El Hassan Essaky
Faculté Polydisciplinaire de Safi
Département Maths-Info

Cours de Probabilité-Statistique

SMA-SMI Semestre 3

First draft

Contents

Chapitre 1

Dénombrement

1.1 Dénombrement.

1.1.1 Quelques définitions

Définition 1.

Disposition : sous ensemble ordonné ou non d'un ensemble.
Disposition sans répétition : c'est une disposition où un élément peut apparaître 0 ou 1 fois.
Disposition avec répétition : un élément peut figurer plus d'une fois.
Disposition ordonnée : l'ordre d'obtention d'un élément est important.
Disposition non-ordonnée : l'ordre d'obtention d'un élément n'est pas important.

Exemple : Prenons un jeu de dé à 6 faces numérotées par $\Omega = \{1;2;3;4;5;6\}$. Après 3 jets, nous obtenons la réalisation $A = (2;5;1)$; nous réitérons les jets et nous obtenons $B = (5;1;2)$. A et B sont équivalents si nous considérons que les dispositions sont non-ordonnées. En revanche, ils ne sont pas équivalents si nous sommes dans le cadre d'une disposition ordonnée.

1.1.2 Principe multiplicatif

Principe: Soit ξ une expérience qui comporte 2 étapes : la 1ère qui a p résultats possibles et chacun de ces résultats donne lieu à q résultats lors de la 2ème. Alors l'expérience a $p \times q$ résultats possibles.
D'une façon plus générale : si un événement peur se produire de p_1 façons et si suivant cet événement un second événement peut se produire de p_2 façons et ainsi de suite jusq'au $k$ème qui peut se produire de p_k façons alors le nombre de façons dont les événements peuvent se produire est $p_1 \times ... \times p_k$.
Exemple :
1. On jette une pièce de monnaie 3 fois. Combien y-a-t-il de résultats possibles? le premier lancer donne lieu à deux résultats possibles Pile (P) ou Face (F). Même chose pour le second et le troisième lancers. D'après le principe multiplicatif il y a $2 \times 2 \times 2 = 8$ résultats possibles.
2. Une urne contient $1R, 1B, 1N, 1V$. On effectue 2 tirages successifs avec remise. Combien y-a-t-il de résultats possibles? 1ér tirage : 4 possibilités. 2ème tirage: 4 possibilités. Il y en a donc 4×4 possibilités.
Conséquence
Si une expérience ξ consiste à répéter n fois de façon indépendante une même experience ξ_0 qui a p résultats possibles, alors ξ a $p^n = p \times p \times ... \times p$ issues possibles.
Exemple : ξ_0: "lancer une pièce de monnaie". On répète cette expérience 3 fois. Alors le nombre de résultats possible est $2^3 = 8$ (ici $p = 2$).

1.1.3 Arrangement sans répétition

Exemple :
On a 3 lettres a, b, c. Combien y-a-t-il de mots de 2 lettres différentes? On a ab, ba, ac, ca, bc, cb : il y a 6 mots de 2 lettres : arrangements de 2 éléments parmi 3.

Définition 2.

Soit un ensemble Ω composé de n éléments. Nous constituons un échantillon E de taille p. On appelle arrangement de p objets toute disposition ordonnée et sans répétition de p objets pris parmi n. Ce nombre est noté : A_n^p. Dans ce cas, on a necessairement : $1 \leq p \leq n$. Si $n < p$, alors $A_n^p = 0$.

Remarque: Réaliser un arrangement sans répétition des éléments de Ω, c'est déterminer un p-uplet $(x_1, ..., x_p)$ d'éléments de Ω deux à deux distincts. C'est aussi définir une application injective d'un ensemble E à p éléments dans Ω à n éléments.

Résultat : Le nombre A_n^p est donné par

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

Proof. Pour le premier objet tiré on a : n possibilités.
Pour le 2ème objet tiré on a : $n-1$ possibilités.
Pour pème objet tiré on a : $n-p+1$ possibilités. Donc par le principe multiplicatif on a

$$n \times (n-1) \times ... \times (n-p+1) = n \times (n-1) \times ... \times (n-p+1) \frac{(n-p) \times ... \times 2 \times 1}{(n-p) \times ... \times 2 \times 1} = \frac{n!}{(n-p)!}.$$

□

Exemple: Pour accéder à une banque de donnéees, il faut 4 lettres (de la langue française) différentes. Combien de mots de passe peut-on avoir? $A_{26}^4 = 258800$.

1.1.4 Arrangement avec répétition

Définition 3.

Soit un ensemble E composé de n éléments. Nous constituons un échantillon E de taille p à partir des éléments de Ω. Si nous avons à choisir p éléments parmi n dans une disposition ordonnée (les places sont distinctes) et avec répétition (on peut choisir le même élément plusieurs fois), on dit qu'on a un arrangement de p éléments parmi n. Le nombre d'arrangement avec répétition est n^p.

Remarque: Réaliser un arrangement avec répétition des éléments de E, c'est aussi définir une application d'un ensemble F à p éléments dans E à n élément.

1.1.5 Permutation sans répétition

Définition 4.

Soit E un ensemble fini de $n \in \mathbb{N}^*$ d'éléments. Une permutation de n éléments de E est une disposition ordonnée sans répétition de ces n éléments. C'est un arrangement sans répétition de n éléments parmi n : $A_n^n = n!$.

Remarque: Réaliser une permutation sans répétition C'est aussi définir une bijection de l'ensemble Ω sur lui-même. L'ensemble des permutations d'un ensemble à n éléments s'appelle le groupe symétrique d'ordre n et se note S_n.

Exemple: De combien de manière peut-on classer 4 individus? $A_4^4 = 4! = 24$.

1.1.6 Combinaison sans répétition

Exemple:
On a 4 éléments a, b, c, d. Les combinaisons de 2 éléments sont $\{a,b\},\{a,c\},\{a,d\},\{b,c\},\{b,d\},\{c,d\}$. Il y en a 6.

Définition 5.

Soient $p \in \mathbb{N}$ et $n \in \mathbb{N}$ tels que $0 \leq p \leq n$. Soit E un ensemble fini de n d'éléments. Une combinaison de p éléments de E parmi n est une disposition non ordonnée de p de E éléments parmi n. (sous ensembles de p éléments parmi n).

Résultat : Le nombre de combinaisons de p éléments parmi n est donné par:

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

Proof. A partir d'une combinaison de p éléments on peut faire $p!$ arrangements. Donc $A_n^p = p!C_n^p$. □

Remarque : La question combien y-a-t-il de combinaisons? peut être traduite par : combien y-a-t-il de façons de tirer p éléments parmi n sans remise et sans prise en compte de l'ordre des éléments tirés.
Propriétés :

1. $C_n^0 = C_n^n = 1$.
2. $C_n^p = C_n^{n-p}$ (complémentaire).
3. $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$ (triangle de Pascal).

Proof. Pour le troisième point, on a

$$\begin{aligned} C_{n-1}^p + C_{n-1}^{p-1} &= \frac{(n-1)!}{p!(n-p-1)!} + \frac{(n-1)!}{(p-1)!(n-p)!} \\ &= \frac{n.(n-1)!}{p!(n-p)!} = C_n^p. \end{aligned}$$

□

Proposition 6.

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k} \quad \text{(formule de Binôme).}$$

Proof. Par récurrence sur n. On a

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)\sum_{k=0}^{n} C_n^k a^k b^{n-k} \\
&= \sum_{k=0}^{n} C_n^k a^{k+1} b^{n-k} + \sum_{k=0}^{n} C_n^k a^k b^{n+1-k} \\
&= \sum_{k=1}^{n+1} C_n^{k-1} a^k b^{n+1-k} + \sum_{k=0}^{n} C_n^k a^k b^{n+1-k} \\
&= \sum_{k=1}^{n} C_n^{k-1} a^k b^{n+1-k} + C_n^n a^{n+1} b^0 + \sum_{k=1}^{n} C_n^k a^k b^{n+1-k} + C_n^0 a^0 b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} (C_n^{k-1} + c_n^k) a^k b^{n+1-k} + b^{n+1} \\
&= \sum_{k=0}^{n+1} C_{n+1}^k a^k b^{n+1-k}.
\end{aligned}$$

□

1.2 Les ensembles

Un ensemble est une collection d'objets appelés éléments.
L'ensemble vide $\emptyset$: il ne contient aucun élément.
Soit Ω un ensemble.
Un ensemble A est un sous ensemble de Ω ou une partie de Ω si tous les éléments de A sont éléments de Ω. On dit aussi que A est une partie de Ω.
L'ensemble des parties de Ω est noté $\mathcal{P}(\Omega)$.
Exemple : l'ensemble des parties de Ω avec $\Omega = \{a, b, c\}$.

$$\mathcal{P}(\Omega) = \{\Omega, \{a, b\}, \{a, c\}, \{b, c\}, \{a\}, \{b\}, \{c\}, \emptyset\}.$$

Définition 7.

Soient $A, B \in \mathcal{P}(\Omega)$.

1. A et B sont disjoints ssi $A \cap B = \emptyset$.
2. $A \cup B = \{x \in \Omega : x \in A \quad \text{ou} \quad x \in B\}$, réunion de A et B.
3. $A \cap B = \{x \in \Omega : x \in A \quad \text{et} \quad x \in B\}$, intersection de A et B.
4. $A^c = \{x \in \Omega : x \notin A\}$, complémentaire de A.
5. $A \backslash B = \{x \in \Omega : x \in A \quad \text{et} \quad x \notin B\}$, différence des ensembles A et B.

Propriétés :

1. $A \cup B = B \cup A$.
2. $A \cap B = B \cap A$
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
5. $(A \cap B)^c = A^c \cap B^c$
6. $(A \cup B)^c = A^c \cup B^c$.

1.3 Notion de cardinal

Si n et p sont deux entiers naturels tels que $p \leq n$, on pose $[p, n] = \{p, p+1, \ldots, n\}$.

Définition 8.

Étant donnés deux ensembles E et F. S'il existe une bijection de E dans F, on dit que E et F sont équipotents.

Exemple: Soit P l'ensemble des entiers pairs et I celui des entiers impairs, alors P et I sont équipotents, en effet l'application $\phi : P \to I : \phi(n) = n+1$ est une bijection.

Définition 9.

On dit qu'un ensemble E est fini s'il existe un entier naturel non nul n tel que E soit équipotent à $[1, n]$. Ce nombre n est appelé le cardinal de E, on écrit $card(E) = n$.

Si E a un nombre fini d'éléments, alors pour tout $A \in \mathcal{P}(E)$, A a également un nombre fini d'éléments.
Propriétés :

1. $card(A) = card(E) - card(A)$.
2. $card(A \cup B) = card(A) + card(B) - card(A \cap B)$
3. $card(A \backslash B) = card(A) - card(A \cap B)$
4. $card(\emptyset) = 0$.

Définition 10.

On dit qu'un ensemble E est dénombrable s'il est en bijection avec une partie de $\mathbb{N}$. Ce qui revient à dire que E est équipotent à une partie de $\mathbb{N}$.

Chapitre 2

Mesure de Probabilités

2.1 Du langage ensembliste à celui des événements

On a une expérience aléatoire (tirage d'une carte, lancement d'un dé, ...) qui est une action qui débouche sur plusieurs résultats possibles. Pour la définir, il s'agit de recenser l'ensemble de ses résultats possibles appelés éventualités ou événements élémentaires. On appelle univers l'ensemble de ses éventualités et il est noté Ω.
Un événement est un ensemble d'éventualités noté A, c'est donc aussi une partie de Ω, $A \subset \mathcal{P}(\Omega)$. On représente un événement par un ensemble.
Exemple : jet de dé = expérience aléatoire
Univers $\Omega = \{1,2,3,4,5,6\}$.
Eventualités $\{1\},\{2\}...\{6\}$.
Evénement A : tomber sur un impair $A = \{1,3,5\}$.
A^c est l'événement réalisé quand A ne l'est pas. $A^c = \{2,4,6\}$.
$A \cup B$ est l'événement qui est réalisé dès que l'un au moins des événements A ou B l'est.
$A \cap B$ est l'événement qui est réalisé dès que les événements A et B le sont.
$A = \{1,3,5\}$. B = tomber sur un nombre ≤ 3.
$A \cup B = \{1,2,3,5\}$
$A \cap B = \{1,3\}$.
Ω est l'événement certain : réalisé à coup sûr.
$\emptyset$ est l'événement impossible : obtenir 7 par exemple.
$A \subset B$ signifie que chaque fois que A est réalisé B l'est aussi.
$A \cap B = \emptyset$: événements incompatibles.
Soit l'ensemble des éventualités ou univers Ω. On arrive au point essentiel de définir la probabilité d'un événement A qui doit mesurer la chance que l'événement A a de se réaliser lorsqu'on effectue une épreuve.
La complexité de la définition d'une probabilité dépend de celle de l'univers Ω: fini, infini dénombrable, infini non dénombrable.

2.2 Espace de cardinal fini

Dans le cas où l'espace Ω est fini $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, la probabilité $\mathbb{P}(A)$ peut se définir pour tout sous ensemble $A \subset \Omega$ et la probabilité $\mathbb{P}$ est donnée par une suite finie $(p_i)_{1\leq i\leq n}$ qui est la suite des probabilités des événements élémentaires $p_i = \mathbb{P}(\{\omega_i\})$.

Définition 11.

