

COURS DE PROBABILITÉ

Licence de Mathématiques
Semestre 6

Kossi Essona GNEYOU ¹

30 avril 2020

1. Professeur à l'Université de Lomé

Sommaire

1	ESPACE PROBABILISÉ – ÉVÉNEMENTS INDÉPENDANTS.	5
1.1	Vocabulaire de probabilité	5
1.1.1	Langage des événements	5
1.1.2	Notion intuitive de probabilité	6
1.1.3	Probabilité - Espace probabilisé	6
1.2	Probabilité Conditionnelle	8
1.2.1	Événements indépendants	9
1.2.2	Indépendance d'une famille d'événements	10
1.2.3	Application : Loi de tout et de rien	11
2	VARIABLES ALÉATOIRES – LOIS DE PROBABILITÉ.	12
2.1	Variable aléatoire - Loïs de probabilité	12
2.2	Moments de variables aléatoires	13
2.3	Inégalités fondamentales	14
2.4	Fonction de répartition - Densité de Probabilité	15
2.5	Lois multivariées	16
2.6	Fonctions caractéristiques	18
2.7	Formule de changement de variables	21
2.8	Indépendance de variables aléatoires	23
2.9	Lois conditionnelles	26
2.9.1	Cas de variables aléatoires réelles discrètes	26
2.9.2	Cas de variables aléatoires réelles continues	26
2.9.3	Espérance conditionnelle	28
3	LOIS DE PROBABILITÉ USUELLES.	31
3.1	Lois discrètes courantes	31
3.1.1	Loi de Bernoulli de paramètre $p : \mathcal{B}(1, p)$	31
3.1.2	Loi binomiale de paramètres n et $p : \mathcal{B}(n, p)$	31
3.1.3	Loi multinomiale de paramètres $n, p_1, \dots, p_r : \mathcal{M}(n, p_1, \dots, p_r)$	31
3.1.4	Loi de Poisson de paramètre $\lambda : \mathcal{P}(\lambda)$	32
3.1.5	Loi de Pascal ou loi géométrique de paramètre $p : \mathcal{G}(p)$	32
3.1.6	Loi binomiale négative de paramètres n et $p : \mathcal{BN}(n, p)$	32
3.1.7	Loi Hypergéométrique de paramètres $N, M, n : \mathcal{H}(N, M, n)$	32
3.2	Lois continues courantes	33
3.2.1	Loi uniforme sur $[a, b] : \mathcal{U}[a, b]$	33
3.2.2	Loi gamma de paramètres p et $\theta : \gamma(p, \theta)$	33
3.2.3	Loi beta de paramètres p et $q : b(p, q)$	33
3.2.4	Loi normale de moyenne m et d'écart-type $\sigma : \mathcal{N}(m, \sigma)$	33
3.2.5	Loi de Cauchy de paramètres $a, b : \text{Cauchy}(a, b)$	34
3.3	Vecteurs Gaussiens	35
3.3.1	Représentation de lois normales sur $\mathbb{R}^n$	36
3.3.2	Singularité et absolue continuité des lois gaussiennes	36

4	CONVERGENCE EN LOI – THÉORÈME CENTRAL LIMITE	39
4.1	Convergence des mesures positives bornée	39
4.2	Le Théorème Centrale Limite	45
5	CONVERGENCES PRESQUE SÛRE DES SUITES ET SÉRIES DE V.A., LOI DES GRANDS NOMBRES	48
5.1	Les quatre types de convergences	48
5.2	Convergence <i>p.s.</i> de séries de v.a. indépendantes – Théorèmes limites forts	53
5.3	Loi forte des grands nombres	55
6	introduction aux processus aléatoires	58
6.1	Processus aléatoires courants	58
6.1.1	Processus de branchement	58
6.1.2	Processus de Poisson	59
6.1.3	Marche aléatoire sur $\mathbb{Z}$	61
6.2	Chaînes de Markov	62
6.2.1	Ergodicité	64
6.2.2	Chaîne de Markov continues	66
7	Exercices d'application	68
7.0.1	Exercices du chapitre 1	68
7.0.2	Exercices du chapitre 2	70
7.0.3	Exercices des chapitres 2, 3	72
7.0.4	Exercices des chapitres 4, 5, 6	74
7.0.5	Exercices, problèmes, devoirs et examens divers	76

Introduction

Le calcul des probabilités semble avoir son origine dans les jeux de hasard. Il fallait savoir décider des mises et des gains pour que les jeux soient équitables ou non. Aujourd'hui, il a investi de très nombreuses activités humaines et son utilisation est nécessaire dans la plupart des activités industrielles. Il est utile par exemple de savoir si la construction d'un appareil est raisonnable, si il a une chance de fonctionner ou si l'assemblage d'un très grand nombre de composants réduira son espérance de vie de manière inacceptable (navette spatiale par exemple). Le calcul des probabilités et les statistiques sont utilisées pour augmenter, souvent de manière importante, la qualité des produits manufacturés, les coûts de production s'en trouvent réduits car le risque de panne est diminué. La transmission de l'information est devenue extrêmement importante et se développe sans cesse. Pour que cette transmission soit utile, il est nécessaire qu'elle soit sans erreur, il faut donc détecter et corriger les erreurs sans en connaître les causes. C'est pourquoi l'information est transmise de manière redondante (bit de parité etc.), ce qui permet d'effectuer les corrections nécessaires. Un exemple extrême de ces techniques est celui des satellites envoyés à plusieurs milliards de kilomètres de la terre. La distance très importante affaiblit les signaux envoyés par ces engins qui nous arrivent extrêmement faibles et fortement bruités. Il est possible, malgré tout, de reconstruire l'information originale et les images reconstruites sont de bonne qualité. Un autre domaine du calcul des probabilités et des statistiques est celui de la compression des données. Ici, il faut utiliser les structures particulières des données à stocker pour en minimiser la taille. Mais ces structures ne sont pas parfaitement établies, elles varient statistiquement, c'est ce qui rend difficile le travail de compression. Les travaux de Shannon et de Kolmogorov sur la théorie de l'information ont ouvert de nouveaux domaines de recherche en probabilité et statistique.

Ce cours de probabilités s'adresse aux étudiants de Licence de mathématiques 3ème année ou semestres 5 et 6. Il est issu de ceux que j'ai faits en Maîtrise de Mathématiques de 1994 à 2006 puis en troisième année de Licence depuis 2007 à l'université de Lomé (UL) et depuis 2010 à l'université de Kara (UK). A l'UK, il est présenté en deux cours de 36h chacun, plus détaillés que celui présenté à l'université de Lomé en 48h.

On suppose de la part des étudiants concernés, une maîtrise des techniques d'analyse mathématiques telle qu'elle est enseignée dans les quatre premiers semestres du système LMD à l'UL et à l'UK. Il est plus aisé de s'initier à la théorie des probabilités dans le contexte de la théorie de la mesure. C'est pourquoi ce cours est généralement programmé avant celui de la théorie de l'intégration.

Le cours comporte 6 chapitres, les 3 premiers constituant le cours de probabilité II et les trois derniers, ceux constituant l'essentiel du cours de probabilité III à l'UK. Des exercices de TD ainsi que des devoirs et examens sont aussi présentés à la fin du cours. Les théorèmes, propositions lemmes et définitions ont une numérotation dépendant du chapitre et du paragraphe dans lequel ils apparaissent.

Il y a aujourd'hui beaucoup de traités de probabilités en librairies ou téléchargeables sur

internet au moins en français ou en anglais. Les références utilisées dans ce cours sont les livres de Métivier, Bouleau, Foata et Fuchs, Cottrelle, Duhamel et Genon-Catalot, Dacunha-Castelle et Duflo, Neveu, Rényi, Ross, Feller, Grimmett et Stirzaker, Tucker, etc.

Chapitre 1

ESPACE PROBABILISÉ – ÉVÉNEMENTS INDÉPENDANTS.

1.1 Vocabulaire de probabilité

La théorie de la probabilité a un langage qui lui est propre et les notations sont celles de *mesure et intégration* dont les connaissances sont indispensables pour suivre avec intérêt les cours appelés Probabilité 2 et Probabilité 3.

Dans ce qui suit, une *expérience aléatoire* (e.a.) est une expérience dont les résultats dépendent du hasard. L'ensemble des résultats possibles d'une e.a. est appelé *univers* et noté Ω . On appelle :

- *Épreuve*, tout essai d'une e.a.

- *évènement*, toute partie (ou sous-ensemble) A de Ω , c'est-à-dire une collection d'éléments de Ω dont la réalisation décrit un fait possible de l'expérience aléatoire. Un évènement réduit à un singleton $\{\omega\}$, $\omega \in \Omega$, est appelé *évènement élémentaire* ou *éventualité*.

La classe de tous les évènements possibles liés à une e.a. est notée $\mathcal{A}$. On impose à $\mathcal{A}$ d'être contenu dans l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω et d'être une *tribu*, c-à-d. un ensemble contenant Ω , stable par complémentation et par réunion dénombrable. Le lecteur est renvoyé au cours de mesure et intégration pour un rappel sur les notions de tribu, tribu engendrée par une classe de parties, tribu borélienne, etc.

On appelle espace probabilisable, tout couple $(\Omega, \mathcal{A})$ lié à une e.a. où Ω est l'ensemble des résultats possibles et $\mathcal{A}$ est une tribu sur Ω dite tribu des évènements de l'expérience aléatoire.

1.1.1 Langage des évènements

On dit qu'un évènement A lié à une expérience aléatoire $\mathcal{E}$ est réalisé si le résultat ω obtenu est un élément de A . Il vient :

- . Ω est l'évènement certain (car il est toujours réalisé au cours d'une e.a.)
- . $\emptyset$ est l'évènement impossible (car $\emptyset$ n'est jamais réalisé).
- . À chaque évènement A correspond l'évènement contraire de A noté "non A " ou A^c ou encore $\bar{A}$.

Soient A et B deux évènements liés à une même e.a. Lorsqu'ils se produisent :

- l'évènement " A implique B " ($A \Rightarrow B$) est noté $A \subset B$;
- l'évènement " A et B " se note $A \cap B$. Si $A \cap B = \emptyset$, les évènements A et B sont dits incompatibles ;
- l'évènement " A ou B " se note $A \cup B$;
- l'évènement " A et non B " se note $A \setminus B$ ou $A \cap \bar{B}$;
- l'évènement " A ou B mais pas les deux" se note $A \Delta B$. Rappelons que $A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$.

Si (A_n) est une suite d'évènements :

- $\cup A_n$ est l'évènement " A ou B " ;
- $\cap A_n$ est l'évènement " A et B " ;
- $\limsup A_n$ ou $\overline{\lim} A_n$ est l'évènement " A se produit une infinité de fois" ;
- $\liminf A_n$ ou $\underline{\lim} A_n$ est l'évènement " A se produit sauf peut-être un nombre fini d'entre eux".

Les évènements $\overline{\lim} A_n$ et $\underline{\lim} A_n$ sont définis par :

$$\overline{\lim} A_n = \bigcap_{n \geq 1} \bigcup_{k \geq n} A_k \text{ et } \underline{\lim} A_n = \bigcup_{n \geq 1} \bigcap_{k \geq n} A_k.$$

On a : $\underline{\lim} A_n \subset \overline{\lim} A_n$; $(\underline{\lim} A_n)^c = \overline{\lim} A_n^c$ et $(\overline{\lim} A_n)^c = \underline{\lim} A_n^c$.

1.1.2 Notion intuitive de probabilité

Soit $\mathcal{E}$ une e.a. Tous les évènements liés à $\mathcal{E}$ n'ont pas la même "chance" de se produire. Soit A un évènement lié à $\mathcal{E}$ c-à-d. $A \in \mathcal{A}$. Répétons n fois l'expérience $\mathcal{E}$ dans les mêmes conditions. Soit r le nombre de fois que A est réalisé au cours des n répétitions. Alors le nombre $f(A) = \frac{r}{n}$ est appelé fréquence d'apparition de A et on a $f(A) \in [0, 1]$.

Si B est un autre évènement tels que A et B soient incompatibles et si B s'est produit s fois au cours des n épreuves, $A \cup B$ s'est produit $r + s$ fois (puisque A et B ne se produisent jamais simultanément). On a donc

$$f(A \cup B) = \frac{r+s}{n} = \frac{r}{n} + \frac{s}{n} = f(A) + f(B).$$

L'évènement certain Ω s'est produit n fois (puisque'il se produit à chaque essai de $\mathcal{E}$). On a donc $f(\Omega) = \frac{n}{n} = 1$. On a également

$$f(\bar{A}) = \frac{n-r}{n} = 1 - \frac{r}{n} = 1 - f(A)$$

et $f(\emptyset) = 0 = 1 - f(\Omega)$. Intuitivement, lorsque n tend vers $+\infty$, $f(A) = \frac{r}{n}$ admet une limite qui ne dépend pas de l'évènement A . Cette limite est appelée probabilité de l'évènement A et est notée $P(A)$.

1.1.3 Probabilité - Espace probabilisé

Définition 1.1.1

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle probabilité sur $(\Omega, \mathcal{A})$, toute application P de $\mathcal{A}$ dans $[0, 1]$ possédant les propriétés suivantes :

- 1) $P(\Omega) = 1$
- 2) $\forall A, B \in \mathcal{A}, A \cap B = \emptyset, P(A \cup B) = P(A) + P(B)$.
- 3) Si $A_1, A_2, \dots$ est une suite d'événements deux à deux incompatibles,

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

On dit que P est une mesure positive σ -finie, de masse totale 1.

On appelle espace probabilisé, tout triplet $(\Omega, \mathcal{A}, P)$ où $(\Omega, \mathcal{A})$ est un espace probabilisable et P est une probabilité sur $(\Omega, \mathcal{A})$.

Propriété 1.1.1

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, A et B deux événements (i.e. $A, B \in \mathcal{A}$). On a

- 1) $P(\emptyset) = 0$; 2) $P(\bar{A}) = 1 - P(A)$; 3) Si $A \subset B, P(A) \leq P(B)$;
- 4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$; 5) Si (A_n) est une suite dans $\mathcal{A}, P(\bigcup_{n \geq 1} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.

Preuve :

- 1) évident. 2) On écrit $\Omega = A \cup \bar{A}$ donc $1 = P(A) + P(\bar{A})$ (car $A \cap \bar{A} = \emptyset$).
- 3) On écrit $B = (B \setminus A) \cup A$ d'où $P(B) = P(B \setminus A) + P(A) \geq P(A)$.
- 4) On écrit $A = (A \setminus B) \cup (A \cap B), B = (B \setminus A) \cup (A \cap B)$ d'où

$$\begin{aligned} P(A) + P(B) &= P(A \setminus B) + P(A \cap B) + P(B \setminus A) + P(A \cap B) \\ &= P((A \setminus B) \cup (B \setminus A) \cup (A \cap B)) + P(A \cap B) \\ &= P(A \cup B) + P(A \cap B) \end{aligned}$$
- 5) On écrit $\bigcup_{n \geq 1} A_n = \bigcup_{n \geq 1} B_n$ avec $B_1 = A_1, B_2 = A_2 \setminus A_1, \dots, B_n = A_n \setminus (\bigcup_{k=1}^{n-1} A_k)$ d'où $P(\bigcup_{n \geq 1} A_n) = P(\bigcup_{n \geq 1} B_n) = \sum_{n=1}^{\infty} P(B_n) \leq \sum_{n=1}^{\infty} P(A_n)$. $\square$

Remarques 1.1.1

- 1) Lorsque Ω est fini, toutes les éventualités ont la même probabilité de se produire. On dit qu'elles sont équiprobables. Dans ce cas $\mathcal{A} = \mathcal{P}(\Omega)$ et $\forall A \in \mathcal{P}(\Omega)$,

$$P(A) = \frac{\text{Card } A}{\text{Card } \Omega} = \frac{\text{nbre de cas favorables à } A}{\text{nbre de cas possibles}}$$

et on retrouve la formule classique rencontrée au moins en terminale.

- 2) Si A est une partie propre de Ω c-à-d. $A \neq \emptyset, A \subset \Omega, A \neq \Omega$, et si $\mathcal{A}$ est la tribu engendré par $A : \mathcal{A} = \sigma(\{A\}) = \{\emptyset, A, \bar{A}, \Omega\}$, alors toute probabilité sur $(\Omega, \mathcal{A})$ est de la forme $P(\emptyset) = 0, P(A) = p, P(\bar{A}) = 1 - p$ et $P(\Omega) = 1$ avec $0 \leq p \leq 1$.
- 3) Si on a une probabilité sur $(\Omega, \mathcal{A})$, on peut aussi définir d'autres probabilités sur $(\Omega, \mathcal{A})$ appelées probabilités conditionnelles.

Exercice 1.1.1

Soit $\mathcal{D}$ une classe de Ω . $\mathcal{D}$ est dite σ -additive ou est un système de Dynkins si elle satisfait :

- 1) $\Omega \in \mathcal{D}$, 2) A et $B \in \mathcal{D}$ et $A \subset B$, alors $B \setminus A \in \mathcal{D}$, 3) Si (A_n) est une suite dans $\mathcal{D}$ telle que $A_i \cap A_j = \emptyset \forall i \neq j$ alors $\bigcup A_n \in \mathcal{D}$.

Une classe $\mathcal{M}$ des parties de Ω est dite monotone si pour toute suite monotone (A_n) dans $\mathcal{M}$, $\lim A_n \in \mathcal{M}$ où $\lim A_n = \bigcup A_n$ si (A_n) est croissante et $\lim A_n = \bigcap A_n$ si (A_n) est décroissante.

Si $\mathcal{C}$ est une classe de parties de Ω , on note par $\mathcal{D}(\mathcal{C})$ (respectivement $\mathcal{M}(\mathcal{C})$) la classe σ -additive (respectivement monotone) engendrée par $\mathcal{C}$.

Montrer que :

- a) Une tribu est une classe σ -additive.
- b) Pour qu'une classe σ -additive soit une tribu, il faut et il suffit qu'elle soit stable par intersection finie.
- c) Si C est une classe de parties de Ω stable par intersection finie, alors
 - c₁) $\sigma(C) \subset \mathcal{D}(C)$; c₂) $\sigma(C) \subset \mathcal{M}(C)$.
 La propriété c) à démontrer est appelée Théorème de la Classe Monotone (TCM).

1.2 Probabilité Conditionnelle

Définition 1.2.1

Soit $(\Omega, \mathcal{A}, P)$ un e.p. , $B \in \mathcal{A}$ tel que $P(B) > 0$ i.e. B est un évènement de probabilité non nulle, et $A \in \mathcal{A}$. On appelle probabilité conditionnelle de A sachant B , la quantité notée $P(A | B)$ donnée par

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

Exercice 1.2.1

Montrer que la fonction d'ensemble sur $\mathcal{A}$, $A \mapsto P(A | B)$ est une probabilité.

Remarque 1.2.1

1) Si $A \cap B = \emptyset$, $P(A \cap B) = 0$; 2) Si $B \subset A$, $P(A|B) = 1$.

La probabilité conditionnelle sachant B : $P^B(\cdot)$ est donc portée par B .

Les deux propriétés suivantes sont des plus importantes propriétés de la probabilité conditionnelle :

Propriété 1.2.1 (Règle de la multiplication)

Pour tous n évènements $A_1, A_2, \dots, A_n$ telle que $P(A_1 A_2 \dots A_n) > 0$, on a

$$P(A_1 A_2 \dots A_n) = P(A_1) \cdot P(A_2 | A_1) P(A_3 | A_1 A_2) \dots P(A_n | A_1 A_2 \dots A_{n-1})$$

où AB signifie $A \cap B$ pour des évènements A et B .

Preuve :

Par récurrence immédiate sur n . e.g., pour $n = 3$, on a si $P(A_1 A_2 A_3) > 0$

$$P(A_1 A_2 A_3) = P(A_3 | A_1 A_2) P(A_1 A_2) = P(A_3 | A_1 A_2) P(A_2 | A_1) P(A_1)$$

puisque $P(A_1) \geq P(A_1 A_2) \geq P(A_1 A_2 A_3) > 0$. □

Propriété 1.2.2 (Théorème de la partition ou règle des probabilités totales)

Soit $(C_n)_{n \geq 1}$ une partition de Ω ou encore un système complet d'évènements tous de probabilité non nulle : $C_i \cap C_j = \emptyset$, $\forall i \neq j$, $\Omega = \bigcup_{n \geq 1} C_n$ et $\forall n \geq 1$, $P(C_n) > 0$. Alors pour tout évènement $A \in \mathcal{A}$, on a

$$P(A) = \sum_{n \geq 1} P(A | C_n) P(C_n).$$

Si de plus $P(A) > 0$ alors (règle de **Bayes**) : $\forall k \geq 1$,

$$P(C_k | A) = \frac{P(A | C_k) P(C_k)}{\sum_{n \geq 1} P(A | C_n) P(C_n)}.$$

1.2.1 Événements indépendants

Définition 1.2.2

Soient A et B deux événements ($A, B \in \mathcal{A}$). On dit que A et B sont indépendants et on note $A \perp B$ si $P(A \cap B) = P(A)P(B)$.

On voit immédiatement que si A et B sont indépendants, il en est de même pour $\bar{A}$ et B , A et $\bar{B}$, $\bar{A}$ et $\bar{B}$.

Remarque 1.2.2

Si A et B sont de probabilités non nulles, $A \perp B$ si, et seulement, $P(A | B) = P(A)$ et $P(B | A) = P(B)$.

Définition 1.2.3

Soit (A_n) une suite d'événements. On dit que les A_i sont indépendants (mutuellement) si et ssi pour toute suite finie $A_{i_1}, A_{i_2}, \dots, A_{i_p}$ d'entre eux, $2 \leq p \leq n$, on a

$$P(A_{i_1}A_{i_2} \cdots A_{i_p}) = \prod_{k=1}^p P(A_{i_k}) \quad (1.1)$$

On remarque que dans le cas d'une suite finie de n événements, il y a $C_n^2 + C_n^3 + \cdots + C_n^n = 2^n - n - 1$ conditions (1.1) à vérifier.

Lemme 1.2.1 (Borel-Cantelli)

Soit $(A_n)_{n \in \mathbb{N}^*}$ une suite d'événements. Deux cas sont possibles en ce qui concerne l'événement $\overline{\lim} A_n$:

- a) Si $\sum_{n \geq 1} P(A_n) < +\infty$ alors $P[\overline{\lim} A_n] = 0$. En d'autres termes, avec probabilité 1, au plus un nombre fini des A_n se produisent si la série précédente converge.
- b) Si les (A_n) sont indépendants et si $\sum_{n \geq 1} P(A_n) = +\infty$ alors $P[\overline{\lim} A_n] = 1$. En d'autres termes, avec probabilité 1, une infinité des A_n se produisent si $\sum_{n \geq 1} P(A_n) = +\infty$.

Preuve :

1) D'après l'inégalité de Boole,

$$P[\overline{\lim} A_n] = P\left[\bigcap_{n \geq 1} \bigcup_{k \geq n} A_k\right] \leq P\left[\bigcup_{k \geq n} A_k\right] \leq \sum_{k \geq n} P(A_k) \xrightarrow{n \rightarrow \infty} 0.$$

2) Comme $(\bigcup_{k \geq n} A_k)$ est décroissante, $\overline{\lim} A_n = \lim_{n \rightarrow \infty} \bigcup_{k \geq n} A_k$. Donc

$$P[\overline{\lim} A_n] = P\left[\lim_{n \rightarrow \infty} \bigcup_{k \geq n} A_k\right] = \lim_{n \rightarrow \infty} P\left[\bigcup_{k \geq n} A_k\right].$$

Soit

$$\begin{aligned} 1 - P[\overline{\lim} A_n] &= \lim_{n \rightarrow \infty} P\left[\bigcap_{k \geq n} \bar{A}_k\right] \stackrel{\text{ce}}{=} \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \prod_{k=n}^m P(\bar{A}_k) \\ &= \lim_{n \rightarrow \infty} \prod_{k=n}^{\infty} (1 - P(A_k)). \end{aligned}$$

Or $\forall k, \log(1 - P(A_k)) \leq -P(A_k)$ et $\sum_{n \geq 1} P(A_n) = +\infty$ donc $\sum_{k=n}^{\infty} \log(1 - P(A_k)) \xrightarrow{n \rightarrow \infty} 0$, d'où $1 - P[\overline{\lim} A_n] = 0$. $\square$

1.2.2 Indépendance d'une famille d'événements

Définition 1.2.4

Soit $(\Omega, \mathcal{A})$ un espace probabilisable et $(\mathcal{A}_i)_{i \in I}$ une collection de famille d'événements telles que $\forall i \in I, \mathcal{A}_i \in \mathcal{A}$.

On dit que cette collection est indépendante si pour toute partie finie J de I , pour tout $A_j \in \mathcal{A}_j, j \in J$: $P\left[\bigcap_{j \in J} A_j\right] = \prod_{j \in J} P(A_j)$.

Lemme 1.2.2

Soit $(C_i)_{i \in I}$ une collection de famille d'événements indépendantes. Alors les classes σ -additives $\mathcal{D}_i$ engendrées par les C_i sont encore indépendantes.

Preuve :

Supposons d'abord I fini, e.g. $I = \{1, 2, \dots, n\}$ et $C_1, C_2, \dots, C_n$ indépendantes. Posons

$$\mathcal{H}_1 = \left\{ A \in \mathcal{A} \mid \forall C_i \in \mathcal{C}_i, i \neq 1, P\left[A \cap \left(\bigcap_{i=2}^n C_i\right)\right] = P(A) \prod_{i=2}^n P(C_i) \right\}.$$

Alors $\mathcal{H}_1$ a les propriétés de classe monotone c'est-à-dire

- (i) $\Omega \in \mathcal{H}_1$
- (ii) Si $A, B \in \mathcal{H}_1, A \subset B$ alors $B \setminus A \in \mathcal{H}_1$ et
- (iii) Si (A_n) est une suite de $\mathcal{H}_1$, croissante alors $A = \lim_{n \rightarrow \infty} A_n = \bigcup_n A_n \in \mathcal{H}_1$ (les vérifications de (i) à (iii) sont faciles).

D'après le théorème de la classe monotone (TCM), $\mathcal{H}_1$ contient $\mathcal{D}_1$.

Donc $A_1 \in \mathcal{D}_1, C_i \in \mathcal{C}_i, i \neq 1$ et on a

$$P\left[A_1 \cap \left(\bigcap_{i=2}^n C_i\right)\right] = P(A_1) \prod_{i=2}^n P(C_i). \quad (1.2)$$

On répète le même argument $n - 1$ fois encore. Alors $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n$ ont la propriété (1.2). On raisonne ainsi pour chaque partie J de I qu'on peut ne plus supposer finie. Si $J \subset I, J$ finie,

$$\forall A_j \in \mathcal{D}_j, j \in J, P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

et les $\mathcal{D}_j$ sont indépendants.

□

Théorème 1.2.1

Si $(C_i)_{i \in I}$ est une collection de familles indépendantes et si chaque C_i est stable par intersection finie, alors les tribus engendrées par les C_i sont encore indépendantes.

Preuve :

Les classes σ -additives $\mathcal{D}_i = \mathcal{D}(C_i)$ engendrée par les C_i sont indépendantes d'après le Lemme 1.2.2 d'où le résultat puisque $\mathcal{D}_i = \mathcal{A}_i$ d'après le TCM. □

Corollaire 1.2.1

Soit $(\mathcal{A}_i)_{i \in I}$ est une famille de tribus indépendantes. Soit $(I_j)_{j \in J}$ une partition de I . Notons $\mathcal{A}_{I_j} = \bigvee_{i \in I_j} \mathcal{A}_i = \sigma\left(\bigcup_{i \in I_j} \mathcal{A}_i\right)$.

Alors les $\mathcal{A}_{I_j}, j \in J$ sont encore indépendantes (principe de regroupement).

Corollaire 1.2.2

Soit $(I, \leq)$ un ensemble préordonné filtrant à droite : $\forall i, j \in I, \exists l \in I \mid i \leq l \text{ et } j \leq l$ (e.g. $(\mathbb{N}, \leq), (\mathbb{R}, \leq)$).

Soit $(\mathcal{A}_i)_{i \in I}$ une famille croissante de sous-tribus de $\mathcal{A}$: $\forall i, j \in I, i \leq j \Rightarrow \mathcal{A}_i \subset \mathcal{A}_j \subset \mathcal{A}$ et soit $\mathcal{C}$ une sous-tribu de $\mathcal{A}$. Si $\forall i \in I, \mathcal{A}_i \perp \mathcal{C}$ alors $\lim \mathcal{A}_i = \bigvee_{i \in I} \mathcal{A}_i \perp \mathcal{C}$.

1.2.3 Application : Loi de tout et de rien

Soit $(\Omega, \mathcal{A}, P)$ un e.p. , $(\mathcal{A}_n)$ une suite indépendante de sous-tribus de $\mathcal{A}$. Posons $\mathcal{F}_n = \bigvee_{k \leq n} \mathcal{A}_k$ e.g. si $(\mathcal{E}_n)$ est une suite d'expériences aléatoires indépendantes et $\mathcal{A}_n$ est la tribu d'évènements relatifs à $\mathcal{E}_n$ alors $\mathcal{F}_n$ est la tribu de tous les évènements relatifs aux n premières expériences) et $\mathcal{G}_n = \bigvee_{k \geq n} \mathcal{A}_k$ est la tribu des évènements relatifs à toutes les expériences aléatoires de rang supérieure ou égale à n .

Alors $(\mathcal{F}_n)$ est croissante et $(\mathcal{G}_n)$ est décroissante vers $\bigcap_n \mathcal{G}_n$. La tribu $\mathcal{A} = \bigcap_n \mathcal{G}_n$ est appelée tribu asymptotique ou tribu de queue. Les évènements de $\mathcal{A}$ sont appelés évènements asymptotiques. Ce sont les évènements qui peuvent s'exprimer par le comportement à l'infini des $\mathcal{E}_n$. On a

Théorème 1.2.2 (Loi de tout et de rien ou loi 0 , 1 de Kolmogorov)

$$\forall A \in \mathcal{A}, P(A) = 0 \text{ ou } 1.$$

Preuve :

Il suffit de montrer que $\mathcal{A} \perp \mathcal{A}$.

D'après le principe de regroupement (Corollaire 1.2.1), $\mathcal{G}_{n+1} = \bigvee_{k \geq n+1} \mathcal{A}_k$ est indépendante de $\mathcal{F}_n = \bigvee_{k \leq n} \mathcal{A}_k$. Donc $\mathcal{A} \subset \mathcal{G}_{n+1} \implies \mathcal{A} \perp \mathcal{F}_n$ c'est-à-dire $\mathcal{A} \perp \bigvee_{k \leq n} \mathcal{A}_k$.

Or $(\mathcal{F}_n)$ converge vers $\bigvee_n \mathcal{F}_n = \sigma(\bigcup_n \mathcal{F}_n) = \sigma(\bigcup_n \mathcal{A}_n) = \bigvee_n \mathcal{A}_n$. Donc en passant à la limite, $\mathcal{A} \perp \bigvee_n \mathcal{A}_n = \mathcal{G}_1$. Comme $\mathcal{A} \subset \mathcal{G}_1$ alors $\mathcal{A} \perp \mathcal{A}$. Par suite, $\forall A \in \mathcal{A}, A \perp A$. D'où $P(A \cap A) = (P(A))^2 = P(A)$ soit $P(A) = 0$ ou $P(A) = 1$. $\square$

Chapitre 2

VARIABLES ALÉATOIRES – LOIS DE PROBABILITÉ.

Soit $\mathcal{E}$ une expérience aléatoire et $(\Omega, \mathcal{A}, P)$ l'espace probabilisé attaché à $\mathcal{E}$. Parfois, pour pouvoir observer un $\omega \in \Omega$ pris au hasard suivant la mesure de probabilité P , on a besoin qu'une fonction de ω puisse être observée. Plus précisément, on considère une "quantité" X pouvant prendre ses valeurs dans un ensemble $\mathcal{X}$ de façon aléatoire et lié à $\mathcal{E}$ c'est-à-dire, la valeur prise par X dépend de l'issue de $\mathcal{E}$. Notons $\mathcal{E}_X$ l'expérience aléatoire attachée à X , $\mathcal{B}$ la tribu des événements observables par X et $(\mathcal{X}, \mathcal{B}, P_X)$ l'espace probabilisé associé. Un événement B lié à $\mathcal{E}_X$ i.e. $B \in \mathcal{B}$ sera réalisé si, et seulement si, à la suite de $\mathcal{E}$ dont ω apparaît $x = X(\omega) \in B$ est observé.

Schématiquement, X est l'application $(\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{B})$ et d'une façon formelle,

$$\begin{aligned} \text{"B réalisé"} &= \{X = x, x \in \mathcal{B}\} \text{ vu au niveau de } \mathcal{E}_X \\ &= \{\omega \in \Omega, X(\omega) \in \mathcal{B}\} \text{ vu au niveau de } \mathcal{E} \\ &= X^{-1}(B) \end{aligned}$$

On va imposer que $\forall B \in \mathcal{B}, X^{-1}(B) \in \mathcal{A}$, d'où l'application $X : (\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{B})$ est une application mesurable de $(\Omega, \mathcal{A})$ dans $(\mathcal{X}, \mathcal{B})$.

2.1 Variable aléatoire - Lois de probabilité

Définition 2.1.1

On appelle variable aléatoire (en abrégé v.a.) sur $(\Omega, \mathcal{A}, P)$, toute application mesurable X de $(\Omega, \mathcal{A})$ dans $(\mathcal{X}, \mathcal{B}) : \forall B \in \mathcal{B}, X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\} = \{X \in B\} \in \mathcal{A}$.

Lorsque :

- $\mathcal{X} = \mathbb{R}$, X est dite variable aléatoire réelle (v.a.r.),
- $\mathcal{X} = \mathbb{R}^n$, X est un n -uplet de v.a.r. dit vecteur aléatoire de $\mathbb{R}^n$,
- $\mathcal{X} = E$ espace vectoriel quelconque, X est dite variable aléatoire,
- $\mathcal{X} = \mathbb{N}, \mathbb{Z}$ ou D avec D dénombrable, X est dite variable aléatoire discrète.

Exercice 2.1.1

1) Soit $A \in \mathcal{A}$. On définit la fonction indicatrice de A notée 1_A par

$$1_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases}$$

Montrer que 1_A est une v.a.r.

2) Soit $X : (\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{B})$ une v.a. et $f : (\mathcal{X}, \mathcal{B}) \rightarrow (\mathcal{Y}, \mathcal{C})$ une fonction mesurable. Montrer que $Y = f(X) = f \circ X$ est une v.a. de $(\Omega, \mathcal{A})$ dans $(\mathcal{Y}, \mathcal{C})$.

Définition 2.1.2

Soit $X : (\Omega, \mathcal{A}) \longrightarrow (\mathcal{X}, \mathcal{B})$ une v.a. On appelle tribu engendrée par X notée $\sigma(X)$, la plus petite tribu sur Ω rendant X mesurable. On a donc $\sigma(X) = \sigma\{X^{-1}(B), B \in \mathcal{B}\}$.

Propriété 2.1.1

Soient X et Y deux v.a.r. sur $(\Omega, \mathcal{A}, P)$. Alors Y est $\sigma(X)$ -mesurable ou encore $\sigma(Y) \subset \sigma(X)$ si, et seulement si, il existe une fonction Borel-mesurable f sur $\mathbb{R}$ telle que $Y = f(X)$.

Preuve :

La condition est suffisante. En effet si $Y = f(X) = f \circ X$, alors $\forall A \in \mathcal{B}_{\mathbb{R}}, Y^{-1}(A) = (f \circ X)^{-1}(A) = X^{-1}(f^{-1}(A))$. Vue que f est Borel-mesurable et X est une v.a.r., $f^{-1}(A) \in \mathcal{B}_{\mathbb{R}}$ et $X^{-1}(f^{-1}(A)) \in \mathcal{A}$. Pour la condition nécessaire, supposons d'abord que $Y = \mathbb{1}_A$ avec $A = X^{-1}(C)$, $C \in \mathcal{B}_{\mathbb{R}}$. Alors $Y = \mathbb{1}_{X^{-1}(C)} = \mathbb{1}_C \circ X = f \circ X = f(X)$ avec $f = \mathbb{1}_C$ et f est Borel-mesurable. Si Y est étagée : $Y = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$ avec $\forall i = 1, 2, \dots, n, A_i = X^{-1}(C_i), C_i \in \mathcal{B}_{\mathbb{R}}$ et on obtient $Y = f(X)$ avec $f = \sum_{i=1}^n a_i \mathbb{1}_{C_i}$.

Les cas Y positive et Y $\sigma(X)$ -mesurable quelconque s'en déduisent par passage à la limite croissante et en écrivant $Y = Y^+ - Y^-$ avec Y^+ et Y^- positives. $\square$

Définition 2.1.3

Soit $X : (\Omega, \mathcal{A}) \longrightarrow (\mathcal{X}, \mathcal{B})$ une v.a. On appelle loi de probabilité de la v.a. X , la probabilité P_X image de P par X définie par

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

La loi de X est dite multivariée si $\mathcal{X} = \mathbb{R}^n, n \geq 2$.

2.2 Moments de variables aléatoires**Définition 2.2.1**

Soit X une v.a.r. sur $(\Omega, \mathcal{A}, P)$. On appelle espérance mathématique de X lorsqu'elle existe, la quantité réelle notée $\mathbb{E}(X)$ ou $\bar{X}$ et donnée par

$$\bar{X} = \mathbb{E}(X) = \int_{\Omega} X dP.$$

Soit $k \in \mathbb{N}^*$. On appelle moment d'ordre k de X lorsqu'il existe, la quantité réelle notées $\mathbb{E}(X^k)$ ou $m_k(X)$ et donnée par

$$m_k(X) = \mathbb{E}(X^k) = \int_{\Omega} X^k dP.$$

On appelle variance (respectivement moment centré d'ordre $k \geq 1$) de X notée $\text{Var}(X)$ ou σ_X^2 (respectivement $\mu_k(X)$) lorsqu'ils existent, les quantités réels donnés par

$$\begin{aligned} \sigma_X^2 &= \text{Var}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \int_{\Omega} (X - \mathbb{E}(X))^2 dP, \\ \mu_k(X) &= \mathbb{E}(X - \mathbb{E}(X))^k = \int_{\Omega} (X - \mathbb{E}(X))^k dP. \end{aligned}$$

On appelle écart-type de X , le réel noté $\sigma_X \in \mathbb{R}^+$ donné par $\sigma_X = \sqrt{\text{Var}(X)}$.

