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 éléction

Définition 1.2:
Opérations sur les ensembles: soient $A, B, C, ...$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 Algebre 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 Algebre de Boole alors $\phi \in \mathcal{F}$
et pour toute suite finie $(A_k), 1 \leq k \leq n$ dans $\mathcal{F}$ on a $\overset{n}{\underset{k=1}{\cap}} A_k \in \mathcal{F}$

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

Définition 1.4: ($\sigma-$algèbre de Boole)
Soit $\mathcal{F} \subset \mathcal{P}(\Omega)$ une famille de parties de Ω. $\mathcal{F}$ est une $\sigma-$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 $\underset{n}{\cup} A_n \in \mathcal{F}$

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

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

Exercices.

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

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

Définition 1.6: (Cardinaux)
Si Ω est un ensemble quelconque on définit le cardinal de Ω par:
$$|\Omega| = \left\{ \begin{array}{c} n \text{ si } \Omega \text{ possede } n \text{ elements} \\ \infty \text{ si } \Omega \text{ est infini} \end{array} \right\}$$

Proposition 1.7:
Si $|\Omega| = n$ alors $|\mathcal{P}(\Omega)| = 2^n \quad (*)$

Preuve: Par recurrence 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 ecrivons 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 Cartesien)
Soient Ω_1, Ω_2 deux ensembles, on définit le produit cartesien $\Omega_1 \times \Omega_2$ 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
$\Omega_1 = \{a_1, a_2, ..., a_m\}, \Omega_2 = \{b_1, b_2, ..., b_n\}$; ainsi on aura:
$\Omega_1 \times \Omega_2 = \{(a_i, b_j), \quad 1 \leq i \leq m, 1 \leq j \leq n\}$ que l'on peut ecrire:

$$\begin{array}{c} (a_1, b_1), (a_1, b_2), ..., (a_1, b_n) \\ (a_2, b_1), (a_2, b_2), ..., (a_2, b_n) \\ . \\ (a_m, b_1), (a_m, b_2), ..., (a_m, b_n) \end{array}$$

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 defini (Définition **1.6**) le cardinal de E par:

$$|E| = \left\{ \begin{array}{c} n \text{ si } E \text{ possede } n \text{ elements} \\ \infty \text{ si } E \text{ est infini} \end{array} \right\}$$

(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 elements vaut $C_n^k = \dfrac{n!}{k!\,(n-k)!}$. Donner le nombre d'arragements d'une partie de k elements. En deduire le nombre total d'arragements de k elements.

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

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

(voir demonstration par recurrence Proposition **1.7**)

4. Montrer par recurrence sur n que $\sum\limits_{k=0}^{n} k = \dfrac{n\,(n+1)}{2}$.

Chapitre 2

ESPACES PROBABILISES

1. Experiences Aléatoires

L'objectif de ce chapitre est de construire un modéle mathematique pour gérer les phénoménes qui évoluent au hasard c'est à dire de façon imprevisible. Il s'agit de techniques mathematiques rigoureuses destinées à mesurer le caractère aléatoire de ces phénoménes en vue de décisions eventuelles concernant ces phénoménes.La construction du modèle commence par une description exhaustive du phénoméne à etudier. C'est l'objet de ce qui suit.

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

-lancer une piece de monnaie, résultats $\{pile, face\}$

-lancer un dé, résultats $\{1, 2, 3, 4, 5, 6\}$

-prevoir la météo un jour donné

-nombre de pieces défecteuses dans un ensemble de pieces fabriquées dans un atelier.

Une experience qui n'est pas alétoire est dite deterministe, son résultat est connu d'avance avec certitude.

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

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

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

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

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

Exemple.(1) Avec l'experience du dé on a $\Omega = \{1, 2, 3, 4, 5, 6\}$ et comme exemple l'evenement 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 experience $\Omega = \{1, 2, 3, ..., 12\}$ et si A est l'evenement "boule de numero multiple de 3" alors $A = \{3, 6, 9, 12\}$

1.4 Interpretation des opérations d'ensembles en termes d'evenements
Soit Ω l'ensemble fondamental d'une experience et soient A, B des evenements:

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

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

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

$A \cup B$ est l'evenement réalisé par l'un au moins des evenements A ou B

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

L'étape qui va suivre cette base descriptive du phénoméne consiste à mesurer le caractère incertain des evenements. C'est une étape importante de la construction du modèle qui doit respecter des conditions strictes d'application. Ce modèle mathematique 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 function
$P : \mathcal{P}(\Omega) \longrightarrow [0,1]$ satisfaisant les conditions:
(a) $P(\Omega) = 1$
$(b).\sigma-$additivité: pour toute suite $(A_n) \subset \mathcal{P}(\Omega)$, telle que:
$A_i \cap A_j = \phi$, $\forall i \neq j$, on a :$P\left(\underset{n}{\cup} A_n\right) = \sum_n P(A_n)$
Le couple Ω, P s'appelle espace probabilisé.