Soit $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ un ensemble à n éléments. On définit une probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega)$ par la donnée d'une suite finie de réels positifs p_i de somme $\sum_{i=1}^{n} p_i = 1$ donnée par $p_i = \mathbb{P}(\{\omega_i\})$. Pour tout $A \subset \Omega, \mathbb{P}(A)$ est alors donnée par

$$\mathbb{P}(A) = \sum_{i,\omega_i \in A}^{n} p_i.$$

Exemple : Dans l'expérience lancement d'un dé on a $\Omega = \{1,2,3,4,5,6\} = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. On pose $p_i = \mathbb{P}(\{\omega_i\})$ et on considère l'événement A : tomber sur un pair $A = \{2,4,6\} = \{\omega_2, \omega_4, \omega_6\}$. On a

$$\mathbb{P}(A) = \sum_{i,\omega_i \in A}^{6} p_i = p_2 + p_4 + p_6 = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

Formellement on a $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$, $A \mapsto \mathbb{P}(A)$.
Remarque : On constate facilement que $\mathbb{P}$ satisfait les propriétés suivantes :

1. $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$.

2. Si $A \cap B = \emptyset$, alors
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$
Plus généralement si $A_1, A_2, ..., A_p$ sont 2 à 2 disjoints, alors
$$\mathbb{P}(A_1 \cup A_2 \cup ... \cup A_p) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + ... + \mathbb{P}(A_p).$$

Proof. Tout d'abord on a,
$$\mathbb{P}(\Omega) = \sum_{i, \omega_i \in \Omega} p_i = \sum_{i=1}^{n} p_i = 1.$$
Comme $\Omega = A \cup A^c$ et $A \cap A^c = \emptyset$, donc
$$1 = \sum_{i, \omega_i \in A} p_i + \sum_{i, \omega_i \in A^c} p_i = \mathbb{P}(A) + \mathbb{P}(A^c).$$
Comme $A \cap B = \emptyset$, on a
$$\mathbb{P}(A \cup B) = \sum_{i, \omega_i \in A \cup B} p_i = \sum_{i, \omega_i \in A} p_i + \sum_{i, \omega_i \in B} p_i = \mathbb{P}(A) + \mathbb{P}(B).$$
□

Remarque : Nous disons qu'il y a équiprobabilité lorsque les probabilités de tous les événements élémentaires sont égales.
Dans ce cas, $\mathbb{P}$ est la probabilité uniforme sur $(\Omega, \mathcal{P}(\Omega))$.
Conséquence : S'il y a équiprobabilité, pour tout événement A, nous avons alors
$$\mathbb{P}(A) = \frac{card(A)}{card(\Omega)} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Exemple :
1. On effectue une partie de pile ou face en trois coups. Quelle est la probabilité d'obtenir face au premier lancer et pile au dernier?
On modélise la situation en prenant $\Omega = \{P, F\}^3$ P désigne pile et F face. $\mathbb{P}$ est définie sur l'ensemble de toutes les parties de Ω. Il y a $8 = 2^3$ triplets de résultats possibles :
$$\{(P;P;P);(P;P;F);(P;F;P);(F;P;P);(F;F;P);(F;P;F);(P;F;F);(F;F;F)$$
Si on suppose la pièce bien équilibrée, à priori chacun de ces triplets est équiprobable. L'événement A cherché se décompose en $(F;F;P);(F;P;P)$. donc
$$\mathbb{P}(A) = \frac{card(A)}{card(\Omega)} = \frac{2}{8} = \frac{1}{4}.$$

2. Une urne contient une boule blanche(B) et une noire (N). On fait 2 tirages avec remise Probabilité d'avoir 2 Noires?
On a $\Omega = \{(N,N);(B,B);(B,N);(N,B)\}$. $card(\Omega) = 4$ (ou alors on applique le principe multiplicatif vu qu'il y a indépendance).
A : "avoir 2 Noires". $A = \{(N,N)\}$. On a
$$\mathbb{P}(A) = \frac{card(A)}{card(\Omega)} = \frac{1}{4}.$$

2.3 Espaces infinis dénombrables

Dans le cas où l'espace Ω est fini $\{\omega_1, \omega_2, ..., \omega_n, ...\}$, la probabilité $\mathbb{P}(A)$ peut se définir pour tout sous ensemble $A \subset \Omega$ et la probabilité $\mathbb{P}$ est donnée par une suite infinie $(p_i)_{i\geq 1}$ qui est la suite des probabilités des événements élémentaires $p_i = \mathbb{P}(\{\omega_i\})$.

Définition 12.

Soit $\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}$ un ensemble infini dénombrable. On définit une probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega)$ par la donnée d'une suite finie de réels positifs p_i de somme $\displaystyle\sum_{i=1}^{\infty} p_i = 1$ donnée par $p_i = \mathbb{P}(\{\omega_i\})$. Pour tout $A \subset \Omega, \mathbb{P}(A)$ est alors donnée par

$$\mathbb{P}(A) = \sum_{i, \omega_i \in A}^{\infty} p_i.$$

On constate encore que $\mathbb{P}$ satisfait les propriétés suivantes :

1. $\mathbb{P}(A) + \mathbb{P}(A^c) = 1.$
2. Si $A \cap B = \emptyset$, alors
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$
   Plus généralement si $A_1, A_2, ..., A_n, ...$ une suite infinie d'éléments sont 2 à 2 disjoints, alors
$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + ... + \mathbb{P}(A_n) + ... = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \quad \text{propriéte de } \sigma-\text{additivité}.$$

Remarque : Noter que l'équiprobabilité n'existe plus lorsque l'espace est dénombrable non fini : si on accorde la même probabilite p à chaque ω_i, par σ-additivité, on doit avoir

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\cup_{i=1}^{+\infty}\{\omega_i\}) = \sum_{i=1}^{\infty} \mathbb{P}(\{\omega_i\}) = \sum_{i=1}^{\infty} p = \infty.$$

2.4 Espace général

On l'a vu dans les cas précédents : la probabilité $\mathbb{P}$ est une fonction qui à un ensemble A associe un nombre compris entre 0 et 1, sa probabilité $\mathbb{P}(A)$. C'est donc une fonction d'ensembles, c'est à dire sur les ensembles. Cette fonction $\mathbb{P}$ doit vérifier un certain nombre de propriétés (poids total égale à 1, σ-additivité). Pour à la fois les satisfaire et être définie avec cohérence, on ne peut pas définir, en général, la probabilité de tous les sous-ensembles $A \subset \Omega$. $\mathbb{P}$ n'est donc pas en général définie sur tout $\mathcal{P}(\Omega)$,. On doit se restreindre à une famille d'événements $\mathcal{F} \subset \mathcal{P}(\Omega)$) qu'on appellera tribu ou de σ-algèbre). On peut alors définir la probabilité $\mathbb{P}$ sur l'ensemble $\mathcal{F}$ qui vérifie les propriétés suivantes:

1. $\emptyset \in \mathcal{F}$
2. si $A \in \mathcal{F}$, $A^c \in \mathcal{F}$
3. si $(A_n)_n$ est une suite de $\mathcal{F}$, alors $\cup_{n=1}^{+\infty} A_n \in \mathcal{F}$.

Exemple : $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ est une tribu. On peut définir une probabilité sur $\mathcal{F}$ par $\mathbb{P}(\emptyset), \mathbb{P}(\Omega) = 1, \mathbb{P}(A) = p \in [0;1]$ et $\mathbb{P}(A^c) = 1 - p$.

Définition 13.

Soient un ensemble Ω, $\mathcal{F}$ une tribu. On appelle probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{F})$ toute application $\mathbb{P}$ de $\mathcal{F}$ dans $[0;1]$ qui vérifie :

1. $\mathbb{P}(\Omega) = 1$.
2. Si $A_1, A_2, ..., A_n, ...$ une suite infinie d'éléments de $\mathcal{F}$ 2 à 2 disjoints, alors

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + ... + \mathbb{P}(A_n) + ... = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \quad \text{propriété de } \sigma - \text{additivité.}$$

On appelle $(\Omega;, \mathcal{F}, \mathbb{P})$ un espace probabilisé ou espace de probabilité.

Remarque : En pratique, pour vérifier que $\mathbb{P}$ est une probabilité, il suffit de vérifier que $\mathbb{P}(A) \geq 0$ pour tout $A \in \mathcal{F}$, que $\mathbb{P}(\Omega) = 1$ et que $\mathbb{P}$ est additive : $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ quand A et B sont disjoints.

Propriétés : On constate encore que $\mathbb{P}$ satisfait les propriétés suivantes :

1. Pour tous $A, B \in \mathcal{F}$, $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$. en effet, on a $\Omega = A \cup A^c$ et $A \cap A^c = \emptyset$. Donc
$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A) + \mathbb{P}(A^c).$$
2. $\mathbb{P}(\emptyset) = 0$.
3. Additivité : Si $A \cap B = \emptyset$, alors
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$
En effet, il suffit de prendre $A_1 = A, A_2 = B, A_i = \emptyset, \forall i \geq 2$.
4. Si $A_1, A_2, ..., A_n \in \mathcal{F}$ 2 à 2 disjoints, alors
$$\mathbb{P}(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$
5. Pour tous $A, B \in \mathcal{F}$ tels que $A \subset B$, on a $\mathbb{P}(A) \leq \mathbb{P}(B)$.
En effet $B = (B \backslash A) \cup A$ où la réunion est disjointe. On a donc
$$\mathbb{P}(B) = \mathbb{P}(A) + P(B \backslash A) \geq \mathbb{P}(A).$$
6. $\forall A, B \in \mathcal{F}$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - P(A \cap B)$. En effet
$$A \cup B = A \cup (B \backslash (A \cap B).$$
On a donc
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \backslash (A \cap B))$$
Or $B = (B \backslash (A \cap B)) \cup (A \cap B)$, donc
$$\mathbb{P}(B) = P(B \cap A) + \mathbb{P}((B \backslash (A \cap B)))$$
ce qui implique que
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \backslash (A \cap B)) = \mathbb{P}(A) + \mathbb{P}(B) - P(B \cap A)$$
7. Propriétés de continuité monotone séquentielle:
   (a) Si $(A_n)_{n \in \mathbb{N}^*}$ est une suite croissante d'événements (i.e. pour tout n, $A_n \subset A_{n+1}$) alors
$$\lim_{n \to +\infty} \mathbb{P}(A_n) = P(\cup_{n=1}^{n} A_n)).$$

(b) Si $(B_n)_{n\in\mathbb{N}^*}$ est une suite décroissante d'événements (i.e. pour tout n, $B_{n+1}\subset B_n$) alors

$$\lim_{n\to+\infty}\mathbb{P}(B_n)=P(\cap_{n=1}^{n}A_n)).$$

En effet, dans le cas croissant, on pose $C_n=A_n\backslash A_{n-1}$, $n\geq 2$ et $C_1=A_1$, donc $\cup_{k=1}^{n}C_k=A_n$. Comme les $(C_n)_n$ sont 2 à 2 disjoints alors

$$\lim_{n\to+\infty}\mathbb{P}(A_n)=\lim_{n\to+\infty}\mathbb{P}(\cup_{k=1}^{n}C_k)=\lim_{n\to+\infty}\sum_{k=1}^{n}\mathbb{P}(C_k)=\sum_{n=1}^{\infty}\mathbb{P}(C_n)=\mathbb{P}(\cup_{n=1}^{\infty}C_n)=\mathbb{P}(\cup_{n=1}^{\infty}A_n).$$

Chapitre 3

ABILITÉS CONDITIONNELLES-NOTION D'INDÉPENDANCE

3.1 Conditionnement

Le conditionnement a pour objet de répondre à la question suivante : comment se modifie la probabilité d'un événement lorsque l'on connaît déjà une information supplémentaire ? Dans la suite, on fixe un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

Définition 14.

Soit A un événement tel que $\mathbb{P}(A) > 0$. Pour tout événement $B \in \mathcal{F}$, on définit la probabilité conditionnelle de B sachant A, noté $\mathbb{P}(B|A)$ par :

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

Signification : $\mathbb{P}(A|B)$ exprime la probabilité de A quand on sait que B est déjà réalisé.
Exemple : On extrait sans remise 2 cartes d'un jeu de 40. Soit l'événement : A= "La première carte tirée est un roi". Quelle est la probabilité que la 2ème carte soit aussi un roi, on note cet événement par B?
Directement on a, $\mathbb{P}(A|B) = \frac{3}{39}$.
Si on veut appliquer la définition, c'est plus compliqué. On a tout d'abord $\mathbb{P}(A) = \frac{4}{40}$.
$\mathbb{P}(A \cap B)$? il s'agit de définir l'univers Ω sur lequel est défini $A \cap B$. Soit l'experience aléatoire ξ : tirer 2 cartes sans remise. On a $card(\Omega) = A_{40}^2$ et $card(A \cap B) = A_4^2$. Ainsi

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{3}{9}.$$

Proposition 15.

Soit dans un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, $A \in \mathcal{F}$ tel que $P(A) > 0$. Alors la fonction P_A definie par :

$$P_A : B \in \mathcal{F} \to P_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \in [0, 1]$$

est une probabilité sur $\mathcal{F}$.

Définition 16.