On a le résultat suivant pour calculer facilement un moment de la v.a. X ou l'espérance d'une fonction de X .

Théorème 2.2.1 (de transfert)

Soit $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathcal{X}, \mathcal{B}, P_X)$ une v.a. et g une fonction mesurable de $(\mathcal{X}, \mathcal{B})$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Alors $g(X)$ est P -intégrable si, et seulement si, g est P_X -intégrable et on a dans ce cas $\mathbb{E}(g(X)) = \int_{\Omega} g(X) dP = \int_{\mathcal{X}} g dP_X$.

Preuve :

Si $g = \mathbb{1}_B$, $B \in \mathcal{B}$ on a

$$\begin{aligned} \int g dP_X &= \int \mathbb{1}_B dP_X = P_X(B) = P(X^{-1}(B)) = \int \mathbb{1}_{X^{-1}(B)} dP = \int \mathbb{1}_B \circ X dP \\ &= \int g(X) dP = \mathbb{E}(g(X)) \end{aligned}$$

On passe au cas g simple par linéarité de l'intégrale, puis au cas g positive par passage à la limite croissante et en appliquant le théorème de la convergence monotone (Théorème de Beppo-Levi) et ensuite, au cas g intégrable en écrivant $g = g^+ - g^-$ avec g^+ et g^- mesurables positives. $\square$

Remarques 2.2.1

- 1) Si la loi P_X de X est absolument continue par rapport à une mesure $\mu : P_X = f \cdot \mu$ avec f mesurable positive telle que $\forall B \in \mathcal{B}$, $P_X(B) = \int_B dP_X = \int_B f d\mu$, alors $\int_{\mathcal{X}} g dP_X = \int_{\mathcal{X}} g(x)f(x) d\mu(x)$.
- 2) Dans le cas où $\mathbb{E}(|X|)$ existe, on dit que X est P -intégrable et on note $X \in L^1(\Omega, \mathcal{A}, P)$ ou simplement $X \in L^1$. Plus généralement, si pour $p \in \mathbb{N}^*$, $\mathbb{E}(|X|^p) < +\infty$, on note $X \in L^p(\Omega, \mathcal{A}, P)$ ou $X \in L^p$.
- 3) L'opérateur $\mathbb{E}$ n'est autre chose que l'intégrale de Lebesgue d'une fonction mesurable par rapport à la mesure P (e.g. $\mathbb{E}$ est linéaire croissante).

2.3 Inégalités fondamentales

Théorème 2.3.1 (Inégalité de Markov)

Soit X une v.a.r. telle que $|X|^r \in L^1$, $r > 0$. Alors $\forall \varepsilon > 0$, $P[|X| > \varepsilon] \leq \frac{1}{\varepsilon^r} \mathbb{E}(|X|^r)$.

Preuve :

$$\begin{aligned} \mathbb{E}(|X|^r) &= \int |X|^r dP = \int_{\{|X| > \varepsilon\}} |X|^r dP + \int_{\{|X| \leq \varepsilon\}} |X|^r dP \\ &\geq \int_{\{|X| > \varepsilon\}} |X|^r dP \geq \varepsilon^r \int_{\{|X| > \varepsilon\}} dP = \varepsilon^r P[|X| > \varepsilon]. \end{aligned}$$

$\square$

Corollaire 2.3.1 (Inégalité de Bienamey-Tchebyshev)

Si $X \in L^2(\Omega, \mathcal{A}, P)$ alors $\forall \varepsilon > 0$, $P[|X - \mathbb{E}(X)| > \varepsilon] \leq \frac{\text{var}(X)}{\varepsilon^2}$.

Preuve :

On applique l'inégalité de Markov à la v.a. $Y = X - \mathbb{E}(X)$ avec $r = 2$. $\square$

Théorème 2.3.2 (Inégalité de Jensen)

Si $X \in L^1(\Omega, \mathcal{A}, P)$ et si $g : \mathbb{R} \longrightarrow \mathbb{R}$ est une fonction convexe telle que $g(X) \in L^1$ alors $g(\mathbb{E}(X)) \leq \mathbb{E}(g(X))$.

Preuve :

Posons $m = \mathbb{E}(X)$. g étant convexe, il existe une droite passant par le point $(m, g(m))$ et située sous le graphe de g c'est-à-dire, une fonction affine

$\varphi(x) = a(x - m) + g(m) \leq g(x) \forall x \in \mathbb{R}$. On a donc $\varphi(X) = a(X - m) + g(m) \leq g(X)$, d'où en prenant l'espérance on obtient $g(m) \leq \mathbb{E}(g(X))$ c'est-à-dire $g(\mathbb{E}(X)) \leq \mathbb{E}(g(X))$. $\square$

Exercice 2.3.1

Reprendre la démonstration de l'inégalité de Jensen en supposant que g est convexe et deux fois dérivable. Indication : écrire la formule de Taylor de g à l'ordre 2 au voisinage de $m = \mathbb{E}(X)$.

2.4 Fonction de répartition - Densité de Probabilité

Définition 2.4.1

Soit X une v.a.r. sur $(\Omega, \mathcal{A}, P)$. On appelle fonction de répartition de X , la fonction F_X définie sur $\mathbb{R}$ par $F_X(x) = P_X([-\infty, x]) = P(X \in [-\infty, x]) = P[X \leq x]$.

On vérifie que F_X caractérise la loi de la v.a. X et a les propriétés suivantes :

- (i) F_X est croissante.
- (ii) $\lim_{x \rightarrow +\infty} F_X(x) = 1$ et $\lim_{x \rightarrow -\infty} F_X(x) = 0$.
- (iii) F_X est continue à droite : $\lim_{h \rightarrow 0^+} F_X(x + h) = F_X(x)$, $\forall x \in \mathbb{R}$.

Théorème 2.4.1

Si F est une fonction sur $\mathbb{R}$ vérifiant les propriétés (i)-(iii) alors il existe un espace probabilisé $(\Omega, \mathcal{A}, P)$ et une v.a.r. X sur $(\Omega, \mathcal{A}, P)$ telle que F soit la fonction de répartition de X .

Preuve :

On pose $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$ la tribu borélienne de $\mathbb{R}$, P la mesure de Lebesgue-Stieltjes déterminée par F sur $\mathcal{A}$: $P([a, b]) = F(b) - F(a)$ et on définit X par $X(\omega) = \omega$, $\forall \omega \in \Omega$. Alors X est une v.a.r. sur $(\Omega, \mathcal{A}, P)$ telle que $\forall x \in \mathbb{R}$, $P[X \leq x] = F(x)$. $\square$

Autres propriétés de F_X :

- Si F_X est continue, on dit que la loi de X est continue ou diffuse, ce qui est équivalent à $P[X = x] = 0$, $\forall x \in \mathbb{R}$.
- S'il existe une fonction mesurable positive f telle que $\forall x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^x f(t) dt$ alors on dit que X ou la loi P_X de X est absolument continue par rapport à la mesure de Lebesgue dt sur $\mathbb{R}$ et f est appelé densité de probabilité de X . La densité de probabilité joue le même rôle que la probabilité élémentaire $P_x = P[X = x]$ dans le cas X discrète, mais $f \not\leq 1$ en général.

Remarque 2.4.1

Par analogie au Théorème 2.4.1, si g est une fonction mesurable positive ou nulle intégrable et si $\int_{-\infty}^{+\infty} g(x) dx = 1$, alors g est une densité de probabilité, celle dont la fonction de répartition absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}$ est G définie par $G(x) = \int_{-\infty}^x g(t) dt$.

Théorème 2.4.2

Si X est une v.a.r. absolument continue, ayant pour densité de probabilité f alors

- i) $P[X = x] = 0, \forall x \in \mathbb{R};$
 ii) $P[a \leq X \leq b] = \int_a^b f(t) dt, \forall a, b \in \mathbb{R}.$

Preuve :

On a $\forall x \in \mathbb{R}, P[X = x] = \int_x^x f(t)dt = 0$, cela signifie que si $a \leq b$, $P[a \leq X \leq b] = P[a < X \leq b] = F_X(b) - F_X(a)$ car $\{a < X \leq b\} = \{X \leq b\} \setminus \{X \leq a\}$ et donc que $F_X(b) - F_X(a) = \int_a^b f(t) dt.$ $\square$

Remarques 2.4.1

1) D'après le Théorème 2.4.2, si X est une v.a.r. et B est un borélien de $\mathbb{R}$ ($B \in \mathcal{B}_{\mathbb{R}}$), alors $P[X \in B] = \int_B dF_X.$

En effet si B est de la forme $B = \bigcup_{i=1}^n (a_i, b_i]$ avec $a_i, b_i \in \mathbb{R}, a_i \leq b_i$ et $n \in \mathbb{N}^*$, alors

$$\begin{aligned} P[X \in B] &= P[X \in \bigcup_{i=1}^n (a_i, b_i)] = \sum_{i=1}^n P[a_i < X \leq b_i] \\ &= \sum_{i=1}^n \int_{(a_i, b_i]} dF_X(x) = \int_{\bigcup_{i=1}^n (a_i, b_i]} dF_X(x) \\ &= \int_B dF_X(x) \end{aligned}$$

et on conclut grâce au théorème de la classe monotone vu que la classe des parties de $\mathbb{R}$ de la forme $\bigcup_{i=1}^n (a_i, b_i]$ est stable par intersection finie.

2) Si g est une fonction Borel-mesurable sur $\mathbb{R}$ alors d'après le Théorème 2.2.1, $\mathbb{E}(g(X))$ existe si et ssi, $\int_{\mathbb{R}} |g(x)| dF_X(x) < +\infty$ et dans ce cas $\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x) dF_X(x).$

3) Si X est absolument continue avec pour densité f_X alors $\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$ et $\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$ dès que ces intégrales sont finies.

2.5 Lois multivariées

Les définitions et résultats précédents se généralisent au cas de vecteurs aléatoires dans $\mathbb{R}^n$:

Définition 2.5.1

Soit $X = (X_1, \dots, X_n)$ un v.a. sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{R}^n$. On appelle fonction de répartition de X , la fonction F_X sur $\mathbb{R}^n$ définie par

$$F_X(x_1, \dots, x_n) = P[X_1 \leq x_1, \dots, X_n \leq x_n], \forall (x_1, \dots, x_n) \in \mathbb{R}^n.$$

F_X détermine la loi de probabilité de X dite loi conjointe du n -uple $(X_1, \dots, X_n)$.

On appelle loi marginale de la v.a.r. X_i du n -uple, la loi de fonction de répartition

$$F_{X_i}(x) = \lim_{\substack{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n \rightarrow +\infty \\ x_i \rightarrow x}} F_X(x_1, \dots, x_n).$$

La loi du n -uple est dite discrète s'il existe une partie dénombrable D de $\mathbb{R}^n$ telle que

$$P_X(D) = \sum_{(x_1, \dots, x_n) \in D} P[X_1 = x_1, \dots, X_n = x_n] = 1.$$

Elle est dite absolument continue s'il existe une fonction Borel-mesurable et positive f sur $\mathbb{R}^n$ telle que $\forall (x_1, \dots, x_n) \in \mathbb{R}^n$,

$$F_X(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) dt_1 \dots dt_n.$$

S'il en est ainsi, on dit que la v.a. X a pour densité

$$f_X(x_1, \dots, x_n) = \frac{\partial^n F_X(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n}.$$

On appelle densité marginale du p -uplet $(X_{i_1}, \dots, X_{i_p})$ extrait d'un n -uplet $(X_1, \dots, X_n)$, $1 \leq p \leq n$, la fonction

$$f_{(X_{i_1}, \dots, X_{i_p})}(x_{i_1}, \dots, x_{i_p}) = \int_{\mathbb{R}^{n-p}} f(x_{i_1}, \dots, x_{i_p}, u_{i_{p+1}}, \dots, u_{i_n}) du_{i_{p+1}} \cdots du_{i_n}.$$

Exemples 2.5.1

1) Soit X et Y deux v.a. sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{Z}$ telles que $\forall (i, j) \in \mathbb{Z}^2, P[X = i, Y = j] = P_{ij}$. Alors la loi conjointe du couple (X, Y) a pour fonction de répartition

$$F_{(X,Y)}(x, y) = \sum_{i \leq x, j \leq y} P_{ij}, \quad \forall (x, y) \in \mathbb{Z}^2.$$

2) Soit (X, Y) un couple de v.a.r. de fonction de répartition conjointe

$$F_{(X,Y)}(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-x-y}, & \text{si } (x, y) \in \mathbb{R}_+^2 \\ 0, & \text{sinon} \end{cases}.$$

Alors les lois marginales de X et de Y ont pour fonction de répartition $F_X(x) = (1 - e^{-x})\mathbb{1}_{\mathbb{R}^+}(x)$ et $F_Y(y) = (1 - e^{-y})\mathbb{1}_{\mathbb{R}^+}(y)$. La densité conjointe est donnée par

$$f_{(X,Y)}(x, y) = \frac{\partial^2 F_{(X,Y)}}{\partial x \partial y} = \begin{cases} e^{-x-y}, & \text{si } (x, y) \in \mathbb{R}_+^2 \\ 0, & \text{sinon} \end{cases}$$

et les densités marginales de X et Y sont données par

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y) dy = e^{-x}\mathbb{1}_{\mathbb{R}^+}(x) = F'_X(x) \quad \text{et} \\ f_Y(y) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y) dx = e^{-y}\mathbb{1}_{\mathbb{R}^+}(y) = F'_Y(y).$$

Théorème 2.5.1

Si $X = (X_1, \dots, X_n)$ est un v.a. dans $\mathbb{R}^n$ admettant pour densité f_X et si A est un borélien de $\mathbb{R}^n$ ($A \in \mathcal{B}_{\mathbb{R}^n}$) alors

$$P[X \in A] = P[(X_1, \dots, X_n) \in A] = \int_A dP_X = \int_{\{(x_1, \dots, x_n) \in A\}} f_X(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

Théorème 2.5.2

Si $X = (X_1, \dots, X_n)$ est un v.a. dans $\mathbb{R}^n$ et $g : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction Borel-mesurable, alors $\mathbb{E}(g(X_1, \dots, X_n))$ existe si et ssi $\int_{\mathbb{R}^n} |g(x_1, \dots, x_n)| dF_X(x_1, \dots, x_n) < +\infty$ et dans ce cas

$$\mathbb{E}[g(X_1, \dots, X_n)] = \int_{\Omega} g(X_1, \dots, X_n) dP = \int_{\mathbb{R}^n} g(x_1, \dots, x_n) dF_X(x_1, \dots, x_n).$$

Exemples 2.5.2

1) Soit f la fonction sur $\mathbb{R}^2$ définie par $f(x, y) = \begin{cases} \frac{1}{ab}, & \text{si } 0 < x < a \text{ et } 0 < y < b \\ 0, & \text{sinon} \end{cases}$.

f est une densité de probabilité car $f \geq 0$ et $\int_{\mathbb{R}^2} f(x, y) dxdy = 1$.

Si (X, Y) est le vecteur aléatoire de $\mathbb{R}^2$ dont la loi a pour densité f , on a

$$P[(X, Y) \in A] = \int \int_A f(x, y) dxdy = \frac{1}{ab} \int_{A \cap B} dxdy = \frac{\int_{A \cap B} dxdy}{ab} = \frac{\text{Aire}(A)}{\text{Aire}(B)}$$

pour tout borélien A , $A \subset B =]0, a[\times]0, b[$.

2) Soit X et Y deux v.a.r. d'espérances finies et soit $a, b \in \mathbb{R}$. Alors

$$\begin{aligned}\mathbb{E}(aX + bY) &= \int_{\Omega} aX + bY dP = \int \int_{\mathbb{R}^2} (ax + by) dF_{(X,Y)}(x, y) \\ &= \int \int_{\mathbb{R}^2} ax dF_{(X,Y)}(x, y) + \int \int_{\mathbb{R}^2} by dF_{(X,Y)}(x, y) \\ &= a \int_{\mathbb{R}} x dF_X(x) + b \int_{\mathbb{R}} y dF_Y(y) = a\mathbb{E}(X) + b\mathbb{E}(Y).\end{aligned}$$

ce qui signifie que l'opérateur $\mathbb{E}$ est linéaire sur l'espace $L^2(\Omega, \mathcal{A}, P)$.

Définition 2.5.2

Soit $X = (X_1, \dots, X_n)$ un v.a. de $\mathbb{R}^n$. Alors l'espérance mathématique de X est le vecteur de $\mathbb{R}^n$ donné par

$$\mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n))'$$

où dorénavant l'opération prime désigne la transposée d'un vecteur ou d'une matrice.

Pour $n = 2$, on appelle covariance du couple (X, Y) le nombre noté $\text{Cov}(X, Y)$ ou σ_{XY} donné par

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

On appelle coefficient de corrélation de X et Y , le nombre noté r_{XY} ou ρ_{XY} donné par

$$r_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

On appelle variance du v.a. $X = (X_1, \dots, X_n)$ de $\mathbb{R}^n$, la matrice carrée d'ordre n notée Σ_X et appelée matrice var-covar ou matrice de dispersion de X donnée par

$$\Sigma_X = \mathbb{E}[(X - \mathbb{E}(X))(X - \mathbb{E}(X))'] = [\text{Cov}(X_i, X_j)]_{1 \leq i, j \leq n}.$$

Remarques 2.5.1

1) En remarquant que $(X, Y) \mapsto \text{Cov}(X, Y)$ est une forme bilinéaire symétrique définie positive sur l'espace préhilbertien $L^2(\Omega, \mathcal{A}, P)$ on a, en appliquant l'inégalité de Schwarz

$$-1 \leq R_{XY} \leq 1.$$

2) On vérifie que Σ_X est une matrice symétrique et positive :

$$\forall a \in \mathbb{R}^n, a' \Sigma_X a \geq 0$$

et invariante par translation par un vecteur constant :

$$\forall a \in \mathbb{R}^n, \Sigma_{X+a} = \Sigma_X.$$

2) On vérifie aussi que si $M \in \mathcal{M}_{p,n}(\mathbb{R})$ et $Y = MX$ alors

$$\mathbb{E}(Y) = M\mathbb{E}(X) \quad \text{et} \quad \Sigma_Y = M\Sigma_X M' \in \mathcal{M}_p(\mathbb{R}).$$

2.6 Fonctions caractéristiques

En considérant dans le Théorème 2.2.1 de transfert les cas particuliers $g_t(x) = e^{<t, x>}$, $g_t(x) = e^{i<t, x>}$, où $\forall t, x \in \mathbb{R}^n$, $n \geq 1$, $<t, x> = \sum_{i=1}^n t_i x_i$ est le produit scalaire canonique des vecteurs x et t de $\mathbb{R}^n$ et $g_s(x) = s^x$, on obtient sous réserve de leurs existences :

Définition 2.6.1

Soit μ une mesure bornée sur $\mathbb{R}^n$. On appelle transformée de Fourier de μ la fonction sur $\mathbb{R}^n$, $t \mapsto \Phi_\mu(t)$ donnée par

$$\Phi_\mu(t) = \int_{\mathbb{R}^n} e^{i\langle t, x \rangle} d\mu(x).$$

On appelle transformée de Laplace (respectivement de Fourier) d'un v.a. X de $\mathbb{R}^n$, la fonction sur $\mathbb{R}^n$, $t \mapsto M_X(t)$ (resp. $t \mapsto \Phi_X(t)$) définie par

$$\begin{aligned} M_X(t) &= \mathbb{E}[e^{\langle t, X \rangle}] = \int_{\mathbb{R}^n} e^{\langle t, x \rangle} dP_X(x) \quad (\text{resp.}) \\ \Phi_X(t) &= \mathbb{E}(e^{i\langle t, X \rangle}) = \int_{\mathbb{R}^n} e^{i\langle t, x \rangle} dP_X(x). \end{aligned}$$

Lorsque X est une v.a.r. à valeurs dans $\mathbb{N}$, on définit la fonction génératrice G_X de X

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k P[X = k], \quad \forall s \in [-1, 1].$$

La transformée de Fourier $\Phi_X(t)$ est donc la transformée de Fourier de la mesure de probabilité P_X . Elle aussi appelée fonction caractéristique de X .

Exercice 2.6.1

1) Montrer que $\forall k \in \mathbb{N}$, $P_k = P[X = k] = \frac{1}{k!} G_X^{(k)}(0)$, en particulier la connaissance de G_X détermine la loi de X .

2) Montrer que $\forall k \in \mathbb{N}^*$, $\mathbb{E}(X^k)$ existe si et ssi, $G_X^{(k)}(1)$ existe et dans ce cas on a

$$\mathbb{E}[X(X-1)\cdots(X-k+1)] = G_X^{(k)}(1),$$

en particulier $\mathbb{E}(X) = G_X'(1)$ et $\mathbb{E}(X^2) = G_X''(1) + G_X'(1)$.

Le nombre $\eta_k = \mathbb{E}[X(X-1)\cdots(X-k+1)]$ lorsqu'il existe, est appelé moment factoriel d'ordre k de X .

Théorème 2.6.1

On a les propriétés suivantes :

1) La fonction caractéristique $\Phi_X(t)$ caractérise la loi de la v.a. X :

$\Phi_X(t) = \Phi_Y(t)$, $\forall t \in \mathbb{R}^n$, $n \geq 1$, si et ssi, X et Y ont même loi.

2) Si $a \in \mathbb{R}$, $b \in \mathbb{R}^n$, on a $\Phi_{aX+b}(t) = e^{i\langle t, b \rangle} \Phi_X(at)$.

3) Si X est une v.a. à valeurs dans $\mathbb{R}^n$, $n \geq 1$ et si X a un moment d'ordre $p : |X|^p \in L^1$, alors Φ_X est de classe C^p et pour $k = (k_1, \dots, k_n)$ tel que $|k|_1 = \alpha \leq p$ on a

$$\frac{\partial^{k_1}}{\partial t_1^{k_1}} \cdots \frac{\partial^{k_n}}{\partial t_n^{k_n}} \Phi_X(0) = \frac{\partial^\alpha \Phi_X}{\partial t_1^{k_1} \partial t_2^{k_2} \cdots \partial t_n^{k_n}}(0) = i^\alpha \mathbb{E}[X_1^{k_1} X_2^{k_2} \cdots X_n^{k_n}]$$

où $|k| = k_1 + \cdots + k_n$ et $0 = (0, \dots, 0)$ est le vecteur nul de $\mathbb{R}^n$.

La démonstration du 1) du théorème 2.6.1 est basée sur le lemme et le théorème suivants.

Lemme 2.6.1

L'ensemble $\mathcal{E}$ des fonctions $x \mapsto e^{i\langle t, x \rangle}$, $t \in \mathbb{R}^p$ de $\mathbb{R}^p$ dans $\mathbb{C}$ est total dans $L^1_{\mathbb{C}}(\mu)$, c-à-dire, l'adhérence de l'espace vectoriel engendré par ces fonctions est égale à $L^1_{\mathbb{C}}(\mu)$.

Preuve :

Il suffit de démontrer que pour tout $Y \in L^1_{\mathbb{C}}(\mu)$, $Y \perp \{x \mapsto e^{i\langle t, x \rangle}\}$ c'est-à-dire $\int Y(x) e^{it'x} d\mu(x) = 0$, $\forall t \in \mathbb{R}^p$ implique $Y = 0$ $\mu - pp$. Ce qui revient à montrer que $\int Y(x) \cos(t'x) d\mu(x) = 0$, $\forall t \in \mathbb{R}^p$, $\Rightarrow Y = 0$ $\mu - pp$.

Soit $\mathcal{A}$ l'algèbre engendrée par les fonctions $x \mapsto \cos(t'x)$, $t \in \mathbb{R}^p$. Il s'agit de montrer que $\forall Y \in L^1_{\mathbb{C}}(\mu)$, $Y \perp \mathcal{A} \Rightarrow Y = 0$ $\mu - pp$. Comme $\mathcal{A}$ est un ensemble de fonctions bornées stables par multiplication, soit $B_n = [-n, n]^p$, $n \in \mathbb{N}^*$. Alors la tribu engendrée par la restriction de $\mathcal{A}$ à B_n est la tribu borélienne sur B_n d'après le théorème de Stone-Weierstrass. D'où $\sigma(\mathcal{A}) = \mathcal{B}_{\mathbb{R}^p}$.

Soit $Y \in \mathcal{A}$ et soit $\mathcal{H} = \{Z \in L^\infty(\mu) \mid \int YZ d\mu = 0\}$. $\mathcal{H}$ est un espace vectoriel contenant $\mathcal{A}$, stable par limite croissante et par limite uniforme. $\mathcal{H}$ contient donc toutes les fonctions $\sigma(\mathcal{A})$ -mesurables bornées, donc aussi toutes les fonctions mesurables bornées, d'après le théorème de la classe monotone. Par conséquent $\forall B \in \mathcal{B}_{\mathbb{R}^p}$, $\mathbb{I}_B \perp Y$ i.e. $\int_B Y d\mu = 0$, ce qui implique $Y = 0$ $\mu - pp$. $\square$

Théorème 2.6.2 (d'unicité)

Soit μ et ν deux mesures bornées sur $\mathbb{R}^p$. Si $\Phi_\mu = \Phi_\nu$ alors $\mu = \nu$.

Preuve :

1) Supposons $\Phi_\mu = \Phi_\nu$ et posons $m = \frac{1}{2}(\mu + \nu)$. Alors $\Phi_\mu = \Phi_\nu = \Phi_m$. Or $m(B) = 0 \Rightarrow \mu(B) = 0$. Donc $\mu \ll m$. D'après le théorème de Radon-Nikodym, il existe $Z \geq 0$ m -intégrable telle que $\mu = Z.m$, d'où $\Phi_\mu = \Phi_\nu \Leftrightarrow \int e^{it'x} Z(x) dm(x) = \int e^{it'x} dm(x)$, $\forall t \in \mathbb{R}^p$. Soit encore $\int e^{it'x} (Z(x) - 1) dm(x) = 0$, $\forall t \in \mathbb{R}^p$. Comme $Z - 1 \in L^1(m)$, alors $Z - 1 = 0$ $m - pp$, i.e. $Z = 1$ $m - pp$. d'après le Lemme 2.6.1 précédent, soit encore $\mu = m$. D'où $\mu = \nu = m$. $\square$

Preuve du Théorème 2.6.1 :

1) On applique le Théorème d'unicité aux mesures de probabilité P_X et P_Y sur $\mathbb{R}^n$.

$$2) \Phi_{aX+b}(t) = \mathbb{E} \left[e^{i\langle t, aX+b \rangle} \right] = e^{i\langle t, b \rangle} \mathbb{E} \left[e^{i\langle t, aX \rangle} \right] = e^{i\langle t, b \rangle} \Phi_X(at).$$

3) Supposons d'abord $p = 1$ et $n = 1$. La fonction $t \mapsto e^{itX(\omega)}$ est de classe C^1 pour tout $\omega \in \Omega$, et on a $\sup_{t \in \mathbb{R}} |iX(\omega) e^{itX(\omega)}| \leq |X(\omega)|$ avec $X \in L^1$ par hypothèse. Le théorème de dérivation sous le signe intégral montre que $\Phi'_X \in C^1$ et que $\Phi'_X(t) = i\mathbb{E}(Xe^{itX})$ soit $\Phi'_X(0) = i\mathbb{E}(X)$. On termine la preuve par récurrence sur p .

Dans les cas particuliers $p = 1, n \geq 1$, le théorème fournit $\mathbb{E}(X_j) = -i \frac{\partial \Phi_X}{\partial t_j}(0, \dots, 0)$, $j = 1, \dots, n$

et pour $p = 2, n \geq 1$ $\mathbb{E}(X_k X_l) = -i \frac{\partial^2 \Phi_X}{\partial t_k \partial t_l}(0, \dots, 0)$, $k, l = 1, \dots, n$. $\square$

Propriété 2.6.1

La fonction caractéristique $\Phi_X(t)$ d'un v.a. X de $\mathbb{R}^n$, $n \geq 1$ vérifie

(i) $\Phi_X(0) = 1$, (ii) $|\Phi_X(t)| \leq 1 \forall t \in \mathbb{R}^n$, (iii) $\Phi_X(t)$ est uniformément continue.

Preuve :

$$(i) \quad \Phi_X(0) = \int_{\mathbb{R}^n} dP_X(x) = P_X(\mathbb{R}^n) = 1.$$

$$(ii) \quad \forall t \in \mathbb{R}, |\Phi_X(t)| \leq \int_{\mathbb{R}^n} |e^{i\langle t, x \rangle}| dP_X(x) = \int_{\mathbb{R}^n} dP_X(x) = 1.$$

(iii) Par l'inégalité des accroissements finis, on a

$$\sup_{|t-s| \leq \eta} |e^{i\langle t, x \rangle} - e^{i\langle s, x \rangle}| \leq \min(2, \eta \|x\|) \quad (2.1)$$

où $\|x\|$ désigne une norme sur $\mathbb{R}^n$. L'expression à droite de (2.1) étant indépendante de t et de s , on a le résultat en appliquant le théorème de la convergence dominée de Lebesgue. $\square$

Théorème 2.6.3 (Formule d'inversion)

Soit X est un v.a. dans $\mathbb{R}^n$, $n \geq 1$ telle que $\Phi_X(t) \in L^1(\mathbb{R}^n)$. Alors la loi de X est absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$ admettant pour densité de probabilité donnée par

$$f_X(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\langle t, x \rangle} \Phi_X(t) dt.$$

Le lecteur pourrait consulter le livre de Rényi pages 294-299 pour une démonstration et une extension dans le cas $n = 1$.

Exercice 2.6.2

1) Montrer que si $X \in L^1$ alors $\mathbb{E}(|X|) = \frac{1}{\pi} \int_0^{+\infty} \frac{1 - \operatorname{Re}(\Phi_X(t))}{t^2} dt$.

Indication : Poser $u = t|X|$ d'où $\int_0^{+\infty} \frac{1 - \cos(t|X|)}{t^2} dt = |X| \int_0^{+\infty} \frac{1 - \cos u}{u^2} du$.

Prendre ensuite l'espérance en utilisant le Théorème de Fubini.

2) Soit X une v.a.r. à valeurs dans $\mathbb{N}$ avec $P_k = P[X = k]$, $\forall k \in \mathbb{N}$. En supposant la fonction caractéristique Φ_X de X connue, calculer P_k , pour tout $k \in \mathbb{N}$.

Solution : $P_j = \frac{1}{2\pi} \int_0^{2\pi} e^{itj} \Phi_X(t) dt$, $j = 0, 1, \dots$

2.7 Formule de changement de variables

Pour des calculs de lois dans des situations compliquées, on a souvent besoin d'effectuer un changement de variables dans les intégrales. Le théorème suivant en est un rappel de cours d'intégration.

Théorème 2.7.1

Soit $\mathcal{D}$ et Δ deux ouverts de $\mathbb{R}^n$, $n \geq 1$ et φ un C^1 -difféomorphisme de $\mathcal{D}$ sur Δ . Alors pour toute fonction borélienne positive ou intégrable g sur Δ on a

$$\int_{\Delta} g(y) dy = \int_{\mathcal{D}} g(\varphi(x)) |J\varphi(x)| dx \quad (2.2)$$

où $J\varphi(x)$ est le déterminant de la matrice jacobienne de $\varphi = (\varphi_1, \dots, \varphi_n)$.

Cas des variables aléatoires

Soit X un v.a. à valeurs dans un ouvert $\mathcal{D}$ de $\mathbb{R}^n$ ayant une densité f_X et soit ψ un C^1 -difféomorphisme de l'ouvert $\mathcal{D}$ dans un ouvert Δ de $\mathbb{R}^n$. On pose $Y = \psi(X)$ et on cherche la loi de Y . On a pour toute fonction mesurable positive g quelconque sur Δ :

$$\begin{aligned}\mathbb{E}(g(Y)) &= \mathbb{E}(g(\psi(X))) = \int_{\Omega} g(\psi(X)) dP = \int_{\mathcal{D}} g(\psi(x)) f_X(x) dx \\ &= \int_{\Delta} g(y) f_X(\psi^{-1}(y)) |J\psi^{-1}(y)| dy \\ &= \int_{\Delta} g(y) f_X(\psi^{-1}(y)) |J\psi^{-1}(y)| \mathbb{1}_{\Delta}(y) dy.\end{aligned}$$

On en déduit que la v.a. Y a pour loi de densité donnée par

$$f_Y(y) = f_X(\psi^{-1}(y)) |J\psi^{-1}(y)| \mathbb{1}_{\Delta}(y).$$

Exemple 2.7.1

Soit $X = (X_1, X_2)$ un couple de v.a.r. de loi conjointe la loi normale $\mathcal{N}(0, I_2)$ sur $\mathbb{R}^2$ définie par la densité $f_X(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}$. On pose $Y = \frac{X_1}{X_2}$ avec $Y = 0$ si $X_2 = 0$. Trouvons la loi de Y .

Pour toute h mesurable positive sur $\mathbb{R}$ on a

$$\mathbb{E}(h(Y)) = \mathbb{E}\left[h\left(\frac{X_1}{X_2}\right)\right] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h\left(\frac{x_1}{x_2}\right) e^{-\frac{1}{2}(x_1^2 + x_2^2)} dx_1 dx_2.$$

Soit le difféomorphisme $\psi : \mathbb{R}^2 \mapsto \mathbb{R}^2$ qui à (x_1, x_2) associe (y, z) avec $y = \frac{x_1}{x_2}$ et $z = x_2$.

Alors ψ^{-1} est le difféomorphisme réciproque qui à (y, z) associe $x_1 = yz$, $x_2 = z$. On a $J\psi^{-1}(y, z) = \frac{D(x_1, x_2)}{D(y, z)} = \begin{vmatrix} z & y \\ 0 & 1 \end{vmatrix} = z$. D'où

$$\begin{aligned}\mathbb{E}(h(Y)) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(y) e^{-\frac{1}{2}(1+y^2)z^2} |z| dy dz \stackrel{\text{Fubini}}{=} \int_{\mathbb{R}} h(y) \left\{ \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\frac{1}{2}(1+y^2)z^2} |z| dy dz \right\} \\ &= \int_{\mathbb{R}} h(y) \frac{dy}{\pi(1+y^2)}.\end{aligned}$$

La loi de Y est donc la loi de densité $f(y) = \frac{1}{\pi(1+y^2)}$, $\forall y \in \mathbb{R}$.

Exemple 2.7.2 (Loi d'une somme de deux v.a.)

Soit X et Y deux v.a.r. sur $(\Omega, \mathcal{A}, P)$ tel que le couple (X, Y) ait pour densité $f_{(X,Y)}(x, y)$. On pose $Z = X + Y$ et on cherche la loi de Z . Z est à valeurs dans $\mathbb{R}$ et on a $\forall z \in \mathbb{R}$,

$$P[Z \leq z] = P[X + Y \leq z] = \int \int_{\{(x,y) \in \mathbb{R}^2 / x+y \leq z\}} f_{(X,Y)}(x, y) dx dy.$$

Posons $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 / x + y \leq z\}$ et soit φ le difféomorphisme de $\mathcal{D}$ dans Δ qui à $(x, y) \in \mathcal{D}$ associe $(u, v) \in \Delta$ avec $u = x$, $v = x + y$ et $\Delta = \mathbb{R} \times]-\infty, z]$.

Le difféomorphisme réciproque φ^{-1} est l'application qui à $(u, v) \in \Delta$ associe $x = u$, $y = v - u$. On a

$$J\varphi^{-1}(u, v) = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

d'où

$$\begin{aligned}
\int_{\mathcal{D}} f_{(X,Y)}(x,y) dx dy &= \int_{\Delta} f_{(X,Y)}(u,v-u) du dv = \int_{\mathbb{R}} \int_{(-\infty,z]} f_{(X,Y)}(u,v-u) du dv \\
&\stackrel{\text{Fubini}}{=} \int_{-\infty}^z \left(\int_{\mathbb{R}} f_{(X,Y)}(u,v-u) du \right) dv \\
&= \int_{-\infty}^z f_Z(v) dv.
\end{aligned}$$

D'où Z a pour loi de densité

$$f_Z(z) = \int_{\mathbb{R}} f(u, z-u) du.$$

Dans le cas particulier où $f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y)$, pour tout $(x,y) \in \mathbb{R}^2$, on obtient

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(u) f_Y(z-u) du \stackrel{\text{def}}{=} f_X * f_Y(z)$$

et la loi de Z est la convolution des lois de X et de Y notée $\mathcal{L}(Z) = \mathcal{L}(X) * \mathcal{L}(Y)$. $f_X * f_Y$ est par définition, la convolution de f_X par f_Y et la convolution est commutative.

Exercice 2.7.1

Soit X_1, X_2, X_3 trois v.a. indépendantes et de même loi normale centrée réduite notée $\mathcal{N}(0,1)$. On pose

$$Y_1 = X_1 + X_2 + X_3, \quad Y_2 = X_1 - X_2, \quad Y_3 = X_1 - X_3.$$

Loi du vecteur $Y = (Y_1, Y_2, Y_3)$?

Solution abrégée : On trouve

$$f_Y(y) = \frac{1}{3(2\pi)^{3/2}} e^{-Q(y)}, \quad \forall y = (y_1, y_2, y_3) \in \mathbb{R}^3$$

où $Q(y)$ est une forme quadratique sur $\mathbb{R}^3$ à préciser.

2.8 Indépendance de variables aléatoires

Rappelons que si $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathcal{X}, \mathcal{B})$ est une v.a, la tribu engendrée par X est la sous-tribu de $\mathcal{A}$ définie par $\sigma(X) = X^{-1}(\mathcal{B}) = \{X^{-1}(B), B \in \mathcal{B}\} \subset \mathcal{A}$.

Définition 2.8.1

Soit $(X_i)_{i \in I}$ une famille de v.a. sur $(\Omega, \mathcal{A}, P)$. On dit qu'elles sont indépendantes et on note $(X_i)_{i \in I} \perp\!\!\!\perp$ si et seulement si, les tribus engendrées $\sigma(X_i)_{i \in I}$ sont indépendantes.

