$, alors $X \rightsquigarrow \mathcal{G}(p)$, où $p = \mathbb{P}(X_0 \in B)$.

Exemple 4.2. On considère une suite infinie de lancers de pièce de monnaie bien équilibrée indépendants. On note X le numéro du premier pile obtenu. Déterminer la loi de X .

Définition 4.1. Soit $X : \Omega \rightarrow \mathbb{N}^*$ une variable aléatoire. On dit que X est sans mémoire si

$$\forall n, k, \mathbb{P}(X > n + k / X > n) = \mathbb{P}(X > k).$$

Propriété "sans mémoire" : Une variable aléatoire de loi géométrique est une variable sans mémoire.

Reciproquement, toute variable aléatoire X à valeurs dans \mathbb{N} , sans mémoire suit une loi géométrique. (Démonstration : à faire au TD).

1.5. Loi Binomial négative $\mathcal{B}_N(r, p)$

- $X \rightsquigarrow \mathcal{B}_N(r, p)$ si $X(\Omega) = \{r, r+1, \dots\}$ et

$$\begin{aligned} \forall k \in X(\Omega), \quad \mathbb{P}(X = k) &= pC_{k-1}^{r-1}p^{r-1}(1-p)^{k-r} \\ &= C_{k-1}^{r-1}p^r(1-p)^{k-r} \end{aligned}$$

- $F_X(x) = \begin{cases} 0, & \text{si } x < r \\ \sum_{i=r}^n C_{i-1}^{r-1}p^r(1-p)^{i-r}, & \text{si } n \leq x < n+1 \end{cases}$

Génération de $\mathcal{B}_N(r, p)$.

$$X \rightsquigarrow \mathcal{B}_N(r, p) \iff X = \sum_{i=1}^r X_i \text{ avec } X_i \rightsquigarrow \mathcal{G}(p), \forall i = 1, \dots, r \text{ indépendantes.}$$

Dans une répétition indépendante de la même épreuve de Bernoulli, la variable aléatoire égale au nombre de répétitions nécessaires à l'obtention du r -ème succès est une variable aléatoire Binomial négative de paramètres r, p , avec $p = \text{Prob}(\text{Succes})$.

$$\{X = k\} = \{\exists I \subset \{1, \dots, k\}, \text{Card}(I) = r \text{ et } \max I = k; i \in I, X_i = 1 \text{ et } i \notin I, X_i = 0\}.$$

$$\mathbb{P}(X = k) = pC_{k-1}^{r-1}p^{r-1}(1-p)^{k-r}$$

- $\mathbb{E}(X) = \sum_{i=1}^r \mathbb{E}(X_i) = \frac{r}{p}, \quad V(X) = \sum_{i=1}^r V(X_i) = \frac{r(1-p)}{p^2}$

- $\varphi_X(t) = \varphi_{\sum_{i=1}^r X_i}(t) = (\varphi_{X_1}(t))^r = \left(\frac{pe^{it}}{1-(1-p)e^{it}}\right)^r, \quad G_X(s) = \left(\frac{ps}{1-(1-p)s}\right)^r$

1.6. Loi de Poisson $\mathcal{P}(\lambda)$

Définition 4.2. Une variable aléatoire X suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$ si :

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k \in \mathbb{N}.$$

On note $X \rightsquigarrow \mathcal{P}(\lambda)$

$$\forall x \in \mathbb{R}, \quad p_X(x) = \frac{\lambda^x}{x!}e^{-\lambda}1_{\mathbb{N}}(x).$$

- $F_X(x) = \begin{cases} 0, & \text{si } x < 0 \\ \sum_{k=0}^n \frac{\lambda^k}{k!}e^{-\lambda}, & \text{si } n \leq x < n+1 \end{cases}$

- $\mathbb{E}(X) = \lambda, \quad V(X) = \lambda$

- $\varphi_X(t) = e^{-\lambda(1-e^{it})}, \quad G_X(s) = e^{-\lambda(1-s)}$

Remarque. Si $X_i \rightsquigarrow \mathcal{P}(\lambda), \forall i = 1, \dots, n$ indépendantes alors $\sum_{i=1}^n X_i \rightsquigarrow \mathcal{P}(n\lambda)$.

Modélisation.

La loi de Poisson est utilisée de façon fondamentale pour modéliser certains phénomènes rares et d'une manière générale dans les problèmes de file d'attente (quand on veut modéliser par exemple le nombre de clients qui se présentent à un guichet, le nombre d'arrivées à un poste de douane dans un intervalle de temps donné, il est naturel d'utiliser une loi de Poisson. Le paramètre λ est le nombre moyen d'arrivées dans l'intervalle de temps considéré.) De même, le nombre d'absents par jour dans une entreprise, le nombre annuel de sinistres par police dans un portefeuille d'assurance à risques homogènes, peuvent être décrits, sous certaines conditions, par une loi de Poisson.

1.7. Loi uniforme sur $\{1, \dots, n\}$; $\mathcal{U}_{\{1, \dots, n\}}$

Définition 4.3. Une variable aléatoire X suit une loi uniforme sur $\{1, \dots, n\}$ avec $n \in \mathbb{N}^*$ si :

$$P(X = k) = \frac{1}{n}, \quad \forall k \in \{1, \dots, n\}.$$

On note $X \rightsquigarrow \mathcal{U}_{\{1, \dots, n\}}$

- $p_X(x) = \frac{1}{n} 1_{\{1, \dots, n\}}(x), \quad \forall x \in \mathbb{R}.$

- $F_X(x) = \begin{cases} 0, & \text{si } x < 1 \\ \frac{k}{n}, & \text{si } k \leq x < k + 1 \\ 1, & \text{si } x \geq n \end{cases}$

- $\mathbb{E}(X) = \frac{1}{n} \sum_{k=1}^n k = \frac{n+1}{2}, \quad \mathbb{E}(X^2) = \frac{1}{n} \sum_{k=1}^n k^2 = \frac{(n+1)(2n+1)}{6}$ et $V(X) = \frac{n^2-1}{12}$

Généralement : $X \rightsquigarrow \mathcal{U}_{\{a_1, \dots, a_n\}} \iff X(\Omega) = \{a_1, \dots, a_n\}$ et $\forall k \in X(\Omega), \quad \mathbb{P}(X = k) = \frac{1}{n}.$

$$\mathbb{E}(X) = \frac{1}{n} \sum_{k=1}^n a_k.$$

Modélisation. Ces variables aléatoires servent à modéliser des expériences aléatoires équiprobables qui consistent à extraire une unité au hasard dans un ensemble contenant n unités : le lancer d'un dé équilibré, choix d'une carte au hasard dans un jeu de cartes...