Soient $A, B \in \mathcal{F}$, avec $P(A) > 0$ et $P(B) > 0$. A et B sont indépendants si et seulement si $\mathbb{P}(A|B) = \mathbb{P}(A)$ ou $\mathbb{P}(B|A) = \mathbb{P}(B)$

Ce qui est le cas dans un tirage avec remise. Le tirage de la première carte n'influe pas sur le tirage de la 2ème.
Caractérisation : $A, B \in \mathcal{F}$ sont indépendants si et seulement si

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

Exemple : Soit une urne contenant 10 blanches, 20 rouges et 30 noires. On tire 2 boules sans remises. Probabilité que la première soit rouge (événement A) et la deuxième soit blanche (événement B) ?
On a $\mathbb{P}(A) = \frac{20}{60} = \frac{1}{3}$ et $\mathbb{P}(B|A) = \frac{10}{59}$

$$\mathbb{P}(A \cap B) = \mathbb{P}(B|A) \times \mathbb{P}(A) = \frac{10}{59} \times \frac{1}{3}.$$

même question avec des tirages avec remises :

$$\mathbb{P}(A \cap B) = P(A) \times P(B) = \frac{1}{3} \times \frac{1}{6} = \frac{1}{18}.$$

3.2 Formule des probabilités totales-Formule de Bayes

Exemple introductif.
Une entreprise utilise 3 machines différentes A, B, C pour fabriquer des pièces.40% sont fabriquées par A, 30% par B et 30% par C. La machine A produit 2% de pièces défectueuses, B 4% et C 5%.

1. On prélève une pièce au hasard. Probabilité qu?elle soit défectueuse.
2. On prélève une pièce. Elle est défectueuse. Probabilité qu?elle vienne de A.
3. On prélève une pièce. Elle est saine. Probabilité qu?elle vienne de C.

Soient A : "être fabriqué par A", B : "être fabriqué par B",.... D : "être défectueuse" et D : "saine". On a $P(A) = 0,4, P(B) = 0,3$, et $P(C) = 0,3$. A, B, C sont tels que $A \cap B = \emptyset$, $A \cap C = \emptyset$, $C \cap B = \emptyset$ et $A \cup B \cup C = \Omega$.
Solution: 1. En appliquant la formule des probabilités totales, on a $P(D) = P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C) = 0,02 \times 0,4 + 0,04 \times 0,3 + 0,05 \times 0,3$.
2. D'après le théorème de Bayes : $P(A|D) = \frac{P(D|A)P(A)}{P(D)}$
Un sac contient 5 boules rouges et 15 boules jaunes. Un deuxième sac contient 9 boules rouges.
Expérience ξ : On tire au hasard 1 boule du 1er sac et on la remet dans le second sac sans regarder la couleur. On tire alors une boule du 2ème sac. Quelle est la probabilité P que cette boule soit rouge?

ξ comporte 2 étapes. La deuxième étape dépend de la 1ère. Si on tire une rouge alors $P = \frac{10}{10} = 1$
Si on tire une jaune alors $P = \frac{9}{10}$.
Soit A = "boule Rouge au 2ème tirage". On cherche $\mathbb{P}(A)$.
B_1 : "la boule tirée du sac 1 est Rouge", B_2 : ""la boule tirée du sac 1 est Jaune". On a

$$B_1 \cup B_2 = \Omega, \quad B_1 \cap B_2 = \emptyset, \quad A = (A \cap B_1) \cup (A \cap B_2)$$

Ainsi

$$\mathbb{P}(A) = \mathbb{P}((A \cap B_1) \cup (A \cap B_2)) = \mathbb{P}(A \cap B_1) + (A \cap B_2) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) = 1 \times \frac{5}{20} + \frac{9}{10} \times \frac{15}{20}$$

Proposition 17.

Formule des probabilités totales
Soient $(B_k)_{k \leq n}$ une suite d'éléments de $\mathcal{F}$, telle que $\cup_{i=1}^n B_i = \Omega$ et $B_i \cap B_j = \emptyset$ pour $i \neq j$ et $\mathbb{P}(B_i) > 0$. Alors pour tout $A \in \mathcal{F}$, on a

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i) \times \mathbb{P}(B_i).$$

Proof. On a $A = A \cap \Omega = \cup_{i=1}^{n}(A \cap B_i)$, ainsi

$$\mathbb{P}(A) == \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{n} \mathbb{P}(A|B_i) \times \mathbb{P}(B_i)$$

□

Remarque : Si $(B_n)_{n\in\mathbb{N}^*} \in \mathcal{F}$ forment une partition dénombrable de Ω en événements avec $\mathbb{P}(B_n) > 0$ pour tout $n \in \mathbb{N}^*$, alors

$$\mathbb{P}(A) = \sum_{i=1}^{+\infty} \mathbb{P}(A|B_i) \times \mathbb{P}(B_i).$$

Exemple : Soit un sac contenant 6 jetons : 5 numérotés 1 et 1 numéroté 2. Soient 2 urnes : U_1 contenant six boules Blanches et quatre Noires et U_2 contenant 8 Blanches et 2 Noires. L'expérience aléatoire comporte 2 étapes. On pioche au hasard dans le sac, on regarde le numéro et on pioche dans l'urne correspondante. Quelle est la probabilité que la boule soit blanche (événement B)?
On a $U_1 \cup U_2 = \Omega$ et $U_1 \cap U2 = \emptyset$, on a donc

$$\mathbb{P}(B) = \mathbb{P}(B \cap U_1) + \mathbb{P}(B \cap U_2) = \mathbb{P}(B|U_1) \times \mathbb{P}(U_1) + \mathbb{P}(B|U_2) \times \mathbb{P}(U_2) = \frac{6}{10} \times \frac{5}{6} + \frac{8}{10} \times \frac{1}{6} = \frac{7}{15}.$$

Théorème 18.

Formule de Bayes
Soient $(B_k)_{k\leq n}$ une suite d'éléments de $\mathcal{F}$, telle que $\cup_{i=1}^{n} B_i = \Omega$ et $B_i \cap B_j = \emptyset$ pour $i \neq j$ et $\mathbb{P}(B_i) > 0$. Alors pour tout $A \in \mathcal{F}$, on a

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k) \times \mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i) \times \mathbb{P}(B_i)}.$$

Proof. On a $\mathbb{P}(B_k|A) = \dfrac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)}$ et on applique la formule des probabilités totales. □

Remarque : Si $(B_n)_{n\in\mathbb{N}*} \in \mathcal{F}$ forment une partition dénombrable de Ω en événements avec $\mathbb{P}(B_n) > 0$ pour tout $n \in \mathbb{N}^*$, alors

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k) \times \mathbb{P}(B_k)}{\sum_{i=1}^{+\infty} \mathbb{P}(A|B_i) \times \mathbb{P}(B_i)}.$$

Exemple : Une entreprise utilise 3 machines différentes A, B ,C pour fabriquer des pièces. 40% sont fabriquées par A, 30 % par B et 30 % par C. La machine A produit 2 % de pièces défectueuses, B 4 % et C 5 %.

1. On prélève une pièce au hasard. Probabilité qu'elle soit défectueuse.
2. On prélève une pièce. Elle est défectueuse. Probabilité qu'elle vienne de A.
3. On prélève une pièe. Elle est saine. Probabilité qu'elle vienne de C.

Soient les événements: A : "être fabriqué par A", B : "être fabriqué par B", C : "être fabriqué par C", D : "être défectueuse" et D^c : "être saine".
On a $\mathbb{P}(A) = 0,4, \mathbb{P}(B) = 0,3$, et $\mathbb{P}(C) = 0,3$.
A, B, C sont tels que $A \cap B = \emptyset$, $A \cap C = \emptyset$, $C \cap B = \emptyset$ et $A \cup B \cup C = \Omega$.
1. En appliquant la formule des probabilités totales, on a

$$\mathbb{P}(D) = \mathbb{P}(D|A) \times \mathbb{P}(A) + \mathbb{P}(D|B) \times \mathbb{P}(B) + \mathbb{P}(D|C) \times \mathbb{P}(C) = 0,02 \times 0,4 + 0,04 \times 0,3 + 0,05 \times 0,3.$$

2. D'après le théorème de Bayes, on a

$$\mathbb{P}(A|D) = \frac{\mathbb{P}(D|A) \times \mathbb{P}(A)}{\mathbb{P}(D)}.$$

3. On a
$$\mathbb{P}(C|D^c) = \frac{\mathbb{P}(D^c|C) \times \mathbb{P}(C)}{\mathbb{P}(D^c)} = \frac{\mathbb{P}(D^c|C) \times \mathbb{P}(C)}{1 - \mathbb{P}(D)}.$$

Chapitre 4

Variables aléatoires

4.1 Variables aléatoires discrètes

4.1.1 Exemple introductif

Soient 2 joueurs A et B. L'un des deux lance un dé:
si 1 ou 6, A donne à B 1dh
si 2,3 ou 5, B donne à A 2dhs
si 4, partie nulle.
Appelons X le gain de A.
X dépend du hasard, plus particulièrement du résultat du lancer de dé. On dira que X est une variable aléatoire car elle dépend du hasard.
Ici l'expérience ξ = lancer de dé. L'univers $\Omega = \{1,2,3,4,5,6\}$ ainsi X dépend des événements de Ω et peut prendre les valeurs $-1, 0, 2$. $X(1) = -1$, $X(6) = -1$, $X(2) = X(3) = X(5) = 2$, $X(4) = 0$.

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

Le dé étant non truqué, les événements élémentaires de Ω sont équiprobables. Ainsi $(\Omega, \mathbb{P})$ est un espace probabilisé $\mathbb{P}(\{i\}) = \frac{1}{6}$, $i = 1,2,3,4,5,6$.
On demande, par exemple, la probabilité que A a de gagner 2dhs. On a

$$\mathbb{P}(X=2) = \mathbb{P}(\omega \in \Omega : X(\omega) = 2) = \mathbb{P}(X^{-1}(2)) = \mathbb{P}(\{2,3,5\}) = \frac{3}{6} = \frac{1}{2}.$$

De plus, $\mathbb{P}(X=-1) = \frac{1}{3}$ et $\mathbb{P}(X=0) = \frac{1}{6}$.
Ainsi à chaque valeur de X, on peut associer une probabilité. Cette correspondance s'appelle loi de probabilité de X représentée par le tableau suivant:

Valeurs de X	-1	0	2
Probabilité	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

Définition 19.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. On appelle variable aléatoire discrète sur $(\Omega, \mathcal{F}, \mathbb{P})$ toute application

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

telle que l'ensemble de ses images $X(\Omega) = \{X(\omega) : \omega \in \Omega)\}$ est une partie au plus dénombrable de $\mathbb{R}$, i.e. fini ou dénombrable. On peut donc numéroter ses éléments par des indices entiers : $X(\Omega) = \{x_1, x_2, ..., x_k; ...\}$.

L'événement $A_k = \{\omega \in \Omega : X(\omega) = x_k\}$ est aussi noté $X^{-1}(\{x_k\})$ (où X^{-1} désigne l'inverse ensembliste) ou encore $\{X = x_k\}$. Dans ce cas on a

$$\sum_{k, x_k \in X(\Omega)} \mathbb{P}(A_k) = \sum_{k, x_k \in X(\Omega)} \mathbb{P}(X = x_k) = \mathbb{P}(X \in \{x_1, x_2, ..., x_k; ...\}) = \mathbb{P}(\Omega) = 1.$$

4.1.2 Loi d'une variable aléatoire discrète

Définition

Définition 20.

Soit X une v.a.d. sur $(\Omega;\mathcal{F};\mathbb{P})$. On lui associe sa loi P_X qui est la probabilité définie sur les parties de $\mathbb{R}$ par

$$\mathbb{P}_X(\{x_k\}) = \mathbb{P}(X = x_k) = \mathbb{P}(\{\omega \in \Omega : X(\omega) = x_k\})$$

Ainsi pour tout $B \subset \mathbb{R}$:

$$\mathbb{P}_X(B) = \sum_{k, x_k \in B} \mathbb{P}(X = x_k).$$

Exemple : Dans l'exemple introductif, on a

Valeurs x_k de X	-1	0	2
Loi de X :$\mathbb{P}(X = x_k)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

Exemples des lois des variables aléatoires discrètes

1. Loi de Bernoulli : Une v.a.r. X est dite variable aléatoire de Bernoulli de paramètre p, (pour $p \in [0,1]$) si elle est à valeurs dans $\{0,1\}$ et si

$$\mathbb{P}_X(\{1\}) = \mathbb{P}(X = 1) = p, \mathbb{P}_X(\{0\}) = \mathbb{P}(X = 0) = 1 - p.$$

   Toute épreuve n'ayant que 2 résultats possibles peut-être considérée comme une situation d'alternative. En d'autres termes, les 2 résultats possibles (le succès S et l'Echec E) sont complémentaires l'un de l'autre.

2. Loi binomiale Une v.a.r. X est dite de loi Binomiale de paramètres n et p (pour $n \in \mathbb{N}^*$ et $p \in [0,1]$) si elle est à valeurs dans $\{0,1,...,n\}$ et si

$$\mathbb{P}(X = k) = \mathbb{P}_X(\{k\}) = C_n^k p^k (1-p)^{n-k},$$

   pour $k = 0,1,...,n$. On écrit $X \sim B(n,p)$.
   C'est une succession d'épreuves de Bernouilli indépendantes.
   Si on réalise n fois successivement et d'une manière indépendante une expérience aléatoire qui a 2 résultats possibles, S (succès) de probabilité p et E (Echec) de probabilité $1-p$, X = "nombre de succès obtenus" est une v.a. binomiale de paramèètre n et p.

3. Loi géométrique : Une v.a.r. X est dite de loi géométrique de paramètre p, pour p compris entre 0 et 1, si elle est à valeurs dans N^* et si

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

   On note $X \sim G(p)$. A la base, il y a toujours l'épreuve de Bernouilli qui a 2 résultats possibles. Mais cette fois, on ne connait pas le nombre d'épreuves. X suit la loi géométrique de paramètre p, si elle est égale au nombre d'épreuves de Bernouilli indépendantes qu'il faut réaliser pour obtenir pour la 1ère fois l'évènement de probabilité p.