On a les critères probabilistes suivants :

Théorème 2.8.1

Soient $X_1, X_2, \dots, X_n$ n v.a. sur $(\Omega, \mathcal{A}, P)$. Elles sont indépendantes si et seulement si, $P_{(X_1, X_2, \dots, X_n)} = \otimes_{i=1}^n P_{X_i}$.

Preuve :

Soit $X_i : (\Omega, \mathcal{A}) \longrightarrow (\mathcal{X}_i, \mathcal{B}_i)$. La loi conjointe de $(X_1, \dots, X_n)$ est la loi produit si et seulement si, pour tout parallélépipède mesurable $B_1 \times B_2 \times \dots \times B_n$, $B_i \in \mathcal{B}_i$, $i = 1, \dots, n$, on a

$$P_{(X_1, X_2, \dots, X_n)}(B_1 \times B_2 \times \dots \times B_n) = P_{X_1}(B_1)P_{X_2}(B_2) \cdots P_{X_n}(B_n)$$

ou encore

$$P[X_1 \in B_1, \dots, X_n \in B_n] = \prod_{i=1}^n P(X_i \in B_i) = \prod_{i=1}^n P(X_i^{-1}(B_i)). \quad (2.3)$$

Or l'indépendance des v.a. X_i équivaut à l'indépendance des tribus $\sigma(X_i)$ qu'elles engendrent, c'est-à-dire, les X_i sont indépendantes si et seulement si (2.3) est vérifiée. $\square$

Corollaire 2.8.1

Les v.a.r. $X_1, \dots, X_n$ sont indépendantes si et seulement si,

$$F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i), \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Preuve :

C'est (2.3) avec $B_i =]-\infty, x_i]$. $\square$

Corollaire 2.8.2

Si les n v.a.r. $X_1, \dots, X_n$ ont pour densités $f_{X_1}, \dots, f_{X_n}$ respectivement alors $(X_1, \dots, X_n) \perp\!\!\!\perp$ si et seulement si, $\forall (x_1, \dots, x_n) \in \mathbb{R}^n$,

$$f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i)$$

où $f_{(X_1, \dots, X_n)}(x_1, \dots, x_n)$ est la densité conjointe de n -uple $(X_1, \dots, X_n)$.

Proposition 2.8.1

Soit $X_1, \dots, X_n, n$ v.a. sur $(\Omega, \mathcal{A}, P)$ à valeurs $(\mathcal{X}_1, \mathcal{B}_1), \dots, (\mathcal{X}_n, \mathcal{B}_n)$ respectivement et soit $g_1, \dots, g_n$ n fonctions mesurables de $(\mathcal{X}_1, \mathcal{B}_1)$ dans $(\mathcal{Y}_1, \mathcal{C}_1), \dots, (\mathcal{X}_n, \mathcal{B}_n)$ dans $(\mathcal{Y}_n, \mathcal{C}_n)$ respectivement. Alors la suite $(X_1, \dots, X_n)$ est indépendante si et ssi $(g(X_1), \dots, g(X_n))$ est indépendante.

Preuve :

La condition est évidemment suffisante. Elle est nécessaire car les tribus engendrées par les v.a. $g_i(X_i)$ sont incluses dans les tribus engendrées par les X_i , $i = 1, 2, \dots, n$. Elles sont donc à fortiori indépendantes en vertu de la définition 2.8.1. $\square$

Théorème 2.8.2

Les v.a.r. $X_1, \dots, X_n$ sur $(\Omega, \mathcal{A}, P)$ sont indépendantes si et ssi, pour toutes fonctions réelles ou complexes Borel-mesurables positives ou bornées $h_1, \dots, h_n$

$$\mathbb{E} \left(\prod_{i=1}^n h_i(X_i) \right) = \prod_{i=1}^n \mathbb{E} (h_i(X_i)).$$

Preuve :

Notons $G(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i)$, $\forall (x_1, \dots, x_n) \in \mathbb{R}^n$ et soit μ_G la mesure de Lebesgue-Stieltjes sur $\mathbb{R}^n$ vérifiant pour tout $a_i, b_i \in \mathbb{R}$, $a_i \leq b_i$

$$\mu_G \left(\prod_{i=1}^n (a_i, b_i] \right) = \prod_{i=1}^n (F_{X_i}(b_i) - F_{X_i}(a_i)).$$

En utilisant le Théorème de Fubini avec le Corollaire 2.8.1 on a

$$\begin{aligned} \prod_{i=1}^n \mathbb{E} (h_i(X_i)) &= \prod_{i=1}^n \int_{-\infty}^{+\infty} h_i(x_i) dF_{X_i}(x_i) = \int_{\mathbb{R}^n} \left(\prod_{i=1}^n h_i(x_i) \right) dG(x_1, \dots, x_n) \\ &= \mathbb{E} \left(\prod_{i=1}^n h_i(X_i) \right) \end{aligned}$$

si et seulement si les X_i sont indépendantes. $\square$

Corollaire 2.8.3

Si $X_1, \dots, X_n$ sont n v.a.r. indépendantes et de carrés intégrables alors

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i).$$

Preuve :

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n X_i\right) &= \mathbb{E}\left[\left(\sum_{i=1}^n X_i - \mathbb{E}\left(\sum_{i=1}^n X_i\right)\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^n (X_i - \mathbb{E}(X_i))^2\right] \\ &= \sum_{i=1}^n \text{Var}(X_i) + \underbrace{\sum_{j \neq k} \mathbb{E}\left[(X_j - \mathbb{E}(X_j))(X_k - \mathbb{E}(X_k))\right]}_{= 0} \end{aligned} \quad \square$$

Corollaire 2.8.4

Soit $X_1, \dots, X_n$, n v.a. sur $(\Omega, \mathcal{A}, P)$ à valeurs $\mathbb{R}^n$, $n \geq 1$ et soient $\Phi_{X_1}, \dots, \Phi_{X_n}$ leurs fonctions caractéristiques respectives. Alors $(X_1, \dots, X_n)$ indépendante si et seulement si, $\forall (t_1, \dots, t_n) \in \mathbb{R}^n$,

$$\Phi_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \prod_{i=1}^n \Phi_{X_i}(t_i).$$

Preuve :

Prendre $h_k(x) = e^{i\langle t, x \rangle}$, $k = 1, \dots, n$ dans le Théorème 2.8.2. $\square$

Corollaire 2.8.5

Si $X_1, \dots, X_n$ sont n v.a. sur $(\Omega, \mathcal{A}, P)$ indépendantes alors la loi de $X_1 + \dots + X_n$ est égale au produit de convolution des lois de X_i :

$$P_{X_1 + \dots + X_n} = P_{X_1} * P_{X_2} * \dots * P_{X_n}.$$

En particulier si les X_i ont des densités f_{X_i} , $i = 1, 2, \dots, n$ et si $(X_1, \dots, X_n)$ indépendante alors

$$f_{X_1 + \dots + X_n} = f_{X_1} * f_{X_2} * \dots * f_{X_n}.$$

Preuve :

Par récurrence sur n . Pour $n = 2$, on a $f_{X_1 + X_2} = f_{X_1} * f_{X_2}$. On passe au rang n en écrivant $X_1 + \dots + X_n = S_{n-1} + X_n$ avec $S_{n-1} = X_1 + \dots + X_{n-1}$. On termine en appliquant l'hypothèse de récurrence et le résultat pour $n = 2$ à $S_{n-1} + X_n$. $\square$

Corollaire 2.8.6

Si les X_i , $i = 1, 2, \dots, n$ sont des v.a. indépendantes alors pour tout $1 < p < n$, les vecteurs $X = (X_1, \dots, X_p)$ et $Y = (X_{p+1}, \dots, X_n)$ sont indépendantes.

Exercice 2.8.1

1) Soit X une v.a. à valeurs $\{0, 1\}$ avec $P[X = 1] = p$ et $P[X = 0] = 1 - p$. Si $X_1, \dots, X_n$ sont n v.a. indépendantes de même loi que X , montrer que $Z = X_1 + \dots + X_n$ a pour loi $\Omega = \{1, 2, \dots, n\}$ et $\forall k \in \Omega$, $P[Z = k] = C_n^k p^k (1 - p)^{n-k}$.

2) Soit X une v.a. à valeurs $\mathbb{R}_+$ admettant pour densité $f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)$, $\lambda \in \mathbb{R}_+^*$. Montrer que si $X_1, \dots, X_n$ sont n v.a. indépendantes de même loi que X alors la v.a. $S_n = X_1 + \dots + X_n$ a pour loi de densité

$$f_{S_n}(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x).$$

3) Montrer que si X et Y sont deux v.a. indépendantes à valeurs $\mathbb{N}$, alors la fonction génératrice de $X + Y$ vérifie $G_{X+Y}(s) = G_X(s) G_Y(s)$.

2.9 Lois conditionnelles

Par analogie aux probabilités conditionnelles, on définit les lois conditionnelles dans les cas simples.

2.9.1 Cas de variables aléatoires réelles discrètes

Soit X et Y deux v.a.r. discrètes. Par analogie à la définition de la probabilité conditionnelle de A sachant B avec $P(B) > 0$, on définit la loi conditionnelle de X sachant $Y = y$ pour tout $y \in \mathbb{R}$ telle que $P(Y = y) \neq 0$ par :

$$P^{Y=y}(X = x) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

La fonction de répartition conditionnelle de X sachant $Y = y$ est définie par

$$F_{X/Y}(x|y) = P^{Y=y}(X \leq x) = \sum_{s \leq x} P^{Y=y}(X = s).$$

Remarques 2.9.1

Lorsque X et Y sont indépendantes, $P(X = x|Y = y) = P(X = x)$ et $F_{X/Y}(x|y) = F_X(x) = P(X \leq x)$.

Exemple 2.9.1

Soient X et Y deux v.a. indépendantes de lois de Poisson de paramètres λ et μ respectivement. On cherche la loi de X sachant $X + Y = n$, $n \in \mathbb{N}$. On a, pour tout $0 \leq k \leq n$,

$$P(X = k|X + Y = n) = \frac{P(X = k, X + Y = n)}{P(X + Y = n)} = \frac{P(X = k, Y = n - k)}{P(X + Y = n)}.$$

Or X et Y sont indépendantes, donc $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$ et $P(X = k, Y = n - k) = P(X = k)P(Y = n - k)$. Par conséquent,

$$P(X = k|X + Y = n) = \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)} = \frac{\frac{\lambda^k e^{-\lambda}}{k!} \frac{\mu^{(n-k)} e^{-\mu}}{(n-k)!}}{\frac{(\lambda + \mu)^n e^{-(\lambda + \mu)}}{n!}}.$$

Après simplification, on obtient

$$P(X = k|X + Y = n) = C_n^k \left(\frac{\lambda}{\lambda + \mu} \right)^k \left(\frac{\mu}{\lambda + \mu} \right)^{n-k}.$$

On en conclut que la loi de X sachant $X + Y = n$ est la loi binomiale de paramètres n et $p = \frac{\lambda}{\lambda + \mu}$.

2.9.2 Cas de variables aléatoires réelles continues

Supposons X et Y continues et soit $f_{X,Y}(x, y)$ la densité conjointe du couple (X, Y) , $f_X(x)$ et $f_Y(y)$ les densités marginales de X et de Y respectivement. Alors il existe des distributions conditionnelles de X sachant Y et de Y sachant X données par

- densité conditionnelle de X sachant $Y = y$:

$$f_X^{Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad \forall y \text{ tq } f_Y(y) \neq 0, \quad (2.4)$$

- densité conditionnelle de Y sachant $X = x$:

$$f_Y^{X=x}(y) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \quad \forall x \text{ tq } f_X(x) \neq 0. \quad (2.5)$$

Preuve de 2.5 :

On part de

$$\begin{aligned} P[Y \leq y | x \leq X < x + dx] &= \frac{P[Y \leq y, x \leq X < x + dx]}{P[x \leq X < x + dx]} \\ &= \frac{\int_{u=x}^{u=x+dx} \int_{-\infty}^{v=y} f_{X,Y}(u, v) du dv}{\int_{u=x}^{u=x+dx} f_X(u) du}. \end{aligned}$$

En divisant le numérateur et le dénominateur par dx et en faisant tendre dx vers 0, on obtient

$$P[Y \leq y | x \leq X < x + dx] \xrightarrow{dx \rightarrow 0} G(y) = \int_{-\infty}^y \frac{f_{X,Y}(x, v)}{f_X(x)} dv$$

Il est clair que $G(y)$ est la f.r. d'une loi continue ayant pour densité

$$g(y) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \quad y \in \mathbb{R},$$

et on appelle G (resp. g) distribution ou f.r. conditionnelle (resp. densité conditionnelle) de Y sachant $X = x$, valable pour tout x telle que $f_X(x) > 0$. $\square$

Remarque 2.9.1

Dans le cas où X et Y sont indépendantes, on a $f_Y^{X=x}(y) = f_Y(y)$.

Exemple 2.9.2

Soit X et Y de densité conjointe

$$f_{X,Y}(x, y) = \begin{cases} 2e^{-x-y}, & \text{si } 0 \leq x \leq y < +\infty \\ 0, & \text{sinon.} \end{cases}$$

$I(A)$ désignant ici la fonction indicatrice de l'événement A , on a

$$f_X(x) = 2e^{-2x}I(x \geq 0), \quad f_Y(y) = 2e^{-y}(1 - e^{-y})I(y \geq 0).$$

La densité conditionnelle de Y sachant $X = x$ est

$$f_Y^{X=x}(y) = \frac{2e^{-x-y}I(0 \leq x \leq y)}{2e^{-2x}I(x \geq 0)} = e^{x-y}I(0 \leq x \leq y).$$

La densité conditionnelle de X sachant $Y = y$ est

$$f_X^{Y=y}(x) = \frac{2e^{-x-y}I(0 \leq x \leq y)}{2e^{-y}(1 - e^{-y})I(y \geq 0)} = \frac{e^{-x}}{1 - e^{-y}}I(0 \leq x \leq y).$$

2.9.3 Espérance conditionnelle

Cas discret :

Pour tout y telle que $P(Y = y) > 0$, on définit l'espérance conditionnelle de X sachant $Y = y$ par

$$\mathbb{E}(X|Y = y) = \sum_x x P^{Y=y}(X = x) = \sum_x x P(X = x|Y = y) \quad (2.6)$$

Exemple 2.9.3

X et Y sont indépendantes de même loi binomiale de paramètres n et p , $\mathcal{B}(n, p)$. On cherche l'espérance conditionnelle de X sachant $X + Y = m$.

On détermine d'abord la loi conditionnelle. Posons $q = 1 - p$. Puisque X et Y sont indépendantes, on sait que $X + Y \rightsquigarrow \mathcal{B}(2n, p)$ et que $P(X = x, Y = y) = P(X = x)P(Y = y)$ (ce qui est faux en général). D'où

$$P(X = k|X + Y = m) = \frac{P(X = k)P(Y = m - k)}{P(X + Y = m)} = \frac{C_n^k p^k q^{n-k} C_n^{m-k} p^{m-k} q^{n-m+k}}{C_{2n}^m p^m q^{2n-m}}.$$

Après simplification, on obtient

$$P(X = k|X + Y = m) = \frac{C_n^k C_n^{m-k}}{C_{2n}^m}.$$

Il en résulte que la loi de X sachant $X + Y = m$ est la loi hypergéométrique $\mathcal{H}(N, n, M)$ avec $N = 2n$, $n = m$ et $M = n$ dont l'espérance mathématique est $n \cdot \frac{M}{N} = m \frac{n}{2n} = \frac{m}{2}$. On en déduit que $\mathbb{E}(X|X + Y = m) = \frac{m}{2}$.

cas continu :

Lorsque les v.a. X et Y sont continues, on définit l'espérance conditionnelle de X sachant $Y = y$ pour tout y telle que $f_Y(y) > 0$ par

$$\mathbb{E}(X|Y = y) = \int_{\mathbb{R}} x f_X^{Y=y}(x) dx. \quad (2.7)$$

Exemple 2.9.4

Soit X et Y de densité conjointe

$$f_{X,Y}(x, y) = \frac{e^y e^{-x/y}}{y} I(x > 0, y > 0).$$

On a d'abord, en se rappelant que $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$,

$$\begin{aligned} f_X^{Y=y}(x) &= \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{f_{X,Y}(x, y)}{\int_0^{+\infty} f_{X,Y}(x, y) dx} \\ &= \frac{y^{-1} e^y e^{-x/y} I(x > 0, y > 0)}{\int_{-\infty}^{+\infty} y^{-1} e^y e^{-x/y} I(x > 0, y > 0) dx} = \frac{1}{y} e^{-\frac{x}{y}} I(y > 0). \end{aligned}$$

Donc pour tout $y > 0$, la loi conditionnelle de X sachant $Y = y$ est la loi exponentielle de paramètre $\frac{1}{y}$ et cette loi a pour espérance y . Ce qui veut dire que

$$\mathbb{E}(X|Y = y) = \int_0^{+\infty} x y^{-1} e^y e^{-x/y} dx = y.$$

Notons que $\mathbb{E}(X|Y)$ est une variable aléatoire $\sigma(Y)$ -mesurable, d'où d'après la proposition 2.1.1

$$\mathbb{E}(X|Y) = f \circ Y = f(Y)$$

avec f fonction réelle mesurable. Cette fonction f est définie par $x \mapsto f(y) = \mathbb{E}(X|Y = y)$. On a

Théorème 2.9.1

$$\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X) \quad (2.8)$$

Si Y est discrète, ce théorème dit que

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \sum_y \mathbb{E}(X|Y = y)P(Y = y) \quad (2.9)$$

et dans le cas cas continu, il dit que

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \int_{\mathbb{R}} \mathbb{E}(X|Y = y)f_Y(y)dy. \quad (2.10)$$

Preuve du Théorème dans le cas X et Y discrètes :

D'après (2.9), il vient

$$\begin{aligned} \sum_y \mathbb{E}(X|Y = y)P(Y = y) &= \sum_y \sum_x xP(X = x|Y = y)P(Y = y) \\ &= \sum_y \sum_x xP(X = x, Y = y) = \sum_x xP(X = x) = \mathbb{E}(X). \end{aligned}$$

Exercices 2.9.1

Ex 1 : Soient X et Y deux v.a.r. continues. Trouver la loi de $X + Y$ et étudier le cas X et Y indépendantes.

Solution :

On a

$$\begin{aligned} P(X + Y \leq z) &= \int_{-\infty}^{+\infty} P(X + Y \leq z|Y = y)f_Y(y)dy \\ &= \int_{-\infty}^{+\infty} P(X \leq z - y|Y = y)f_Y(y)dy \\ &= \int_{-\infty}^{+\infty} F_{X|Y}(z - y|y)f_Y(y)dy \end{aligned}$$

où $F_{X|Y}(x|y)$ désigne la f.r. de la loi conditionnelle de X sachant $Y = y$.

Dans le cas d'indépendance, on a $F_{X|Y}(x|y) = F_X(x)$, d'où

$$P(X + Y \leq z) = \int_{-\infty}^{+\infty} F_X(z - y)f_Y(y)dy = \int_{-\infty}^{+\infty} F_X(z - y)f_Y(y)dy$$

et par le théorème de dérivation sous le signe intégral, on obtient la densité $f_{X+Y}(z)$ de la v.a. $X + Y$:

$$f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_Y(y)f_X(z - y)dy = f_X \star f_Y(z)$$

où $f \star g$ est la convolution de f par g .

Ex 2 : On note par $N(t)$ le nombre de personnes arrivées dans une gare à l'heure $t \geq 0$ et par Y l'heure d'arrivée du train. On admet que $N(t)$ suit la loi de Poisson de paramètre λt notée $\mathcal{P}(\lambda t)$ et que Y est indépendante de $N(t)$, $\forall t \geq 0$. Calculer $\mathbb{E}(N(Y))$.

Solution :

$\forall t \geq 0$, on a par indépendance et le fait que $N(t) \rightsquigarrow \mathcal{P}(\lambda t)$,

$$\mathbb{E}(N(t)|Y = t) = \mathbb{E}(N(t)) = \lambda t.$$

D'où

$$\mathbb{E}(N(Y)) = \mathbb{E}(\mathbb{E}(N(Y)|Y)) = \mathbb{E}(\lambda Y) = \lambda \mathbb{E}(Y).$$

Soit $\mathbb{E}(N(Y)) = \lambda \frac{T}{2}$ si Y suit la loi uniforme sur $[0, T]$ notée $\mathcal{U}[0, T]$.

Ex 3 : Soient $X_1, X_2, \dots, X_n$ des v.a.r. indépendantes et de même loi. Soit N une v.a. à valeurs entières indépendante des X_i . On pose $Y = \sum_{i=1}^N X_i$. Déterminer $\mathbb{E}(Y)$.

Solution :

En tenant compte de l'indépendance des (X_i) et de N , puis du fait que les X_i sont iid, on a

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}(\mathbb{E}(Y|N)) = \sum_{n=0}^{+\infty} \mathbb{E}(Y|N = n)P(N = n) = \sum_{n=0}^{+\infty} \mathbb{E}\left(\sum_{i=1}^n X_i\right)P(N = n) \\ &= \sum_{n=0}^{+\infty} \sum_{i=1}^n \mathbb{E}(X_i)P(N = n) = \sum_{n=0}^{+\infty} n\mathbb{E}(X_1)P(N = n) \\ &= \left(\sum_{n=0}^{+\infty} nP(N = n)\right)\mathbb{E}(X_1) = \mathbb{E}(N)\mathbb{E}(X_1). \end{aligned}$$

Ex 4 : Une urne contient b boules blanches et n boules noires. On tire les boules une à une jusqu'à l'apparition de la première boule blanche. Nombre moyen de boules noires tirées ?

Solution :

Soit X la v.a. égale au nombre de boules noires tirées, et soit $M_{b,n} = \mathbb{E}(X)$. Soit Y la v.a. définie par

$$Y = \begin{cases} 1, & \text{si la 1ère boule tirée est blanche} \\ 0, & \text{sinon} \end{cases}$$

On a

$$M_{b,n} = \mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y = 1)P(Y = 1) + \mathbb{E}(\mathbb{E}(X|Y = 0)P(Y = 0)).$$

Or $\mathbb{E}(\mathbb{E}(X|Y = 1)) = 0$, $\mathbb{E}(\mathbb{E}(X|Y = 0)) = 1 + M_{b,n-1}$ et $P(Y = 0) = \frac{n}{b+n}$, donc

$$M_{b,n} = \mathbb{E}(X) = \frac{n}{b+n}(1 + M_{b,n-1})$$

On trouve facilement

$$\begin{aligned} M_{b,1} &= \frac{1}{b+1}(1 + M_{b,0}) = \frac{1}{b+1} \\ M_{b,2} &= \frac{2}{b+2}(1 + M_{b,1}) = \frac{2}{b+1}. \end{aligned}$$

et par récurrence sur n ,

$$M_{b,n} = \mathbb{E}(X) = \frac{n}{b+1}.$$

Chapitre 3

LOIS DE PROBABILITÉ USUELLES.

Dans ce qui suit, le lecteur peut bien vérifier s'il le désire, que la loi indiquée est bien une loi de probabilité. Les valeurs numériques fournies pourraient aussi être vérifiées à titre d'exercice.

3.1 Lois discrètes courantes

3.1.1 Loi de Bernoulli de paramètre $p : \mathcal{B}(1, p)$

On dit qu'une v.a.r. X suit une loi de Bernoulli de paramètre p , $0 < p < 1$ et on note $X \sim \mathcal{B}(1, p)$ si $\mathcal{X} = X(\Omega) = \{0, 1\}$ et on a $P[X = 1] = p$ et $P[X = 0] = q = 1 - p$. On a $E(X) = p$ et $Var(X) = p(1 - p) = pq$.

Exemple 3.1.1

Soit $\mathcal{E}$ une e.a. et A un événement attaché à $\mathcal{E}$ de probabilité $p = P(A)$, $0 < p < 1$ et X la v.a. égale à 1 si A est réalisé et 0 sinon. Alors $P[X = 1] = P(A) = p$ et $P[X = 0] = P(\bar{A}) = q$ et $X \sim \mathcal{B}(1, p)$.

3.1.2 Loi binomiale de paramètres n et $p : \mathcal{B}(n, p)$

Une v.a.r. X suit une loi binomiale de paramètres n et p , $0 \leq p \leq 1$ et on note $X \sim \mathcal{B}(n, p)$ si $\mathcal{X} = X(\Omega) = \{0, 1, \dots, n\}$ et $\forall k = 0, 1, \dots, n$,

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

On a $E(X) = np$ et $Var(X) = np(1 - p) = npq$.

Par exemple si on répète n fois de suite l'expérience $\mathcal{E}$ précédente dans les mêmes conditions et de façon indépendante, et si X est la v.a. égale au nombre de fois que l'événement A est réalisé au cours des n épreuves, on a $X(\Omega) = \{0, 1, \dots, n\}$ et $\forall k \in X(\Omega)$, $P[X = k] = C_n^k p^k (1 - p)^{n-k}$ et $X \sim \mathcal{B}(n, p)$.

On remarque que $X = \sum_{i=1}^n X_i$ avec $X_i = \mathbb{1}_A \sim \mathcal{B}(1, p)$ et $X_1, \dots, X_n$ indépendantes. La loi $\mathcal{B}(n, p)$ est donc la loi de la somme de n v.a. indépendantes et de même loi $\mathcal{B}(1, p)$.

3.1.3 Loi multinomiale de paramètres $n, p_1, \dots, p_r : \mathcal{M}(n, p_1, \dots, p_r)$

Une v.a. X suit une loi multinomiale de paramètres $n, p_1, \dots, p_r, p_i \geq 0, i = 1, \dots, r$ et $\sum_{i=1}^r p_i = 1$ et on note $X \sim \mathcal{M}(n, p_1, \dots, p_r)$, si $\forall k = (k_1, \dots, k_r) \in \mathbb{N}^r, \sum_{i=1}^r k_i = n$,

$$P[X = k] = \frac{n!}{k_1! k_2! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}.$$

Par exemple, dans une population il y a des individus de types $1, 2, \dots, r$, chaque individu étant de type i avec probabilité p_i telles que $p_1 + \dots + p_r = 1$. On tire au hasard n individus. Soit X_i = nombre d'individus de type i tirés. Alors $X = (X_1, \dots, X_r)$ et $X \sim \mathcal{M}(n, p_1, \dots, p_r)$. On a

$$\mathbb{E}(X_i) = np_i \quad \text{et} \quad \text{Cov}(X_i, X_j) = -np_i q_j \quad \text{avec} \quad q_j = 1 - p_j, \quad 1 \leq i, j \leq r.$$

3.1.4 Loi de Poisson de paramètre $\lambda : \mathcal{P}(\lambda)$

Une v.a. X suit une loi de Poisson de paramètre $\lambda > 0$ et on note $X \sim \mathcal{P}(\lambda)$ si $X(\Omega) = \mathbb{N}$ et $\forall k \in \mathbb{N}$,

$$P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}.$$

On a $\mathbb{E}(X) = \lambda$ et $\text{Var}(X) = \lambda$.

La loi de Poisson est en faite une sorte de loi binomiale pour laquelle n est assez grand et p petite voisine de 0. Par exemple si p est la probabilité de rencontrer un étudiant le jour du marché Assi-yéyé d'Agoè, la loi du nombre d'étudiants présents au marché ce jour est une loi de Poisson de paramètre np si n est le nombre total de gens présents au marché ce jour.

3.1.5 Loi de Pascal ou loi géométrique de paramètre $p : \mathcal{G}(p)$

Une v.a. X suit une loi de Pascal de paramètre p , $0 < p < 1$ et on note $X \sim \mathcal{G}(p)$ si $X(\Omega) = \mathbb{N}^*$ et $\forall k \in \mathbb{N}^*$,

$$P[X = k] = p(1 - p)^{k-1} = pq^{k-1}.$$

Par exemple en revenant à l'expérience aléatoire $\mathcal{E}$ du premier paragraphe, la v.a. égale au nombre d'essais nécessaires pour obtenir pour la première fois l'événement A , suit la loi de Pascal de paramètre $p = P(A)$. On a $\mathbb{E}(X) = \frac{1}{p}$ et $\text{Var}(X) = \frac{q}{p^2}$. Remarquons que la loi géométrique peut être définie sur $\mathbb{N}$. Dans ce cas la v.a. X est le nombre d'échecs nécessaires pour obtenir pour la première fois un succès.

3.1.6 Loi binomiale négative de paramètres n et $p : \mathcal{BN}(n, p)$

Une v.a. X suit une loi binomiale négative de paramètres n et p , $0 < p < 1$, si $X(\Omega) = \{n, n+1, \dots\}$ et $\forall k \geq n$,

$$P[X = k] = C_{k-1}^{n-1} p^n (1-p)^{k-n} = C_{n+k-1}^{n-1} p^n (1-p)^k \mathbb{1}_{\{k \geq n\}}.$$

Par exemple la v.a. égale au nombre d'essais nécessaires pour obtenir n réalisations de l'événement A suit une loi $\mathcal{BN}(n, p)$.

3.1.7 Loi Hypergéométrique de paramètres $N, M, n : \mathcal{H}(N, M, n)$

Une v.a. X suit une loi hypergéométrique de paramètres N, M et n si $X(\Omega) = \{0, 1, \dots, n\}$, $n \leq N$ et $\forall k \in X(\Omega)$,

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}.$$

Par exemple N = taille de la population estudiantine de l'UL, M le nombre d'étudiants étrangers. On interroge n étudiants choisis au hasard sur le campus de l'UL. La loi de la v.a. X égale au nombre d'étudiants étrangers présents parmi les n étudiants interrogés est une loi hypergéométrique de paramètres N, M, n .

Notons que parfois les paramètres sont N, n, p où p est la proportion des M étudiants étrangers au sein du campus. On a $\mathbb{E}(X) = np$ et $\text{Var}(X) = \frac{N-n}{N-1} npq$, avec $q = 1 - p$.

3.2 Lois continues courantes

3.2.1 Loi uniforme sur $[a, b] : \mathcal{U}[a, b]$

On peut supposer $a = 0$, et $b = 1$. On dit qu'une v.a. X suit une loi uniforme sur $[0, 1]$ ou simplement une loi uniforme, si X est continue et a pour densité

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

On a $\mathbb{E}(X) = \frac{1}{2}$ et $\text{Var}(X) = \frac{1}{12}$.

Plus généralement, X suit une loi uniforme sur l'intervalle $[a, b]$, $a < b$ et on note $X \sim \mathcal{U}[a, b]$ si la v.a. $U = \frac{X - a}{b - a}$ suit la loi $\mathcal{U}[0, 1]$. On en déduit que

$$f_X(x) = \frac{1}{b - a} \mathbb{1}_{[a,b]}(x), \quad \mathbb{E}(X) = \frac{a + b}{2} \quad \text{et} \quad \text{var}(X) = \frac{(b - a)^2}{12}.$$

3.2.2 Loi gamma de paramètres p et $\theta : \gamma(p, \theta)$

X suit une loi de gamma de paramètres $p > 0$, $\theta > 0$ et on note $X \sim \gamma(p, \theta)$, si X est continue et a pour densité

$$f_X(x) = \frac{\theta^p}{\Gamma(p)} x^{p-1} e^{-\theta x} \mathbb{1}_{\mathbb{R}^+}(x)$$

où $\Gamma(s) = \int_0^{+\infty} x^{s-1} e^{-x} dx$, $\forall s > 0$. On a

$$\mathbb{E}(X) = \frac{p}{\theta}, \quad \text{Var}(X) = \frac{p}{\theta^2} \quad \text{et} \quad \forall k \in \mathbb{N}^*, \quad \mathbb{E}(X^k) = \frac{\Gamma(p + k)}{\theta^k \Gamma(p)}.$$

Cas particulier : Pour $p = 1$, la loi $\gamma(1, \theta)$ est appelée loi exponentielle de paramètre $\theta > 0$ notée $\mathcal{E}(\theta)$. $X \sim \mathcal{E}(\theta)$ si X est continue et a pour densité

$$f_X(x) = \theta e^{-\theta x} \mathbb{1}_{\mathbb{R}^+}(x).$$

3.2.3 Loi beta de paramètres p et $q : b(p, q)$

$X \sim b(p, q)$, $p > 0$, $q > 0$ si X est continue et a pour densité

$$f_X(x) = \frac{x^{p-1} (1 - x)^{q-1}}{\beta(p, q)} \mathbb{1}_{[0,1]}(x)$$

où $\beta(p, q) = \int_0^1 x^{p-1} (1 - x)^{q-1} dx$. On a

$$\mathbb{E}(X) = \frac{p}{p + q}, \quad \text{Var}(X) = \frac{pq}{(p + q)^2 (p + q + 1)}.$$

3.2.4 Loi normale de moyenne m et d'écart-type $\sigma : \mathcal{N}(m, \sigma)$

$X \sim \mathcal{N}(m, \sigma)$, $m \in \mathbb{R}$ et $\sigma > 0$, si X est continue et admet pour densité

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

La courbe représentative de $f_X(x)$ est symétrique par rapport à la droite d'équation $x = m$.

Dans le cas particulier $m = 0$ et $\sigma = 1$, la loi $\mathcal{N}(0, 1)$ est appelée loi normale centrée réduite ou loi de Laplace-Gauss. Sa densité est

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

Elle a une représentation graphique de la forme

~~μ~~

Soit $\Phi(x)$ la fonction de répartition de la loi $\mathcal{N}(0, 1)$. $\Phi(x) = \int_{-\infty}^x f_X(t) dt$. On a la propriété suivante résultant de la symétrie de $\Phi(x)$: $\Phi(-u) = 1 - \Phi(u)$.

Donc si $u_\alpha \in \mathbb{R}$ vérifie $\Phi(u_\alpha) = \alpha \in [0, 1]$, on a en valeur absolue $u_\alpha = u_{1-\alpha}$.

Notons que $X \rightsquigarrow \mathcal{N}(m, \sigma)$ si et ssi, $U = \frac{X - m}{\sigma} \rightsquigarrow \mathcal{N}(0, 1)$ ou encore si et ssi $X = \sigma U + m$ avec $U \rightsquigarrow \mathcal{N}(0, 1)$. D'où $\mathbb{E}(X) = m$ et $\sigma_X^2 = \text{Var}(X) = \sigma^2$.

3.2.5 Loi de Cauchy de paramètres a, b : Cauchy (a, b)

$X \rightsquigarrow \text{Cauchy}(a, b)$ si X est continue et a pour densité

$$f_X(x) = \frac{b}{\pi(b^2 + (x - a)^2)}.$$

Les cas $a = 0$ et $b = 1$ sont les plus courants. Comme on le voit bien, la loi de Cauchy n'a aucun moment.

Le tableau suivant fournit un résumé utile de certaines valeurs caractéristiques courantes que le lecteur peut retenir, sinon retrouver par calcul :

$X \rightsquigarrow$	$\mathcal{B}(1, p)$	$\mathcal{B}(n, p)$	$\mathcal{P}(\lambda)$	$\mathcal{G}(p)$	$\mathcal{BN}(n, p)$	$\mathcal{U}[0, 1]$	$\mathcal{E}(\lambda)$	$\gamma(n, \lambda)$	$\mathcal{N}(0, 1)$
$\mathbb{E}(X)$	p	np	λ	$\frac{1}{p}$	$\frac{n}{p}$	$\frac{1}{2}$	$\frac{1}{\lambda}$	$\frac{n}{\lambda}$	0
σ_X^2	pq	npq	λ	$\frac{q}{p^2}$	$\frac{nq}{p^2}$	$\frac{1}{12}$	$\frac{1}{\lambda^2}$	$\frac{n}{\lambda^2}$	1
$G_X(s)$	$(q + ps)$	$(q + ps)^n$	$e^{\lambda(s-1)}$	$\frac{ps}{1-qs}$	$\left(\frac{ps}{1-qs}\right)^n$	$\times$	$\times$	$\times$	$\times$
$\varphi_X(t)$	$q + pe^{it}$	$(q + pe^{it})^n$	$e^{\lambda(e^{it}-1)}$	$\frac{pe^{it}}{1-qe^{it}}$	$\left(\frac{pe^{it}}{1-qe^{it}}\right)^n$	$\frac{e^{it}-1}{it}$	$\frac{\lambda}{\lambda-it}$	$\left(\frac{\lambda}{\lambda-it}\right)^n$	$e^{-\frac{t^2}{2}}$

Par exemple,

- 1) $X \rightsquigarrow \mathcal{U}[a, b] \iff X = (b - a)U + a$ avec $U \rightsquigarrow \mathcal{U}[0, 1]$. Donc $\Phi_X(t) = e^{ita}\Phi_U((b - a)t)$, soit $\Phi_X(t) = \frac{e^{itb} - e^{ita}}{(b - a)it}$.
- 2) $X \rightsquigarrow \mathcal{N}(m, \sigma) \iff X = \sigma U + m$ avec $U \rightsquigarrow \mathcal{N}(0, 1)$. Donc $\Phi_X(t) = e^{itm}\varphi_U(\sigma t)$, soit $\Phi_X(t) = e^{itm}e^{-\frac{\sigma^2 t^2}{2}}$.
- 3) Si $X \rightsquigarrow \text{Cauchy}(0, 1)$, on a en utilisant le théorème des résidus $\Phi_X(t) = e^{-|t|}$.

3.3 Vecteurs Gaussiens

Définition 3.3.1

Soit X un v.a. dans $\mathbb{R}^n$. On dit que X est normal ou gaussien s'il existe $m \in \mathbb{R}^n$ et $\Gamma \in \mathcal{M}_n(\mathbb{R})$ symétrique et positive telle que

$$\Phi_X(t) = e^{i\langle t, m \rangle} e^{-\frac{1}{2}\langle \Gamma t, t \rangle}, \quad \forall t \in \mathbb{R}^n$$

où $\langle t, m \rangle = \sum_{i=1}^n t_i m_i$ et $\langle \Gamma t, t \rangle = t' \Gamma t = \sum_{1 \leq i, j \leq n} \gamma_{ij} t_i t_j$, γ_{ij} étant les éléments de la matrice Γ . On note $X \rightsquigarrow \mathcal{N}(\vec{m}, \Gamma)$.

