

THÉORIE DES PROBABILITÉS

La Théorie des Probabilités constitue un ensemble de modèles mathématiques destinés à l'étude des phénomènes dont l'évolution est gérée par le hasard.

L'homme est souvent confronté à des expériences dont il ne connaît pas les résultats avec certitude. Afin de maîtriser ce genre de phénomènes il adopte une méthode subjective qui consiste à mesurer le degré de certitude (ou d'incertitude), des issues auxquelles il fait face, par un nombre. Il appelle ce nombre "probabilité de réalisation de l'issue considérée". Cette méthode de mesure de l'incertain est différente de celle qui consiste à mesurer par exemple une distance entre deux points donnés. Parmi les expériences gérées par le hasard, l'une des plus simples est celle qui consiste à jeter une pièce de monnaie, donnant lieu à deux résultats possibles: pile ou face. Pour maîtriser ce type d'expérience on mesure chacun des résultats par un nombre que l'on convient d'appeler "probabilité de réalisation du résultat". L'objectif de la Théorie des Probabilités est de donner à cette mesure de l'incertain un contenu mathématique rigoureux permettant une maîtrise sans faille de son utilisation. Dans cet exposé nous avons adopté une présentation simple et intuitive de cette Théorie sans faire appel à la Théorie de la mesure et de l'intégration.

Chapitre 1

THÉORIE DES ENSEMBLES

1. Algèbre d'ensembles - Dénombrement

Définition 1.1: Soit Ω un ensemble arbitraire:
on note $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω , ϕ désigne l'ensemble vide.

Exemples:

- (a) ensemble de nombres: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$
- (b) ensemble de communes
- (c) ensemble de résultats d'une élection

Définition 1.2:

Opérations sur les ensembles: soient A, B, C, \dots des parties de Ω :
union $A \cup B = \{\omega : \omega \in A \text{ ou } \omega \in B\}$
intersection $A \cap B = \{\omega : \omega \in A \text{ et } \omega \in B\}$
complement $A^c = \{\omega : \omega \notin A\}$
inclusion $A \subset B$ qui signifie que A est un sous ensemble de B .

Propriétés 1.3:

$A \subset B \iff B^c \subset A^c \quad (A^c)^c = A$
 $(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$
 $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, et $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Définition 1.4: (Algèbre de Boole)

Soit $\mathcal{F} \subset \mathcal{P}(\Omega)$ une famille de parties de Ω . \mathcal{F} est une Algèbre de Boole si elle satisfait les conditions:

- (a) $\Omega \in \mathcal{F}$
- (b) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
- (c) Si A, B sont dans \mathcal{F} alors $A \cup B \in \mathcal{F}$

Remarque 1.5:

si \mathcal{F} est une Algèbre de Boole alors $\phi \in \mathcal{F}$

et pour toute suite finie $(A_k), 1 \leq k \leq n$ dans \mathcal{F} on a $\bigcap_{k=1}^n A_k \in \mathcal{F}$

Exemple:

- (a) $\mathcal{P}(\Omega)$ est une algèbre
- (b) $\mathcal{F} = \{\Omega, \phi\}$ est une algèbre

Définition 1.4: (σ -algèbre de Boole)

Soit $\mathcal{F} \subset \mathcal{P}(\Omega)$ une famille de parties de Ω . \mathcal{F} est une σ -algèbre de Boole si elle satisfait les conditions:

- (a) $\Omega \in \mathcal{F}$
- (b) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
- (c) Si (A_n) est une suite dans \mathcal{F} alors $\bigcup_n A_n \in \mathcal{F}$

le couple (Ω, \mathcal{F}) formé d'un ensemble Ω et d'une σ -algèbre de parties de Ω s'appelle espace probabilisable.

Il est évident que $\mathcal{P}(\Omega)$ est une σ -algèbre et que toute σ -algèbre est une algèbre. Voir exercice 2 ci-dessous pour un exemple d'algèbre qui n'est pas une σ -algèbre.

Exercices.

1. Soient $\mathcal{F}_1, \mathcal{F}_2$, 2 algèbres (resp. σ -algèbres) sur Ω
Montrer que $\mathcal{F}_1 \cap \mathcal{F}_2$ est une algèbre (resp. σ -algèbre).

2. Soit Ω un ensemble non vide. On considère les familles de parties:
 $\mathcal{F} = \{A \subset \Omega : A \text{ ou } A^c \text{ dénombrable}\}$ et $\mathcal{A} = \{A \subset \Omega : A \text{ ou } A^c \text{ fini}\}$
Montrer que \mathcal{F} est une σ -algèbre et que \mathcal{A} est une algèbre
Si Ω est infini \mathcal{A} n'est pas une σ -algèbre.

Définition 1.6: (Cardinaux)

Si Ω est un ensemble quelconque on définit le cardinal de Ω par:

$$|\Omega| = \begin{cases} n & \text{si } \Omega \text{ possède } n \text{ éléments} \\ \infty & \text{si } \Omega \text{ est infini} \end{cases}$$

Proposition 1.7:

Si $|\Omega| = n$ alors $|\mathcal{P}(\Omega)| = 2^n$ (*)

Preuve: Par récurrence sur n :

- (a) pour $n = 0$, on a $\Omega = \phi$ et $2^0 = 1$
la propriété (*) est vraie puisque $\mathcal{P}(\Omega) = \{\phi\}$.
- (b) supposons (*) vraie pour n et montrons qu'elle est vraie pour $n + 1$
supposons $|\Omega| = n + 1$ et écrivons Ω sous la forme $\Omega = \Omega_1 \cup \{\omega\}$ avec $|\Omega_1| = n$
et $\omega \notin \Omega_1$. Dans ce cas on a $\mathcal{P}(\Omega) = \{A \cup \{\omega\}, A \in \mathcal{P}(\Omega_1)\} \cup \mathcal{P}(\Omega_1)$ et
 $\{A \cup \{\omega\}, A \in \mathcal{P}(\Omega_1)\} \cap \mathcal{P}(\Omega_1) = \phi$ on déduit que $|\mathcal{P}(\Omega)| = 2^n + 2^n = 2^{n+1}$

Définition 1.8: (Produit Cartésien)

Soient Ω_1, Ω_2 deux ensembles, on définit le produit cartésien $\Omega_1 \times \Omega_2$ par:

$$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

La notation (ω_1, ω_2) est celle d'un couple, i.e $(\omega_1, \omega_2) \neq (\omega_2, \omega_1)$ si $\omega_1 \neq \omega_2$

Proposition 1.9:

Pour tous ensembles finis Ω_1, Ω_2 on a $|\Omega_1 \times \Omega_2| = |\Omega_1| \times |\Omega_2|$

Preuve: Posons $|\Omega_1| = m, |\Omega_2| = n$, avec

$\Omega_1 = \{a_1, a_2, \dots, a_m\}, \Omega_2 = \{b_1, b_2, \dots, b_n\}$; ainsi on aura:

$\Omega_1 \times \Omega_2 = \{(a_i, b_j), 1 \leq i \leq m, 1 \leq j \leq n\}$ que l'on peut écrire:

$$\begin{aligned} & (a_1, b_1), (a_1, b_2), \dots, (a_1, b_n) \\ & (a_2, b_1), (a_2, b_2), \dots, (a_2, b_n) \\ & \vdots \\ & (a_m, b_1), (a_m, b_2), \dots, (a_m, b_n) \end{aligned}$$

il est clair que l'on a au total $m.n$ couples, donc $|\Omega_1 \times \Omega_2| = |\Omega_1| \times |\Omega_2|$. ■

Exercices.

