



**1<sup>ère</sup> année**

# **PROBABILITES**

**2009-2010**



# TABLE DES MATIERES

<b>1.</b>	<b>Eléments de base des probabilités</b>	<b>page 1</b>
1.1	Définitions	1
1.2	Propriétés d'une probabilité	4
1.3	Exemples de probabilités	8
1.4	Rappels de dénombrement	10
1.5	Probabilités conditionnelles	14
<b>2.</b>	<b>Variables aléatoires</b>	<b>17</b>
2.1	Définitions et généralités	17
2.2	Variables aléatoires réelles discrètes	21
2.3	Variables aléatoires réelles absolument continues	23
2.4	Lois usuelles	27
2.5	Fonctions d'une variable aléatoire réelle	37
<b>3.</b>	<b>Moments d'une variable aléatoire réelle</b>	<b>41</b>
3.1	Moments d'une variable aléatoire réelle discrète	41
3.2	Moments d'une variable aléatoire réelle absolument continue	44
3.3	Moments d'une fonction d'une variable aléatoire réelle	46
3.4	Propriétés des moments d'une variable aléatoire réelle	47
3.5	Moments des lois usuelles	49
<b>4.</b>	<b>Vecteurs aléatoires</b>	<b>55</b>
4.1	Vecteurs aléatoires discrets	55
4.2	Vecteurs aléatoires continus	61
4.3	Vecteurs aléatoires mixtes	70
<b>5.</b>	<b>Indépendance et corrélation</b>	<b>71</b>
5.1	Rappels : variables aléatoires indépendantes	71
5.2	Vecteurs aléatoires à composantes indépendantes	71
5.3	Corrélation linéaire de deux variables aléatoires	74
5.4	Caractéristiques des vecteurs aléatoires	78
5.5	Vecteurs gaussiens	80
5.6	Somme de variables aléatoires indépendantes	82
<b>6.</b>	<b>Distributions conditionnelles</b>	<b>89</b>
6.1	Distribution conditionnelle de $X$ sachant $X \in A$	89
6.2	Distribution conditionnelle de $X$ sachant $Y$	93
<b>7.</b>	<b>Fonctions génératrices et caractéristiques</b>	<b>101</b>
7.1	Fonctions génératrices des probabilités d'une variable entière	101
7.2	Fonctions génératrices des moments	104
7.3	Fonctions caractéristiques	107
7.4	Fonctions caractéristiques des lois usuelles	111
7.5	Fonctions caractéristiques de vecteurs aléatoires	113

<b>8.</b>	<b>Suites de variables aléatoires</b>	<b>117</b>
8.1	Convergences des suites de variables aléatoires	117
8.2	Lois des grands nombres	119
8.3	Théorème de la limite centrée	120
<b>9.</b>	<b>Statistiques sur un échantillon aléatoire</b>	<b>123</b>
9.1	Définitions	123
9.2	Distribution de la moyenne aléatoire d'un échantillon	124
9.3	Distribution de la variance aléatoire d'un échantillon	125
9.4	Cas des échantillons gaussiens	127
<b>10.</b>	<b>Notions d'estimation</b>	<b>129</b>
10.1	Généralités	129
10.2	Détermination d'un estimateur de maximum de vraisemblance	131
10.3	Estimation par intervalle de confiance	135
	<b>TABLES NUMERIQUES</b>	<b>141</b>

# 1. Éléments de base des probabilités

## 1.1. Définitions

### Définition 1.1.a : *Théorie des probabilités*

La théorie des probabilités est la branche des mathématiques qui donne un cadre formel à l'étude des phénomènes aléatoires (ou considérés comme tels).

### Définition 1.1.b : Phénomène *aléatoire* (ou *indéterministe*, ou *stochastique*)

Phénomène qui, répété dans des conditions identiques, peut donner des résultats différents.

### Définition 1.1.c : *Univers des possibles* (ou “*ensemble fondamental*”, ou “*référentiel*”)

Souvent noté  $\Omega$ , il désigne l'ensemble des résultats possibles que l'on choisit de considérer pour un phénomène aléatoire.

### Exemples 1.1.d :

1°) Phénomène : la trajectoire d'un dé lancé sur une table ;

Résultat considéré : la face supérieure du dé  $\Rightarrow \Omega = \llbracket 1, 6 \rrbracket$  ;

2°) Phénomène : la trajectoire de deux dés lancés sur une table ;

Résultat considéré : les faces supérieures des dés  $\Rightarrow \Omega = \{ (1,1), \dots (6,6) \}$  ;

3°) Phénomène : la trajectoire de deux dés lancés sur une table ;

Résultat considéré : la somme des faces supérieures des dés  $\Rightarrow \Omega = \llbracket 2, 12 \rrbracket$  ;

4°) Phénomène : la consultation d'un site internet ;

Résultat considéré : le nombre d'accès au site durant une période de temps donnée ;

Le nombre maximum d'accès étant inconnu et théoriquement non borné :  $\Omega = \mathbb{N}$  ;

5°) Phénomène : la consultation d'un site internet ;

Résultat considéré : la durée d'une consultation (en minutes) ;

La durée maximale d'une consultation étant inconnue et théoriquement non bornée :

$$\Omega = \mathbb{R}_+^* ;$$

**Définition 1.1.e : Evénement**

Un événement est un sous-ensemble de l'univers des possibles,  $\Omega$ .

Un événement est dit *réalisé* si le résultat du phénomène aléatoire appartient à cet événement.

**Exemples 1.1.f :**

1°)  $\Omega = \llbracket 1, 6 \rrbracket$  ;  $A = \{ 2, 4, 6 \}$  : l'événement  $A$  est réalisé si le lancé du dé donne un résultat pair ;

2°)  $\Omega = \llbracket 2, 12 \rrbracket$  ;  $B = \{ 7 \}$  : l'événement  $B$  est réalisé si la somme des faces des deux dés est égale à 7 ;

3°)  $\Omega = \mathbb{R}_+^*$  ;  $C = ] 10 ; +\infty [$  : l'événement  $C$  est réalisé si la consultation dure plus de dix minutes ;

**Définitions 1.1.g : Evénements *certain, impossible, simple, composé***

$\Omega$  (considéré ici comme sous-ensemble) est l'événement *certain*.

$\emptyset$  est l'événement *impossible*.

Tout singleton est appelé événement *simple*.

Tout sous-ensemble de cardinalité supérieure à 1 est appelé événement *composé*.

**Définitions 1.1.h : Opérations ensemblistes sur les événements**

Soient  $A$  et  $B$  deux événements :

$\overline{A} = \Omega - A$  est l'événement *contraire* à  $A$  ;

$A \cap B$  est la *conjonction* de  $A$  et  $B$  : il est réalisé si  $A$  et  $B$  sont réalisés ;

$A \cup B$  est la *disjonction (non exclusive)* de  $A$  et  $B$  : il est réalisé si  $A$  ou  $B$  sont réalisés ;

$A \oplus B = (A \cup B) - (A \cap B)$  est la *disjonction exclusive* de  $A$  et  $B$  : il est réalisé si  $A$  ou bien  $B$  est réalisé ;

Si  $A \cap B = \emptyset$ ,  $A$  et  $B$  sont dits *disjoints* (ou *incompatibles*, ou *en contradiction*) : leur réalisation simultanée est impossible ;

Si  $A \subset B$ , la réalisation de  $A$  implique celle de  $B$  ;

**Rappels 1.1.i : Règles de Morgan**

$$\overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \overline{A_i} ; \quad \overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \overline{A_i} .$$

**Définition 1.1.k :** Famille d'événements probabilisable et tribu

Pour qu'une famille d'événements  $T \subset \mathcal{P}(\Omega)$  (ensemble des parties de  $\Omega$ ) soit *probabilisable*, il est nécessaire que :

- elle soit stable par complémentation :  $\forall A \in T, \bar{A} \in T$  ;
- elle soit stable pour l'union (et l'intersection) finie ou dénombrable :

$$\forall (A_n)_{n \in N} \in T^N, \bigcup_{n \in IN} A_n \in T \text{ et } \bigcap_{n \in IN} A_n \in T ;$$

- $\Omega \in T$  et  $\emptyset \in T$ .

Mais il est suffisant que :

- elle soit stable par complémentation :  $\forall A \in T, \bar{A} \in T$  ;
- elle soit stable pour l'union finie ou dénombrable :

$$\forall (A_n)_{n \in N} \in T^N, \bigcup_{n \in IN} A_n \in T ;$$

- $\emptyset \in T$ .

Une telle famille probabilisable est appelée une *tribu* sur l'ensemble  $\Omega$ .

**Définition 1.1.l :** Tribu engendrée par une famille d'événements

La tribu  $T$  engendrée par une famille d'événements  $(A_i)_{i \in I}$  est l'ensemble des parties de  $\Omega$  obtenu en effectuant l'union de zéro, un, deux, ...,  $n = \text{Card}(I)$  événement(s)  $A_i$ .

**Exemples 1.1.m :** pour  $\Omega = \llbracket 1, 6 \rrbracket$  :

- la tribu engendrée par  $\{\emptyset, \Omega\}$  est  $\{\emptyset, \Omega\}$  (tribu *grossière*) ;
- la tribu engendrée par  $\{\{1, 2, 3\}, \{4, 5, 6\}\}$  est  $\{\emptyset, \{1, 2, 3\}, \{4, 5, 6\}, \Omega\}$  ;
- la tribu engendrée par  $\{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$  est  $\{\emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}, \{1, 2, 3, 4\}, \{1, 2, 5, 6\}, \{3, 4, 5, 6\}, \Omega\}$  (de cardinal  $2^3 = 8$ ) ;
- la tribu engendrée par  $\{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$  est  $\mathcal{P}(\Omega)$  (de cardinal  $2^6 = 64$ ).

**Remarque 1.1.n :**

Dans la pratique, sauf cas particuliers, les familles probabilisables choisies seront :

- pour  $\Omega$  fini ou dénombrable :  $\mathcal{P}(\Omega)$ , appelée tribu discrète, qui est engendrée par l'ensemble des singletons de  $\Omega$  ;
- pour  $\Omega$  non dénombrable : la tribu engendrée par les ouverts de  $\mathbb{R}$  (tribu borélienne).

**Définition 1.1.o : Espace probabilisable**

Si  $T$  est une famille probabilisable (une tribu) sur  $\Omega$ , alors le couple  $(\Omega, T)$  est appelé un *espace probabilisable*.

**Définition 1.1.p : Probabilité**

On appelle probabilité sur un espace probabilisable  $(\Omega, T)$  toute application  $\mathbf{P}$  de  $T$  dans  $[0 ; 1]$  telle que :

- $\mathbf{P}(\Omega) = 1$  ;
- pour toute famille finie ou dénombrable d'événements disjoints,  $(A_i)_{i \in I}$  :

$$\mathbf{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mathbf{P}(A_i) \text{ (propriété d'additivité).}$$

**Définition 1.1.q : Espace probabilisé**

Si  $\mathbf{P}$  est une probabilité définie sur un espace probabilisable  $(\Omega, T)$ , alors le triplet  $(\Omega, T, \mathbf{P})$  est appelé un *espace probabilisé*.

**1.2. Propriétés d'une probabilité**

Dans toute cette section, nous considérerons un espace probabilisé  $(\Omega, T, \mathbf{P})$ .

**Proposition 1.2.a :**

Etant donnée une probabilité  $\mathbf{P}$ , on a  $\mathbf{P}(\emptyset) = 0$ .

Démonstration :

Comme  $\Omega$  et  $\emptyset$  sont évidemment disjoints,  $\mathbf{P}(\Omega \cup \emptyset) = \mathbf{P}(\Omega) + \mathbf{P}(\emptyset)$ .

Or,  $\Omega \cup \emptyset = \Omega$ , donc  $\mathbf{P}(\Omega) + \mathbf{P}(\emptyset) = \mathbf{P}(\Omega)$ , d'où  $\mathbf{P}(\emptyset) = 0$ .

**Propositions 1.2.b :**

Etant donnés deux événements  $A$  et  $B$ , on a les propriétés suivantes :

1.  $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$  ;
2.  $\mathbf{P}(A \setminus B) = \mathbf{P}(A) - \mathbf{P}(A \cap B)$  ;
3.  $B \subset A \Rightarrow \mathbf{P}(B) \leq \mathbf{P}(A)$ .
4.  $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$  ;



### Démonstrations :

1.  $A$  et  $\bar{A}$  sont disjoints et tels que  $A \cup \bar{A} = \Omega$ , donc on a :

$$\mathbf{P}(A \cup \bar{A}) = \mathbf{P}(A) + \mathbf{P}(\bar{A}) = 1, \text{ d'où le résultat.}$$

2.  $A \cap B$  et  $A \setminus B$  sont disjoints et tels que  $(A \cap B) \cup (A \setminus B) = A$ , donc on a :

$$\mathbf{P}[(A \cap B) \cup (A \setminus B)] = \mathbf{P}(A \cap B) + \mathbf{P}(A \setminus B) = \mathbf{P}(A), \text{ d'où le résultat.}$$

3. C'est une conséquence du point précédent : si  $B \subset A$ , alors  $A \cap B = B$ , et comme

$$\mathbf{P}(A \setminus B) \geq 0, \text{ on a } \mathbf{P}(A) - \mathbf{P}(B) \geq 0, \text{ d'où le résultat.}$$

4.  $(A \setminus B)$  et  $B$  sont disjoints et tels que  $(A \setminus B) \cup B = A \cup B$ , d'où :

$$\mathbf{P}[(A \setminus B) \cup B] = \mathbf{P}(A \setminus B) + \mathbf{P}(B) = \mathbf{P}(A \cup B), \text{ d'où le résultat d'après le point 2.}$$

### **Proposition 1.2.c :**

Pour toute famille finie d'événements  $(A_i)_{1 \leq i \leq n}$ , on a :  $\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbf{P}(A_i)$ .

### Démonstration :

La propriété se démontre par récurrence : elle est triviale pour  $n = 1$  ; pour  $n = 2$ , comme d'après le point ci-dessus on a  $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$ , il en découle immédiatement que  $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$  ; l'héritage se démontre en réappliquant le cas  $n = 2$  entre  $\bigcup_{i=1}^n A_i$  et  $A_{n+1}$ .

### **Théorème 1.2.d : Théorème des probabilités totales**

Si  $(B_i)_{i \in I}$  est une partition de  $\Omega$ , alors  $\forall A \in \mathcal{T}, \mathbf{P}(A) = \sum_{i \in I} \mathbf{P}(A \cap B_i)$ .

### Démonstration :

$$\mathbf{P}(A) = \mathbf{P}(A \cap \Omega) = \mathbf{P}\left(A \cap \bigcup_{i \in I} B_i\right) = \mathbf{P}\left(\bigcup_{i \in I} A \cap B_i\right). \text{ Or } (A \cap B_i)_{i \in I} \text{ est une partition de } A,$$

d'où le résultat.

Rappel :  $(B_i)_{i \in I}$  est une partition de  $\Omega$  s.s. :

-  $\forall (i, j) \in I^2, (i \neq j) \Rightarrow B_i \cap B_j = \emptyset$  ;

-  $\bigcup_{i \in I} B_i = \Omega$ .

**Proposition 1.2.e :** *Formule (du crible) de (Sylvester-)Poincaré*

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{J \in \mathcal{P}_k(\{1, \dots, n\})} \mathbf{P}\left(\bigcap_{j \in J} A_j\right), \text{ où } \mathcal{P}_k(E) = \{P \in \mathcal{P}(E) / \text{Card}(P) = k\}.$$

Démonstration :

On démontre la formule par récurrence sur  $n$ . Elle est triviale pour  $n = 1$ , et connue pour  $n = 2$ . Supposons la vraie pour un  $n \in \mathbb{N}^*$ , et soient  $n+1$  événements  $A_1, \dots, A_{n+1}$ . On a :

$$\begin{aligned} \mathbf{P}\left(\bigcup_{i=1}^{n+1} A_i\right) &= \mathbf{P}\left(\left(\bigcup_{i=1}^n A_i\right) \cup A_{n+1}\right) = \mathbf{P}\left(\bigcup_{i=1}^n A_i\right) + \mathbf{P}(A_{n+1}) - \mathbf{P}\left(\left(\bigcup_{i=1}^n A_i\right) \cap A_{n+1}\right) \\ &= \mathbf{P}\left(\bigcup_{i=1}^n A_i\right) + \mathbf{P}(A_{n+1}) - \mathbf{P}\left(\bigcup_{i=1}^n (A_i \cap A_{n+1})\right) \\ &= \sum_{k=1}^n (-1)^{k-1} \sum_{J \in \mathcal{P}_k(\{1, \dots, n\})} \mathbf{P}\left(\bigcap_{j \in J} A_j\right) + \mathbf{P}(A_{n+1}) - \sum_{k=1}^n (-1)^{k-1} \sum_{J \in \mathcal{P}_k(\{1, \dots, n\})} \mathbf{P}\left(\left(\bigcap_{j \in J} A_j\right) \cap A_{n+1}\right) \\ &= \sum_{k=1}^{n+1} (-1)^{k-1} \sum_{\substack{I \in \mathcal{P}_k(\{1, \dots, n+1\}) \\ n+1 \notin I}} \mathbf{P}\left(\bigcap_{i \in I} A_i\right) + \sum_{l=1}^{n+1} (-1)^{l-1} \sum_{\substack{J \in \mathcal{P}_l(\{1, \dots, n+1\}) \\ n+1 \in J}} \mathbf{P}\left(\bigcap_{j \in J} A_j\right) \\ &= \sum_{k=1}^{n+1} (-1)^{k-1} \sum_{J \in \mathcal{P}_k(\{1, \dots, n+1\})} \mathbf{P}\left(\bigcap_{j \in J} A_j\right). \end{aligned}$$

**Proposition 1.2.f :** *Propriété de continuité de la probabilité*

Pour toute famille infinie dénombrable d'événements  $(A_n)_{n \geq 0}$ , telle que, pour tout  $n \in \mathbb{N}$ ,

$$A_n \subset A_{n+1} \text{ alors : } \mathbf{P}\left(\bigcup_{n=0}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mathbf{P}(A_n).$$

Démonstration :

Posons  $B_k = A_{k+1} - A_k$ , pour  $k \in \mathbb{N}$ .

Les événements  $B_k$  sont disjoints, et comme  $A_n = A_0 \cup \bigcup_{k=0}^{n-1} B_k$ ,  $n \in \mathbb{N}$ , on a :

$$\bigcup_{n=0}^{\infty} A_n = A_0 \cup \bigcup_{k=0}^{\infty} B_k.$$

$$\begin{aligned} \text{On a alors } \mathbf{P}\left(\bigcup_{n=0}^{\infty} A_n\right) &= \mathbf{P}(A_0) + \sum_{k=0}^{\infty} \mathbf{P}(B_k) = \mathbf{P}(A_0) + \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} \mathbf{P}(B_k) \\ &= \lim_{n \rightarrow \infty} \left( \mathbf{P}(A_0) + \sum_{k=0}^{n-1} \mathbf{P}(B_k) \right) = \lim_{n \rightarrow \infty} \mathbf{P}(A_n) \end{aligned}$$

**Proposition 1.2.g :**

Pour toute famille infinie dénombrable d'événements  $(A_n)_{n \geq 0}$ , on a :

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

Démonstration :

Si la série  $\sum_{n \geq 0} \mathbf{P}(A_n)$  diverge, le résultat est trivial.

On considère la suite d'événements  $(C_k)_{k \geq 0}$ , avec  $C_k = \bigcup_{n=0}^k A_n$ .

Cette suite est croissante, au sens de l'inclusion. La proposition précédente peut donc s'appliquer à celle-ci et la suite des  $\mathbf{P}(C_k)$  a donc une limite quand  $k$  tend vers l'infini.

Il a été établi précédemment que pour tout  $k \in \mathbb{N}$ ,  $\mathbf{P} \left( \bigcup_{n=0}^k A_n \right) \leq \sum_{n=0}^k \mathbf{P}(A_n)$ .

En considérant donc que la série  $\sum_{n \geq 0} \mathbf{P}(A_n)$  converge, le résultat est établi par passage à la limite de l'inégalité ci-dessus.

**Proposition 1.2.h : Propriété de continuité de la probabilité**

Pour toute famille infinie dénombrable d'événements  $(A_n)_{n \geq 0}$ , on a :

Si  $A_i \supset A_{i+1}$  pour tout  $i$ , alors  $\mathbf{P} \left( \bigcap_{n=0}^{\infty} A_n \right) = \lim_{n \rightarrow \infty} \mathbf{P}(A_n)$ .

Démonstration :

Posons  $B_n = A_0 - A_n$ , pour  $n \in \mathbb{N}^*$ . Les  $B_n$  forment une suite croissante, et on a :

$$\mathbf{P}(A_0) = \mathbf{P}(B_n) + \mathbf{P}(A_n) \geq \mathbf{P}(B_n).$$

Ainsi, la suite des  $\mathbf{P}(B_n)$  est croissante et majorée, donc  $\lim_{n \rightarrow \infty} \mathbf{P}(B_n)$  existe.

$$\begin{aligned} \text{De plus, d'après la proposition 1.2.e : } \lim_{n \rightarrow \infty} \mathbf{P}(B_n) &= \mathbf{P} \left( \bigcup_{n=1}^{\infty} B_n \right) = \mathbf{P} \left( A_0 - \bigcap_{n=1}^{\infty} A_n \right) \\ &= \mathbf{P}(A_0) - \mathbf{P} \left( \bigcap_{n=1}^{\infty} A_n \right). \end{aligned}$$

$$\text{Donc, } \mathbf{P} \left( \bigcap_{n=1}^{\infty} A_n \right) = \mathbf{P}(A_0) - \lim_{n \rightarrow \infty} \mathbf{P}(B_n) = \lim_{n \rightarrow \infty} [\mathbf{P}(A_0) - \mathbf{P}(B_n)] = \lim_{n \rightarrow \infty} \mathbf{P}(A_n).$$

## 1.3. Exemples de probabilité

### 1.3.1 $\Omega$ fini

Pour  $\Omega = \{x_i\}_{i \in \{1, \dots, n\}}$ , on peut définir une probabilité vérifiant :

$$\forall i \in \llbracket 1, n \rrbracket, \mathbf{P}(\{x_i\}) = p_i, \text{ avec } \sum_{i=1}^n p_i = 1.$$

Pour tout événement  $A = \bigcup_{i \in I} \{x_i\}$ , avec  $I \subset \llbracket 1, n \rrbracket$ , on a alors :  $\mathbf{P}(A) = \sum_{i \in I} p_i$ .

**Cas particulier de l'équiprobabilité :**  $p_1 = p_2 = \dots = p_n = \frac{1}{n} = \frac{1}{\text{Card}(\Omega)}$

Pour tout événement  $A = \bigcup_{i \in I} \{x_i\}$ , avec  $I \subset \llbracket 1, n \rrbracket$ , on a  $\mathbf{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$ .

Dans ce cas particulier, les calculs de probabilités se ramènent très souvent à des problèmes de dénombrement (il faut déterminer  $\text{Card}(A)$ ).