Théorème 3.3.1

Un v.a. $X = (X_1, \dots, X_n)$ dans $\mathbb{R}^n$ est gaussien si et seulement si, toute combinaison linéaire des X_i (en incluant les constantes) suit une loi normale sur $\mathbb{R}$.

Preuve :

Supposons X gaussien et soit $c_1, c_2, \dots, c_n \in \mathbb{R}$, $c = (c_1, \dots, c_n)$. Posons $Y = \sum_{k=1}^n c_k X_k = \langle c, X \rangle$.

$$\begin{aligned} \forall t \in \mathbb{R}, \quad \Phi_Y(t) &= \mathbb{E}\left(e^{it \sum_{k=1}^n c_k X_k}\right) = \mathbb{E}\left(e^{i\langle tc, X \rangle}\right) \\ &= \Phi_X(tc) \\ &= e^{i\langle m, tc \rangle} e^{-\frac{1}{2}\langle \Gamma(tc), tc \rangle} = e^{it\langle m, c \rangle} e^{-\frac{1}{2}t^2 \langle \Gamma c, c \rangle} \end{aligned}$$

avec $\langle m, c \rangle = \sum_{i=1}^n c_i m_i$. $\Phi_Y(t)$ est donc la f.c. d'une loi normale sur $\mathbb{R}$ de moyenne $\langle m, c \rangle$ et

de variance $\sigma^2 = \langle \Gamma c, c \rangle = \sum_{i,j=1}^n \gamma_{ij} c_i c_j$, γ_{ij} , $1 \leq i, j \leq n$ étant les éléments de la matrice Γ .

En effet, si $c_i = 1$ et $c_j = 0 \quad \forall j \neq i$, $Y = X_i \rightsquigarrow \mathcal{N}(m_i, \gamma_{ii})$ avec $m_i = \mathbb{E}(X_i)$ et $\gamma_{ii} = \text{Var}(X_i)$.

si $c_i = c_j = 1$ et $c_k = 0 \quad \forall k \neq i, j$, $Y = X_i + X_j$ et $Y \rightsquigarrow \mathcal{N}(m_i + m_j, \sigma_{ij})$ avec $m_i + m_j = \mathbb{E}(X_i) + \mathbb{E}(X_j)$ et $\sigma_{ij} = \text{Var}(X_i + X_j) = \text{Var}(X_i) + \text{Var}(X_j) + 2\text{Cov}(X_i, X_j) = \gamma_{ii} + \gamma_{jj} + 2\gamma_{ij}$. Il en résulte que $[\gamma_{ij}]_{1 \leq i, j \leq n} = \Gamma$.

Réciproquement soit X un v.a. de $\mathbb{R}^n$, $X = (X_1, \dots, X_n)$ telle que toute combinaison linéaire des X_i soit gaussienne sur $\mathbb{R}$.

Posons $m_i = \mathbb{E}(X_i)$, $\gamma_{ij} = \text{Cov}(X_i, X_j)$, $m = (m_i)_{i=1, \dots, n}$ et $\Gamma = [\gamma_{ij}]_{1 \leq i, j \leq n}$. Soit $t = (t_1, \dots, t_n) \in \mathbb{R}^n$.

Par hypothèse $\langle t, X \rangle = \sum_{i=1}^n t_i X_i$ est gaussienne sur $\mathbb{R}$ de moyenne $\sum_{i=1}^n t_i \mathbb{E}(X_i) = \sum_{i=1}^n t_i m_i = \langle m, t \rangle$ et de variance

$$\sigma^2 = \sum_{i,j=1}^n \text{Cov}(t_i X_i, t_j X_j) = \sum_{i,j=1}^n t_i t_j \text{Cov}(X_i, X_j) = \sum_{i,j=1}^n \gamma_{ij} t_i t_j = \langle \Gamma t, t \rangle.$$

De plus $\langle \Gamma t, t \rangle = \sigma^2 \geq 0$. D'où Γ est du type positif. Γ est aussi symétrique et on a $\mathbb{E}(e^{i\langle t, X \rangle}) = \Phi_{\langle t, X \rangle}(1) = e^{i\langle m, t \rangle} e^{-\frac{1}{2}\langle \Gamma t, t \rangle}, \forall t \in \mathbb{R}^n$. X est donc gaussien avec $X \sim \mathcal{N}(\vec{m}, \Gamma)$. $\square$

Corollaire 3.3.1

Si $X = (X_1, \dots, X_n) \sim \mathcal{N}(\vec{m}, \Gamma)$ alors $\vec{m} = (\mathbb{E}(X_i))_{1 \leq i \leq n}$ et $\Gamma = [\text{Cov}(X_i, X_j)]_{1 \leq i, j \leq n}$.

Théorème 3.3.2

Soit X un vecteur gaussien de $\mathbb{R}^n$: $X \sim \mathcal{N}(\vec{m}, \Gamma)$ et soit $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ une fonction affine : $g(x) = \varphi(x) + b$ avec $\varphi \in L(\mathbb{R}^n)$ et $b \in \mathbb{R}^p$. Alors $g(X)$ est gaussien de loi $\mathcal{N}(g(m), \varphi \Gamma \varphi^t)$.

Preuve :

$$\begin{aligned} \forall s \in \mathbb{R}^p, \Phi_{g(X)}(s) &= \mathbb{E}(e^{i\langle g(X), s \rangle}) = \mathbb{E}(e^{i\langle \varphi(X) + b, s \rangle}) = e^{i\langle b, s \rangle} \mathbb{E}(e^{i\langle \varphi(X), s \rangle}) \\ &= e^{i\langle b, s \rangle} \mathbb{E}(e^{i\langle X, {}^t\varphi s \rangle}) = e^{i\langle b, s \rangle} \Phi_X({}^t\varphi s) \\ &= e^{i\langle \varphi(m) + b, s \rangle} e^{-\frac{1}{2}\langle \varphi \Gamma \varphi^t s, s \rangle} = e^{i\langle g(m), s \rangle} e^{-\frac{1}{2}\langle \varphi \Gamma \varphi^t s, s \rangle} \\ &= \Phi_{\mathcal{N}(g(m), \varphi \Gamma \varphi^t)}(s). \end{aligned}$$

D'où $g(X) \sim \mathcal{N}(g(m), \varphi \Gamma \varphi^t)$ d'après le théorème d'unicité. $\square$

3.3.1 Représentation de lois normales sur $\mathbb{R}^n$

Soit $m \in \mathbb{R}^n$, Γ une matrice carrée d'ordre n et $X \sim \mathcal{N}(\vec{m}, \Gamma)$. Si Γ est du type positif de rang $p < n$, Γ est semblable à une matrice diagonale c'est-à-dire, il existe une base orthonormée de $\mathbb{R}^n$ dans laquelle Γ est semblable à $\tilde{\Gamma} = \text{diag}(\lambda_1, \dots, \lambda_p, 0, \dots, 0)$ avec $\lambda_i > 0, \forall i = 1, 2, \dots, p$. Soit P la matrice de passage. P est orthogonale c'est-à-dire $P^{-1} = {}^tP$ et on a $\Gamma = P \tilde{\Gamma} {}^tP$. Posons $Y = {}^tP(X - m)$. Y est gaussienne comme fonction linéaire d'un vecteur gaussien. On a $\mathbb{E}(Y) = 0$ et

$$\begin{aligned} \Phi_Y(t) &= e^{-\frac{1}{2}\langle {}^tP\Gamma P t, t \rangle} = e^{-\frac{1}{2}\langle \tilde{\Gamma} t, t \rangle} = e^{-\frac{1}{2}\sum_{i=1}^p \lambda_i t_i^2} = \prod_{i=1}^p e^{-\frac{1}{2}\lambda_i t_i^2} \\ &= \Phi_{\mathcal{N}(0, \lambda_1)}(t_1) \times \Phi_{\mathcal{N}(0, \lambda_2)}(t_2) \times \dots \times \Phi_{\mathcal{N}(0, \lambda_p)}(t_p) \times \Phi_{\delta_0}(t_{p+1}) \times \dots \times \Phi_{\delta_0}(t_n) \\ &= \Phi_{\mathcal{N}(0, \lambda_1) \otimes \mathcal{N}(0, \lambda_2) \otimes \dots \otimes \mathcal{N}(0, \lambda_p) \otimes \delta_0 \otimes \dots \otimes \delta_0}(t), \forall t \in \mathbb{R}^n. \end{aligned}$$

Donc $Y = (Y_1, \dots, Y_p, 0, \dots, 0)$ avec $Y_1, Y_2, \dots, Y_p$ indépendantes telles que pour tout $i = 1, 2, \dots, p$, $Y_i \sim \mathcal{N}(0, \sigma_i^2)$, où $\sigma_i^2 = \lambda_i$ est une valeur propre de la matrice carrée symétrique Γ . On a donc $X = PY + m$.

En conclusion, si X est un v.a. gaussien dans $\mathbb{R}^n$ c-à-d. $X \sim \mathcal{N}(\vec{m}, \Gamma_X)$, $\vec{m} \in \mathbb{R}^n$ et $\Gamma_X \in \mathcal{M}_n(\mathbb{R})$ symétrique positive et de rang p , il existe une suite de p v.a.r. $Y_1, Y_2, \dots, Y_p$ indépendantes, de loi $\mathcal{N}(0, \lambda_i)$, $i = 1, 2, \dots, p$ respectivement où les λ_i sont les valeurs propres non nulles de Γ_X et une matrice orthogonale P telles que $Y = (Y_1, \dots, Y_p, 0, \dots, 0)$ et $X = PY + m$. Réciproquement, on a

Théorème 3.3.3

$\forall \vec{m} \in \mathbb{R}^n$ et $\forall \Gamma \in \mathcal{M}_n(\mathbb{R})$ symétrique positive, il existe un v.a. X dans $\mathbb{R}^n$ telle que $X \sim \mathcal{N}(\vec{m}, \Gamma)$.

3.3.2 Singularité et absolue continuité des lois gaussiennes

Rappelons qu'une mesure ν est dite étrangère à une mesure μ (on dit aussi ν est μ -singulière) s'il existe N tel que $\mu(N) = 0$ et $\nu(N) = 1$. En général toute mesure discrète sur $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ est

étrangère à la mesure de Lebesgue sur $\mathbb{R}^n$, $n \geq 1$. D'autre part on dit que ν est absolument continue par rapport à μ s'il existe une fonction mesurable positive f telle que $\frac{d\nu}{d\mu} = f$ (Théorème de Radon-Nikodym) et f est appelée densité de ν par rapport à μ .

Théorème 3.3.4

Soit X un vecteur gaussien dans $\mathbb{R}^n$ de moyenne $\vec{m}$ et de variance Γ_X .

- 1) Si $\text{rg}(\Gamma_X) < n$, alors X reste presque sûrement dans un sous-espace affine de $\mathbb{R}^n$ de dimension $\text{rg}(\Gamma_X)$ et P_X est étrangère à la mesure de Lebesgue sur $\mathbb{R}^n$.
- 2) Si $\text{rg}(\Gamma_X) = n$ alors X admet une densité f_X par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$ donnée par

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Gamma_X}} \exp \left\{ -\frac{1}{2} \langle \Gamma_X^{-1}(x - m), (x - m) \rangle \right\}, \quad \forall x \in \mathbb{R}^n.$$

Preuve :

1) Soit $p = \text{rg}(\Gamma_X) < n$ et $Y = (Y_1, \dots, Y_p, 0, \dots, 0) \in E_p = \mathbb{R}^p \times \{0\}^{n-p}$ telle que $X = PY + m$ comme dans la représentation précédente. Alors X reste presque sûrement dans $V_X = PE_p + m$ qui est une variété de dimension p . La loi de X est donc portée par V_X et donc $\lambda(V_X) = 0$ où λ est la mesure de Lebesgue sur $\mathbb{R}^n$. Par suite P_X est étrangère à λ .

2) Si $p = n$, on a $Y = (Y_1, \dots, Y_n)$ avec $Y_i \in \mathcal{N}(0, \sigma_i^2)$, $\sigma_i^2 = \lambda_i$, $i = 1, \dots, n$ et $(Y_1, \dots, Y_n)$ indépendante. Donc Y admet pour densité

$$f_Y(y) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Gamma_X}} \exp \left\{ -\frac{1}{2} \langle \tilde{\Gamma}_X^{-1} y, y \rangle \right\}, \quad \forall y = (y_1, \dots, y_n) \in \mathbb{R}^n.$$

Mais $X = PY + m = \Theta(Y) \implies Y = \Theta^{-1}(X) = {}^tP(X - m)$ et $\Gamma_X = P \tilde{\Gamma}_X {}^tP$.

Comme P est orthogonale donc $\det P = 1$, P est un isomorphisme. D'après la formule de changement de variable, $\forall x \in \mathbb{R}^n$

$$\begin{aligned} f_X(x) &= f_Y(\Theta^{-1}(x)) |J_{\Theta^{-1}}(x)| = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Gamma_X}} \exp \left\{ -\frac{1}{2} \langle \tilde{\Gamma}_X^{-1} {}^tP(x - m), {}^tP(x - m) \rangle \right\} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Gamma_X}} \exp \left\{ -\frac{1}{2} \langle P \tilde{\Gamma}_X^{-1} {}^tP(x - m), (x - m) \rangle \right\} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Gamma_X}} \exp \left\{ -\frac{1}{2} \langle \Gamma_X^{-1}(x - m), (x - m) \rangle \right\}. \end{aligned}$$

□

Remarque 3.3.1

Pour $n = 1$, X gaussienne sur $\mathbb{R}$ de moyenne $m \in \mathbb{R}$ et de variance $\Gamma = \sigma^2$ si et seulement,

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

Pour $\sigma = 0$, $\Phi_X(t) = e^{itm} = \Phi_{\delta_m}(t)$ et $X \rightsquigarrow \delta_m$ où δ_m est la mesure de Dirac en m , par conséquent $X = m$ p.s. On dit que X ou la loi de X est dégénérée.

Exercice 3.3.1

Soit X un vecteur gaussien dans $\mathbb{R}^n$, $X = (X_1, \dots, X_n)$ et soit I et J deux parties disjointes de $\{1, 2, \dots, n\}$. Montrer que si $\forall i \in I$ et $\forall j \in J$, $\text{Cov}(X_i, X_j) = 0$ alors $(X_i)_{i \in I}$ et $(X_j)_{j \in J}$ sont

indépendantes.

indication

Posons $Y = (X_1, X_2, \dots, X_p, X_{p+1}, \dots, X_n) = (U, V)$ avec $U = (X_1, X_2, \dots, X_p)$ et $V = (X_{p+1}, \dots, X_n)$.

Alors $\Gamma_Y = \begin{bmatrix} \Gamma_U & 0 \\ 0 & \Gamma_V \end{bmatrix}$ et $\forall t \in \mathbb{R}^n, t = (u, v), u \in \mathbb{R}^p, v \in \mathbb{R}^{n-p}, \Phi_Y(t) = \Phi_Y(u, v) = \Phi_U(u) \cdot \Phi_V(v)$.

Chapitre 4

CONVERGENCE EN LOI – THÉORÈME CENTRAL LIMITE

4.1 Convergence des mesures positives bornée

Soient E un espace métrique (e.g. $E = \mathbb{R}^n$), et soient $C_k(E)$, $C_0(E)$, $C_b(E)$ respectivement l'espace des fonctions sur E continues à support compact, tendant vers 0 à l'infini, bornées. On note par $\mathcal{B}_E$ la tribu borélienne de E et par $\mathcal{M}(E)$ l'ensemble des mesures positives et bornées sur $(E, \mathcal{B}_E)$. On a

Définition 4.1.1

Soit (μ_n) une suite d'éléments de $\mathcal{M}(E)$ et $\mu \in \mathcal{M}(E)$. On dit que

- (i) μ_n tend vers μ vaguement si $\forall f \in C_k(E)$, $\int f d\mu_n \rightarrow \int f d\mu$.
- (ii) μ_n tend vers μ faiblement si $\forall f \in C_0(E)$, $\int f d\mu_n \rightarrow \int f d\mu$.
- (iii) μ_n tend vers μ étroitement si $\forall f \in C_b(E)$, $\int f d\mu_n \rightarrow \int f d\mu$.

On remarque que (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i), mais, il n'y a pas réciprocity en général (par exemple $n\delta_n$ tend vers 0 vaguement mais non faiblement, et δ_n tend vers 0 faiblement mais non étroitement).

Dans toute la suite, on notera par $\|f\| = \sup_{x \in E} |f(x)|$ et éventuellement par $\mu(f) = \int f d\mu$. On rappelle qu'un espace vectoriel $\mathcal{H}$ est total si l'ensemble des combinaisons linéaires des éléments de $\mathcal{H}$ forme un sous-espace vectoriel dense.

Propriété 4.1.1

Soient $\mathcal{H}$ un sous-ensemble total de $C_0(E)$, (μ_n) une suite de $\mathcal{M}(E)$ et $\mu \in \mathcal{M}(E)$. Si $\forall f \in \mathcal{H}$, i) $\sup_n \mu_n(1) < \infty$ et ii) $\int f d\mu_n \rightarrow \int f d\mu$, alors μ_n converge vers μ faiblement.

Preuve

Soit V le s.e.v. engendré par $\mathcal{H}$. Alors pour tout $f \in C_0(E)$ et tout $g \in V$

$$\begin{aligned} \left| \int f d\mu_n - \int f d\mu \right| &\leq \left| \int f d\mu_n - \int g d\mu_n \right| + \left| \int g d\mu_n - \int g d\mu \right| \\ &\quad + \left| \int g d\mu - \int f d\mu \right| \\ &\leq \left| \int f d\mu_n - \int g d\mu_n \right| + \|f - g\|(\mu_n(1) - \mu(1)) \end{aligned}$$

On peut choisir g assez voisin de f pour que le second terme à droite de la dernière inégalité soit petit puis, pour g fixé, choisir n assez grand pour que le premier terme soit petit.

Propriété 4.1.2

Si $\mu_n \rightarrow \mu$ vaguement et si $\mu_n(1) \rightarrow \mu(1)$ alors, $\mu_n \rightarrow \mu$ étroitement.

Preuve

Soit $f \in C_b(E)$. Alors pour $\varphi \in C_k(E)$ telle que $0 \leq \varphi \leq 1$,

$$\begin{aligned} \left| \int f d\mu_n - \int f d\mu \right| &\leq \left| \int f d\mu_n - \int f\varphi d\mu_n \right| + \left| \int f\varphi d\mu_n - \int f\varphi d\mu \right| \\ &\quad + \left| \int f\varphi d\mu - \int f d\mu \right| \end{aligned}$$

Mais

$$\left| \int f\varphi d\mu - \int f d\mu \right| \leq \int |f|(1-\varphi) d\mu \leq \|f\|(\mu(1) - \mu(\varphi))$$

et d'après les hypothèses,

$$\left| \int f\varphi d\mu - \int f d\mu \right| \leq \overline{\lim} \|f\|(\mu_n(1) - \mu_n(\varphi)) = \|f\|(\mu(1) - \mu(\varphi)).$$

Étant donné $\varepsilon > 0$, on peut choisir $\varphi \in C_k(E)$ telle que $0 \leq \varphi \leq 1$ et $\mu(1) - \mu(\varphi) < \frac{\varepsilon}{3\|f\|}$ (ce qui est toujours possible car si $\varphi \in C_k(\mathbb{R}^d) \nearrow 1$, $\int (1-\varphi) d\mu \rightarrow 0$). Alors $\left| \int f\varphi d\mu - \int f d\mu \right| < \frac{\varepsilon}{3}$ et si $n \geq n_1$, $\left| \int f\varphi d\mu_n - \int f d\mu_n \right| < \frac{\varepsilon}{3}$. Il suffit maintenant de choisir $n \geq n_2$ pour avoir $\left| \int f\varphi d\mu_n - \int f d\mu \right| < \frac{\varepsilon}{3}$ car $f\varphi \in C_k(E)$. D'où pour $n \geq \max(n_1, n_2)$, $\left| \int f\varphi d\mu - \int f d\mu \right| < \varepsilon$.

Corollaire 4.1.1

Si $\mu_n \rightarrow \mu$ vaguement et si $\sup_n \mu_n(1) < +\infty$, alors $\mu_n \rightarrow \mu$ faiblement.

NOTE

Appelons par $\mathcal{M}^a(E)$ l'ensemble des mesures positives sur E de masse $\leq a$. Alors on peut montrer que $\mathcal{M}^a(E)$ muni de la topologie de la convergence étroite est métrisable et si E est compacte, $\mathcal{M}^a(E)$ muni de cette topologie est compact. Aussi, a-t-on le critère de faible compacité suivant :

Corollaire 4.1.2

Si μ_n est une suite de mesures bornées sur $\mathbb{R}^p$ telle que $\sup_n \mu_n(1) \leq A < +\infty$ alors il existe une sous-suite (μ_{n_k}) et une mesure bornée μ telle que $\mu_{n_k} \xrightarrow{n} \mu$ faiblement.

Preuve

Voir dans Dacunha-Castelle T1 page 82.

Rappelons les résultats suivants concernant chaque élément de $\mathcal{M}(E)$ avant d'établir des critères de convergence étroite.

Théorème 4.1.1

Toute mesure finie μ sur $(E, \mathcal{B}_E)$ est régulière c'-à-d. $\forall A \in \mathcal{B}_E$ et $\forall \varepsilon > 0$, il existe un fermé F et un ouvert G de E telles que $F \subset A \subset G$ et $\mu(G - F) < \varepsilon$.

Preuve

Plus tard.

Ce théorème implique alors que la mesure μ est déterminée par les valeurs $\mu(F)$, F fermé parcourant E .

Théorème 4.1.2

Pour tout fermé F de E et tout $\varepsilon > 0$, il existe une fonction $f \in C_b(E)$ telle que $f(x) = 1$ si $x \in F$, $f(x) = 0$ si $d(x, F) > \varepsilon$ et $0 \leq f(x) \leq 1 \forall x \in E$ (f peut être prise uniformément continue).

Preuve

Soit la fonction φ de $\mathbb{R}$ dans lui-même définie par

$$(1) \quad \varphi(t) = \begin{cases} 1 & \text{si } t \leq 0 \\ 1 - t & \text{si } 0 \leq t \leq 1 \\ 0 & \text{si } t \geq 1 \end{cases}$$

Alors la fonction f définie par $f(x) = \varphi\left(\frac{1}{\varepsilon}d(x, F)\right)$ répond à la question.

Graphe de f si $F = [a, b]$:

$$\text{Graph of } f \text{ for } F = [a, b]$$

Dans ce qui suit ∂A désigne le bord de A et une partie A de E telle que $\mu(\partial A) = 0$ est appelée μ -ensemble de continuité. On notera que ∂A est un fermé de E .

Le théorème suivant donne les conditions de convergence étroite.

Théorème 4.1.3 (Portmanteau)

Soit (μ_n) une suite d'éléments de $\mathcal{M}(E)$, $\mu \in \mathcal{M}(E)$. Les propriétés suivantes sont équivalentes :

- (i) $\mu_n \rightarrow \mu$ étroitement.
- (ii) $\lim_n \int f d\mu_n = \int f d\mu$, pour toute fonction réelle f uniformément continue et bornée sur E .
- (iii) $\forall F$ fermé de E , $\mu(F) \geq \overline{\lim}_n \mu_n(F)$.
- (iv) $\forall G$ ouvert de E , $\mu(G) \leq \underline{\lim}_n \mu_n(G)$.
- (v) $\lim_n \mu_n(A) = \mu(A)$ pour tout μ -ensemble de continuité $A \subset E$.

Preuve

On va démontrer le théorème en établissant les implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i), (iii) $\iff$ (iv) et (iii) $\iff$ (v).

(i) $\Rightarrow$ (ii) est évident.

(ii) $\Rightarrow$ (iii) : Soient F un fermé et $\delta > 0$. Pour tout ε arbitrairement petit, comme les ouverts $G = \{x \mid d(x, F) < \varepsilon\}$ décroissent vers F quand $\varepsilon \searrow 0$, on a $\mu(G) < \mu(F) + \delta$ d'après le Théorème 4.1.1. D'autre part, si f est comme dans la démonstration du Théorème 4.1.2, alors f est uniformément continue sur E , $f(x) = 1$ si $x \in F$, $f(x) = 0$ si $x \in G^c$ et $0 \leq f(x) \leq 1$, $\forall x$. D'où

$$\mu_n(F) = \int_F f d\mu_n \leq \int f d\mu_n \quad (4.1)$$

$$\int f d\mu = \int_G f d\mu \leq \mu(G) < \mu(F) + \delta. \quad (4.2)$$

L'hypothèse (ii) ensemble avec (4.1) et (4.2) entraînent

$$\overline{\lim}_n \mu_n(F) \leq \lim_{n \rightarrow +\infty} \int f d\mu_n = \int f d\mu \leq \mu(F) + \delta$$

et comme δ est arbitraire, on a bien (iii).

(iii) $\Rightarrow$ (i) : Soit $f \in C_b(E)$. On va montrer d'abord que

$$\overline{\lim}_n \int f d\mu_n \leq \int f d\mu. \quad (4.3)$$

Par une transformation linéaire de f (avec un coefficient de linéarité strictement positif, on peut se ramener au cas où $0 < f < 1$ (e.g. on fait $f \leftarrow \frac{M-f}{2M}$, avec M telle que $\|f\| \leq M$). Pour un entier positif k fixé, posons

$$F_i = \left\{x \mid \frac{i}{k} \leq f(x)\right\} \quad i = 0, 1, \dots, k. \quad (4.4)$$

F_i est fermé pour tout i et comme $0 < f < 1$, on a pour toute mesure μ

$$\sum_{i=1}^k \frac{i-1}{k} \mu \left\{x \mid \frac{i-1}{k} \leq f(x) < \frac{i}{k}\right\} \leq \int f d\mu \quad (4.5)$$

$$< \sum_{i=1}^k \frac{i}{k} \mu \left\{x \mid \frac{i-1}{k} \leq f(x) < \frac{i}{k}\right\}. \quad (4.6)$$

La somme à droite de la dernière inégalité est égale à

$$\sum_{i=1}^k \frac{i}{k} \{\mu(F_{i-1}) - \mu(F_i)\} = \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \mu(F_i) \quad (4.7)$$

et une transformation similaire de la somme à gauche de (4.5) entraîne

$$\frac{1}{k} \sum_{i=1}^k \frac{i}{k} \mu(F_i) \leq \int f d\mu < \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \mu(F_i). \quad (4.8)$$

Si donc (iii) est vérifiée alors $\overline{\lim}_n \mu_n(F_i) \leq \mu(F_i)$ pour tout i , d'où en appliquant l'inégalité à droite de (4.8) à μ_n et celle de gauche à μ , on a

$$\overline{\lim}_n \int f d\mu_n \leq \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \overline{\lim}_n \mu_n(F_i) \leq \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \mu \left\{f \geq \frac{i}{k}\right\} \leq \frac{1}{k} + \int f d\mu.$$

En faisant tendre k vers $+\infty$, on établit bien (4.3). Enfin, on applique (4.3) à $-f$ et obtient

$$\underline{\lim} \int f d\mu_n \geq \int f d\mu. \quad (4.9)$$

(4.3) et (4.9) établissent bien le (i).

(iii) $\Leftrightarrow$ (iv) résulte facilement du passage au complémentaire.

(iii) $\Rightarrow$ (v) : Soit $A \in \mathcal{B}_E$. Comme $\mathring{A} \subset A \subset \overline{A}$, on a

$$\mu(\overline{A}) \geq \overline{\lim}_n \mu_n(\overline{A}) \geq \overline{\lim}_n \mu_n(A) \geq \underline{\lim}_n \mu_n(A) \geq \underline{\lim}_n \mu_n(\mathring{A}) \geq \mu(\mathring{A}).$$

Si $\mu(\partial A) = 0$ alors $\mu(\mathring{A}) = \mu(A) = \mu(\overline{A})$. En conclusion $\mu_n(A) \rightarrow \mu(A)$.

(v) $\Rightarrow$ (iii) : Comme $\partial \{x/d(x, F) \leq \delta\} \subset \{x/d(x, F) = \delta\}$, les $\partial \{x/d(x, F) \leq \delta\}$ sont distincts pour des δ distincts. Ils sont donc de mesures nulles sauf pour un ensemble dénombrable d'entre eux. Soit alors une suite positive δ_k tendant vers 0 telle que la suite $F_k = \{x/d(x, F) \leq \delta_k\}$ sont de bords de mesures nulles. Alors $\limsup \mu_n(F) \leq \lim \mu_n(F_k) = \mu(F_k)$ pour chaque k . Si F est fermé alors $F_k \searrow F$ d'où (iii). Ce qui achève la démonstration du Théorème 4.1.3.

Contre exemple de (i) $\Leftrightarrow$ (v)

Soit $\mu_n = \delta_{\frac{1}{n}}$ et $\mu = \delta_0$. On a $\mu_n \xrightarrow{e} \mu$ mais si $A =]0, 1[$, $\mu_n(A) = 1 \nrightarrow \mu(A) = 0$.

Dans le cas $E = \mathbb{R}^d$ on a les résultats suivants :

Théorème 4.1.4 (Un critère de convergence étroite)

Pour que $\mu_n \rightarrow \mu$ étroitement, il faut et il suffit que pour tout $t \in \mathbb{R}^d$, $\varphi_{\mu_n}(t)$ tende vers $\varphi_\mu(t)$ où $\varphi_v(t)$ désigne la transformée de Fourier de v au point t .

Pour tout $\sigma > 0$, posons

$$g_\sigma(u) = (2\pi\sigma^2)^{-\frac{d}{2}} \exp\left(-\frac{\|u\|^2}{2\sigma^2}\right), \quad \forall u \in \mathbb{R}^d. \quad (4.10)$$

On vérifie immédiatement (voir aussi dans Chapitre 2) que, $e^{-\frac{\|t\|^2}{2}} = \int_{\mathbb{R}^d} e^{i\langle t, u \rangle} g_1(u) du$ d'où,

$$e^{-\frac{\|x-y\|^2}{2\sigma^2}} = \int_{\mathbb{R}^d} g_1(u) e^{i\langle x-y, \frac{u}{\sigma} \rangle} du = \sigma^d \int_{\mathbb{R}^d} g_1(\sigma u) e^{i\langle x-y, u \rangle} du. \quad (4.11)$$

Alors la démonstration du Théorème 4.1.4 est basée sur les lemmes suivants.

Lemme 4.1.1

Soit μ une mesure bornée sur $\mathbb{R}^d$, alors

$$\int g_\sigma(x-y) d\mu(x) = (2\pi)^{-\frac{d}{2}} \int \varphi_\mu(u) g_1(\sigma u) e^{-i\langle u, y \rangle} du.$$

Preuve

Puisque $|g_1(\sigma u)e^{i\langle u, x-y \rangle}| \leq ce^{-\frac{\sigma^2}{2}\|u\|^2}$ et que cette dernière expression est $d\mu(x).du$ -intégrable, on a par le Théorème de Fubini,

$$\int_{\mathbb{R}^d} g_\sigma(x-y)d\mu(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} d\mu(x) \int_{\mathbb{R}^d} g_1(\sigma u)e^{i(x-y)'u} du \quad (4.12)$$

$$= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{ix'u} d\mu(x) \right) g_1(\sigma u)e^{-iy'u} du \quad (4.13)$$

où le signe prime désigne la transposée d'un vecteur.

Lemme 4.1.2

L'espace vectoriel engendré par les applications $x \mapsto g_\sigma(x-y)$ lorsque σ parcourt $\mathbb{R}_+^*$ et $y \in \mathbb{R}^d$ est dense dans $C_0(\mathbb{R}^d)$.

Preuve

En effet on vérifie aisément que cet espace est une sous-algèbre de $C_0(\mathbb{R}^d)$ qui sépare les points (y compris le point à l'infini). On conclut en appliquant le Théorème de Stone-Weierstrass.

Preuve du Théorème 4.1.4

La condition est évidemment nécessaire. Réciproquement on observe d'abord que $\mu_n(1) = \varphi_{\mu_n}(0) \xrightarrow{n \rightarrow \infty} \varphi_\mu(0) = \mu(1)$. Puis d'après le Lemme 4.1.1,

$$\int g_\sigma(x-y)d\mu_n(x) = (2\pi)^{-\frac{d}{2}} \int \varphi_{\mu_n}(u).g_1(\sigma u)e^{-i\langle u, y \rangle} du$$

et d'après le théorème de Lebesgue, cette dernière intégrale tend vers

$$(2\pi)^{-\frac{d}{2}} \int \varphi_\mu(u).g_1(\sigma u)e^{-i\langle u, y \rangle} du = \int g_\sigma(x-y) d\mu(x).$$

On a donc $\int f d\mu_n \rightarrow \int f d\mu$ pour tout $f(x) = g_\sigma(x-y)$, $\sigma > 0$ et $y \in \mathbb{R}^d$. Or l'espace vectoriel engendré par ces fonctions est dense dans $C_0(\mathbb{R}^d)$ (Lemme 4.1.2). On en conclut, utilisant le fait que $\mu_n(1) \xrightarrow{n \rightarrow \infty} \mu(1)$ et donc que $\sup_n \mu_n(1) < +\infty$ et que $\int f d\mu_n \rightarrow \int f d\mu$ pour tout $f \in C_0(\mathbb{R}^d)$, d'après la Proposition 4.1.1 et enfin que μ_n tend vers μ étroitement d'après la Proposition 4.1.2.

Théorème 4.1.5 (Paul Lévy)

Soit (μ_n) une suite de mesures positives finies sur $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$. Alors

- 1) Si $\mu_n \rightarrow \mu$ étroitement alors $\varphi_{\mu_n} \rightarrow \varphi_\mu$ simplement.
- 2) Si $\varphi_{\mu_n} \rightarrow \varphi$ simplement et si φ est continue en 0 alors il existe $\mu \in \mathcal{M}(\mathbb{R}^d)$ tel que $\mu_n \rightarrow \mu$ étroitement et on a $\varphi = \varphi_\mu$.

En particulier $\mu_n \rightarrow \mu$ étroitement si et seulement si $\varphi_{\mu_n} \xrightarrow{s} \varphi_\mu$.

Preuve

Plus tard.

Définition 4.1.2

Soit X_n une suite de v.a. On dit que X_n converge en loi vers une v.a. X et on note $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$, si la loi de probabilités μ_n de X_n converge étroitement vers la loi de probabilité μ de X . Donc

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff \forall f \in C_b(E), \mathbb{E}(f(X_n)) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(f(X)).$$

Note : C'est par abus de langage qu'on dit " X_n tend en loi vers X ". Ce n'est pas la suite de v.a. X_n qui converge mais c'est plutôt la suite de leurs lois. La v.a. X n'est pas définie de façon unique mais seulement sa loi.

4.2 Le Théorème Centrale Limite

Théorème 4.2.1 (TCL)

Soit $X_1, X_2, \dots, X_n$ une suite de v.a. indépendantes et identiquement distribuées (i.i.d.) à valeurs dans $\mathbb{R}^d$ telle que $\mathbb{E}\|X_1\|^2 < +\infty$. Alors en posant $m = \mathbb{E}(X_1)$, $D = \text{Var}(X)$ (vecteur espérance et matrice de dispersion) et $S_n = X_1 + X_2 + \dots + X_n$, la loi de $\frac{S_n - nm}{\sqrt{n}}$ converge étroitement vers la loi $\mathcal{N}_d(0, D)$.

En particulier si $d = 1$, en posant $\sigma^2 = \text{Var}(X_1) = \mathbb{E}(X_1 - m)^2$ on a

$$P\left\{a \leq \frac{S_n - nm}{\sigma \sqrt{n}} \leq b\right\} \xrightarrow[n \rightarrow +\infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

Preuve

Soit $\varphi(u)$ la fonction caractéristique (f.c.) de $X_1 - m$. D'après la Proposition 5.1.2 φ est de classe C^2 et $\frac{\partial \varphi}{\partial u_i}(0) = 0, i = 1, \dots, d$ et $\left[\frac{\partial^2 \varphi}{\partial u_i \partial u_j}(0), i, j = 1, \dots, d\right] = -D$. D'où d'après la formule de Taylor-Young,

$$\varphi(u) = 1 - \frac{1}{2} u' D u + \|u\|^2 \varepsilon(u) \text{ avec } \lim_{\|u\| \rightarrow 0} \varepsilon(u) = 0.$$

Or la f.c. $\varphi_n(u)$ de $\frac{S_n - nm}{\sqrt{n}}$ est $\left(\varphi\left(\frac{u}{\sqrt{n}}\right)\right)^n$. On a donc (en passant par log),

$$\lim_{n \rightarrow +\infty} \varphi_n(u) = \lim_{n \rightarrow +\infty} \left[1 - \frac{1}{2n} [u' D u + 2\|u\|^2 \varepsilon\left(\frac{u}{\sqrt{n}}\right)]\right]^n = \exp\left(-\frac{1}{2} u' D u\right)$$

d'où le résultat puisque $\exp\left(-\frac{1}{2} u' D u\right)$ est la f.c. de la loi $\mathcal{N}_d(0, D)$.

Remarque 4.2.1

L'hypothèse $\mathbb{E}\|X_1\| < \infty$ est essentielle. Sans elle il peut y avoir convergence vers d'autres lois que la loi normale.

Exemple d'application

$X \rightsquigarrow \mathcal{B}(n, p)$ et on se propose d'évaluer $P\left[\left|\frac{X}{n} - p\right| > \alpha\right]$. Comme X peut s'écrire $X = X_1 + \dots + X_n$ avec $X_1, \dots, X_n$ i.i.d. $\mathcal{B}(1, p)$, $\mathbb{E}(X_1) = p$ et $\text{Var}(X_1) = p(1-p)$. on a d'après le TCL $\frac{X - np}{\sqrt{np(1-p)}} \rightsquigarrow \mathcal{N}(0, 1)$ pour n assez grand. D'où

$$P\left[\left|\frac{X}{n} - p\right| > \alpha\right] = P\left[\frac{|X - np|}{\sqrt{np(1-p)}} > \frac{\alpha \sqrt{n}}{\sqrt{p(1-p)}}\right] \cong \frac{2}{\sqrt{2\pi}} \int_{\alpha_n}^{+\infty} e^{-\frac{t^2}{2}} dt$$

avec $\alpha_n = \frac{\alpha \sqrt{n}}{\sqrt{p(1-p)}}$ (ceci, en utilisant la symétrie par rapport à 0 de la loi $\mathcal{N}(0, 1)$).