3. Si E est un ensemble quelconque on a défini (Définition 1.6) le cardinal de E par:

$$|E| = \begin{cases} n & \text{si } E \text{ possède } n \text{ éléments} \\ \infty & \text{si } E \text{ est infini} \end{cases}$$

(a) Montrez que $|A \cup B| = |A| + |B| - |A \cap B|$

(b) On rappelle que si $|\Omega| = n$, le nombre de parties de Ω ayant k éléments vaut $C_n^k = \frac{n!}{k!(n-k)!}$. Donner le nombre d'arrangements d'une partie de k

éléments. En déduire le nombre total d'arrangements de k éléments.

(c) Montrer que $|\mathcal{P}(\Omega)| = 2^n$, en utilisant le binôme de Newton:

$$\sum_{k=0}^n C_n^k a^k b^{n-k} = (a+b)^n.$$

(voir démonstration par récurrence Proposition 1.7)

4. Montrer par récurrence sur n que $\sum_{k=0}^n k = \frac{n(n+1)}{2}$.

Chapitre 2

ESPACES PROBABILISES

1. Experiences Aléatoires

L'objectif de ce chapitre est de construire un modèle mathématique pour gérer les phénomènes qui évoluent au hasard c'est à dire de façon imprévisible. Il s'agit de techniques mathématiques rigoureuses destinées à mesurer le caractère aléatoire de ces phénomènes en vue de décisions éventuelles concernant ces phénomènes. La construction du modèle commence par une description exhaustive du phénomène à étudier. C'est l'objet de ce qui suit.

1.1 Une expérience aléatoire est une expérience dont le résultat ne peut être connu d'avance avec certitude. Comme exemple on peut citer:

- lancer une pièce de monnaie, résultats $\{pile, face\}$
- lancer un dé, résultats $\{1, 2, 3, 4, 5, 6\}$
- prévoir la météo un jour donné
- nombre de pièces défectueuses dans un ensemble de pièces fabriquées dans un atelier.

Une expérience qui n'est pas aléatoire est dite déterministe, son résultat est connu d'avance avec certitude.

1.2. Ensemble fondamental d'une expérience aléatoire, c'est l'ensemble de tous les résultats possibles de l'expérience.

On note cet ensemble $\Omega = \{\omega : \omega \text{ résultat possible}\}$

1.3. Evénement, c'est par définition une partie de Ω

$\mathcal{P}(\Omega)$ est l'ensemble de tous les événements associés à l'expérience

l'événement A se réalise si on observe un résultat $\omega \in A$.

Exemple.(1) Avec l'expérience du dé on a $\Omega = \{1, 2, 3, 4, 5, 6\}$ et comme exemple l'événement A : "face paire" $A = \{2, 4, 6, \}$

(2) Une urne contient 12 boules numérotées de 1 à 12. On tire une boule au hasard. Dans cette expérience $\Omega = \{1, 2, 3, \dots, 12\}$ et si A est l'événement "boule de numéro multiple de 3" alors $A = \{3, 6, 9, 12\}$

1.4 Interpretation des opérations d'ensembles en termes d'événements

Soit Ω l'ensemble fondamental d'une expérience et soient A, B des événements:

Ω est l'événement certain car tous les résultats ω le réalisent

ϕ est l'événement impossible car aucun ω ne le réalise

$A \cap B$ est l'événement réalisé par l'un des ω réalisant A et B simultanément.

$A \cup B$ est l'événement réalisé par l'un au moins des événements A ou B

A^c est l'événement réalisé par l'un des résultats ω qui ne réalisent pas A

L'étape qui va suivre cette base descriptive du phénomène consiste à mesurer le caractère incertain des événements. C'est une étape importante de la construction du modèle qui doit respecter des conditions strictes d'application. Ce modèle mathématique est défini par la notion de mesure de probabilité.

2. Mesure de Probabilité

2.1 Définition

Une mesure de probabilité sur un ensemble Ω est une fonction

$P : \mathcal{P}(\Omega) \rightarrow [0, 1]$ satisfaisant les conditions:

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

(b). σ -additivité: pour toute suite $(A_n) \subset \mathcal{P}(\Omega)$, telle que:

$$A_i \cap A_j = \emptyset, \forall i \neq j, \text{ on a } P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$$

Le couple Ω, P s'appelle espace probabilisé.

Remarques

1. Pour des raisons techniques il arrive que P soit définie sur une σ -algèbre $\mathcal{F} \subset \mathcal{P}(\Omega)$ (voir cours de théorie de la mesure).

2. Il est facile de voir que $P(\emptyset) = 0$, en effet on a

$$\Omega = \Omega \cup \emptyset \text{ et } P(\Omega) = P(\Omega) + P(\emptyset) \implies P(\emptyset) = 0$$

on peut aussi remarquer que $P(A^c) = 1 - P(A)$, en effet on a $\Omega = A \cup A^c$ et $A \cap A^c = \emptyset$ donc $P(\Omega) = P(A) + P(A^c) = 1$, d'où $P(A^c) = 1 - P(A)$

2.2 Espace probabilisé discret:

L'espace Ω, P est discret si Ω est une suite finie ou infinie:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \dots\} \text{ avec } \sum_n P(\omega_n) = 1$$

Exemples.

(1) $\Omega = \{\text{pile}, \text{face}\}$ $P(\text{pile}) = \alpha$, $P(\text{face}) = 1 - \alpha$

(2) $\Omega = \{1, 2, 3, 4, 5, 6\}$ (expérience du dé)

$P(i) = \frac{1}{6}$ pour $1 \leq i \leq 6$ (probabilité uniforme)

si $A = \text{"face paire"}$ alors $A = \{2, 4, 6\}$

$$\text{on a } P(A) = P\{2, 4, 6\} = P(2) + P(4) + P(6) = \frac{3}{6} = \frac{1}{2}$$

(3) Une urne contient 12 boules numérotées de 1 à 12. On tire une boule au hasard. Dans cette expérience $\Omega = \{1, 2, 3, \dots, 12\}$ et $A = \{3, 6, 9, 12\}$ est l'événement "No. multiple de 3". On définit P en posant:

$P(\omega) = x$ pour tout numéro pair ω

$P(\omega) = y$ pour tout numéro impair ω

$$\text{on a alors } P(\Omega) = 6x + 6y = 1 \text{ donc } x + y = \frac{1}{6}$$

$$\text{dans ce cas } P(A) = P(3) + P(6) + P(9) + P(12) = 2y + 2x = \frac{2}{6} = \frac{1}{3}$$

(4) Considérons l'expérience: lancer un dé 2 fois.

On a $\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$ et supposons $P =$ probabilité uni-

forme $P(i, j) = \frac{1}{36}, \forall (i, j) \in \Omega$; calculons la probabilité de $A = \{(i, j) \in \Omega : i = j\} = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$

$$\text{on a } P(A) = P(1, 1) + P(2, 2) + P(3, 3) + P(4, 4) + P(5, 5) + P(6, 6) = \frac{6}{36} = \frac{1}{6};$$

alors $A^c = \{(i, j) \in \Omega : i \neq j\}$ et $P(A^c) = 1 - P(A) = 1 - \frac{1}{6} = \frac{5}{6}$

3. Probabilité conditionnelle-Indépendance

3.1 Définition. Soit Ω, P un espace probabilisé. (Définition 2.1)

La probabilité $P(A)$ d'un événement A peut être réévalué en tenant compte de la réalisation d'un autre événement B . Cette réévaluation utilise la notion de probabilité conditionnelle définie par $P(A|B) = \frac{P(A \cap B)}{P(B)}$, $P(B) \neq 0$

lire "probabilité de A si B "

l'exemple qui suit justifie cette définition:

soit $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ avec P = probabilité uniforme

si $A \subset \Omega$ alors $P(A) = \frac{|A|}{n}$. Soient A, B 2 événements; parmi les résultats qui réalisent B , $A \cap B$ sont ceux qui réalisent A donc si B se réalise, la probabilité

de réaliser A vaut $\frac{|A \cap B|}{|B|}$ que l'on peut écrire $\frac{\frac{|A \cap B|}{n}}{\frac{|B|}{n}} = \frac{P(A \cap B)}{P(B)}$