#### Exemples 1.3.1.a :

1°) Phénomène : la trajectoire d'un dé lancé sur une table ;

Résultat considéré : la face supérieure du dé ;

$$\Omega = \llbracket 1, 6 \rrbracket, \text{ et } \mathbf{P}(\{1\}) = \dots = \mathbf{P}(\{6\}) = \frac{1}{6}.$$

2°) Phénomène : la trajectoire de deux dés lancés sur une table ;

Résultat considéré : les faces supérieures des dés ;

$$\Omega = \{ (1,1), \dots (6,6) \} \text{ et } \mathbf{P}(\{(1,1)\}) = \dots = \mathbf{P}(\{(6,6)\}) = \frac{1}{36};$$

3°) Phénomène : la trajectoire de deux dés lancés sur une table ;

Résultat considéré : la somme des faces supérieures des dés ;

$$\Omega = \llbracket 2, 12 \rrbracket \text{ et } \mathbf{P}(\{2\}) = \mathbf{P}(\{12\}) = \frac{1}{36}, \mathbf{P}(\{3\}) = \mathbf{P}(\{11\}) = \frac{2}{36}, \dots$$

$$\mathbf{P}(\{6\}) = \mathbf{P}(\{8\}) = \frac{5}{36}, \text{ et } \mathbf{P}(\{7\}) = \frac{6}{36}.$$

### 1.3.2 $\Omega$ infini dénombrable

Pour  $\Omega = \{x_i\}_{i \in \mathbb{N}}$  on peut définir une probabilité vérifiant :

$$\forall i \in \mathbb{N}, \mathbf{P}(\{x_i\}) = p_i.$$

En conséquence, on a  $\mathbf{P}(\Omega) = \sum_{i=0}^{+\infty} p_i = 1$ ,

et pour tout événement  $A = \bigcup_{i \in I} \{x_i\}$ , avec  $I \subset \mathbb{N}$ ,  $\mathbf{P}(A) = \sum_{i \in I} p_i$ .

#### Exemple 1.3.2.a :

Phénomène : la consultation d'un site internet ;

Résultat considéré : le nombre d'accès au site durant une période de temps donnée

$$\Omega = \mathbb{N} \text{ et } \forall n \in \mathbb{N}, \mathbf{P}(\{n\}) = \frac{\lambda^n}{n!} e^{-\lambda} (\lambda \in \mathbb{R}_+^*);$$

### 1.3.3 $\Omega$ infini non dénombrable

Pour  $\Omega$  intervalle de  $\mathbb{R}$ , on peut définir une probabilité vérifiant :

$$\forall (a, b) \in \Omega^2, \mathbf{P}([a; b]) = \mathbf{P}([a; b[) = \mathbf{P}(]a; b]) = \mathbf{P}(]a; b[) = \int_a^b f(t) dt,$$

où  $f$  est une fonction définie sur  $\Omega$ , positive, intégrable et telle que  $\int_{\Omega} f(t) dt = 1$ .

#### Exemple 1.3.3.a :

Phénomène : la consultation d'un site internet ;

Résultat considéré : la durée d'une consultation (en minutes) ;

$$\Omega = \mathbb{R}_+^* \text{ et } f(t) = \lambda e^{-\lambda t} (\lambda \in \mathbb{R}_+^*);$$

$$\text{Pour } C = ]10; +\infty[, \mathbf{P}(C) = \left[ -e^{-\lambda t} \right]_{10}^{+\infty} = e^{-10\lambda}.$$

## 1.4. Rappels de dénombrement

### 1.4.1. Arrangements

#### Définition 1.4.1.a : Arrangements

Soient deux ensembles  $E$  et  $F$ , avec  $\text{Card}(E) = p$  et  $\text{Card}(F) = n$ .

Le nombre d'*arrangements* est le nombre d'injections de  $E$  dans  $F$  (on doit avoir  $n \geq p$ ) :

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

#### Exemples 1.4.1.b :

1°) C'est le nombre de manières de placer  $p$  jetons dans  $n$  cases, chaque case ne pouvant contenir qu'un seul jeton.

2°) Dans le cas particulier,  $n = p$ , c'est le nombre de bijections de  $E$  dans  $F$  ( $n!$ ), donc par exemple le nombre de permutations possibles de  $n$  objets.

### 1.4.2. Combinaisons

#### Définition 1.4.2.a : Combinaisons

Le nombre de *combinaisons* de  $p$  éléments parmi  $n$  est le nombre de manières de choisir  $p$  éléments parmi  $n$ .

Pour  $p \in \llbracket 0, n \rrbracket$ ,  $C_n^p = \frac{n!}{p!(n-p)!}$  et, par convention,  $C_n^p = 0$  sinon.

Remarque :  $C_n^p$  est également noté  $\binom{n}{p}$ .

**Propriétés 1.4.2.b :**

1.  $C_n^0 = C_n^n = 1$  ;
2.  $C_n^p = C_n^{n-p}$  ;
3.  $C_n^p + C_n^{p+1} = C_{n+1}^{p+1}$  ;

Le triangle de Pascal découle de cette dernière propriété :

$n \setminus p$	0	1	2	3	...
0	1				
1	1	1			
2	1	2	1		
3	1	3	3	1	
$\vdots$	$\vdots$				$\ddots$

Démonstration :

Les deux premières propriétés se vérifient immédiatement.

$$\begin{aligned}
 \text{Si } p \in \llbracket 0, n-1 \rrbracket : C_n^p + C_n^{p+1} &= \frac{n!}{p!(n-p)!} + \frac{n!}{(p+1)!(n-p-1)!} \\
 &= \frac{n!}{(p+1)!(n-p)!} ((p+1) + (n-p)) \\
 &= \frac{(n+1)!}{(p+1)!((n+1)-(p+1))!} = C_{n+1}^{p+1}.
 \end{aligned}$$

$$\text{Si } p < -1 \text{ ou } p > n : C_n^p = C_n^{p+1} = C_{n+1}^{p+1} = 0.$$

$$\text{Si } p = -1 : C_n^p + C_n^{p+1} = C_n^0 = 1 \text{ et } C_{n+1}^{p+1} = C_{n+1}^0 = 1.$$

$$\text{Si } p = n : C_n^p + C_n^{p+1} = C_n^n = 1 \text{ et } C_{n+1}^{p+1} = C_{n+1}^{n+1} = 1.$$

Remarque : si on considère un élément  $e$  particulier,  $C_n^p$  est le nombre de parties à  $p$  éléments contenant  $e$  et  $C_n^{p+1}$  le nombre de parties à  $p+1$  éléments ne contenant pas  $e$ .

**Exemple 1.4.2.c :**

Le nombre de mains de 8 cartes (pour un jeu de 32 cartes) ayant 3 as est :  $C_4^3 \times C_{28}^5$

### 1.4.3. Autres dénombrements

**Proposition 1.4.3.a :** *Nombre d'applications de  $E$  dans  $F$*

Etant donnés deux ensembles finis  $E$  et  $F$  de cardinaux respectifs  $n$  et  $p$ , le nombre d'applications de  $E$  dans  $F$  est  $p^n$ .

**Proposition 1.4.3.b :** *Nombre d'applications surjectives de  $E$  dans  $F$*

Etant donnés deux ensembles finis  $E$  et  $F$  de cardinaux respectifs  $n$  et  $p$ , le nombre d'applications surjectives de  $E$  dans  $F$  est le nombre  $S(n, p)$  défini par la relation de récurrence

$$S(n+1, p) = p ( S(n, p-1) + S(n, p) ).$$

De plus, on a : si  $n < p$   $S(n, p) = 0$  ; si  $p = 1$ ,  $S(n, p) = 1$ .

Les premiers de ces nombres sont donnés dans le tableau ci-dessous :

$n \setminus p$	1	2	3	4	5	6	7
1	1	0	0	0	0	0	0
2	1	2	0	0	0	0	0
3	1	6	6	0	0	0	0
4	1	14	36	24	0	0	0
5	1	30	150	240	120	0	0
6	1	62	540	1560	1800	720	0
7	1	126	1806	8400	16800	15120	5040
8	1	254	5796	40824	126000	191520	141120
9	1	510	18150	186480	834120	1905120	2328480
10	1	1022	55980	818520	5103000	16435440	29635200
11	1	2046	171006	3498000	29607600	129230640	322494480
12	1	4094	519156	14676024	165528000	953029440	3162075840
13	1	8190	1569750	60780720	901020120	6711344640	2,8806E+10
14	1	16382	4733820	249401880	4809004200	4,5674E+10	2,4862E+11

Remarques :

1°) Si  $n = p$ , on a  $S(n, p) = n!$ .

2°) Si  $p = 2$ , on a  $S(n, p) = 2^n - 2$ , car seules les deux fonctions constantes ne sont pas surjectives.

Démonstration :

Chaque surjection peut être définie à partir d'une partition de  $E$  en  $p$  parties et d'une association de chacune de ces parties à un élément de  $F$ , image commune à tous les éléments de cette partie. En notant  $P(n, p)$  le nombre de partitions d'un ensemble de  $n$  éléments en  $p$  parties, on a donc :

$$S(n, p) = p! P(n, p).$$

Par ailleurs, on a la formule de récurrence  $P(n+1, p) = P(n, p-1) + p P(n, p)$ .



En effet, parmi les  $P(n+1, p)$  partitions d'un ensemble de  $n+1$  éléments,  $x_1, \dots, x_{n+1}$ , il y a celles qui contiennent le singleton  $\{x_{n+1}\}$  et celles qui ne le contiennent pas. Le nombre de partitions de la première catégorie est donné par le nombre de partitions de  $\{x_1, \dots, x_n\}$  en  $p-1$  parties, tandis que le nombre de partitions de la seconde catégorie est donné par le nombre de partitions de  $\{x_1, \dots, x_n\}$  en  $p$  parties multiplié par  $p$  parce que l'élément  $x_{n+1}$  peut être placée dans n'importe laquelle des  $p$  parties.

#### **Exemples 1.4.3.c :**

Le nombre de manières de placer  $n$  jetons différenciés dans  $p$  cases, est  $p^n$  si chaque case peut contenir un nombre quelconque de jetons et  $S(n, p)$  si chaque case doit contenir au moins un jeton.

#### **Proposition 1.4.3.d (admise) :**

Le nombre de sommes de  $n$  entiers égales à  $p$  est  $N(n, p) = C_{n+p-1}^{p-1}$ .

Remarque :  $N(n, p)$  est le nombre de manières de placer  $p-1$  "séparateurs" entre  $n$  objets.

Le nombre de sommes de  $p$  entiers non nuls égales à  $n$  est  $N^*(n, p) = C_{n-1}^{p-1}$ .

#### **Exemples 1.4.3.e :**

Le nombre de manières de placer  $p$  jetons indifférenciés dans  $n$  cases, est  $N(n, p)$  si chaque case peut contenir un nombre quelconque de jetons et  $N^*(n, p)$  si chaque case doit contenir au moins un jeton.

#### **Proposition 1.4.3.f : Nombre de parties d'un ensemble $E$**

Etant donné un ensemble fini  $E$  de cardinal  $n$ , le nombre de parties de  $E$  est  $2^n$ .

#### Démonstration :

Pour  $k \in \llbracket 0, n \rrbracket$ , le nombre de parties de  $E$  de cardinal  $k$  est  $C_n^k$ . Le nombre total de parties

est donc  $\sum_{k=0}^n C_n^k = 2^n$  d'après la formule du binôme.

## 1.5. Probabilités conditionnelles

### 1.5.1. Définitions

Dans toute cette partie, nous considérons des événements appartenant à un espace probabilisé  $(\Omega, T, P)$ .

#### Définition 1.5.1.a : Probabilité conditionnelle

Soient  $A$  et  $B$  deux événements tels que  $P(B) \neq 0$ . On appelle probabilité conditionnelle de  $A$  sachant  $B$  (ou de “ $A$  si  $B$ ”, ou de “ $A$  relativement à  $B$ ”, etc.), la valeur :

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

$P_B(A)$  est la probabilité pour que  $A$  se réalise sachant que  $B$  est réalisé.

#### Proposition 1.5.1.b :

$P_B$  est une probabilité sur  $(\Omega, T, P)$ .

#### Démonstration :

D’une part on a bien  $P_B(\Omega) = 1$ , et d’autre part, pour toute famille finie ou dénombrable d’événements disjoints,  $(A_i)_{i \in I}$  :

$$P_B\left(\bigcup_{i \in I} A_i\right) = \frac{P\left[\left(\bigcup_{i \in I} A_i\right) \cap B\right]}{P(B)} = \frac{P\left[\bigcup_{i \in I} (A_i \cap B)\right]}{P(B)} = \frac{\sum_{i \in I} P(A_i \cap B)}{P(B)} = \sum_{i \in I} P_B(A_i).$$

#### Exemples 1.5.1.c :

Les proportions de pièces défectueuses fabriquées par trois machines  $M_1$ ,  $M_2$  et  $M_3$  sont respectivement égales à 0.3 %, 0.8 % et 1 %.

$\Omega = \{\text{“pièce correcte”}, \text{“pièce défectueuse”}\} \times \{\text{“pièce fabriquée par la machine } M_1\text{”}, \text{“pièce fabriquée par la machine } M_2\text{”}, \text{“pièce fabriquée par la machine } M_3\text{”}\}$  ;

Soit  $D$  l’événement “la pièce fabriquée est défectueuse”, et  $(M_i)_{i \in \{1,2,3\}}$  les événements “la pièce est fabriquée par la machine  $M_i$ ”.

On a :  $P(D | M_1) = 0.003$  ;  $P(D | M_2) = 0.008$  ;  $P(D | M_3) = 0.01$  .

**Théorème 1.5.1.d : Reformulation du théorème des probabilités totales**

Soit  $(\Omega, T, \mathbf{P})$  un espace probabilisé.

Si  $(B_i)_{i \in I}$  est une partition de  $\Omega$ , alors :

$$\forall A \in T, \mathbf{P}(A) = \sum_{i \in I} \mathbf{P}(A \cap B_i) = \sum_{i \in I} (\mathbf{P}(A | B_i) \times \mathbf{P}(B_i))$$

**Proposition 1.5.1.e : Formule des probabilités composées**

$$\mathbf{P}\left(\bigcap_{i=1}^n A_i\right) = \left[ \prod_{i=2}^n \mathbf{P}\left(A_i \mid \bigcap_{j=1}^{i-1} A_j\right) \right] \times \mathbf{P}(A_1)$$

Démonstration :

La démonstration peut être effectuée par récurrence sur  $n$ .

Pour  $n = 2$ , on a bien  $\mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_2 | A_1) \times \mathbf{P}(A_1)$ .

Hypothèse de récurrence : pour  $n \geq 2$ , on a  $\mathbf{P}\left(\bigcap_{i=1}^n A_i\right) = \left[ \prod_{i=2}^n \mathbf{P}\left(A_i \mid \bigcap_{j=1}^{i-1} A_j\right) \right] \times \mathbf{P}(A_1)$  ;

$$\begin{aligned} \text{Alors : } \mathbf{P}\left(\bigcap_{i=1}^{n+1} A_i\right) &= \mathbf{P}\left(A_{n+1} \cap \bigcap_{i=1}^n A_i\right) = \mathbf{P}\left(A_{n+1} \mid \bigcap_{i=1}^n A_i\right) \times \mathbf{P}\left(\bigcap_{i=1}^n A_i\right) \\ &= \mathbf{P}\left(A_{n+1} \mid \bigcap_{i=1}^n A_i\right) \times \left[ \prod_{i=2}^n \mathbf{P}\left(A_i \mid \bigcap_{j=1}^{i-1} A_j\right) \right] \times \mathbf{P}(A_1) = \left[ \prod_{i=2}^{n+1} \mathbf{P}\left(A_i \mid \bigcap_{j=1}^{i-1} A_j\right) \right] \times \mathbf{P}(A_1) \end{aligned}$$

**1.5.2. Formule de Bayes****Proposition 1.5.2.a : Formules de Bayes**

$$1^\circ) \mathbf{P}(B | A) = \frac{\mathbf{P}(A | B) \times \mathbf{P}(B)}{\mathbf{P}(A)}$$

2°) Si  $(B_i)_{i \in I}$  est une partition de  $\Omega$ , alors :

$$\mathbf{P}(B_k | A) = \frac{\mathbf{P}(A | B_k) \times \mathbf{P}(B_k)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A | B_k) \times \mathbf{P}(B_k)}{\sum_{i \in I} (\mathbf{P}(A | B_i) \times \mathbf{P}(B_i))}$$

Démonstration :

$$1^\circ) \mathbf{P}(A | B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \Rightarrow \mathbf{P}(A \cap B) = \mathbf{P}(A | B) \times \mathbf{P}(B) \quad (1)$$

$$\mathbf{P}(B | A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} \Rightarrow \mathbf{P}(A \cap B) = \mathbf{P}(B | A) \times \mathbf{P}(A) \quad (2)$$

La première formule découle des égalités (1) et (2).

**Exemple 1.5.2.b :**

Si  $M_1$  (respectivement  $M_2, M_3$ ) effectue 50 % (respectivement 35 %, 15 %) de la production, on a :  $\mathbf{P}(M_1) = 0.5$  ;  $\mathbf{P}(M_2) = 0.35$  ;  $\mathbf{P}(M_3) = 0.15$  .

$$\begin{aligned} \text{Alors : } \mathbf{P}(M_1 | D) &= \frac{\mathbf{P}(D | M_1) \times \mathbf{P}(M_1)}{\sum_{i=1}^3 (\mathbf{P}(D | M_i) \times \mathbf{P}(M_i))} \\ &= \frac{0.003 \times 0.5}{0.003 \times 0.5 + 0.008 \times 0.35 + 0.01 \times 0.15} \approx 0.26 . \end{aligned}$$

**1.5.3. Notion d'indépendance****Définition 1.5.3.a :** *Indépendance de deux événements*

Deux événements  $A$  et  $B$  sont dits *indépendants* si et seulement si  $\mathbf{P}_B(A) = \mathbf{P}(A)$  ou bien, ce qui est équivalent :

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

**Définition 1.5.3.b :** *Indépendance mutuelle d'une famille d'événements*

Les événements d'une famille  $(A_i)_{i \in I}$  sont dits *mutuellement indépendants* si et seulement si :

$$\forall J \subset I, \mathbf{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbf{P}(A_j)$$

**Proposition 1.5.3.c :**

Si des événements sont mutuellement indépendants, alors ils sont indépendants deux à deux, mais la réciproque est fausse.

Démonstration :

L'implication dans le sens direct est immédiate.

Pour montrer que la réciproque est fausse, il suffit de trouver un contre-exemple. En voici un : Considérons un dé tétraédrique dont une face est rouge, une face verte, une face bleue et la quatrième face tricolore, rouge, verte et bleue et notons  $R$ , (respectivement  $V$  et  $B$ ) les événements “la face contre la table contient du rouge” (respectivement “du vert”, “du bleu”). Les événements  $R$ ,  $V$  et  $B$  sont deux à deux indépendants mais ne sont pas mutuellement indépendants.

## 2. Variables aléatoires

Dans tout cette section  $(\Omega, T, P)$  désigne un espace probabilisé.

### 2.1. Définitions et généralités

#### Définition 2.1.a : Variable aléatoire

Une variable aléatoire est une application de  $\Omega$  dans un ensemble  $E$ .

#### Exemples 2.1.b :

1°)  $\Omega = \{ \text{pile, face} \}$ ;  $E = \{ 1, -1 \}$ ;  $X(\text{pile}) = 1$  et  $X(\text{face}) = -1$ .

2°)  $\Omega = \mathbb{R}$ ;  $E = \{ \text{négatif, nul, positif} \}$ ;  $X(\omega) = \text{négatif si } \omega < 0, X(\omega) = \text{nul si } \omega = 0 \text{ et } X(\omega) = \text{positif si } \omega > 0$ .

#### Définition 2.1.c : Variable certaine

Une variable aléatoire  $X$  telle que  $X(\Omega)$  est réduit à un singleton  $\{a\}$  est dite *certaine*.

#### Définition 2.1.d : Variable aléatoire réelle (v.a.r)

Toute variable aléatoire  $X$  telle que  $X(\Omega) \subset \mathbb{R}$  est appelée une *variable aléatoire réelle* que nous désignerons, dans la suite de ce document sous l'abréviation "v.a.r".

#### Définition 2.1.e : Variable mesurable ; Probabilité image

On définit sur  $E$ , une tribu  $T'$  : ainsi  $(E, T')$  est un espace probabilisable, et tout élément  $B$  de  $T'$  peut être désigné comme un événement.

Pour tout événement  $B$  de  $T'$ , on notera  $X^{-1}(B)$  l'ensemble  $\{ \omega \in \Omega ; X(\omega) \in B \}$ .

Une variable aléatoire est dite *mesurable* si et seulement si :

$$\forall B \in T', X^{-1}(B) \in T.$$

Si une variable aléatoire  $X$  est mesurable, on peut définir, sur  $(E, T')$ , la *probabilité image* de  $P$  par  $X$ , notée  $P_X$ , par :

$$\forall B \in T', P_X(B) = P(X^{-1}(B)).$$

**Remarque 2.1.f :**

Nous ne considérerons que des variables aléatoires mesurables. Dans ces conditions,  $X^{-1}(B)$  sera toujours un événement de  $T$ .

**Exemple 2.1.g :**

Avec  $E = \mathbb{R}$ , pour qu'une v.a.r.  $X$  soit mesurable, il est nécessaire que  $\forall x \in E$ ,

$$X^{-1}([-\infty, x]) = \{ \omega \in \Omega ; X(\omega) \leq x \} \in T,$$

car nous voulons que  $\mathbf{P}(\{ \omega \in \Omega ; X(\omega) \leq x \})$  existe.

On démontre que cette condition est suffisante pour que  $X$  soit mesurable.

**Notations 2.1.h :**

L'ensemble  $X^{-1}(B) = \{ \omega \in \Omega ; X(\omega) \in B \}$  sera noté “ $X \in B$ ”. On a donc :

$$\mathbf{P}_X(B) = \mathbf{P}(X^{-1}(B)) = \mathbf{P}(\{ \omega \in \Omega ; X(\omega) \in B \}) = \mathbf{P}(X \in B).$$

Par exemple :

- pour  $B = [a ; b]$ ,  $\mathbf{P}_X([a ; b]) = \mathbf{P}(\{ \omega \in \Omega ; a \leq X(\omega) \leq b \}) = \mathbf{P}(a \leq X \leq b)$  ;
- pour  $B = \{x\}$ ,  $\mathbf{P}_X(\{x\}) = \mathbf{P}(\{ \omega \in \Omega ; X(\omega) = x \}) = \mathbf{P}(X = x)$ .

“ $X \leq x$ ”, “ $X = x$ ”, “ $a \leq X \leq b$ ”, etc. désignent donc des sous ensembles de  $\Omega$ .

**Définition 2.1.i : Variables aléatoires indépendantes**

Soient  $X_1$  et  $X_2$  deux variables aléatoires mesurables, et soient  $T_1$  et  $T_2$  des tribus définies sur  $X_1(\Omega)$  et sur  $X_2(\Omega)$ .  $X_1$  et  $X_2$  sont dites *indépendantes* si et seulement si pour tout couple  $(B_1, B_2)$  de  $T_1 \times T_2$ ,  $X_1^{-1}(B_1)$  et  $X_2^{-1}(B_2)$  sont deux événements indépendants de  $\Omega$ .

On a donc  $\mathbf{P}((X_1 \in B_1) \cap (X_2 \in B_2)) = \mathbf{P}(X_1^{-1}(B_1) \cap X_2^{-1}(B_2))$

$$= \mathbf{P}(X_1^{-1}(B_1)) \cdot \mathbf{P}(X_2^{-1}(B_2)) = \mathbf{P}(X_1 \in B_1) \cdot \mathbf{P}(X_2 \in B_2)$$

**Définition 2.1.j :** *Fonction de répartition d'une variable aléatoire réelle.*

La fonction de répartition d'une variable aléatoire  $X$ , notée  $F_X$ , est définie par :

$$F_X : \mathbb{R} \rightarrow [0 ; 1]$$

$$x \mapsto \mathbf{P} ( X \leq x )$$

Remarque : s'il n'y a pas d'ambiguïté cette fonction pourra être notée  $F$ .