Remarques

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

2. Il est facile de voir que $P(\phi) = 0$, en effet on a
$\Omega = \Omega \cup \phi$ et $P(\Omega) = P(\Omega) + P(\phi) \Longrightarrow P(\phi) = 0$
on peut aussi remarquer que $P(A^c) = 1 - P(A)$, en effet on a $\Omega = A \cup A^c$ et $A \cap A^c = \phi$ 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:
$\Omega = \{\omega_1, \omega_2, \omega_3, ...\}$ avec $\sum_n P(\omega_n) = 1$

Exemples.

(1) $\Omega = \{pile, face\}$ $P(pile) = \alpha$, $P(face) = 1 - \alpha$
(2) $\Omega = \{1, 2, 3, 4, 5, 6\}$ (experience du dé)
$P(i) = \dfrac{1}{6}$ pour $1 \leq i \leq 6$ (probabilité uniforme)
si A : "face paire" alors $A = \{2, 4, 6, \}$
on a $P(A) = P\{2, 4, 6, \} = P(2) + P(4) + P(6) = \dfrac{3}{6} = \dfrac{1}{2}$

(3) Une urne contient 12 boules numérotées de 1 à 12. On tire une boule au hasard. Dans cette experience $\Omega = \{1, 2, 3, ..., 12\}$ et $A = \{3, 6, 9, 12\}$ est l'evenement "No. multiple de 3". On définit P en posant:
$P(\omega) = x$ pour tout numero pair ω
$P(\omega) = y$ pour tout numero impair ω
on a alors $P(\Omega) = 6x + 6y = 1$ donc $x + y = \dfrac{1}{6}$
dans ce cas $P(A) = P(3) + P(6) + P(9) + P(12) = 2y + 2x = \dfrac{2}{6} = \dfrac{1}{3}$

(4) Considérons l'experience: 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 uniforme $P(i, j) = \dfrac{1}{36}, \forall (i, j) \in \Omega$; calculons la probabilité de $A = \{(i, j) \in \Omega : i = j\} = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$
on a $P(A) = P(1, 1) + P(2, 2) + P(3, 3) + P(4, 4) + P(5, 5) + P(6, 6) = \dfrac{6}{36} = \dfrac{1}{6}$;
alors $A^c = \{(i, j) \in \Omega : i \neq j\}$ et $P(A^c) = 1 - P(A) = 1 - \dfrac{1}{6} = \dfrac{5}{6}$

3. Probabilité conditionnelle-Indépendance

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

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

lire "probabilité de A si B "

l'exemple qui suit justifie cette définition:

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

si $A \subset \Omega$ alors $P(A) = \dfrac{|A|}{n}$. Soient A, B 2 evenements; parmi les résultats qui réalisent B, $A \cap B$ sont ceux qui réalisent A donc si B se réalise , la probabilté de réaliser A vaut $\dfrac{|A \cap B|}{|B|}$ que l'on peut écrire $\dfrac{\frac{|A \cap B|}{n}}{\frac{|B|}{n}} = \dfrac{P(A \cap B)}{P(B)}$

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 evenements A et B sont indépendants si

$P(A|B) = P(A)$,autrement dit $P(A \cap B) = P(A).P(B)$

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

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

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

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

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

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

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

Le cas des evenements $[A^c, B]$ 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 experience $\Omega = \{1, 2, 3, ..., 12\}$ et $A = \{3, 6, 9, 12\}$ est l'evenement "No. multiple de 3". On définit P en posant:

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

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

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

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

il est facile de montrer que les evenements 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 evenements A, B on a:

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

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

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

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

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

3.5 Définition. Soit $\{A_i, i\in I\}$ une famille arbitraire d'evenements on dit que ces evenements sont indépendants dans leur ensemble si:
pour toute partie finie $J\subset I$ on a $P\left(\underset{j\in J}{\cap}A_j\right)=\underset{j\in J}{\Pi}P(A_j)$