1.8. $X \rightsquigarrow \mathcal{L}$, avec $X(\Omega) = \mathbb{N}$

$$\mathbb{P}(X = x) = u(x) \geq 0, \quad \sum_{x=1}^{\infty} u(x) = 1.$$

X v.a discret \iff il existe $(a_i)_{i \in I}$ fini ou dénombrable telle que

$$\mathbb{P}(X = x) = p_i, \quad \sum_{i \in I} p_i = 1.$$

Toute série à termes positives (u_n) convergente permet d'obtenir une variable aléatoire de loi $P_X : \mathbb{P}(X = n) = \frac{u_n}{\sum u_n}.$

2. Lois usuelles continues

2.1. Loi uniforme sur $[a, b], a < b \in \mathbb{R}$

- $X \rightsquigarrow \mathcal{U}_{[a,b]} \iff f_X(x) = \frac{1}{b-a} 1_{[a,b]}(x).$
- $F_X(x) = \int_{-\infty}^x f_X(t)dt = \begin{cases} 0, & \text{si } x < a \\ \int_a^x \frac{dt}{b-a}, & \text{si } a \leq x < b \\ 1, & \text{si } x \geq b \end{cases} = \begin{cases} 0, & \text{si } x < a \\ \frac{x-a}{b-a}, & \text{si } a \leq x < b \\ 1, & \text{si } x \geq b \end{cases}$
- $\mathbb{E}(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}$
- $\varphi_X(t) = \frac{e^{itb} - e^{ita}}{it(b-a)},$

NB Si $a = -\alpha, b = \alpha, \alpha > 0 \implies \varphi_X(t) = \frac{e^{it\alpha} - e^{-it\alpha}}{2it\alpha} = \frac{\sin(t\alpha)}{t\alpha}.$

Modélisation. Ces variables aléatoires servent à modéliser des expériences aléatoires équiprobables qui consistent à tirer un point au hasard dans un segment ou dans un intervalle, de temps par exemple.

Les lois uniformes continues sont très utiles de par la facilité des calculs qu'elles permettent et pour leur utilisation dans les simulations.

2.2. Loi exponentielle $\xi(\theta), \theta > 0$

- $X \rightsquigarrow \xi(\theta) \iff f_X(x) = \theta e^{-\theta x} 1_{]0,+\infty[}(x).$
- $F_X(x) = \int_{-\infty}^x f_X(t)dt = \begin{cases} 0, & \text{si } x \leq 0 \\ 1 - e^{-\theta x}, & \text{si } x > 0 \end{cases}$
- $\mathbb{E}(X^k) = \theta \int_0^{+\infty} e^{-\theta x} x^k dx.$

Cette quantité se calcule à l'aide de la fonction Gamma :

Définition 4.4. Nous appelons fonction Gamma la fonction définie par :

$$\Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx, \quad \forall a \in \mathbb{R}_+^*.$$

Propriété 4.1. La fonction Gamma satisfait aux relations suivantes :

$$\Gamma(a+1) = a\Gamma(a), \quad \forall a \in \mathbb{R}_+^*, \quad \Gamma(n+1) = n!, \quad \forall n \in \mathbb{N}, \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Remarque. La fonction Gamma est une généralisation directe de la notion de factorielle pour des nombres réels non entiers. Le logiciel R permet d'obtenir des valeurs de gamma.

On montre que :

$$\int_0^{+\infty} e^{-\rho x} x^{a-1} dx = \frac{\Gamma(a)}{\rho^a}, \quad \forall \rho, a \in \mathbb{R}_+^*.$$

Donc, $\mathbb{E}(X^k) = \theta \frac{\Gamma(k+1)}{\theta^{k+1}} = \frac{k!}{\theta^k}$, par conséquent,

$$\mathbb{E}(X) = \frac{1}{\theta}, \quad \mathbb{E}(X^2) = \frac{2}{\theta^2} \text{ et } V(X) = \frac{1}{\theta^2}$$

- $\varphi_X(t) = \frac{\theta}{\theta - it}$,

Définition 4.5. Soit $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire. On dit que X est sans mémoire si

$$\forall s, t \in \mathbb{R}_+, \mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s).$$

Propriété "sans mémoire" : Une variable aléatoire de loi exponentielle est une variable sans mémoire.

Reciproquement, toute variable aléatoire X à valeurs dans \mathbb{N} , sans mémoire suit une loi exponentielle. (Démonstration : à faire au TD).

Modélisation. A cause de la propriété d'absence de mémoire, la loi exponentielle est utilisée quand on veut modéliser la durée de vie d'un composant électronique, ou le temps entre les arrivées de deux clients successifs dans une file d'attente (il y a en particulier des liens très forts entre les lois exponentielles et les lois de Poisson).

Le paramètre $\theta = \frac{\lambda}{\tau}$, où λ est le nombre moyen d'arrivées dans l'intervalle de temps de longueur τ . C'est donc le nombre moyen d'arrivées par unité de temps.

2.3. Loi Gamma $\gamma(a, \rho)$, $a > 0, \rho > 0$

Définition 4.6. Une variable aléatoire continue X suit une loi Gamma, ou eulerienne de paramètres $a, \rho \in \mathbb{R}_+^*$ si elle admet pour densité de probabilité la fonction

$$f_X(x) = \frac{\rho^a}{\Gamma(a)} e^{-\rho x} x^{a-1} 1_{]0, +\infty[}(x) = \begin{cases} \frac{\rho^a}{\Gamma(a)} e^{-\rho x} x^{a-1}, & \text{si } x > 0 \\ 0, & \text{si } x \leq 0, \end{cases}$$

et note $X \rightsquigarrow \gamma(a, \rho)$.

- $F_X(x) = \int_{-\infty}^x f_X(t) dt = \begin{cases} 0, & \text{si } x \leq 0 \\ \frac{\rho^a}{\Gamma(a)} \int_0^x e^{-\rho t} t^{a-1} dt, & \text{si } x > 0 \end{cases}$

- $\mathbb{E}(X^k) = \frac{\rho^a}{\Gamma(a)} \int_0^{+\infty} e^{-\rho x} x^{a+k-1} dx = \frac{\rho^a}{\Gamma(a)} \cdot \frac{\Gamma(a+k)}{\rho^{a+k}} = \frac{\Gamma(a+k)}{\Gamma(a)\rho^k}$, par conséquent,

$$\mathbb{E}(X) = \frac{\Gamma(a+1)}{\Gamma(a)\rho} = \frac{a}{\rho}, \quad \mathbb{E}(X^2) = \frac{\Gamma(a+2)}{\Gamma(a)\rho^2} = \frac{a(a+1)}{\rho^2} \text{ et } V(X) = \frac{a}{\rho^2}$$

- $\varphi_X(t) = \left(\frac{\rho}{\rho - it}\right)^a$,

Remarque. $\gamma(1, \rho) = \xi(\rho)$. Si $X \rightsquigarrow \gamma(a, \rho)$ alors $Y = \rho X \rightsquigarrow \gamma(a, 1)$.

$X \rightsquigarrow \gamma(n, \rho) \iff X = \sum_{i=1}^n X_i$ avec $X_i \rightsquigarrow \gamma(1, \rho)$, $\forall i = 1, \dots, n$ indépendantes.