**Proposition 2.1.k :**

La fonction de répartition de toute variable aléatoire  $X$ ,  $F_X$ , possède les propriétés fondamentales suivantes :

1. Elle prend des valeurs entre 0 et 1 :  $\forall x \in X(\Omega), 0 \leq F_X(x) \leq 1$  ;
2. Elle est croissante :  $x_1 \leq x_2 \Leftrightarrow F_X(x_1) \leq F_X(x_2)$  ;
3. Elle est continue à droite :  $\lim_{h \rightarrow 0} F_X(x+h) = F_X(x)$  ;
4. Sa limite en  $-\infty$  vaut 0 :  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  ;
5. Sa limite en  $+\infty$  vaut 1 :  $\lim_{x \rightarrow +\infty} F_X(x) = 1$ .

Démonstration :

1.  $F_X(x)$  est la valeur d'une probabilité, donc est à valeur dans  $[0 ; 1]$  ;
2.  $x_1 \leq x_2 \Rightarrow ]-\infty ; x_1] \subset ]-\infty ; x_2] \Rightarrow \mathbf{P}_X(]-\infty ; x_1]) \leq \mathbf{P}_X(]-\infty ; x_2])$  ;
3.  $\lim_{h \rightarrow 0} F_X(x+h) = \lim_{n \rightarrow \infty} F_X\left(x + \frac{1}{n}\right) = \lim_{n \rightarrow \infty} \mathbf{P}_X\left(\left]-\infty ; x + \frac{1}{n}\right]\right)$   

$$= \mathbf{P}_X\left(\bigcap_{n=1}^{\infty} \left]-\infty ; x + \frac{1}{n}\right]\right) = \mathbf{P}_X(]-\infty ; x]) = F_X(x) ;$$
4. On a  $\lim_{x \rightarrow -\infty} F_X(x) = \lim_{n \rightarrow \infty} F_X(-n) = \lim_{n \rightarrow \infty} \mathbf{P}_X(]-\infty ; -n]) = \mathbf{P}_X\left(\bigcap_{n=1}^{\infty} \left]-\infty ; -n]\right)\right)$ .

Or  $\bigcap_{n=1}^{\infty} \left]-\infty ; -n]\right) = \emptyset$ , d'où  $\lim_{x \rightarrow -\infty} F_X(x) = \mathbf{P}_X(\emptyset) = 0$  ;

5. On a  $\lim_{x \rightarrow +\infty} F_X(x) = \lim_{n \rightarrow \infty} F_X(n) = \lim_{n \rightarrow \infty} \mathbf{P}_X(]-\infty ; n]) = \mathbf{P}_X\left(\bigcup_{n=1}^{\infty} \left]-\infty ; n]\right)\right)$ .

Or  $\bigcup_{n=1}^{\infty} \left]-\infty ; n]\right) = \mathbb{R}$ , d'où  $\lim_{x \rightarrow +\infty} F_X(x) = \mathbf{P}_X(\mathbb{R}) = 1$ .

**Proposition 2.1.1 :**

Soient une variable aléatoire  $X$ , de fonction de répartition  $F$ , et deux nombres réels  $a$  et  $b$  tels que  $a \leq b$ . Alors :

$$1^\circ) \mathbf{P}(X < a) = \lim_{\substack{x \rightarrow a \\ x < a}} F(x) = \lim_{x \rightarrow a^-} F(x) = F(a^-);$$

$$2^\circ) \mathbf{P}(a < X \leq b) = F(b) - F(a);$$

$$3^\circ) \mathbf{P}(a \leq X \leq b) = F(b) - F(a^-);$$

$$4^\circ) \mathbf{P}(a < X < b) = F(b^-) - F(a);$$

$$5^\circ) \mathbf{P}(X = a) = F(a) - F(a^-).$$

Démonstration :

1°) Soit  $(a_n)_{n \in \mathbb{N}}$  une suite numérique croissante telle que  $\lim_{n \rightarrow \infty} a_n = a$ .

Comme  $\forall n \in \mathbb{N}, ]-\infty, a_n[ \subset ]-\infty, a_{n+1}[$ , d'après la propriété **1.2.e**, dite de continuité de la

$$\begin{aligned} \text{probabilité, on a : } \mathbf{P}(X < a) &= \mathbf{P}(]-\infty; a[) = \mathbf{P}\left(\bigcup_{n=0}^{+\infty} ]-\infty; a_n[\right) = \lim_{n \rightarrow \infty} \mathbf{P}(]-\infty; a_n[) \\ &= \lim_{n \rightarrow \infty} F(a_n) = F(a^-). \end{aligned}$$

$$\begin{aligned} 2^\circ) \mathbf{P}(a < X \leq b) &= \mathbf{P}_X(]-\infty; b[ - ]-\infty; a]) \\ &= \mathbf{P}_X(]-\infty; b[) - \mathbf{P}_X(]-\infty; a]) = F(b) - F(a). \end{aligned}$$

$$\begin{aligned} 3^\circ) \mathbf{P}(a \leq X \leq b) &= \mathbf{P}_X(]-\infty; b[ - ]-\infty; a[) \\ &= \mathbf{P}_X(]-\infty; b[) - \mathbf{P}_X(]-\infty; a[) = F(b) - F(a^-). \end{aligned}$$

$$\begin{aligned} 4^\circ) \mathbf{P}(a < X < b) &= \mathbf{P}_X(]-\infty; b[ - ]-\infty; a]) \\ &= \mathbf{P}_X(]-\infty; b[) - \mathbf{P}_X(]-\infty; a]) = F(b^-) - F(a). \end{aligned}$$

$$5^\circ) \mathbf{P}(X = a) = \mathbf{P}(a \leq X \leq a) = \mathbf{P}_X(]-\infty; a[ - ]-\infty; a[) = F(a) - F(a^-).$$



## 2.2. Variables aléatoires réelles discrètes

### Définition 2.2.a : Variable aléatoire réelle discrète

Une v.a.r.  $X$  est dite *discrète* lorsque l'ensemble image de  $\Omega$  par  $X$ ,  $X(\Omega)$ , est fini ou dénombrable.

### Exemples 2.2.b :

1°)  $\Omega = \{ \text{pile, face} \}$ ;  $X(\Omega) = \{ 1, -1 \}$ ;  $X(\text{pile}) = 1$  et  $X(\text{face}) = -1$ ;

$X$  est une v.a.r. discrète finie.

2°) On lance une pièce  $N$  fois de suite :  $\Omega = \{ \text{pile, face} \}^N$ ;

$X$  associe à chaque événement de  $\Omega$  le nombre de "pile" obtenus :  $X(\Omega) = \llbracket 0, N \rrbracket$ ;

$X$  est une v.a.r. discrète finie.

3°) On lance une pièce jusqu'à ce que l'on obtienne un "pile" :  $\Omega = \bigcup_{k \geq 0} \{ \text{face} \}^k \times \{ \text{pile} \}$ ;

$X$  associe à chaque événement de  $\Omega$  le nombre de lancer qui lui correspond :  $X(\Omega) = \mathbb{N}^*$ ;

$X$  est une v.a.r. discrète dénombrable.

### Définition 2.2.c : Loi d'une variable aléatoire réelle discrète

La loi d'une v.a.r. discrète est définie par :

- a) l'ensemble  $X(\Omega) = \{ x_k, k \in K \}$ , avec  $K \subset \mathbb{N}$ ;
- b) pour chaque valeur  $x_k \in X(\Omega)$ , la probabilité  $p_k = \mathbf{P}(X = x_k)$ .

Remarque : Deux v.a.r. différentes, mais qui ont la même loi sont dites *identiquement distribuées* (ou *équiréparties*).

### Exemples 2.2.d : basé sur l'exemple 2.2.b

1°)  $\Omega = \{ \text{pile, face} \}$ ;  $X(\Omega) = \{ 1, -1 \}$ ;  $\mathbf{P}(X = 1) = p$ ;  $\mathbf{P}(X = -1) = 1 - p$ ;

2°)  $\Omega = \{ \text{pile, face} \}^N$ ;  $X(\Omega) = \llbracket 0, N \rrbracket$ ;  $\forall k \in X(\Omega)$ ,  $\mathbf{P}(X = k) = \binom{N}{k} p^k (1 - p)^{N-k}$ ;

3°)  $\Omega = \bigcup_{n \geq 0} \{ \text{face} \}^n \times \{ \text{pile} \}$ ;  $X(\Omega) = \mathbb{N}^*$ ;  $\forall k \in X(\Omega)$ ,  $\mathbf{P}(X = k) = (1 - p)^{k-1} p$

**Définition 2.2.e : Fonction de répartition**

Soit  $X$  une v.a.r. discrète dont la loi de probabilité est donnée par  $X(\Omega) = \{x_k, k \in K\}$  et par  $\forall x_k \in X(\Omega), \mathbf{P}(X = x_k) = p_k$ .

La fonction de répartition de  $X$  est alors :

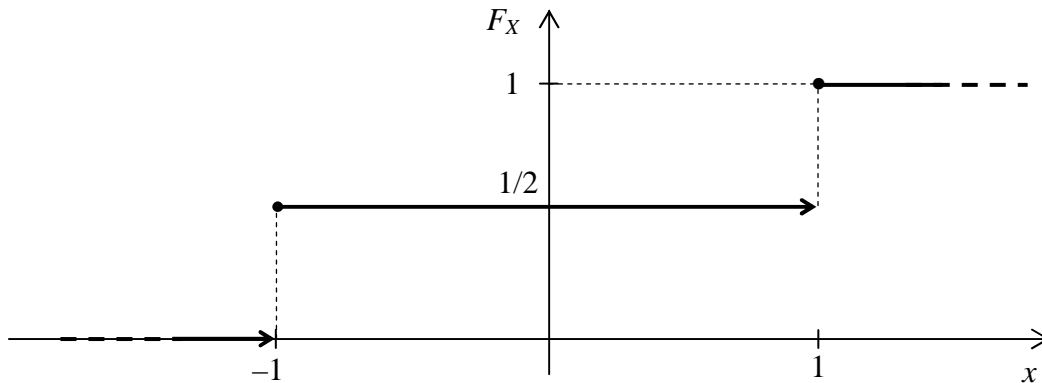
$$F_X(x) = \mathbf{P}(X \leq x) = \sum_{\substack{k \in K \\ x_k \leq x}} \mathbf{P}(X = x_k) = \sum_{\substack{k \in K \\ x_k \leq x}} p_k.$$

Cette fonction est une fonction en escalier, continue à droite, et ayant un nombre fini ou dénombrable de points de discontinuité qui sont les points  $x_k$ .

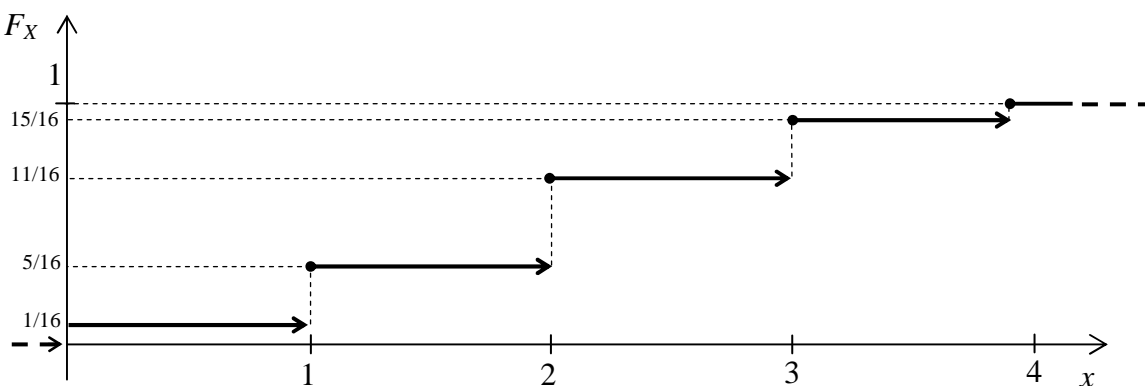
Pour ces points, on a :  $\mathbf{P}(X = x_k) = F(x_k) - F(x_{k-1})$ .

**Exemples 2.2.f : basé sur l'exemple 2.2.b (avec  $p = 1/2$ )**

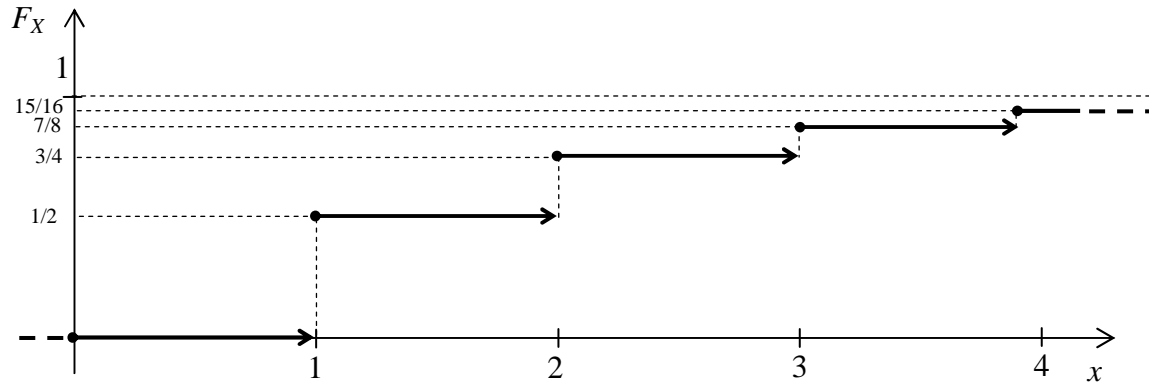
1°)  $\Omega = \{\text{pile, face}\}; X(\Omega) = \{1, -1\}; \mathbf{P}(X = 1) = p; \mathbf{P}(X = -1) = 1 - p;$



2°)  $\Omega = \{\text{pile, face}\}^N; X(\Omega) = \llbracket 0, N \rrbracket; \forall k \in X(\Omega), \mathbf{P}(X = k) = C_N^k p^k (1 - p)^{N-k};$  avec  $N = 4$  :



$$3^\circ) \Omega = \bigcup_{n \geq 0} \{ \text{face} \}^n \times \{ \text{pile} \} ; X(\Omega) = \mathbb{N}^* ; \forall k \in X(\Omega), \mathbf{P}(X = k) = (1 - p)^{k-1} p ;$$



### 2.3. Variables aléatoires réelles absolument continues

**Définition 2.3.a :** *Variable aléatoire réelle absolument continue*

Une v.a.r. est absolument continue si et seulement si sa fonction de répartition  $F$  est continûment dérivable par morceaux sur  $\mathbb{R}$ .

**Définition 2.3.b :** *Densité de probabilité*

On appelle *densité de probabilité* sur  $\mathbb{R}$  toute fonction  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$  vérifiant :

1.  $f$  est positive ou nulle ;
2.  $f$  est continue par morceaux sur  $\mathbb{R}$  ;
3. l'intégrale  $\int_{-\infty}^{+\infty} f(t) dt$  converge et vaut 1.

**Proposition-définition 2.3.c :** *Densité de probabilité d'une v.a.r. absolument continue*

Toute variable aléatoire absolument continue admet une fonction de densité de probabilité, qui est la dérivée de la fonction de répartition.

Démonstration :

Soit  $X$  une variable aléatoire absolument continue, de fonction de répartition  $F_X$ .

Pour tout  $x \in X(\Omega)$ , la densité de probabilité de  $X$  en  $x$ , notée  $f_X(x)$  est définie par :

$$f_X(x) = \lim_{h \rightarrow 0} \frac{1}{h} \mathbf{P}(x \leq X \leq x + h).$$

Comme  $\mathbf{P}(x \leq X \leq x + h) = F_X(x + h) - F_X(x)$  et que  $F_X$  est dérivable, on a bien  $f_X = F_X'$ .

**Définition 2.3.d : Loi d'une variable aléatoire réelle absolument continue**

La loi d'une variable aléatoire absolument continue  $X$  est définie par l'ensemble  $X(\Omega)$  et par sa fonction de densité.

Remarque : Deux variables aléatoires différentes, mais qui ont la même loi sont dites *identiquement distribuées* (ou *équiréparties*).

**Définition 2.3.e : Support de la fonction de densité**

Etant donnée une variable aléatoire absolument continue  $X$  de fonctions de densité et de répartition  $f$  et  $F$  respectivement. L'ensemble  $X(\Omega)$  (qui est donc infini non dénombrable) est appelé le *support* de la fonction de densité.

La fonction de densité s'annule en dehors de son support, c'est-à-dire sur  $\mathbb{R} \setminus X(\Omega)$  (ainsi que sur les éventuels points de discontinuité ou de non dérivabilité de  $F$ ).

On a :

$$\int_{-\infty}^{+\infty} f(t) dt = \int_{X(\Omega)} f(t) dt = 1.$$
**Proposition 2.3.f :**

Etant donnés deux réels  $a$  et  $b$  tels que  $a \leq b$ , on a :

$$\begin{aligned} 1^\circ) \mathbf{P}(a \leq X \leq b) &= \mathbf{P}(a \leq X < b) = \mathbf{P}(a < X \leq b) = \mathbf{P}(a < X < b) \\ &= F(b) - F(a) = \int_a^b f(t) dt ; \end{aligned}$$

$$2^\circ) \mathbf{P}(X \geq a) = \int_a^{+\infty} f(t) dt ;$$

$$3^\circ) \mathbf{P}(X \leq b) = \int_{-\infty}^b f(t) dt$$

**Définition 2.3.g : Événement presque impossible ; presque certain**

Un événement *presque impossible* est un événement non impossible (*i.e.* non vide), mais de probabilité nulle. Un événement *presque certain* est un événement non certain (*i.e.* strictement inclus dans  $\Omega$ ), mais de probabilité égale à 1.

**Proposition 2.3.h :**

Etant donnée une variable aléatoire absolument continue  $X$  et un réel  $x \in X(\Omega)$ , l'événement  $X = x$  est presque impossible.

Démonstration :

$$\forall x \in X(\Omega), \mathbf{P}(X = x) = \mathbf{P}(x \leq X \leq x) = F(x) - F(x^-) = F(x) - F(x) = 0, \text{ et } \{x\} \neq \emptyset.$$

**Exemples 2.3.i :**

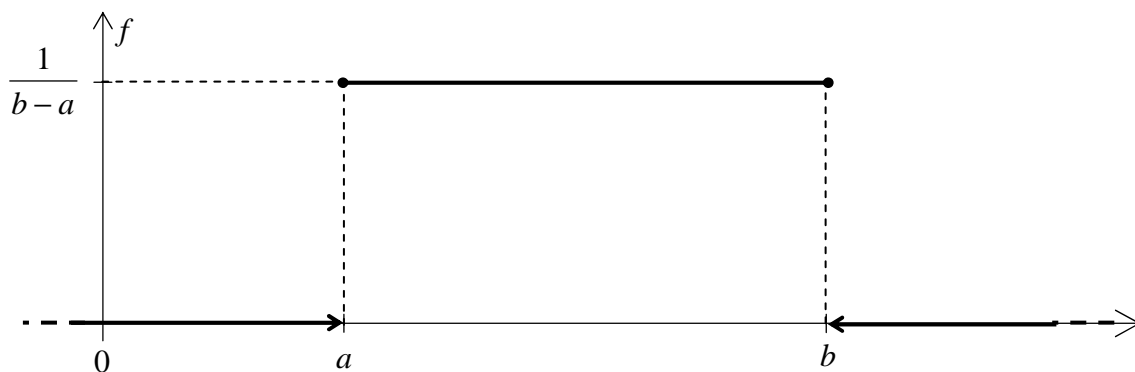
1°) Loi *uniforme* sur un segment  $[a ; b]$  de  $\mathbb{R}$ , non réduit à un singleton ( $a < b$ ) :

On a  $X(\Omega) = [a ; b]$ , et la fonction de densité est :

$$f: \mathbb{R} \rightarrow \mathbb{R}_+$$

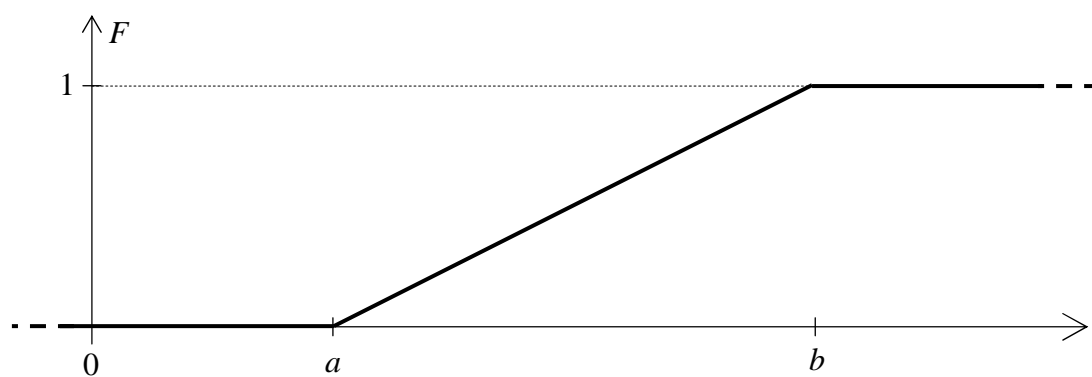
$$t \mapsto \frac{1}{b-a} \text{ si } t \in [a ; b], \text{ et } 0 \text{ sinon.}$$

Fonction de densité :



La fonction de répartition d'une loi uniforme sur  $[a ; b]$  est :

$$F : x \mapsto \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a < x < b \\ 1 & \text{si } x \geq b \end{cases}$$

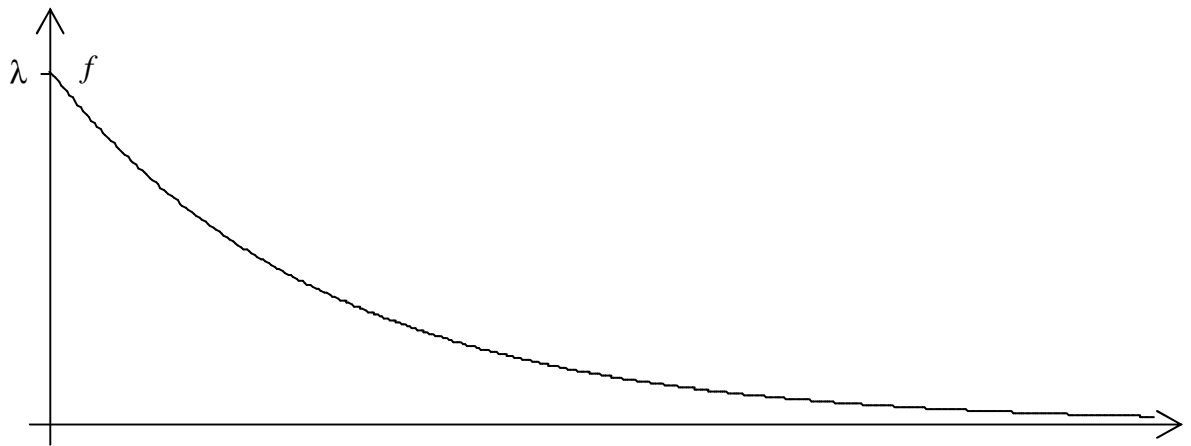


2°) Loi *exponentielle* de paramètre  $\lambda \in \mathbb{R}_+^*$ .