Si la réalisation de A ne dépend pas de celle de B on aura $P(A|B) = P(A)$, cela conduit à la définition suivante:

3.2 Définition. Les événements A et B sont indépendants si

$$P(A|B) = P(A), \text{ autrement dit } P(A \cap B) = P(A) \cdot P(B)$$

3.3. Proposition. Si les événements A et B sont indépendants alors:

$[A, B^c], [A^c, B], [A^c, B^c]$ sont des couples d'événements indépendants

Preuve On a $P(A \cap B) = P(A) \cdot P(B)$

Montrons que A, B^c sont indépendants c'est à dire que $P(A \cap B^c) = P(A) \cdot P(B^c)$

on a $A = (A \cap B) \cup (A \cap B^c)$ puisque $\Omega = B \cup B^c$ et $A = A \cap \Omega$

donc $P(A) = P(A \cap B) + P(A \cap B^c) = P(A) \cdot P(B) + P(A \cap B^c)$

$P(A \cap B^c) = P(A) - P(A) \cdot P(B) = P(A) \cdot (1 - P(B)) = P(A) \cdot P(B^c)$.

Le cas des événements $[A^c, B]$ et $[A^c, B^c]$ se démontre de la même façon.

Exemple.

Une urne contient 12 boules numérotées de 1 à 12. On tire une boule au hasard. Dans cette expérience $\Omega = \{1, 2, 3, \dots, 12\}$ et $A = \{3, 6, 9, 12\}$ est l'événement "No. multiple de 3". On définit P en posant:

$P(\omega) = x$ pour tout numéro pair ω

$P(\omega) = y$ pour tout numéro impair ω

on a alors $P(\Omega) = 6x + 6y = 1$ donc $x + y = \frac{1}{6}$

dans ce cas $P(A) = P(3) + P(6) + P(9) + P(12) = 2y + 2x = \frac{2}{6} = \frac{1}{3}$

il est facile de montrer que les événements suivants $A = \{2, 4, 6, 8, 10, 12\}$ et $B = \{3, 6, 9, 12\}$ sont indépendants $\forall 0 < x, y < 1$,

3.4. Proposition. Pour tous événements A, B on a:

(a) $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$

(b) $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)}$ (Formule de Bayes)

Preuve Il est clair que $P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$
 et $P(A \cap B^c) = P(A|B^c) P(B^c)$, $P(B \cap A^c) = P(B|A^c) \cdot P(A^c)$
 (a) $P(A) = P(A \cap B) + P(A \cap B^c) = P(A|B) \cdot P(B) + P(A|B^c) P(B^c)$

$$(b) P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B \cap A) + P(B \cap A^c)}$$

ensuite on remplace $P(B \cap A)$ par $P(B|A) \cdot P(A)$ et $P(B \cap A^c)$ par $P(B|A^c) \cdot P(A^c)$.

3.5 Définition. Soit $\{A_i, i \in I\}$ une famille arbitraire d'événements on dit que ces événements sont indépendants dans leur ensemble si:

pour toute partie finie $J \subset I$ on a $P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$

Chapitre 3

VARIABLES ALEATOIRES (V.A)

1. Définitions-Loi d'une variable aléatoire

1.1 Définition. Soit Ω, P un espace probabilisé
une variable aléatoire sur Ω, P est une fonction

$$X : \Omega \longrightarrow \mathbb{R}, \quad \omega \in \Omega \longrightarrow X(\omega) \in \mathbb{R}$$

$X(\Omega)$ est l'ensemble des valeurs de X

C'est une représentation numérique des résultats d'une expérience.

On utilise les notations suivantes:

$$\{X = x\} = \{\omega \in \Omega : X(\omega) = x\}$$

$$\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}$$

$$\{a < X < b\} = \{\omega \in \Omega : a < X(\omega) < b\}$$

1.2 Définition.

Soit X une variable aléatoire sur Ω, P

on définit la loi de probabilité de X par la fonction

$$f : \mathbb{R} \longrightarrow [0, 1], x \longrightarrow f(x) = P\{X = x\} = P\{\omega \in \Omega : X(\omega) = x\}$$

1.3 Définition. (Fonction de Répartition)

Soit X une V.A définie sur l'espace Ω, P . La fonction de répartition de X est définie par:

$F : \mathbf{R} \longrightarrow [0, 1], F(x) = P(X \leq x) = P(\omega \in \Omega : X(\omega) \leq x)$. Il n'est pas difficile de montrer que:

(a) F est croissante

(b) F est continue à droite i.e $\lim_{\substack{x \rightarrow a \\ x > a}} F(x) = F(a) \forall a \in \mathbf{R}$

(c) F admet une limite à gauche i.e $\lim_{\substack{x \rightarrow a \\ x < a}} F(x) = l$

(d) $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$

1.4 Exemples.

(a) On lance 3 fois une pièce de monnaie: p pile, f face

$\Omega = \{ppp, fpp, pfp, ppf, pff, fpf, ffp, fff\}$, P = probabilité uniforme

$X(\omega)$ = nombre de piles dans ω . $X(\Omega) = \{0, 1, 2, 3\}$

loi de X :

$$f(0) = P(X = 0) = P\{fff\} = \frac{1}{8}$$

$$f(1) = P(X = 1) = P\{pff, fpf, ffp\} = \frac{3}{8}$$

$$f(2) = P(X = 2) = P\{fpp, pfp, ppf\} = \frac{3}{8}$$

$$f(3) = P(X = 3) = P\{ppp\} = \frac{1}{8}$$

(b) Une urne contient 15 boules:

7 boules de 1kg, 5 boules de 3kg, 3 boules de 5kg

on tire un boule ω au hasard et on note son poids $X(\omega)$

on a Ω = l'ensemble des 15 boules, P = probabilité uniforme
 $X : \Omega \longrightarrow \{1, 3, 5\}$ $\omega \longrightarrow X(\omega)$ = poids de ω
 $P(X = 1) = \frac{7}{15}$, $P(X = 3) = \frac{5}{15}$, $P(X = 5) = \frac{3}{15}$

2. Variables aléatoires discrètes

2.1 Définition.

Soit X une V.A définie sur l'espace Ω , P on dit que X est discrète si l'ensemble de ses valeurs $X(\Omega)$ est une suite finie ou infinie de nombres réels; on écrit: $X(\Omega) = \{x_1, x_2, x_3, \dots\}$ et la loi de X est donnée par $f(x_n) = P(X = x_n)$ $n = 1, 2, 3, \dots$

2.1 Exemples. Soit Ω, P un espace probabilisé

(a) Variable de Bernoulli: $\mathcal{B}(p)$, $0 < p < 1$

$X(\Omega) = \{0, 1\}$, $P(X = 1) = p$, $P(X = 0) = 1 - p$

(b) Variable Binomiale: $\mathcal{B}(n, p)$, $n \geq 0$ $0 < p < 1$

$X(\Omega) = \{0, 1, 2, 3, \dots, n\}$, $P(X = k) = C_n^k \cdot p^k \cdot (1 - p)^{n-k}$, $k = 0, 1, 2, \dots, n$.

pour vérifier que $\sum_{k=0}^n P(X = k) = 1$, on utilise le Binôme de Newton

$$\sum_{k=0}^n C_n^k a^k b^{n-k} = (a + b)^n \text{ avec } a = p, b = 1 - p$$

(c) Loi géométrique $\mathcal{G}(p)$, $0 < p < 1$

$X(\Omega) = \{1, 2, 3, \dots\}$, $P(X = k) = p \cdot (1 - p)^{k-1}$, $k = 1, 2, 3, \dots$

$$\sum_{k=1}^{\infty} p \cdot (1 - p)^{k-1} = 1$$

(d) Loi de Poisson $\mathcal{P}(\lambda)$, $\lambda > 0$

$X(\Omega) = \{0, 1, 2, 3, \dots\}$, $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0, 1, 2, 3, \dots$

$$\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = 1 \text{ (utiliser la relation } \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda} \text{)}$$

3. Esperance-Variance de V.A discrètes

On considère une V.A discrete X sur un espace Ω, P avec:

$$X(\Omega) = \{x_1, x_2, x_3, \dots\} \text{ et la loi de } X \text{ donnée par } f(x_n) = P(X = x_n) \\ n = 1, 2, 3, \dots$$

3.1 Définition.

L'esperance de X est définie par:

$$E(X) = \sum_{n \geq 1} x_n \cdot P(X = x_n) = \sum_{n \geq 1} x_n \cdot f(x_n)$$

on dit que X est d'esperance finie si $-\infty < E(X) < +\infty$.