4. Loi de Poisson : Une v.a.r. X est dite de loi de Poisson de paramètre λ, si elle est à valeurs dans $\mathbb{N}$ et si

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

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

   Soit X une variable binomiale $B(n,p)$. On va supposer que $p \approx 0$ et que $n \to +\infty$ de manière à ce que $np \to \lambda$ Quel modèle va suivre cette variable aléatoire ?

4.1.3 Espérance d'une variable aléatoire

Définition 21.

Soit X une v.a. discrète prenant une infinité de valeurs $\{x_1, x_2, ... x_n, ...\}$ si

$$\sum_{k=1}^{\infty} |x_k| \mathbb{P}(X = x_k) < +\infty,$$

on définit l'espérance de X par

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} x_k \mathbb{P}(X = x_k).$$

Exemple : Dans l'exemple introductif, on a

Valeurs x_k de X	-1	0	2
Loi de X :$\mathbb{P}(X = x_k)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

Donc

$$\mathbb{E}(X) = \sum_{k=1}^{3} x_k \mathbb{P}(X = x_k) = -1 \times \frac{1}{3} + 0 \times \frac{1}{6} + 2 \times \frac{1}{2} = \frac{2}{3}.$$

Exemple 1. *Soit $X \sim B(n, p)$, dont la loi est donnée par :*

$$\mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k},$$

($p \in [0, 1]$ fixé). Alors

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{n} k \mathbb{P}(X = k) \\
&= \sum_{k=0}^{n} k C_n^k p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n(n-1)!}{(k-1)!(n-1-(k-1))!} p^k (1-p)^{n-k} \\
&= n \sum_{q=0}^{n-1} C_{n-1}^q p^{q+1} (1-p)^{n-1-q} \\
&= np(p+1-p)^{n-1} = np\,.
\end{aligned}$$

Exemple 2. *Soit $X \sim \mathcal{P}(\lambda)$, $\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$. Alors*

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{+\infty} \frac{k \lambda^k e^{-\lambda}}{k!} \\
&= \sum_{k=1}^{+\infty} \frac{1}{(k-1)!} \lambda^k e^{-\lambda} \\
&= \sum_{q=0}^{+\infty} \frac{\lambda^{q+1} e^{-\lambda}}{q!} = \lambda.
\end{aligned}$$

Remarque 1. *1. Si X est constante ($\exists c \in \mathbb{R}$, $\forall \omega \in \Omega X(\omega) = c$), son espérance est $\mathbb{E}(X) = c$.*
2. (Propriétés de l'espérance mathématique)
Si X et Y sont deux v.a.d. admettant des espérances alors on a les propriétés suivantes:

1. *Pour tout réel a, $\mathbb{E}(aX) = a\mathbb{E}(X)$.*

2. *Linéarité:*
$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

3. *Croissance : si X et Y sont telles que $X \le Y$ (c'est à dire $\forall \omega \in \Omega, X(\omega) \le Y(\omega)$ alors $\mathbb{E}(X) \le \mathbb{E}(Y)$.*

Le premier point est clair : en effet si $X(\Omega) = \{x_1, ..., x_k, ...\}$ alors $aX(\Omega) = \{ax_1, ..., ax_k, ...\}$ et comme $\mathbb{P}(aX = ax_k) = \mathbb{P}(X = x_k)$, on a d'abord

$$\sum_{k=1}^{\infty} |ax_k| \mathbb{P}(aX = ax_k) = |a| \sum_{k=1}^{\infty} |x_k| \mathbb{P}(X = x_k) < +\infty.$$

par hypothèse de l'existence de $\mathbb{E}[X]$. Puis la même chose sans les valeurs absolues donne :

$$\mathbb{E}(aX) = \sum_{k=1}^{\infty} ax_k \mathbb{P}(aX = ax_k) = a \sum_{k=1}^{\infty} x_k \mathbb{P}(X = x_k) = a\mathbb{E}(X).$$

Pour le deuxième point : posons $Z = X + Y$, supposons que X et Y prennent un nombre fini de valeurs $\{x_1, ..., x_n\}$ et $\{y_1, ..., y_p\}$ alors Z prend aussi un nombre fini de valeurs $\{z_k = x_i + y_j; i = 1, ..., n; j = 1, ..., p\}$ et la loi de Z est donnée par

$$\mathbb{P}(Z = z_k) = \sum_{(i,j): x_i + y_j = z_k} \mathbb{P}(X = x_i, Y = y_j).$$

On a alors

$$\begin{aligned}
\mathbb{E}(Z) &= \sum_k z_k \mathbb{P}(Z = z_k) \\
&= \sum_k z_k \sum_{(i,j): x_i + y_j = z_k} \mathbb{P}(X = x_i, Y = y_j) \\
&= \sum_k \sum_{(i,j): x_i + y_j = z_k} (x_i + y_j) \mathbb{P}(X = x_i, Y = y_j)
\end{aligned}$$

Définition 22.

Soit $r \in \mathbb{N}^*$, on appelle moment d'ordre r de la v.a. discrète X le nombre

$$\mathbb{E}|X|^r = \sum_{k=1}^{\infty} |x_k|^r \mathbb{P}(X = x_k).$$

On peut alors définir aussi

$$\mathbb{E}(X^r) = \sum_{k=1}^{\infty} (x_k)^r \mathbb{P}(X = x_k).$$

où $X(\Omega) = \{x_1, ..., x_k, ...\}$

Théorème 23.

(Théorème de transfert) Soient Soit X une v.a. discrète prenant une infinité de valeurs $\{x_1, x_2, ... x_n, ...\}$ et F une fonction numérique sur $\mathbb{R}$ (ou dont l'ensemble de définition contient au moins l'ensemble des valeurs $X(\Omega)$ de X). Alors si $\mathbb{E}|F(X)|$ existe ($\sum_{k=1}^{\infty} |F(x_k)|\mathbb{P}(X = x_k) < \infty$), on a

$$\mathbb{E}[F(X) = \sum_{k=1}^{\infty} F(x_k)\mathbb{P}(X = x_k).$$

4.1.4 Variance d'une variable aléatoire

Définition 24.

Soit X une v.a. de domaine $X(\Omega) = \{x_1, ..., x_k, ...\}$ et avec un moment d'ordre 2. On appelle respectivement variance de X et écart-type de X les quantités

$$\begin{aligned} Var(X) &= \mathbb{E}\Big[(X - \mathbb{E}(X)^2\Big] = \sum_{k=1}^{\infty} (x_k - \mathbb{E}(X))^2 \mathbb{P}(X = x_k) \\ &= \sigma_X = \sqrt{Var(X)}. \end{aligned}$$

Proposition 25.

(Translation et changement d'échelle)
Si X a un moment d'ordre 2 alors

$$Var(aX) = a^2 Var(X), \quad Var(X + b) = Var(X).$$

Proof. On a

$$Var(aX) = \mathbb{E}[(aX - \mathbb{E}[aX])^2] = \mathbb{E}[a^2(X - \mathbb{E}[X])^2] = a^2\mathbb{E}[(X - E[X])^2] = a^2 Var(X).$$

et

$$Var(X + b) = \mathbb{E}[(X + b - \mathbb{E}[X + b])^2] = \mathbb{E}[(X + b - \mathbb{E}[X] - b)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2] = Var(X).$$

□

Proposition 26.

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

4.1.5 Fonction de répartition

Définition 27.

Soit X une v.a. discrète prenant une infinité de valeurs $\{x_1, x_2, ... x_n, ...\}$ telle que $x_1 < x_2 < .... < x_n$. On appelle fonction de répartition de la v.a.r. X, la fonction F_X définie sur $\mathbb{R}$ par

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

On a aussi pour une v.a. X discrète

$$F_X(x) = \sum_{\substack{k, x_k \in X(\Omega) \\ x_k \leq x}} \mathbb{P}(X = x_k).$$

Théorème 28.

La fonction de répartition F_X d'une v.a.r. X satisfait les propriétés suivantes :

1. $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.
2. $0 \leq F_X(x) \leq 1$ pour tout x dans $\mathbb{R}$,
3. La fonction F_X est croissante,
4. la fonction F_X est constante sur chaque intervalle $[x_k, x_{k+1}[$ (avec un saut $p_{k+1} = F_X(x_{k+1}) - F_X(x_k) = \mathbb{P}(X = x_{k+1})$).
5. On a $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$.
6. La fonction F_X est continue à droite.

Proof. Pour la propriété 1, on a

$$\mathbb{P}(a < X \leq b) = 1 - \mathbb{P}(\{X \leq a\} \cup \{X > b\}) = 1 - (\mathbb{P}(\{X \leq a\}) + \mathbb{P}(\{X > b\})) = \mathbb{P}(\{X \leq b\} - \mathbb{P}(\{X \leq a\} = F_X(b) - F_X(a).$$

La propriété 2 est évidente puisque la probabilité de n'importe quel événement est toujours positive et inférieure à 1.
3. On a si $x \leq y$

$$F_X(y) - F_X(x) = \mathbb{P}(x < X \leq y) = \sum_{\substack{k, x_k \in X(\Omega) \\ x < x_k \leq y}} \mathbb{P}(X = x_k) \geq 0.$$

4. Si $x \leq y$ et $x, y \in [x_k, x_{k+1}[$, montrons que $F_X(y) = F_X(x)$. En effet

$$F_X(y) - F_X(x) = \mathbb{P}(x < X \leq y) = \sum_{\substack{i, x_i \in X(\Omega) \\ x_k \leq x < x_i \leq y < x_{k+1}}} \mathbb{P}(X = x_i) = 0.$$

car $\{(X = x_i\} = \emptyset$ si $x_k < x_i < x_{k+1}$. Si $y = x_{k+1}$ et $x = x_k$ alors

$$p_{k+1} = F_X(x_{k+1}) - F_X(x_k = \sum_{\substack{i, x_i \in X(\Omega) \\ x_k < x_i \leq x_{k+1}}} \mathbb{P}(X = x_i) = \mathbb{P}(X = x_{k+1}).$$

5. On a

$$\lim_{x \to -\infty} F_X(x) = \lim_{x \to -\infty} \sum_{\substack{k, x_k \in X(\Omega) \\ x_k \leq x}} \mathbb{P}(X = x_k) = 0$$

car pour $x \leq \inf_k(x_k)$, la somme est vide donc nulle.

$$\lim_{x\to+\infty} F_X(x) = \lim_{x\to+\infty} \sum_{\substack{k,x_k\in X(\Omega)\\ x_k\leq x}} \mathbb{P}(X = x_k) = 1$$

car pour $x \geq \sup_k(x_k)$, la somme devient

$$\sum_{\substack{k,x_k\in X(\Omega)\\ x_k\leq x}} \mathbb{P}(X = x_k) = \sum_{k,x_k\in X(\Omega)} \mathbb{P}(X = x_k) = 1.$$

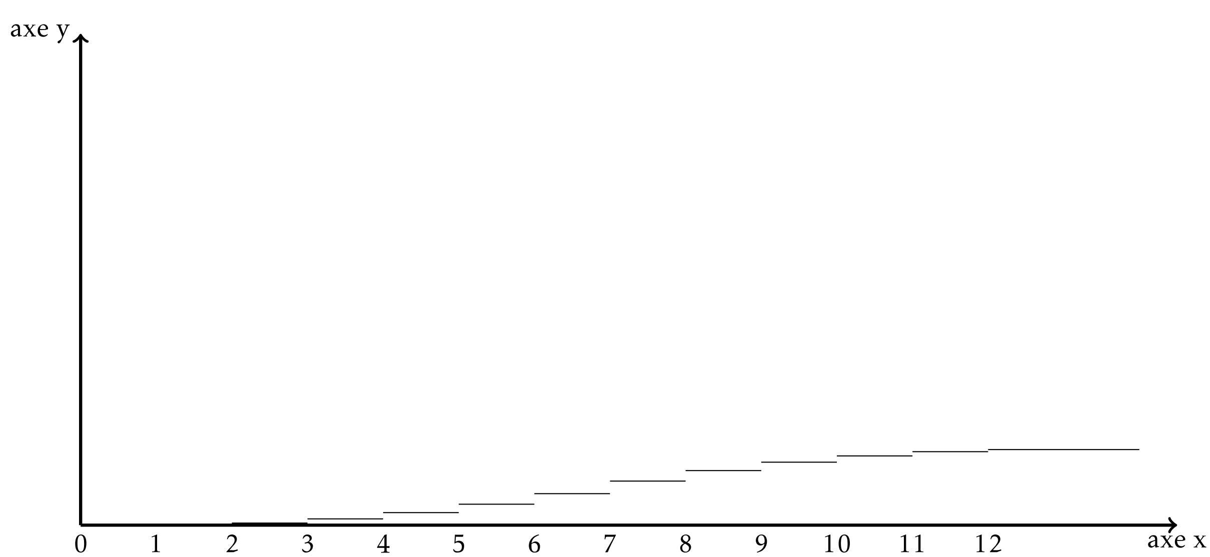
$\square$

Exemple : Soit X la variable aléatoire qui donne la somme des faces obtenues en lançant deux fois un dé à six faces bien équilibré. La loi de X est donnée par l'ensemble des valeurs possibles $S(X) = \{2;3;4;5;6;7;8;9;10;11;12\}$ et les probabilités associées

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Valeurs x_k de X	2	3	4	5	6	7	8	9	10	11	12
Loi de X :$\mathbb{P}(X = x_k)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

La fonction de répartition est alors donnée par :



4.2 Variables aléatoires continues

Définition 29.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. On appelle variable aléatoire réelle (v.a.r.) sur $(\Omega, \mathcal{F}, \mathbb{P})$ toute application

$$X: \quad \Omega \longrightarrow \mathbb{R}$$
$$\omega \mapsto X(\omega).$$

telle que pour tout intervalle I de $\mathbb{R}$, l'ensemble $X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\}$ fait partie de la famille $\mathcal{F}$, i.e $X^{-1}(I) \in \mathcal{F}$.