On a  $X(\Omega) = \mathbb{R}_+$ , et la fonction de densité est :

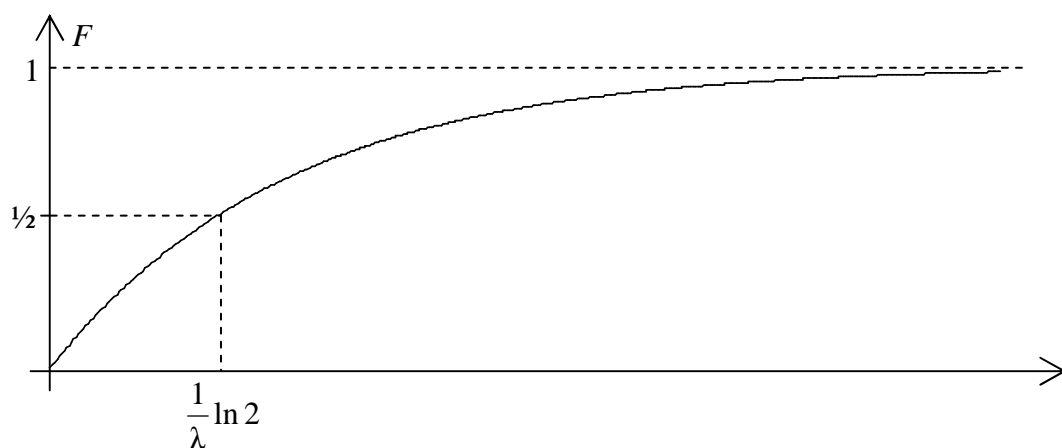
$$f: \mathbb{R} \rightarrow \mathbb{R}_+ \\ t \mapsto \lambda e^{-\lambda t} \text{ si } t \geq 0, \text{ et } 0 \text{ si } t < 0.$$

Fonction de densité :



La fonction de répartition d'une loi exponentielle de paramètre  $\lambda \in \mathbb{R}_+^*$ , est :

$$F: x \mapsto \int_{-\infty}^x f(t) dt = \begin{cases} 1 - e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}.$$



## 2.4. Lois usuelles

### 2.4.1. Lois discrètes

#### Définition 2.4.1.a : Loi certaine

Une variable aléatoire  $X$  suit une loi *certaine* de valeur  $a \in \mathbb{R}$ , et on note  $X \sim \mathcal{C}(a)$ , si et seulement si :

1.  $X(\Omega) = \{ a \}$  ;
2.  $\mathbf{P}(X = a) = 1$ .

#### Définition 2.4.1.b : Loi uniforme

Une variable aléatoire  $X$  suit une loi *uniforme* sur une partie  $\{ x_1, \dots, x_N \}$  de  $\mathbb{R}$ , et on note  $X \sim \mathcal{U}_{\{x_1, \dots, x_N\}}$ , si et seulement si :

1.  $X(\Omega) = \{ x_1, \dots, x_N \}$  ;
2.  $\mathbf{P}(X = x_1) = \dots = \mathbf{P}(X = x_N) = \frac{1}{N}$ .

#### Définition 2.4.1.c : Loi de Bernoulli

Une variable aléatoire  $X$  suit une loi de *Bernoulli* de paramètre  $p$ , et on note  $X \sim \mathcal{B}(p)$ , si et seulement si :

3.  $X(\Omega) = \{ 0, 1 \}$  ;
4.  $\mathbf{P}(X = 1) = p$  et  $\mathbf{P}(X = 0) = 1 - p = q$ .

#### Exemple 2.4.1.d :

Toute épreuve aléatoire dont le résultat est binaire (réussite ou échec, pile ou face, etc.) est appelée une épreuve de Bernoulli. La variable aléatoire qui vaut 1 si l'épreuve réussit et 0 si elle échoue suit une loi de Bernoulli dont le paramètre  $p$  est la probabilité de réussite de l'épreuve de Bernoulli "associée" à cette variable.

**Définition 2.4.1.e : Loi binomiale**

Une variable aléatoire  $X$  suit une loi *binomiale* de paramètres  $n$  et  $p$ , et on note  $X \sim \mathcal{B}(n, p)$ , si et seulement si :

1.  $X(\Omega) = \llbracket 0, n \rrbracket$  ;
2.  $\forall k \in X(\Omega), \mathbf{P}(X = k) = C_n^k p^k (1-p)^{n-k}$ .

Remarque : La loi binomiale de paramètres 1 et  $p$  est une loi de Bernoulli.

**Exemple 2.4.1.f :**

Si l'on répète  $n$  fois une épreuve de Bernoulli de paramètre  $p$ , la variable aléatoire qui compte le nombre de succès obtenus sur ces  $n$  épreuves suit une loi binomiale de paramètres  $n$  et  $p$ .

**Définition 2.4.1.g : Loi de Pascal ; loi géométrique**

Soient  $r \in \mathbb{N}^*$ , et  $p \in ]0 ; 1[$ . Une variable aléatoire  $X$  suit une loi de *Pascal* de paramètres  $r$  et  $p$ , et on note  $X \sim \mathcal{P}(r, p)$ , si et seulement si :

1.  $X(\Omega) = \llbracket r, +\infty \rrbracket$  ;
2.  $\forall k \in X(\Omega), \mathbf{P}(X = k) = C_{k-1}^{r-1} p^r q^{k-r}$ , avec  $q = 1 - p$ .

La loi de Pascal de paramètres 1 et  $p$  est appelée *loi géométrique*.

**Exemple 2.4.1.h :**

Si l'on répète une épreuve de Bernoulli de paramètre  $p$ , la variable aléatoire qui compte le nombre d'essais nécessaires pour atteindre la  $r^{\text{ème}}$  réussite, suit une loi de Pascal de paramètres  $r$  et  $p$ . Une v.a.r. qui suit une loi géométrique compte donc le nombre d'essais nécessaires pour atteindre la première réussite. On observe que ce nombre est évidemment minoré par  $r$ , mais n'est pas majoré.

**Définition 2.4.1.i : Loi binomiale négative**

Soient  $r \in \mathbb{N}^*$ , et  $p \in ]0 ; 1[$ . Une variable aléatoire  $X$  suit une loi *binomiale négative* de paramètres  $r$  et  $p$ , et on note  $X \sim \mathcal{Bn}(r, p)$ , si et seulement si :

1.  $X(\Omega) = \mathbb{N}$  ;
2.  $\forall k \in \mathbb{N}, \mathbf{P}(X = k) = C_{k+r-1}^r p^r q^k$ , avec  $q = 1 - p$ .



**Exemple 2.4.1.j :**

Si l'on répète une épreuve de Bernoulli de paramètre  $p$ , la variable aléatoire qui compte le nombre d'échecs qui précèdent la  $r^{\text{ème}}$  réussite, suit une loi binomiale négative de paramètres  $r$  et  $p$ . On observe d'une part, que ce nombre peut être nul, mais qu'il n'est pas majoré, et d'autre part, que si on lui ajoute  $r$ , on obtient le nombre d'essais nécessaires pour atteindre la  $r^{\text{ème}}$  réussite, qui suit une loi de Pascal de paramètres  $r$  et  $p$ . Donc, si  $N_{\text{échecs}} \in \mathcal{Bn}(r, p)$  et  $N_{\text{essais}} \in \mathcal{P}(r, p)$ , on a  $N_{\text{essais}} = N_{\text{échecs}} + r$ , d'où :

$$\mathbf{P}(N_{\text{échecs}} = k) = \mathbf{P}(N_{\text{essais}} = k + r) = C_{k+r-1}^{r-1} p^r q^k = C_{k+r-1}^k p^r q^k.$$

**Définition 2.4.1.k : Loi hypergéométrique**

Soient  $N, n \in \mathbb{N}^*$ , et  $p \in ]0; 1[$  tel que  $Np$  est entier. On note  $N_1 = Np$  et  $N_2 = N - N_1$ .

Une variable aléatoire  $X$  suit une loi *hypergéométrique* de paramètres  $N, n$  et  $p$ , et on note  $X \sim \mathcal{H}(N, n, p)$ , si et seulement si :

1.  $X(\Omega) = [\max(0, n - N_2), \min(n, N_1)]$  ;
2.  $\forall k \in [\max(0, n - N_2), \min(n, N_1)]$ ,  $\mathbf{P}(X = k) = \frac{C_{N_1}^k C_{N_2}^{n-k}}{C_N^n}$ .

**Exemple 2.4.1.l :**

Si l'on considère une population d'effectif  $N$  comportant une proportion  $p$  d'individus ayant un caractère donné, la v.a.r. qui donne le nombre de tels individus que l'on obtient en effectuant  $n$  tirages aléatoires sans remise dans cette population est distribuée suivant une loi hypergéométrique de paramètres  $N, n$  et  $p$ . On notera en particulier que la plus petite valeur prise par cette variable est bien  $n - N_2$  si  $n > N_2$  (et évidemment 0 sinon), et que la plus grande valeur est  $n$  si  $n \leq N_1$  et  $N_1$  sinon.

**Proposition 2.4.1.m : Convergence de la loi hypergéométrique vers une loi binomiale**

Etant donnée une suite de variables aléatoires hypergéométriques  $(X_N)_{N \in \mathbb{N}}$  de paramètres  $N, n$  et  $p$ , on a :

$$\lim_{N \rightarrow +\infty} \mathbf{P}(X_N = k) = C_n^k p^k q^{n-k}.$$

Ce résultat, qui est la *convergence en loi* de la loi hypergéométrique vers une loi binomiale, montre qu'une loi binomiale de paramètres  $n$  et  $p$  devient une bonne approximation d'une loi hypergéométrique de paramètres  $N, n$  et  $p$  quand  $N$  tend vers l'infini.

Démonstration :

Pour  $k$  fixé dans  $X_N(\Omega)$ , on a, en notant  $q = 1 - p$  :

$$\begin{aligned} \mathbf{P}(X_N = k) &= \frac{C_{N_1}^k C_{N_2}^{n-k}}{C_N^n} = \frac{(Np)(Np-1)\dots(Np-k+1) (Nq)(Nq-1)\dots(Nq-n+k+1)}{k! \frac{N(N-1)\dots(N-n+1)}{n!}} \\ &= \frac{n!}{k!(n-k)!} \frac{(Np)(Np-1)\dots(Np-k+1) (Nq)(Nq-1)\dots(Nq-n+k+1)}{N(N-1)\dots(N-n+1)}. \end{aligned}$$

Or  $\frac{(Np)(Np-1)\dots(Np-k+1) (Nq)(Nq-1)\dots(Nq-n+k+1)}{N(N-1)\dots(N-n+1)}$  est équivalent, quand  $N$  tend

vers l'infini à  $\frac{(Np)^k (Nq)^{n-k}}{N^n} = p^k q^{n-k}$ .

Donc  $\lim_{N \rightarrow +\infty} \mathbf{P}(X_N = k) = C_n^k p^k q^{n-k}$ .

**Définition 2.4.1.n :** *Loi de Poisson*

Une variable aléatoire  $X$  suit une loi de *Poisson* de paramètre  $\lambda > 0$ , et on note  $X \sim \mathcal{P}(\lambda)$ , si et seulement si :

1.  $X(\Omega) = \mathbb{N}$ ;
2.  $\forall k \in \mathbb{N}, \mathbf{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ .

**Proposition 2.4.1.o :** *Convergence de la loi binomiale vers une loi de Poisson*

Etant donnée une suite de variables aléatoires binomiales  $(X_n)_{n \in \mathbb{N}}$  de paramètres  $n$  et  $p_n$  où  $p_n$  est le terme général d'une suite numérique  $(p_n)_{n \in \mathbb{N}}$  telle que  $\lim_{n \rightarrow +\infty} n p_n = \lambda \in \mathbb{R}$ , on a :

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

Ce résultat, qui est la convergence en loi de la loi binomiale vers une loi de Poisson, montre qu'une loi de Poisson de paramètre  $\lambda$  devient une bonne approximation d'une loi binomiale de paramètres  $n$  et  $p$  quand  $n$  tend vers l'infini,  $p$  tend vers 0, et le produit  $n \times p$  tend vers la constante  $\lambda$ .

Démonstration :

$$\begin{aligned} \text{Pour } k \text{ fixé on a, } \forall n \geq k : \mathbf{P}(X_n = k) &= C_n^k (p_n)^k (1 - p_n)^{n-k} \\ &= \frac{n(n-1)\dots(n-k+1)}{k!} (p_n)^k (1 - p_n)^{n-k}. \end{aligned}$$

$$\text{Or } (1 - p_n)^{n-k} \underset{k \text{ fixé}}{\underset{n \rightarrow \infty}{\sim}} e^{n \ln(1-p_n)} \text{ et pour } p_n \in V(0) : n \ln(1-p_n) \underset{n \rightarrow \infty}{\sim} np_n \rightarrow \lambda.$$

$$\text{Donc } (1 - p_n)^{n-k} \underset{n \rightarrow \infty}{\rightarrow} e^{-\lambda}.$$

$$\text{D'autre part : } n(n-1)\dots(n-k+1)(p_n)^k \underset{k \text{ fixé}}{\underset{n \rightarrow \infty}{\sim}} n^k (p_n)^k = (np_n)^k \underset{n \rightarrow \infty}{\sim} \lambda^k.$$

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

## 2.4.2. Lois continues

**Définition 2.4.2.a :** *Loi uniforme* (cf. Exemple 4.3.i 1°))

Une variable aléatoire  $X$  suit une loi *uniforme* sur un segment  $[a; b]$  de  $\mathbb{R}$ , et on note  $X \sim \mathcal{U}_{[a; b]}$ , si et seulement si elle admet pour densité la fonction :

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R}_+ \\ t &\mapsto \frac{1}{b-a} \text{ si } t \in [a; b], \text{ et } 0 \text{ sinon.} \end{aligned}$$

$$\text{Sa fonction de répartition est } F : x \mapsto \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a < x < b \\ 1 & \text{si } x \geq b \end{cases}.$$

**Définition 2.4.2.b :** *Loi exponentielle de paramètre  $\lambda$*  (cf. Exemple 4.3.i 2°))

Une variable aléatoire suit une loi *exponentielle* de paramètre  $\lambda > 0$ , et on note  $X \sim \mathcal{E}(\lambda)$ , si et seulement si sa fonction de densité est :

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R}_+ \\ t &\mapsto \lambda e^{-\lambda t} \text{ si } t \geq 0, \text{ et } 0 \text{ si } t < 0. \end{aligned}$$

$$\text{Sa fonction de répartition est } F : x \mapsto \int_{-\infty}^x f(t) dt = \begin{cases} 1 - e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}.$$

**Définition 2.4.2.c : Loi normale centrée réduite**

Une variable aléatoire suit une loi *normale centrée réduite*, et on note  $X \sim \mathcal{N}(0, 1)$ , si et seulement si sa fonction de densité est :

$$f: \mathbb{R} \rightarrow \mathbb{R}$$
$$t \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

**Remarques :**

1°) La loi normale est également appelée loi de *Gauss*, loi *gaussienne*, voire parfois loi de *Laplace-Gauss*.

2°) On a bien :  $\int_{-\infty}^{+\infty} f(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} e^{-\frac{t^2}{2}} dt \underset{t=\sqrt{2}u}{=} \frac{2}{\sqrt{\pi}} \int_0^{+\infty} e^{-u^2} du = 1.$

(Rappel : L'intégrale  $\int_0^{+\infty} e^{-t^2} dt$  est appelée l'*intégrale de Gauss* et vaut  $\frac{\sqrt{\pi}}{2}$ )

3°) Comme le montre, la figure ci-dessous, la fonction de densité est paire. La distribution gaussienne est dite *symétrique*.

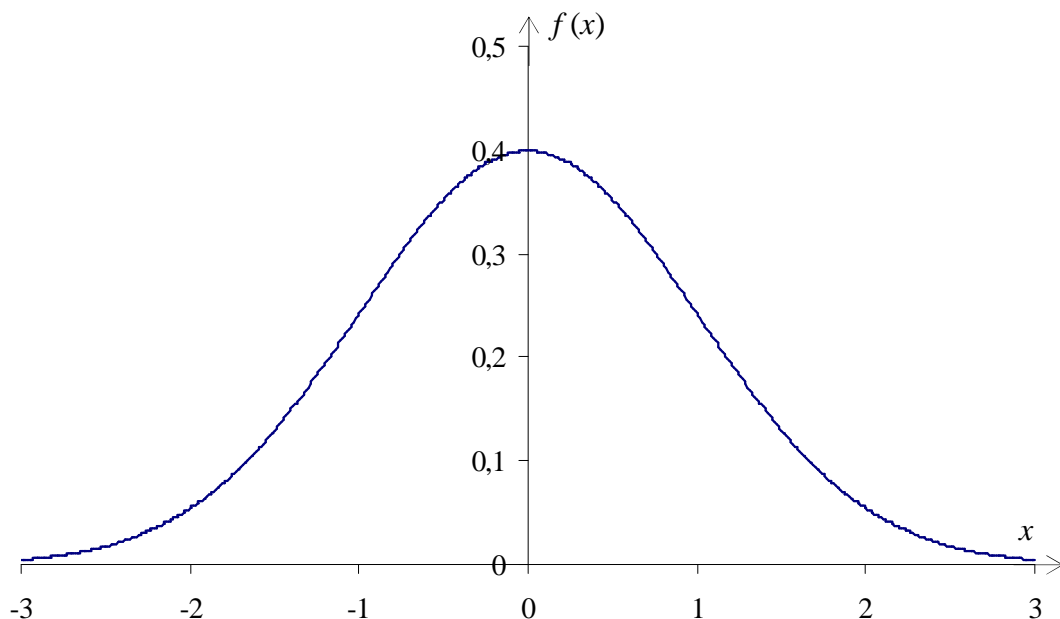


Figure 5.2.a : Fonction de densité de la loi normale centrée réduite

4°) La fonction de répartition de la loi normale centrée réduite  $F: x \mapsto \int_{-\infty}^x f(t) dt$  n'a pas d'expression formelle. Elle est donc calculée par intégration numérique, et la figure ci-après en donne une représentation graphique.

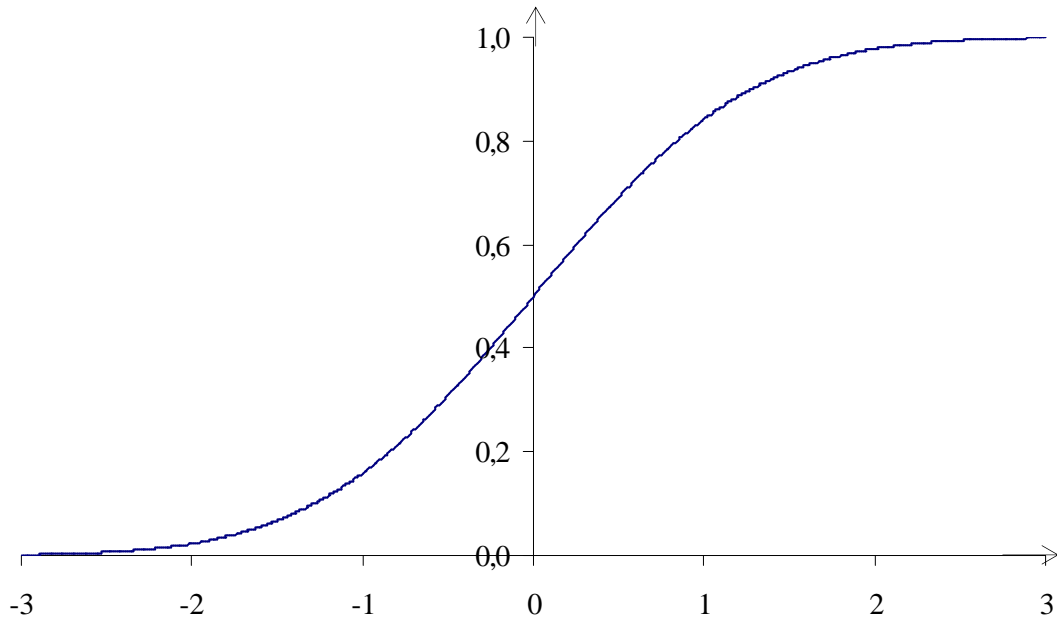


Figure 5.2.b : Fonction de répartition de la loi normale centrée réduite

Les valeurs de cette fonction de répartition doivent donc être lues dans une table. Cette table donne avec 4 décimales de précision, pour des valeurs positives de  $x$  croissant par pas de 0.01, les valeurs de  $F(x)$ . Les valeurs de  $F(x)$  pour des valeurs négatives de  $x$  se déduisent en utilisant la parité de la fonction de densité qui implique que  $\forall x \in \mathbb{R}, F(-x) = 1 - F(x)$ . On remarquera également que  $F(0) = 0,5$ .

Une autre table, dite table des *fractiles*, donne avec 4 décimales de précision, pour des valeurs de  $p$  comprises entre 0 et 1, et croissant par pas de 0.001, les valeurs  $x$  telles que  $F(x) = p$ . Ces valeurs sont appelées les *fractiles* de la loi normale centrée réduite (et  $x$  tel que  $F(x) = p$  est appelé le fractile d'ordre  $p$ ). On remarquera que, d'après la propriété 2., si  $F(x) = p$ , alors  $F(-x) = 1 - p$ , et donc que  $F^{-1}(p) = -F^{-1}(1 - p)$ . La table 2 ne donne donc que les valeurs absolues de  $F^{-1}(p)$ , le signe étant négatif pour les valeurs de  $p$  comprises entre 0 et 0.499 (obtenues avec la colonne de gauche et la ligne du bas), et positif pour les valeurs de  $p$  comprises entre 0.501 et 1 (obtenues avec la colonne de droite et la ligne du haut).

**Définition 2.4.2.d :** Loi normale de paramètres  $\mu$  et  $\sigma$

Une variable aléatoire suit une loi normale de paramètres  $\mu$  et  $\sigma$ , où  $\mu \in \mathbb{R}$  et  $\sigma \in \mathbb{R}_+^*$ , et on note  $X \sim \mathcal{N}(\mu, \sigma)$ , si et seulement si sa fonction de densité est :

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

$$t \mapsto \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

Remarques :

Les paramètres  $\mu$  et  $\sigma$  sont respectivement désignés par les termes d'*espérance* et d'*écart-type*. On utilise aussi le terme de *variance* pour désigner le carré  $\sigma^2$ . Tous ces termes seront définis dans la section 2.6.

**Proposition 2.4.2.e** (admise) : *Fonction affine d'une loi normale*

Toute fonction affine d'une loi normale est une loi normale. Plus précisément, étant donné  $(a, b) \in \mathbb{R}^* \times \mathbb{R}$  et  $X$  une loi normale de paramètres  $\mu$  et  $\sigma$ , alors la loi  $Y = aX + b$  est une loi normale de paramètres  $a\mu + b$  et  $|a|\sigma$  :

$$(a, b) \in \mathbb{R}^* \times \mathbb{R}, X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow Y = aX + b \sim \mathcal{N}(a\mu + b, a^2 \sigma^2).$$

**Proposition 2.4.2.f** (admise) : *Somme de deux lois normales indépendantes*

La somme de deux lois normales indépendantes de paramètres  $\mu_1, \sigma_1$  et  $\mu_2, \sigma_2$  respectivement, est une loi normale de paramètres  $\mu_1 + \mu_2$  et  $\sqrt{\sigma_1^2 + \sigma_2^2}$ .

**Définition 2.4.2.g** : *Loi log-normale*

Une variable aléatoire suit une loi *log-normale*, de paramètres  $\lambda$  et  $\zeta$ , et on note  $X \sim \mathcal{LN}(\lambda, \zeta)$ , si et seulement si sa fonction de densité est :

$$f: \mathbb{R}_+ \rightarrow \mathbb{R}$$
$$t \mapsto \frac{1}{\zeta \sqrt{2\pi}} \frac{1}{t} e^{-\frac{1}{2} \left( \frac{\ln t - \lambda}{\zeta} \right)^2}.$$

**Proposition 2.4.2.h** :

Le logarithme (népérien) d'une loi log-normale, de paramètres  $\lambda$  et  $\zeta$  est une loi normale de paramètres  $\lambda$  et  $\zeta$  :  $U \sim \mathcal{LN}(\lambda, \zeta) \Rightarrow X = \ln U \sim \mathcal{N}(\lambda, \zeta)$ .

Donc une loi log-normale est l'exponentielle d'une loi normale.