3.2 Exemples. Soit Ω, P un espace probabilisé

(a) Variable de Bernoulli: $\mathcal{B}(p), 0 < p < 1$

$$X(\Omega) = \{0, 1\}, P(X = 1) = p, P(X = 0) = 1 - p$$

$$E(X) = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = p$$

(b) Variable Binomiale: $\mathcal{B}(n, p), n \geq 0, 0 < p < 1$

$$X(\Omega) = \{0, 1, 2, 3, \dots, n\}, P(X = k) = C_n^k \cdot p^k \cdot (1 - p)^{n-k}, k = 0, 1, 2, \dots, n.$$

$$E(X) = \sum_{k=0}^n k \cdot C_n^k \cdot p^k \cdot (1 - p)^{n-k} = n \cdot p \text{ (voir calculs en Annexe)}$$

(c) Loi géométrique $\mathcal{G}(p), 0 < p < 1$

$$X(\Omega) = \{1, 2, 3, \dots\}, P(X = k) = p \cdot (1 - p)^{k-1}, k = 1, 2, 3, \dots$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot p \cdot (1 - p)^{k-1} = \frac{1}{p} \text{ (voir calculs en Annexe)}$$

(d) Loi de Poisson $\mathcal{P}(\lambda), \lambda > 0$

$$X(\Omega) = \{0, 1, 2, 3, \dots\}, P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, 3, \dots$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \text{ (voir calculs en Annexe)}$$

3.3 Propriétés de l'Esperance.

$$(1) \forall a, b \in R, E(aX + bY) = a \cdot E(X) + b \cdot E(Y)$$

$$(2) |E(X)| \leq E(|X|)$$

$$(3) X \leq Y \implies E(X) \leq E(Y)$$

3.4 Définition. (Variance)

On considère une V.A discrete X sur un espace Ω, P avec:

$$X(\Omega) = \{x_1, x_2, x_3, \dots\} \text{ et la loi de } X \text{ donnée par } f(x_n) = P(X = x_n) \\ n = 1, 2, 3, \dots$$

La variance de X est définie par:

$$V(X) = E \left[(X - E(X))^2 \right]$$

$$\sigma_X = \sqrt{V(X)} \text{ s'appelle l'écart type}$$

3.3 Propriétés de la Variance.

On pose $E(X) = \mu$

$$(a) V(X) = E(X^2) - \mu^2$$

$$(b) \forall a, b \in R, V(aX + b) = a^2 \cdot V(X)$$

3.4 Variance de Lois Discrètes usuelles.

(a) Variable de Bernoulli: $\mathcal{B}(p)$, $0 < p < 1$

$$X(\Omega) = \{0, 1\}, P(X = 1) = p, P(X = 0) = 1 - p$$

$$E(X) = 1.P(X = 1) + 0.P(X = 0) = p$$

$$V(X) = p(1 - p)$$

(b) Variable Binomiale: $\mathcal{B}(n, p)$, $n \geq 0$ $0 < p < 1$

$$X(\Omega) = \{0, 1, 2, 3, \dots, n\}, P(X = k) = C_n^k \cdot p^k \cdot (1 - p)^{n-k}, k = 0, 1, 2, \dots, n.$$

$$E(X) = \sum_{k=0}^n k \cdot C_n^k \cdot p^k \cdot (1 - p)^{n-k} = n \cdot p \quad (\text{voir calculs en Annexe})$$

$$V(X) = np(1 - p)$$

(c) Loi géométrique $\mathcal{G}(p)$, $0 < p < 1$

$$X(\Omega) = \{1, 2, 3, \dots\}, P(X = k) = p \cdot (1 - p)^{k-1}, k = 1, 2, 3, \dots$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot p \cdot (1 - p)^{k-1} = \frac{1}{p} \quad (\text{voir calculs en Annexe})$$

$$V(X) = \frac{1 - p}{p^2}$$

(d) Loi de Poisson $\mathcal{P}(\lambda)$, $\lambda > 0$

$$X(\Omega) = \{0, 1, 2, 3, \dots\}, P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, 3, \dots$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \quad (\text{voir calculs en Annexe})$$

$$V(X) = \lambda$$

4. Variables aléatoires absolument continues

4.1 Définition.

Soit X une V.A définie sur l'espace Ω, P on dit que X est absolument continue si sa fonction de répartition F (Définition 1.3) est de la forme

$$F(x) = \int_{-\infty}^x f(x) \cdot dx = P(\omega \in \Omega : X(\omega) \leq x) \text{ avec } f \text{ fonction réelle positive}$$

Riemann intégrable et $\int_{-\infty}^{\infty} f(x) \cdot dx = 1$. Pour simplifier nous supposons la

fonction f continue, de sorte que F soit dérivable et $F'(x) = f(x)$, dans ce cas $F(b) - F(a) = \int_a^b f(x) \cdot dx = P(a \leq X \leq b) = P(\omega \in \Omega : X(\omega) \in [a, b])$ et

$$P(X = x) = \int_x^x f(x) \cdot dx = 0, \forall x.$$

La fonction f s'appelle densité de la variable aléatoire X .

4.2 Exemples.

(a) **La loi uniforme $\mathcal{U} [a, b]$ de densité donnée par:**

$$f(x) = \frac{1}{b-a} \text{ si } x \in [a, b]$$

$$f(x) = 0 \text{ si } x \notin [a, b]$$

Fonction de répartition donnée par:

$$F(x) = 0 \text{ si } x < a$$

$$F(x) = \frac{x-a}{b-a} \text{ si } x \in [a, b]$$

$$F(x) = 1 \text{ si } x > b$$

(b) **La loi exponentielle $\mathcal{E}(\lambda)$, $\lambda > 0$, de densité donnée par:**

$$f(x) = \lambda e^{-\lambda x}, \text{ si } x \geq 0$$

$$f(x) = 0, \text{ si } x < 0$$

Fonction de répartition:

$$F(x) = 0, \text{ si } x < 0$$

$$F(x) = 1 - e^{-\lambda x}, \text{ si } x \geq 0$$

(c) **Loi Normale ou Loi de Gauss $\mathcal{N}(m, \sigma)$, de densité donnée par:**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \text{ avec } x \in \mathbf{R}$$

Cas particulier: **Loi Normale centrée réduite avec $\sigma = 1, m = 0$:**

$$\text{densité donnée par: } f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ avec } x \in \mathbf{R}$$

4.3 Esperance et Variance de V.A. absolument continues

Soit X une V.A absolument continue de densité $f(x)$

(a) On définit l'esperance de X par:

$$E(X) = \int_{-\infty}^{+\infty} x.f(x).dx = \mu, \text{ l'existence de } E(X) \text{ equivaut à l'integrabilité}$$

de la fonction $x.f(x)$

(b) On définit la variance de X par:

$$V(X) = E[(X - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 .f(x) .dx, \text{ l'existence de } V(X) \text{ equiv-}$$

aut à l'integrabilité de la fonction $(x - \mu)^2 .f(x)$

(Rappel: $V(X) = E(X^2) - \mu^2$, écart type $\sigma = \sqrt{V(X)}$)