Modélisation. La loi Gamma peut décrire des phénomènes de durée de vie, en assurance pour l'étude du temps écoulé entre deux sinistres dans des portefeuilles à risques hétérogènes ou encore pour prendre en compte cette hétérogénéité. En général pour des distributions fortement asymétriques avec une décroissance rapide en queue de distribution, une loi Gamma peut être un bon modèle.

2.4. Loi normale $\mathcal{N}(m, \sigma^2)$ (ou loi de Gauss)

Définition 4.7. Une variable aléatoire continue X suit une loi normale ou encore loi de Laplace-Gauss de paramètres $\sigma \in \mathbb{R}_+^*$ et $m \in \mathbb{R}$ si elle admet pour densité de probabilité la fonction

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-m)^2}, \quad \forall x \in \mathbb{R}.$$

On note $X \rightsquigarrow \mathcal{N}(m, \sigma^2)$.

Si $m = 0$ et $\sigma = 1$, on dit que X suit la loi normale centrée et réduite, ou encore la loi Normale standard $\mathcal{N}(0, 1)$.

Important à retenir :

$$(1) \int_{-\infty}^{+\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

$$(2) \int_{-\infty}^{+\infty} e^{-\alpha(x-t)^2} dx = \int_{-\infty}^{+\infty} e^{-\alpha x^2} dx$$

$$(3) \int_{-\infty}^{+\infty} e^{-\alpha x^2 + \beta x} dx = e^{\frac{\beta^2}{4\alpha}} \sqrt{\frac{\pi}{\alpha}}$$

Soit $X \rightsquigarrow \mathcal{N}(m, \sigma^2)$. Alors

- $\mathbb{E}(X) = m, \quad V(X) = \sigma^2$
- $\varphi_X(t) = e^{itm - \frac{1}{2}\sigma^2 t^2}$

En pratique, on calcule les probabilités relatives aux intervalles à l'aide la fonction de répartition de la loi normale centrée et réduite $\mathcal{N}(0, 1)$ qui dispose d'une table, appelée table de loi normale centrée et réduite :

Cette fonction de répartition est souvent notée π ou \mathcal{N} . On a donc

$$\pi(x) = \mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$$

Soit $X \rightsquigarrow \mathcal{N}(m, \sigma^2)$. Alors $T = \frac{X-m}{\sigma} \rightsquigarrow \mathcal{N}(0, 1)$ et $\pi(t) = \mathbb{P}(T \leq t)$.

Remarque : $\pi(-\infty) = 0$, $\pi(+\infty) = 1$, $\pi(0) = 0.5$ et $\forall t \in \mathbb{R}$, $\pi(-t) = 1 - \pi(t)$.

Pour tout $t \in \mathbb{R}$, $\mathbb{P}(|T| \leq t) = 2\pi(t) - 1$ et $\mathbb{P}(|T| \geq t) = 2(1 - \pi(t))$.

Exemple :

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}\left(\frac{a-m}{\sigma} \leq T \leq \frac{b-m}{\sigma}\right) = \pi\left(\frac{b-m}{\sigma}\right) - \pi\left(\frac{a-m}{\sigma}\right).$$

Modélisation. Ces lois apparaissent dans le théorème central limite, que nous verrons plus tard. Elles sont souvent utiliser pour modéliser de petites erreurs ou variations aléatoires autour d'une quantité déterministe.

2.5. Loi de Cauchy

Définition 4.8. Une variable aléatoire continue X suit une loi de Cauchy, de paramètres $\in \mathbb{R}_+^*$ et $m \in \mathbb{R}$ si elle admet pour densité de probabilité la fonction

$$f_X(x) = \frac{\alpha}{\pi(\alpha^2 + (x-m)^2)}, \quad \forall x \in \mathbb{R}.$$

On note $X \rightsquigarrow \text{Cauchy}(m, \alpha)$

Une variable aléatoire continue X suivant une loi de Cauchy n'a aucun moments.

- Pour $m = 0$, $\varphi_X(t) = e^{-\alpha|t|}$.

Chapitre 5

Vecteurs aléatoires

1. Introduction

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Pour tout entier $n \geq 1$, on note $\mathcal{B}_{\mathbb{R}^n}$ la tribu engendrée par les ouverts de \mathbb{R}^n . Cette tribu est appelée tribu borelienne de \mathbb{R}^n .

Définition 5.1. On appelle vecteur aléatoire à n dimensions ou variable aléatoire à n dimensions toute application \mathbb{X} de (Ω, \mathcal{F}) dans $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ telle pour tout borélien $B \in \mathcal{B}_{\mathbb{R}^n}$,

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

On note par $\{\mathbb{X}(\omega) \in B\}$ l'événement $\{\omega \in \Omega, \mathbb{X}(\omega) \in B\}$.

On note

$$\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \quad \text{ou} \quad \mathbb{X} = (X_1, \dots, X_n)$$

où pour tout $1 \leq i \leq n$, X_i est une variable aléatoire réel ou à une dimension définie sur $(\Omega, \mathcal{F}) : X_i : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

2. Loi conjointe d'un vecteur aléatoire

Définition 5.2. Soit $\mathbb{X} : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}^n$ un vecteur aléatoire. On appelle loi conjointe de \mathbb{X} la probabilité $\mathbb{P}_{\mathbb{X}} : \mathcal{B}_{\mathbb{R}^n} \rightarrow [0, 1]$ définie par $\mathbb{P}_{\mathbb{X}}(B) = \mathbb{P}(\mathbb{X}^{-1}(B)) = \mathbb{P}(\mathbb{X} \in B)$.

Autrement dit, c'est l'application qui permet d'associer à tout n -uplets d'intervalles B_1, \dots, B_n de \mathbb{R} , la probabilité de l'événement $(X_1 \in B_1, \dots, X_n \in B_n)$.

2.1. Fonction de repartition

Si $a = (a_1, \dots, a_n)$ est un élément de \mathbb{R}^n , on pose :

$$\Delta(a) = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 \leq a_1, \dots, x_n \leq a_n\}.$$

Définition 5.3. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire. On appelle fonction de répartition de \mathbb{X} la fonction $F_{\mathbb{X}}$ définie sur \mathbb{R}^n par :
pour tout x dans \mathbb{R}^n ,

$$F_{\mathbb{X}}(x) = \mathbb{P}_{\mathbb{X}}(\Delta(x)).$$

On a alors

$$\begin{aligned} F_{\mathbb{X}}(x) &= \mathbb{P}_{\mathbb{X}}(\Delta(x)) \\ &= \mathbb{P}(\mathbb{X} \in \Delta(x)) \\ &= \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) \end{aligned}$$

pour tout $x = (x_1, \dots, x_n)$ dans \mathbb{R}^n

2.2. Vecteur aléatoire discret

Définition 5.4. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire. On dit que \mathbb{X} est discret si $\mathbb{X}(\Omega)$ est fini ou dénombrable.