Démonstration :

Tout d'abord, comme  $U(\Omega) = \mathbb{R}_+, X(\Omega) = \mathbb{R}$ , et :

$$F_X(x) = \mathbf{P}(X \leq x) = \mathbf{P}(\ln U \leq x) = \mathbf{P}(U \leq e^x) \Rightarrow f_X(x) = e^x f_U(e^x) = \frac{1}{\zeta \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \lambda}{\zeta} \right)^2}.$$

**Définition 2.4.2.i : Fonction  $\Gamma$** 

On définit la fonction  $\Gamma$  par :  $\Gamma : x \mapsto \int_0^{+\infty} t^{x-1} e^{-t} dt$ .

**Propositions 2.4.2.j (admises) :**

1°)  $\Gamma(1) = 1$  ;

2°)  $\forall x > 0, \Gamma(x+1) = x \Gamma(x)$ , donc  $\forall k \in \mathbb{N}^*, \Gamma(k) = (k-1) !$

3°)  $\Gamma(1/2) = \sqrt{\pi}$ .

**Définition 2.4.2.k : Loi gamma de paramètre  $r$** 

Une variable aléatoire  $X$  suit une loi *gamma* de paramètre  $r > 0$ , et on note  $X \sim \Gamma(r)$  si elle admet pour densité la fonction :

$$f: \mathbb{R} \rightarrow \mathbb{R}_+$$

$$t \mapsto \frac{1}{\Gamma(r)} t^{r-1} e^{-t} \text{ si } t > 0, \text{ et } 0 \text{ sinon.}$$

**Remarque :**

Une loi gamma de paramètre  $r = 1$  est une loi exponentielle de paramètre  $\lambda = 1$ .

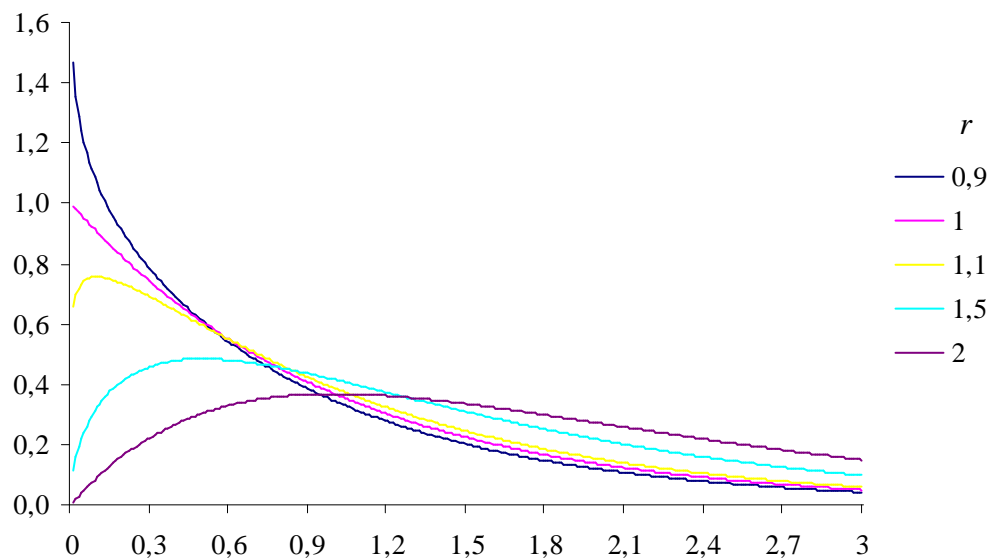


Figure 5.2.c :

Fonctions de densité de lois gammas pour quelques valeurs du paramètre  $r$ .

**Définition 2.4.2.1 :** *Loi gamma de paramètre  $\alpha$  et  $\beta$*

Une variable aléatoire  $X$  suit une loi gamma de paramètre  $\alpha$  et  $\beta$  ( $\alpha > 0$  et  $\beta > 0$ ), et on note  $X \sim \Gamma(\alpha, \beta)$  si elle admet pour densité la fonction :

$$f: \mathbb{R} \rightarrow \mathbb{R}_+ \\ t \mapsto \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \text{ si } t > 0, \text{ et } 0 \text{ sinon.}$$

**Propositions 2.4.2.m** (admises) :

$$1^\circ) X \sim \Gamma(\alpha, 1) \Leftrightarrow X \sim \Gamma(\alpha).$$

$$2^\circ) X \sim \Gamma(1, \lambda) \Leftrightarrow X \sim \mathcal{E}(\lambda).$$

$$3^\circ) X \sim \mathcal{N}(0, 1) \Rightarrow \frac{X^2}{2} \sim \Gamma\left(\frac{1}{2}, 1\right);$$

**Propositions 2.4.2.n** (admise) : *Somme de lois gammas*

La somme de lois gammas ayant la même valeur de leur second paramètre est une loi gamma.

$$\text{Plus précisément : } \forall i \in \llbracket 1, n \rrbracket, X_i \sim \Gamma(\alpha_i, \lambda) \Rightarrow \sum_{i=1}^n X_i \sim \Gamma\left(\sum_{i=1}^n \alpha_i, \lambda\right).$$

**Corollaire-définition 5.2.o** (admise) : *Somme de lois exponentielles ; Loi d'Erlang de paramètre  $n$*

Etant donné  $n \in \mathbb{N}^*$ , la somme de  $n$  lois exponentielles de paramètre  $\lambda$  est une loi gamma de

$$\text{paramètres } n \text{ et } \lambda : \quad \forall i \in \llbracket 1, n \rrbracket, X_i \sim \mathcal{E}(\lambda) \Rightarrow \sum_{i=1}^n X_i \sim \Gamma(n, \lambda).$$

La loi  $\Gamma(n, \lambda)$  est également appelée loi d'Erlang de paramètre  $n$ .

Sa densité est donnée par :

$$f: \mathbb{R} \rightarrow \mathbb{R}_+ \\ t \mapsto \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} \text{ si } t > 0, \text{ et } 0 \text{ sinon.}$$



## 2.5. Fonctions d'une variable aléatoire réelle

### 2.5.1. Variables aléatoires réelles discrètes

**Définition 2.5.1.a :** *Fonction d'une variable aléatoire réelle discrète*

Soit  $X$  une v.a.r. discrète et  $\varphi$  une fonction à valeurs réelles définie en tout point de  $X(\Omega) = \{x_k : k \in K\}$ . On définit alors  $\varphi(X)$  par :

$$\begin{aligned}\varphi(X) : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto \varphi(X(\omega)),\end{aligned}$$

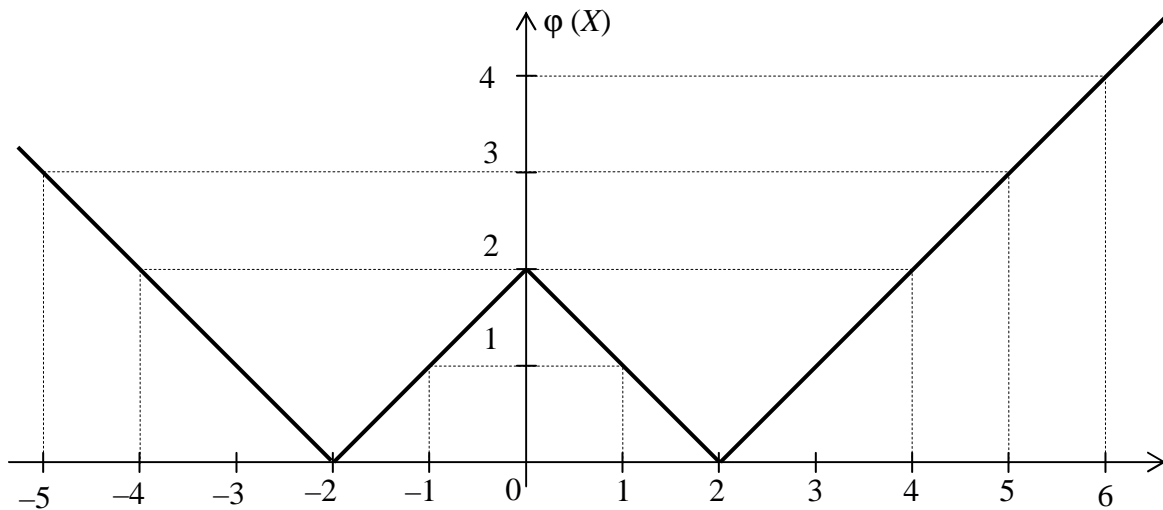
par sa loi :

1.  $[\varphi(X)](\Omega) = \varphi[X(\Omega)] = \varphi[\{x_k : k \in K\}] = \{y : \exists k \in K : \varphi(x_k) = y\}$  ;
2.  $\mathbf{P}(\varphi(X) = y) = \sum_{k \in K : \varphi(x_k) = y} \mathbf{P}(X = x_k)$ .

$\varphi(X)$  est donc une v.a.r. discrète.

Notation : On pourra noter  $\varphi : X \mapsto \varphi(X)$ .

**Exemple 2.5.1.b :**  $\varphi : X \mapsto ||X| - 2|$ , avec  $X(\Omega) = \llbracket -5, 6 \rrbracket$ .



Si l'on note  $Y = \varphi(X)$ , la loi de  $Y$  est donnée par :

1.  $Y(\Omega) = \llbracket 0, 4 \rrbracket$  ;
2.  $\mathbf{P}(Y = 0) = \mathbf{P}(X \in \{-2, 2\})$ ,  $\mathbf{P}(Y = 1) = \mathbf{P}(X \in \{-3, -1, 1, 3\})$ ,  
 $\mathbf{P}(Y = 2) = \mathbf{P}(X \in \{-4, 0, 4\})$ ,  $\mathbf{P}(Y = 3) = \mathbf{P}(X \in \{-5, 5\})$ ,  
 $\mathbf{P}(Y = 4) = \mathbf{P}(X = 6)$ .

### 2.5.2. Variables aléatoires réelles absolument continues

**Définition 2.5.2.a :** *Fonction d'une variable aléatoire réelle absolument continue*

Soit  $X$  une v.a.r. absolument continue admettant une densité  $f$ , et  $\varphi$  une fonction à valeurs réelles définie sur  $X(\Omega)$ , ou sur un sous ensemble  $B$  de  $X(\Omega)$  tel que  $\mathbf{P}(X \notin B) = 0$ . Alors  $\varphi(X)$ , définie par :

$$\begin{aligned}\varphi(X) : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto \varphi(X(\omega)),\end{aligned}$$

est une v.a.r. si et seulement si :

$$\forall x \in \mathbb{R}, \mathbf{P}(\{\omega \in \Omega : \varphi(X(\omega)) \leq x\}) = \mathbf{P}(\varphi(X) \leq x) \text{ existe.}$$

**Proposition 2.5.2.b :**

Une condition suffisante pour que  $Y = \varphi(X)$  soit une v.a.r. absolument continue est que sa fonction de répartition  $F_Y$  soit continûment dérivable par morceaux sur  $\mathbb{R}$ .

On a alors  $f_Y(x) = F_Y'(x)$  si  $F_Y$  est dérivable en  $x$ , et  $f_Y(x) = 0$  sinon.

**Exemples 2.5.2.c :**

1°)  $\varphi : X \mapsto Y = |X|$ , avec  $X(\Omega) = \mathbb{R}$ ;

On a  $Y(\Omega) = \mathbb{R}_+$ ;

Pour  $x \geq 0$ ,  $F_Y(x) = \mathbf{P}(|X| \leq x) = \mathbf{P}(-x \leq X \leq x) = F_X(x) - F_X(-x)$ .

$F_X$  est continue sur  $\mathbb{R}$ , donc  $F_Y$  est continue sur  $] -\infty ; 0 [$  et sur  $] 0 ; +\infty [$ .

De plus  $\lim_{x \rightarrow 0^-} F_Y(x) = \lim_{x \rightarrow 0^+} F_Y(x) = 0$ , donc  $F_Y$  est continue sur  $\mathbb{R}$ .

Enfin  $F_Y$  est continûment dérivable sur  $\mathbb{R}^*$ , donc la densité  $f_Y$  de  $Y$  est :

$$f_Y : x \mapsto f_Y(x) = 0 \text{ si } x \leq 0, \text{ et } f_Y(x) = f_X(x) + f_X(-x) \text{ sinon.}$$

2°)  $\varphi : X \mapsto Y = \frac{1}{X}$ , avec  $X(\Omega) = \mathbb{R}$ ;

$\varphi : x \mapsto \frac{1}{x}$  est définie sur le sous ensemble  $B = \mathbb{R}^*$  de  $X(\Omega)$ , et l'on a bien :

$$\mathbf{P}(X \notin B) = \mathbf{P}(X = 0) = 0,$$

ce qui permet de “négliger” l'événement “ $X = 0$ ”.

On a  $Y(\Omega) = \mathbb{R}^*$ .

$$\text{Si } x < 0 : F_Y(x) = \mathbf{P} \left( \frac{1}{X} \leq x \right) = \mathbf{P} \left( \frac{1}{x} \leq X < 0 \right) = F_X(0) - F_X \left( \frac{1}{x} \right);$$

$$\text{Si } x > 0 : F_Y(x) = \mathbf{P} \left( X < 0 \text{ ou } X \geq \frac{1}{x} \right) = \mathbf{P}(X < 0) + \mathbf{P} \left( X \geq \frac{1}{x} \right) = F_X(0) + 1 - F_X \left( \frac{1}{x} \right);$$

$$\text{Si } x = 0 : F_Y(x) = \mathbf{P} \left( \frac{1}{X} \leq 0 \right) = \mathbf{P}(X < 0) = F_X(0).$$

$F_X$  est continue sur  $\mathbb{R}$ , donc  $F_Y$  est continue sur  $] -\infty ; 0 [$  et sur  $] 0 ; +\infty [$ .

De plus  $\lim_{x \rightarrow 0^-} F_Y(x) = \lim_{x \rightarrow 0^+} F_Y(x) = F_X(0)$ , donc  $F_Y$  est continue sur  $\mathbb{R}$ .

Enfin  $F_Y$  est continûment dérivable sur  $\mathbb{R}^*$ , donc la densité  $f_Y$  de  $Y$  est :

$$f_Y : x \mapsto f_Y(x) = \frac{1}{x^2} f_X \left( \frac{1}{x} \right).$$

### Proposition 2.5.2.d : Application à la loi normale

Toute fonction affine d'une loi normale est une loi normale. Plus précisément, étant donnés  $(a, b) \in \mathbb{R}^* \times \mathbb{R}$  et  $X$  une loi normale de paramètres  $\mu$  et  $\sigma$ , alors la loi  $Y = aX + b$  est une loi normale de paramètres  $a\mu + b$  et  $|a|\sigma$  :

$$(a, b) \in \mathbb{R}^* \times \mathbb{R}, X \sim \mathcal{N}(\mu, \sigma) \Rightarrow Y = aX + b \sim \mathcal{N}(a\mu + b, |a|\sigma).$$

Démonstration :

$$X(\Omega) = \mathbb{R} \Rightarrow Y(\Omega) = \mathbb{R}.$$

1<sup>er</sup> cas :  $a > 0$

$$\forall x \in \mathbb{R}, F_Y(x) = \mathbf{P}(aX + b \leq x) = \mathbf{P}\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right), \text{ donc } F_Y \text{ est continûment}$$

$$\text{dérivable sur } \mathbb{R} \text{ et } f_Y(x) = \frac{1}{a} f_X\left(\frac{x-b}{a}\right);$$

2<sup>ème</sup> cas :  $a < 0$

$$\forall x \in \mathbb{R}, F_Y(x) = \mathbf{P}(aX + b \leq x) = \mathbf{P}\left(X \geq \frac{x-b}{a}\right) = 1 - F_X\left(\frac{x-b}{a}\right), \text{ donc } F_Y \text{ est continûment}$$

$$\text{dérivable sur } \mathbb{R} \text{ et } f_Y(x) = -\frac{1}{a} f_X\left(\frac{x-b}{a}\right).$$

$$\forall a \in \mathbb{R}^*, \text{ on a } f_Y(x) = \frac{1}{|a|} f_X\left(\frac{x-b}{a}\right) = \frac{1}{|a|\sigma\sqrt{2\pi}} e^{-\frac{\left(\frac{x-b}{a} - \mu\right)^2}{2\sigma^2}} = \frac{1}{|a|\sigma\sqrt{2\pi}} e^{-\frac{(x-(a\mu+b))^2}{2(|a|\sigma)^2}},$$

donc  $Y \in \mathcal{N}(a\mu + b, |a|\sigma)$ .

### Corollaire 2.5.2.e : Centrage et réduction d'une loi normale

Si  $X$  suit une loi normale de paramètres  $\mu$  et  $\sigma$ , alors la loi  $X^* = \frac{X-\mu}{\sigma}$  suit une loi gaussienne centrée réduite.

Démonstration :

$X - \mu \sim \mathcal{N}(0, \sigma)$ , puisque avec  $a = 1$  et  $b = -\mu$  :  $\mu' = a\mu + b = 0$  et  $\sigma' = |a| \sigma = \sigma$ .

Enfin,  $X^* \sim \mathcal{N}(0, 1)$ , puisque avec  $a = 1/\sigma$  et  $b = 0$ ,  $a\mu' + b = 0$  et  $|a| \sigma' = |a| \sigma = 1$ .

### 2.5.3. Indépendance de deux fonctions de variables aléatoires

#### Proposition 2.5.3.a :

Soient  $X_1$  et  $X_2$  deux variables aléatoires mesurables indépendantes, et soient  $T_1$  et  $T_2$  des tribus définies sur  $X_1(\Omega)$  et sur  $X_2(\Omega)$ . Soient  $\varphi_1$  et  $\varphi_2$  deux fonctions réelles définies respectivement sur  $X_1(\Omega)$  et  $X_2(\Omega)$ . Si  $\varphi_1$  (respectivement  $\varphi_2$ ) sont telles que l'image réciproque de tout borélien de  $\mathbb{R}$  est un événement de  $T_1$  (respectivement de  $T_2$ ), alors les v.a.r  $\varphi_1(X_1)$  et  $\varphi_2(X_2)$  sont indépendantes.

Démonstration :

Soient  $B_1'$  et  $B_2'$  deux boréliens de  $\mathbb{R}$ , et soient  $B_1$  et  $B_2$  leurs images réciproques respectives par  $\varphi_1$  et  $\varphi_2$  :  $B_1 = \{ x \in X_1(\Omega) \text{ t.q. } \varphi_1(x) \in B_1' \}$  et  $B_2 = \{ x \in X_2(\Omega) \text{ t.q. } \varphi_2(x) \in B_2' \}$ .

En notant  $Y_1 = \varphi_1(X_1)$  et  $Y_2 = \varphi_2(X_2)$  on a alors :

$$Y_1^{-1}(B_1') = \{ \omega \in \Omega \text{ t.q. } \varphi_1(X_1(\omega)) \in B_1' \} = \{ \omega \in \Omega \text{ t.q. } X_1(\omega) \in B_1 \} = X_1^{-1}(B_1),$$

et de même  $Y_2^{-1}(B_2') = X_2^{-1}(B_2)$ .

On en déduit, d'après la définition 1.c que les variables  $Y_1$  et  $Y_2$  sont indépendantes.

### 3. Moments d'une variable aléatoire réelle

**Proposition 3.a :** *Formule du binôme négatif*

Cette formule, très utile pour les calculs présentés dans ce chapitre, établit que pour tout  $x \in \mathbb{R}$  tel que  $|x| < 1$ , on a :

$$\sum_{k=r}^{+\infty} \binom{k}{r} x^{k-r} = \frac{1}{(1-x)^{r+1}}$$

**Propositions 3.b :** *Quelques formules utiles aux calculs des moments d'une v.a.r*

1°) En appliquant la formule du binôme négatif avec  $r = 1$ , on obtient :

$$\sum_{k=1}^{+\infty} k x^k = \frac{x}{(1-x)^2}$$

2°) En appliquant la formule du binôme négatif avec  $r = 2$ , on obtient

$$\sum_{k=1}^{+\infty} k^2 x^k = \frac{x(1+x)}{(1-x)^3}$$

Démonstration :

$$\sum_{k=1}^{+\infty} k^2 x^k = x^2 \sum_{k=2}^{+\infty} k(k-1) x^{k-2} + \sum_{k=1}^{+\infty} k x^k = \frac{2x^2}{(1-x)^3} + \frac{x}{(1-x)^2} = \frac{x(1+x)}{(1-x)^3}.$$

#### 3.1. Moments d'une variable aléatoire réelle discrète

**Définition 3.1.a :** *Moment d'ordre  $r$*

Soit une v.a.r. discrète telle que  $X(\Omega) = \{x_k, k \in K\}$ , et  $r \in \mathbb{N}$ .

Si la série de terme général  $x_k^r \times \mathbf{P}(X = x_k)$  converge, le *moment d'ordre  $r$*  de  $X$  est

$$m_r[X] = \sum_{k \in K} x_k^r \times \mathbf{P}(X = x_k).$$

Remarque :

Si la série de terme général  $x_k^r \times \mathbf{P}(X = x_k)$  ne converge pas, le *moment d'ordre  $r$*  de  $X$  n'existe pas.

**Définition 3.1.b : Espérance mathématique**

L'espérance mathématique d'une v.a.r. est son moment d'ordre 1.

Pour une v.a.r discrète elle est égale à :

$$m_1 [ X ] = \mathbf{E} [ X ] = \sum_{k \in K} x_k \times \mathbf{P} ( X = x_k ) .$$

**Exemples** (à partir des exemples des paragraphes 3.1.1 et 3.1.2 et en prenant  $p = 1/2$ ) :

$$1^\circ) \Omega = \{ \text{pile, face} \}; X(\Omega) = \{ 1, -1 \}; \mathbf{P} ( X = 1 ) = \mathbf{P} ( X = -1 ) = \frac{1}{2};$$

$$\text{Alors } \mathbf{E} [ X ] = -1 \times \frac{1}{2} + 1 \times \frac{1}{2} = 0 .$$

$$2^\circ) \Omega = \{ \text{pile, face} \}^N; X(\Omega) = \llbracket 0, N \rrbracket; \forall k \in X(\Omega), \mathbf{P} ( X = k ) = C_N^k p^k (1-p)^{N-k};$$

$$\text{Avec } N = 4, \mathbf{P} ( X = 0 ) = \mathbf{P} ( X = 4 ) = \frac{1}{16}, \mathbf{P} ( X = 1 ) = \mathbf{P} ( X = 3 ) = \frac{4}{16} \text{ et } \mathbf{P} ( X = 2 ) = \frac{6}{16} .$$

$$\text{Alors } \mathbf{E} [ X ] = 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16} = \frac{32}{16} = 2 .$$

$$3^\circ) \Omega = \bigcup_{n \geq 0} \{ \text{face} \}^n \times \{ \text{pile} \}; X(\Omega) = \mathbb{N}^* .$$

$$\forall k \in X(\Omega), \mathbf{P} ( X = k ) = (1-p)^{k-1} p, \text{ soit, avec } p = \frac{1}{2}, \mathbf{P} ( X = k ) = \frac{1}{2^k} .$$

$$\text{Alors } \mathbf{E} [ X ] = \sum_{k=1}^{+\infty} k \left( \frac{1}{2} \right)^k = \frac{1/2}{(1-1/2)^2} = 2 .$$

**Cas particuliers :**

1°) Si  $\mathbf{E} [ X ] = 0$ , la variable  $X$  est dite *centrée* ;

2°) Si la série de terme général  $x_k \times \mathbf{P} ( X = x_k )$  ne converge pas, alors  $\mathbf{E} [ X ]$  n'existe pas.

Ceci est par exemple le cas pour  $X$  tel que  $X(\Omega) = \mathbb{N}^*$  et  $\mathbf{P} ( X = k ) = \frac{6}{\pi^2} \times \frac{1}{k^2}$  ;

3°) Pour une variable certaine telle que  $X(\Omega) = \{a\}$ , alors  $\mathbf{E} [ X ] = a$ .