4.4 Exemples. (Rappel: $V(X) = E(X^2) - \mu^2$, avec $E(X) = \mu$)

(a) **La loi uniforme** $\mathcal{U}[a, b]$: $E(X) = \int_a^b \frac{x}{b-a}.dx = \frac{a+b}{2}$

$$V(X) = \int_a^b \frac{x^2}{b-a}.dx - \left(\frac{a+b}{2}\right)^2$$

(b) **La loi exponentielle** $\mathcal{E}(\lambda)$, $\lambda > 0$: $E(X) = \int_0^{+\infty} x.\lambda e^{-\lambda x}.dx = \frac{1}{\lambda}$

$$V(X) = \int_0^{+\infty} x^2.\lambda e^{-\lambda x}.dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

(c) **Loi Normale** $\mathcal{N}(m, \sigma)$: $E(X) = \int_{-\infty}^{+\infty} x.\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}} = m$

$$V(X) = \int_{-\infty}^{+\infty} x^2.\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}} - m^2 = \sigma^2.$$

5. Inégalités

5.1. Définition. Soit A une partie de \mathbf{R}

(a) La fonction indicatrice de A est définie par:

$$I_A(x) = 1 \text{ si } x \in A \text{ et } I_A(x) = 0 \text{ si } x \in A^c.$$

(b) Si f est une fonction de \mathbf{R} dans \mathbf{R} on note la fonction définie par:

$$(f.I_A)(x) = f(x) \text{ si } x \in A$$

$$(f.I_A)(x) = 0 \text{ si } x \in A^c$$

5.2. Définition. Soit X une V.A sur Ω, P , et $A \subset \Omega$

$$(X.I_A)(\omega) = X \text{ si } \omega \in A$$

$$(X.I_A)(\omega) = 0 \text{ si } \omega \in A^c$$

5.3. Définition. (Intégrale locale)

(a) Soit f une fonction de \mathbf{R} dans \mathbf{R} et soit A une partie de \mathbf{R} , l'intégrale

locale de la fonction f sur A est définie par $\int_A f(x) .dx = \int_{\mathbf{R}} (f.I_A)(x) .dx$.

A titre d'exemple on peut prendre $A =]a, +\infty[$

on obtient $\int_A f(x) .dx = \int_a^{+\infty} f(x) .dx$.

(b) Soit X une V.A sur Ω, P , et $A \subset \Omega$ l'espérance locale de la V.A X sur A est définie par $E_A(X) = E(X.I_A)$

A titre d'exemple on peut prendre $A = \{\omega \in \Omega : X(\omega) \geq a\}$

L'application de l'espérance locale pour la V.A $|X|$ permet de prouver l'inégalité de Markov ci-dessous

Le but de la suite est de donner certaines inégalités utiles dans les applications. Les V.A considérées sont supposées avoir une espérance finie.

Inégalité de Markov

Soit X une variable aléatoire telle que $E(|X|) < \infty$ alors $\forall \lambda > 0$ on a:

$$P(|X| > \lambda) \leq \frac{E(|X|)}{\lambda}$$

Inégalité de Jensen

$$|E(X)| \leq E(|X|)$$

Plus généralement si $f : \mathbf{R} \rightarrow \mathbf{R}$ est une fonction convexe telle que

$$E(|f(X)|) < \infty, \text{ alors: } f(E(X)) \leq E(f(X))$$

Inégalité de Tchebychev

Soit $\lambda > 0$ et soit X une V.A de variance finie alors

$$P(|X - E(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$$

6. Convergence de suites de variables aléatoires

Approximations

6.1 Types de convergence

Soit X_n une suite de V.A. sur Ω, P , de fonction de répartition F_n

(a) **Convergence en probabilité:**

X_n converge en probabilité vers la V.A X si:

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0 \text{ (notation } X_n \xrightarrow{P} X)$$

(b) **Convergence presque sûre**

X_n converge presque sûrement vers la V.A X si: $\exists A \subset \Omega$

avec $P(A) = 0$ et $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega), \forall \omega \notin A$ (notation $X_n \xrightarrow{p.s} X$)

(c) **Convergence en Loi**

X_n converge en loi vers la V.A X de fonction de répartition F si:

$$\lim_{n \rightarrow \infty} F_n(t) = F(t) \text{ en tout point de continuité } t \text{ de } F \text{ (notation } X_n \xrightarrow{\mathcal{L}} X)$$

6.2. Approximation de la loi Binomiale par la loi de Poisson

Soit X_n une suite de V.A de loi Binomiale $\mathcal{B}(n, p_n), n \geq 0 \quad 0 < p_n < 1$

Supposons que $\lim_{n \rightarrow \infty} np_n = \lambda > 0$

X_n converge en loi vers la V.A X de Poisson de parametre λ

i.e $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, 3, \dots$

6.3. Theoreme

Soit X_n, X une suite de V.A. sur Ω, P , alors on a:

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

La reciproque de ces implications est fausse en general

6.3. Theoreme: Loi des grands nombres

(a) **Loi faible:**

Soit X_n , une suite de V.A. sur Ω, P

supposons les X_n independantes et de meme loi

(donc de meme esperance $= \mu$) alors:

(a) **Loi faible:**

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n} \text{ converge en probabilité vers } \mu.$$

i.e $P(|M_n - \mu| > \epsilon) \rightarrow 0, \forall \epsilon > 0$

(b) **Loi forte:**

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n} \text{ converge presque sûrement vers } \mu.$$

(Les V.A. X_n sont independantes si $\forall n \geq 1$ et pour tous

$\{A_1, A_2, \dots, A_n\} \subset \mathcal{P}(\Omega)$:

$$P((X_1 \in A_1) \cap (X_2 \in A_2) \cap \dots \cap P(X_n \in A_n)) = \prod_{j=1}^n P(X_j \in A_j))$$

7. Theorem de la limite centrale

Rappel: **Loi Normale**

Loi Normale ou Loi de Gauss $\mathcal{N}(m, \sigma)$, de densité donnée par:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \text{ avec } x \in \mathbf{R}$$

Cas particulier: **Loi Normale centrée réduite** avec $\sigma = 1, m = 0$:

$$\text{densité donnée par: } f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ avec } x \in \mathbf{R}$$

Theoreme:

Soit X_n , une suite de V.A. sur Ω, P

supposons les X_n independantes et de meme loi

d'esperance μ et d'ecart type σ finis, $\sigma \neq 0$, alors on a:

$$\text{avec } M_n = \frac{X_1 + X_2 + \dots + X_n}{n},$$

$\frac{M_n - \mu}{\sigma/\sqrt{n}}$ converge en loi vers **Loi Normale centrée réduite**

avec $\sigma = 1, m = 0$ de densité donnée par: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ avec $x \in \mathbf{R}$

ce qui signifie: $\lim_{n \rightarrow \infty} .P \left(\frac{M_n - \mu}{\sigma/\sqrt{n}} \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} .dx$

Annexe 1

Esperance et Variance de V A absolument continues

Rappel

Soit X une V.A définie sur l'espace Ω, P on dit que X est absolument continue si sa fonction de répartition $F(x) = P(\omega \in \Omega : X(\omega) \leq x)$ est de la forme

$$F(x) = \int_{-\infty}^x f(x) .dx \text{ avec } f \text{ fonction réelle positive Riemann intégrable et } \int_{-\infty}^{\infty} f(x) .dx =$$

1. Pour simplifier nous supposerons la fonction f continue, de sorte que F soit dérivable et $F'(x) = f(x)$, dans ce cas $F(b) - F(a) = \int_a^b f(x) .dx =$

$$P(a \leq X \leq b) = P(\omega \in \Omega : X(\omega) \in [a, b]) \text{ et } P(X = x) = \int_x^x f(x) .dx = 0, \forall x.$$

La fonction f s'appelle densité de la variable aléatoire X .