La loi $\mathbb{P}_{\mathbb{X}}$ de \mathbb{X} est alors une loi discrète. Elle est entièrement définie par les probabilités élémentaires $p_{\mathbb{X}}(x_1, \dots, x_n)$ où (x_1, \dots, x_n) parcourt $\mathbb{X}(\Omega)$:

$$p_{\mathbb{X}}(x_1, \dots, x_n) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n).$$

2.3. Vecteur aléatoire absolument continu

Définition 5.5. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire. On dit que \mathbb{X} est absolument continu si sa loi de probabilité $\mathbb{P}_{\mathbb{X}}$ possède une densité relativement à la mesure de Lebesgue sur \mathbb{R}^n , c-à-d, il existe une fonction positive et intégrable f telle que l'on ait pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$,

$$F_{\mathbb{X}}(x_1, \dots, x_n) = \int_{\Delta(x_1, \dots, x_n)} f_{\mathbb{X}}(t_1, \dots, t_n) dt_1 \dots dt_n = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_{\mathbb{X}}(t_1, \dots, t_n) dt_1 \dots dt_n.$$

On a alors

$$\frac{\partial^n F_{\mathbb{X}}(x_1, \dots, x_n)}{\partial x_1 \dots \partial x_n} = f_{\mathbb{X}}(x_1, \dots, x_n)$$

en tout point $(x_1, \dots, x_n) \in \mathbb{R}^n$ où f est continue.

La fonction f est appelée alors la densité (conjointe) de probabilité de \mathbb{X} .

Rq 1. La condition

$$\mathbb{P}_{\mathbb{X}}(\mathbb{R}^n) = 1$$

se traduit par

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$$

Rq 2. Tout fonction

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

positive, intégrable telle que

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$$

est la densité de probabilité d'une loi de probabilité unique.

3. Loïs marginales

Définition 5.6. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire de loi $\mathbb{P}_{\mathbb{X}}$. Les loïs marginales de la loi conjointe $\mathbb{P}_{\mathbb{X}}$ sont les loïs de probabilité \mathbb{P}_{X_i} des variables aléatoires X_i , $1 \leq i \leq n$. \mathbb{P}_{X_i} est loi marginale d'ordre i .

- Dans le cas discret, cette loi est donnée par

$$p_{X_i}(x_i) = \mathbb{P}(X_i = x_i) = \sum_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n} \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$$

- Dans le cas continu, cette loi est déterminée par la densité de X_i appelée densité marginale :

$$f_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} f_{\mathbb{X}}(x_1, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n$$

ou par la fonction de répartition de X_i appelée fonction de répartition marginale :

$$F_{X_i}(x_i) = \lim_{\substack{x_j \rightarrow +\infty \\ j \neq i}} F_{\mathbb{X}}(x_1, \dots, x_n)$$

4. Espérance mathématique et matrice de covariance

Définition 5.7 (Espérance mathématique). Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire.

L'espérance de \mathbb{X} est $\mathbb{E}(\mathbb{X}) = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_n) \end{pmatrix}$, où

$$\text{cas discret : } \mathbb{E}(X_i) = \sum_{x_i} x_i \mathbb{P}(X_1 = x_i) = \sum_{x_1, \dots, x_n} x_i \mathbb{P}(X_1 = x_1, \dots, X_n = x_n).$$

$$\text{cas continu : } \mathbb{E}(X_i) = \int_{\mathbb{R}} x_i f_{X_i}(x_i) dx_i = \int_{\mathbb{R}^n} x_i f_{\mathbb{X}}(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Exemple 5.1. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ un vecteur aléatoire de densité

$$f_{\mathbb{X}}(x_1, x_2) = C x_1 x_2 \mathbf{1}_{\{0 \leq x_1 \leq x_2 \leq 1\}}.$$

1. Déterminer C
2. Déterminer les lois marginales
3. Calculer $\mathbb{E}(\mathbb{X})$.

Théorème 5.1. Soit \mathbb{X} un vecteur aléatoire à valeur dans \mathbb{R}^n .

- Si \mathbb{X} est discret, $\mathbb{P}_{\mathbb{X}}$ est la loi de \mathbb{X} ssi pour toute fonction $h : \mathbb{N}^n \rightarrow \mathbb{R}$ positive ou bornée,

$$\mathbb{E}(h(\mathbb{X})) = \sum_{x_1, \dots, x_n} h(x_1, \dots, x_n) \mathbb{P}(X_1 = x_1, \dots, X_n = x_n).$$

- Si \mathbb{X} est continu, $f_{\mathbb{X}}$ est la densité de \mathbb{X} ssi pour toute fonction $h : \mathbb{R}^n \rightarrow \mathbb{R}$ positive ou bornée,

$$\mathbb{E}(h(\mathbb{X})) = \int_{\mathbb{R}^n} h(x_1, \dots, x_n) f_{\mathbb{X}}(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Proposition 5.1. Soit \mathbb{X} un vecteur aléatoire à valeur dans \mathbb{R}^n de densité $f_{\mathbb{X}}$ de support D . Soit $\varphi : D \rightarrow \Delta$ une fonction bijective et différentielle. Soit $\mathbb{Y} = \varphi(\mathbb{X})$. Alors la densité de \mathbb{Y} est

$$f_{\mathbb{Y}}(y) = f_{\mathbb{X}}(\varphi^{-1}(y)) |J_{\varphi^{-1}(y)}| \mathbf{1}_{\Delta},$$

où $|J_{\varphi^{-1}(y)}|$ est le jacobien de φ^{-1} au point y .

$$J_{\varphi^{-1}(y)} = \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \Big|_y.$$

Soit $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto h(x_1, \dots, x_n) = \begin{pmatrix} h_1(x_1, \dots, x_n) \\ \vdots \\ h_n(x_1, \dots, x_n) \end{pmatrix}$, alors le jacobien de h au point a est

$$J_{h(a)} = \frac{\partial(h_1, \dots, h_n)}{\partial(x_1, \dots, x_n)} \Big|_a = \begin{vmatrix} \frac{\partial h_1}{\partial x_1}(a) & \dots & \frac{\partial h_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial x_1}(a) & \dots & \frac{\partial h_n}{\partial x_n}(a) \end{vmatrix}.$$

Définition 5.8 (Variance-Covariance).

a-) Soit \mathbb{X} un vecteur aléatoire à valeurs dans \mathbb{R}^n . On appelle matrice de covariance ou matrice de dispersion de \mathbb{X} la matrice carré d'ordre n , $\Sigma_{\mathbb{X}} = (\sigma_{ij})_{1 \leq i, j \leq n}$,

où $\sigma_{ij} = Cov(X_i, X_j) = \mathbb{E} \left[(X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j)) \right] = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$.

$$\Sigma_{\mathbb{X}} = \mathbb{E} \left[(\mathbb{X} - \mathbb{E}(\mathbb{X})) (\mathbb{X} - \mathbb{E}(\mathbb{X}))^t \right] = \mathbb{E}(\mathbb{X}\mathbb{X}^t) - \mathbb{E}(\mathbb{X})[\mathbb{E}(\mathbb{X})]^t.$$

b-) Soit \mathbb{X} et \mathbb{Y} deux vecteurs aléatoires à valeurs resp. dans \mathbb{R}^n et \mathbb{R}^p .