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

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

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

C'est une représentation numérique des résultats d'une experience.
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 répartitionde 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 $\lim\limits_{\substack{x \to a \\ x > a}} .F(x) = F(a)\ \forall a \in \mathbf{R}$

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

(d) $\lim\limits_{x \to -\infty} .F(x) = 0,\ \lim\limits_{x \to \infty} .F(x) = 1$

1.4 Exemples.

(a) On lance 3 fois une piece 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 :

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

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

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

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

(b) Une urne contient 15 boules:
7 boules de 1kg, 5 boules de 3kg, 3 boules de 5kg
on tire un boule ω au hasardet 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
$P(X = 1) = \dfrac{7}{15}$, $P(X = 3) = \dfrac{5}{15}$, $P(X = 5) = \dfrac{3}{15}$

2. Variables alétoires discrétes

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

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: $\mathcal{B}(n, p)$, $n \geq 0$ $0 < p < 1$
$X(\Omega) = \{0, 1, 2, 3..., n\}$, $P(X = k) = C_n^k.p^k.(1 - p)^{n-k}$, $k = 0, 1, 2, ..., n.$
pour vérifier que $\sum\limits_{k=0}^{n} .P(X = k) = 1$, on utilise le Binome de Newton
$\sum\limits_{k=0}^{n} C_n^k \; a^k b^{n-k} = (a + b)^n$ avec $a = p$, $b = 1 - p$
(c) Loi géometrique $\mathcal{G}(p)$, $0 < p < 1$
$X(\Omega) = \{1, 2, 3...\}$, $P(X = k) = p.(1 - p)^{k-1}$, $k = 1, 2, 3, ...$
$\sum\limits_{k=1}^{\infty} .p.(1 - p)^{k-1} = 1$

(d) Loi de Poisson $\mathcal{P}(\lambda)$, $\lambda > 0$
$X(\Omega) = \{0, 1, 2, 3...\}$, $P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k = 0, 1, 2, 3, ...$
$\sum\limits_{k=0} .e^{-\lambda}\dfrac{\lambda^k}{k!} = 1$ (utiliser la relation $\sum\limits_{k=0} .\dfrac{\lambda^k}{k!} = e^{\lambda}$)

3. Esperance-Variance de V.A discrétes

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

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

3.1 Définition.

L'esperance de X est définie par:

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

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: $\mathcal{B}(p), 0 < p < 1$

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

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

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

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

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

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

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

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

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

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

$E(X) = \sum\limits_{k=1}^{\infty} .k.e^{-\lambda}\dfrac{\lambda^k}{k!} = \lambda$ (voir calculs en Annexe)

3.3 Propriétés de l'Esperance.

(1) $\forall a, b \in R$, $E(aX + bY) = a.E(X) + b.E(Y)$
(2) $|E(X)| \leq E(|X|)$
(3) $X \leq Y \implies E(X) \leq E(Y)$

3.4 Définition. (Variance)

On considére une V.A discrete X sur un espace Ω, P avec:
$X(\Omega) = \{x_1, x_2, x_3....\}$ et la loi de X donnée par $f(x_n) = P(X = x_n)$ $n = 1, 2, 3...$
La variance de X est définie par:

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

$\sigma_X = \sqrt{V(X)}$ 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) $\forall a, b \in R$, $V(aX + b) = a^2.V(X)$

3.4 Variance de Lois Discrétes 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: $\mathcal{B}(n, p)$, $n \geq 0$ $0 < p < 1$

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

$E(X) = \sum\limits_{k=0}^{n} .k.C_n^k\, p^k.(1 - p)^{n-k} = n.p$ (voir calculs en Annexe)

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

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

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

$E(X) = \sum\limits_{k=1}^{\infty} .k.p.(1 - p)^{k-1} = \dfrac{1}{p}$ (voir calculs en Annexe)

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

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

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

$E(X) = \sum\limits_{k=1}^{\infty} .k.e^{-\lambda}\dfrac{\lambda^k}{k!} = \lambda$ (voir calculs en Annexe)

$V(X) = \lambda$

4. Variables alétoires 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 $F(x) = \int_{-\infty}^{x} f(x).dx = P(\omega \in \Omega : X(\omega) \leq x)$ avec f fonction réélle 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 $F(b) - F(a) = \int_{a}^{b} f(x).dx = P(a \leq X \leq b) = P(\omega \in \Omega : X(\omega) \in [a,b])$ et $P(X = x) = \int_{x}^{x} f(x).dx = 0, \forall x$.