Exemples.

(a) **La loi uniforme $\mathcal{U} [a, b]$ de densité donnée par:**

$$f(x) = \frac{1}{b-a} \text{ si } x \in [a, b]$$

$$f(x) = 0 \text{ si } x \notin [a, b]$$

Fonction de répartition donnée par:

$$F(x) = 0 \text{ si } x < a$$

$$F(x) = \frac{x-a}{b-a} \text{ si } x \in [a, b]$$

$$F(x) = 1 \text{ si } x > b$$

(b) **La loi exponentielle $\mathcal{E}(\lambda)$, $\lambda > 0$, de densité donnée par:**

$$f(x) = \lambda e^{-\lambda x}, \text{ si } x \geq 0$$

$$f(x) = 0, \text{ si } x < 0$$

Fonction de répartition:

$$F(x) = 0, \text{ si } x < 0$$

$$F(x) = 1 - e^{-\lambda x}, \text{ si } x \geq 0$$

(c) **Loi Normale ou Loi de Gauss $\mathcal{N}(m, \sigma)$, de densité donnée par:**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \text{ avec } x \in \mathbf{R}$$

Cas particulier: **Loi Normale centrée réduite avec $\sigma = 1, m = 0$:**

$$\text{densité donnée par: } f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ avec } x \in \mathbf{R}$$

Esperance et Variance de V.A. absolument continus

Soit X une V.A. absolument continue de densité $f(x)$

(a) On définit l'esperance de X par:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx = \mu$$

(b) On définit la variance de X par:

$$V(X) = E[(X - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot f(x) \cdot dx$$

(Rappel: $V(X) = E(X^2) - \mu^2$), écart type $\sigma = \sqrt{V(X)}$

Calculs de $E(X)$, $V(X)$ pour certaines variables absolument continus

1. Variable uniforme $\mathcal{U}[a, b]$ de densité donnée par:

$$f(x) = \frac{1}{b-a} \text{ si } x \in [a, b]$$

$$f(x) = 0 \text{ si } x \notin [a, b]$$

$$(a) E(X) = \int_a^b x \cdot \frac{1}{b-a} \cdot dx = \left[\frac{x^2}{2} \right]_a^b \cdot \frac{1}{b-a} = \left(\frac{b^2 - a^2}{2} \right) \cdot \frac{1}{b-a} = \frac{a+b}{2}$$

$$E(X) = \frac{a+b}{2}$$

pour calculer la variance on utilise la relation:

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

$$E(X^2) = \int_a^b x^2 \cdot \frac{1}{b-a} \cdot dx = \left[\frac{x^3}{3} \right]_a^b \cdot \frac{1}{b-a} = \left(\frac{b^3 - a^3}{3} \right) \cdot \frac{1}{b-a}$$

mais on a $b^3 - a^3 = (b-a) \cdot (b^2 + ab + a^2)$, on obtient:

$$E(X^2) = \frac{(b^2 + ab + a^2)}{3}$$

$$V(X) = \frac{(b^2 + ab + a^2)}{3} - \left(\frac{a+b}{2} \right)^2 = \frac{(b^2 + ab + a^2)}{3} - \frac{a^2 + b^2 + 2ab}{4}$$

$$V(X) = \frac{4 \cdot (b^2 + ab + a^2) - 3 \cdot (a^2 + b^2 + 2ab)}{12} = \frac{(b-a)^2}{12}$$

(b) loi exponentielle $\mathcal{E}(\lambda)$, $\lambda > 0$, de densité donnée par:

$$f(x) = \lambda e^{-\lambda x}, \text{ si } x \geq 0$$

$$f(x) = 0, \text{ si } x < 0$$

$$E(X) = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} \cdot dx = \lambda \cdot \int_0^{\infty} x \cdot e^{-\lambda x} \cdot dx$$

on exprime $x.e^{-\lambda x}$ avec la dérivée de $e^{-\lambda x}$ par rapport à λ :

$x.e^{-\lambda x} = -\frac{de^{-\lambda x}}{d\lambda}$ on obtient

$$E(X) = \lambda \int_0^{\infty} x.e^{-\lambda x}.dx = \lambda \int_0^{\infty} -\frac{de^{-\lambda x}}{d\lambda}.dx$$

d'après une propriété de l'intégrale en théorie de la mesure on peut écrire:

$$\int_0^{\infty} -\frac{de^{-\lambda x}}{d\lambda}.dx = -\frac{d}{d\lambda} \int_0^{\infty} e^{-\lambda x}.dx = -\frac{d}{d\lambda} -\frac{1}{\lambda} \cdot [e^{-\lambda \cdot x}]_0^{\infty} = -\frac{d}{d\lambda} \left(\frac{1}{\lambda} \right) = \frac{1}{\lambda^2}$$

cela donne $E(X) = \lambda \cdot \frac{1}{\lambda^2} = \frac{1}{\lambda}$.

Pour calculer la variance on évalue $E(X^2) = \lambda \int_0^{\infty} x^2.e^{-\lambda x}.dx$

en intégrant par parties; on trouve $V(X) = E(X^2) - (E(X))^2 = \frac{1}{\lambda^2}$.

Annexe 2 Inégalités

5.1. Définition. Soit A une partie de \mathbf{R}

(a) La fonction indicatrice de A est définie par:

$$I_A(x) = 1 \text{ si } x \in A \text{ et } I_A(x) = 0 \text{ si } x \in A^c.$$

son esperance vaut $E(I_A) = P(A)$

(b) Si f est une fonction de \mathbf{R} dans \mathbf{R} on note la fonction définie par:

$$(f.I_A)(x) = f(x) \text{ si } x \in A$$

$$(f.I_A)(x) = 0 \text{ si } x \in A^c$$

5.2. Définition. Soit X une V.A sur Ω, P , et $A \subset \Omega$

$$(X.I_A)(\omega) = X \text{ si } \omega \in A$$

$$(X.I_A)(\omega) = 0 \text{ si } \omega \in A^c$$

5.3. Définition. (Intégrale locale)

(a) Soit f une fonction de \mathbf{R} dans \mathbf{R} et soit A une partie de \mathbf{R} , l'intégrale

locale de la fonction f sur A est définie par $\int_A f(x) .dx = \int_{\mathbf{R}} (f.I_A)(x) .dx$.

A titre d'exemple on peut prendre $A =]a, +\infty[$

on obtient $\int_A f(x) .dx = \int_a^{+\infty} f(x) .dx$.

(b) Soit X une V.A sur Ω, P , et $A \subset \Omega$ l'esperance locale de la V.A X sur A est définie par $E_A(X) = E(X.I_A)$

A titre d'exemple on peut prendre $A = \{\omega \in \Omega : X(\omega) \geq a\}$

L'application de l'esperance locale pour la V.A $|X|$ permet de prouver l'inegalite de Markov ci-dessous

Le but de la suite est de donner certaines inégalités utiles dans les applications.

Les V.A considérées sont supposées avoir une esperance finie.