**Définition 3.1.c : Moment centré d'ordre  $r$** 

Soit une v.a.r discrète telle que  $X(\Omega) = \{ x_k, k \in K \}$ , telle que son espérance mathématique existe et est égale à  $m$ , et soit  $r \in \mathbb{N}$ .

Si la série de terme général  $(x_k - m)^r \times \mathbf{P} ( X = x_k )$  converge, le *moment centré d'ordre  $r$*  de  $X$

$$\text{est : } \mu_r [ X ] = \sum_{k \in K} (x_k - m)^r \times \mathbf{P} ( X = x_k ) .$$

**Définition 3.1.d : Variance**

La *variance* d'une v.a.r est son moment centré d'ordre 2.

Pour une v.a.r discrète elle est égale à :

$$\mu_2 [X] = \mathbf{V} [X] = \sum_{k \in K} (x_k - m)^2 \times \mathbf{P} (X = x_k).$$

**Remarques :**

1°) Si  $\mathbf{V} [X] = 1$ , la variable  $X$  est dite *réduite*.

2°) Si la série de terme général  $(x_k - m)^r \times \mathbf{P} (X = x_k)$  ne converge, alors  $\mathbf{V} [X]$  n'existe pas.

Ceci est par exemple le cas pour  $X$  tel que  $X(\Omega) = \mathbb{N}^*$  et  $\mathbf{P} (X = k) = \frac{2k}{k^4 + k^2 + 1}$ .

3°) Pour une variable certaine,  $\mathbf{V} [X] = 0$ .

**Exemples 3.1.e** (reprenant l'exemple 4.2.b avec  $p = 1/2$ ) :

1°)  $\Omega = \{ \text{pile, face} \}$ ;  $X(\Omega) = \{ 1, -1 \}$ ;  $\mathbf{P} (X = 1) = \mathbf{P} (X = -1) = \frac{1}{2}$ ;  $\mathbf{E} [X] = 0$ ;

Alors  $\mathbf{V} [X] = 1 \times \frac{1}{2} + 1 \times \frac{1}{2} = 1$ .

2°)  $\Omega = \{ \text{pile, face} \}^N$ ;  $X(\Omega) = \llbracket 0, N \rrbracket$ ;  $\forall k \in X(\Omega)$ ,  $\mathbf{P} (X = k) = C_N^k p^k (1-p)^{N-k}$ ;

Avec  $N = 4$ ,  $\mathbf{E} [X] = 2$ ;

Alors  $\mathbf{V} [X] = 4 \times \frac{1}{16} + 1 \times \frac{4}{16} + 0 \times \frac{6}{16} + 1 \times \frac{4}{16} + 4 \times \frac{1}{16} = \frac{16}{16} = 1$ .

3°)  $\Omega = \bigcup_{n \geq 0} \{ \text{face} \}^n \times \{ \text{pile} \}$ ;  $X(\Omega) = \mathbb{N}^*$ .

$\forall k \in X(\Omega)$ ,  $\mathbf{P} (X = k) = (1-p)^{k-1} p$ , soit, avec  $p = \frac{1}{2}$ ,  $\mathbf{P} (X = k) = \frac{1}{2^k}$ .

$$\begin{aligned} \mathbf{E} [X] = 2, \text{ donc } \mathbf{V} [X] &= \sum_{k=1}^{+\infty} (k-2)^2 \left( \frac{1}{2} \right)^k = \sum_{k=1}^{+\infty} k^2 \left( \frac{1}{2} \right)^k - 4 \sum_{k=1}^{+\infty} k \left( \frac{1}{2} \right)^k + 4 \sum_{k=1}^{+\infty} \left( \frac{1}{2} \right)^k \\ &= \frac{1/2 (1+1/2)}{(1-1/2)^3} - 4 \frac{1/2}{(1-1/2)^2} + 4 \\ &= 6 - 8 + 4 = 2. \end{aligned}$$

**Définition 3.1.f : Ecart-type ; Coefficient de variation**

L'*écart-type* d'une v.a.r est  $\sigma [X] = \sqrt{\mathbf{V} [X]}$ . Le *coefficient de variation* est le rapport de l'écart-type sur l'espérance mathématique.

### 3.2. Moments d'une variable aléatoire réelle absolument continue

#### Définition 3.2.a : Moment d'ordre $r$

Soit une v.a.r absolument continue de densité  $f$  sur  $X(\Omega)$ , et  $r \in \mathbb{N}$ .

Si l'intégrale  $\int_{X(\Omega)} t^r f(t) dt$  est absolument convergente, i.e. si l'application  $t \mapsto t^r f(t)$  est intégrable sur  $X(\Omega)$ , le *moment d'ordre  $r$*  de  $X$  est :

$$m_r [X] = \int_{X(\Omega)} t^r f(t) dt .$$

#### Remarques :

Si l'intégrale  $\int_{X(\Omega)} t^r f(t) dt$  ne converge pas absolument, alors le *moment d'ordre  $r$*  de  $X$  n'existe pas.

#### Définition 3.2.b : Espérance mathématique

L'*espérance mathématique* d'une v.a.r est son moment d'ordre 1.

Pour une variable aléatoire absolument continue elle est égale à :

$$m_1 [X] = \mathbf{E} [X] = \int_{X(\Omega)} t f(t) dt .$$

#### Remarques :

1°) Si  $\mathbf{E} [X] = 0$ , la variable  $X$  est dite *centrée* ;

2°) Si l'intégrale  $\int_{X(\Omega)} t f(t) dt$  ne converge pas absolument, alors  $\mathbf{E} [X]$  n'existe pas. Ceci est

par exemple le cas pour la loi de Cauchy, de densité  $t \mapsto \frac{1}{\pi(1+t^2)}$  sur  $\mathbb{R}$ , car  $\lim_{A \rightarrow +\infty} \int_0^A \frac{t dt}{\pi(1+t^2)} =$

$\lim_{A \rightarrow +\infty} \frac{1}{2\pi} \ln(1+A^2) = +\infty$ , et donc l'intégrale  $\int_{-\infty}^{+\infty} t f(t) dt$  ne peut pas converger absolument.

#### Exemples 3.2.c :

1°) Si  $X$  est une loi uniforme sur  $[a ; b]$  :

$$\mathbf{E} [X] = \frac{1}{b-a} \int_a^b t dt = \frac{a+b}{2} ;$$

2°) Si  $X$  est une loi exponentielle de paramètre  $\lambda$  :

$$\mathbf{E} [X] = \int_0^{+\infty} \lambda t e^{-\lambda t} dt = \left[ -t e^{-\lambda t} \right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda t} dt = \frac{1}{\lambda} .$$



**Définition 3.2.d : Moment centré d'ordre  $r$** 

Soit une v.a.r absolument continue de densité  $f$  sur  $X(\Omega)$ , telle que son espérance mathématique existe et est égale à  $m$ , et soit  $r \in \mathbb{N}$ .

Si l'intégrale  $\int_{X(\Omega)} (t-m)^r f(t) dt$  est absolument convergente, le *moment centré d'ordre  $r$*  de  $X$

est :

$$\mu_r [X] = \int_{X(\Omega)} (t-m)^r f(t) dt .$$

**Définition 3.2.e : Variance**

La *variance* d'une v.a.r est son moment centré d'ordre 2.

Pour une variable aléatoire absolument continue elle est égale à :

$$\mu_2 [X] = V [X] = \int_{X(\Omega)} (t-m)^2 f(t) dt .$$

**Remarques :**

1°) Si  $V [X] = 1$ , la variable  $X$  est dite *réduite* ;

2°) Si  $\int_{X(\Omega)} (t-m)^2 f(t) dt$  n'est pas absolument convergente, alors  $V [X]$  n'existe pas.

Ceci est par exemple le cas pour la loi de Pareto de densité  $t \mapsto \frac{\alpha}{a} \left(\frac{a}{t}\right)^{\alpha+1}$  sur  $[a ; +\infty[$ , pour

$\alpha = 2$  et  $a = 1$ .

3°) Pour une variable certaine,  $V [X] = 0$ .

**Exemples 3.2.f :**

1°) Si  $X$  est une loi uniforme sur  $[a ; b]$  :  $E [X] = m = \frac{a+b}{2}$  ;

$$V [X] = \frac{1}{b-a} \int_a^b (t-m)^2 dt = \frac{1}{b-a} \int_{\frac{a-b}{2}}^{\frac{b-a}{2}} t^2 dt = \frac{2}{b-a} \left[ \frac{t^3}{3} \right]_0^{\frac{b-a}{2}} = \frac{(b-a)^2}{12} .$$

2°) Si  $X$  est une loi exponentielle de paramètre  $\lambda$  :  $E [X] = m = \frac{1}{\lambda}$  ;

$$\begin{aligned} V [X] &= \int_0^{+\infty} \left(t - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda t} dt = \int_0^{+\infty} t^2 \lambda e^{-\lambda t} dt - \frac{2}{\lambda} \int_0^{+\infty} t \lambda e^{-\lambda t} dt + \frac{1}{\lambda^2} \int_0^{+\infty} \lambda e^{-\lambda t} dt \\ &= \int_0^{+\infty} t^2 \lambda e^{-\lambda t} dt - \frac{1}{\lambda^2} \text{ et } \int_0^{+\infty} t^2 \lambda e^{-\lambda t} dt = \left[ -t^2 e^{-\lambda t} \right]_0^{+\infty} + 2 \int_0^{+\infty} t e^{-\lambda t} dt = \frac{2}{\lambda^2} , \text{ d'où } V [X] = \frac{1}{\lambda^2} . \end{aligned}$$

**Définition 3.2.g : Ecart-type ; Coefficient de variation**

L'écart-type d'une v.a.r est  $\sigma [ X ] = \sqrt{V [ X ]}$ . Le coefficient de variation est le rapport de l'écart-type sur l'espérance mathématique.

**3.3. Moments d'une fonction d'une variable aléatoire réelle****Proposition 3.3.a :**

Soient  $X$  une v.a.r discrète telle que  $X(\Omega) = \{ x_k : k \in K \}$ , et  $\varphi$  une fonction à valeurs réelles définie en tout point  $x_k$ . Alors l'espérance mathématique de  $\varphi(X)$  existe si et seulement si la série de terme général  $\varphi(x_k) \times P(X = x_k)$  converge, et elle est égale à :

$$E[\varphi(X)] = \sum_{k \in K} \varphi(x_k) P(X = x_k).$$

Démonstration :

On a  $X(\Omega) = \{ x_k : k \in K \}$  et  $[\varphi(X)](\Omega) = \{ y_l : l \in L \}$ . Alors :

$$\begin{aligned} E[\varphi(X)] &= \sum_{l \in L} y_l P(\varphi(X) = y_l) \\ &= \sum_{l \in L} y_l \left( \sum_{k \in K_l} P(X = x_k) \right), \text{ avec } K_l = \{ k \in K : \varphi(x_k) = y_l \} \subset K, \text{ et } \bigcup_{l \in L} K_l = K \\ &= \sum_{l \in L} \sum_{k \in K_l} \varphi(x_k) P(X = x_k) = \sum_{k \in \bigcup_{l \in L} K_l} \varphi(x_k) P(X = x_k) = \sum_{k \in K} \varphi(x_k) P(X = x_k). \end{aligned}$$

**Proposition 3.3.b (admise) :**

Soient  $X$  une v.a.r absolument continue de densité  $f$ , et  $\varphi$  une fonction à valeurs réelles définie sur  $X(\Omega)$  (ou sur un sous ensemble  $B$  de  $X(\Omega)$  tel que  $P(X \notin B) = 0$ ).

Alors l'espérance de  $\varphi(X)$  est

$$E[\varphi(X)] = \int_{X(\Omega)} \varphi(x) f(x) dx,$$

si cette intégrale est absolument convergente, et n'existe pas sinon.

Conséquences :

On peut donc écrire  $m_r[X] = E[X^r]$ ,  $\mu_r[X] = E[(X - E[X])^r]$ , et donc, en particulier :

$$V[X] = E[(X - E[X])^2].$$

### 3.4. Propriétés des moments d'une variable aléatoire réelle

#### **Proposition 3.4.a :** *Linéarité de l'espérance*

Soient  $X$  une v.a.r, et  $a$  et  $b$  deux nombres réels. On a alors :

$$\mathbf{E} [ aX + b ] = a \mathbf{E} [ X ] + b .$$

Démonstration :

Si  $X$  est une variable discrète telle que  $X(\Omega) = \{ x_k : k \in K \}$ , il s'agit d'une application de la proposition 7.3.a :

$$\begin{aligned} \mathbf{E} [ aX + b ] &= \sum_{k \in K} (a x_k + b) \mathbf{P} ( X = x_k ) \\ &= a \sum_{k \in K} x_k \mathbf{P} ( X = x_k ) + b \sum_{k \in K} \mathbf{P} ( X = x_k ) = a \mathbf{E} [ X ] + b . \end{aligned}$$

Si  $X$  est une variable absolument continue de densité  $f$ , il s'agit d'une application de la proposition 7.3.b :

$$\mathbf{E} [ aX + b ] = \int_{X(\Omega)} (ax + b) f(x) dx = a \int_{X(\Omega)} x f(x) dx + b \int_{X(\Omega)} f(x) dx = a \mathbf{E} [ X ] + b .$$

Cas particulier :

Si  $a = 0$ , on a alors  $\mathbf{E} [ b ] = b$ , où  $b$  désigne, dans la partie gauche de l'égalité la variable certaine telle  $X(\Omega) = \{ b \}$ .

#### **Proposition 3.4.b** (admise) : *Linéarité de l'espérance*

Soient  $X$  et  $Y$  deux v.a.r discrètes admettant chacune une espérance, et soient  $a$  et  $b$  deux nombres réels. Alors la variable aléatoire  $aX + bY$  admet une espérance, et on a :

$$\mathbf{E} [ aX + bY ] = a \mathbf{E} [ X ] + b \mathbf{E} [ Y ] .$$

#### **Proposition 3.4.c :** *Formule de Kœnig*

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

Démonstration :

$$\begin{aligned} \mathbf{E} [ ( X - \mathbf{E} [ X ] )^2 ] &= \mathbf{E} [ X^2 - 2 \mathbf{E} [ X ] X + \mathbf{E} [ X ]^2 ] \\ &= \mathbf{E} [ X^2 ] + \mathbf{E} [ -2 \mathbf{E} [ X ] X ] + \mathbf{E} [ \mathbf{E} [ X ]^2 ] \\ &= \mathbf{E} [ X^2 ] - 2 \mathbf{E} [ X ] \mathbf{E} [ X ] + \mathbf{E} [ X ]^2 \end{aligned}$$

**Proposition 3.4.d :**

Soient  $X$  une v.a.r, et  $a$  et  $b$  deux nombres réels. On a alors :

$$\mathbf{V} [ aX + b ] = a^2 \mathbf{V} [ X ] .$$

Démonstration :

$$\begin{aligned} \mathbf{V} [ aX + b ] &= \mathbf{E} [ (aX + b)^2 ] - \mathbf{E} [ aX + b ]^2 \\ &= \mathbf{E} [ a^2 X^2 + 2abX + b^2 ] - ( a \mathbf{E} [ X ] + b )^2 \\ &= a^2 \mathbf{E} [ X^2 ] + 2ab \mathbf{E} [ X ] + b^2 - a^2 \mathbf{E} [ X ]^2 - 2ab \mathbf{E} [ X ] - b^2 \\ &= a^2 ( \mathbf{E} [ X^2 ] - \mathbf{E} [ X ]^2 ) = a^2 \mathbf{V} [ X ] . \end{aligned}$$

Cas particuliers :

1.  $a = 0 : \mathbf{V} [ b ] = 0$ , où  $b$  désigne, dans la partie gauche de l'égalité la variable certaine telle  $X(\Omega) = \{ b \}$ .
2.  $a = 1 : \mathbf{V} [ X + b ] = \mathbf{V} [ X ]$ .

**Théorème 3.4.e : Inégalité de Bienaymé-Tchebychev**

Soit  $X$  une v.a.r admettant une espérance et une variance. Alors :

$$\forall \varepsilon > 0, \mathbf{P} ( | X - \mathbf{E} [ X ] | \geq \varepsilon ) \leq \frac{\mathbf{V}[X]}{\varepsilon^2} .$$

Démonstration :

Si  $X$  est une variable discrète telle que  $X(\Omega) = \{ x_k : k \in K \}$ , en notant  $m = \mathbf{E} [ X ]$  et  $B = \{ x_k \in X(\Omega) : | x_k - m | \geq \varepsilon \}$ , on a :

$$\begin{aligned} \mathbf{V} [ X ] &= \sum_{x_k \in X(\Omega)} ( x_k - m )^2 \mathbf{P} ( X = x_k ) \geq \sum_{x_k \in B} ( x_k - m )^2 \mathbf{P} ( X = x_k ) \\ &\geq \varepsilon^2 \sum_{x_k \in B} \mathbf{P} ( X = x_k ) = \varepsilon^2 \mathbf{P} ( B ), \end{aligned}$$

$$\text{d'où } \mathbf{P} ( B ) = \mathbf{P} ( | X - \mathbf{E} [ X ] | \geq \varepsilon ) \leq \frac{\mathbf{V}[X]}{\varepsilon^2} .$$

Si  $X$  est une variable absolument continue de densité  $f$  sur  $X(\Omega)$ , en notant  $m = \mathbf{E} [ x ]$  et  $B = \{ x \in X(\Omega) : | x - m | \geq \varepsilon \}$ , on a :

$$\begin{aligned} \mathbf{V} [ X ] &= \int_{X(\Omega)} ( x - m )^2 f(x) dx \geq \int_B ( x - m )^2 f(x) dx \\ &\geq \varepsilon^2 \int_B f(x) dx = \varepsilon^2 \mathbf{P} ( B ), \end{aligned}$$

$$\text{d'où } \mathbf{P} ( B ) = \mathbf{P} ( | X - \mathbf{E} [ X ] | \geq \varepsilon ) \leq \frac{\mathbf{V}[X]}{\varepsilon^2} .$$

Remarque :

L'inégalité de Bienaymé-Tchebychev peut également s'écrire :

$$\forall k > 0, \mathbf{P}(|X - \mathbf{E}[X]| \geq k\sigma) \leq \frac{1}{k^2} \text{ avec } \sigma = \sqrt{\mathbf{V}[X]}$$

### 3.5. Moments des lois usuelles

#### 3.5.1. Lois discrètes

Partout où la notation  $p$  sera utilisée, on notera  $q = 1 - p$ .

**Proposition 3.5.1.a :** *Moments d'une loi certaine de valeur  $a$*

Si  $X \sim \mathcal{C}(a)$ , alors :  $\mathbf{E}[X] = a$  et  $\mathbf{V}[X] = 0$ .

**Proposition 3.5.1.b :** *Moments d'une loi discrète uniforme sur  $\llbracket 1, N \rrbracket$*

Si  $X \sim \mathcal{U}_{\llbracket 1, N \rrbracket}$ , alors :  $\mathbf{E}[X] = \frac{N+1}{2}$  et  $\mathbf{V}[X] = \frac{N^2-1}{12}$ .

Démonstration :

$$\mathbf{E}[X] = \sum_{k=1}^N \frac{1}{N} \times k = \frac{1}{N} \times \frac{N(N+1)}{2} = \frac{N+1}{2} ;$$

$$\mathbf{V}[X] = \sum_{k=1}^N \frac{1}{N} \times k^2 - \left(\frac{N+1}{2}\right)^2 = \frac{1}{N} \times \frac{N(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} = \frac{N^2-1}{12} .$$

**Proposition 3.5.1.c :** *Moments d'une loi de Bernoulli*

Si  $X \sim \mathcal{B}(p)$ , alors :  $\mathbf{E}[X] = p$  ;  $\mathbf{V}[X] = pq$ .

Démonstration :

$$\mathbf{E}[X] = 1 \times p + 0 \times (1-p) = p ;$$

$$\mathbf{V}[X] = 1^2 \times p + 0^2 \times (1-p) - p^2 = p - p^2 = p(1-p) = pq ;$$

**Proposition 3.5.1.d :** *Moments d'une loi binomiale*

Si  $X \sim \mathcal{B}(n, p)$ , alors :  $\mathbf{E}[X] = np$  ;  $\mathbf{V}[X] = npq$ .

Démonstration :

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k} = \sum_{k=1}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^n np \frac{(n-1)!}{(k-1) [(n-1)-(k-1)]!} p^{k-1} (1-p)^{(n-1)-(k-1)} = np \sum_{k=0}^{n-1} C_{n-1}^k p^k (1-p)^{n-1-k} = np ; \\
\mathbf{V}[X] &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 \text{ et } \mathbf{E}[X^2] = \mathbf{E}[X(X-1)] + \mathbf{E}[X], \\
\text{et } \mathbf{E}[X(X-1)] &= \sum_{k=0}^n k(k-1) C_n^k p^k (1-p)^{n-k} = \sum_{k=2}^n k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^n n(n-1)p^2 \frac{(n-2)!}{(k-2) [(n-2)-(k-2)]!} p^{k-2} (1-p)^{(n-2)-(k-2)} \\
&= n(n-1)p^2 \sum_{k=0}^{n-2} C_{n-2}^k p^k (1-p)^{n-2-k} = n^2 p^2 - n p^2 , \\
\text{d'où } \mathbf{V}[X] &= n^2 p^2 - n p^2 + np - n^2 p^2 = np(1-p).
\end{aligned}$$

**Proposition 3.5.1.e : Moments d'une loi de Pascal**

Si  $X \sim \mathcal{P}(r, p)$ , alors :  $\mathbf{E}[X] = \frac{r}{p}$  ;  $\mathbf{V}[X] = \frac{rq}{p^2}$ .

Démonstration :

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{k=r}^{+\infty} k C_{k-1}^{r-1} p^r q^{k-r} = p^r \sum_{k=r}^{+\infty} \frac{k(k-1)!}{(r-1)!(k-r)!} q^{k-r} = r p^r \sum_{k=r}^{+\infty} \binom{k}{r} q^{k-r} = \frac{r p^r}{(1-q)^{r+1}} = \frac{r}{p} . \\
\mathbf{V}[X] &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \mathbf{E}[X(X+1)] - \mathbf{E}[X] - \mathbf{E}[X]^2 . \\
\mathbf{E}[X(X+1)] &= \sum_{k=r}^{+\infty} k(k+1) C_{k-1}^{r-1} p^r q^{k-r} = p^r \sum_{k=r}^{+\infty} \frac{(k+1)k(k-1)!}{(r-1)!(k-r)!} q^{k-r} \\
&= r(r+1) p^r \sum_{k=r}^{+\infty} \binom{k+1}{r+1} q^{k-r} \\
&= r(r+1) p^r \sum_{k=r+1}^{+\infty} \binom{k}{r+1} q^{k-(r+1)} \\
&= \frac{r(r+1) p^r}{(1-q)^{r+2}} = \frac{r(r+1)}{p^2} .
\end{aligned}$$

Finalement  $\mathbf{V}[X] = \frac{r(r+1)}{p^2} - \frac{r}{p} - \frac{r^2}{p^2} = \frac{rq}{p^2}$ .

**Proposition 3.5.1.f : Moments d'une loi binomiale négative**

Si  $X \sim \mathcal{Bn}(r, p)$ , alors :  $\mathbf{E}[X] = \frac{rq}{p}$  ;  $\mathbf{V}[X] = \frac{rq}{p^2}$ .

Démonstration :

Si  $X_{\text{bn}} \sim \mathcal{Bn}(r, p)$  et  $X_P \in \mathcal{P}(r, p)$ , on a  $X_{\text{bn}} = X_P - r$ , d'où :

$$\mathbf{E}[X_{\text{bn}}] = \mathbf{E}[X_P] - r = \frac{r}{p} - r = \frac{rq}{p} \text{ et } \mathbf{V}[X_{\text{bn}}] = \mathbf{V}[X_P] = \frac{rq}{p^2}.$$

**Proposition liminaire 3.5.1.g : Formule de Vandermonde**

Cette formule s'écrit :  $\sum_{k=0}^n C_{n_1}^k C_{n_2}^{n-k} = C_{n_1+n_2}^n$ .