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:**

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

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

Fonction de répartition donnée par:

$F(x) = 0$ si $x < a$

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

$F(x) = 1$ si $x > b$

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

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

$f(x) = 0$, si $x < 0$

Fonction de répartition:

$F(x) = 0$, si $x < 0$

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

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

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

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

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

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

(b) On définit la variance de X par:

$V(X) = E\left[(X-\mu)^2\right] = \int_{-\infty}^{+\infty} (x-\mu)^2 .f(x).dx$, l'existence de $V(X)$ equivaut à l'integrabilité de la fonction $(x-\mu)^2 .f(x)$

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

4.4 Exemples. (Rappel: $V(X) = E\left(X^2\right) - \mu^2$), 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:
$I_A(x) = 1$ si $x \in A$ et $I_A(x) = 0$ si $x \in A^c$.
(b) Si f est une fonction de $\mathbf{R}$ dans $\mathbf{R}$ on note la fonction définie par:
$(f.I_A)(x) = f(x)$ si $x \in A$
$(f.I_A)(x) = 0$ si $x \in A^c$

5.2.Définition. Soit X une V.A sur Ω, P, et $A \subset \Omega$
$(X.I_A)(\omega) = X$ si $\omega \in A$
$(X.I_A)(\omega) = 0$ si $\omega \in A^c$

5.3.Définition. (Integrale locale)

(a) Soit f une fonction de $\mathbf{R}$ dans $\mathbf{R}$ et soit A une partie de $\mathbf{R}$, l'integrale locale de la fonction f sur A est définie par $\int\limits_A f(x).dx = \int\limits_{\mathbf{R}} (f.I_A)(x).dx$.

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

on obtient $\int\limits_A f(x).dx = \int\limits_a^{+\infty} f(x).dx$.

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

Inégalité de Jensen

$|E(X)| \leq E(|X|)$
Plus géneralement si $f : \mathbf{R} \longrightarrow \mathbf{R}$ est une fonction convexe telle que $E(|f(X)|) < \infty$, 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:

$\forall \epsilon > 0,\ \lim\limits_{n\to\infty} .P\left(|X_n - X| > \epsilon\right) = 0$ (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 $\lim\limits_{n\to\infty} .X_n(\omega) = X(\omega)$, $\forall \omega \notin A$ (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:

$\lim\limits_{n\to\infty} .F_n(t) = F(t)$ en tout point de continuité t de F(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 $\lim\limits_{n\to\infty} .np_n = \lambda > 0$

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

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

6.3.Theoreme

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

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

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:**

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

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

(b) **Loi forte:**

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

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

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

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

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 $\sigma = 1, m = 0$:**

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

Theoreme:

Soit 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:

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

$\frac{M_n - \mu}{\sigma/\sqrt{n}}$ 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: $\lim\limits_{n \to \infty} .P\left(\frac{M_n - \mu}{\sigma/\sqrt{n}} \leq x\right) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} .dx$

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 $F(x) = \int_{-\infty}^{x} f(x).dx$ avec f fonction réélle 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 $F(b) - F(a) = \int_{a}^{b} f(x).dx = P(a \leq X \leq b) = P(\omega \in \Omega : X(\omega) \in [a,b])$ 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:**

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

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

Fonction de répartition donnée par:

$F(x) = 0$ si $x < a$

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

$F(x) = 1$ si $x > b$

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

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

$f(x) = 0$, si $x < 0$

Fonction de répartition:

$F(x) = 0$, si $x < 0$

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

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

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

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

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

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:

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

(b) On définit la variance de X par:

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

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

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

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

$f(x) = \dfrac{1}{b-a}$ si $x \in [a,b]$
$f(x) = 0$ si $x \notin [a,b]$

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

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

pour calculer la variance on utilise la relation:

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

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

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

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

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

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

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

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

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

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

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

$$E(X) = \lambda.\int\limits_0^\infty x.e^{-\lambda x}.dx = \lambda.\int\limits_0^\infty -\frac{de^{-\lambda x}}{d\lambda}.dx$$

d'aprés une propriété de l'integrale en théorie de la mesure on peut écrire:

$$\int\limits_0^\infty -\frac{de^{-\lambda x}}{d\lambda}.dx = -\frac{d}{d\lambda}\int\limits_0^\infty e^{-\lambda x}.dx = -\frac{d}{d\lambda} - \frac{1}{\lambda}.\left[e^{-\lambda.x}\right]_0^\infty = -\frac{d}{d\lambda}\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda^2}$$

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

Pour calculer la variance on evalue $E\left(X^2\right) = \lambda.\int\limits_0^\infty x^2.e^{-\lambda x}.dx$

en integrant par parties; on trouve $V(X) = E\left(X^2\right) - \left(E(X)\right)^2 = \dfrac{1}{\lambda^2}$.