4.2.1 Loi d'une variable aléatoire continue

A nouveau, à chaque v.a.r., on associe sa loi. Elle définit une probabilité sur $\mathbb{R}$.

Définition

Définition 30.

Soit X une v.a.r. sur $(\Omega; \mathcal{F}; \mathbb{P})$. On lui associe sa loi P_X en posant: pour tout intervalle I de $\mathbb{R}$

$$\mathbb{P}_X(I) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in I\}) = \mathbb{P}(X^{-1}(I))$$

Il est facile de vérifier qu'il s'agit bien d'une probabilité sur $\mathbb{R}$: en effet, d'abord

$$\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(\Omega) = 1.$$

Puis si $(I_n)_{n\in\mathbb{N}}$ est une famille d'intervalles de $\mathbb{R}$ deux à deux disjoints, on a :

$$X^{-1}(\cup_{n=1}^{\infty} I_n) = \{\omega \in \Omega : X(\omega) \in \cup_{n=1}^{\infty} I_n\} = \cup_{n=1}^{\infty}\{\omega \in \Omega : X(\omega) \in I_n\} = \cup_{n=1}^{\infty} X^{-1}(I_n).$$

Comme les événements $(\{X \in I_n\})_{n\in\mathbb{N}}$, sont deux à deux disjoints, il suit par σ-additivité de $\mathbb{P}$:

$$\mathbb{P}_X(\cup_{n=1}^{\infty} I_n) = \mathbb{P}(X^{-1}(\cup_{n=1}^{\infty} I_n)) = \mathbb{P}(\cup_{n=1}^{\infty} X^{-1}(I_n)) = \sum_{n=1}^{\infty} \mathbb{P}(X^{-1}(I_n)) = \sum_{n=1}^{\infty} \mathbb{P}_X(I_n).$$

4.2.2 Fonction de répartition

Définition 31.

On appelle fonction de répartition de la v.a.r. X, la fonction F_X définie sur $\mathbb{R}$ par

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

La fonction de répartition F_X d'une v.a.r. X jouit des mêmes propriétés que celles des v.a. discrètes qu'on énonce de la même façon.

Théorème 32.

La fonction de répartition F_X d'une v.a.r. X satisfait les propriétés suivantes :

1. $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.
2. $0 \leq F_X(x) \leq 1$ pour tout x dans $\mathbb{R}$,
3. La fonction F_X est croissante,
4. la fonction F_X est continue à droite,
5. On a $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$.

F_X vérifiant les propriétés 2-4, est appelée la fonction de répartition de la v.a.r. X.

Proof. Pour la propriété 1, on a

$$\mathbb{P}(a < X \leq b) = 1 - \mathbb{P}(\{X \leq a\} \cup \{X > b\}) = 1 - (\mathbb{P}(\{X \leq a\}) + \mathbb{P}(\{X > b\})) = \mathbb{P}(\{X \leq b\} - \mathbb{P}(\{X \leq a\} = F_X(b) - F_X(a).$$

La propriété 2 est évidente puisque la probabilité de n'importe quel événement est toujours positive et inférieure à 1.
Pour établir 3, considérons x et x' deux réels tels que $x \leq x'$. On a bien sûr l'inclusion

$$]-\infty, x] \subset]-\infty, x'],$$

et alors

$$F_X(x) = \mathbb{P}_X(]-\infty, x]) \leq \mathbb{P}_X(]-\infty, x']) = F_X(x').$$

Pour prouver 4, considérons une suite $(h_n)_n$ de réels décroissante vers 0. Pour tout x dans $\mathbb{R}$, on a :

$$\mathbb{P}_X(]x, x + h_n]) = F_X(x + h_n) - F_X(x).$$

Or la suite d'intervalles $(]x, x + h_n])_n$ est décroissante avec n. Ainsi il vient :

$$\lim_{n \to +\infty} \mathbb{P}_X(]x, x + h_n]) = \mathbb{P}_X(\bigcap_{n=1}^{+\infty}]x, x + h_n]) = \mathbb{P}_X(\emptyset) = 0.$$

F est donc continue à droite.
Pour 5, considérons la suite d'intervalles $(]-\infty, -n])_n$ décroissante vers $\emptyset$ quand n tend vers $+\infty$. On a :

$$\lim_{x \to -\infty} F_X(x) = \lim_{n \to +\infty} F_X(-n) = \mathbb{P}_X(\bigcap_{n=1}^{+\infty}]-\infty, -n]) = \mathbb{P}_X(\emptyset) = 0.$$

considérons la suite d'intervalles $(]-\infty, n])_n$ croissante vers $\mathbb{R}$ quand n tend vers $+\infty$. On a :

$$\lim_{x \to +\infty} F_X(x) = \lim_{n \to +\infty} F_X(n) = \mathbb{P}_X(\bigcup_{n=1}^{+\infty}]-\infty, n]) = \mathbb{P}_X(\mathbb{R}) = 1.$$

□

4.2.3 Variables aléatoires réelles à densité

La loi d'une v.a.r. est à densité f si pour tout intervalle de $\mathbb{R}$, la probabilité d'appartenance de X à cet intervalle s'exprime comme l'intégrale de f sur cet intervalle.

Définition 33.

Une fonction $f : \mathbb{R} \to \mathbb{R}$ est appelée densité de probabilité si

1. f est positive : en tout point t où elle est définie $f(t) \geq 0$,
2. f est intégrable sur $\mathbb{R}$ d'intégrale 1 :

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

Définition 34.

La v.a.r. X suit la loi de densité f si

$$\forall [a;b] \subset \mathbb{R}; \quad \mathbb{P}(X \in [a;b]) = \int_a^b f(t)dt$$

Notons que pour une v.a. X de densité f, la probabilité que X vaille un point est 0, car c'est une intégrale sur un intervalle réduit à un point :

$$\mathbb{P}(X = a) = \mathbb{P}(X \in [a;a]) = \int_a^a f(t)dt = 0.$$

Par conséquent

$$\mathbb{P}(X \in [a;b]) = \mathbb{P}(X \in]a;b]) = \mathbb{P}(X \in]a;b[) = \mathbb{P}(X \in [a;b[), \quad \mathbb{P}(X \leq x) = \mathbb{P}(X < x).$$

Lorsqu'elle existe la densité f est naturellement reliée à la fonction de répartition F_X.

Proposition 35.

Si X est une v.a.r. de densité f, sa fonction de répartition F_X vérifie

1. $\forall x \in \mathbb{R}$, $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)dt$.
2. F_X est continue sur $\mathbb{R}$.
3. Si f est continue au point x_0, alors F_X est dérivable en x_0 de dérivée $F'_X(x_0) = f(x_0)$.

Proof. 1. Soit $x \in \mathbb{R}$, et soit $A = \{\omega \in \Omega : X(\omega) \in]-\infty, x]\}$ et $A_n = \{\omega \in \Omega : X(\omega) \in]-n, x]\}$

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \int_{-n}^{x} f(t)dt = \int_{-\infty}^{x} f(t)dt$$

2. F_X est continue à droite, il suffit de montrer qu'elle est continue à gauche. En effet, considérons une suite $(h_n)_n$ de réels décroissante vers 0. Pour tout x dans $\mathbb{R}$, on a la suite d'intervalles $(]-\infty, x-h_n])_n$ est croissante avec n vers $]-\infty, x[$. Ainsi il vient :

$$\lim_{n\to+\infty} F_X(x-h_n) = \lim_{n\to+\infty} \mathbb{P}_X(]-\infty, x-h_n]) = \mathbb{P}_X(\bigcup_{n=1}^{+\infty}]-\infty, x-h_n]) = \mathbb{P}_X(X < x) = F_X(x).$$

F est donc continue à gauche.

3. Comme par hypothèse f est continue en x_0, elle est définie sur tout un voisinage de x_0. Par continuité de f en x_0, on a:
$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in]x_0 - \delta, x_0 + \delta[: |f(x) - f(x_0)| < \varepsilon.$$
Pour tout h tel que $0 < |h| < \delta$, on a alors
$$F_X(x_0 + h) - F_X(x_0) = \int_{x_0}^{x_0+h} f(t)dt.$$
D'où
$$|F_X(x_0 + h) - F_X(x_0) - hf(x_0)| = |\int_{x_0}^{x_0+h} (f(t) - f(x_0))dt| \leq |\int_{x_0}^{x_0+h} |f(t) - f(x_0)|dt| \leq |h|\varepsilon.$$
En divisant par h puis en faisant h vers 0, on constate que F_X est dérivable en x_0, de dérivée $F'_X(x_0) = f(x_0)$.

□

4.2.4 Espérance d'une variable aléatoire continue à densité

Définition 36.

Si X est une v.a.r. de densité f telle que l'intégrale $\int_{-\infty}^{+\infty} |x|f(x)dx$ converge, on appelle espérance de X le réel (fini)
$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

Remarque 2. *(Propriétés de l'espérance mathématique)*
Si X et Y sont deux v.a.c. admettant des espérances alors on a les propriétés suivantes:

1. *Pour tout réel a,* $\mathbb{E}(aX) = a\mathbb{E}(X)$.
2. *Linéarité:*
$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$
3. *Croissance : si X et Y sont telles que* $X \leq Y$ *(c'est à dire* $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$ *alors* $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

Proof. On prouve seulement le 1. pour $a > 0$:
$$\mathbb{E}(aX) = \int_{-\infty}^{+\infty} axf(x)dx = a\int_{-\infty}^{+\infty} xf(x)dx = a\mathbb{E}(X).$$
3. Si $X \geq 0$, alors $\mathbb{P}(X \leq 0) = \int_{-\infty}^{0} f(t)dt = 0$. Comme f est positive alors $f = 0$ pour $t < 0$. Donc
$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_{0}^{+\infty} xf(x)dx \geq 0.$$
maintenant on applique cette propriété à la v.a. positive $Y - X$ puis on utilise la linéarité de l'espérance.

□

Théorème 37.

(Théorème de transfert) Soient Soit X une v.a.c. à valeurs dans $\mathbb{R}$ et F une fonction continue numérique sur $\mathbb{R}$. Alors si $\mathbb{E}[F(X)]$ existe, on a
$$\mathbb{E}[F(X)] = \int_{\mathbb{R}} F(x)f(x)dx.$$

Définition 38.

Soit X une v.a. c. avec un moment d'ordre 2 i.e. $\mathbb{E}(X^2) < \infty$. On appelle respectivement variance de X et écart-type de X les quantités

$$\begin{aligned} Var(X) &= \mathbb{E}\left[(X - \mathbb{E}(X)^2\right] \\ &= \sigma_X = \sqrt{Var(X)}. \end{aligned}$$

Proposition 39.

(Translation et changement d'échelle)
Si X a un moment d'ordre 2 alors

$$Var(aX) = a^2 Var(X), \quad Var(X+b) = Var(X).$$

Proof. On a

$$Var(aX) = \mathbb{E}[(aX - \mathbb{E}[aX])^2] = \mathbb{E}[a^2(X - \mathbb{E}[X])^2] = a^2\mathbb{E}[(X - E[X])^2] = a^2 Var(X).$$

et

$$Var(X+b) = \mathbb{E}[(X + b - \mathbb{E}[X+b])^2] = \mathbb{E}[(X + b - \mathbb{E}[X] - b)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2] = Var(X).$$

□

Proposition 40.

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

Théorème 41.

(Inégalité de Markov)
Si X est une v.a.r. positive à densité ayant une espérance alors pour tout $a > 0$

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Proof. Si on note f la densité de la v.a. X, on a $f(x) = 0$ si $x < 0$ car X est à valeurs positives. Puis

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} xf(x)dx = \int_0^a xf(x)dx + \int_a^{+\infty} xf(x)dx \\ &\geq \int_a^{+\infty} xf(x)dx \geq a\int_a^{+\infty} f(x)dx \\ &\geq a(1 - \int_{-\infty}^a f(x)dx) = a(1 - \mathbb{P}(X \leq a)) = a\mathbb{P}(X \geq a). \end{aligned}$$

□

4.3 Lois à densité classiques

La fonction indicatrice d'un ensemble A est

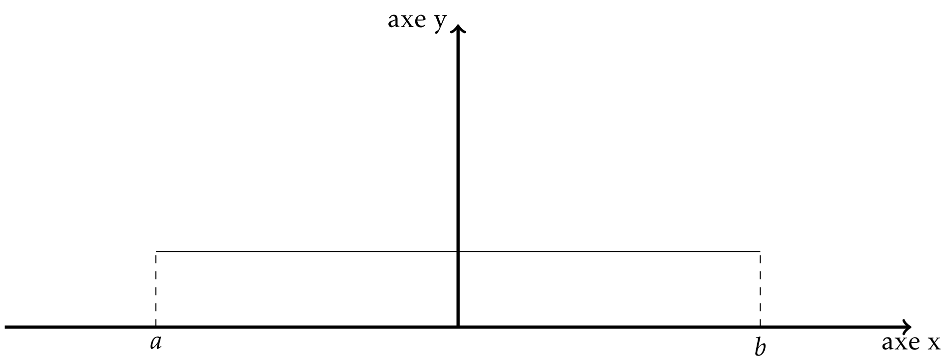
$$1_A(x) = \begin{cases} 1 & \text{si} \quad x \in A \\ 0 & \text{si} \quad x \notin A \end{cases}$$

4.3.1 Loi uniforme

Définition 42.