On appelle covariance de \mathbb{X} et \mathbb{Y} la matrice de taille (n, p) ,

$$Cov(\mathbb{X}, \mathbb{Y}) = \left(Cov(X_i, Y_j) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}.$$

$$Cov(\mathbb{X}, \mathbb{Y}) = \mathbb{E} \left[(\mathbb{X} - \mathbb{E}(\mathbb{X})) (\mathbb{Y} - \mathbb{E}(\mathbb{Y}))^t \right] = \mathbb{E}(\mathbb{X}\mathbb{Y}^t) - \mathbb{E}(\mathbb{X})[\mathbb{E}(\mathbb{Y})]^t.$$

On a $\Sigma_{\mathbb{X}} = Cov(\mathbb{X}, \mathbb{X})$, c'est pour cette raison qu'on l'appelle abusivement variance de $\mathbb{X} : V(\mathbb{X})$.

Propriété 5.1. Soit \mathbb{X} et \mathbb{Y} deux vecteurs aléatoires à valeurs resp. dans \mathbb{R}^n et \mathbb{R}^p . soit $B \in \mathbb{R}^p$ et A une matrice de taille (p, n) telle que $\mathbb{Y} = A\mathbb{X} + B$. Alors

- $\mathbb{E}(\mathbb{Y}) = A\mathbb{E}(\mathbb{X}) + B$
- $\Sigma_{\mathbb{Y}} = A\Sigma_{\mathbb{X}}A^t$.

Propriété 5.2. Soit \mathbb{X} et \mathbb{Y} deux vecteurs aléatoires à valeurs dans \mathbb{R}^n . Alors

- $V(\mathbb{X} + \mathbb{Y}) = V(\mathbb{X}) + V(\mathbb{Y}) + 2Cov(\mathbb{X}, \mathbb{Y})$
- $Cov(\mathbb{X}, \mathbb{Y}) = Cov(\mathbb{Y}, \mathbb{X})$
- $\Sigma_{\mathbb{X}}$ est une matrice symétrique semi-définie positive i.e $\forall x \in \mathbb{R}^n, x^t \Sigma_{\mathbb{X}} x \geq 0$.

Définition 5.9 (Fonction caractéristique-Fonction génératrice).

a-) Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire. On appelle fonction caractéristique de \mathbb{X} la

fonction $\varphi_{\mathbb{X}} : \mathbb{R}^n \rightarrow \mathbb{C}$ définie par :

$$\forall a \in \mathbb{R}^n, \varphi_{\mathbb{X}}(a) = \mathbb{E}(e^{i\langle a, \mathbb{X} \rangle}) \text{ où } \langle a, \mathbb{X} \rangle = a^t \mathbb{X} = \sum_{k=1}^n a_k X_k.$$

b-) Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire à valeurs dans \mathbb{N}^n . On appelle fonction génératrice de \mathbb{X} la fonction $G_{\mathbb{X}} : [-1, 1]^n \rightarrow \mathbb{R}$ définie par :

$$\forall s \in \mathbb{N}^n, \quad G_{\mathbb{X}}(s) = G_{\mathbb{X}}(s_1, \dots, s_n) = \mathbb{E}\left(\prod_{i=1}^n s_i^{X_i}\right).$$

Exemple 5.2. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ un vecteur aléatoire de densité

$$f_{\mathbb{X}}(x_1, x_2, x_3) = C(x_2 - x_3)e^{-x_1} \mathbf{1}_{\{0 < x_3 < x_2 < x_1\}}.$$

1. Déterminer C

2. Déterminer la densité du vecteur $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$

3. Déterminer les lois marginales

4. Calculer $\mathbb{E}(\mathbb{X})$ et $\Sigma_{\mathbb{X}}$.

5. Indépendance et conditionnement

Définition 5.10 (Indépendance). Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire. Les composantes X_1, X_2, \dots, X_n sont indépendantes si la loi conjointe $\mathbb{P}_{\mathbb{X}}$ est le produit des lois marginales $\mathbb{P}_{X_1}, \mathbb{P}_{X_2}, \dots, \mathbb{P}_{X_n}$ i.e. $\mathbb{P}_{\mathbb{X}} = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2} \otimes \dots \otimes \mathbb{P}_{X_n}$.

Autrement dit $\mathbb{P}(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \cdot \mathbb{P}(X_2 \in B_2) \cdots \mathbb{P}(X_n \in B_n)$

où $\forall i \in \{1, 2, \dots, n\}, B_i \in \mathcal{B}_{\mathbb{R}}$

Théorème 5.2. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire. Il y a équivalence entre

(i) X_1, X_2, \dots, X_n sont indépendantes

(ii) $F_{\mathbb{X}}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdots F_{X_n}(x_n)$

(iii) $f_{\mathbb{X}}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdots f_{X_n}(x_n)$

(iv) $\forall g_1, g_2, \dots, g_n$ mesurables positives, $g_1 \circ X_1, g_2 \circ X_2, \dots, g_n \circ X_n$ sont indépendantes

(v) $\varphi_{\mathbb{X}}(a_1, \dots, a_n) = \varphi_{X_1}(a_1) \cdot \varphi_{X_2}(a_2) \cdots \varphi_{X_n}(a_n)$.

Corollaire 5.1. Si X_1, X_2, \dots, X_n sont indépendants alors

(1) $\mathbb{E}(X_1 \cdot X_2 \cdots X_n) = \mathbb{E}(X_1) \cdot \mathbb{E}(X_2) \cdots \mathbb{E}(X_n)$

$$(2) \Sigma_{\mathbb{X}} = \begin{pmatrix} \sigma_{11} & & 0 \\ & \ddots & \\ 0 & & \sigma_{nn} \end{pmatrix}$$

$$(3) V(a^t \mathbb{X}) = a^t \Sigma_{\mathbb{X}} a, \quad \forall a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

$$(4) \varphi_{\sum X_i}(t) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) \cdots \varphi_{X_n}(t), \quad \forall t \in \mathbb{R}.$$

Proposition 5.2. Soit X_1, X_2, \dots, X_n des v.a.r indépendantes alors

$$(1) \text{ Si } X_i \rightsquigarrow \mathcal{P}(\lambda_i), \quad i = 1, \dots, n \text{ alors } \sum_{i=1}^n X_i \rightsquigarrow \mathcal{P}\left(\sum_{i=1}^n \lambda_i\right).$$

$$(2) \text{ Si } X_i \rightsquigarrow \gamma(a_i, \rho), \quad i = 1, \dots, n \text{ alors } \sum_{i=1}^n X_i \rightsquigarrow \gamma\left(\sum_{i=1}^n a_i, \rho\right).$$

$$(3) \text{ Si } X_i \rightsquigarrow \mathcal{N}(m_i, \sigma_i^2), \quad i = 1, \dots, n \text{ alors } \sum_{i=1}^n X_i \rightsquigarrow \mathcal{N}\left(\sum_{i=1}^n m_i, \sum_{i=1}^n \sigma_i^2\right).$$

Définition 5.11 (Loi conditionnelle). Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire.