Annexe 2
Inégalités

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

(a) La fonction indicatrice de A est définie par:
$I_A(x) = 1$ si $x \in A$ et $I_A(x) = 0$ 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:
$(f.I_A)(x) = f(x)$ si $x \in A$
$(f.I_A)(x) = 0$ si $x \in A^c$

5.2.Définition. Soit X une V.A sur Ω, P, et $A \subset \Omega$
$(X.I_A)(\omega) = X$ si $\omega \in A$
$(X.I_A)(\omega) = 0$ si $\omega \in A^c$

5.3.Définition. (Integrale locale)

(a) Soit f une fonction de $\mathbf{R}$ dans $\mathbf{R}$ et soit A une partie de $\mathbf{R}$, l'integrale locale de la fonction f sur A est définie par $\int\limits_A f(x).dx = \int\limits_{\mathbf{R}} (f.I_A)(x).dx$.
A titre d'exemple on peut prendre $A =]a, +\infty[$
on obtient $\int\limits_A f(x).dx = \int\limits_a^{+\infty} f(x).dx$.
(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:
$P(|X| > \lambda) \leq \dfrac{E(|X|)}{\lambda}$

solution.
on a $|X| \geq |X|.I_{(|X|>\lambda)} > \lambda.I_{(|X|>\lambda)}$
$\Longrightarrow E(|X|) \geq E\left(\lambda.I_{(|X|>\lambda)}\right) = \lambda.P(|X| > \lambda)$
par la croissance de l'esperance
comme on a $E\left(I_{(|X|>\lambda)}\right) = P(|X| > \lambda)$ (Definition **5.1** (a))
on deduit $P(|X| > \lambda) \leq \dfrac{E(|X|)}{\lambda}$

Inégalité de Jensen

$|E(X)| \leq E(|X|)$
Plus géneralement si $f : \mathbf{R} \longrightarrow \mathbf{R}$ est une fonction convexe telle que $E(|f(X)|) < \infty$, alors: $f(E(X)) \leq E(f(X))$

solution.
Pour toute V.A X on définit les V.A:
$X^+ = \sup(X, 0)$ et $X^- = \inf(-X, 0)$
il n'est pas difficile de montrer que:
$X = X^+ - X^-$, $\quad |X| = X^+ + X^-$, d'où par linéarité de l'esperance:
$E(X) = E(X^+) - E(X^-)$, $\quad E(|X|) = E(X^+) + E(X^-)$, on obtient:
$|E(X)| = |E(X^+) - E(X^-)| \leq |E(X^+)| + |E(X^-)| = E(|X|)$

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

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

$|X - E(X)|^2 \geq |X - E(X)|^2 .I_{(|X-E(X)| \geq \lambda)} \geq \lambda^2.I_{(|X-E(X)| \geq \lambda)}$
$\Longrightarrow E|X - E(X)|^2 \geq \lambda^2.E\left(I_{(|X-E(X)| \geq \lambda)}\right)$

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

Annexe 3

Convergence de suites de variables

A. Variables aléatoires independandantes

Soit (X_n) une suite de variables aléatoires
Les v.a X_n sont dites independantes si $\forall k \geq 1$ et pour tous $\{A_1, A_2, ..., A_k\} \subset \mathcal{P}(\Omega)$:
$P((X_1 \in A_1) \cap (X_2 \in A_2) \cap ... \cap P(X_k \in A_k)) = \prod_{j=1}^{k} P(X_j \in A_j)$.