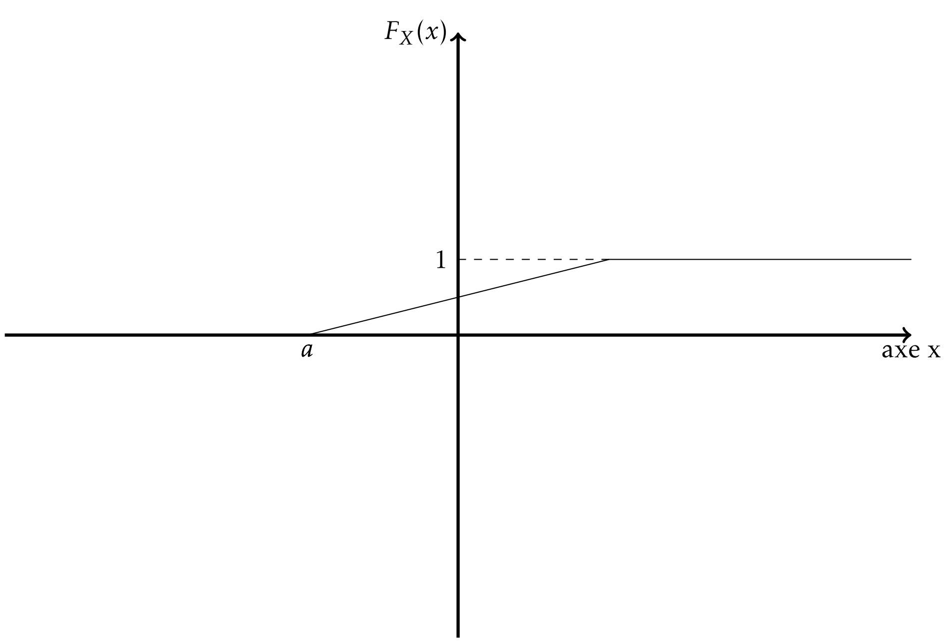
La v.a.r. X suit une loi uniforme sur l'intervalle $[a;b]$ $(a<b)$ si elle a une densité f donné par

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



Sa fonction de répartition est affine par morceaux :

$$F_X(x)=\int_{-\infty}^{x} f(t)dt=\begin{cases} 0 \quad \text{si} \quad x\leq a \\ \dfrac{x-a}{b-a} \quad \text{si} \quad x\in[a,b] \\ 1 \quad \text{si} \quad x\geq b \end{cases}$$



4.3.2 Loi exponentielle

Définition 43.

Soit a un réel strictement positif. La v.a.r. X suit une loi exponentielle de paramètre λ si elle admet pour densité:

$$f(x) = \lambda e^{-\lambda x} 1_{[0,+\infty[}(x).$$

On note $X \backsim \mathcal{E}(\lambda)$.

Sa fonction de répartition est donnée par :

$$F_X(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0 \quad \text{si} \quad x \leq 0 \\ 1 - e^{-\lambda x} \quad \text{si} \quad x \geq 0 \end{cases}$$

En pratique, à la place de la fonction de répartition, on utilise souvent la fonction de survie G d'une v.a.r. de loi exponentielle

$$G(x) = 1 - F_X(x) = \begin{cases} 1 \quad \text{si} \quad x \leq 0 \\ e^{-\lambda x} \quad \text{si} \quad x \geq 0 \end{cases}$$

Les lois exponentielles sont souvent utilisées pour modéliser des temps d'attente ou des durées de vie. Par exemple, les temps d'attente à partir de maintenant du prochain tremblement de terre, de la prochaine panne d'un appareil, de la prochaine désintégration dans un réacteur nucléaire suivent des lois exponentielles.
Une propriété intéressante de ce type de loi est l'absence de mémoire. Cette propriété caractérise les lois exponentielles.

Théorème 44.

Si la v.a.r. X suit une loi exponentielle alors elle vérifie la propriété d'absence de mémoire : pour tous $s, t \geq 0$

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s).$$

Autrement dit si X survit jusqu'en t, sa survie pendant encore s unités de temps est la même qu'une survie de durée s depuis le départ : tout se passe comme si, ce qui se passe de 0 à t est oublié pour survivre encore s unités de temps. C'est à comparer, par exemple, avec la vie humaine qui a une mémoire : pour un homme de 60 ans, la probabilité de vivre encore 30 ans n'est pas la même que pour celle d'un nouveau né (de 0 an).

Proof. On a

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s) = \frac{\mathbb{P}(\{X > s + t\} \cap \{X > t\})}{\mathbb{P}(X > s)} = \frac{\mathbb{P}(\{X > s + t\})}{\mathbb{P}(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbb{P}(X > s).$$

□

4.3.3 Loi de Cauchy

Définition 45.

Une variable aléatoire réelle suit une loi de Cauchy de paramètre $a > 0$ si elle admet pour densité :

$$f(x) = \frac{1}{\pi} \frac{1}{a^2 + x^2}.$$

On note $X \backsim \mathcal{C}(a)$.

4.3.4 Loi normale ou gaussienne

Définition 46.

On dit que la v.a.r. X suit une loi gaussienne ou normale $\mathcal{N}(m,\sigma^2)$ si elle a pour densité la fonction : pour tout $x \in \mathbb{R}$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x-m)^2}{2\sigma^2}}$$

La loi normale standard $\mathcal{N}(0,1)$ est celle de densité

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\dfrac{x^2}{2}}.$$

Remarque : Cette loi est fondamentale en théorie des probabilités et en statistique : c'est la loi limite de la moyenne dans une suite infinie d'épreuves répétées indépendantes.

Proposition 47.

Si la v.a.r. X suit une loi $\mathcal{N}(m,\sigma^2)$, alors $Y := \dfrac{X-m}{\sigma}$ suit la loi $\mathcal{N}(0,1)$.
La v.a. Y s'appelle la v.a. centrée réduite associée à X.

Proof. Calculons pour $a < b$ quelconques $\mathbb{P}(a \le Y \le b)$:

$$\mathbb{P}(a \le Y \le b) = \mathbb{P}(a \le \frac{X-m}{\sigma} \le b) = \mathbb{P}(\sigma a + m \le X \le \sigma b + m) = \int_{\sigma a+m}^{\sigma b+m} \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x-m)^2}{2\sigma^2}} dx$$

Il suffit alors de faire le changement de variable $s = \dfrac{t-m}{\sigma}$ pour obtenir

$$\mathbb{P}(a \le Y \le b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\dfrac{x^2}{2}} dx.$$

Donc Y suit la loi $\mathcal{N}(0,1)$. □

Chapitre 5

Couple de variables aléatoires

Dans des situations où interviennent plusieurs variables aléatoires, le calcul de la probabilité d'un événement dont la réalisation dépend des valeurs de ces variables doit faire intervenir ces variables considérées dans leur ensemble et non chacune isolément. Cela amène ainsi à étudier une nouvelle notion : celle de vecteur aléatoire.

5.1 Exemple Introductif

On effectue deux tirages successifs sans remise dans une urne contenant n boules numérotées 1 à n.
On note :
X la v.a.r. égale au plus petit numéro obtenu.
Y la v.a.r. égale au plus grand numéro obtenu.
L'univers $\Omega = \{\text{arrangment de 2 éléments}\}$. $Card(\Omega) = A_n^2 = n(n-1)$.

$$(X,Y)(\Omega) = \{(i,j) : 1 \leq i < j \leq n\} \subset X(\Omega) \times Y(\Omega).$$

On s'interesse à la loi conjointe du couple (X,Y) : $\mathbb{P}(X=i, Y=j) = \mathbb{P}(\{X=i\} \cap \{Y=j\})$ avec $i<j$ et $i \in \{1,\dots,n-1\}$ et $j \in \{2,\dots,n\}$. Il y a deux cas possibles (i,j) ou (j,i) et donc

$$\mathbb{P}(X=i, Y=j) = \frac{card(\{X=i, Y=j\})}{card(\Omega)} = \frac{2}{n(n-1)}.$$

On vérifie ici que

$$\sum_{1 \leq i < j \leq n} \mathbb{P}(X=i, Y=j) = C_n^2 \frac{2}{n(n-1)} = 1.$$

On peut aussi calculer les lois marginales $\mathbb{P}(X=i)$ et $\mathbb{P}(Y=j$.

5.2 Généralités

Définition 48.

Soient X, Y des v.a. définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. L'application

$$\begin{aligned} &\Omega \longrightarrow \mathbb{R}^2 \\ &\omega \mapsto (X(\omega), Y(\omega)). \end{aligned}$$

est appelée couple aléatoire, on le note (X,Y). Les variables aléatoires X et Y sont alors appelées ses marginales.

Le couple aléatoire (X,Y) permet de transporter la probabilité $\mathbb{P}$ de l'espace Ω sur l'espace $\mathbb{R}^2$. Rappelons qu'un produit cartésien $A \times B$ de deux ensembles A et B de $\mathbb{R}$ désigne l'ensemble suivant de $\mathbb{R}^2$:

$$A \times B = \{(a,b) : a \in A \quad \text{et} \quad b \in B\}.$$

Définition 49.

La loi $\mathbb{P}_{(X,Y)}$ du couple (X,Y) est la probabilité définie sur l'ensemble des produits d'intervalles $I \times J$ de $\mathbb{R}^2$ par $\forall I, J$ intervalles de $\mathbb{R}$

$$\mathbb{P}_{(X,Y)}(I \times J) = \mathbb{P}(\{\omega \in \Omega : (X(\omega), Y(\omega)) \in I \times J\})$$

Remarque : Dans le cas de v.a. X, Y discrètes, il est facile de voir que la loi du couple (X,Y) est caractérisée par les probabilités ponctuelles $\mathbb{P}_{(X,Y)}(\{x_i,y_j\}) = \mathbb{P}(X = x_i, Y = y_j)$ pour tout $x_i \in X(\Omega), y_j \in Y(\Omega)$ dans les domaines de valeurs de X et de Y.

5.2.1 Couple aléatoire discret

Rappelons encore que pour X, Y v.a. discrètes, les lois de X et de Y sont définies sur toutes les parties de $\mathbb{R}$, celle du couple sur toutes les parties de $\mathbb{R}^2$. Le résultat suivant montre qu'on retrouve la loi des v.a. marginales à partir de celle d'un couple.

Proposition 50.

Si (X,Y) est un couple aléatoire de v.a. discrètes de domaine $(X,Y)(\Omega) = \{(x_1,y_1),...,(x_i,y_i),...\}$, les domaines des marginales X, Y s'obtiennent par projection :

$$X(\Omega) = p_1((X,Y)(\Omega)) = \{x_1,...,x_i,...\} \quad Y(\Omega) = p_2((X,Y)(\Omega)) = \{y_1,...,y_i,...\}$$

où p_1, p_2 sont les première et seconde projections

$$p_1 : \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x,y) \mapsto x. \end{array} \qquad p_2 : \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x,y) \mapsto y. \end{array}$$

Les lois marginales $\mathbb{P}_X, \mathbb{P}_Y$ (i.e. les lois de X et de Y: ses marginales) sont données par :

$$\begin{array}{ll} \forall x_k \in X(\Omega) & \mathbb{P}(X = x_k) = \displaystyle\sum_{j, y_j \in Y(\Omega)} \mathbb{P}(X = x_k, Y = y_j) \\ \forall y_j \in Y(\Omega) & \mathbb{P}(Y = y_j) = \displaystyle\sum_{k, x_k \in X(\Omega)} \mathbb{P}(X = x_k, Y = y_j) \end{array}$$

Proof. Il sufft de faire la preuve pour le domaine et les probabilités ponctuelles de X. Or pour k fixé $\{X = x_k\}$ est la réunion de la famille dénombrable d'événements deux à deux disjoints $\{\mathbb{P}(X = x_k, Y = y_j)\}$ pour tous les j tels que $y_j \in Y(\Omega)$ car $\{Y = y_j\}$ forme une partition de Ω. On conclut alors par σ-additivité de $\mathbb{P}$

$$\begin{aligned} \mathbb{P}(X = x_k) &= \mathbb{P}(\{X = x_k\} \cap \Omega) = \mathbb{P}(\{X = x_k\} \cap \left(\cup_{j=1}^{+\infty} \{Y = y_j\} \right) = \mathbb{P}(\cup_{j=1}^{+\infty} \left(\{X = x_k\} \cap \{Y = y_j\} \right) \\ &= \sum_{j, y_j \in Y(\Omega)} \mathbb{P}(X = x_k, Y = y_j). \end{aligned}$$

□

Exemple : On donne le tableau des probabilités ponctuelles $\mathbb{P}(X = x_i; Y = y_j)$ d'un vecteur aléatoire discret (X,Y) :

$X \mid Y$	$y_1 = -1$	$y_2 = 2$	$y_= 3$	$y_4 = 5$	
$x_1 = 0$	0.1	0.05	0.15	0.0	0.3
$x_2 = 2$	0.05	0.2	0.05	0.1	0.4
$x_3 = 3$	0.1	0	0.1	0.1	0.3
	0.25	0.25	0.3	0.2	1

On en déduit la loi de X : $X(\Omega) = \{0; 2; 3\}$ et

$$\mathbb{P}(X = 0) = 0.3 \quad \mathbb{P}(X = 2) = 0.4 \quad \mathbb{P}(X = 3) = 0.3.$$

et celle de Y : $Y(\Omega) = \{-1; 2; 3; 5\}$ et

$$\mathbb{P}(Y = -1) = 0.25 \quad \mathbb{P}(Y = 2) = 0.25 \quad \mathbb{P}(Y = 3) = 0.3 \quad \mathbb{P}(Y = 5) = 0.2.$$

5.2.2 Variables aléatoires indépendantes

Définition 51.