Soit $\mathbb{Y} = \begin{pmatrix} X_{i_1} \\ \vdots \\ X_{i_p} \end{pmatrix}$ et $\mathbb{Z} = \begin{pmatrix} X_{i_{p+1}} \\ \vdots \\ X_{i_n} \end{pmatrix}$ deux groupes de composantes de \mathbb{X} .

On appelle loi conditionnelle de \mathbb{Y} sachant $\mathbb{Z} = z$, lorsqu'elle existe, la loi de probabilité définie par

$$\mathbb{P}_{\mathbb{Y}|\mathbb{Z}=z}(B) = \lim_{dz \rightarrow 0} \mathbb{P}(\mathbb{Y} \in B | z < \mathbb{Z} \leq z + dz).$$

On a aussi

$$\mathbb{P}_{\mathbb{Y}|\mathbb{Z}=z}(B) = \lim_{dz \rightarrow 0} \frac{\mathbb{P}(\mathbb{Y} \in B, \mathbb{Z} \in]z, z + dz])}{\mathbb{P}(\mathbb{Z} \in]z, z + dz])}.$$

Fonction de repartition conditionnelle

$$F_{\mathbb{Y}|\mathbb{Z}=z}(y) = \mathbb{P}_{\mathbb{Y}|\mathbb{Z}=z}(Y_1 \leq y_1, \dots, Y_p \leq y_p) = \lim_{dz \rightarrow 0} \frac{\mathbb{P}(Y_1 \in]-\infty, y_1], \dots, Y_p \in]-\infty, y_p], \mathbb{Z} \in]z, z + dz])}{\mathbb{P}(\mathbb{Z} \in]z, z + dz])}.$$

Donc

$$F_{\mathbb{Y}|\mathbb{Z}=z}(y) = \frac{\frac{\partial^{n-p}}{\partial z_1 \dots \partial z_{n-p}} F_{\left(\frac{\mathbb{Y}}{\mathbb{Z}}\right)}(y, z)}{f_{\mathbb{Z}}(z)}.$$

Densité conditionnelle

$$f_{\mathbb{Y}|\mathbb{Z}=z}(y) = \frac{f_{\left(\frac{\mathbb{Y}}{\mathbb{Z}}\right)}(y, z)}{f_{\mathbb{Z}}(z)} \quad (\mathbb{Y} \text{ continu}). \text{ Si } \mathbb{Y} \text{ est discret, } f_{\mathbb{Y}|\mathbb{Z}=z}(y) = \mathbb{P}(\mathbb{Y} = y | \mathbb{Z} = z).$$

Théorème 5.3 (Fondamental). Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur aléatoire. La densité conjointe est liée aux densités conditionnelles par

$$f_{\mathbb{X}}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdot f_{X_2|X_1=x_1}(x_2) \cdots f_{X_n|X_1=x_1, \dots, X_{n-1}=x_{n-1}}(x_n)$$

Théorème 5.4. Soit $\mathbb{X} = \begin{pmatrix} \mathbb{Y} \\ \mathbb{Z} \end{pmatrix}$. Alors \mathbb{Y} et \mathbb{Z} sont indépendants ssi $f_{\mathbb{Z}|\mathbb{Y}=y}(z) = f_{\mathbb{Z}}(z)$.

Définition 5.12 (Espérance conditionnelle).

Soit $\mathbb{X} = \begin{pmatrix} \mathbb{Y} \\ \mathbb{Z} \end{pmatrix}$ avec $\mathbb{Y} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$ et $\mathbb{Z} = \begin{pmatrix} X_{p+1} \\ \vdots \\ X_n \end{pmatrix}$.

• On appelle espérance conditionnelle de \mathbb{Y} sachant $\mathbb{Z} = z$ la loi moyenne

$$\mathbb{E}_{\mathbb{Z}=z}(\mathbb{Y}) = \mathbb{E}(\mathbb{Y}|\mathbb{Z} = z) = \int_{\mathbb{R}^p} y d\mathbb{P}_{\mathbb{Y}|\mathbb{Z}=z}(y) = \varphi(z) = \begin{pmatrix} \varphi_1(z) \\ \vdots \\ \varphi_p(z) \end{pmatrix}$$

avec $\varphi_i(z) = \mathbb{E}(Y_i|\mathbb{Z} = z)$

• On appelle espérance conditionnelle de \mathbb{Y} sachant \mathbb{Z} l'expression $\mathbb{E}(\mathbb{Y}|\mathbb{Z}) = \varphi(\mathbb{Z})$.

$\varphi(\mathbb{Z})$ est définie par $\varphi(\mathbb{Z})(\omega) = \varphi(z) = \mathbb{E}(\mathbb{Y}|\mathbb{Z} = z)$.

Remarque.

Si $B \in \mathcal{B}^p$, $\mathbb{Y} = \mathbf{1}_B$ et $\mathbb{Z} = \begin{pmatrix} Z_1 \\ \vdots \\ Z_{n-p} \end{pmatrix}$, alors

(1) Si \mathbb{Z} est continu, on a

$$\mathbb{E}(\mathbb{Y}|\mathbb{Z} = z) = \mathbb{P}(B|\mathbb{Z} = z) = \lim_{dz \rightarrow 0} \mathbb{P}(B|z < \mathbb{Z} \leq z + dz) = \lim_{dz \rightarrow 0} \frac{\mathbb{P}(B \cap [z < \mathbb{Z} \leq z + dz])}{\mathbb{P}(\mathbb{Z} \in]z, z + dz])}$$

(2) Si \mathbb{Z} est discret, alors

$$\mathbb{E}(\mathbb{Y}|\mathbb{Z} = z) = \mathbb{P}(B|\mathbb{Z} = z) = \frac{\mathbb{P}(B \cap [\mathbb{Z} = z])}{\mathbb{P}(\mathbb{Z} = z)}$$

Exemple 5.3. Soit X_1, \dots, X_n n variables aléatoires indépendantes de même loi $\mathcal{P}(\lambda)$. On pose $Y = \mathbf{1}_{[X_1=x_1]}$ et $Z = \sum_{i=1}^n X_i$.