Inégalité de Markov

Soit X une variable aléatoire telle que $E(|X|) < \infty$ alors $\forall \lambda > 0$ on a:

$$P(|X| > \lambda) \leq \frac{E(|X|)}{\lambda}$$

solution.

$$\text{on a } |X| \geq |X| .I_{(|X|>\lambda)} > \lambda .I_{(|X|>\lambda)}$$

$$\implies E(|X|) \geq E(\lambda .I_{(|X|>\lambda)}) = \lambda .P(|X| > \lambda)$$

par la croissance de l'esperance

$$\text{comme on a } E(I_{(|X|>\lambda)}) = P(|X| > \lambda) \text{ (Definition 5.1 (a))}$$

$$\text{on deduit } P(|X| > \lambda) \leq \frac{E(|X|)}{\lambda}$$

Inégalité de Jensen

$$|E(X)| \leq E(|X|)$$

Plus généralement si $f : \mathbf{R} \rightarrow \mathbf{R}$ est une fonction convexe telle que

$$E(|f(X)|) < \infty, \text{ alors: } f(E(X)) \leq E(f(X))$$

solution.

Pour toute V.A X on définit les V.A:

$$X^+ = \sup(X, 0) \text{ et } X^- = \inf(-X, 0)$$

il n'est pas difficile de montrer que:

$$\begin{aligned} X &= X^+ - X^-, & |X| &= X^+ + X^-, \text{ d'où par linéarité de l'esperance:} \\ E(X) &= E(X^+) - E(X^-), & E(|X|) &= E(X^+) + E(X^-), \text{ on obtient:} \\ |E(X)| &= |E(X^+) - E(X^-)| \leq |E(X^+)| + |E(X^-)| = E(|X|) \end{aligned}$$

Pour le cas general avec une fonction convexe f

voir cours de théorie de la mesure.

Inégalité de Tchebychev

Soit $\lambda > 0$ et soit X une V.A de variance finie alors

$$P(|X - E(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$$

solution.

On utilise la même technique que celle appliquée dans l'inégalité de Markov:

$$\begin{aligned} |X - E(X)|^2 &\geq |X - E(X)|^2 \cdot I_{(|X - E(X)| \geq \lambda)} \geq \lambda^2 \cdot I_{(|X - E(X)| \geq \lambda)} \\ \implies E|X - E(X)|^2 &\geq \lambda^2 \cdot E(I_{(|X - E(X)| \geq \lambda)}) \end{aligned}$$

mais $E|X - E(X)|^2 = V(X)$, et $E(I_{(|X - E(X)| \geq \lambda)}) = P(|X - E(X)| \geq \lambda)$
 finalement on obtient l'inégalité recherchée:

$$P(|X - E(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}.$$

Annexe 3

Convergence de suites de variables

A. Variables aléatoires indépendantes

Soit (X_n) une suite de variables aléatoires

Les v.a X_n sont dites indépendantes si $\forall k \geq 1$ et pour tous

$\{A_1, A_2, \dots, A_k\} \subset \mathcal{P}(\Omega)$:

$$P((X_1 \in A_1) \cap (X_2 \in A_2) \cap \dots \cap (X_k \in A_k)) = \prod_{j=1}^k P(X_j \in A_j).$$

Propriété fondamentale:

Si les v.a X_n sont indépendantes et d'espérances finies alors:

$$(*) \quad \forall n \geq 2 \text{ on a } E(X_1 \cdot X_2 \cdot \dots \cdot X_n) = E(X_1) \cdot E(X_2) \cdot \dots \cdot E(X_n)$$

On donne ici une preuve dans le cas de 2 v.a discrètes X, Y :

supposons donc $X(\Omega) = \{x_1, x_2, \dots\}$, $Y(\Omega) = \{y_1, y_2, \dots\}$

$$\text{ona } E(X \cdot Y) = \sum_n \sum_k x_n \cdot y_k \cdot P(X = x_n, Y = y_k) = \sum_n \sum_k x_n \cdot y_k \cdot P(X = x_n) \cdot P(Y = y_k)$$

où on a utilisé l'indépendance de X et Y dans la relation

$$P(X = x_n, Y = y_k) = P(X = x_n) \cdot P(Y = y_k)$$

remarquons enfin que $\sum_n \sum_k x_n \cdot y_k \cdot P(X = x_n) \cdot P(Y = y_k)$

$$= \sum_n x_n \cdot P(X = x_n) \cdot \sum_k y_k \cdot P(Y = y_k) = E(X) \cdot E(Y)$$

Remarques:

1. Il est possible de démontrer la propriété (*) par récurrence sur n en utilisant des vecteurs aléatoires qui ne figurent pas au programme.
2. La propriété (*) est aussi vraie pour les v.a absolument continues mais nécessite des techniques d'intégration multiple.

B. Types de convergence

Soit X_n une suite de V.A. sur Ω, P , de fonction de répartition F_n

(a) Convergence en probabilité:

X_n converge en probabilité vers la V.A X si:

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0 \text{ (notation } X_n \xrightarrow{P} X)$$

(b) Convergence presque sûre

X_n converge presque sûrement vers la V.A X si: $\exists A \subset \Omega$

avec $P(A) = 0$ et $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega), \forall \omega \notin A$ (notation $X_n \xrightarrow{p.s.} X$)

(c) Convergence en Loi

X_n converge en loi vers la V.A X de fonction de répartition F si:

$$\lim_{n \rightarrow \infty} F_n(t) = F(t) \text{ en tout point de continuité } t \text{ de } F \text{ (notation } X_n \xrightarrow{\mathcal{L}} X)$$

Approximation de la loi Binomiale par la loi de Poisson

Soit X_n une suite de V.A de loi Binomiale $\mathcal{B}(n, p_n)$, $n \geq 0$ $0 < p_n < 1$:

$$P(X_n = k) = C_n^k \cdot (p_n)^k \cdot (1 - p_n)^{n-k}$$

Supposons que $\lim_{n \rightarrow \infty} np_n = \lambda > 0$

X_n converge en loi vers la V.A X de Poisson de parametre λ

$$\text{i.e } P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, 3, \dots$$

Démonstration:

Choisissons l'entier n suffisamment grand pour remplacer p_n par $\frac{\lambda}{n}$

dans ce cas $P(X_n = k) = C_n^k \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}$, pour n grand

et cela implique:

$$P(X_n = k) = \frac{n!}{k!(n-k)!} \cdot \frac{\lambda^k}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} \cdot \frac{n!}{(n-k)!n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k}$$

il est facile de voir que $\lim_{n \rightarrow \infty} \frac{n!}{(n-k)!n^k} = 1$ et $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = 1$

il reste à calculer $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n$; on pose $x = -\frac{n}{\lambda}$

on obtient $\left(1 - \frac{\lambda}{n}\right)^n = \left(1 + \frac{1}{x}\right)^{-\lambda x}$ d'autre part il est bien connu que

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e \text{ donc } \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^{-\lambda x} = e^{-\lambda}$$

finalement $\lim_{n \rightarrow \infty} \frac{n!}{(n-k)!n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} = e^{-\lambda}$ et

$$\lim_{n \rightarrow \infty} P(X_n = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}.$$

Theoreme

Soit X_n, X une suite de V.A. sur Ω, P , alors on a:

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

La reciproque de ces implications est fausse en general

C. Loi des grands nombres

(a) Loi faible:

Soit X_n , une suite de V.A. sur Ω, P

supposons les X_n independantes et de meme loi

(donc de meme esperance $=\mu$) alors:

$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ converge en probabilité vers μ .

$$\text{i.e } P(|M_n - \mu| > \epsilon) \longrightarrow 0, \forall \epsilon > 0$$

(b) **Loi forte:**

$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ converge presque sûrement vers μ .