Les v.a. discrètes X et Y sont indépendantes si

$$\forall x_k \in X(\Omega),\ \forall y_j \in Y(\Omega) \quad \mathbb{P}(X = x_k, Y = y_j) = \mathbb{P}(X = x_k) \times \mathbb{P}(Y = y_j)$$

Remarque : Une conséquence importante : si on connait les lois de X et de Y, des variables supposées indépendantes, on peut reconstruire la loi du couple $(X;Y)$ à partir des marginales par la relation précédente. Insistons sur le fait que ce n'est pas vrai en général quand X et Y ne sont pas indépendantes.
Exemple : 1. Dans l'exemple précédent, X et Y ne sont pas indépendantes car par exemple

$$\mathbb{P}(X = 2, Y = 2) = 0.2 \neq 0.4 \times 0.25 = 0.1 = \mathbb{P}(X = 2) \times \mathbb{P}(Y = 2)$$

2. On donne le tableau de la loi d'un couple (X, Y) en donnant les probabilités ponctuelles $\mathbb{P}(X = x_i, Y = y_j)$:

$X \mid Y$	$y_1 = -1$	$y_2 = 2$	$y_= 3$	
$x_1 = 0$	0.12	0.08	0.20	0.4
$x_2 = 2$	0.18	0.12	0.30	0.6
	0.3	0.2	0.5	1

On vérifie ici que X et Y sont indépendantes car pour tout $k = 1;2$ et $j = 1;2;3$, on a

$$\mathbb{P}(X = x_k, Y = y_j) = \mathbb{P}(X = x_k) \times \mathbb{P}(Y = y_j).$$

5.2.3 Covariance de deux variables aléatoires discrètes

La covariance permet d'estimer la dépendance entre deux variables aléatoires.

Définition 52.

La covariance de deux variables X et Y de carré intégrable est

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

L'espérance $\mathbb{E}(XY)$ est calculée à partir de la loi jointe de (X, Y) :

$$\mathbb{E}(XY) = \sum_{(i,j)\in(X(\Omega),Y(\Omega))} x_i y_j P(X = x_i, Y = y_j)$$

Théorème 53.

Si X et Y sont indépendantes, alors

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Et par suite cov(X, Y) =0.

Proof. On a

$$\mathbb{E}(XY) = \sum_{(i,j)\in(X(\Omega),Y(\Omega))} x_i y_j P(X = x_i, Y = y_j) = \sum_{i\in X(\Omega)} x_i P(X = x_i) \sum_{j\in Y(\Omega)} y_j P(Y = y_j)$$

□

Exercice : Soit X une v.a.r discrète dont la loi est donnée par :

Valeurs de X	-2	-1	0	1	2
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$

On considère la variable $Y = X^2$.

1. Déterminer la loi de Y ainsi que celle du couple (X, Y).
2. X et Y sont-elle indépendantes ?
3. Calculer $Cov(X, Y)$ et faire une remarque sur ce résultat.

Remarque : Soient X et Y deux v.a. Alors

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y).$$

Définition 54.

Le coefficient de corrélation linéaire est défini pour des variables non constantes par :

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)var(Y)}}.$$

Remarque : On a toujours $\rho(X, Y) \in [-1, 1]$. Plus $|\rho(X, Y)|$ est proche de 1 plus la dépendance entre les variables X et Y est forte. En effet, : on étudie le polynôme

$$P(\lambda) = \mathbb{E}(X^2)\lambda^2 - 2\mathbb{E}|XY|\lambda + \mathbb{E}(Y^2).$$

Ce polynôme se factorise en

$$P(\lambda) = \mathbb{E}[(\lambda|X| - |Y|)^2 \geq 0.$$

Il n'admet donc pas deux racines réelles distinctes, ce qui signifie que son discriminant (réduit) Δ est négatif. Ce qui donne

$$|\mathbb{E}(XY)| \leq \mathbb{E}|XY| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}.$$

On applique donc cette dernière inégalité avec $(X - \mathbb{E}(X))$ et $(Y - \mathbb{E}(Y))$.
Remarque : Si X et Y sont indépendantes, alors $Cov(X, Y) = 0$ et donc $\rho(X, Y) = 0$. On a par conséquent

$$Var(X + Y) = Var(X) + Var(Y).$$

Proposition 55.

Soient X et Y deux variables aléatoires discrètes indépendantes, f et g deux fonctions dont les domaines de définition contiennent respectivement $X(\Omega)$ et $Y(\Omega)$. Alors les variables aléatoires discrètes $f(X)$ et $g(Y)$ sont indépendantes.

Proposition 56.

Si X et Y sont deux v. a. d. indépendantes, la loi de la variable aléatoire $X + Y$ est donnée par :

$$\mathbb{P}(X + Y = n) = \sum_{i+j=n} \mathbb{P}(X = i)\mathbb{P}(Y = j) = \sum_{i=0}^{n} \mathbb{P}(X = i)\mathbb{P}(Y = n - i).$$

Proof. Comme X et Y sont à valeurs entières, il en est de même pour $X+Y$ dont la loi sera caractérisée par les $\mathbb{P}(X+Y=n)$. Pour les calculer, il suffitde décomposer l'événement $\{X+Y=n\}$ en la réunion disjointe de tous les événements $\{X=i, Y=j\}$ tels que $i+j=n$. □

5.3 Annexe : Convergence des lois

5.3.1 Lois hypergéométriques

En pratique lorsque l'on tire un échantillon de taille n parmi une population de taille N, le bon sens veut que l'on ne prenne pas 2 fois le même individu. Cela signifie que le tirage se fait sans remise. Les v.a de Bernouilli associées aux différents éléments de l'échantillon sont alors dépendants. Le schéma binomiale n'est plus adapté. Le schéma hypergéométrique est une succession d'épreuves de Bernouilli non indépendantes.
Alors que la loi binomiale intervient dans les tirages avec remise, la loi hypergéométrique correspond aux tirages sans remise.
Exemple Dans une production totale de N objets dont M sont défectueux, on prélève au hasard un échantillon de n objets (tirage sans remise). Soit X le nombre aléatoire d'objets défectueux dans l'échantillon. Quelle est sa loi?
On peut prendre comme espace l'ensemble de tous les échantillons possibles (toutes les parties à n éléments d'un ensemble de cardinal N) muni de l'équiprobabilité. Chaque échantillon a ainsi une probabilité $\frac{1}{C_N^n}$ d'être choisi. Les échantillons (événements élémentaires) réalisant l'événement $\{X = k\}$ sont ceux qui contiennent k objets défetueux et $n-k$ objets non défectueux. Ceci n'est réalisable que si $0 \leq k \leq M$ et $0 \leq n-k \leq N-M$. Dénombrons ces échantillons. On les forme en choisissant k objets défectueux dans une sous-population de M et en complétant par $n-k$ objets non défectueux choisis dans une sous population de $N-M$. Il y en a donc $C_M^k \times C_{N-M}^{n-k}$. Finalement :

$$\mathbb{P}(X=k) = \frac{C_M^k \times C_{N-M}^{n-k}}{C_N^n}, 0 \leq k \leq M, \quad 0 \leq n-k \leq N-M.$$

Définition 57.

Si dans une population de taille N, on a 2 types de populations:
M individus type 1.
$N-M$ individus type 2.
On fait n tirages sans remise dans la population et la v.a. X = nombre d'individus de type 1 dans l'échantillon. La loi de X est donc définie par

$$\mathbb{P}(X=k) = \frac{C_M^k \times C_{N-M}^{n-k}}{C_N^n}, 0 \leq k \leq M, \quad 0 \leq n-k \leq N-M.$$

Cette loi s'appelle loi hypergéométrique de paramètres N, M et n.
Notation : $X \sim \mathcal{H}(N, M, n)$.

Théorème 58.

On suppose que N tend vers $+\infty$ en vérifiant la condition :

$$\lim_{N\to+\infty} \frac{M}{N} = p, \quad 0 < p < 1.$$

Alors, pour n fixé, la loi hypergéométrique $\mathcal{H}(N, M, n)$ converge vers la loi binomiale $\mathcal{B}(n, p)$, i.e.

$$\forall k = 0, ..., n, \quad \lim_{N\to+\infty} \frac{C_M^k \times C_{N-M}^{n-k}}{C_N^n} = C_n^k p^k (1-p)^{n-k}.$$

On dit que la loi hypergéométrique converge vers la loi binomiale quand N tend vers $+\infty$.

Proof. Remarquons d'abord que comme p est strictement positif, l'hypothèse implique que M tend vers $+\infty$ avec N ; il en va de même pour $N - M$ puisque $p < 1$.
Pour n et k fixés, posons :

$$p_N = \frac{C_M^k \times C_{N-M}^{n-k}}{C_N^n} = C_n^k \frac{M!}{(M-k)!} \times \frac{(N-M)!}{((N-M)-(n-k))!} \times \frac{(N-n)!}{N!}.$$

Quand N tend vers $+\infty$ on a donc

$$\frac{M!}{(M-k)!} \sim M^k$$
$$\frac{(N-M)!}{((N-M)-(n-k))!} \sim (N-M)^{n-k}$$
$$\frac{(N-n)!}{N!} \sim \frac{1}{N^n}.$$

Donc

$$p_N \sim C_n^k \frac{M^k \times (N-M)^{n-k}}{N^n} = C_n^k (\frac{M}{N})^k (\frac{N-M}{N})^{n-k}.$$

□

Donc, puisque par hypothèse $\lim_{N\to+\infty} \frac{M}{N} = p$ et $\lim_{N\to+\infty} \frac{N-M}{N} = 1-p$

$$\lim_{n\to+\infty} p_N = \lim_{n\to+\infty} C_n^k (\frac{M}{N})^k (\frac{N-M}{N})^{n-k} = C_n^k (\frac{M}{N})^k (\frac{N-M}{N})^{n-k}.$$

Remarqque : Dans la pratique, on utilise l'approximation ds que $\frac{n}{N} < 0.1$.

5.3.2 Loi géométriques

Exemple: (Un problème de temps d'attente).
Considérons une suite infinie d'épreuves répétées indépendantes avec même probabilité de succès $p \in]0,1[$. Soit X le numéro (aléatoire) de la première épreuve où l'on obtient un succès. Si l'on n'obtient jamais de succès, on conviendra que $X = +\infty$. Calculer $\mathbb{P}(X = k)$ pour tout $k \in N^*$.
En notant $R_i = \{$ succès à la ième épreuve$\}$, on a :

$$\{X = k\} = \{\text{échec aux (k - 1) premières et succès à la kème}\} = (\cap_{i=1}^{k-1} R_i^c) \cap R_k$$

D'où par indépendance des épreuves :

$$\mathbb{P}\{X = k\} = \mathbb{P}\Big[(\cap_{i=1}^{k-1} R_i^c) \cap R_k\Big] = \prod_{i=1}^{k-1} (\mathbb{P}(R_i^c)) \times \mathbb{P}(R_k) = (1-p)^{k-1} \times p.$$

Définition 59.

Une v.a.r. X est dite de loi géométrique de paramètre p, pour p compris entre 0 et 1, si elle est à valeurs dans N^* et si
$$\mathbb{P}(X = k) = (1-p)^{k-1}p.$$
On note $X \sim G(p)$.

5.3.3 Loi de Poisson

Définition 60.

Une v.a.r. X est dite de loi de Poisson de paramètre λ, si elle est à valeurs dans $\mathbb{N}$ et si
$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$
On note $X \sim \mathcal{P}(\lambda)$.

Une des raisons de l'importance de cette loi est le théorème de convergence de la loi binomiale vers la loi de Poisson.

Théorème 61.

Si $(p_n)_{n\geq 1}$ est une suite de réels de $[0,1]$ vérifiant
$$\lim_{n\to+\infty} np_n = \lambda > 0,$$
alors :
$$\lim_{n\to+\infty} C_n^k p_n^k (1-p_n)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}.$$
On dit que la loi binomiale convergence vers la loi de Poisson.

Proof. On pose $np_n = \lambda u_n$ avec $\lim_{n\to+\infty} u_n = 1$ et donc
$$C_n^k p_n^k (1-p_n)^{n-k} = \frac{n!}{(n-k)!k!}(\frac{\lambda}{n})^k u_n^k (1-\frac{\lambda u_n}{n})^{n-k}.$$
Mais on a les limites suivantes
$$\lim_{n\to+\infty} \frac{n!}{(n-k)!n^k} = 1$$
$$\lim_{n\to+\infty} u_n^k = 1$$
$$\lim_{n\to+\infty} (1-\frac{\lambda u_n}{n})^{n-k} = e^{-\lambda},$$
ce qui donne le résultat voulu. □

Lorsqu'un événement a une faible probabilité p d'apparition lors d'une épreuve de Bernouilli et si l'on répète un grand nombre de fois cette épreuve (n) le nombre total de réalisations de l'événement considéré suit à peu près une loi de Poisson de paramètre $m = np$ Dans la pratique : $N > 50$ et $p < 0.1$.

Chapitre 6

Introduction à la statistique descriptive

La statistique est l'étude des populations, dont les éléments sont des individus; le plus souvent on n'étudie pas toute la population, mais seulement un échantillon de celle-ci. L'effectif d'un échantillon est le nombre d'individus qui le composent.

Définition 62.

On appelle caractère statistique simple tout application $X : P \to \mathbb{R}$, avec P un ensemble fini appelé population, tout élément de P s'appelle individu.