1. Calculer $\mathbb{E}(Y|Z = x)$.

2. Calculer $\mathbb{E}(X_1|Z)$.

Théorème 5.5. Soit $\mathbb{X} = \begin{pmatrix} Y \\ Z \end{pmatrix}$. Alors pour $g : \mathbb{R}^p \rightarrow \mathbb{R}$, on a

(1) $\mathbb{E}(g(\mathbb{Y})) = \int_{\mathbb{R}^p} g(y) d\mathbb{P}_{\mathbb{Y}}(y) = \mathbb{E}(\mathbb{E}_Z(g(\mathbb{Y}))) = \mathbb{E}(\mathbb{E}(g(\mathbb{Y})|Z))$

(2) $\mathbb{E}(AY|Z) = A\mathbb{E}(Y|Z)$

(3) $\mathbb{E}(g(Y)|Y) = g(Y)$, \mathbb{P}_Y - p.s.

(4) $Var(Y) = \mathbb{E}(Var(Y|Z)) + Var(\mathbb{E}(Y|Z))$

Exercice 5.1. Soit $\begin{pmatrix} Y \\ Z \end{pmatrix}$ un vecteur aléatoire tel que $\mathbb{P}_{Y|Z=z} = \mathcal{P}(z)$ et $Z \rightsquigarrow \gamma(1, 1)$.

1. Déterminer la densité (conjointe) de $\begin{pmatrix} Y \\ Z \end{pmatrix}$.

2. Calculer $\mathbb{E}(Y|Z)$, $\mathbb{E}(Z|Y)$ et $\mathbb{E}(Z^{-Y})$.

6. Vecteurs aléatoires Gaussiens

6.1. Généralités

Définition 5.13. $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ est un vecteur gaussien (vecteur multinormal, ou variable

multinormale) si toute combinaison linéaire de ses composantes $a^t\mathbb{X} = \sum_{i=1}^n a_i X_i$ est une

variable aléatoire normale. On note $\mathbb{X} \rightsquigarrow \mathcal{N}(m, \Sigma)$ pour dire que $\mathbb{E}(\mathbb{X}) = m = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}$ et

$Var(\mathbb{X}) = \Sigma$ i.e. $Y = a^t\mathbb{X} \rightsquigarrow \mathcal{N}(a^tm, a^t\Sigma a)$

Exemple : Si X_1, \dots, X_n sont indépendantes et gaussiennes, alors $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ est un vecteur gaussien.

Proposition 5.3. Si $\mathbb{X} \rightsquigarrow \mathcal{N}(m, \Sigma)$ alors chaque composante X_i suit une loi normale

$\mathcal{N}(m_i, \sigma_{ii})$ avec $m = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}$ et $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq n}$.

Théorème 5.6. Si $\mathbb{X} \rightsquigarrow \mathcal{N}(m, \Sigma)$ alors $\varphi_{\mathbb{X}}(a) = e^{ia^tm - \frac{1}{2}a^t\Sigma a}$

Théorème 5.7. Soit $\mathbb{X} \rightsquigarrow \mathcal{N}(m, \Sigma)$.

Σ est une matrice diagonale ssi les v.a. X_1, X_2, \dots, X_n sont indépendantes.

Théorème 5.8 (Fondamental).

(i) $\forall m \in \mathbb{R}^n$, $\int_{\mathbb{R}^n} e^{-\frac{1}{2}(x-m)^t A(x-m)} dx = (2\pi)^{\frac{n}{2}} (\det A)^{-\frac{1}{2}}$ si A est symétrique définie positive.

(ii) Si $\mathbb{X} \rightsquigarrow \mathcal{N}(m, \Sigma)$ avec Σ régulière, alors \mathbb{X} admet pour densité

$$f_{\mathbb{X}}(x) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(x-m)^t \Sigma^{-1}(x-m)}.$$

(iii) Si $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ a pour densité

$$f_{\mathbb{X}}(x) = K e^{-\frac{1}{2} \sum_{i,j} a_{ij} (x_i - m_i)(x_j - m_j)}$$

alors \mathbb{X} est un vecteur multinormal et $\mathbb{E}(\mathbb{X}) = m = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}$, $Var(\mathbb{X}) = \Sigma = (\sigma_{ij})$ avec

$$A = (a_{ij}) = \Sigma^{-1} \text{ et } K = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} = \sqrt{\frac{\det A}{(2\pi)^n}}$$

Théorème 5.9 (Loi conditionnelle).

Soit $\mathbb{X} \rightsquigarrow \mathcal{N}(m, \Sigma)$ non dégénérée partitionnée en deux v.a. \mathbb{Y} et \mathbb{Z} à p et $n - p$ composantes resp. Alors

(i) $\mathbb{P}_{\mathbb{Y}|\mathbb{Z}=z} = \mathcal{N}(m_{\mathbb{Y}} + \Sigma_{12} \Sigma_{22}^{-1}(z - m_{\mathbb{Z}}), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$ avec $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$

$Var(\mathbb{Y}) = \Sigma_{11}$, $Var(\mathbb{Z}) = \Sigma_{22}$, $Cov(\mathbb{Y}, \mathbb{Z}) = \Sigma_{12}$, $\Sigma_{21} = \Sigma_{12}^t$ et $\mathbb{E}(\mathbb{Y}) = m_{\mathbb{Y}}$.

(ii) $\mathbb{E}(\mathbb{Y}|\mathbb{Z}) = \mathbb{E}(\mathbb{Y}) + \Sigma_{12} \Sigma_{22}^{-1}(\mathbb{Z} - \mathbb{E}(\mathbb{Z}))$

6.2. Lois dérivées d'un vecteur Gaussien

Définition 5.14 (Loi de Khi-deux).

Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \rightsquigarrow \mathcal{N}(0, I_n)$.

Alors $\|\mathbb{X}\|^2 = \mathbb{X}^t \mathbb{X} = \sum_{i=1}^n X_i^2$ suit une loi de Khi-deux à n degrés de liberté notée $\chi^2(n)$.

Proposition 5.4. Soit $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \rightsquigarrow \mathcal{N}(0, \Sigma)$. Si Σ est inversible alors $\mathbb{X}^t \Sigma^{-1} \mathbb{X} \rightsquigarrow \chi^2(n)$.

Remarque.

Si $X_i \rightsquigarrow \mathcal{N}(0, 1)$ alors $X_i^2 \rightsquigarrow \gamma(\frac{1}{2}, \frac{1}{2})$. Donc $\chi^2(n) = \gamma(\frac{n}{2}, \frac{1}{2})$.

Si $Y_1 \rightsquigarrow \chi^2(n_1)$ et $Y_2 \rightsquigarrow \chi^2(n_2)$ et si Y_1 et Y_2 sont indépendantes alors $Y_1 + Y_2 \rightsquigarrow \chi^2(n_1 + n_2)$.

Définition 5.15 (Loi de Student).

Soit $X \rightsquigarrow \mathcal{N}(0, 1)$ et $Y \rightsquigarrow \chi^2(n)$. On suppose que X et Y sont indépendantes.