Démonstration:

(a) **Loi faible:**

On utilise l'inégalité de Tchebychev:

Soit $\lambda > 0$ et soit X une V.A de variance finie alors

$$P(|X - E(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$$

on l'applique à la v.a $X = M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$

puisqu'on les v.a sont indépendantes et de même loi on a:

$$E(X_i) = \mu, \forall i \text{ et } E(M_n) = \frac{n \cdot \mu}{n} = \mu \text{ et de plus } V(M_n) = \frac{\sum_{i=1}^n V(X_i)}{n^2}$$

$$\text{avec } V(X_i) = \sigma^2, \forall i \implies V(M_n) = \frac{\sigma^2}{n}$$

l'inégalité de Tchebychev donne:

$$P(|M_n - E(M_n)| \geq \epsilon) \leq \frac{V(M_n)}{\epsilon^2} = \frac{\frac{\sigma^2}{n}}{\epsilon^2} = \frac{\sigma^2}{n \cdot \epsilon^2} \longrightarrow 0 \text{ si } n \longrightarrow \infty.$$

(b) **Loi forte:**

La démonstration de la loi forte est plus élaborée et nécessite des techniques avancées de théorie des probabilités qui ne sont pas au programme.

Annexe 4

9. Soit Y une variable binomiale de parametre $0 < p < 1$:

$$P(Y = k) = C_n^k \cdot p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n.$$

(a) Calculer son esperance $E(Y)$

(b) Donner un exemple d'experience gérée par Y .

Rappel: **Formule du Binome de Newton:**

$$\forall a, b \text{ dans } \mathbb{R}, (a+b)^n = \sum_{k=0}^n C_n^k \cdot a^k \cdot b^{n-k}$$

solution. (a) $E(Y) = \sum_{k=0}^n k \cdot P(Y = k) = \sum_{k=1}^n k \cdot C_n^k \cdot p^k (1-p)^{n-k}$ (*)

$$\text{on a } k \cdot C_n^k = k \cdot \frac{n!}{k!(n-k)!} = k \cdot \frac{n!}{k \cdot (k-1)!(n-k)!} =$$

$$\frac{n \cdot (n-1)!}{(k-1)!((n-1)-(k-1)!)} \text{ puisque } n-k = (n-1) - (k-1) \text{ on obtient:}$$

$$k \cdot C_n^k = n \cdot C_{n-1}^{k-1}, \text{ on remplace } k \cdot C_n^k \text{ par } n \cdot C_{n-1}^{k-1} \text{ dans la formule (*), donc:}$$

$$E(Y) = \sum_{k=1}^n k \cdot C_n^k \cdot p^k (1-p)^{n-k} = \sum_{k=1}^n n \cdot C_{n-1}^{k-1} \cdot p^k (1-p)^{n-k} =$$

$$n \cdot p \cdot \sum_{k=1}^n C_{n-1}^{k-1} \cdot p^{k-1} (1-p)^{(n-1)-(k-1)}, \text{ mais d'après}$$

la **Formule du Binome de Newton**

$$\text{on a } \sum_{k=1}^n C_{n-1}^{k-1} \cdot p^{k-1} (1-p)^{(n-1)-(k-1)} = (p + (1-p))^{n-1} = 1$$

par consequent on obtient $E(Y) = n \cdot p$.

(b) Exemple d'experience gérée par Y :

Le pourcentage de pièces défectueuses d'un procédé de fabrication industrielle vaut $0 < p < 1$. Cela signifie que la probabilité qu'une piece fabriquée soit défectueuse vaut p . Si on considère au hazard un ensemble de n pièces fabriquées par ce procédé le nombre Y de pièces défectueuses contenues dans cet ensemble suit une loi binomiale de paramètres n et p .

10. Probabilité Produit

Soient Ω_1, P_1 et Ω_2, P_2 deux espaces probabilisés. On définit

sur l'ensemble $\Omega = \Omega_1 \times \Omega_2$ la probabilité produit P de P_1 et P_2 par:

si $A \subset \Omega$ et $A = A_1 \times A_2$, avec $A_1 \subset \Omega_1, A_2 \subset \Omega_2, P(A) = P_1(A_1) \cdot P_2(A_2)$.

On note $P = P_1 \otimes P_2$

Pour une piece donnée on suppose: $P(\text{pile}) = \alpha$.

On lance 2 fois cette piece. Si l'experience est gérée par

la probabilité produit $\pi = P \otimes P$, déterminer la probabilité des evenements $(p, p), (p, f), (f, p), (f, f)$.

solution.

$$\pi(p, p) = \pi(p) \cdot \pi(p) = \alpha^2, \quad \pi(p, f) = \pi(p) \cdot \pi(f) = \alpha \cdot (1-\alpha)$$

$\pi(f, p) = \pi(f) \cdot \pi(p) = (1 - \alpha) \cdot \alpha$, $\pi(f, f) = \pi(f) \cdot \pi(f) = (1 - \alpha)^2$
 on vérifie bien que:
 $\pi(p, p) + \pi(p, f) + \pi(f, p) + \pi(f, f) = \alpha^2 + 2\alpha \cdot (1 - \alpha) + (1 - \alpha)^2 = 1$

11. (a) On lance un dé 2 fois au hasard. Si chaque jet est géré par la probabilité uniforme π , déterminer la probabilité produit $P(\omega_1, \omega_2) = \pi \otimes \pi(\omega_1, \omega_2)$ de chaque résultat (ω_1, ω_2) .
 (b) Soit S la somme des résultats obtenus: calculer $P(S = 7)$, $P(S = 8)$, $P(S = 9)$

solution.

(a) On a $\Omega = \{(\omega_1, \omega_2) : 1 \leq \omega_1 \leq 6, 1 \leq \omega_2 \leq 6\}$
 $\pi(\omega_1) = \pi(\omega_2) = \frac{1}{6}$, $\forall \omega_1, \omega_2$ compris entre 1 et 6 puisque chaque jet est géré par la probabilité uniforme π .
 Donc $P(\omega_1, \omega_2) = \pi \otimes \pi(\omega_1, \omega_2) = \pi(\omega_1) \times \pi(\omega_2) = \frac{1}{36}$.
 Par conséquent P est la probabilité uniforme sur Ω .

(b) $S(\omega_1, \omega_2) = \omega_1 + \omega_2$
 $(S = 7) = \{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)\}$
 $P(S = 7) = P\{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)\} = \frac{6}{36} = \frac{1}{6}$
 on calcule $P(S = 8)$, $P(S = 9)$ de façon analogue.

12. Soient X, Y 2 V.A de Ω dans \mathbf{R} définies sur Ω, P et de même loi. Montrer que $E(X) = E(Y)$ et $V(X) = V(Y)$ (considérer le cas discret et le cas absolument continu et remarquer que $X(\Omega) = Y(\Omega)$).

solution. X, Y ont la même loi:

(1) Cas discret: $X(\Omega) = Y(\Omega) = \{v_1, v_2, \dots\}$ et $P(X = v_n) = P(Y = v_n)$
 $E(X) = \sum_{v_n} v_n \cdot P(X = v_n) = \sum_{v_n} v_n \cdot P(Y = v_n) = E(Y)$

Pour l'égalité des variance on utilise la relation établie en cours:
 $V(X) = E(X^2) - (E(X))^2$, $V(Y) = E(Y^2) - (E(Y))^2$
 avec $E(X^2) = \sum_{v_n} v_n^2 \cdot P(X = v_n)$ et $E(Y^2) = \sum_{v_n} v_n^2 \cdot P(Y = v_n)$
 on déduit que $V(X) = V(Y)$

(2) Cas absolument continu: $X(\Omega) = Y(\Omega) = I$ interval de \mathbb{R}
 X, Y ont la même loi, c'est à dire la même densité $\varphi(t), t \in I$

$$\begin{aligned}
 E(X) &= E(Y) = \int_I t \cdot \varphi(t) \cdot dt, & E(X^2) &= E(Y^2) = \int_I t^2 \cdot \varphi(t) \cdot dt \\
 V(X) &= E(X^2) - (E(X))^2, & V(Y) &= E(Y^2) - (E(Y))^2 \\
 &= \int_I t^2 \cdot \varphi(t) \cdot dt - \left(\int_I t \cdot \varphi(t) \cdot dt \right)^2.
 \end{aligned}$$