Elle s'obtient en identifiant les coefficients des termes de degré  $n$  entre le produit  $(1+X)^{n_1} \times (1+X)^{n_2}$  et le polynôme  $(1+X)^{n_1+n_2}$ .

**Proposition 3.5.1.h : Moments d'une loi hypergéométrique**

Si  $X \sim \mathcal{H}(N, n, p)$ , alors :  $\mathbf{E}[X] = np$  ;  $\mathbf{V}[X] = \frac{N-n}{N-1} npq$

Démonstration :

$$\mathbf{E}[X] = \sum_{k=0}^n k \mathbf{P}(X=k) = \sum_{k=1}^n \frac{k C_{N_1}^k C_{N_2}^{n-k}}{C_N^n}.$$

$$\text{Or, } k C_{N_1}^k = k \frac{N_1!}{k!(N_1-k)!} = N_1 \frac{(N_1-1)!}{(k-1)![(N_1-1)-(k-1)]!} = N_1 C_{N_1-1}^{k-1}.$$

$$\text{On en déduit que } \mathbf{E}[X] = \frac{N_1}{C_N^n} \left( \sum_{k=1}^n C_{N_1-1}^{k-1} C_{N_2}^{n-1-(k-1)} \right) = \frac{N_1}{C_N^n} \left( \sum_{k=0}^{n-1} C_{N_1-1}^k C_{N_2}^{n-1-k} \right).$$

En appliquant la formule de Vandermonde, on obtient :

$$\mathbf{E}[X] = \frac{N_1}{C_N^n} C_{N-1}^{n-1} = N_1 \frac{(N-n)!n!}{N!} \frac{(N-1)!}{(n-1)!(N-n)!} = \frac{N_1 n}{N} = np, \text{ où } p = \frac{N_1}{N}.$$

$$\mathbf{V}[X] = \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \mathbf{E}[X(X-1)] + \mathbf{E}[X] - \mathbf{E}[X]^2.$$

$$\mathbf{E}[X(X-1)] = \sum_{k=0}^n k(k-1) \mathbf{P}(X=k) = \sum_{k=2}^n \frac{k(k-1) C_{N_1}^k C_{N_2}^{n-k}}{C_N^n}.$$

$$\text{Or, comme } k(k-1) C_{N_1}^k = k(k-1) \frac{N_1!}{k!(N_1-k)!} = N_1(N_1-1) \frac{(N_1-2)!}{(k-2)![(N_1-2)-(k-2)]!}$$

$= N_1 (N_1 - 1) C_{N_1-2}^{k-2}$ , on en déduit que :

$$\begin{aligned} \mathbf{E} [ X (X - 1) ] &= \frac{N_1 (N_1 - 1)}{C_N^n} \left( \sum_{k=2}^n C_{N_1-2}^{k-2} C_{N_2}^{n-2-(k-2)} \right) = \frac{N_1 (N_1 - 1)}{C_N^n} \left( \sum_{k=0}^{n-2} C_{N_1-2}^k C_{N_2}^{n-2-k} \right) \\ &= \frac{N_1 (N_1 - 1)}{C_N^n} C_{N-2}^{n-2} = N_1 (N_1 - 1) \frac{(N - n)! n!}{N!} \frac{(N - 2)!}{(n - 2)! (N - n)!} = \frac{N_1 (N_1 - 1) n (n - 1)}{N (N - 1)} \\ &= \frac{pN (pN - 1) n (n - 1)}{N (N - 1)} = \frac{p (pN - 1) n (n - 1)}{(N - 1)}. \end{aligned}$$

$$\begin{aligned} \text{Finalement } \mathbf{V} [ X ] &= \frac{p (pN - 1) n (n - 1)}{(N - 1)} + np - (np)^2 \\ &= \frac{p (pN - 1) n (n - 1) + np (N - 1) - n^2 p^2 (N - 1)}{(N - 1)} \\ &= \frac{np}{N - 1} ((pN - 1) (n - 1) + (N - 1) - np (N - 1)) \\ &= \frac{np}{N - 1} (pNn - pN - n + 1 + N - 1 - npN + np) = \frac{np}{N - 1} (-pN - n + N + np) \\ &= \frac{np}{N - 1} (N - n) (1 - p) = \frac{N - n}{N - 1} npq \text{ où } q = 1 - p. \end{aligned}$$

**Proposition 3.5.1.i : Moments d'une loi de Poisson**

Si  $X \sim \mathcal{P}(\lambda)$ , alors :  $\mathbf{E} [ X ] = \mathbf{V} [ X ] = \lambda$ .

Démonstration :

$$\begin{aligned} \mathbf{E} [ X ] &= \sum_{k=1}^{+\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{+\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} = \lambda. \\ \mathbf{V} [ X ] &= \mathbf{E} [ X^2 ] - \mathbf{E} [ X ]^2 = \mathbf{E} [ X (X - 1) ] + \mathbf{E} [ X ] - \mathbf{E} [ X ]^2 \\ &= \sum_{k=2}^{+\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} + \lambda - \lambda^2 = \lambda^2 \sum_{k=2}^{+\infty} e^{-\lambda} \frac{\lambda^{k-2}}{(k-2)!} + \lambda - \lambda^2 = \lambda. \end{aligned}$$



### 3.5.2. Lois continues

**Proposition 3.5.2.a :** Moments d'une loi uniforme sur un segment  $[a ; b]$

Si  $X \sim \mathcal{U}_{[a ; b]}$ , alors :  $\mathbf{E}[X] = \frac{a+b}{2}$  ;  $\mathbf{V}[X] = \frac{(b-a)^2}{12}$ .

Démonstration :

$$\mathbf{E}[X] = \frac{1}{b-a} \int_a^b t \, dt = \frac{a+b}{2} ;$$

$$\mathbf{V}[X] = \frac{1}{b-a} \int_a^b (t-m)^2 \, dt = \frac{1}{b-a} \int_{\frac{a-b}{2}}^{\frac{b-a}{2}} t^2 \, dt = \frac{2}{b-a} \left[ \frac{t^3}{3} \right]_0^{\frac{b-a}{2}} = \frac{(b-a)^2}{12} .$$

**Proposition 3.5.2.b :** Moments d'une loi exponentielle de paramètre  $\lambda$

Si  $X \sim \mathcal{E}(\lambda)$  alors :  $\mathbf{E}[X] = \frac{1}{\lambda}$  ;  $\mathbf{V}[X] = \frac{1}{\lambda^2}$ .

Démonstration :

$$\mathbf{E}[X] = \int_0^{+\infty} \lambda t e^{-\lambda t} \, dt = \left[ -t e^{-\lambda t} \right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda t} \, dt = \frac{1}{\lambda} .$$

$$\begin{aligned} \mathbf{V}[X] &= \int_0^{+\infty} \left( t - \frac{1}{\lambda} \right)^2 \lambda e^{-\lambda t} \, dt = \int_0^{+\infty} t^2 \lambda e^{-\lambda t} \, dt - \frac{2}{\lambda} \int_0^{+\infty} t \lambda e^{-\lambda t} \, dt + \frac{1}{\lambda^2} \int_0^{+\infty} \lambda e^{-\lambda t} \, dt \\ &= \int_0^{+\infty} t^2 \lambda e^{-\lambda t} \, dt - \frac{1}{\lambda^2} \text{ et } \int_0^{+\infty} t^2 e^{-\lambda t} \, dt = \left[ -t^2 e^{-\lambda t} \right]_0^{+\infty} + 2 \int_0^{+\infty} t e^{-\lambda t} \, dt = \frac{2}{\lambda^2} , \end{aligned}$$

d'où  $\mathbf{V}[X] = \frac{1}{\lambda^2}$ .

**Proposition 3.5.2.c :** Moments d'une loi normale centrée réduite

Si  $X \sim \mathcal{N}(0, 1)$  alors :  $\mathbf{E}[X] = 0$  ;  $\mathbf{V}[X] = 1$ .

Démonstration :

$$\mathbf{E}[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} \, dt = \lim_{A \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} \int_{-A}^{+A} t e^{-\frac{t^2}{2}} \, dt = 0 : X \text{ est donc centrée;} ,$$

$$\begin{aligned} \mathbf{V}[X] &= \mathbf{E}[X^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} \, dt = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} t^2 e^{-\frac{t^2}{2}} \, dt \\ &= \sqrt{\frac{2}{\pi}} \left[ -t e^{-\frac{t^2}{2}} \right]_0^{+\infty} + \sqrt{\frac{2}{\pi}} \int_0^{+\infty} e^{-\frac{t^2}{2}} \, dt = \sqrt{\frac{2}{\pi}} \frac{\sqrt{2\pi}}{2} = 1 : X \text{ est donc réduite.} \end{aligned}$$

**Proposition 3.5.2.d : Moments d'une loi normale**

Si  $X \sim \mathcal{N}(\mu, \sigma)$  alors :  $\mathbf{E} [ X ] = \mu ; \mathbf{V} [ X ] = \sigma^2$ .

Démonstration :

En général, comme  $\forall (\sigma, \mu) \in \mathbb{R}_+^* \times \mathbb{R}, X \sim \mathcal{N}(\mu, \sigma) \Rightarrow X^* = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$ , alors :

$\forall (\sigma, \mu) \in \mathbb{R}_+^* \times \mathbb{R}, X \sim \mathcal{N}(\mu, \sigma) \Rightarrow \mathbf{E} [ X ] = \mathbf{E} [ \sigma X^* + \mu ] = \sigma \mathbf{E} [ X^* ] + \mu = \mu ;$

$\forall (\sigma, \mu) \in \mathbb{R}_+^* \times \mathbb{R}, X \sim \mathcal{N}(\mu, \sigma) \Rightarrow \mathbf{V} [ X ] = \mathbf{V} [ \sigma X^* + \mu ] = \sigma^2 \mathbf{V} [ X^* ] = \sigma^2 .$

**Proposition 3.5.2.e : Moments d'une loi log-normale**

Si  $U \sim \mathcal{LN}(\lambda, \zeta)$  alors :  $\mathbf{E} [ U ] = e^{\lambda + \frac{\zeta^2}{2}} ; \mathbf{V} [ U ] = (e^{\zeta^2} - 1) e^{2\left(\lambda + \frac{\zeta^2}{2}\right)}$ .

**Proposition 3.5.2.f : Moments d'une loi gamma de paramètre  $r$**

Si  $X \sim \Gamma(r)$  alors :  $\mathbf{E} [ X ] = \mathbf{V} [ X ] = r$

**Proposition 3.5.2.g : Moments d'une loi gamma de paramètres  $\alpha$  et  $\beta$**

Si  $X \sim \Gamma(r)$  alors :  $\mathbf{E} [ X ] = \frac{\alpha}{\beta} ; \mathbf{V} [ X ] = \frac{\alpha}{\beta^2} .$

## 4. Vecteurs aléatoires

### Définition 4.a : Vecteur aléatoire

Etant donnés  $m \in \mathbb{N}^*$ , et  $X_1, \dots, X_m$ ,  $m$  v.a.r, on définit le vecteur aléatoire  $V = (X_1, \dots, X_m)$  par l'application de  $\Omega$  dans  $\mathbb{R}^m$ :  $\omega \mapsto V(\omega) = (X_1(\omega), \dots, X_m(\omega))$ .

### Définition 4.b : Fonction de répartition d'un vecteur aléatoire

Etant donné un vecteur aléatoire  $V = (X_1, \dots, X_m)$ , la fonction de répartition  $F_V$  de ce vecteur est définie en chaque point  $(x_1, \dots, x_m) \in V(\Omega)$  par :

$$F_V(x_1, \dots, x_m) = \mathbf{P}(X_1 \leq x_1 \cap \dots \cap X_m \leq x_m)$$

## 4.1. Vecteurs aléatoires discrets

### 4.1.1 Définitions

#### Définition 4.1.1.a : Vecteur aléatoire discret ; loi conjointe

Si  $X_1, \dots, X_m$  sont  $m$  variables aléatoires discrètes, on définit un vecteur aléatoire discret  $V = (X_1, \dots, X_m)$  par sa loi, qui est elle-même définie par :

1. l'ensemble  $V(\Omega) \subset \mathbb{R}^m$ ;
2. pour chaque  $m$ -uplet  $(x_1, \dots, x_m) \in X_1(\Omega) \times \dots \times X_m(\Omega)$ , la probabilité

$$\mathbf{P}[(X_1, \dots, X_m) = (x_1, \dots, x_m)] = \mathbf{P}((X_1 = x_1) \cap \dots \cap (X_m = x_m)).$$

Cette loi est également appelée la *loi conjointe* des variables aléatoires  $X_1, \dots, X_m$ .

#### Exemples 4.1.1.b :

$$1^\circ) V = (X, Y), \text{ avec } V(\Omega) = \llbracket 1, 6 \rrbracket^2, \text{ et } \mathbf{P}((X, Y) = (n, m)) = \frac{1}{36}.$$

$$2^\circ) V = (X, Y), \text{ avec } V(\Omega) = \mathbb{N}^2, \text{ et } \mathbf{P}((X, Y) = (n, m)) = \frac{1}{2^{n+1}} \frac{e^{-1}}{m!}.$$

$$3^\circ) V = (X, Y), \text{ avec } V(\Omega) = \mathbb{N}^2, \text{ et } \mathbf{P}((X, Y) = (n, m)) = \frac{e^{-n} n^m}{m! 2^{n+1}}.$$

$$4^\circ) V = (X, Y), \text{ avec } V(\Omega) = \{(n, m) \in \mathbb{N}^2 \text{ t. q. } m \leq n\}, \text{ et } \mathbf{P}((X, Y) = (n, m)) = \frac{1}{(n+1) 2^{n+1}}$$

**Définition 4.1.1.c : Lois marginales**

Etant donné un vecteur aléatoire discret  $(X_1, \dots, X_m)$ , on désigne par *lois marginales* les lois des composantes  $X_i, i \in \llbracket 1, m \rrbracket$ , de ce vecteur. Elles sont définies par :

1.  $X_i(\Omega)$  ;
2.  $\mathbf{P}(X_i = x_i) = \sum_{j \neq i} \sum_{k_j \in X_j(\Omega)} \mathbf{P}((X_1, \dots, X_m) = (k_1, \dots, k_{i-1}, x_i, k_{i+1}, \dots, k_m))$ .

**Cas particulier :  $m = 2$** 

Etant donnés  $V = (X, Y)$ ,  $V(\Omega) = \{(x_i, y_j), i \in I, j \in J\}$  et  $\mathbf{P}((X, Y) = (x_i, y_j)) = p_{ij}$ , les lois marginales sont définies par :

1.  $X(\Omega)$  et  $\mathbf{P}(X = x_i) = \sum_{j \in J} p_{ij}$ , noté  $p_{i.}$  ;
2.  $Y(\Omega)$  et  $\mathbf{P}(Y = y_j) = \sum_{i \in I} p_{ij}$ , noté  $p_{.j}$  ;

**Exemples 4.1.1.b, suite :**

$$1^\circ) \mathbf{P}(X = n) = \sum_{m=1}^6 \mathbf{P}((X, Y) = (n, m)) = \frac{1}{6}, \text{ et } \mathbf{P}(Y = m) = \sum_{n=1}^6 \mathbf{P}((X, Y) = (n, m)) = \frac{1}{6}.$$

$$2^\circ) \mathbf{P}(X = n) = \sum_{m=0}^{+\infty} \mathbf{P}((X, Y) = (n, m)) = \sum_{m=0}^{+\infty} \frac{1}{2^{n+1}} \frac{e^{-1}}{m!} = \frac{1}{2^{n+1}}, \text{ et}$$

$$\mathbf{P}(Y = m) = \sum_{n=0}^{+\infty} \mathbf{P}((X, Y) = (n, m)) = \sum_{n=0}^{+\infty} \frac{1}{2^{n+1}} \frac{e^{-1}}{m!} = \frac{e^{-1}}{m!}.$$

$$3^\circ) \mathbf{P}(X = n) = \sum_{m=0}^{+\infty} \mathbf{P}((X, Y) = (n, m)) = \sum_{m=0}^{+\infty} \frac{e^{-n} n^m}{m! 2^{n+1}} = \frac{1}{2^{n+1}}, \text{ et}$$

$$\mathbf{P}(Y = m) = \sum_{n=0}^{+\infty} \mathbf{P}((X, Y) = (n, m)) = \sum_{n=0}^{+\infty} \frac{e^{-n} n^m}{m! 2^{n+1}}.$$

$$4^\circ) \mathbf{P}(X = n) = \sum_{m=0}^{+n} \mathbf{P}((X, Y) = (n, m)) = \sum_{m=0}^n \frac{1}{(n+1) 2^{n+1}} = \frac{1}{2^{n+1}}, \text{ et}$$

$$\mathbf{P}(Y = m) = \sum_{n=m}^{+\infty} \mathbf{P}((X, Y) = (n, m)) = \sum_{n=m}^{+\infty} \frac{1}{(n+1) 2^{n+1}}.$$

**Définition 4.1.1.d : Fonction de répartition d'un vecteur aléatoire discret**

Etant donné un vecteur aléatoire discret  $V = (X_1, \dots, X_m)$ , la fonction de répartition  $F_V$  de ce vecteur est définie en chaque point  $(x_1, \dots, x_m) \in V(\Omega)$  par :

$$F_V(x_1, \dots, x_m) = \mathbf{P}(X_1 \leq x_1 \cap \dots \cap X_m \leq x_m) = \sum_{k_1 \leq x_1} \dots \sum_{k_m \leq x_m} \mathbf{P}(X_1 = k_1 \cap \dots \cap X_m = k_m).$$

**4.1.2 Fonction d'un vecteur aléatoire discret****Proposition 4.1.2.a : loi d'une fonction scalaire d'un vecteur aléatoire discret**

Etant donnés un vecteur aléatoire discret  $V = (X_1, \dots, X_m)$  et une fonction  $\varphi$  à valeurs réelles définie en tout point de  $V(\Omega)$ , l'application :

$$Y = \varphi(X_1, \dots, X_m) : \Omega \rightarrow \mathbb{R}$$

$$\omega \mapsto \varphi(X_1(\omega), \dots, X_m(\omega)),$$

est une v.a.r discrète dont la loi est définie par :

1.  $Y(\Omega) = \varphi(V(\Omega))$  ;
2. pour tout  $y \in Y(\Omega)$ ,  $\mathbf{P}(Y = y) = \sum_{\substack{x_1 \in X_1(\Omega), \dots, x_m \in X_m(\Omega) \\ \varphi(x_1, \dots, x_m) = y}} \mathbf{P}[V = (x_1, \dots, x_m)]$ .

**Proposition 4.1.2.b : loi de la somme de variables aléatoires discrètes**

Etant données  $m$  v.a.r discrètes  $X_1, \dots, X_m$ , la loi de  $Y = X_1 + \dots + X_m$  est définie par :

1.  $Y(\Omega) = X_1(\Omega) + \dots + X_m(\Omega)$  ;
2.  $\forall y \in Y(\Omega)$ ,  $\mathbf{P}(Y = y) = \sum_{\substack{x_1 \in X_1(\Omega), \dots, x_m \in X_m(\Omega) \\ x_1 + \dots + x_m = y}} \mathbf{P}[V = (x_1, \dots, x_m)]$ .

où  $V$  désigne le vecteur aléatoire discret  $V = (X_1, \dots, X_m)$ .

**Proposition 4.1.2.c : loi de la somme de deux variables aléatoires discrètes**

Etant données deux v.a.r discrètes  $X$  et  $Y$ , la loi de  $Z = X + Y$  est définie par :

$$\begin{aligned}
 1. \quad Z(\Omega) &= X(\Omega) + Y(\Omega) ; \\
 2. \quad \forall z \in Z(\Omega), \mathbf{P}(Z = z) &= \sum_{\substack{x \in X(\Omega), y \in Y(\Omega) \\ x + y = z}} \mathbf{P}((X, Y) = (x, y)) \\
 &= \sum_{\substack{x \in X(\Omega), y \in Y(\Omega) \\ x + y = z}} \mathbf{P}((X, Y) = (x, y)) \\
 &= \sum_{x \in X(\Omega)} \mathbf{P}((X, Y) = (x, z - x)) \\
 &= \sum_{y \in Y(\Omega)} \mathbf{P}((X, Y) = (z - y, y)).
 \end{aligned}$$

**Exemples 4.1.2.d :**

En reprenant les vecteurs aléatoires  $V = (X, Y)$  des exemples **4.1.1.b**, et en notant  $Z = X + Y$ , on obtient :

$$\begin{aligned}
 1^\circ) \quad Z(\Omega) &= \llbracket 2, 12 \rrbracket \text{ et } \mathbf{P}(Z = z) = \sum_{\substack{n, m \in \{1, \dots, 6\}^2 \\ n + m = z}} \mathbf{P}[(X, Y) = (n, m)] \\
 &= \sum_{n=1}^6 \mathbf{P}[(X, Y) = (n, z - n)]
 \end{aligned}$$

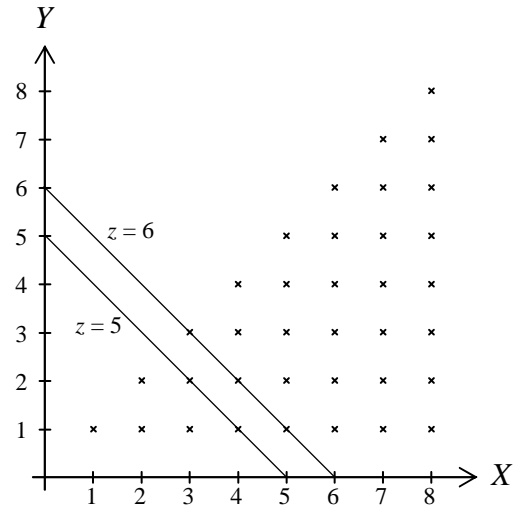
$$\text{Pour } z \leq 7, \mathbf{P}(Z = z) = \sum_{n=1}^{z-1} \frac{1}{36} = \frac{z-1}{36}, \text{ et pour } z \geq 8, \mathbf{P}(Z = z) = \sum_{n=z-6}^6 \frac{1}{36} = \frac{13-z}{36}.$$

$$\begin{aligned}
 2^\circ) \quad Z(\Omega) &= \mathbb{N}, \text{ et } \mathbf{P}(Z = z) = \sum_{\substack{n, m \in \mathbb{N}^2 \\ n + m = z}} \mathbf{P}[(X, Y) = (n, m)] \\
 &= \sum_{n=0}^z \mathbf{P}[(X, Y) = (n, z - n)] = \sum_{n=0}^z \frac{1}{2^{n+1}} \frac{e^{-1}}{(z - n)!} \\
 &= \sum_{m=0}^z \mathbf{P}[(X, Y) = (z - m, m)] = \sum_{m=0}^z \frac{1}{2^{z-m+1}} \frac{e^{-1}}{m!}
 \end{aligned}$$

$$\begin{aligned}
 3^\circ) \quad Z(\Omega) &= \mathbb{N}, \text{ et } \mathbf{P}(Z = z) = \sum_{\substack{n, m \in \mathbb{N}^2 \\ n + m = z}} \mathbf{P}[(X, Y) = (n, m)] \\
 &= \sum_{n=0}^z \mathbf{P}[(X, Y) = (n, z - n)] = \sum_{n=0}^z \frac{1}{2^{n+1}} \frac{e^{-n} n^{z-n}}{(z - n)!} \\
 &= \sum_{m=0}^z \mathbf{P}[(X, Y) = (z - m, m)] = \sum_{m=0}^z \frac{1}{2^{z-m+1}} \frac{e^{-(z-m)} (z - m)^m}{m!}
 \end{aligned}$$

4°)  $Z(\Omega) = \mathbb{N}$ , et

$$\begin{aligned} \mathbf{P}(Z = z) &= \sum_{\substack{n \in \mathbb{N}, m \in \{0, \dots, n\} \\ n+m=z}} \mathbf{P}[(X, Y) = (n, m)] \\ &= \sum_{n=\lceil \frac{z}{2} \rceil}^z \frac{1}{(n+1) 2^{n+1}} = \sum_{m=0}^{\lfloor \frac{z}{2} \rfloor} \frac{1}{(z-m+1) 2^{z-m+1}} \end{aligned}$$



**Proposition 4.1.2.e :** *Application à la loi binomiale*

Si  $X = \sum_{i=1}^n X_i$ , avec  $\forall i \in \llbracket 1, n \rrbracket, X_i \sim \mathcal{B}(p)$ , alors  $X \sim \mathcal{B}(n, p)$ .