Une table de la loi $\mathcal{N}(0, 1)$ montre que $\frac{1}{\sqrt{2\pi}} \int_u^{+\infty} e^{-\frac{t^2}{2}} dt \neq 0.025$ si $u = 1,96$. D'où pour $n = 400$ et $p = \frac{1}{2}$, $P\left[\left|\frac{X}{400} - \frac{1}{2}\right| > 0.05\right] \cong 0.05$.

Lorsque les v.a. X_n ont des densités régulières, on peut renforcer le Théorème 4.2.1.

Théorème 4.2.2 (Théorème de la limite locale)

Soit $X_1, \dots, X_n$ une suite de v.a. i.i.d. à valeurs dans $\mathbb{R}^d$ de f.c. φ_{X_1} telles que $\mathbb{E}\|X_1\|^2 < +\infty$ et $\varphi_{X_1} \in L^1$. Alors D est inversible, les v.a. $\frac{S_n - nm}{\sqrt{n}}$ ont des densités continues et ces densités convergent uniformément sur $\mathbb{R}^d$ vers la densité de la loi $\mathcal{N}_d(0, D)$.

La démonstration du Théorème 4.2.2 est basée sur les lemmes suivants :

Lemme 4.2.1

Soit X un v.a. à valeurs dans $\mathbb{R}^d$ tel que $\mathbb{E}\|X\|^2 < \infty$ et $\varphi = \varphi_X \in L^2$. Alors X a une densité continue, $D(X)$ est inversible et

(i) Il existe $\rho > 0, a > 0$ tel que $|\varphi(t)| \leq e^{-at^2}$ pour $|t| \leq \rho$,

(ii) Pour tout $\rho > 0$, $\sup_{|t| \geq \rho} |\varphi(t)| < 1$.

Preuve

X a une densité continue d'après le Théorème d'inversion. Si $D = D(X)$ n'était pas inversible, il existerait $t \neq 0$ tel que $t'Dt = 0$, d'où $\mathbb{E}(t'X)^2 = 0$ i.e. $t'X = 0$ p.s. et X ne peut avoir de densité. Remarquons que ceci entraîne l'existence de $k > 0$ et $K > 0$ tels que

$$k\|t\|^2 \leq t'Dt \leq K\|t\|^2 \text{ pour tout } t \text{ (e.g. } k = \inf S_p(D) \text{ et } K = \rho(D)). \quad (4.14)$$

(i) Posons $\psi(t) = e^{-a\|t\|^2} \varphi(t)$ où a sera précisé ultérieurement. Un calcul facile montre que $\psi(0) = 1$, $\frac{\partial \psi}{\partial t_i}(0) = 0$, $\left[\frac{\partial^2 \psi}{\partial t_i \partial t_j}(0), i, j\right] = aI - D$ et de (4.14) on en déduit l'existence de $a > 0$, $m > 0$ et $M > 0$ telles que

$$m\|t\|^2 \leq t'(aI - D)t \leq M\|t\|^2 \text{ pour tout } t \in \mathbb{R}^d. \quad (4.15)$$

Alors $\psi(t) = 1 - \frac{1}{2}t'(aI - D)t + \|t\|^2 \lambda(t)$ avec $|\lambda(t)| \rightarrow 0$ si $\|t\| \rightarrow 0$ et pour $M\|t\|^2 < 1$,

$$|\psi(t)| \leq \left|1 - \frac{1}{2}t'(aI - D)t + \|t\|^2 \lambda(t)\right| \leq 1 - \|t\|^2 \left(\frac{m}{2} - \lambda(t)\right)$$

et donc $|\psi(t)| < 1$ pour $\|t\| \leq \rho$.

(ii) Remarquons d'abord que pour $t \neq 0$, $|\varphi(t)| < 1$; sinon il existerait $s \in \mathbb{R}$ tel que $\mathbb{E}(e^{it'X}) = e^{is}$ donc $\mathbb{E}(e^{i(t'X-s)} - 1) = 0$ et en prenant la partie réelle, $\mathbb{E} \cos((t'X - s)) = 1$ i.e. $\cos(t'X - s) = 1$ p.s., soit $X \in \{x / t'x = s + 2k\pi\}$. Mais cet ensemble est de mesure nulle, ce qui contredit le fait que X ait une densité. On a donc pour tout $R > 0, \rho > 0$, $\sup \{|\varphi(t)|, \rho \leq \|t\| \leq R\} < 1$. On conclut grâce au lemme suivant.

Lemme 4.2.2

$|\varphi(t)| \rightarrow 0$ quand $t \rightarrow +\infty$.

Preuve

Il s'agit de montrer que $\varphi(t) = \int e^{it'x} f(x) dx \longrightarrow 0$ si $\|t\| \rightarrow +\infty$ pour $f \in L^1$.

Si $f \in C_K \cap C^1$, c'est un calcul facile. Ensuite, on observe que $|\hat{f}(t) - \hat{g}(t)| \leq \|f - g\|_1$ (où $\hat{f}$ est la transformée de Fourier de $f : \hat{f}(t) = \int e^{it'x} f(x) dx$) et $C_K \cap C^1$ est dense dans L^1 .

Preuve du Théorème 4.2.2

1- Ceci étant, soit $\varphi(t)$ la f.c. de $X_1 - \mathbb{E}(X_1)$. Comme $\varphi(t) = e^{-it'm} \varphi_{X_1}(t) \in L^1$, on déduit du Lemme 4.2.1 que $X_1 - \mathbb{E}(X_1)$ a une densité continue $f(x)$, donc $S_n - nm$ a pour densité $f * f * \dots * f$ et donc $\frac{1}{\sqrt{n}}(S_n - nm)$ a une densité continue et que D est inversible.

2- Montrons d'abord que la f.c. de $\frac{S_n - nm}{\sqrt{n}}$, à savoir $\psi_n(t) = \left(\varphi\left(\frac{t}{\sqrt{n}}\right)\right)^n$ tend vers $\exp\left(-\frac{1}{2}t'Dt\right)$ dans L^1 . Pour cela on écrit

$$\psi_n(t) = \psi_n(t) \mathbb{1}_{\{\|t\| \leq \rho \sqrt{n}\}} + \psi_n(t) \mathbb{1}_{\{\|t\| > \rho \sqrt{n}\}}$$

le ρ étant fourni par le Lemme 4.2.1 (i). Alors en chaque t ,

$$\psi_n(t) \mathbb{1}_{\{\|t\| \leq \rho \sqrt{n}\}} \xrightarrow{n \rightarrow \infty} \exp\left(-\frac{1}{2}t'Dt\right) \quad (\text{d'après le T.C.L.}) \quad \text{et}$$

$$\left|\varphi\left(\frac{t}{\sqrt{n}}\right)\right|^n \mathbb{1}_{\{\|t\| \leq \ell \sqrt{n}\}} \leq \left\{\exp\left(-a\left\|\frac{t}{\sqrt{n}}\right\|^2\right)\right\}^n = \exp(-a\|t\|^2) \in L^1 \quad (\text{Lemme 4.2.1})$$

Donc $\psi_n(t) \mathbb{1}_{\{\|t\| \leq \rho \sqrt{n}\}} \xrightarrow{n \rightarrow \infty} \exp\left(-\frac{1}{2}t'Dt\right)$ dans L^1 .

D'autre part, si $c = \sup \{|\varphi(t)| ; \|t\| \geq \rho\} < 1$ (Lemme 4.2.1 (ii)), on a

$$\int_{\{\|t\| \leq \ell \sqrt{n}\}} \left|\varphi\left(\frac{t}{\sqrt{n}}\right)\right|^n dt \leq c^{n-1} \int_{\mathbb{R}^d} \left|\varphi\left(\frac{t}{\sqrt{n}}\right)\right|^n dt = c^{n-1} n^{\frac{d}{2}} \int_{\mathbb{R}^d} |\varphi(t)| dt \xrightarrow{n \rightarrow \infty} 0$$

donc $\psi_n(t) \mathbb{1}_{\{\|t\| > \ell \sqrt{n}\}} \xrightarrow{n \rightarrow \infty} 0$ dans L^1 et on a le résultat cherché.

3- Mais alors en notant par $f_n(x)$ la densité de $\frac{1}{\sqrt{n}}(S_n - nm)$ et par $g(x)$ celle de $\mathcal{N}_d(0, D)$, on a grâce au Théorème d'inversion

$$f_n(x) - g(x) = (2\pi)^{-d} \int e^{-it'x} \left[\psi_n(t) - \exp\left(-\frac{1}{2}t'Dt\right) \right] dt$$

d'où

$$\|f_n - g\|_\infty \leq (2\pi)^{-d} \int \left| \psi_n(t) - \exp\left(-\frac{1}{2}t'Dt\right) \right| dt \xrightarrow{n \rightarrow \infty} 0.$$

(Voir dans Rényi page 421 pour un énoncé simplifié dans le cas $d = 1$).

Chapitre 5

CONVERGENCES PRESQUE SÛRE DES SUITES ET SÉRIES DE V.A., LOI DES GRANDS NOMBRES

Soient $X_1, X_2, \dots, X_n, \dots$ une suite de v.a. à valeurs dans $\mathbb{R}^d$ définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$. On notera par $|x|$ la norme sur $\mathbb{R}^d$.

5.1 Les quatre types de convergences

Définition 5.1.1

- 1- On dit que (X_n) tend vers X en probabilité et on note $X_n \xrightarrow[n \rightarrow \infty]{P} X$ si $\forall \varepsilon > 0$,
$$\lim_{n \rightarrow \infty} P[|X_n - X| > \varepsilon] = 0.$$
- 2- (X_n) tend vers X presque sûrement et on note $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ s'il existe $\Omega_0 \subset \Omega$, $P(\Omega_0) = 0$ tel que
$$\forall \omega \notin \Omega_0, X_n(\omega) \xrightarrow[n \rightarrow \infty]{} X(\omega).$$

En d'autres termes, si $\forall \varepsilon > 0$ $P[\limsup_{n \rightarrow \infty} |X_n - X| > \varepsilon] = 0$.
- 3- (X_n) tend vers X en moyenne d'ordre r et on note $X_n \xrightarrow[n \rightarrow \infty]{L^r} X$ si $X_n, X \in L^r(\Omega, \mathcal{A}, P)$ et
$$E(|X_n - X|^r) \xrightarrow[n \rightarrow \infty]{} 0.$$
- 4- (X_n) tend vers X en loi et on note $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X$ si $\forall x$, $P[X = x] = 0$,
$$\lim_{n \rightarrow \infty} P[X_n < x] = P[X < x].$$

Théorème 5.1.1

Si $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ alors $X_n \xrightarrow[n \rightarrow \infty]{P} X$.

Preuve

Par hypothèse, il existe $\Omega_0 \subset \Omega$, $P(\Omega_0) = 0$ tel que $\forall \omega \notin \Omega_0$, $\forall \eta > 0$, $\exists n \in \mathbb{N}$ tel que $\forall k \geq n$, $|X_k(\omega) - X(\omega)| < \eta$ ou encore, $\forall \eta > 0$, $\forall \omega \notin \Omega_0$, $\exists n \in \mathbb{N}$ tel que $\omega \in A_n^\eta = \{\omega \mid \forall k \geq n, |X_k(\omega) - X(\omega)| \leq \eta\}$. Mais on peut écrire $A_n^\eta = \bigcap_{k \geq n} \{|X_k - X| \leq \eta\}$. Donc $\eta > 0$ étant donné, $\Omega_0^c \subset \bigcup_{n=1}^{\infty} A_n^\eta$ donc $\bigcap_{n=1}^{\infty} A_n^{\eta^c} = \bigcap_{n=1}^{\infty} \left\{ \bigcup_{k \geq n} [|X_k - X| > \eta] \right\} \subset \Omega_0$ d'où $0 \leq P\left(\bigcap_{n=1}^{\infty} A_n^{\eta^c}\right) \leq P(\Omega_0) = 0$. La suite d'ensemble $A_n^{\eta^c}$, décroissante est donc d'intersection de probabilité nulle. D'où $\lim_{n \rightarrow \infty} P[A_n^{\eta^c}] = 0$. Or $[|X_n - X| > \eta] \subset A_n^{\eta^c}$. Donc $\forall \eta > 0$, $0 \leq \lim_{n \rightarrow \infty} P[|X_n - X| > \eta] \leq \lim_{n \rightarrow \infty} P[A_n^{\eta^c}] = 0$ et (X_n) tend vers X en probabilité.

Théorème 5.1.2

Si $X_n \xrightarrow[n \rightarrow \infty]{P} X$ alors il existe une sous-suite (X_{n_k}) extraite de la suite (X_n) telle que $X_{n_k} \xrightarrow[k \rightarrow \infty]{p.s.} X$.

Preuve

Soit $\varepsilon_n > 0$ tel que $\sum \varepsilon_n < \infty$. Par hypothèse, il existe un entier n_1 tel que $\forall n \geq n_1, P[|X_n - X| \geq \varepsilon_1] < \varepsilon_1$ et par récurrence, $\forall k \geq 2$ il existe $n_k > n_{k-1}$ tel que $P[|X_{n_k} - X| \geq \varepsilon_k] < \varepsilon_k$. Alors pour tout $k \geq 1$, $P[|X_{n_{k+1}} - X_{n_k}| \geq \varepsilon_k] \leq P[|X_{n_{k+1}} - X| \geq \frac{1}{2}\varepsilon_k] + P[|X_{n_k} - X| \geq \frac{1}{2}\varepsilon_k] \leq \varepsilon_k$.

Posons $A_k = [|X_{n_{k+1}} - X_{n_k}| \geq \varepsilon_k]$. Alors $\sum_{k \geq 0} P[A_k] < \sum_{k \geq 0} \varepsilon_k < \infty$.

D'après le lemme de Borel-Cantelli, $P[A_n \text{ i.o.}] = 0$ ou encore $P[A_n \text{ i.o.}]^c = 1$.

Soit $\varepsilon > 0$ et $\omega \in [A_n \text{ i.o.}]^c$. Alors il existe un entier k_0 tel que pour tout $k \geq k_0$, $|X_{n_k}(\omega) - X_{n_{k+1}}(\omega)| < \varepsilon_k$. Soit $k_1 > k$ tel que $\sum_{k \geq k_1} \varepsilon_k < \varepsilon$.

Alors pour tout $p > q > k_1$ on a :

$$|X_{n_p}(\omega) - X_{n_q}(\omega)| \leq \sum_{m=q}^{p-1} |X_{n_m}(\omega) - X_{n_{m+1}}(\omega)| \leq \sum_{m=q}^{p-1} \varepsilon_m < \sum_{m=k_1}^{\infty} \varepsilon_m < \varepsilon \text{ ce qui prouve que } X_{n_k} \xrightarrow[k \rightarrow \infty]{p.s.} X.$$

Théorème 5.1.3

Si $X_n \xrightarrow[n \rightarrow \infty]{P} X$ alors $\mathcal{L}(X_n) \xrightarrow[n \rightarrow \infty]{} \mathcal{L}(X)$.

Preuve

Il s'agit de montrer que pour toute fonction continue φ tendant vers 0 à l'infini, $\lim_n \int \varphi(x) P_{X_n}(dx) = \int \varphi(x) P_X(dx)$.

Soient φ satisfaisant aux hypothèses ci-dessus. Alors φ est bornée et nécessairement uniformément continue sur $\mathbb{R}^d$. Soit $\delta > 0$ arbitraire. Alors

$$\begin{aligned} \left| \int (\varphi(X) - \varphi(X_n)) dP \right| &\leq \int_{|X-X_n| \geq \delta} |\varphi(X) - \varphi(X_n)| dP + \int_{|X-X_n| < \delta} |\varphi(X) - \varphi(X_n)| dP \\ &\leq 2\|\varphi\| P[|X - X_n| \geq \delta] + \int_{|X-X_n| < \delta} |\varphi(X) - \varphi(X_n)| dP. \end{aligned}$$

où $\|\varphi\| = \sup_{x \in \mathbb{R}^d} |\varphi(x)|$. Soit $\varepsilon > 0$. φ étant uniformément continue, on peut choisir $\delta > 0$ tel que

$$\forall n, |X - X_n| < \delta \implies |\varphi(X) - \varphi(X_n)| < \frac{\varepsilon}{2}.$$

La convergence en probabilité entraîne pour le δ choisi, l'existence de N tel que

$$\forall n \geq N, P[|X - X_n| > \delta] \leq \frac{\varepsilon}{4\|\varphi\|}.$$

En définitive, $\forall \varepsilon > 0$, $\exists N$ tel que $\forall n > N$, $\left| \int (\varphi(X) - \varphi(X_n)) dP \right| < \varepsilon$.

Donc $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X$.

Théorème 5.1.4

Si $X_n \xrightarrow[n \rightarrow \infty]{L^r} X$ alors $X_n \xrightarrow[n \rightarrow \infty]{P} X$.

Preuve

Par hypothèse $\lim_{n \rightarrow +\infty} \int |X_n - X|^r dP = 0$. Soit alors $\delta > 0$. On a

$$\begin{aligned} \int |X_n - X|^r dP &\geq \int_{|X-X_n| \geq \delta} |X_n - X|^r dP + \int_{|X-X_n| < \delta} |X_n - X|^r dP \\ &\geq \delta^r P[|X - X_n| \geq \delta] + \int_{|X-X_n| < \delta} |X_n - X|^r dP \\ &\geq \delta^r P[|X - X_n| \geq \delta] \end{aligned}$$

d'où $\lim_{n \rightarrow +\infty} P[|X - X_n| \geq \delta] \leq \lim_{n \rightarrow +\infty} \frac{1}{\delta^r} \int |X_n - X|^r dP = 0$ et $X_n \xrightarrow[n \rightarrow \infty]{P} X$.

Définition 5.1.2

On dit que la suite (X_n) est équi-intégrable (ou uniformément intégrable)

si $\forall \varepsilon > 0, \exists M > 0$ tel que $\forall n \in \mathbb{N}, \int_{\{|X_n| > M\}} |X_n| dP < \varepsilon$

ou encore si $\lim_{a \rightarrow +\infty} \sup_{n \in \mathbb{N}} \int_{\{|X_n| > a\}} |X_n| dP = 0$.

Remarque 5.1.1

Si Y est une v.a. intégrable alors pour tout $\varepsilon > 0$, il existe toujours $M > 0$ tel que $\int_{\{|Y| > M\}} |Y| dP < \varepsilon$.

En effet si $Y \in L^1$, Y est fini p.p. et pour presque tout $\omega \in \Omega$, il existe $n \in \mathbb{N}^*$ tel que $n \geq |Y(\omega)|$ d'où $P\left[\bigcap_n \{|Y| > n\}\right] = 0$. Les $\{|Y| > n\}$ formant une suite décroissante d'ensembles dont la probabilité tend vers 0, on a bien

$$\lim_n \int_{\{|Y| > n\}} |Y| dP = \lim_n \int \mathbb{1}_{\{|Y| > n\}} |Y| dP = 0 \text{ d'où } \forall \varepsilon > 0, \exists n \text{ tel que } \int_{\{|Y| > n\}} |Y| dP \leq \varepsilon.$$

Lemme 5.1.1

Soient ν et P deux mesures positives et bornées telle ν soit absolument continue par rapport à P . Alors $\forall \varepsilon > 0, \exists \eta > 0$ tel que $P(F) < \eta \implies \nu(F) < \varepsilon$ (1)

Preuve

Remarquons d'abord que si les hypothèses du Lemme sont vérifiées alors $\forall F \in \mathcal{A}, P(F) = 0 \implies |\nu|(F) = 0$ (2)

où $|\nu|$ est la variation totale de ν . (Rappelons que l'on a $\nu = \nu^+ - \nu^-$ et $|\nu| = \nu^+ + \nu^-$ où ν^+ et ν^- sont les variations positives et négatives de ν). En effet, soit Ω^+ l'ensemble portant ν^+ et Ω^- l'ensemble portant ν^- . Pour tout F tel que $P(F) = 0$, on a $P(F \cap \Omega^+) = 0$ et $P(F \cap \Omega^-) = 0$. D'où $|\nu|(F) = \nu(F \cap \Omega^+) + \nu(F \cap \Omega^-) = 0$.

Montrons alors que (2) $\implies$ (1). Si (1) était faux, (2) le serait. En effet supposons qu'il existe $\varepsilon > 0$ tel que pour tout n , il existe $F_n \in \mathcal{A}$ tel que $P(F_n) < \frac{1}{2^n}$ et $|\nu|(F_n) \geq \varepsilon$. On aurait ainsi

$$P\left(\limsup F_n\right) = P\left(\lim_n \left[\bigcup_{p \geq n} F_p\right]\right) = \lim_n P\left[\bigcup_{p \geq n} F_p\right] \leq \lim_n \sum_{p=n}^{\infty} P[F_p] < \lim_n \frac{1}{2^{n-1}} = 0 \text{ alors que } |\nu|\left(\limsup F_n\right) = \lim_n |\nu|\left(\bigcup_{p \geq n} F_p\right) \geq \varepsilon.$$

Il existerait donc un $F \in \mathcal{A}$ (e.g. $F = \limsup F_n$) tel que $P(F) = 0$ et $|\nu|(F) \geq \varepsilon$ et (2) serait faux.

Théorème 5.1.5

Si $X_n \xrightarrow[n \rightarrow \infty]{L^1} X$ alors (X_n) est équi-intégrable.

Preuve

Soit (X_n) tel que $X_n \xrightarrow[n \rightarrow \infty]{L^1} X$ et soit $M_0 > 0$ arbitraire. Alors

$$(3) \int_{\{|X_n| \geq M_0\}} |X_n| dP \geq \int_{\{|X_n| \geq M_0\}} |X_n - X| dP + \int_{\{|X_n| \geq M_0\}} |X| dP$$

d'où $M_0 P\left[\left|X_n\right| \geq M_0\right] \leq \int_{\{|X_n| \geq M_0\}} |X_n| dP \leq \int |X_n - X| dP + \int_{\{|X_n| \geq M_0\}} |X| dP$. Or $\lim_n \int |X_n - X| dP = 0$ et $X \in L^1$. Donc le dernier terme de la suite des inégalités ci-dessus est borné par un nombre

$$K < \infty \text{ et alors } P\left[\left|X_n\right| \geq M_0\right] \leq \frac{K}{M_0}.$$

Soit maintenant $\varepsilon > 0$. Si on applique le Lemme 5.1.1 à la mesure de densité $|X|$ par rapport à P , on peut trouver M_0 tel que $P\left[\left|X_n\right| \geq M_0\right] \leq \frac{K}{M_0} < \eta$ et par suite $\int_{\{|X_n| \geq M_0\}} |X| dP < \varepsilon$. D'après

la convergence dans L^1 , $\forall \varepsilon > 0$, il existe n_ε tel que $\forall n \geq n_\varepsilon$,

$\int |X_n - X| dP \leq \varepsilon$. Alors de (3) il vient : $\forall \varepsilon > 0, \forall n \geq n_\varepsilon, \int_{\{|X_n| \geq M_0\}} |X_n| dP \leq 2\varepsilon$. Or d'après la remarque précédente pour $k < n_\varepsilon$ il existe M_k tel que $\int_{\{|X_k| \geq M_k\}} |X_k| dP \leq 2\varepsilon$. Si on pose $M = \max(M_0, M_1, \dots, M_{n_\varepsilon})$, on aura $\forall \varepsilon > 0, \exists M > 0$ tel que $\forall n \int_{\{|X_n| \geq M\}} |X_n| dP \leq 2\varepsilon$. Cqfd.

Théorème 5.1.6

Si $X_n \xrightarrow[n \rightarrow \infty]{P} X$ et (X_n) est équi-intégrable alors $X_n \xrightarrow[n \rightarrow \infty]{L^1} X$.

Preuve

Soit (X_n) équi-intégrable et convergente en probabilité vers X . Alors $\int |X| dP \leq \liminf \int |X_n| dP \leq \sup_n \int |X_n| dP < \infty$ d'après le lemme de Fatou.

Soit $\delta > 0$ arbitraire. $\int |X - X_n| dP = \int_{\{|X - X_n| \geq \delta\}} |X - X_n| dP + \int_{\{|X - X_n| < \delta\}} |X - X_n| dP$.

Or $\int_{\{|X - X_n| < \delta\}} |X - X_n| dP \leq \delta P[|X - X_n| < \delta] \leq \delta$ et

$\int_{\{|X - X_n| \geq \delta\}} |X - X_n| dP \leq \int_{\{|X - X_n| \geq \delta\}} |X| dP + \int_{\{|X - X_n| \geq \delta\}} |X_n| dP$.

Alors soit $M > 0$. On a

$\int |X - X_n| dP \leq \delta + \int_{\{|X - X_n| \geq \delta\}} |X| dP + \int_{\{|X - X_n| \geq \delta, |X_n| \geq M\}} |X_n| dP + \int_{\{|X - X_n| \geq \delta, |X_n| < M\}} |X_n| dP$ d'où

$$\int |X - X_n| dP \leq \delta + \int_{\{|X - X_n| \geq \delta\}} |X| dP + \int_{\{|X_n| \geq M\}} |X_n| dP + MP[|X - X_n| \geq \delta]. \quad (5.1)$$

Soit $\varepsilon > 0$. Comme (X_n) est équi-intégrable, il existe $M > 0$ tel que $\forall n, \int_{\{|X_n| \geq M\}} |X_n| dP \leq \frac{\varepsilon}{4}$.

Choisissons alors $\delta \leq \frac{\varepsilon}{4}$. On sait qu'il existe η tel que

$P[|X - X_n| \geq \delta] \leq \eta \implies \int_{\{|X - X_n| \geq \delta\}} |X| dP \leq \frac{\varepsilon}{4}$ (Lemme 5.1.1).

La convergence en probabilité entraîne l'existence d'un n_ε tel que si $n \geq n_\varepsilon$,

alors $P[|X - X_n| \geq \delta] \leq \eta$ et $P[|X - X_n| \geq \delta] \geq \frac{\varepsilon}{4M}$.

Alors de (5.1) il vient $\forall \varepsilon > 0, \exists n_\varepsilon$ tel que $\forall n \geq n_\varepsilon, \int |X - X_n| dP \leq \varepsilon$ et la suite (X_n) converge vers X dans L^1 .

Le lecteur intéressé pourrait démontrer en exercice les autres implications dans le schéma à connaître ci-dessous :

Si (X_n) est équi-intégrable et $X_n \xrightarrow[n \rightarrow \infty]{P} X$ alors $X_n \xrightarrow[n \rightarrow \infty]{L^1} X$

Un des résultats utile en probabilité et statistique est celui qui dit, dans un cas particulier, que si (X_n) et (Y_n) sont deux suites de v.a. telles que $X_n \xrightarrow[n \rightarrow \infty]{P} X$ et $Y_n \xrightarrow[n \rightarrow \infty]{P} Y$ et si f est une fonction continue sur $\mathcal{X} * \mathcal{X}$ alors $f(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{P} f(X, Y)$. Pour établir ce résultat on a :

Théorème 5.1.7

Une condition nécessaire et suffisante pour que $X_n \xrightarrow[n \rightarrow \infty]{P} X$ est que toute sous-suite de (X_n) contienne une sous-suite qui converge presque sûrement vers X .

Preuve

Si $X_n \xrightarrow[n \rightarrow \infty]{P} X$ alors d'après la définition même de la convergence en probabilité, toute sous-suite (X_{n_k}) de (X_n) converge en probabilité vers X . La condition devient alors nécessaire d'après le Théorème 5.1.2.

Pour montrer qu'elle est suffisante, supposons que (X_n) ne tende pas vers X en probabilité. Alors il existe $\varepsilon_0 > 0$ et $\delta_0 > 0$ tel que $\limsup_{k \rightarrow \infty} P[|X_{n_k} - X| \geq \varepsilon_0] = \delta_0 > 0$ et aussi il existe une sous-suite (X_{n_k}) de (X_n) tel que $P(|X_{n_k} - X| \geq \varepsilon_0) \xrightarrow[k \rightarrow \infty]{} \delta_0 > 0$.

Par conséquent, ni (X_{n_k}) ni aucune de ses sous-suites ne converge en probabilité vers X . Donc aucune sous-suite de (X_{n_k}) ne peut converger p.s. vers X .

Théorème 5.1.8

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^q$ une fonction continue sur un borélien B de $\mathbb{R}^2$ tel que $P[X \in B] = 1$. Si $X_n \xrightarrow[n \rightarrow \infty]{P} X$ alors $f(X_n) \xrightarrow[n \rightarrow \infty]{P} f(X)$.

Preuve

Soit $(f(X_{n_j}))$ une sous-suite de la suite $(f(X_n))$. D'après le Théorème 5.1.7, on a seulement besoin de montrer qu'il existe une sous-suite $(f(X_{k_j}))$ de la sous-suite $(f(X_{n_j}))$ qui converge p.s.

Comme $X_{n_j} \xrightarrow[n \rightarrow \infty]{P} X$, il existe une sous-suite (X_{k_j}) de la sous-suite (X_{n_j}) tel que $X_{k_j} \xrightarrow[k \rightarrow \infty]{p.s.} X$. Soit alors $A = \{X_{k_j} \xrightarrow[k \rightarrow \infty]{} X\} \cap \{X \in B\}$. Il est clair que $P(A) = 1$ et $\omega \in A \implies X_{k_j}(\omega) \xrightarrow[k \rightarrow \infty]{} X(\omega) \in B$. f étant continue sur B , pour tout $\omega \in A$, $f(X_{k_j}(\omega)) \xrightarrow[k \rightarrow \infty]{} f(X(\omega))$. En d'autres termes $f(X_{k_j}) \xrightarrow[k \rightarrow \infty]{p.s.} f(X)$ donc $f(X_n) \xrightarrow[n \rightarrow \infty]{P} f(X)$.

Corollaire 5.1.1

Soit (X_n) et (Y_n) deux suites de v.a. telles que $X_n \xrightarrow[n \rightarrow \infty]{P} X$ et $Y_n \xrightarrow[n \rightarrow \infty]{P} Y$.

Si f est une fonction mesurable sur $\mathbb{R}^2$ tel que (X, Y) appartienne p.s. à l'ensemble de continuité de f alors $f(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{P} f(X, Y)$.

La démonstration du corollaire est basée sur le Lemme suivant :

Lemme 5.1.2

Soit X et Y deux v.a. à valeurs dans $\mathbb{R}^p$. Alors $\forall \varepsilon > 0$,

$$P[|X + Y| \geq \varepsilon] \leq P[|X| \geq \frac{\varepsilon}{2}] + P[|Y| \geq \frac{\varepsilon}{2}].$$

Preuve

Vérifions d'abord que $\{|X + Y| \geq \varepsilon\} \subset \{|X| \geq \frac{1}{2}\varepsilon\} \cup \{|Y| \geq \frac{1}{2}\varepsilon\}$. Si $|X(\omega) + Y(\omega)| \geq \varepsilon$ alors d'après l'inégalité triangulaire $|X(\omega)| + |Y(\omega)| \geq \varepsilon$. Or si la somme de deux nombres est supérieure ou égale à ε alors l'un au moins des deux nombres est supérieur ou égale à $\frac{\varepsilon}{2}$. Donc $|X(\omega)| \geq \frac{1}{2}\varepsilon$ ou $|Y(\omega)| \geq \frac{1}{2}\varepsilon$ d'où le résultat. Le lemme en résulte en prenant les probabilités dans chaque membre de l'inclusion.

Preuve du Corollaire 5.1.1

Montrons d'abord que si $X_n \xrightarrow[n \rightarrow \infty]{P} X$, $Y_n \xrightarrow[n \rightarrow \infty]{P} Y$ alors $(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{P} (X, Y)$. Notons par 0 la v.a. égale p.s. à 0. On a alors $|(X_n, Y_n) - (X, Y)| \leq |(X_n, 0) - (X, 0)| + |(0, Y_n) - (0, Y)|$. Comme $|(X_n, 0) - (X, 0)| = |X_n - X|$ et $|(0, Y_n) - (0, Y)| = |Y_n - Y|$, alors $(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{P} (X, Y)$ d'après le Lemme 5.1.2, et le Corollaire 5.1.1 résulte immédiatement du Théorème 5.1.8.

Corollaire 5.1.2

Si $X_n \xrightarrow[n \rightarrow \infty]{P} X$, $Y_n \xrightarrow[n \rightarrow \infty]{P} Y$ alors

- i) $aX_n + bY_n \xrightarrow[n \rightarrow \infty]{P} aX + bY \quad \forall a, b \in \mathbb{R}$
- ii) $\frac{1}{X_n} \xrightarrow[n \rightarrow \infty]{P} \frac{1}{X}$ si $P[X_n \neq 0] = P[X \neq 0] = 1$
- iii) $X_n Y_n \xrightarrow[n \rightarrow \infty]{P} XY$.

5.2 Convergence p.s. de séries de v.a. indépendantes – Théorèmes limites forts

Soit (X_n) est une suite de v.a. et $S_n = X_1 + X_2 + \dots + X_n$. On dira que la série $\sum X_n$ converge p.s. si la suite des sommes partielles (S_n) converge p.s..

On dira que la série $\sum X_n$ converge en loi s'il existe une distribution de probabilité F tel que $F_{S_n} \xrightarrow[n \rightarrow +\infty]{e} F$.

De même $\sum X_n$ converge en probabilité ou en moyenne d'ordre r s'il existe une v.a. X tel que $S_n \xrightarrow[n \rightarrow +\infty]{P} X$ ou $S_n \xrightarrow[n \rightarrow +\infty]{L^r} X$. L'outil fondamental pour les théorèmes limites forts est

Théorème 5.2.1 (Inégalité de Kolmogorov)

Soit $X_1, \dots, X_n$ des v.a. indépendantes de second moment fini. Alors

$$P\left[\max_{1 \leq k \leq n} |S_k - \mathbb{E}(S_k)| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2} \sum_{j=1}^n \text{Var}(X_j).$$

Preuve

Sans perte de généralité, on peut supposer $\mathbb{E}(X_k) = 0$ pour tout k . Posons $A_0 = \Omega$ et $\forall k \geq 1$,

$$A_k = \{\max_{1 \leq j \leq k} |S_j| < \varepsilon\} = \bigcap_{j=1}^k \{|S_j| < \varepsilon\} \quad (5.2)$$

$$B_k = A_{k-1} \cap A_k^c = \left(\bigcap_{j=1}^{k-1} \{|S_j| < \varepsilon\}\right) \cap \{|S_k| \geq \varepsilon\}. \quad (5.3)$$

B_k est l'évènement " S_k est la 1^{ère} somme partielle dont la valeur absolue est supérieure ou égale à ε " et il est facile de vérifier que

$$A_n^c = \{\max_{1 \leq j \leq n} |S_j| \geq \varepsilon\} = \bigcup_{k=1}^n B_k.$$

En outre $B_1, B_2, \dots, B_n$ sont des évènements disjoints et pour tout $1 \leq k \leq n$, on a

$$\begin{aligned} \int_{B_k} S_n^2 dP &= \mathbb{E}(S_n \mathbb{1}_{B_k})^2 = \mathbb{E}((S_n - S_k) \mathbb{1}_{B_k} + S_k \mathbb{1}_{B_k})^2 \\ &= \mathbb{E}((S_n - S_k) \mathbb{1}_{B_k})^2 + 2\mathbb{E}((S_n - S_k) S_k \mathbb{1}_{B_k}) + \mathbb{E}(S_k \mathbb{1}_{B_k})^2. \end{aligned}$$

Comme les v.a. $S_n - S_k$ et $S_k \mathbb{1}_{B_k}$ sont indépendantes et $\mathbb{E}(X_j) = 0$ pour $1 \leq j \leq n$, on a $\mathbb{E}((S_n - S_k) S_k \mathbb{1}_{B_k}) = \mathbb{E}(S_n - S_k) \mathbb{E}(S_k \mathbb{1}_{B_k}) = 0$ d'où $\int_{B_k} S_n^2 dP \geq \mathbb{E}(S_k \mathbb{1}_{B_k})^2 \geq \varepsilon^2 P(B_k)$. Il en résulte que

$$\sum_{k=1}^n \text{Var}(X_k) = \text{Var}(S_n) \geq \int_{A_n^c} S_n^2 dP = \sum_{k=1}^n \int_{B_k} S_n^2 dP = \varepsilon^2 P\left(\bigcup_{k=1}^n B_k\right) \text{ cqfd.}$$

Théorème 5.2.2

Soit (X_n) une suite de v.a. indépendantes telle que $\mathbb{E}(X_k^2) < \infty$ pour tout k . Si $\sum_{k=1}^{\infty} \text{Var}(X_k) < \infty$ alors

$$\sum_{n=1}^{\infty} (X_n - \mathbb{E}(X_n)) \text{ converge p.s.}$$

Preuve

Posons $S_n = \sum_{j=1}^n \mathbb{E}(X_j - \mathbb{E}(X_j))$. Alors $\overline{\lim} S_n$ et $\underline{\lim} S_n$ sont finies p.s. En effet, pour tout $m \geq 1$, $\overline{\lim}_{n \rightarrow \infty} |S_n| \leq \overline{\lim}_{n \rightarrow \infty} |S_n - S_m| + |S_m| \leq \sup_{n \geq m} |S_n - S_m| + |S_m|$ et d'après l'inégalité de Kolmogorov (Théorème 5.2.1) on a

$$\begin{aligned} P\left[\sup_{n \geq m} |S_n - S_m| \geq \varepsilon\right] &\leq P\left[\bigcup_{n=m}^{\infty} \{|S_n - S_m| \geq \varepsilon\}\right] \\ &= \lim_{M \rightarrow +\infty} P\left[\bigcup_{n=m}^M \{|S_n - S_m| \geq \varepsilon\}\right] = \lim_{M \rightarrow +\infty} P\left[\max_{m \leq n \leq M} |S_n - S_m| \geq \varepsilon\right] \\ &\leq \frac{1}{\varepsilon^2} \sum_{n=m}^{\infty} \text{Var}(X_n) < \infty \quad \text{d'après l'hypothèse.} \end{aligned}$$

Donc en faisant tendre ε vers $+\infty$, la 1^{ère} probabilité tend vers 0 et on a montré que $\overline{\lim} S_n$ et de même $\underline{\lim} S_n$ sont finies p.s.