Les caractères des individus peuvent être qualitatifs (par exemple le prénom, la nationalité, ...) ou quantitatifs (l'âge, la taille, les revenus mensuels...). Les caractères quantitatifs peuvent être discrets (la pointure de chaussures, le nombre de personnes au foyer, ...) ou continus (la taille, la superficie d?une région, ...).

Définition 63.

On appelle modalité toute valeur $x_i \in X(P)$ telle que $X(P) = \{x_1, ..., x_n\}$, avec n le nombre des modalités différentes.

Pour faciliter l'étude pour des caractères continus on peut regrouper les valeurs en classes, c'est à dire en intervalles deux à deux disjoints $[x_i, x_{i+1}[$. Pour chaque classe $[x_i, x_{i+1}[$ on désigne par $c_i = \dfrac{x_i + x_{i+1}}{2}$ le centre de la classe et $x_{i+1} - x_i$ est l'amplitude de la classe.

Définition 64.

Une série statistique est un ensemble de couples (x_i, n_i), où les x_i sont les modalités du caractère X et les n_i le nombre de fois où la valeur x_i apparaît. L'effectif total de l'échantillon est donc $n = \sum_{i=1}^{n} n_i$. On appelle fréquence d'apparition de x_i le nombre $f_i = \dfrac{n_i}{n}$, où n_i est l'effectif correspond au modalité x_i.
Les effectifs cumulés N_i et les fréquences cumulées F_i sont donnés par

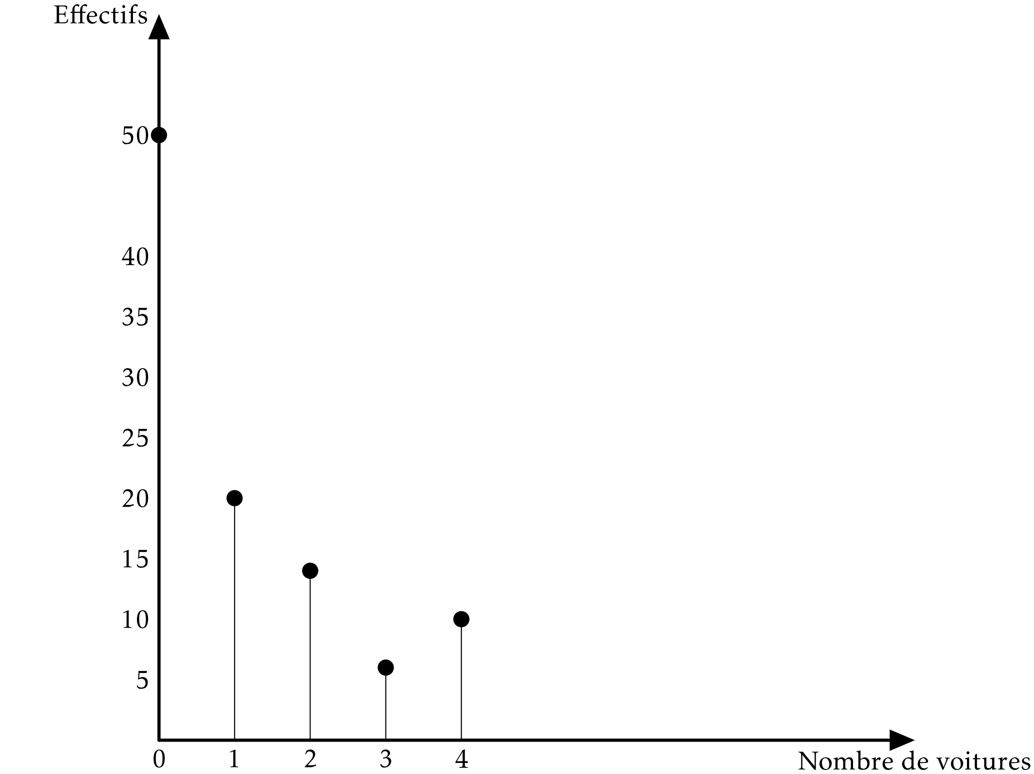
$$N_i = \sum_{j=1}^{i} n_j, \quad F_i = \sum_{j=1}^{i} f_j.$$

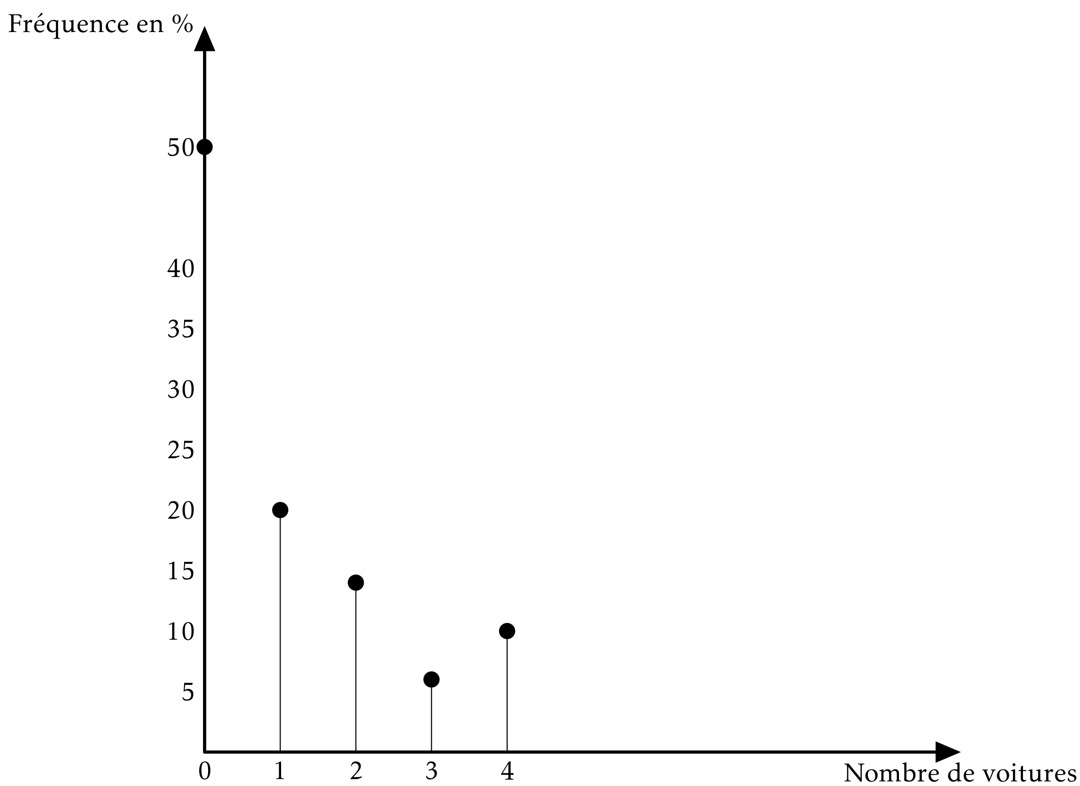
Exemple 1 : sur un échantillon de cent pièces tirées de la production journalière d'une usine, on compte le nombre de défauts constatés : Ici X est la variable qui correspond au nombre de défauts:

Nombre de défauts	0	1	2	3	4
Effectifs n_i	50	20	14	6	10
Effectifs cumulés N_i	50	70	84	90	100
Fréquences f_i	0.5	0.2	0.14	0.06	0.1
Fréquences cumulées F_i	0.5	0.7	0,84	0.9	1

Ici les valeurs (modalités) de X sont donc $x_0 = 0, x_1 = 1, ..., x_4 = 4$ d'effectifs respectifs $n_0 = 50, n_1 = 20, n_2 = 14, n_3 = 6, n_4 = 10$. L'effectif total est de $n = 100$.

Représentation graphique :
Diagramme en bâtons :





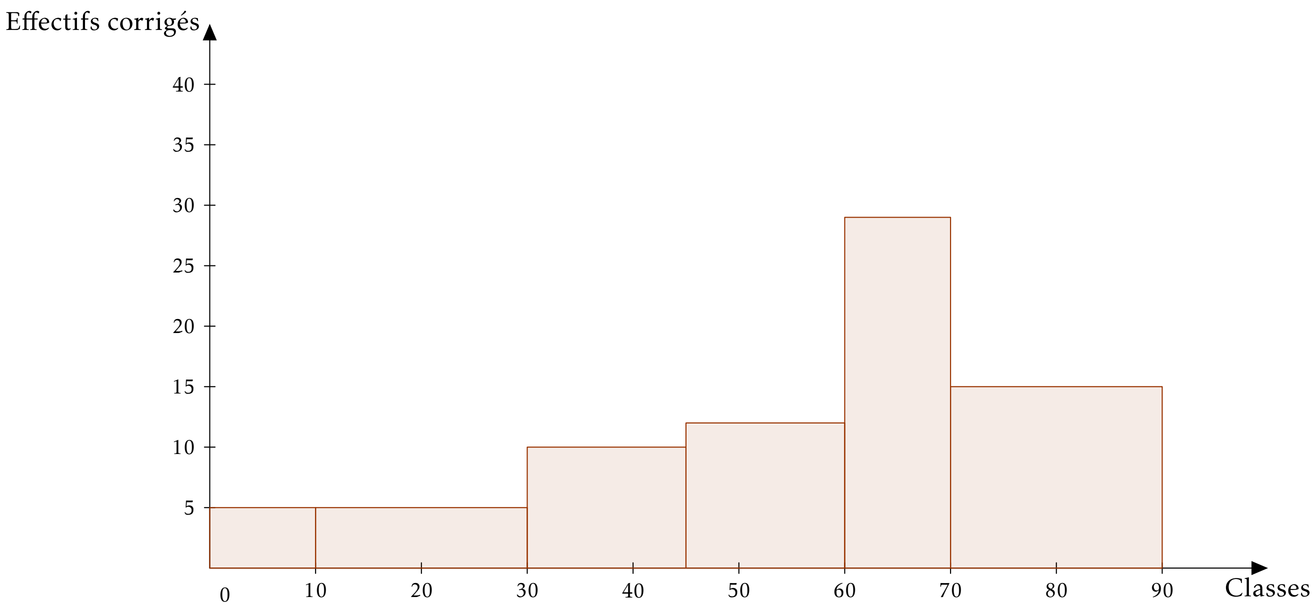
Exemple 2 : Pour une période d'une heure et demi dans un parking, on a compté le nombre de voitures rentrant et on a obtenu les valeurs suivantes:

Temps (min)	[0; 10[	[10; 30[	[30; 45[	[45; 60[	[60; 70[	[70; 90[
Nombre de voiture	5	10	15	18	29	30
Fréquences	0.04	0.09	0.88	0.16	0.27	0.28

L'effectif total est $n = 107$. On calcule l'amplitude de la classe. Ici (10, 20, 15,15, 10, 20).
Pour tracer l'histogramme on place en abscisse les différentes classes, ici [0;10[, [10;30[, [30;45[, [45;60[, [60;70[, [70;90[. Pour chaque classe on calcule alors la hauteur du rectangle correspondant : en général on se ramène à la plus petite amplitude : Ici 10.

Temps (en min)	[0;10[	[10;30[	[30;45[	[45;60[	[60;70[	[70;90[
Nombre de voitures	5	10	15	18	29	30
Amplitude	10	20	15	15	10	20
Effectifs corrigés	$\frac{5\times 10}{10} = 5$	$\frac{10\times 10}{20} = 5$	$\frac{15\times 10}{15} = 10$	$\frac{18\times 10}{15} = 12$	$\frac{29\times 10}{10} = 29$	$\frac{30\times 10}{20} = 15$
Fréquences	0.04	0.09	0.14	0.168	0.27	0.29
Fréq. cumulées	0.04	0.13	0.27	0.438	0.708	1
Centre de classe c_i	5	20	37.5	52.5	65	80
$n_i c_i$	25	200	562.5	945	1885	2400

Histogramme :



6.1 Caractéristiques de position

La moyenne arithmétique :
la moyenne d'une série statistique (x_i, n_i) est le nombre

$$\overline{x} = \frac{\sum_{i=1}^{p} n_i x_i}{n}.$$

Dans le cas où les valeurs sont regroupées en classes, on peut déterminer une valeur approchée de la moyenne en remplaçant pour le calcul chaque classe par son milieu.

$$\overline{x} = \frac{\sum_{i=1}^{p} n_i c_i}{n}.$$

La variance et l'écart-type de la série statistique sont données par :

$$Var(X) = \frac{\sum_{i=1}^{p} n_i (x_i - \overline{x})^2}{n}.$$
$$\sigma_X = \sqrt{Var(X)}.$$

le mode est la valeur la plus fréquente (la valeur de plus grand effectif) d'une série statistique ; pour une série répartie en classe on parle de classe modale (la classe de plus grand effectif après correction des effectifs).
La médiane d'une série statistique, notée Me, est le nombre qui partage la population en deux parties de même effectif : les individus dont la valeur du caractère est inférieure à la médiane, et les individus dont la valeur du caractère est supérieure à la médiane. Quand les données sont regroupées en classes on parle de classe modale.
La médiane est la valeur pour laquelle la fréquence cumulée est égale à 50%.
Exemple
Dans l'exemple 1, le mode est 0 et la médiane est aussi 0. Dans l'exemple 2, la classe modale est $[60,70[$ et la classe médiane est $[60,70[$. On calcule la médiane par interpolation linéaire :

$$\begin{array}{ccc} 60 & Me & 70 \\ \hline 0.438 & 0.50 & 0.708 \end{array} \longrightarrow$$

$$\frac{0.50 - 0.438}{Me - 60} = \frac{0.708 - 0.438}{70 - 60}$$

Ainsi $Me = 60,70$