Alors $T = \frac{X}{\sqrt{\frac{Y}{n}}}$ suit une loi de student à n degrés de liberté. On note $T \rightsquigarrow T(n)$.

Définition 5.16 (Loi de Fisher-Snédecor).

Soit $Y_1 \rightsquigarrow \chi^2(n_1)$ et $Y_2 \rightsquigarrow \chi^2(n_2)$. On suppose que Y_1 et Y_2 sont indépendantes.

Alors $F = \frac{\frac{Y_1}{n_1}}{\frac{Y_2}{n_2}} = \frac{Y_1}{Y_2} \cdot \frac{n_2}{n_1}$ suit une loi de Fisher-Snédecor à n_1 et n_2 degrés de liberté.

On note $F \rightsquigarrow F(n_1, n_2)$.

Chapitre 6

Convergences

1. Convergence en loi

Définition 6.1. Soit $(X_n)_{n \geq 0}$ une suite de v.a et X une v.a toutes définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, de fonctions de répartitions $(F_n)_{n \geq 0}$ et F . On dit que (X_n) converge en loi vers X si $(F_n(x))_{n \geq 0}$ converge vers $F(x)$ en tout point x de continuité de F . On note $X_n \xrightarrow{\mathcal{L}} X$

Exemple 6.1.

(1) $X_n \rightsquigarrow \mathcal{N}(0, \frac{1}{n}) \implies X_n \xrightarrow{\mathcal{L}} X = 0$.

(2) $X_n \rightsquigarrow \mathcal{U}_{[a-\frac{1}{n}, b+\frac{1}{n}]} \implies X_n \xrightarrow{\mathcal{L}} X \rightsquigarrow \mathcal{U}_{[a,b]}$.

Proposition 6.1. Soit $(X_n)_{n \geq 0}$ une suite de v.a et X une v.a toutes définies sur $(\Omega, \mathcal{F}, \mathbb{P})$, de fonctions caractéristiques $(\varphi_{X_n})_{n \geq 0}$ et φ_X .

a) Alors $X_n \xrightarrow{\mathcal{L}} X \iff \forall t \in \mathbb{R}, \varphi_{X_n}(t) \longrightarrow \varphi_X(t)$

b) $X_n \xrightarrow{\mathcal{L}} X \iff \mathbb{E}(h(X_n)) \longrightarrow \mathbb{E}(h(X))$, pour toute fonction h bornée et continue par morceaux.

La proposition suivante assure que la famille des lois gaussiennes est stable pour la convergence en loi.

Proposition 6.2. Soit $(X_n)_{n \geq 0}$ une suite de v.a et X une v.a gaussiennes de loi $\mathcal{N}(m_n, \sigma_n^2)$.

La suite converge en loi si et seulement si $m_n \xrightarrow{n \rightarrow +\infty} m \in \mathbb{R}$ et $\sigma_n \xrightarrow{n \rightarrow +\infty} \sigma \in [0, +\infty[$.

Et la loi limite est la loi gaussienne $\mathcal{N}(m, \sigma^2)$.

2. Convergence en probabilité

Définition 6.2. Soit $(X_n)_{n \geq 0}$ une suite de v.a et X une v.a toutes définies sur $(\Omega, \mathcal{F}, \mathbb{P})$.

On dit que (X_n) converge en probabilité vers X si $\forall \varepsilon > 0, \mathbb{P}(|X_n - X| > \varepsilon) \longrightarrow 0$.

On note $X_n \xrightarrow{\mathbb{P}} X$

Proposition 6.3. Soit $(X_n)_{n \geq 0}$ une suite de v.a et X une v.a toutes définies sur $(\Omega, \mathcal{F}, \mathbb{P})$.

a) Alors $X_n \xrightarrow{\mathbb{P}} X \implies X_n \xrightarrow{\mathcal{L}} X$

b) $X_n \xrightarrow{\mathcal{L}} a \in \mathbb{R} \implies X_n \xrightarrow{\mathbb{P}} a$.

3. Convergence presque sure

Définition 6.3. On dit que $(X_n)_{n \geq 0}$ converge presque surement vers X si

$\mathbb{P}(\{\omega : X_n(\omega) \rightarrow X(\omega)\}) = 1$. Ce qui équivaut à $\mathbb{P}(\{\omega : X_n(\omega) \not\rightarrow X(\omega)\}) = 0$.

On note $X_n \xrightarrow{p.s.} X$

Proposition 6.4.

$X_n \xrightarrow{p.s.} X \implies X_n \xrightarrow{\mathbb{P}} X$.

4. Théorèmes limites

Théorème 6.1 (Loi faible des grands nombres). Soit $(X_n)_{n \geq 0}$ une suite de v.a telle

que $\mathbb{E}(X_i) = m$ et $Cov(X_i, X_j) = \delta_{ij}\sigma^2$. Alors $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} m$

Théorème 6.2 (Loi forte des grands nombres). Soit $(X_n)_{n \geq 0}$ une suite de v.a indé-

pendantes de même loi avec $\mathbb{E}(X_i) = m$, alors $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p.s.} m$

Théorème 6.3 (Limite centrale). Soit $(X_n)_{n \geq 0}$ une suite de v.a indépendantes de même

loi avec $\mathbb{E}(X_i) = m$ et $Var(X_i) = \sigma^2$. Alors

$\sqrt{n}(\bar{X}_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$.

$\sqrt{n} \left(\frac{\bar{X}_n - m}{\sigma} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$, où $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

Théorème 6.4. Soit $(X_n)_{n \geq 0}$ une suite de v.a indépendantes telle que

$\sqrt{n}(X_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$.

Si $g : \mathbb{R} \rightarrow \mathbb{R}$ est dérivable, alors $\sqrt{n}(g(X_n) - g(m)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, [g'(m)]^2 \sigma^2)$

4.1. Forme multidimensionnelle

Théorème 6.5 (Limite centrale pour les vecteurs aléatoires). Soit $(\mathbb{X}_n)_{n \geq 0}$ une suite de vecteur aléatoires indépendants de même loi avec $\mathbb{E}(\mathbb{X}_i) = m$ et $Var(\mathbb{X}_i) = \Sigma$. Alors

$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \mathbb{X}_i - m \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma)$.

Théorème 6.6. Soit $(X_n)_{n \geq 0}$ une suite de vecteur aléatoires indépendants à valeurs dans \mathbb{R}^p telle que $\sqrt{n}(\mathbb{X}_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma)$. Si $g : \mathbb{R}^p \rightarrow \mathbb{R}^q$ est différentiable, alors

$$\sqrt{n}(g(X_n) - g(m)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, G^t \Sigma G) \text{ avec } G = J_g(m)^t = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(m) & \cdots & \frac{\partial g_q}{\partial x_1}(m) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_1}{\partial x_p}(m) & \cdots & \frac{\partial g_q}{\partial x_p}(m) \end{pmatrix}_{(p \times q)}$$