Pour montrer que S_n converge p.s. nous avons seulement besoin de montrer que $\forall \varepsilon > 0$, $P[\limsup S_n - \liminf S_n \geq \varepsilon] = 0$.

Pour le faire, observons d'abord que si (a_n) est une suite de nombres réels alors $\sup |a_n| = \max\{|\sup a_n|, |\inf a_n|\}$ et $|\sup a_n| \geq \varepsilon$ ou $|\inf a_n| \geq \varepsilon$ entraîne $\sup |a_n| \geq \varepsilon$. Donc $\forall m \in \mathbb{N}^*, \forall \varepsilon > 0$,

$$\begin{aligned} \{\limsup S_n - \liminf S_n \geq 2\varepsilon\} &\subset \{\sup_{n \geq m} S_n - \inf_{n \geq m} S_n \geq 2\varepsilon\} \\ &\subset \{|\sup_{n \geq m} S_n - S_m| \geq \varepsilon\} \cup \{|S_m - \inf_{n \geq m} S_n| \geq \varepsilon\} \\ &\subset \{\sup_{n \geq m} |S_n - S_m| \geq \varepsilon\}. \end{aligned}$$

Alors d'après l'inégalité de Kolmogorov

$$P[\limsup S_n - \liminf S_n \geq 2\varepsilon] \leq \frac{1}{\varepsilon^2} \sum_{j=m}^{\infty} \text{Var}(X_j) \xrightarrow{m \rightarrow +\infty} 0$$

d'où le résultat.

Corollaire 5.2.1

Si (X_n) est une suite de v.a. indépendantes telles que $\sum_n \text{Var}(X_n) < \infty$ et $\sum_n \mathbb{E}(X_n) < \infty$ alors $\sum_n X_n$ converge p.s.

Preuve

C'est une conséquence immédiate du Théorème 5.2.2 et du fait que la somme de deux séries convergentes est convergente.

5.3 Loi forte des grands nombres

La loi forte des grands nombres est le nom donné à une classe de théorèmes traitant de la convergence presque sûre d'une suite de moyennes arithmétiques de v.a. Avant d'énoncer le théorème le plus important, commençons par la série de lemmes suivants.

Lemme 5.3.1

Si (a_n) est une suite de nombres réels tendant vers a quand $n \rightarrow +\infty$,

alors $\frac{a_1 + a_2 + \cdots + a_n}{n} \xrightarrow{n \rightarrow +\infty} a$.

Preuve

C'est d'après un simple exercice d'analyse de DEUG connu sous le nom de convergence au sens de Césaro. Le lecteur peut donc s'y référer.

Lemme 5.3.2 (Kronecker)

Si $\sum_{n=1}^{\infty} b_n$ est une série convergente de nombres réels alors $\frac{1}{n} \sum_{k=1}^n kb_k \xrightarrow{n \rightarrow +\infty} 0$.

Preuve

Posons $b = \sum_{k=1}^{\infty} b_k$. D'après le Lemme 5.3.1

$$\frac{b_1 + (b_1 + b_2) + \cdots + (b_1 + b_2 + \cdots + b_n)}{n} \xrightarrow{n \rightarrow +\infty} b.$$

Or

$$\frac{nb_1 + (n-1)b_2 + \cdots + b_n}{n} = \frac{n+1}{n} \sum_{k=1}^n b_k - \frac{1}{n} \sum_{k=1}^n kb_k \xrightarrow{n \rightarrow +\infty} b$$

d'où le résultat.

Lemme 5.3.3

Si (X_n) est une suite de v.a. indépendantes telle que $\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow[n \rightarrow +\infty]{p.s.} X$ alors X est une constante p.s.

Preuve

Pour chaque k fixé $\frac{X_k}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} 0$. Donc l'hypothèse implique

$$\frac{X_k + X_{k+1} + \cdots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} X \quad \forall k$$

et donc X est mesurable par rapport à la tribu asymptotique $\mathcal{A}_\infty$ engendrée par les X_n : $\mathcal{A}_\infty = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \dots)$ d'où le résultat.

Lemme 5.3.4

Si X est une v.a. alors $\mathbb{E}(X)$ existe si, et seulement si $\sum_{n \geq 0} P[|X| \geq n] < \infty$.

Preuve

La conclusion du lemme découle des deux inégalités suivantes :

$$\int |X| dP = \int \left(\sum_{n=1}^{\infty} |X| \mathbb{1}_{\{n-1 \leq |X| < n\}} \right) dP \leq \sum_{n=1}^{\infty} nP[n-1 \leq |X| < n] = \sum_{n=0}^{\infty} P[|X| \geq n]$$

$$\text{et } \int |X| dP \geq \sum_{n=0}^{\infty} nP[|X| \geq n] - 1.$$

Théorème 5.3.1 (Kolmogorov)

Si (X_n) est une suite de v.a. indépendantes tel que $\mathbb{E}(X_n^2) < \infty$ pour tout n et $\sum_{k=1}^{\infty} \frac{\text{Var}(X_k)}{k^2} < \infty$, alors

$$\frac{1}{n} \sum_{k=1}^{\infty} \frac{(X_k - \mathbb{E}(X_k))}{k} \xrightarrow[n \rightarrow +\infty]{p.s.} 0.$$

Preuve

Comme $\sum_{k=1}^{\infty} \frac{\text{Var}(X_k)}{k^2} < \infty$ alors d'après le Théorème 5.2.2 $\sum_{k=1}^{\infty} \frac{(X_k - \mathbb{E}(X_k))}{k}$ converge p.s. et d'après le Lemme 5.3.2 de Kronecker, on peut conclure que $\frac{1}{n} \sum_{k=1}^n (X_k - \mathbb{E}(X_k))$ tend vers 0 p.s.

Théorème 5.3.2 (Loi forte des grands nombres de Kolmogorov)

Soit (X_n) une suite de v.a. i.i.d. Alors $\frac{1}{n} \sum_{k=1}^n X_k$ converge p.s. si, et seulement si $\mathbb{E}(X_1) < \infty$ et dans ce

$$\text{cas } \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = \mathbb{E}(X_1) \text{ p.s.}$$

Preuve

Supposons que $\frac{1}{n} \sum_{k=1}^n X_k$ converge p.s. Posons $S_n = \sum_{k=1}^n X_k$. Alors $\frac{X_n}{n} = \frac{S_n}{n} - \frac{S_{n-1}}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} 0$,

soit $\frac{X_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} 0$. Donc $P\left[\frac{|X_n|}{n} \geq 1 \text{ i.o.}\right] = 0$ ou encore $P[|X_n| \geq n \text{ i.o.}] = 0$. Alors d'après le lemme de Borel-Cantelli, comme les événements $\{|X_n| \geq n\}$ sont indépendants, $\sum_{n=1}^{\infty} P[|X_n| \geq n] < \infty$.

Mais $P[|X_n| \geq n] = P[|X_1| \geq n]$. Donc $\sum_{n=1}^{\infty} P[|X_1| \geq n] < \infty$ et $\mathbb{E}(X_1)$ existe d'après le Lemme 5.3.4.

Réciproquement supposons que $\mathbb{E}(X_1)$ existe. Alors si on note $Y_n = X_n \mathbb{1}_{\{|X_n| < n\}}$, on a : $P[Y_n \neq X_n] = P[|X_n| \geq n] = P[|X_1| \geq n]$ d'où d'après le Lemme 5.3.4, $\sum_{n=1}^{\infty} P[Y_n \neq X_n] < \infty$ ce qui implique $P[Y_n \neq X_n \text{ i.o.}] = 0$ d'après le lemme de Borel-Cantelli. Il suffira donc de montrer que $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = \mathbb{E}(X_1)$ p.s.

Comme $\mathbb{E}(Y_n) = \mathbb{E}[X_n \mathbb{1}_{\{|X_n| < n\}}] \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(X_1)$ (convergence dominée)

alors $\frac{\mathbb{E}(Y_1) + \mathbb{E}(Y_2) + \dots + \mathbb{E}(Y_n)}{n} \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(X_1)$ d'après le Lemme 5.3.1.

Il suffit alors de montrer que $\frac{1}{n} \sum_{k=1}^n (Y_k - \mathbb{E}(Y_k)) \xrightarrow[n \rightarrow +\infty]{p.s.} 0$. D'après le Théorème 5.3.1, il suffirait

de montrer que $\sum_{k=1}^{\infty} \frac{\text{Var}(Y_k)}{k^2} < \infty$.

Comme $\text{Var}(Y_k) = \mathbb{E}(Y_k^2) - (\mathbb{E}(Y_k))^2 \leq \mathbb{E}(Y_k^2)$, il suffirait de montrer tout simplement que

$$\sum_{k=1}^{\infty} \frac{\mathbb{E}(Y_k^2)}{k^2} < \infty. \text{ Or } \mathbb{E}(Y_k^2) = \mathbb{E}(X_1^2 \mathbb{1}_{\{|X_1| < k\}}) \text{ donc on va montrer que } \sum_{k=1}^{\infty} \frac{\mathbb{E}(X_1^2 \mathbb{1}_{\{|X_1| < k\}})}{k^2} < \infty. \text{ D'après}$$

le théorème de Beppo-Levi on pourra montrer seulement que $\int \left(\sum_{k=1}^{\infty} \frac{X_1^2}{k^2} \mathbb{1}_{\{|X_1| < k\}} \right) dP < \infty$. Mais alors pour $m = 1, 2, \dots$ on a

$$\begin{aligned}
\int \left(\sum_{k=1}^{\infty} \frac{X_1^2}{k^2} \mathbb{1}_{\{|X_1| < k\}} \right) \mathbb{1}_{\{m-1 \leq |X_1| \leq m\}} dP &= \int \left(\sum_{k=m}^{\infty} \frac{X_1^2}{k^2} \right) \mathbb{1}_{\{m-1 \leq |X_1| < m\}} dP \\
&\leq \int m^2 \left(\sum_{k=m}^{\infty} k^{-2} \right) \mathbb{1}_{\{m-1 \leq |X_1| < m\}} dP \\
&\leq m^2 \frac{1}{m-1} P[m-1 \leq |X_1| < m] \\
&< 2mP[m-1 \leq |X_1| < m].
\end{aligned}$$

Donc $\int \left(\sum_{k=1}^{\infty} \frac{X_1^2}{k^2} \mathbb{1}_{\{|X_1| < k\}} \right) dP \leq 2 \sum_{m=1}^{\infty} mP[m-1 \leq |X_1| < m] \leq 2(\mathbb{E}|X_1| + 1) < \infty$. Ce qui achève la démonstration.

Chapitre 6

introduction aux processus aléatoires

Dans cette partie du cours, nous nous intéressons à des phénomènes dont les observations dépendent du temps. En effet, le temps joue une part essentielle dans la vie au cours de laquelle beaucoup de quantités se développent de façon aléatoire lorsque le temps passe, entraînant des modèles de calculs de probabilités qui peuvent vite se compliquer. De telles quantités aléatoires dépendant du temps sont appelées processus aléatoires ou stochastiques qui sont de types différents.

De façon formelle, on appelle processus aléatoire une famille de v.a. $X = \{X(t) : t \in T\}$ à valeurs dans un espace $\mathcal{X}$ où T est un ensemble quelconque. Lorsque $T \subset \mathbb{N}$ on dit que le processus X est à temps discret et lorsque T est un intervalle de $\mathbb{R}^+$ on dit que X est un processus à temps continu. Lorsque $\mathcal{X} = \mathbb{R}$, X est dit processus réel. L'application $T \rightarrow \mathcal{X}$, $t \mapsto X(t)$ est appelée trajectoire du processus et pour tout $t \in T$, $\omega \mapsto X(t, \omega)$ est appelé état du processus au temps t . Nous étudierons ici quatre exemples de processus aléatoires courants : le processus de Poisson, la marche aléatoire, le processus de vie et de mort et les chaînes de Markov classiques.

6.1 Processus aléatoires courants

6.1.1 Processus de branchement

Le processus de branchement est un processus qui modélise la croissance d'une population qui s'auto-reproduit. Dans ce qui suit, le terme *nomade* désignera l'objet capable de se reproduire selon les règles suivantes :

- A la date $n = 0$, il n'existe qu'un nomade. Celui-ci vit durant une unité de temps et meurt au temps $n = 1$ en laissant une famille de descendants eux aussi nomades, ayant chacun la même propriété (1ère génération). Chacun des nomades de la 1ère génération vit durant une unité de temps et meurt au temps $n = 2$ en laissant des descendants de la 2ème génération de nomades.

- Ce processus de vie et de mort continue aux temps ultérieurs $n = 3, 4, \dots$

Si pour tout $n \geq 1$, on connaît le nombre Z_n de nomades à la n ème génération, alors on sait quoi dire au sujet du développement de la population des nomades considérés (e.g., étude de la croissance d'une colonie de bactéries en biologie, étude de la probabilité d'extinction d'un nom de famille, etc.)

6.1.2 Processus de Poisson

Supposons qu'à partir de l'instant $t = 0$, on observe un phénomène aléatoire tels qu'un appel téléphonique au standard d'un hôtel ou l'arrivée d'un client dans un magasin.

Soit N_t le nombre de phénomènes observés dans l'intervalle de temps $]0, t[$ avec $N_0 = 0$. $(N_t)_{t \geq 0}$ est à valeurs dans $\mathbb{N}$ vérifiant les hypothèses suivantes :

- 1- $\forall 0 \leq s \leq t, N_s \leq N_t$ ($t \mapsto N_t$ est croissante).
- 2- $\forall 0 \leq s$, le nombre de phénomènes $N_t - N_s$ observés dans l'intervalle $]s, t]$ est indépendant du nombre de phénomènes N_s observés dans l'intervalle $]0, s]$ (on dit que le processus $(N_t)_{t \geq 0}$ est à accroissements indépendants).
- 3- $\forall 0 \leq s$, la loi de $N_t - N_s$ ne dépend que de la longueur de l'intervalle $]s, t]$: $\mathcal{L}(N_t - N_s) = \mathcal{L}(N_{t-s})$ (on dit que le processus $(N_t)_{t \geq 0}$ est à accroissements stationnaires).
- 4- On ne peut observer deux phénomènes ou plus simultanément :
 $P[N_{t+h} - N_t > 1] = o(h)$ quand $h \rightarrow 0$ (on dit que le processus $(N_t)_{t \geq 0}$ est localement continu).

Théorème 6.1.1

Il existe un réel $\lambda > 0$ appelé taux d'arrivée ou intensité du processus telle que

$\forall t \geq 0, N_t \sim \mathcal{P}(\lambda t)$ (loi de Poisson de paramètre λt).

Définition 6.1.1

Le processus $(N_t)_{t \geq 0}$ satisfaisant les hypothèses 1 à 4 est appelé processus de Poisson d'intensité λ .

Preuve

La démonstration du Théorème 3.3.1 est basé sur le lemme suivant :

Lemme 6.1.1

(i) $\forall h \geq 0, P[N_h = 1] = \lambda h + o(h)$.

(ii) $\forall h \geq 0, P[N_h = 0] = 1 - \lambda h + o(h)$.

Preuve

Posons $P_k(t) = P[N_t = k], k \geq 0$.

- Calculons $P_0(t)$: on a $\forall t, h \geq 0$,

$$\begin{aligned} P_0(t+h) &= P[N_{t+h} = 0] = P[N_t = 0, N_{t+h} - N_t = 0] = P[N_t = 0] = P[N_h = 0] \\ &= P_0(t)P_0(h) \end{aligned}$$

D'autre part $P_0(t)$ est une fonction décroissante : si $0 \leq s \leq t, N_s \leq N_t$ et $N_t = 0 \implies N_s = 0$ d'où $P[N_t = 0] \leq P[N_s = 0]$.

$P_0(t)$ vérifie donc

a) $0 \leq P_0(t) \leq 1$, b) $P_0(t+h) = P_0(t)P_0(h) \forall t, h \geq 0$, c) $P_0(t)$ décroissante.

Montrons qu'alors $\forall t \geq 0, P_0(t) = (P_0(1))^t$.

- si $t \in \mathbb{N}$, $P_0(t) = P_0(\underbrace{1 + 1 + \dots + 1}_{n \text{ fois}}) = (P_0(1))^t$ d'après b)

- si $t \in \mathbb{Q}^+$, $t = \frac{p}{q}, p, q \in \mathbb{N}^*$.

$$\text{D'où } qt = p \implies P_0(\underbrace{t + t + \dots + t}_{q \text{ fois}}) = P_0(p) = (P_0(1))^p = (P_0(t))^q.$$

$$\text{D'où } P_0(t) = (P_0(1))^{\frac{p}{q}} = (P_0(1))^t.$$

- si $t \in \mathbb{R}^+ \setminus \mathbb{Q}^+$, il existe (s_n) et (t_n) dans $\mathbb{Q}^+$ tels que $s_n \nearrow t$ et $t_n \searrow t$, d'où $P_0(t_n) \leq P_0(t) \leq P_0(s_n)$ d'après c); soit $(P_0(1))^{t_n} \leq (P_0(1))^t \leq (P_0(1))^{s_n}$.

En passant à la limite quand n tend vers $+\infty$, on obtient $P_0(t) = (P_0(1))^t$

- si $P_0(1)$ était nulle, on aurait $P_0(t) = 0 \forall t \geq 0$, donc aussi $P_0(0) = 0$ c'est-à-dire $P[N_0 = 0] = 0$ ce qui est contraire à l'hypothèse $N_0 = 0$.
- si $P_0(1) = 1$, on aurait $P_0(t) = 1 \forall t \geq 0$, donc aussi $P[N_\infty = 0] = 1$ et $N_\infty = 0$ p.s. Ce qui est impossible car contraire à l'hypothèse $N_t \geq 0$.

On a donc $0 < P_0(t) < 1$. Donc il existe $\lambda \in \mathbb{R}_+^*$ telle que $P_0(t) = e^{-\lambda t} \forall t \geq 0$. Donc $\forall h \geq 0$, $P[N_h = 0] + P[N_h = 1] + P[N_h > 1] = 1$. On a d'après l'hypothèse 4) et (ii), $P[N_h = 1] = 1 - P[N_h = 0] - P[N_h > 1] = \lambda h + o(h)$, d'où le (i).

Preuve du Théorème 3.3.1.

$$\forall k \geq 1, P[N_{t+h} = k] = \sum_{j=0}^k P[N_t = j, N_{t+h} - N_t = k - j] = \sum_{j=0}^k P[N_t = j] P[N_h = k - j]$$

En commençant la sommation par $j = k$,

$$P[N_{t+h} = k] = P_k(t)P_0(h) + P_{k-1}(t)P[N_h = 1] + \sum_{j=0}^{k-2} P[N_t = j] P[N_h = k - j]. \text{ Soit, en tenant compte}$$

de l'hypothèse 4) :

$$P_k(t+h) = P_k(t)(1 - \lambda h + o(h)) + P_{k-1}(t)(\lambda h + o(h)) + o(h) \text{ ou encore}$$

$$\frac{1}{h} [P_k(t+h) - P_k(t)] = \lambda P_{k-1}(t) - \lambda P_k(t) + o(1). \text{ En faisant tendre } h \text{ vers } 0, \text{ on obtient, } P_k(t) \text{ est dérivable à droite et } P'_k = \lambda P_{k-1} - \lambda P_k$$

Or $P_k(0) = P_0(k) = 0 \forall k \geq 1$ et $P_0(t) = e^{-\lambda t}$. On a donc

$$(1) \quad \begin{cases} P'_k = \lambda P_{k-1} - \lambda P_k \\ P_k(0) = 0 \\ P_0(t) = e^{-\lambda t} \end{cases}, \quad \forall k \geq 1$$

Posons $P_k(t) = q_k(t)e^{-\lambda t}$ on a $P'_k(t) = q'_k(t)e^{-\lambda t} - \lambda P_k(t)$. Or $P'_k = \lambda q_{k-1}e^{-\lambda t} - \lambda P_k$. Donc $q'_k(t) = \lambda q_{k-1}(t)$. En résumé (1) équivaut à

$$(2) \quad \begin{cases} P_k(t) = q_k(t)e^{-\lambda t} \\ q'_k(t) = \lambda q_{k-1}(t) \\ q_k(0) = P_k(0) = 0 \\ q_0(t) = 1 \end{cases}, \quad \forall k \geq 1$$

Le système (2) donne $q'_1(t) = \lambda \implies q_1(t) = \lambda t$

$$q'_2(t) = \lambda^2 t \implies q_2(t) = \frac{(\lambda t)^2}{2}$$

et par récurrence : $\forall k \geq 1, q_k(t) = \frac{(\lambda t)^k}{k!}$

$$\text{Soit } P[N_t = 0] = e^{-\lambda t} \text{ et } P[N_t = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad \forall k \geq 1$$

Extension (2^{nde} définition du Processus de Poisson)

Soit T_i l'instant d'arrivée du $i^{\text{ème}}$ phénomène :

$$T_i = \inf \{t > 0 \mid N_t = i\} \quad (1)$$

$$\text{Alors } 0 = T_0 \leq T_1 \leq \dots \leq T_n \leq \dots \text{ et } N_t = i \iff t \in [T_i, T_{i+1}[\quad (2)$$

$(T_n)_{n \geq 0}$ est une suite de v.a. dont les valeurs déterminent complètement le processus N . En effet, si on connaît les T_i alors N_t est défini par (2) et (2) peut se résumer à $N_t = \max \{n \geq 0 \mid$

$$T_n \leq t\} \text{ ou encore } N_t = \sum_{k=0}^{\infty} \mathbb{1}_{\{T_k \leq t\}} \text{ et } \{T_k \leq t\} = \{N_t \geq n\} \quad (4)$$

Les v.a. $X_i = T_i - T_{i-1}$, $i = 1, 2, \dots$, sont appelées durées inter-arrivées du processus.

Théorème 6.1.2

Les durées inter-arrivées d'un Processus de Poisson d'intensité $\lambda > 0$, $X_1, X_2, \dots, X_n, \dots$, sont des v.a. indépendantes et de même loi exponentielle de paramètre λ , $\mathcal{E}(\lambda)$.

Propriété d'absence de mémoire

Une v.a. positive X est dite avoir la propriété d'absence de mémoire si et seulement si, $\forall u, v \geq 0$, $P[X > u + v | X > u] = P[X > v]$.

Exercice 6.1.1

Soit X une v.a.r. positive. Montrer que X a la propriété d'absence de mémoire si et seulement si X est distribuée exponentiellement.

indication :

Pour la condition suffisante, poser $G(u) = P[X > u]$. Alors de (5), on a :

$G(u + v) = G(u) G(v) \forall u, v \geq 0$. En remarquant que $G(u)$ est une fonction décroissante, en déduire que $G(u)$ est de la forme $G(u) = e^{-\lambda u}$, $\lambda \in \mathbb{R}_+^*$.

Preuve du Théorème 3.3.2. – indication

Commencer par étudier la loi de X_1 : on a $P[X_1 > u] = P[N_u = 0] = e^{-\lambda u}$, $\forall u \geq 0$.

Traiter ensuite X_2 puis X_3 etc.

6.1.3 Marche aléatoire sur $\mathbb{Z}$

Une particule se trouve à l'instant $t = 0$ à l'origine 0 d'un axe gradué $0, \pm 1, \pm 2, \dots$

A chaque instant $t \in \{1, 2, 3, \dots\}$ elle avance d'un cran à droite avec probabilité p , $0 < p < 1$ ou recule vers la gauche avec probabilité $q = 1 - p$. Soit $S_0 = 0$ et $\forall n \geq 1$, S_n la position de la particule à la date n . On a

$$S_{n+1} = \begin{cases} S_n + 1 & \text{avec probabilité } p \\ S_n - 1 & \text{avec probabilité } q \end{cases}$$

On suppose que les déplacements d'une position à l'autre sont indépendants. On peut donc écrire $S_n = S_0 + X_1 + \dots + X_n$, $n = 0, 1, 2, \dots$ où S_0 est la position initiale de la particule sur l'axe et $X_1, X_2, \dots$, sont des v.a. indépendantes prenant chacune la valeur $+1$ avec probabilité p et la valeur -1 avec probabilité q .

Le processus $(S_n)_{n \geq 0}$ est appelé marche aléatoire simple. Elle est dite symétrique si $p = q = \frac{1}{2}$ et antisymétrique sinon.

Cet exemple est un modèle de la marche aléatoire parmi tant d'autres tels un jeu dans un casino où S_0 représente la fortune initiale du joueur et S_n sa fortune à l'issue du $n^{\text{ème}}$ jeu, le développement d'une épidémie (marche aléatoire sur $\mathbb{Z}^n$) ou l'indice de prix dans un marché financier.

Exercice 1

Soit $(S_n)_{n \geq 0}$ la marche aléatoire sur $\mathbb{Z}$ précédente.

1) Soit $U_n = P[S_n = S_0]$ (probabilité que la particule revienne à sa position initiale à la date n). Montrer que $U_n = 0$ si n impair et $U_n = C_{2m}^m p^m q^m$ si $n = 2m$ est pair.

2) Montrer que $\forall k \geq 0$, $P[S_n = k | S_0 = 0] = C_n^{\frac{1}{2}(n+k)} p^{\frac{1}{2}(n+k)} q^{\frac{1}{2}(n-k)}$

indication : remarquer que $\forall i \geq 1$, $\frac{1}{2}X_{i+1} \rightsquigarrow \mathcal{B}(1, p)$.

3) Soit $T_0 = \{n > 0 | S_n = 0\}$ (T_0 est la première fois que la particule revient à sa position initiale après la date 0).

Montrer que $P[T_0 < +\infty \mid S_0 = 0] = 1 - |p - q|$.

indication : $\{T_0 < +\infty\} = \bigcup_{n \geq 1} \{S_n = 0\}$.

Soit $B_n = \{T_0 = n\} = \{S_1 \neq 0, \dots, S_{n-1} \neq 0, S_n = 0\}$.

Alors $A_n = \{S_n = 0\} = \bigcup_{k=1}^{\infty} A_n \cap B_k$.

Montrer que $P[A_n \cap B_k] = P(B_k) P(A_{n-k})$, $\forall 1 \leq k \leq n$.

En déduire, en posant $f_n = P(B_n)$ et $U_n = P(A_n)$, que $U_n = \sum_{k=1}^n f_k U_{n-k}$.

4) Calculer $P[S_{2n+1} = 1 \mid S_0 = 0]$.

5) Montrer que $\sum_{n=0}^{\infty} U_n$ converge si $p \neq q$ et diverge si $p = q$.

En déduire qu'une M.A.S. (marche aléatoire simple) antisymétrique revisite presque sûrement une infinité de fois sa position initiale. On pourra se souvenir de la formule de Stirling : $n! \simeq n^{n+\frac{1}{2}} e^{-n} \sqrt{2n}$ pour tout n assez grand.

Exercice 2 (M.A. sur $\mathbb{Z}^2$)

Soit une M.A. en dimension 2 où la particule se déplace sur les points

$\{(i, j), i, j = 0, \pm 1, \pm 2, \dots\}$ du plan.

Soit p, q, r, s tels que $0 < p, q, r, s < 1$ et $p + q + r + s = 1$. On suppose que si la particule est à la position (i, j) à la date n , sa position à la date $n + 1$ est $(i + 1, j)$ avec probabilité p , $(i, j + 1)$ avec probabilité q , $(i - 1, j)$ avec probabilité r et $(i, j - 1)$ avec probabilité s et que les déplacements successifs entre les points sont indépendants.

Soit $S_0 = (0, 0)$ et S_n la position de la particule à la date n . On a $\forall n \geq 0$

$$S_{n+1} = \begin{cases} S_n + (1, 0) \text{ avec proba. } p \\ S_n + (0, 1) \text{ avec proba. } q \\ S_n - (1, 0) \text{ avec proba. } r \\ S_n - (0, 1) \text{ avec proba. } s \end{cases}$$

Soit $U_n = P[S_n = S_0]$. Montrer que

$$U_n = \begin{cases} 0 & \text{si } n \text{ impair} \\ \sum_{k=0}^n \frac{(2n)!}{(k!)^2 ((n-k)!)^2} (pr)^k (qs)^{n-k} & \text{si } n = 2m \text{ est pair} \end{cases}$$

6.2 Chaînes de Markov

Soit $X = (X_n)_{n \geq 0}$ un processus aléatoire à valeurs dans un espace fini ou dénombrable E servant à modéliser l'état d'un système au cours du temps, X_n désignant cet état au temps n .

Dans ce cadre, nous dirons que le système se trouve dans l'état $i \in E$ au temps n si $X_n = i$.

On dira que le processus $X = (X_n)_{n \geq 0}$ est une chaîne de Markov homogène à espace d'états E , si, partant de tout état i de E , la probabilité que la chaîne passe immédiatement à l'état j de E est constante au cours du temps. Plus formellement :

Définition 6.2.1 (Chaînes de Markov homogène)

Soit $X = (X_n)_{n \geq 0}$ un processus aléatoire à valeurs dans un espace fini ou dénombrable E . On dira que X est une chaîne de Markov homogène (CMH) si, pour tous états $i, j, i_0, i_1, \dots, i_{n-1} \in E$ et pour tout $n \in \mathbb{N}^*$,

$$P[X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0] = P[X_{n+1} = j \mid X_n = i]. \quad (6.1)$$

Les probabilités $P(i, j) = P[X_{n+1} = j | X_n = i] = P[X_1 = j | X_0 = i]$, $(i, j) \in E^2$ sont appelées probabilités de transition de la CMH.

Les probabilités de transitions $P_{i,j} = P(i, j)$ vérifient

$$P_{i,j} \geq 0 \quad \text{et} \quad \forall i \in E, \quad \sum_{j \in E} P_{i,j} = 1 \quad (6.2)$$

Lorsque E est fini, avec $\text{Card}(E) = M$, il est commode de disposer de la matrice de transition $\mathbb{P} = [P_{i,j}]$, $1 \leq i, j \leq M$ de la CMH X . La connaissance de transitions $\mathbb{P}$ d'une CMH $(X_n)_{n \in \mathbb{N}}$ et de la loi initiale de X_0 détermine la loi de toute la chaîne.

Exemple 6.2.1

On admet que le fait qu'il ait plu ou non un jour donné est la seule considération à prendre en compte pour prévoir s'il pleuvra le lendemain. Plus précisément s'il ne pleut pas aujourd'hui, il pleuvra demain avec probabilité α et s'il ne pleut pas aujourd'hui, il ne pleuvra pas demain avec probabilité β , $0 < \alpha, \beta < 1$. En convenant de dire que le système est dans l'état 1 s'il ne pleut pas et dans l'état 2 s'il pleut, la situation peut être représentée par une CMH à deux états $E = \{1, 2\}$ et les probabilités de transitions $P_{11} = P[X_{n+1} = 1 | X_n = 1] = 1 - \alpha$, $P_{12} = P[X_{n+1} = 2 | X_n = 1] = \alpha$, $P_{21} = P[X_{n+1} = 1 | X_n = 2] = \beta$, $P_{22} = P[X_{n+1} = 2 | X_n = 2] = 1 - \beta$. La matrice de transitions est donc

$$\mathbb{P} = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}$$

Soit $(X_n)_{n \in \mathbb{N}}$ une CMH. La probabilité de passer d'un état i à un état j en une transition est $P_{i,j} = P[X_{n+1} = j | X_n = i]$. Mais la probabilité que la chaîne passe de i à j en m transitions s'écrit

$$P_{ij}^{(m)} = P^{(m)}(i, j) = P[X_{n+m} = j | X_n = i] \quad (6.3)$$

e.g.

$$\begin{aligned} P_{ij}^{(2)} &= P^{(2)}(i, j) = P[X_{n+2} = j | X_n = i] = P[X_2 = j | X_0 = i] \\ &= \sum_{k \in E} P[X_2 = j | X_1 = k] P[X_1 = k | X_0 = i] \\ &= \sum_{k \in E} P(k, j) P(i, k) = \sum_{k \in E} P_{ik} P_{kj} \end{aligned}$$

Donc $P_{ij}^{(2)}$ est l'élément de la ligne i et de la colonne j de la matrice $\mathbb{P}^2$. Plus généralement

$$\begin{aligned} P_{ij}^{(m)} &= P^{(m)}(i, j) = P[X_{n+m} = j | X_n = i] = P[X_m = j | X_0 = i] \\ &= \sum_{k \in E} P(k, j) P^{(m-1)}(i, k) = \sum_{k \in E} P_{ik}^{(m-1)} P_{kj} \end{aligned}$$

l'élément de la ligne i et de la colonne j de la matrice $\mathbb{P}^m = \mathbb{P}^{m-1} \mathbb{P} = \mathbb{P} \mathbb{P}^{m-1}$.

Théorème 6.2.1 (Chapmann-Kolmogorov)

La matrice de transition en n étapes est égale à la puissance n -ième de la matrice de transition en une étape.

Remarque 6.2.1

Avec l'intérêt des puissances m -ième de la matrice de transition via le Théorème de Chapman-Kolmogorov 6.2.1, on a, pour tout $n \in \mathbb{N}$,

$$P[X_n = j] = \sum_{i \in E} P[X_n = j | X_0 = i] P[X_0 = i] \quad (6.4)$$

Donc en notant par μ la loi de X_0 (loi initiale de la chaîne), l'équation précédente s'écrit

$$P[X_n = j] = \sum_{i \in E} \mu_i P_{ij}^{(n)} \stackrel{\text{Déf}}{=} \mu \mathbb{P}^n(j), \quad \forall n \geq 1 \quad (6.5)$$

où $\mu_i = \mu(i) = P[X_0 = i]$, $\forall i \in E$. En d'autres termes $\mathcal{L}(X_n) = \mu \mathbb{P}^n$.

6.2.1 Ergodicité

Pour une chaîne de Markov homogène, il est intéressant de savoir dans quelle mesure le passé influe sur le futur. Cela revient à s'intéresser au comportement limite des $P_{ij}^{(n)}$ lorsque $n \rightarrow +\infty$. Il apparaît que pour un grand nombre de CMH, la suite $P_{ij}^{(n)}$ converge vers une limite π_j ne dépendant que de l'état j . En d'autres termes, la probabilité de se trouver à l'état j après n transitions pour n assez grand est approximativement π_j , indépendamment de l'état i de départ. On peut montrer qu'une condition suffisante pour qu'une CMH possède cette propriété est qu'il existe un entier $n > 0$ telles que

$$P_{ij}^{(m)} > 0, \quad \forall (i, j) \in E^2 \quad (6.6)$$

Les chaînes de Markov qui vérifient la condition (6.6) sont dites ergodiques. Or d'après le Théorème de Chapman-Kolmogorov 6.2.1

$$P_{ij}^{(n+1)} = \sum_{k \in E} P_{ik}^n P_{kj},$$

Donc lorsque $n \rightarrow +\infty$, les chaînes de Markov ergodiques vérifient $\pi_j = \sum_{k \in E} \pi_k P_{kj}$, $\forall j \in E$, c'est-à-dire

$$\pi \mathbb{P} = \pi \quad (6.7)$$

Comme $\sum_{j \in E} P_{ij}^n = 1$, on aussi $\sum_{j \in E} \pi_j = 1$.

Les résultats précédent sont rassemblés dans le théorème suivant

Théorème 6.2.2 (Ergodique)

Pour toute chaîne de Markov ergodique $\pi_j = \lim_{n \rightarrow +\infty} P_{ij}^n$ existe et les π_j sont les seules solutions non négatives du système d'équation

$$\begin{cases} \pi = \pi \mathbb{P} \\ \sum_{j \in E} \pi_j = 1. \end{cases} \quad (6.8)$$

π est dite loi stationnaire de la chaîne de Markov.

Remarque 6.2.2

L'étude de la convergence des P_{ij}^n ou d'une manière équivalente, de la matrice $\mathbb{P}^n$ revient à étudier le comportement des valeurs propres de $\mathbb{P}$. En effet $\mathbb{P}$ est semblable à une matrice diagonale ou triangulaire. Donc si $\mathbb{P}$ a une valeur propre λ avec $|\lambda| > 1$, la matrice $\mathbb{P}^n$ ne converge pas. On notera que 1 est valeur propre (vp) triviale de $\mathbb{P}$. La chaîne est donc ergodique si et ssi toutes les valeurs propres de $\mathbb{P}$ sont à l'intérieur du cercle unité.

Exemple 1

Soit la chaîne dont la matrice de transition est $\mathbb{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

La matrice semblable à $\mathbb{P}$ est $\mathbb{P}' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Donc les valeurs propres sont toutes de module

1. $\mathbb{P}^n$ ne peut donc converger et la chaîne n'est pas ergodique. On a d'ailleurs $\mathbb{P}^{2n} = I$ et $\mathbb{P}^{2n+1} = \mathbb{P}$.

2) Soit une chaîne à 4 états 1, 2, 3, et 4, dont la matrice de transition est $\mathbb{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Alors $\mathbb{P} = Q\mathbb{P}'Q^{-1}$ avec $\mathbb{P}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \end{pmatrix}$ et $Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & 1 & 1 \\ \frac{1}{3} & \frac{2}{3} & 1 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$

avec $Q^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{6} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix}$.

On a donc

$$\mathbb{P}^n = Q\mathbb{P}'^nQ^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{2^{n+1}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \frac{(-1)^n}{2^{n+1}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{3} & 1 & -1 & \frac{1}{3} \\ \frac{1}{3} & -1 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Par conséquent

$$\pi = \lim_{n \rightarrow +\infty} \mathbb{P}^n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Interprétation :

Si la chaîne se trouve initialement à l'une des extrémités du segment $[1, 4]$, elle y reste presque sûrement. Si elle se trouve à l'état 2, elle a deux chances sur trois d'aboutir tôt ou tard à l'état 1, et une chance sur trois d'aboutir à l'état extrémité 4. Par contre, elle n'a aucune chance de se stabiliser en un état intermédiaire. On a un résultat analogue si le point de départ est l'état 3.

On peut aussi dire que la chaîne est attirée par les extrémités absorbantes 1 et 4, qu'elle finit toujours par rejoindre. Elle a naturellement plus de chance d'aboutir à l'extrémité la plus proche de sa position initiale.