Propriété fondamentale:
Si les v.a X_n sont independantes et d'esperances finies alors:
$(*) \quad \forall n \geq 2$ on a $E(X_1.X_2....X_n) = E(X_1).E(X_2)....E(X_n)$
On donne ici une preuve dans le cas de 2 v.a discretes X, Y:
supposons donc $X(\Omega) = \{x_1, x_2, ...\}$, $Y(\Omega) = \{y_1, y_2, ...\}$
ona $E(X.Y) = \sum_n \sum_k x_n.y_k.P(X = x_n, Y = y_k) = \sum_n \sum_k x_n.y_k.P(X = x_n).P(Y = y_k)$
où on a utilisé l'independance de X et Y dans la relation
$P(X = x_n, Y = y_k) = P(X = x_n).P(Y = y_k)$
remarquons enfin que $\sum_n \sum_k x_n.y_k.P(X = x_n).P(Y = y_k)$
$= \sum_n x_n.P(X = x_n).\sum_k y_k.P(Y = y_k) = E(X).E(Y)$

Remarques:
1. Il est possible de démontrer la propriété $(*)$ par recurrence 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 necessite des techniques d'integration 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:
$\forall \epsilon > 0,\ \lim_{n \to \infty} .P(|X_n - X| > \epsilon) = 0$ (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 $\lim_{n \to \infty} .X_n(\omega) = X(\omega)$, $\forall \omega \notin A$ (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:
$\lim_{n \to \infty} .F_n(t) = F(t)$ en tout point de continuité t de F (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 $\mathcal{B}(n, p_n)$, $n \geq 0 \quad 0 < p_n < 1$:

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

Supposons que $\lim\limits_{n \to \infty} .np_n = \lambda > 0$
X_n converge en loi vers la V.A X de Poisson de parametre λ
i.e $P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k = 0, 1, 2, 3, ...$

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

dans ce cas $P(X_n = k) = C_n^k.\left(\dfrac{\lambda}{n}\right)^k.\left(1-\dfrac{\lambda}{n}\right)^{n-k}$, pour n grand
et cela implique:

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

il est facile de voir que $\lim\limits_{n \longrightarrow \infty} .\dfrac{n!}{(n-k)!.n^k} = 1$ et $\lim\limits_{n \longrightarrow \infty} .\left(1-\dfrac{\lambda}{n}\right)^{-k} = 1$

il reste à calculer $\lim\limits_{n \longrightarrow \infty} .\left(1-\dfrac{\lambda}{n}\right)^n$; on pose $x = -\dfrac{n}{\lambda}$

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

$\lim\limits_{x \longrightarrow \infty} .\left(1+\dfrac{1}{x}\right)^x = e$ donc $\lim\limits_{x \longrightarrow \infty} .\left(1+\dfrac{1}{x}\right)^{-\lambda x} = e^{-\lambda}$

finalement $\lim\limits_{n \longrightarrow \infty} .\dfrac{n!}{(n-k)!.n^k}.\left(1-\dfrac{\lambda}{n}\right)^n.\left(1-\dfrac{\lambda}{n}\right)^{-k} = e^{-\lambda}$ et

$\lim\limits_{n \longrightarrow \infty} .P(X_n = k) = \dfrac{\lambda^k}{k!}.e^{-\lambda}$.

Theoreme
Soit X_n, X une suite de V.A.sur Ω, P, alors on a:
$X_n \xrightarrow{p.s} X \Longrightarrow X_n \xrightarrow{P} X \Longrightarrow X_n \xrightarrow{\mathcal{L}} X$
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:

$M_n = \dfrac{X_1 + X_2 + ... + X_n}{n}$ converge en probabilité vers μ.
i.e $P(|M_n - \mu| > \epsilon) \longrightarrow 0, \forall \epsilon > 0$

(*b*) **Loi forte:**

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

Demonstration:

(*a*) **Loi faible:**

On utilise l'inégalité de Tchebychev:

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

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

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

puisque les v.a sont independantes et de même loi on a:

$E\left(X_i\right) = \mu, \forall i$ et $E\left(M_n\right) = \dfrac{n.\mu}{n} = \mu$ et de plus $V\left(M_n\right) = \dfrac{\sum\limits_{i=1}^{n} V\left(X_i\right)}{n^2}$

avec $V\left(X_i\right) = \sigma^2, \forall i \Longrightarrow V\left(M_n\right) = \dfrac{\sigma^2}{n}$

l'inégalité de Tchebychev donne:

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

(*b*) **Loi forte:**

La demonstration de la loi forte est plus elaborée et necessite
des techniques avancées de théorie des probabilités
qui ne sont pas au programme.