Exemple 2

Soit une chaîne de Markov à trois états 1, 2 et 3, dont la matrice de transitions est

$$\mathbb{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{5} & \frac{1}{10} & \frac{1}{10} \end{pmatrix}.$$

La diagonalisation de $\mathbb{P}$ donne comme valeurs propres $\lambda_1 = 1$, $\lambda_2 = \frac{1}{2}$ et $\lambda_3 = \frac{1}{5}$. La matrice de passage Q et son inverse sont données par :

$$Q = \begin{pmatrix} 1 & 1 & -5 \\ 1 & 2 & 5 \\ 1 & -2 & 1 \end{pmatrix} \quad \text{et} \quad Q^{-1} = \begin{pmatrix} \frac{4}{12} & \frac{3}{12} & \frac{5}{12} \\ \frac{2}{18} & \frac{3}{18} & \frac{-5}{18} \\ \frac{-4}{36} & \frac{3}{36} & \frac{1}{36} \end{pmatrix}.$$

On a donc

$$\lim_{n \rightarrow +\infty} \mathbb{P}^n = \lim_{n \rightarrow +\infty} Q \mathbb{P}^n Q^{-1} = \begin{pmatrix} \frac{4}{12} & \frac{3}{12} & \frac{5}{12} \\ \frac{4}{12} & \frac{3}{12} & \frac{5}{12} \\ \frac{4}{12} & \frac{3}{12} & \frac{5}{12} \end{pmatrix}.$$

Interprétation

La chaîne en question comporte trois éventualités. Notons les E_1 , E_2 et E_3 .

Lorsque la chaîne est en E_1 elle a les probabilités $\frac{1}{2}$, $\frac{1}{4}$ et $\frac{1}{4}$ de rester en E_1 ou d'aller à E_2 ou à E_3 respectivement, etc.

Lorsque la chaîne fait un grand nombre n de transitions à partir d'un instant k , les probabilités que la chaîne soit en E_1 , E_2 et E_3 respectivement à l'instant $n + k$, sont lorsque $n \rightarrow +\infty$, $\frac{4}{12}$, $\frac{3}{12}$, $\frac{5}{12}$ respectivement. Ces probabilités sont bien indépendantes des états de départ à l'instant k considéré.

6.2.2 Chaîne de Markov continues

On dit que $X = (X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov homogène continue si en plus d'avoir les propriétés de Markov homogène précédentes, les X_n sont des variables aléatoires réelles continues et de même loi admettant une densité de probabilité notée $f(x)$. La loi de probabilité conditionnelle de X_{n+k} sachant $X_n = x$ admet donc une densité $f^{(k)}(y|x)$ indépendante de n . Enfin la loi conditionnelle de X_{n+k} ($k \geq 1$) sachant les valeurs prise par les X_i , $i \leq n$ dépend uniquement de la valeur récente prise par X_n . La loi du couple (X_n, X_{n+k}) est alors définie par sa densité

$$f^{(k)}(y|x)f(x). \quad (6.9)$$

De même par la règle de la multiplication, celle du triplet $(X_n, X_{n+k}, X_{n+k+m})$ est définie par la densité

$$f^{(m)}(z|y)f^{(k)}(y|x)f(x). \quad (6.10)$$

Il en résulte que la loi conditionnelle de X_{n+k+m} sachant $X_n = x$ est donnée par

$$f^{(k+m)}(z|x) = \int_{-\infty}^{+\infty} f^{(m)}(z|y)f^{(k)}(y|x)dy. \quad (6.11)$$

En particulier $f^{(1)}(y|x) = f(y|x)$ est la densité (ou noyau) de transition de x à y en une étape et plus généralement

$$f^{(n+1)}(z|x) = \int_{-\infty}^{+\infty} f^{(n)}(y|x)f(z|y)dy. \quad (6.12)$$

Exemple

Supposons que le passage de X_n à X_{n+1} suive une loi normale et que chacun des X_n obéisse à une loi normale réduite.

Posons

$$f(y|x) = \frac{1}{\sqrt{2\pi} \sqrt{1-r^2}} e^{-\frac{1}{2} \frac{(y-rx)^2}{1-r^2}}.$$

Cherchons $f^{(k)}(z|x)$. On a d'abord

$$f^{(2)}(z|x) = \int_{-\infty}^{+\infty} f(z|y)f(y|x)dy = \frac{1}{2\pi(1-r^2)} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \frac{(y-rx)^2 + (z-ry)^2}{1-r^2}} dy.$$

On obtient avec les résultats sur l'intégrale de Gauss

$$f^{(2)}(z|x) = \frac{1}{\sqrt{2\pi} \sqrt{1-r^2}} e^{-\frac{1}{2} \frac{(z-r^2x)^2}{1-r^2}}.$$

On établit enfin par récurrence sur n que $f^{(n)}(z|x)$ s'obtient en remplaçant r par r^n dans l'expression de $f(z|x)$, r étant le coefficient de corrélation entre X_n et X_{n+1} . Si $|r| < 1$, $r^n \rightarrow 0$ quand $n \rightarrow +\infty$ et donc

$$f^{(n)}(z|x) \rightarrow f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

On retrouve la propriété ergodique étudiée dans le cas discret, mais cette propriété ergodique n'est pas assurée si $r = \pm 1$.

Si $r = 1$ par exemple, $f^{(n)}(z|x)$ ne dépend plus de n . Dans ce cas, on $X_{n+1} = X_n$ presque sûrement. X_1 est aléatoire mais son choix étant fait, les X_n successifs sont bien déterminés. De même si $r = -1$, on a $X_2 = -X_1$, $X_3 = X_1$, $X_4 = -X_1$, etc.

Notons que l'espace d'états E d'une chaîne de Markov au lieu d'être discret, peut dépendre d'un temps continu t . Au lieu donc de X_n , on écrit X_t ou $X(t)$ et $X(t)$ est une fonction aléatoire de t . $X(t)_{t \in T}$ est un processus de Markov si la loi de $X(t+h)$ sachant $X(s)$, $s \leq t$ est identique à la loi de $X(t+h)$ sachant $X(t)$, $\forall h > 0$. Les relations de Chapman-Kolmogorov

$$\begin{aligned} P_{ij}^{n+m} &= \sum_{k \in E} P_{ik}^n P_{kj}^m \\ f^{\alpha+\beta}(z|x) &= \int_{-\infty}^{+\infty} f^\alpha(z|y) f^\beta(y|x) dy \end{aligned} \quad (6.13)$$

restent valables, mais ici α et β sont des variables non négatives continues qui peuvent être aussi dérivables. Moyennant quelques hypothèses de régularité, la relation fonctionnelle intégrale ci-dessus peut alors être remplacée par une équation aux dérivées partielles de Kolmogorov. Ce qui est du niveau Master 2.

Chapitre 7

Exercices d'application

7.0.1 Exercices du chapitre 1

Exercice 1

On considère deux événements A et B liés à une même expérience aléatoire.

- 1) On donne : $P(A) = \frac{3}{8}$, $P(A \cap B) = \frac{1}{4}$, $P(A \cup B) = \frac{7}{8}$. Calculer $P(B)$ et $P(A \cap \bar{B})$.
- 2) On donne : $P(A) = \frac{1}{2}$, $P(A \cup B) = \frac{3}{4}$, $P(\bar{B}) = \frac{5}{8}$. Calculer $P(A \cap B)$, $P(\bar{A} \cap \bar{B})$, $P(\bar{A} \cup \bar{B})$ et $P(A \setminus B)$.

Exercice 2

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

- 1) $\forall (A, B, C) \in \mathcal{A}^3, \bar{B} \cap \bar{C} \subset \bar{A} \Rightarrow P(A) \leq P(B) + P(C)$;
- 2) $\forall (A, B) \in \mathcal{A}^2, P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$;
- 3) $\forall (A_1, A_2, \dots, A_n) \in \mathcal{A}^n, P(\bigcap_{i=1}^n A_i) \geq P(A_j) - \sum_{i \neq j} P(\bar{A}_i)$ pour tout $j = 1, 2, \dots, n$.

Exercice 3

Soient $A_1, A_2, \dots, A_n$ une suite d'événements liés à une même expérience aléatoire. Démontrer la formule de Poincaré :

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) + \dots + (-1)^{n-1} P(A_1 \cap \dots \cap A_n).$$

Applications :

1) n couples "légitimes" participent à une soirée. Ils se séparent puis se reforment par couples au hasard.

Quelle est la probabilité qu'au moins un couple "légitime" se retrouve ?

Limite de cette probabilité quand $n \rightarrow +\infty$?

2) Soient n et p , ($p \geq n$) deux entiers strictement positifs. On lance au hasard p boules numérotées de 1 à p sur un plan constitué de n cases numérotées de 1 à n . Chaque boule tombe dans une case et chaque case peut en recevoir plusieurs.

a) Calculer la probabilité qu'au moins une case soit vide.

b) En déduire que le nombre $N_{p,n}$ de surjections d'un ensemble à p éléments sur un ensemble

à n éléments est donné par : $N_{p,n} = \sum_{k=0}^n (-1)^k C_n^k (n-k)^p$.

Exercice 4

On suppose que les numéros de téléphone à 8 chiffres peuvent comporter à chaque position n'importe quel chiffre de 0 à 9 et que tous les numéros sont équiprobables.

Calculer la probabilité qu'un numéro à 8 chiffres soit tel que :

1. Le numéro est formé avec deux chiffres distincts et deux seulement.
2. Le numéro est formé de deux 0, deux 4 et quatre 6.
3. Le numéro comporte deux 5 et deux seulement.
4. La somme des chiffres du numéro est égal à 12.
5. Les chiffres du numéro forme une suite strictement croissante.
6. Les chiffres du numéro forme une suite croissante au sens large.

Exercice 5

On suppose que l'année compte n jours. Quelle est la probabilité pour que dans une assemblée de k personnes ($k < n$) au moins deux personnes aient leurs jours d'anniversaire en commun ?

Evaluer cette probabilité lorsque n et k tendent vers $+\infty$ tel que $\frac{k}{n} \rightarrow \alpha$. A.N. $n = 365$ $k = 19$.

Exercice 6

Un distributeur de prospectus (tous identiques) fait son travail de façon distraite et a distribué au hasard n prospectus dans les p boîtes à lettres d'un immeuble supposées numérotées de 1 à p .

- a) Quelle est la probabilité q_1 pour que dans les boîtes il y ait respectivement $r_1, r_2, \dots, r_p$ prospectus $r_1 + r_2 + \dots + r_p = n$?
- b) Quelle est la probabilité q_2 pour qu'aucune boîte ne soit vide ?

Exercice 7

Avec les données suivantes, dans quels cas les événements A et B sont-ils indépendants ?

	$P(A)$	$P(B)$	$P(A \cup B)$
cas 1	0,1	0,9	0,91
cas 2	0,4	0,6	0,76
cas 3	0,5	0,3	0,73

Exercice 8

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $(A, B, C) \in \mathcal{A}^3$ tels que $P(A) > 0, P(B) > 0$ et $P(C) > 0$. On suppose que A et $B \cup C$ sont indépendants, A et $B \cap C$ sont indépendants, B et $C \cap A$ sont indépendants, C et $A \cap B$ sont indépendants. Montrer que les trois événements sont indépendants.

Exercice 9

Soient A et B deux événements liés à une même expérience aléatoire.

- 1) On suppose que A et B sont incompatibles et de probabilités non nulles. A et B peuvent-ils être indépendants ?
- 2) On suppose que A et B vérifient $P(B/A) = P(B/\bar{A})$. Montrer qu'ils sont indépendants.
- 3) On suppose que A et B sont incompatibles et de probabilités non nulles, et que C est un autre événement telles que $P(C/A) = P(C/B)$. Montrer que $P(C/A) = P(C/B) = P(C/A \cup B)$.

Exercice 10

On lance trois fois de suite une pièce de monnaie parfaite. Soient les événements suivants :

A = " pile apparaît au 1^{er} lancer " ;

B = " Au moins deux piles apparaissent sur les trois lancers " et

C = " On obtient le même résultat sur les trois lancers " .

Les trois événements A, B, C sont-ils deux à deux indépendants ?

Exercice 11

Dans une association, trois bénévoles A, B, C collent des timbres sur des enveloppes. Le travail est effectué avec diligence, A traitant 50% de l'ensemble, B traitant 30% et C traitant 20% . Il est également effectué dans l'allégresse mais aussi dans l'étourderie puisque A omet de mettre un timbre dans 3% des cas, B dans 2% et C dans 1% . Un adhérent se plaint que la lettre qu'il vient de recevoir ne comporte pas de timbre.

Quelle est la probabilité que l'enveloppe de cette lettre ait été traitée par A ? par B ? par C ?

Exercice 12

Soient $A_1, A_2, \dots, A_n$ une suite d'événements liés à une même expérience aléatoire. Démontrer que

$$P(A_1.A_2...A_n) = P(A_1).P(A_2/A_1).P(A_3/A_1A_2) \cdots P(A_n/A_1.A_2...A_{n-1}).$$

Applications : Une urne contient a boules blanches et b boules rouges. On en extrait successivement et avec remise n boules de l'urne.

a) Probabilité que les k premières boules extraites soient blanches et les $n-k$ dernières rouges ?

b) Probabilité d'obtenir k boules blanches au cours des n tirages ?

c) On décide maintenant qu'avant de procéder au tirage suivant on rajoute dans l'urne, c nouvelles boules de même couleur que celle qui a été précédemment tirée et remise dans l'urne.

Quelle est la probabilité d'obtenir k boules blanches au cours des n tirages ?

7.0.2 Exercices du chapitre 2

Exercice 1

Pour tout $s \in \mathbb{R}_+$, on pose $\Gamma(s) = \int_0^{+\infty} x^{s-1} e^{-x} dx$.

1) a) Montrer que $\Gamma(s)$ est bien définie, b) Calculer $\Gamma\left(\frac{1}{2}\right)$, c) Montrer que $\forall s > 1, \Gamma(s) = (s-1)\Gamma(s-1)$, d) Donner l'expression de $\Gamma\left(\frac{2k+1}{2}\right)$ pour tout $k \in \mathbb{N}^*$ en fonction de $\Gamma\left(\frac{1}{2}\right)$.

2) Soit X une variable aléatoire réelle positive telle que $P[X \leq x] = 1 - \exp(-\frac{1}{2}x^2)$. Calculer $E(X)$ et $Var(X)$.

3) Soit $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \left(\frac{4x^2}{\alpha^3 \sqrt{\pi}}\right) \exp(-\frac{x^2}{\alpha^2})$ avec $\alpha \in \mathbb{R}_+$.

a) Montrer que f est une densité de probabilité sur $\mathbb{R}_+$.

b) Soit X la variable aléatoire dont la loi a pour densité f . Calculer $E(X)$ et $Var(X)$.

c) Soit X, Y deux variables aléatoires indépendantes et de même loi de densité f . On pose $Z = \frac{X}{Y}$.

Calculer $P[z \leq Z < z + dz]$ et en déduire la densité de Z .

Exercice 2

Soit X une v.a.r. positive absolument continue telle que $E(X)$ existe.

- a) Montrer que $\mathbb{E}(X) = \int_0^{+\infty} P[X \geq x] dx$.
 b) Montrer que si X est à valeurs dans $\mathbb{N}$ alors $\mathbb{E}(X) = \sum_{n \geq 0} P[X > n]$.

Exercice 3

Soit X une v.a.r. absolument continue dont la densité f est une fonction continue sur $\mathbb{R}$ et telle que $\int_{-\infty}^{+\infty} x^2 f(x) dx < +\infty$.

- a) Montrer que $\lim_{x \rightarrow +\infty} x^2 P[|X| \geq x] = 0$.
 b) Montrer que $\int_0^{+\infty} t P[|X| \geq t] dt$ converge et que $\int_0^{+\infty} t P[|X| \geq t] dt = \frac{1}{2} \mathbb{E}(X^2)$.

Exercice 4

Soit la fonction de $\mathbb{R}$ dans lui-même

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \sin x & \text{si } 0 < x < \frac{\pi}{2} \\ 1 & \text{si } x \geq \frac{\pi}{2} \end{cases}$$

- a) Montrer que F est une fonction de répartition. Déterminer si elle existe, la densité de sa loi.
 b) En appliquant l'inégalité de Bienaymé-Tchebichev, montrer que pour tout a , $0 < a < 1$, $\cos(1-a) \geq 1 - \frac{\pi-3}{a^2}$.

Exercice 5

Soit A un point du plan et Δ une droite ne passant pas par A et située à une distance a de A . On trace au hasard une droite D passant par A . Soit X la position du point d'intersection de D et Δ .

Loi de probabilité de la v.a. X et son espérance mathématique $\mathbb{E}(X)$?

Exercice 6

Soit un plan horizontal où sont tracées des droites parallèles équidistantes à la droite constante d . On jette au hasard sur le plan une aiguille de longueur $l < d$ (aiguille de Buffon). Quelle est la probabilité que cette aiguille rencontre une des droites parallèles?

Exercice 7

On prend au hasard deux nombres réels a et b , $-k \leq a \leq +k$, $-s \leq b \leq +s$. Quelle est la probabilité pour que l'équation du second degré $x^2 + 2ax + b = 0$ ait deux racines réelles.

Exercice 8

Une personne vous a donné un rendez-vous entre 0h et 1h. Elle viendra peut-être vous voir dans cette période avec une probabilité p $0 < p < 1$. Cela fait t minutes que vous l'attendez et elle n'est pas encore arrivée. Quelle est la probabilité p_t qu'elle vienne?

Exercice 9 Soit (X, Y) un couple de variables aléatoires réelles continues dont la loi conjointe admet pour densité

$$f(x, y) = \begin{cases} \frac{3y}{a^3} & \text{si } x+y < a, y-x < a \text{ et } y > 0 \\ 0 & \text{sinon} \end{cases}$$

où $a \in \mathbb{R}_+^*$. Calculer $\text{cov}(X, Y)$. Les variables aléatoires X et Y sont-elles indépendantes?

Exercice 10

Détermine k telle que $f(x, y) = k \mathbb{I}_{\{x^2 + y^2 \leq 1\}}$ soit une densité de probabilité.

Densités marginales du couple (X, Y) dont la loi a pour densité f ? $Cov(X, Y)$? X et Y sont-elles indépendantes?

Exercice 11

La fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x, y) = e^{-y} \mathbb{I}_{\{0 \leq x \leq y < +\infty\}}$ est-elle une densité de probabilité? Si oui, déterminer la et les lois marginales du couple (X, Y) dont la densité est f . X et Y sont-elles indépendantes?

Exercice 12 Soient $U = ST$ et $V = \frac{S}{T}$ où S et T sont deux variables aléatoires indépendantes et de même loi uniforme sur $[0, 1]$. Déterminer la loi de probabilité de (U, V) et les lois marginales de U et de V .

Exercice 13

Soient X et Y deux v.a.r. indépendantes suivant les lois $\mathcal{E}(1/2)$ et $\mathcal{U}[0, 2\pi]$ respectivement. On pose $U = \sqrt{X} \cos Y$ et $V = \sqrt{X} \sin Y$.
Loi de (U, V) ? Lois marginales de U et de V ? U et V sont-elles indépendantes?

Exercice 14

Soit $X \rightsquigarrow \mathcal{B}(n, p)$. Sachant que $X = k$, $Y \rightsquigarrow \mathcal{B}(n - k, r)$, $0 \leq p \leq 1$, $0 \leq r \leq 1$.
Quelle est la loi de probabilité de la v.a. $S = X + Y$?

Exercice 15

Le nombre X de pannes d'une machine suit une loi de Poisson dont on ne connaît pas exactement le paramètre λ . On sait tout de même que λ peut prendre seulement les valeurs 1, 3 et 4 avec des probabilités respectives 0,6; 0,3 et 0,1.
a) Calculer $\mathbb{E}(X)$ et $\text{Var}(X)$.
b) On effectue un essai de la machine et on constate 3 pannes. Comment cela modifie-t-il les probabilités que λ vaille 1, 3 ou 4?
c) Compte tenu de ces modifications, calculer $\mathbb{E}(X)$ et $\text{Var}(X)$.

Exercice 16

Le nombre N d'oeufs pondus par un insecte suit une loi de Poisson de paramètre $\lambda > 0$. Les oeufs se développent de façon indépendante et chacun d'eux donne naissance à un nouvel insecte avec probabilité a , $0 < a < 1$.
1) Loi du nombre X d'insectes à la première génération?
2) On observe k insectes à la première génération. Quelle est la probabilité qu'il y ait eu n oeufs pondus?

7.0.3 Exercices des chapitres 2, 3

Exercice 1

Soient X_1, X_2 deux v.a.r. indépendantes. On suppose que X_1 et X_2 suivent les lois $\mathcal{B}(n, p)$ et $\mathcal{B}(m, p)$ respectivement, $0 \leq p \leq 1$.
1) Trouver la fonction génératrice $G_S(s)$ de la v.a. $S = X_1 + X_2$. En déduire la loi de la variable S .
2) Même questions si $X_1 \rightsquigarrow \mathcal{P}(\lambda)$ et $X_2 \rightsquigarrow \mathcal{P}(\mu)$.
3) Soit $X_n \rightsquigarrow \mathcal{B}(n, p)$. Trouver $\lim_{n \rightarrow +\infty} G_{X_n}(s)$ lorsque $p \rightarrow 0$ et $np \rightarrow \lambda$. Qu'a-t-on prouvé?

Exercice 2

Dans un jeu de pile ou face, la probabilité d'obtenir pile est p , $0 \leq p \leq 1$ et celle de face est $q = 1 - p$. Soit S_n le nombre de jets nécessaires pour obtenir n fois pile.

- 1) Déterminer la fonction génératrice de la v.a. S_n ? (on pourra considérer d'abord le cas $n = 1$).
- 2) En déduire la loi de S_n .

Exercice 3

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. indépendantes, de même loi, de fonction caractéristique Φ . Soit Y une v.a. à valeurs dans $\mathbb{N}$, indépendante de la suite (X_n) et de fonction génératrice Ψ .

On définit $Z = \sum_{k=1}^Y X_k$.

- a) Déterminer la fonction caractéristique Φ de Z .
- b) Si $Y \rightsquigarrow \mathcal{P}(\lambda)$, déterminer Φ . (On dit que Z suit une loi de poisson composée).

Exercice 4

1) Soit U et V deux v.a.r. indépendantes et de même loi $\mathcal{N}(0, 1)$. Déterminer la fonction caractéristique (f.c.) $\Phi_{UV}(t)$ de UV .

2) Soient X_1, X_2, X_3, X_4 des v.a.r. indépendantes et de même loi $\mathcal{N}(0, 1)$. Trouver la f.c. $\Phi_T(t)$ de $T = X_1X_2 - X_3X_4$. Montrer que T a une densité et la calculer. Calculer $\mathbb{E}(T^n)$, $n \in \mathbb{N}^*$.

3) Soient X et Y deux v.a. indépendantes et de même loi $\mathcal{N}(0, 1)$. Montrer que $\frac{X - Y}{\sqrt{2}}$ et $\frac{X + Y}{\sqrt{2}}$

sont indépendantes et de même loi. Quelle est la loi de $\frac{X^2 - Y^2}{2}$ et celle de $\frac{X^2 + Y^2}{2}$?

4) Soient X_1, X_2, X_3 des v.a. indépendantes et de même loi $\mathcal{N}(0, 1)$. Trouver la f.c. $\Phi_{(X_1X_2, X_1X_3)}(s, t)$. Montrer que $\text{cov}(X_1X_2, X_1X_3) = 0$ mais que X_1X_2 et X_1X_3 ne sont pas indépendantes.

Exercice 5

Soit X_1, X_2 deux v.a.r. indépendantes et identiquement distribuées, de même fonction caractéristique $\Phi(t)$. On suppose que X_1 a une densité $f(x)$. En exprimant $\Phi_{X_1 - X_2}$ et $f_{X_1 - X_2}$, montrer que si $\Phi \in L^2(\mathbb{R})$ alors

$$\int_{-\infty}^{+\infty} f^2(u) du = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\Phi(t)|^2 dt$$

(Théorème de Plancherel).

Exercice 6

Soit $\Phi(t) = \frac{1 - it}{1 + t^2}$. Montrer que Φ est une fonction caractéristique d'une v.a. X que l'on précisera. Loi de X ?

Exercice 7

On rappelle la formule suivante dite formule d'inversion. Si $\varphi(t)$ est la fonction caractéristique d'une distribution $F(x)$ continue aux points a et b , $a < b$ alors

$$F(b) - F(a) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left[\varphi(t) \frac{e^{-ita} - e^{-itb}}{2it} - \varphi(-t) \frac{e^{ita} - e^{itb}}{2it} \right] dt. \quad (7.1)$$

a) Trouver F quand $\varphi(t) = e^{-\theta^2 t^2}$.

b) Déduire de (7.1) que si $F(x)$ est absolument continue avec pour densité $f(x) = F'(x)$ et si $\int_{-\infty}^{+\infty} |\varphi(t)| dt < +\infty$, alors

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \varphi(t) e^{-itx} dt. \quad (7.2)$$

On pourra étudier la parité de la fonction $t \mapsto \varphi(t)e^{-itx}$ et remarquer que

$$f(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x-h)}{2h}.$$

Exercice 8

X et Y sont deux variables aléatoires indépendantes et de même loi $\mathcal{U}[0, 1]$.

Loi de $S = X + Y$? Même question si $X \rightsquigarrow \mathcal{U}[-1, 1]$ (résultat : loi triangulaire)

Exercice 9

Soient $X_1, X_2, \dots, X_n$ n variables aléatoires indépendantes et de même loi. Calculer l'espérance mathématique de T_n dans chacun des cas suivants :

1. La loi commune est la loi exponentielle $\mathcal{E}(\lambda)$ et $T_n = \frac{n}{\sum_{i=1}^n X_i}$

2. la loi commune est la loi de Pareto de densité :

$$f(x) = \frac{\alpha - 1}{x^\alpha} \mathbb{I}_{[1, +\infty[}(x), \quad \alpha > 1 \quad \text{et} \quad T_n = 1 + \frac{n}{\sum_{i=1}^n \log X_i}$$

Exercice 10

Soient $X_1, X_2, \dots, X_n$ n variables aléatoires indépendantes et de même loi. On note $F(x)$ et $f(x)$, leur fonction de répartition et densité de probabilité communes.

1) Loi de $X_{(n)} = \max(X_1, X_2, \dots, X_n)$ et de $X_{(1)} = \min(X_1, X_2, \dots, X_n)$.

2) On note $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ le réarrangement par ordre croissant des variables aléatoires $X_1, X_2, \dots, X_n : X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$. On voudrait chercher la loi de $X_{(k)}$, $k \geq 1$. Soit la v.a. $N(x)$ définie par $N(x) = \text{card} \{i, i = 1, 2, \dots, n / X_i \leq x\}$.

a) Loi de $N(x)$?

b) Exprimer l'événement $\{X_{(k)} < x\}$ en fonction de $N(x)$. En déduire $P[X_{(k)} < x]$.

c) Démontrer que $\forall k, 1 \leq k \leq n$, et $\forall p, 0 < p < 1$, on a

$$\frac{1}{\beta(k, n-k+1)} \int_0^p y^{k-1} (1-y)^{n-k} dy = \sum_{j=k}^n C_n^j p^j (1-p)^{n-j}$$

où

$$\beta(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

d) En déduire la densité $f_k(x)$ de la v.a. $X_{(k)}$. Quelle est la loi de $X_{(k)}$ dans le cas où $X \rightsquigarrow \mathcal{U}[0, 1]$?

7.0.4 Exercices des chapitres 4, 5, 6

Exercice 1

Etudier la convergence vague de la suite de mesures :

a) $d\mu_n(x) = \sin nx dx$,

b) $d\mu_n(x) = \sin^2 nx dx$,

c) $\mu_n = \delta_{a_n}$ où (a_n) est une suite dans $\mathbb{R}$.

Exercice 2

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. telle que $X_n \xrightarrow{\mathcal{L}} X$.

- 1) X_n est absolument continue avec pour densité $f_n(x)$. La loi limite de X_n est-elle absolument continue et possède-t-elle une densité ?
- 2) Aucune des X_n n'est absolument continue et $X_n \xrightarrow{\mathcal{L}} X$. La loi limite peut-elle être absolument continue ?

Exercice 3

Soit (X_n) une suite de v.a.r. indépendantes et de même loi

$P[X_n = 1] = P[X_n = -1] = \frac{1}{2}$. Etudier la convergence en loi de la suite (Y_n) définie par $Y_n = \sum_{k=1}^n \frac{X_k}{2^k}$.

Exercice 4

Soit $t \in \mathbb{R}$ et $f \in C_b(\mathbb{R})$. Etudier la limite lorsque $n \rightarrow +\infty$ de la suite $Y_n = e^{-nt} \sum_{k \in \mathbb{N}} \frac{(nt)^k}{k!} f(\frac{k}{n})$.

Exercice 5

Soit $X \rightsquigarrow \gamma(n, \lambda)$. Loi limite de la suite $Y_n = \frac{\lambda}{\sqrt{n}} X - \sqrt{n}$?

Exercice 6

En utilisant le Théorème centrale limite, calculer $\lim_{n \rightarrow +\infty} e^{-n} \sum_{k=0}^n \frac{n^k}{k!}$.

Exercice 7

Soit (X_n) une suite de v.a.r. absolument continues ayant pour densité $f_n(x) = \frac{n}{\pi} \frac{1}{1 + n^2 x^2}$, $x \in \mathbb{R}$. Montrer que (X_n) converge en probabilité. Est-elle convergente en moyenne d'ordre $r \geq 1$?

Exercice 8

Soit (X_n) une suite de v.a.r. Montrer que :

- 1) Si (X_n) est uniformément bornée alors $X_n \xrightarrow{P} 0 \iff \mathbb{E}|X_n| \rightarrow 0$.
- 2) Si $f \in C_b(\mathbb{R})$ et continue en $x_0 \in \mathbb{R}$ alors $X_n \xrightarrow{P} x_0 \implies \mathbb{E}(f(X_n)) \rightarrow f(x_0)$.
- 3) Soit (X_n) une suite de v.a.r. indépendantes et de même loi $\mathcal{U}[0, 1]$ et soient $f, g \in C([0, 1], \mathbb{R})$ avec $f \geq 0$ et $g > 0$. Montrer que

$$\lim_{n \rightarrow +\infty} \int_0^1 \cdots \int_0^1 \frac{f(x_1) + \cdots + f(x_n)}{g(x_1) + \cdots + g(x_n)} dx_1 \cdots dx_n = \frac{\int_0^1 f(x) dx}{\int_0^1 g(x) dx}.$$

Exercice 9

En utilisant la loi des grands nombres, calcule $\lim_{n \rightarrow +\infty} I_n$ où

$$I_n = \int_0^1 \cdots \int_0^1 f\left(\frac{x_1 + \cdots + x_n}{n}\right) dx_1 \cdots dx_n$$

et $f \in C([0, 1], \mathbb{R})$.

Exercice 10

Soit $X_1, X_2, \dots, X_n$ n variables aléatoires indépendantes identiquement distribuées de loi $\mathcal{E}(\lambda)$, $\lambda > 0$. On pose $S_k = X_1 + \cdots + X_k$, $S_0 = 0$.

1- Loi du vecteur aléatoire $(S_1, S_2, \dots, S_n)$? En déduire la loi de la v.a. S_n .

2- Pour tout $t \in \mathbb{R}_+$, on pose $N_t = \sum_{n=1}^{\infty} \mathbb{I}_{\{S_n \leq t\}}$. Loi de N_t ?

Exercice 11

Soit X la durée de vie aléatoire d'un appareil électronique, continue, de f.r. F et de densité f . On suppose qu'après une panne, l'appareil est réparé et la durée de vie après réparation a même loi que X . Soit $X_1, \dots, X_n$ les durées de vie jusqu'à la $n^{\text{ième}}$ panne, indépendantes. On définit $S_n = X_1 + \dots + X_n$ le moment de la $n^{\text{ième}}$ panne; N_t le nombre de pannes observées dans l'intervalle $[0, t]$, $N(0) = 0$; F_n et f_n la f.r. et la densité de S_n .

Montrer que :

- $P[N_t = n] = F_n(t) - F_{n+1}(t) \quad \forall n = 0, 1, 2, \dots$
- $H(t) = \mathbb{E}N(t) = \sum_{n=1}^{+\infty} F_n(t)$
- Soit $W_t = S_{N(t)+1} - t$ le temps écoulé entre le moment t et la première panne après t .
Montrer que si f_n et $\sum f_n$ sont continues sur $\mathbb{R}_+$ alors
 $F_{W_t}(s) = P[W_t \leq s] = F(t+s) - \int_0^t (1 - F(t+s-x))h(x)dx, s \geq 0$ où $h(x) = H'(x)$.
- Lois de $N(t)$ et de W_t si les X_i sont de loi $\mathcal{E}(\lambda)$, $\lambda > 0$?

7.0.5 Exercices, problèmes, devoirs et examens divers

Problème 1

Soit X et Y deux variables aléatoires réelles dont la loi conjointe a pour densité

$$f(x, y) = \begin{cases} e^{-(x+y)}, & \text{si } 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{sinon} \end{cases}$$

- Les v.a. X et Y sont-elles indépendantes ?

On pose $Z = \frac{X}{Y}$. Calculer $P[Z \leq z]$ pour tout $z > 0$. En déduire que la v.a. Z admet une densité de probabilité qu'on calculera.

Problème 2

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soit $A, B, C \in \mathcal{A}$ et $(A_n)_{n \geq 1}$ une suite dans $\mathcal{A}$.

- Montrer que :

- $\overline{B} \cap \overline{C} \subset \overline{A} \Rightarrow P(A) \leq P(B) + P(C)$;
- $P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$;
- $P(\bigcap_{i=1}^n A_i) \geq P(A_j) - \sum_{i \neq j} P(\overline{A}_i)$ pour tout $j = 1, 2, \dots, n$.

- On désigne par A l'évènement "**une infinité des A_n se réalisent**".

- Exprimer A en fonction des A_n .
- Montrer que $\sum_{n=1}^{+\infty} P(A_n) < +\infty \Rightarrow P(A) = 0$.
- Montrer que $P(A) = 1 \iff P(\bigcap_{k \geq n} \overline{A}_k) = 0$.
- Montrer que si les A_n sont indépendants alors $\sum_{n=1}^{+\infty} P(A_n) = +\infty \Rightarrow P(A) = 1$.
- Que peut-on dire finalement de $P(A)$ lorsque les A_n forment une suite d'évènements ?

Problème 3

- Soit U une v.a. de loi uniforme sur $[-\frac{\pi}{2}, +\frac{\pi}{2}]$. Loi de la v.a. $X = \tan U$?

2) Soit X une v.a. continue admettant pour densité de probabilité

$$f(x) = \frac{K}{e^x + e^{-x}}, \quad \forall x \in \mathbb{R}.$$

Trouver la constante K .

3) Soient X et Y deux v.a. indépendantes suivant la loi de Poisson de paramètres α et β respectivement. Trouver la loi conditionnelle de X sachant $X + Y = n$, $n \in \mathbb{N}^*$.

4) Soient X et Y deux v.a. dont la loi conjointe a pour densité

$$f(x, y) = \begin{cases} 2e^{-x-y} & \text{si } 0 \leq x \leq y < +\infty \\ 0 & \text{sinon} \end{cases}$$

Trouver la densité conditionnelle de la loi de X sachant $Y = y$ et calculer $\mathbb{E}[X|Y = y]$.

5) Soit X une v.a. suivant la loi binomiale de paramètre n et p . Calculer $\mathbb{E}(g(X))$ avec $g(x) = \frac{1}{x+1}$. On donnera le résultat sous forme simplifiée.

Examen S6 UL 2011

N.B. : Le candidat traitera obligatoirement 3 exercices au choix parmi les quatre proposés.

Exercice 1

1) Le délégué d'une classe téléphone à ses n camarades pour les prévenir du report d'un DS. Chacun d'eux a la probabilité $p \in [0, 1]$ d'être au bout du fil. N'ayant pas pu joindre tout le monde, il rappelle quelques minutes plus tard. Chacun des camarades qu'il n'avait pas pu joindre a la probabilité $r \in [0, 1]$ d'être au bout du fil. Soit X la v.a. égale au nombre de ceux qu'il a pu joindre au premier tour d'appel et Y la v.a. égale au nombre de camarades joints au second tour d'appel.

a) Loi de X ? Loi de Y sachant $X = k$? Loi de $S = X + Y$?

b) A.N. $n = 35$, $p = 0.029$, $r = 0.09$. $\mathbb{E}(S)$ et $\text{Var}(S)$?

2) Soit X une v.a.r. En utilisant le Théorème de Fubini, calculer $\mathbb{E}\left(\frac{1}{1+X}\right)$ dans chacun des cas suivant : a) $X \sim \mathcal{P}(\lambda)$, $\lambda > 0$; b) $X \sim \mathcal{B}(n, p)$; c) Calculer dans le cas b) la limite de cette expression quand $n \rightarrow +\infty$, $p \rightarrow 0$ et $np \rightarrow \lambda$. Commenter ce résultat.

Exercice 2

Soit $X_1, \dots, X_n$ n v.a.r indépendantes et de même loi $\mathcal{E}(\lambda)$, $\lambda > 0$. On pose $S_0 = 0$ et pour tout $k \geq 0$, $S_k = X_1 + \dots + X_k$.

1) Loi de $(S_1, \dots, S_n)$? En déduire la loi de S_n pour tout $n \in \mathbb{N}^*$.

2) Pour tout $t \in \mathbb{R}_+$, on pose $N(t) = \sum_{i=1}^{+\infty} I(S_n \leq t)$ où $I(A)$ désigne l'indicatrice d'un événement

A. Loi de la v.a. $N(t)$?

On pourra remarquer que pour tout $k \in \mathbb{N}^*$, $\{N(t) = k\} = \{S_n \leq t < S_{n+1}\}$.