Annexe 4

9. Soit Y une variable binomiale de parametre $0 < p < 1$:
$P(Y=k) = C_n^k.p^k(1-p)^{n-k}$, $k = 0, 1, 2..., n$.
(a) Calculer son esperance $E(Y)$
(b) Donner un exemple d'experience gérée par Y.

Rappel: **Formule du Binome de Newton:**
$\forall a, b$ dans $\mathbb{R}$, $(a+b)^n = \sum_{k=0}^{n} C_n^k.a^k.b^{n-k}$

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

on a $k.C_n^k = k.\frac{n!}{k!(n-k)!} = k.\frac{n!}{k.(k-1)!(n-k)!} =$
$\frac{n.(n-1)!}{(k-1)!((n-1)-(k-1))!}$ puisque $n-k = (n-1)-(k-1)$ on obtient:
$k.C_n^k = n.C_{n-1}^{k-1}$, on remplace $k.C_n^k$ par $n.C_{n-1}^{k-1}$ dans la formule $(*)$, donc:
$E(Y) = \sum_{k=1}^{n} k.C_n^k.p^k(1-p)^{n-k} = \sum_{k=1}^{n} n.C_{n-1}^{k-1}.p^k(1-p)^{n-k} =$
$n.p.\sum_{k=1}^{n} C_{n-1}^{k-1}.p^{k-1}(1-p)^{(n-1)-(k-1)}$, mais d'aprés

la **Formule du Binome de Newton**
on a $\sum_{k=1}^{n} C_{n-1}^{k-1}.p^{k-1}(1-p)^{(n-1)-(k-1)} = (p+(1-p))^n = 1$

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

(b) Exemple d'experience gérée par Y :
Le pourcentage de pièces défecteuses d'un procédé de fabrication industrielle vaut $0 < p < 1$. Cela signifie que la probabilité qu'une piece fabriquée soit défecteuse vaut p Si on considère au hazard un ensemble de n pièces fabriquées par ce procédé le nombre Y de pièces défecteuses 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 $A_1 \subset \Omega_1, A_2 \subset \Omega_2$, $P(A) = P_1(A_1).P_2(A_2)$.
On note $P = P_1 \otimes P_2$
Pour une piece donnée on suppose: $P(pile) = \alpha$.
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.
$\pi(p,p) = \pi(p).\pi(p) = \alpha^2$, $\pi(p,f) = \pi(p).\pi(f) = \alpha.(1-\alpha)$

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

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 $\Omega = \{(\omega_1,\omega_2) : 1 \le \omega_1 \le 6,\ \ 1 \le \omega_2 \le 6\}$
$\pi(\omega_1) = \pi(\omega_2) = \dfrac{1}{6},\ \ \forall\omega_1,\omega_2$ compris entre 1 et 6 puisque chaque
jet est géré par la probabilité uniforme π.
Donc $P(\omega_1,\omega_2) = \pi\otimes\pi(\omega_1,\omega_2) = \pi(\omega_1)\times\pi(\omega_2) = \dfrac{1}{36}$.
Par consequent P est la probabilité uniforme sur Ω.
(b) $S(\omega_1,\omega_2) = \omega_1 + \omega_2$
$(S=7) = \{(1,6),(6,1),(2,5),(5,2),(3,4),(4.3)\}$
$P(S=7) = P\{(1,6),(6,1),(2,5),(5,2),(3,4),(4.3)\} = \dfrac{6}{36} = \dfrac{1}{6}$
on calcule $P(S=8)$, $P(S=9)$ de façon analogue.

12. Soient X,Y 2 $V.A$ de Ω dans $\mathbf{R}$ definies sur Ω, P et de méme 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 méme loi:
(1) Cas discret: $X(\Omega) = Y(\Omega) = \{v_1, v_2, ...\}$ et $P(X=v_n) = P(Y=v_n)$
$E(X) = \sum\limits_{v_n} v_n.P(X=v_n) = \sum\limits_{v_n} v_n.P(Y=v_n) = E(Y)$

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

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

$$E(X) = E(Y) = \int_I t.\varphi(t).dt, \qquad E(X^2) = E(Y^2) = \int_I t^2.\varphi(t).dt$$

$$V(X) = E(X^2) - (E(X))^2, \qquad V(Y) = E(Y^2) - (E(Y))^2$$

$$= \int_I t^2.\varphi(t).dt - \left(\int_I t.\varphi(t).dt\right)^2.$$