Exercice 3

1) Soit $F : \mathbb{R} \rightarrow \mathbb{R}$ définie par :

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ \sin x & \text{si } 0 \leq x \leq \frac{\pi}{2} \\ 1 & \text{si } x > \frac{\pi}{2} \end{cases}$$

a) Montrer que $x \mapsto F(x)$ est une fonction de répartition et déterminer si elle existe, la densité de probabilité de sa loi.

b) Soit X la variable aléatoire associée à la f.r. F . En appliquant l'inégalité de Bienamey-Tchebycheff à la v.a. X démontrer que

$$\forall x \in [0, 1], \quad \cos(1 - x) \geq 1 - \frac{\pi - 3}{x^2}$$

2) On suppose dans cette question que F est donnée par

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

a) Sans aucun calcul, dire ce que vaut $F(0.5)$.

b) Montrer que

$$\forall x \in \mathbb{R}_+, \quad \left(\frac{1}{x} - \frac{1}{x^3}\right)e^{-\frac{x^2}{2}} \leq \sqrt{2\pi}(1 - F(x)) \leq \frac{1}{x}e^{-\frac{x^2}{2}}.$$

On pourra comparer les dérivées des fonctions à gauche et à droite de ces inégalités.

Exercice 4

1) Soit X une v.a. de loi $\mathcal{N}(\mu, 1)$. Montrer que la fonction génératrice de moments (f.g.m.) de $X^2 : M_{X^2}(s) = \mathbb{E}(e^{sX^2})$ est donnée par

$$M_{X^2}(s) = \frac{1}{\sqrt{1 - 2s}} \exp\left(\frac{\mu^2 s}{1 - 2s}\right).$$

2) Soit $X_1, \dots, X_n$ n v.a.r indépendantes telles que pour tout $i = 1, \dots, n$, $X_i \sim \mathcal{N}(\mu_i, 1)$. On pose $Y = \sum_{i=1}^n X_i^2$. Montrer que la fonction caractéristique $\Phi_Y(t)$ de Y est donnée par

$$\Phi_Y(t) = \frac{1}{(1 - 2it)^{\frac{n}{2}}} \exp\left(\frac{\theta^2 it}{1 - 2it}\right)$$

où $\theta^2 = \mu_1^2 + \dots + \mu_n^2$. On pourra calculer d'abord la f.g.m. de Y . La loi de Y est dite loi de Chi-carré décentrée à n d.d.l. de paramètre de décentralisation θ notée $\chi^2(n; \theta)$.

Contrôle de Probabilités S6 UL - mai 2011. Durée 1H

Exercice 1

On jette 3 dés homogènes. Calculer les probabilités, des événements suivants :

- 1) Les trois dés montrent la face 6.
- 2) Au moins un dé montre la face 6.
- 3) Aucun dé ne montre les faces 5 ou 6.
- 4) La somme des trois dés est inférieure ou égale à 6.
- 5) On obtient que des 5 ou des 6 sachant que la somme est supérieure ou égale à 7.

Exercice 2

Un quart d'une population a été vacciné contre une certaine épidémie. On a trouvé qu'un malade sur cinq avait été vacciné et que parmi les personnes vaccinées, un sur douze a contracté la maladie. On choisit au hasard une personne dans cette population. Quelle est la probabilité

- 1) qu'elle soit malade ?
- 2) Qu'elle soit malade sachant qu'elle n'a pas été vaccinée ?

Exercice 3

Soit $M = \{1, 2, \dots, n\}$ et soit $\{i_1, i_2, \dots, i_k\}$ un k -uplet d'éléments de M . Pour $k = 0, 1, 2, \dots, n$, donner le cardinal de chacun des sous ensembles d'éléments de M suivants :

A_k l'ensemble des k -uplets d'éléments de M sans répétition,

B_k l'ensemble des k -uplets d'éléments de M avec répétition,

C_k l'ensemble des k -uplets d'éléments ordonnés de M sans répétition,

D_k l'ensemble des k -uplets d'éléments ordonnés de M .

Exercice 4

Soit $\Omega = \mathbb{R}$ et les événements $A_n = \left[0, 1 + \frac{(-1)^{n+1}}{n}\right]$, $n = 1, 2, \dots$

Décrire les événements $A^* = \limsup A_n$ et $A_* = \liminf A_n$.

La suite (A_n) est-elle convergente?

Exercice 5

Soit X une variable aléatoire réelle.

1) On suppose que X suit une loi binomiale $\mathcal{B}(n, p)$, $0 < p < 1$ définie par $\forall k = 0, 1, \dots, n$, $P[X = k] = C_n^k p^k (1-p)^{n-k}$. Calculer $\mathbb{E}(g(X))$ pour

a) $g(x) = e^{itx}$, $t \in \mathbb{R}$; b) $g(x) = \frac{1}{1+X}$. On simplifiera le résultat obtenu.

2) On suppose que la loi de X admet une densité donnée par

$$f(x) = \frac{C}{e^x + e^{-x}}, \quad \forall x \in \mathbb{R}.$$

Déterminer la constante C .

Examen et devoirs S6 UK 2013

Exercice 1

Soient $A_1, A_2, \dots, A_n$ une suite d'événements liés à une même expérience aléatoire.

1) Démontrer que

$$P(A_1.A_2...A_n) = P(A_1).P(A_2/A_1).P(A_3/A_1A_2) \cdots P(A_n/A_1.A_2...A_{n-1}).$$

où AB signifie $A \cap B$.

2) Une urne contient a boules blanches et b boules rouges. On en extrait successivement et avec remise n boules de l'urne.

a) Soit X la v.a. égale au nombre de boules blanches obtenues au cours de ces n tirages. Calculer la probabilité que $X = k$, pour tout $k \in \{0, 1, \dots, n\}$.

b) On décide maintenant qu'avant de procéder au tirage suivant on rajoute dans l'urne, c nouvelles boules de même couleur que celle qui a été précédemment tirée et remise dans l'urne. Calculer encore la probabilité que $X = k$, pour tout $k \in \{0, 1, \dots, n\}$.

c) Dans le cas a) calculer $\mathbb{E}(X)$ et $\text{Var}(X)$. On pourra poser $p = \frac{a}{a+b}$.

Exercice 2

Soit X et Y deux variables aléatoires réelles dont la loi conjointe a pour fonction de répartition

$$F(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y)}, & \text{si } x \geq 0, y \geq 0 \\ 0, & \text{sinon} \end{cases}$$

a) Calculer $\text{cov}(X, Y)$. Les v.a. X et Y sont-elles indépendantes?

b) Calculer $P[X < 1, Y \geq 1]$.

Exercice 3

Soit X et Y deux v.a.r. indépendantes et de même loi de densité

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

On pose $Z = \frac{X}{Y}$. Trouver la densité de la v.a. Z et la loi suivie par Z ?

Exercice 4

Soit X un vecteur gaussien de $\mathbb{R}^3$ de loi $\mathcal{N}(\vec{m}, \Gamma)$. On suppose que $\vec{m} = 0$ et

$$\Gamma = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

- Trouver un opérateur linéaire f de $\mathbb{R}^3$ dans lui-même telle que $Y = f(X)$ soit un vecteur aléatoire dont les composantes sont indépendantes.
- La loi de X est-elle absolument continue? Si oui, donner la densité $f_X(x)$ pour tout $x \in \mathbb{R}^3$. Sinon expliquer pourquoi.

Exercices 5

On dispose de trois urnes U_1, U_2, U_3 contenant chacune des boules blanches (B), noires (N) et rouges (R) avec les compositions :

U_1 : 1 B, 2 N, 3 R; U_2 : 2 B, 1 N, 1 R; U_3 : 4 B, 5 N, 3 R.

On prend une urne au hasard, on tire 2 boules de cette urne sans remise et on constate que l'une est blanche et l'autre est rouge. Quelles sont les probabilités q_1, q_2 et q_3 que l'urne tirée ait été U_1, U_2, U_3 respectivement?

Exercices 6

Soit X une v.a.r. absolument continue dont la loi de probabilité est donnée par la densité

$$f(x) = \lambda e^{-|x|}, \quad \lambda \in \mathbb{R}_+^*.$$

- Déterminer λ et la fonction de répartition F de la v.a. X .
- Déterminer la fonction de répartition et la densité de la v.a. $Y = X^2$.

UNIVERSITÉ DE LOMÉ
FACULTÉ DES SCIENCES

ANNÉE UNIVERSITAIRE
2008/2009

Licence de Mathématiques (SMIII)- DSn°3
Épreuve de Probabilités 1. Durée : 2H

N.B. : Le candidat traite obligatoirement le problème 1 et au choix un des problèmes 2 et 3.

Problème 1

Soit X et Y deux v.a.r de loi conjointe continue de densité

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right\}$$

où ρ est le coefficient de corrélation de X et Y .

1) On pose $U = X + Y, V = X - Y$. Montrer que U et V sont indépendantes et trouver leurs lois marginales.

2) On pose $Z = \frac{Y - \rho X}{\sqrt{1-\rho^2}}$.

- a) Montrer que X et Z sont indépendantes, chacune de loi normale centrée réduite.
b) Montrer que $P[X > 0, Y > 0] = \frac{1}{4} + \frac{1}{2\pi} \arctan \frac{\rho}{\sqrt{1-\rho^2}}$.

Problème 2

Soit X une v.a.r dont la loi a pour fonction caractéristique (f.c.)

$$\Phi_X(t) = (1 + |t|)e^{-|t|}, \quad t \in \mathbb{R}.$$

- 1) Discuter l'existence des moments d'ordre $k \in \mathbb{N}^*$ de la v.a. X .
2) On pose $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$ où $X_1, \dots, X_n$ sont n v.a.r indépendantes, chacune de même loi que X .
a) Déterminer la f.c. de Y_n et étudier sa limite éventuelle lorsque n tend vers $+\infty$. Que peut-on en conclure ?
b) Le résultat obtenu dans la question 2a) était-il prévisible ? si oui expliquer pourquoi.
3) On pose $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$
a) Déterminer la f.c. de Z_n et étudier sa limite éventuelle lorsque n tend vers $+\infty$. Que peut-on en conclure ?
b) Le résultat obtenu dans la question 3a) était-il prévisible ? si oui expliquer pourquoi.
4) Montrer que X est une v.a. absolument continue et déterminer sa densité de probabilité.
5) On pose enfin $Y = \arctan X$. Loi de la v.a. Y et existence de moments d'ordre $k \in \mathbb{N}^*$ de Y ?

Problème 3

Soit $(a, b) \in \mathbb{R}^+ \times \mathbb{R}$ et la loi de probabilité sur $\mathbb{R}$ notée $L(a, b)$ définie par sa fonction de répartition :

$$F(x) = \begin{cases} \frac{1}{2}e^{\frac{x-b}{a}} & \text{si } x \leq b, \\ 1 - \frac{1}{2}e^{-\frac{x-b}{a}} & \text{si } x > b. \end{cases}$$

- 1) Montrer que $F(x)$ est la fonction de répartition d'une loi absolument continue et déterminer la densité de probabilité correspondante.
2) Soit X une variable aléatoire réelle de loi $L(a, b)$. Calculer $\mathbb{E}(X)$ et $\text{Var}(X)$. Soit Y la variable aléatoire centrée réduite associée à X .
3) Déterminer la densité de Y et sa fonction caractéristique. En déduire, pour tout $k \in \mathbb{N}^*$, le moment d'ordre k de Y : $m_k = \mathbb{E}(Y^k)$.
4) Déterminer la fonction caractéristique de X .
5) Soient X_1 et X_2 deux variables aléatoires réelles indépendantes chacune, de loi $L(1, 0)$. On pose $S = X_1 + X_2$. Déterminer la densité de probabilité de S ainsi que $\mathbb{E}(S)$, $\text{Var}(S)$ et la fonction caractéristique de S .
6) On pose $Z = \text{partie entière de } |X_1|$. Déterminer la loi de probabilité de Z puis sa fonction génératrice et en déduire, pour tout $k \in \mathbb{N}^*$, le moment factoriel d'ordre k : $m_{[k]} = \mathbb{E}(\prod_{i=0}^{k-1} (Z - i))$.

Université de Kara

FaST - Semestre VI

Année 2011-2012

Probabilités III - DS. Durée 2h

Exercice 1

Soit X une v.a.r. intégrable et soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$. On pose $\mu = \mathbb{E}(X)$.

- 1) On suppose que f est convexe et deux fois dérivable. En utilisant le développement de Taylor, montrer que

$$f(X) \geq f(\mu) + f'(\mu)(X - \mu).$$

En déduire que si les v.a. X et $f(X)$ ont chacune une espérance finie alors on a $\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$.

2) On suppose dans cette question que f est paire, positive et croissante.

a) Montrer que pour toute constante positive c ,

$$P[|X| > c] \leq \frac{\mathbb{E}(f(X))}{f(c)}.$$

b) En déduire que pour tout $p > 0$

$$P[|X - \mathbb{E}(X)| > c] \leq \frac{\mathbb{E}|X - \mathbb{E}(X)|^p}{f(c)}.$$

Exercice 2

1) Soit X une v.a. suivant la loi normale centrée réduite. Montrer que pour tout $x > 0$,

$$\left(\frac{1}{x} - \frac{1}{x^3}\right) \exp\left(-\frac{1}{2}x^2\right) \leq \sqrt{2\pi}P[X \geq x] \leq \frac{1}{x} \exp\left(-\frac{1}{2}x^2\right).$$

2) Soit $(X_n)_{n \geq 0}$ une suite de v.a.r. absolument continues ayant pour densité f_n pour tout n et soit X une v.a.r. absolument continue de densité f .

a) Montrer que si $f_n \rightarrow f$ p.p. alors (X_n) tend vers X en loi. (Pour cela on pourra montrer que $\forall x \in \mathbb{R}, P[X_n \leq x] \rightarrow P[X \leq x]$).

b) Étudier la réciproque en considérant la suite $f_n(x) = (1 + \sin 2n\pi x)\mathbb{I}_{[0,1]}(x)$.

Exercice 3

Soit X_n une suite de v.a.r.

a) On suppose que la v.a. X_n prend ses valeurs dans l'ensemble $\left\{0, \frac{1}{n}, \dots, \frac{k}{n}, \dots, \frac{n-1}{n}, 1\right\}$ avec $\forall k, P[X_n = \frac{k}{n}] = \frac{1}{n+1}$. Étudier la convergence en loi de la suite (X_n) .

b) On suppose que X_n suit la loi $\gamma(n, \lambda)$. On pose $Y_n = \frac{\lambda}{\sqrt{n}}X_n - \sqrt{n}$. Étudier la convergence en loi de la suite (Y_n) .

c) On suppose que les X_n sont indépendantes et de même loi $\mathcal{U}[1, 2]$. Calculer $\mathbb{E}\left(\frac{X_1^2 + \dots + X_n^2}{X_1 + \dots + X_n}\right)$.

UNIVERSITÉ DE LOMÉ
FACULTÉ DES SCIENCES

ANNÉE UNIVERSITAIRE
2007/2008

Session de juillet 2008 - Examen : Licence de Mathématiques (SMIII)

Épreuve de Probabilités 1. Durée : 2H

Exercice : 8 pts

Soient X et Y deux v.a.r discrètes vérifiant les conditions suivantes :

- la loi de Y est la loi binomiale de paramètre n et $p, n \in \mathbb{N}^*$ et $p \in]0, 1[$.

- X ne peut prendre que les valeurs 0 et 1 et $\forall k \in \{0, 1, \dots, n\} : P[X = 1|Y = k] = \frac{k}{n}$.

a) Loi de X ? b) Calculer $\mathbb{E}(XY)$ c) Loi de $Y - 1$ sachant l'événement $\{X = 1\}$?

Problème : 12pts

Soit X une v.a.r dont la loi a pour fonction caractéristique (f.c.)

$$\Phi_X(t) = (1 + |t|)e^{-|t|}, \quad t \in \mathbb{R}.$$

1) Discute l'existence des moments d'ordre $k \in \mathbb{N}^*$ de la v.a. X .

2) On pose $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$ où $X_1, \dots, X_n$ sont n v.a.r indépendantes, chacune de même loi que X .

a) Détermine la f.c. de Y_n et étudie sa limite éventuelle lorsque n tend vers $+\infty$. Que peut-on

en conclure ?

b) Le résultat obtenu dans la question 2a) était-il prévisible ? si oui expliquer pourquoi.

3) On pose $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$

a) Détermine la f.c. de Z_n et étudie sa limite éventuelle lorsque n tend vers $+\infty$. Que peut-on en conclure ?

b) Le résultat obtenu dans la question 3a) était-il prévisible ? si oui expliquer pourquoi.

4) Montre que X est une v.a. absolument continue et détermine sa densité de probabilité.

5) On pose enfin $Y = \arctan X$. Loi de la v.a. Y et existence de moments d'ordre $k \in \mathbb{N}^*$ de Y ?

UNIVERSITÉ DE LOMÉ
FACULTÉ DES SCIENCES

ANNÉE UNIVERSITAIRE
2007/2008

Licence de Mathématiques (SMIII)- Examen : Session de rattrapage
Epreuve de Probabilités 1. Durée : 3H

N.B. : Le candidat traitera deux exercices au choix parmi les exercices 1, 2, 3 et obligatoirement l'exercice n° 4.

Exercice 1 : 8 pts

Soient $X_1, \dots, X_n$ une suite de v.a. à valeurs dans $\mathbb{N}^*$, indépendantes et de même loi de fonction génératrice (f.g.) $G(s)$ et soit N une v.a. à valeurs dans $\mathbb{N}^*$, indépendante de la suite (X_n) et de f.g. $H(s)$.

1) On pose $Z = \sum_{i=1}^N X_i$. Trouve en fonction de G et H la f.g. de la v.a. Z et en déduit $\mathbb{E}(Z)$ et $\text{Var}(Z)$.

2) A.N. : Le nombre d'accidents de la route enregistrés par an au Togo est une v.a. de moyenne μ et de variance σ^2 . Le nombre de sinistres lors de chaque accident est une v.a. de moyenne ν et de variance ρ^2 . On suppose que les v.a. nombres de sinistres lors de divers accidents sont indépendantes du nombre d'accidents. Calculer la moyenne et l'écart-type du nombre de sinistres enregistrés par an au Togo.

Exercice 2 : 8 pts

Soit X et Y deux v.a.r de loi conjointe continue de densité

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right\}$$

1) On pose $U = X + Y$, $V = X - Y$. Montrer que U et V sont indépendantes et trouver leurs lois marginales.

2) On pose $Z = \frac{Y - \rho X}{\sqrt{1-\rho^2}}$.

a) Montrer que X et Z sont indépendantes, chacune de loi normale centrée réduite.

b) En déduire que $P[X > 0, Y > 0] = \frac{1}{4} + \frac{1}{2\pi} \arctan \frac{\rho}{\sqrt{1-\rho^2}}$.

Exercice 3 : 8pts

Soit $(X_n)_{n \geq 1}$ une suite de v.a.r définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ telle que pour tout $n \geq 1$, $X_n \rightsquigarrow \mathcal{B}(n, x)$ (loi binomiale de paramètres n et $x \in]0, 1[$).

1) Montrer que la suite $(\frac{X_n}{n})_{n \geq 1}$ converge en moyenne quadratique vers x lorsque $n \rightarrow +\infty$.

2) Montrer que $\forall n \geq 1, \forall x \in]0, 1[$ et $\forall \delta > 0$

$$\sum_{\{k \in \mathbb{N} / 0 \leq k \leq n \text{ et } |\frac{k}{n} - x| > \delta\}} C_n^k x^k (1-x)^{n-k} \leq \frac{1}{4n\delta^2}$$

3) Soit f une fonction continue sur l'intervalle $[0, 1]$ à valeurs réelles. Pour tout $n \in \mathbb{N}^*$ et tout

$x \in]0, 1[$, on pose $f_n(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}$.

Vérifier que $f_n(x)$ est un polynôme (polynôme de Bernstein) et démontrer que $\lim_{n \rightarrow +\infty} \sup_{0 \leq x \leq 1} |f_n(x) - f(x)| = 0$.

En déduire le Théorème de Stone-Weierstrass : toute fonction continue sur $[0, 1]$ à valeurs réelles est limite uniforme d'une suite de polynômes.

Exercice 4 : 4 pts

Soit $(X_n)_{n \geq 1}$ une suite de v.a.r telle que pour tout $n \geq 1$, $X_n \rightsquigarrow \mathcal{E}(\frac{1}{n})$ (loi exponentielle de paramètre $\frac{1}{n}$). On pose $Y_n = X_n - [X_n]$ où $[.]$ désigne la partie entière de "." Montrer que la suite (Y_n) converge en loi vers une v.a. absolument continue dont on précisera la loi.

UNIVERSITE DE LOME
FACULTE DES SCIENCES

ANNEE UNIVERSITAIRE
2008/2009

Devoir de Probabilités - Licence de Mathématiques (SMIII)
29 janvier 2009 - Durée : 1H

N.B. : Le candidat traitera obligatoirement un exercice au choix parmi les deux exercices proposés.

Exercice 1

Soient (X, Y) un couple de v.a.r discrètes de loi conjointe donnée par

$$P[X = x, Y = y] = \begin{cases} \left(\frac{1}{2}\right)^{x+y} & \text{si } x = 1, 2, \dots, \infty, y = 1, 2, \dots, \infty \\ 0 & \text{sinon} \end{cases}$$

- 1) Déterminer la loi marginale de X et celle de Y . Les v.a. X et Y sont-elles indépendantes ?
- 2) Déterminer la fonction de répartition (f.r.) de X : $F_X(x) = P[X \leq x]$, $x \in \mathbb{R}$.
- 3) Déterminer $P[X > a, Y < b]$, $a, b \in \mathbb{R}$.
- 4) Déterminer $P[\min(X, Y) < c]$, $c \geq 1$.
- 5) Déterminer $P[X = Y]$ et $P[Y > X]$.

Exercice 2

Soit X une v.a.r admettant pour densité de probabilité la fonction f définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} x + 1 & \text{si } |x| \leq k \\ 0 & \text{si } |x| \geq k \end{cases}$$

- 1) Déterminer k . et trouver la f.r. de la v.a. X .
- 2) Calculer $\mathbb{E}(X^2)$.
- 3) On pose $Y = X^2$. Trouver la loi de Y . On donnera la f.r. et la densité de cette loi.
- 4) Calculer $\mathbb{E}(Y)$. Que constatez-vous ?

UNIVERSITÉ DE LOMÉ
FACULTÉ DES SCIENCES

ANNÉE UNIVERSITAIRE
2008/2009

Examen du premier semestre - SMIII
Épreuve de Probabilité I - Durée 2H

N.B. : Le candidat traitera obligatoirement les exercices 1 et 2 et deux exercices au choix parmi les exercices 3, 4 et 5.

Exercice 1

1) Soit X une v.a.r. de loi continue de fonction de répartition $F(x)$ et de densité $f(x) = F'(x)$. Soit $X_1, X_2, \dots, X_n$ des v.a.r indépendantes et de même loi que X . On pose $U = \min(X_1, X_2, \dots, X_n)$ et $V = \max(X_1, X_2, \dots, X_n)$

Donner sans détailler les calculs, les fonctions de répartitions respectives des v.a. U et V .

2) Dans l'univers considéré comme une boule de centre O et de rayon R , N astres sont distribués de façon aléatoire et indépendante.

a) Quelle est la probabilité pour que la distance entre le centre de l'univers et l'astre le plus proche ne soit pas inférieure à un nombre a , $0 < a < R$?

b) Quelle est la limite de cette probabilité lorsque $R \rightarrow +\infty$ et $\frac{N}{R^3} \rightarrow \frac{4\pi\lambda}{3}$ où λ est une constante positive? On rappelle que le volume $v(r)$ d'une boule de centre O et de rayon r est égale à $\frac{4\pi r^3}{3}$.

Exercice 2

Le nombre N d'oeufs pondus par un insecte est une v.a. suivant une loi de Poisson de paramètre $\lambda > 0$. Les oeufs se développent de façon indépendante et chacun d'eux donne naissance à un nouvel insecte avec une probabilité p , $0 < p < 1$.

1) Loi de probabilité du nombre X d'insectes à la première génération?

2) On observe k insectes à la première génération. Quelle est la probabilité qu'il y ait eu n oeufs pondus?

Exercice 3

Soit (X, Y) un couple de v.a. continues dont la loi conjointe admet pour densité

$$f(x, y) = \begin{cases} \frac{3y}{a^3} & \text{si } x + y < a, y - x < a \text{ et } y > 0 \\ 0 & \text{sinon} \end{cases}$$

où a est un réel strictement positif. Calculer $\text{cov}(X, Y)$. Les v.a. X et Y sont-elles indépendantes?

Exercice 4

Soient U et V deux v.a. indépendantes de même loi uniforme sur $[0, 1]$. On pose $X = \sqrt{-2 \log U} \cos 2\pi V$ et $Y = \sqrt{-2 \log U} \sin 2\pi V$. Montrer que les v.a. X et Y sont indépendantes et trouver leurs lois.

Exercice 5

Soit X une v.a. suivant la loi de densité

$$f(x) = \begin{cases} \frac{\alpha-1}{x^\alpha} & \text{si } x \geq 1 \\ 0 & \text{sinon} \end{cases}$$

où α est un réel strictement supérieur à 1.

Soient $X, X_1, X_2, \dots, X_n$ des v.a.r indépendantes et de même loi que X . On pose

$$Z_n = 1 + \frac{n}{\sum_{i=1}^n \log X_i}.$$

Calculer $\mathbb{E}(Z_n)$ et trouver une v.a. Θ_n satisfaisant $\mathbb{E}(\Theta_n) = \alpha$. On pourra poser $Y_i = \log X_i$.

Exercice 1 :

On dispose de 3 urnes U_1, U_2, U_3 contenant chacun des boules blanches (B), noires (N) et rouges (R) avec les compositions suivantes :

U_1 : 1B, 2N, 3R;

U_2 : 2B, 1N, 1R;

U_3 : 4B, 5N, 3R;

On prend une urne au hasard, on tire sans remise 2 boules de cette urne et on constate que l'une est blanche et l'autre rouge. Quelle est la probabilité p_i que l'urne ait été $U_i, i = 1, 2, 3$?

Exercice 2

Soit X une variable aléatoire réelle (v.a.r.) absolument continue dont la loi de probabilité est définie par la densité

$$f(x) = ke^{-|x|}, \quad k \in \mathbb{R}_+^*.$$

1) Déterminer k ainsi que la fonction de répartition (f.r.) F de X .

2) Déterminer successivement la f.r. puis la densité de probabilité de chacune des v.a.r. suivantes :

$$Y = e^X, \quad Z = X^2, \quad T = \frac{1}{X} \quad \text{et} \quad V = \int_{-\infty}^X \frac{1}{2} e^{|x|} dx.$$

Exercice 3

1) Soit (X, Y) un couple de v.a.r. admettant pour densité de probabilité

$$f(x, y) = ce^{-\frac{x^2+y^2}{2}}, \quad \forall x, y \in \mathbb{R}.$$

a) Déterminer c .

b) Trouver les lois marginales de X et de Y . Les v.a.r. X et Y sont-elles indépendantes?

c) Calculer $\text{cov}(X, Y)$.

d) Déterminer les lois des v.a. $U = X + Y$ et $V = X - Y$. U et V sont-elles indépendantes?

2) Soient $X_1, X_2, \dots, X_n$ des v.a.r. indépendantes de même loi $\mathcal{N}(0, 1)$. Soient $a_1, a_2, \dots, a_n$ et

$b_1, b_2, \dots, b_n$ des réels. Montrer que les v.a. $Y = \sum_{i=1}^n a_i X_i$ et $Z = \sum_{i=1}^n b_i X_i$ sont indépendantes si

et seulement si $\sum_{i=1}^n a_i b_i = 0$.

NB : On attachera de l'importance à la justification et à la précision des réponses.

UNIVERSITÉ DE KARA
FACULTÉ DES SCIENCES ET TECHNIQUES

ANNÉE UNIVERSITAIRE
2009/2010

NB : Le candidat traitera obligatoirement les exercices et au choix un des deux problèmes proposés.

Exercices

1) Soit U une v.a. de loi uniforme sur $[0, 1]$. Loi de la v.a. $Y = -\frac{1}{\lambda} \log(1 - U)$, $\lambda > 0$?

2) Soient X et Y deux v.a.r. de loi conjointe définie par la densité

$$f(x, y) = \begin{cases} e^{-y}, & \text{si } 0 \leq x \leq y; \\ 0, & \text{sinon.} \end{cases}$$

Calculer $\text{cov}(X, Y)$. X et Y sont-elles indépendantes? Calculer le coefficient de corrélation linéaire $\rho(X, Y)$ entre X et Y .

Problème 1

Soit X et Y deux v.a.r indépendantes et de même loi $\mathcal{N}(0, 1)$. On définit la v.a. $Z = \frac{X}{Y}$.

- 1) a) Loi de Z ? Z a-t-elle des moments de tous ordres?
- b) Comparer $P[Z > 1]$ et $P[\mathcal{N}(0, 1) > 1]$. ϕ étant la f.r. de la loi $\mathcal{N}(0, 1)$, on donne $\phi(1) = 0.8413$.
- 2) $Z_1, \dots, Z_n$ étant n v.a.r. indépendantes et de même loi que Z , quelle est la loi de $\bar{Z}_n = \frac{1}{n} \sum_{i=1}^n Z_i$? (On donne la fonction caractéristique de Z : $\phi_Z(t) = e^{-|t|}$).
- 3) On pose $U = |X|$.
 - a) Trouver la loi de la v.a. U . On pourra calculer $\mathbb{E}(g(U))$ pour toute fonction borélienne positive ou bornée.
 - b) Calculer $\mathbb{E}(U)$.

Problème 2

Soit $X_1, X_2, \dots, X_n$, n v.a.r. de loi absolument continues.

- 1) On suppose que ces variables aléatoires sont indépendantes et de même loi $\mathcal{N}(0, 1)$.

- a) Calculer la fonction caractéristique (f.c.) de la v.a. $Y_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n X_k$.

En déduire la loi de Y_n .

- b) Soit N une v.a. à valeurs dans $\mathbb{N}^*$, indépendante des variables X_i . On pose $Y = \frac{1}{\sqrt{N}} \sum_{k=1}^N X_k$.

déterminer la f.c. de Y ainsi que sa loi. (On pourra poser $e^{itY} = \sum_{n \geq 1} e^{itY} \mathbb{1}_{\{N=n\}}$).

- 2) On suppose dans cette question que pour tout $n \geq 1$, X_n a pour densité $f_n(x) = \frac{n^2}{2} e^{-n^2|x|}$, $\forall x \in \mathbb{R}$.

- a) Calculer pour tout $n \in \mathbb{N}^*$, $P[|X_n| > n^{-\frac{3}{2}}]$.

- b) Déterminer $P[\limsup\{|X_n| > n^{-\frac{3}{2}}\}]$.

(On pourra considérer les événements $A_n = \{\omega \in \Omega / |X_n(\omega)| > n^{-\frac{3}{2}}\}$).

UNIVERSITE DE LOME
FACULTE DES SCIENCES

ANNEE UNIVERSITAIRE
2005/2006

Licence de Mathématiques (SMIII) : session de septembre 2006

Examen : Épreuve de probabilités Durée : 2H, coeff. 2.

N.B. : Le candidat traitera obligatoirement les questions de cours et au choix un des deux problèmes proposés.

Questions de cours

- 1) Énoncer sans démonstration la loi forte des grands nombres pour une suite de variables aléatoires indépendantes et de même loi.
- 2) Soient X et Y deux v.a. définies sur un même espace probabilisé et possédant des variances finies non nulles.
 - a) Donner la définition du coefficient de corrélation $r(X, Y)$ de (X, Y) .
 - b) Si (X, Y) est de loi normale dans $\mathbb{R}^2$, expliquer pourquoi X et Y sont indépendantes lorsque $r(X, Y) = 0$. On pourra utiliser le critère de la densité ou celui des fonctions caractéristiques.
- 3) Soit μ et $(\mu_n)_{n \geq 1}$ des probabilités sur $\mathbb{R}$. On rappelle qu'on dit que la suite $(\mu_n)_{n \geq 1}$ converge

en loi vers μ si pour toute fonction continue, bornée, réelle ou complexe sur $\mathbb{R}$, on a

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f(x) \mu_n(dx) = \int_{\mathbb{R}} f(x) \mu(dx).$$

Donner sans démonstration, en termes de transformés de Fourier, une CNS (Condition Nécessaire et Suffisante) de convergence en loi et rappeler le nom du théorème ainsi énoncé.

4) Énoncer sans démonstration une CS de convergence d'une suite de v.a.r $(X_n)_{n \geq 1}$ vers une v.a.r X :

- a) en probabilité ;
- b) presque sûrement.

Problème 1

Soit $(X_n)_{n \geq 1}$, une suite de v.a.r indépendantes et de même loi $\mathcal{N}(m, \sigma)$. On pose

$$\varphi(u) = \mathbb{E}[e^{iu(X_1 - m)}] \text{ et } S_n = X_1 + X_2 + \dots + X_n.$$

1) Calculer $\varphi(u)$. On pourra utiliser la fonction caractéristique de la loi $\mathcal{N}(m, \sigma)$:

$$\varphi_{\mathcal{N}(m, \sigma)}(t) = e^{itm - \frac{1}{2}\sigma^2 t^2}$$

2) Soit $a \in \mathbb{R}$. On pose $B_a^- = \{S_n - nm \leq a\}$ et $B_a^+ = \{S_n - nm \geq a\}$.

1) En utilisant la monotonie de la fonction exponentielle, montrer que

$$P[B_a^s] \leq e^{-ua} (\varphi(u))^n, \forall u \in \mathbb{R}^s \text{ pour } s = +, -.$$

3) En déduire que $\forall \epsilon > 0, P[|\frac{S_n}{n} - m| \geq \epsilon] \leq 2 \exp(-\frac{n\epsilon^2}{2\sigma^2})$

On pourra écrire l'événement $\{|S_n - nm| \geq n\epsilon\}$ comme réunion de B_a^- et de B_a^+ avec a convenable.

4) En choisissant $\epsilon = \sigma \sqrt{2A \frac{\text{Log} n}{n}}$, avec $A > 1$ et $n > 1$, montrer que $\frac{S_n}{n}$ converge presque sûrement vers m et que $|\frac{S_n}{n} - m| = O\left(\frac{\text{Log} n}{n}\right)$ p.s.

Problème 2

Soit X la durée de vie aléatoire d'un appareil électronique, continue de fonction de répartition F et de densité f . On suppose qu'après une panne, l'appareil est réparé et la durée de vie après réparation a même loi de probabilité que X .

Soit $X_1, X_2, \dots, X_n$ les durées de vie indépendantes, de l'appareil jusqu'à la n -ième panne. On définit $S_n = X_1 + X_2 + \dots + X_n$ le moment de la n -ième panne, $N(t)$ le nombre de pannes observées dans l'intervalle $[0, t]$, F_n et f_n la f.r. et la densité de la v.a. S_n .

1) Montrer que

$$a) P[N(t) = n] = F_n(t) - F_{n+1}(t), \forall n \geq 1.$$

$$b) H(t) = \mathbb{E}(N(t)) = \sum_{n=1}^{+\infty} F_n(t)$$

2) Soit $W(t) = S_{N(t)} - t$ le temps écoulé entre le moment t et la première panne après t . Montrer que si f_n et $\sum f_n$ sont continues sur $\mathbb{R}^+$, alors

$$F_{W(t)}(s) = P[W(t) \leq s] = F(t+s) - \int_0^t (1 - F(t+s-x)) dH(x), \forall s \geq 0.$$

3) Lois respectives de $N(t)$ et de $W(t)$ si les X_i suivent la loi exponentielle de paramètre $\lambda > 0$?

Problème 1 (6pts)

Soit $n \geq 2$ et soient $U_1, \dots, U_n$, une suite de v.a.r. indépendantes et de même loi uniforme sur $[0, 1]$. On pose $X = \min(U_1, \dots, U_n)$ et $Y = \max(U_1, \dots, U_n)$.

On fixe a et b tels que $0 \leq a \leq b \leq 1$.

1) Dessiner le sous-ensemble du plan $E_{a,b} = \{(x, y) \in \mathbb{R}^2 / 0 \leq a \leq x \leq y \leq b \leq 1\}$.

2) Calculer $G(a, b) = P[0 \leq a \leq X \leq Y \leq b \leq 1]$.

3) Montrer que $P[X \leq a, Y \leq b] = G(0, b) - G(a, b)$.

En déduire la densité du couple (X, Y) .

4) Soit $Z = Y - X$. Trouver la densité de la loi du couple (X, Z) .

Problème 2 (6pts)

Soit (X_n) une suite de v.a. indépendantes et de même loi

$P[X_n = 1] = p$ et $P[X_n = -1] = 1 - p$, $0 \leq p \leq 1$.

On pose $S_n = X_1 + X_2 + \dots + X_n$.

1) Calculer $P[S_n = 0]$, $E(S_n)$ et $\text{Var}(S_n)$.

2) Calculer $P[\limsup_{n \rightarrow +\infty} \{S_n = 0\}]$.

3) Montrer que pour tout $\epsilon > 0$, $P[|\frac{S_n}{n} - p(1 - p)| > \epsilon] \leq \frac{1}{n\epsilon^2}$.

En déduire que $\frac{S_n}{n}$ converge en probabilité vers $p(1 - p)$.

Problème 3 (6pts)

Le nombre N d'oeufs pondus par un insecte suit la loi de Poisson de paramètre $\lambda > 0$. Les oeufs se développent de façon indépendante et chacun d'eux donne naissance à un nouvel insecte avec probabilité a , $0 < a < 1$.

1) Loi du nombre X d'insectes à la première génération ?

2) On observe k insectes à la première génération. Quelle est la probabilité pour qu'il y ait eu n oeufs pondus ?

Problème 4 (6pts)

1) Soit X une v.a.r de fonction de répartition F continue et strictement croissante.

a) Loi de la v.a. $F(X)$?

b) Réciproquement soit U une v.a. de loi uniforme sur $[0, 1]$. Loi de $F^{-1}(U)$?

2) Soient U et V deux v.a. indépendantes telles que U suit la loi uniforme sur $[0, 2\pi]$ et V suit la loi exponentielle de paramètre $\frac{1}{2}$. On pose

$X = \sqrt{V} \cos U$ et $Y = \sqrt{V} \sin U$

Loi de (X, Y) , de X et de Y respectivement ?

3) Déduire des questions précédentes, comment pourrait-on simuler un échantillon de n variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{N}(0, 1)$ lorsqu'on dispose d'un générateur de "nombres au hasard" dans l'intervalle $[0, 1]$. Un algorithme et un programme en Pascal donnent un bonus de 2 points mais ne sont pas exigés du candidat. On donnera le nom "RANDOM" à la procédure de génération de nombre au hasard.