Démonstration :

$$\begin{aligned} \forall k \in \llbracket 0, n \rrbracket, \mathbf{P}(X = k) &= \sum_{\substack{x_1, \dots, x_n \in \{0,1\} \\ x_1 + \dots + x_n = k}} \mathbf{P}((X_1 = x_1) \cap \dots \cap (X_n = x_n)) \\ &= \sum_{\substack{x_1, \dots, x_n \in \{0,1\} \\ x_1 + \dots + x_n = k}} \mathbf{P}(X_1 = x_1) \dots \mathbf{P}(X_n = x_n) \\ &= \sum_{\substack{x_1, \dots, x_n \in \{0,1\} \\ x_1 + \dots + x_n = k}} p^k (1-p)^{n-k} \\ &= C_n^k p^k (1-p)^{n-k}. \end{aligned}$$

**Proposition 4.1.2.f :** *Espérance de  $Y = \varphi(X_1, \dots, X_n)$*

Etant donné un vecteur aléatoire discret  $V = (X_1, \dots, X_m)$  et une fonction  $\varphi$  à valeurs réelles définie en tout point de  $V(\Omega)$ , soit la variable aléatoire  $Y = \varphi(X_1, \dots, X_m)$ . On a alors :

$$\mathbf{E}[Y] = \sum_{(x_1, \dots, x_m) \in X_1(\Omega) \times \dots \times X_m(\Omega)} \varphi(x_1, \dots, x_m) \mathbf{P}[(X_1 = x_1) \cap \dots \cap (X_m = x_m)],$$

sous réserve que cette somme soit définie.

**Théorème 4.1.2.g : Linéarité de l'espérance**

Soient  $X$  et  $Y$  deux v.a.r discrètes admettant chacune une espérance, et soient  $a$  et  $b$  deux nombres réels. Alors la variable aléatoire  $aX + bY$  admet une espérance, et on a :

$$\mathbf{E} [aX + bY] = a \mathbf{E} [X] + b \mathbf{E} [Y].$$

**Démonstration :**

Posons  $X(\Omega) = \{x_i, i \in I\}$  et  $Y(\Omega) = \{y_j, j \in J\}$ . Soit  $V = (X, Y)$  et  $\phi$  la fonction à valeurs réelles définie en tout point de  $V(\Omega)$  par  $\phi(V) = aX + bY$ .

On a alors :

$$\begin{aligned} \mathbf{E} [aX + bY] &= \sum_{(i,j) \in I \times J} (ax_i + by_j) \mathbf{P} [(X = x_i) \cap (Y = y_j)] \\ &= a \sum_{(i,j) \in I \times J} x_i \mathbf{P} [(X = x_i) \cap (Y = y_j)] + b \sum_{(i,j) \in I \times J} y_j \mathbf{P} [(X = x_i) \cap (Y = y_j)] \\ &= a \sum_{i \in I} \left[ x_i \sum_{j \in J} \mathbf{P} [(X = x_i) \cap (Y = y_j)] \right] + b \sum_{j \in J} \left[ y_j \sum_{i \in I} \mathbf{P} [(X = x_i) \cap (Y = y_j)] \right] \\ &= a \sum_{i \in I} x_i \mathbf{P} (X = x_i) + b \sum_{j \in J} y_j \mathbf{P} (Y = y_j) \\ &= a \mathbf{E} [X] + b \mathbf{E} [Y]. \end{aligned}$$

**4.1.3 Loi multinomiale****Définition 4.1.3.a : Loi multinomiale**

Un vecteur aléatoire discret  $(X_1, \dots, X_r)$  suit une loi multinomiale de paramètres  $n, p_1, \dots, p_r$ , si et seulement si  $X_1(\Omega) = \dots = X_r(\Omega) = \llbracket 0, n \rrbracket$  et pour tout  $(k_1, \dots, k_r) \in \llbracket 0, n \rrbracket^r$  :

$$\mathbf{P} ((X_1 = k_1) \cap \dots \cap (X_r = k_r)) = \begin{cases} 0 & \text{si } k_1 + \dots + k_r \neq n \\ \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r} & \text{si } k_1 + \dots + k_r = n \end{cases}$$

**Exemple 4.1.3.b :**

Si l'on considère une population d'effectif  $N$  comportant  $r$  catégories d'individus  $c_1, \dots, c_r$ , et si l'on note respectivement  $N_i$  et  $p_i$  le nombre et la proportion d'individus appartenant à la catégorie  $c_i$ , on a :  $\forall i \in \llbracket 1, r \rrbracket, p_i = \frac{N_i}{N}, N_1 + \dots + N_r = N$  et  $p_1 + \dots + p_r = 1$ .

Le vecteur aléatoire  $(X_1, \dots, X_r)$  tel que  $X_i$  donne le nombre d'individus de la catégorie  $c_i$  que l'on obtient en effectuant  $N$  tirages aléatoires avec remise dans cette population est distribué suivant une loi multinomiale de paramètres  $N, p_1, \dots, p_r$ .



En effet, si  $k_1, \dots, k_r$  sont des entiers tels que  $k_1 + \dots + k_r = N$ , alors la probabilité d'obtenir, pour chaque  $i \in \llbracket 1, r \rrbracket$ ,  $k_i$  individus de catégorie  $c_i$  dans un ordre donné est

$$\left(\frac{N_1}{N}\right)^{k_1} \dots \left(\frac{N_r}{N}\right)^{k_r} = \frac{N_1^{k_1} \dots N_r^{k_r}}{N^r} = p_1^{k_1} \dots p_r^{k_r},$$

et par ailleurs, le nombre de façons d'ordonner les  $k_1$  tirages donnant les individus de catégorie 1, ..., les  $k_r$  tirages donnant les individus de catégorie  $r$ , est  $C(n; k_1, \dots, k_r)$

$$\begin{aligned} &= C_n^{k_1} C_{n-k_1}^{k_2} \dots C_{k_{r-1}+k_r}^{k_{r-1}} C_{k_r}^{k_r} = \frac{n!}{k_1!(n-k_1)!} \times \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \times \dots \times \frac{(k_{r-1}+k_r)!}{k_{r-1}!k_r!} \times 1 \\ &= \frac{n!}{k_1! \dots k_r!}, \text{ d'où } \mathbf{P}(X_1 = k_1, \dots, X_n = k_n) = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}. \end{aligned}$$

## 4.2. Vecteurs aléatoires continus

### 4.2.1 Définitions

#### Définition 4.2.1.a : Vecteur aléatoire continu

Un vecteur aléatoire  $V = (X_1, \dots, X_m)$  est dit continu si (sauf, éventuellement, sur un nombre fini de domaines de  $V(\Omega)$  sur lesquels l'intégrale de toute fonction intégrable est nulle) sa fonction de répartition  $F_V$  est continue et possède une dérivée d'ordre  $n$  :

$$F_V^{(n)} = \frac{\partial^n F}{\partial x_1 \dots \partial x_n}.$$

#### Définition 4.2.1.b : Densité de probabilité sur $\mathbb{R}^m$

1. On appelle densité de probabilité sur  $\mathbb{R}^m$  toute fonction  $f$  de  $\mathbb{R}^m$  dans  $\mathbb{R}$  vérifiant :  
 $f$  est positive ;
2.  $f$  est continue sur  $\mathbb{R}^m$  sauf sur une union finie de domaines  $D_k$  de  $\mathbb{R}^m$  sur lesquels l'intégrale de toute fonction intégrable est nulle ;
3.  $\iiint_{\mathbb{R}^m} f(t_1, \dots, t_m) dt_1 \dots dt_m = 1$ .

**Proposition-définition 4.2.1.c :** *Densité de probabilité d'un vecteur aléatoire continu*

Tout vecteur aléatoire continu admet une fonction de densité de probabilité, qui est la dérivée d'ordre  $n$  de la fonction de répartition. On a donc :

$$f_V = \frac{\partial^n F}{\partial x_1 \dots \partial x_n} \text{ ou } F_V(x_1, \dots, x_m) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_m} f_V(t_1, \dots, t_m) dt_m \dots dt_1.$$

Démonstration (pour le cas  $n = 2$ ) :

Notons  $F_V$  la fonction de répartition de  $V$ .

En un point  $v$  de  $V(\Omega)$ , la fonction de densité de probabilité de  $V$ , notée  $f_V$ , est définie comme le rapport de la probabilité pour que  $V$  appartienne à un hypercube infinitésimal  $[v; v + dx_1[ \times \dots \times [v; v + dx_m[$  au volume  $dx_1 \times \dots \times dx_m$  de cet hypercube :

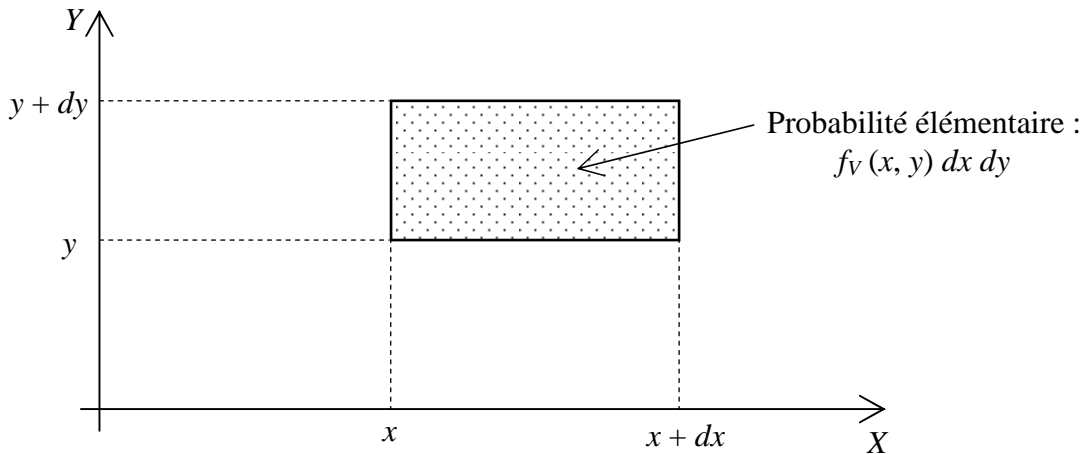
$$f_V(x_1, \dots, x_m) = \lim_{dx_1, \dots, dx_m \rightarrow 0} \frac{\mathbf{P}(V \in [v; v + dx_1[ \times \dots \times [v; v + dx_m[ )}{dx_1 \dots dx_m}.$$

Pour  $n = 2$ , en posant  $V = (X, Y)$ ,  $f_V$  est donc définie par :

$$f_V(x, y) = \lim_{\substack{dx \rightarrow 0 \\ dy \rightarrow 0}} \frac{\mathbf{P}((x \leq X \leq x + dx) \cap (y \leq Y \leq y + dy))}{dx dy}.$$

Or  $\mathbf{P}((x \leq X \leq x + dx) \cap (y \leq Y \leq y + dy))$

$$= F_V(x + dx, y + dy) - F_V(x + dx, y) - F_V(x, y + dy) + F_V(x, y).$$



Donc :

$$\begin{aligned} f_V(x, y) &= \lim_{dx \rightarrow 0} \frac{\lim_{dy \rightarrow 0} \left( \frac{F_V(x + dx, y + dy) - F_V(x + dx, y)}{dy} \right) - \lim_{dy \rightarrow 0} \left( \frac{F_V(x, y + dy) - F_V(x, y)}{dy} \right)}{dx} \\ &= \lim_{dx \rightarrow 0} \frac{\frac{\partial F_V}{\partial y}(x + dx, y) - \frac{\partial F_V}{\partial y}(x, y)}{dx} = \frac{\partial^2 F_V}{\partial x \partial y}(x, y). \end{aligned}$$

**Définition 4.2.1.d : Loi conjointe**

La loi d'un vecteur aléatoire continu  $V = (X_1, \dots, X_m)$  est définie par l'ensemble  $V(\Omega)$  et par sa fonction de densité.

Cette loi est également appelée la *loi conjointe* des variables aléatoires  $X_1, \dots, X_m$ .

**Définition 4.2.1.e : Support de la fonction de densité**

Pour un vecteur aléatoire continu  $V = (X_1, \dots, X_m)$ , l'ensemble  $V(\Omega)$  est le *support* de la fonction de densité.

Cette fonction de densité ne s'annule qu'en un nombre fini de domaines de  $V(\Omega)$  sur lesquels l'intégrale de toute fonction intégrable est nulle, et elle s'annule sur  $\mathbb{R}^m \setminus V(\Omega)$ .

On a donc  $\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_V(t_1, \dots, t_m) dt_m \dots dt_1 = \int_{V(\Omega)} f_V(t) dt = 1$ .

**Proposition 4.2.1.f (admise) :**

Etant donné un vecteur aléatoire continu  $V = (X_1, \dots, X_m)$ , de densité  $f$ , on a :

$$\forall D \subset \mathbb{R}^m, \mathbf{P}(V \in D) = \iint \dots \int_D f(t_1, \dots, t_m) dt_1 \dots dt_m.$$

**Exemples 4.2.1.g :**

$$1^\circ) m = 2, V(\Omega) = [0; 1]^2, \text{ et } f: (x, y) \mapsto f(x, y) = \begin{cases} 1 \text{ si } (x, y) \in [0; 1]^2 \\ 0 \text{ sinon} \end{cases} \quad (\text{loi uniforme}) ;$$

$$2^\circ) m = 2, V(\Omega) = [0; 1]^2, \text{ et } f: (x, y) \mapsto f(x, y) = \begin{cases} 4xy \text{ si } (x, y) \in [0; 1]^2 \\ 0 \text{ sinon} \end{cases} ;$$

$$3^\circ) m = 2, V(\Omega) = [0; 1]^2, \text{ et } f: (x, y) \mapsto f(x, y) = \begin{cases} x + y \text{ si } (x, y) \in [0; 1]^2 \\ 0 \text{ sinon} \end{cases} ;$$

$$4^\circ) m = 2, V(\Omega) = \{ (x, y) \in \mathbb{R}^2 \text{ t.q. } x^2 + y^2 \leq 1 \}, \text{ et } f: (x, y) \mapsto f(x, y) = \begin{cases} \frac{1}{\pi} \text{ si } (x, y) \in V(\Omega) \\ 0 \text{ sinon} \end{cases} ;$$

**Définition 4.2.1.h : Lois marginales**

Etant donné un vecteur aléatoire continu  $V = (X_1, \dots, X_m)$ , de densité  $f$ , on désigne par *lois marginales* les lois des composantes  $X_i$ ,  $i \in \llbracket 1, m \rrbracket$ . Elles sont données par :

1.  $X_i(\Omega)$  ;
2.  $f_i(x_i) = \int \dots \int_{\mathbb{R}^{m-1}} f(t_1, \dots, t_{i-1}, x_i, t_{i+1}, \dots, t_m) dt_1 \dots dt_{i-1} dt_{i+1} \dots dt_m$ .

Les fonctions  $f_i$  sont appelées les fonctions de densité *marginales* du vecteur  $V$ .

**Cas particulier :  $m = 2$** 

Etant donné un vecteur aléatoire continu  $(X, Y)$ , de densité  $f$ , les lois marginales sont données par :

1.  $X(\Omega)$  et  $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$  ;
2.  $Y(\Omega)$  et  $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$ .

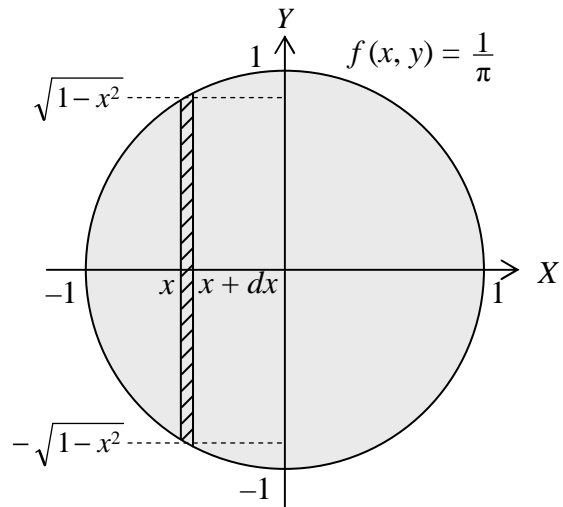
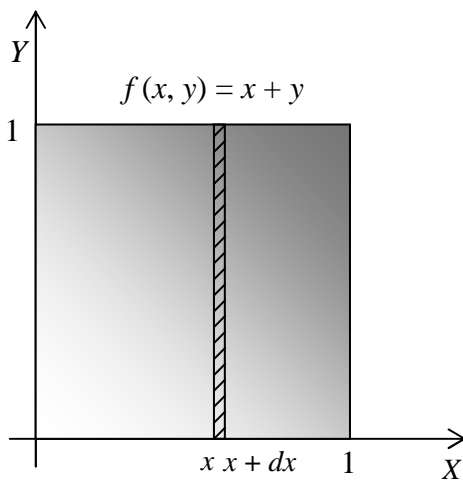
**Exemples 4.2.1.g, suite :**

$$1^\circ) f_X(x) = \int_0^1 dy = 1 \text{ et } f_Y(y) = \int_0^1 dx = 1 ;$$

$$2^\circ) f_X(x) = \int_0^1 4xy dy = 2x \text{ et } f_Y(y) = \int_0^1 4xy dx = 2y ;$$

$$3^\circ) f_X(x) = \int_0^1 (x+y) dy = \left[ xy + \frac{y^2}{2} \right]_0^1 = x + \frac{1}{2} \text{ et } f_Y(y) = \int_0^1 (x+y) dx = \left[ \frac{x^2}{2} + xy \right]_0^1 = y + \frac{1}{2} ;$$

$$4^\circ) f_X(x) = \frac{1}{\pi} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy = \frac{2}{\pi} \sqrt{1-x^2} \text{ et } f_Y(y) = \frac{1}{\pi} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx = \frac{2}{\pi} \sqrt{1-y^2}.$$



#### 4.2.2 Fonction d'un vecteur aléatoire continu

**Proposition 4.2.2.a :** *Loi d'une fonction vectorielle d'un vecteur aléatoire absolument continu (changement de variables)*

Soient un vecteur aléatoire continu  $V = (X_1, \dots, X_m)$  de fonction de densité  $f_V$ , et une fonction  $\varphi$  définie sur  $V(\Omega) \subset \mathbb{R}^m$  (ou sur un sous ensemble  $A$  de  $V(\Omega)$  tel que  $\mathbf{P}(V \notin A) = 0$ ) et à valeurs dans  $\mathbb{R}^m$ .

Si  $\varphi$  est un  $C^1$ -difféomorphisme de  $\mathbb{R}^m$ , c'est-à-dire une application de  $\mathbb{R}^m$  dans  $\mathbb{R}^m$  inversible telle que  $\varphi$  et  $\psi = \varphi^{-1}$  sont continûment différentiables, alors l'application

$$W = \varphi(V) : \Omega \rightarrow \mathbb{R}^m$$

$$\omega \mapsto \varphi(X_1(\omega), \dots, X_m(\omega)),$$

est un vecteur aléatoire continu défini par  $W(\Omega) = \varphi(V(\Omega))$  et par sa fonction de densité  $f_W$  qui est telle que  $f_W(y) = f_V(\psi(y)) \times |J_\psi(y)|$ , pour  $y \in W(\Omega)$  et  $f_W(y) = 0$  sinon.

$J_\psi(y)$  désigne le jacobien de  $\psi$  : si  $\psi(y_1, \dots, y_m) = (\psi_1(y_1, \dots, y_m), \dots, \psi_m(y_1, \dots, y_m))$ , alors

$$J_\psi(y) = \det \begin{pmatrix} \frac{\partial \psi_1}{\partial y_1} & \dots & \frac{\partial \psi_m}{\partial y_1} \\ \vdots & & \vdots \\ \frac{\partial \psi_1}{\partial y_m} & \dots & \frac{\partial \psi_m}{\partial y_m} \end{pmatrix}.$$

Démonstration :

Soit  $B \subset W(\Omega)$ .

Alors  $\mathbf{P}(W \in B) = \mathbf{P}(\varphi(V) \in B) = \mathbf{P}(V \in \varphi^{-1}(B) = \psi(B))$

$$= \int_{\psi(B)} f_V(x_1, \dots, x_m) dx_1 \dots dx_m = \int_{\psi(B)} f_V(x) dx.$$

Puisque  $\psi$  est bijective, le changement de variable  $x = \psi(y)$  donne

$$\mathbf{P}(W \in B) = \int_B f_V(\psi(y)) |J_\psi(y)| dy$$

ce qui établit le résultat :  $f_W(y) = f_V(\psi(y)) \times |J_\psi(y)|$ .

**Application 4.2.2.b :** *Transformation polaire d'un vecteur cartésien du plan*

On prend  $m = 2$ ,  $V = (X, Y)$  et  $W = (R, \Theta)$ , avec  $R(\Omega) = \mathbb{R}_+$ ,  $\Theta(\Omega) = [0; 2\pi[$  et

$$\begin{cases} X = R \cos \Theta \\ Y = R \sin \Theta \end{cases}.$$

On a  $V = \psi(W)$  avec  $\psi : (\rho, \theta) \mapsto (\rho \cos \theta, \rho \sin \theta)$ .

$\psi$  est un  $C^1$ -difféomorphisme de  $\mathbb{R}^2$ , et on a  $J_\psi(y) = \begin{vmatrix} \cos \theta & \sin \theta \\ -\rho \sin \theta & \rho \cos \theta \end{vmatrix} = \rho$ , d'où

$$f_W(\rho, \theta) = \rho \times f_V(\rho \cos \theta, \rho \sin \theta).$$

**Proposition 4.2.2.c :** *Loi d'une fonction scalaire d'un vecteur aléatoire absolument continu*

Soient un vecteur aléatoire continu  $V = (X_1, \dots, X_m)$  de densité  $f$ , et une fonction  $\varphi$  à valeurs réelles définie sur  $X_1(\Omega) \times \dots \times X_m(\Omega)$  (ou sur un sous ensemble  $A$  de  $X_1(\Omega) \times \dots \times X_m(\Omega)$  tel que  $\mathbf{P}(V \notin A) = 0$ ).

L'application :  $Y = \varphi(X_1, \dots, X_m) : \Omega \rightarrow \mathbb{R}$

$$\omega \mapsto \varphi(X_1(\omega), \dots, X_m(\omega)),$$

est une v.a.r absolument continue dont la loi est définie par :

1.  $Y(\Omega) = \varphi(X_1(\Omega), \dots, X_m(\Omega))$  ;
2.  $\forall y \in Y(\Omega), \mathbf{P}(Y \leq y) = \iint \dots \int_{D_y} f(t_1, \dots, t_m) dt_1 \dots dt_m,$

avec  $D_y = \{(x_1, \dots, x_m) \in \mathbb{R}^m \text{ tels que } \varphi(x_1, \dots, x_m) \leq y\}$ .

**Proposition 4.2.2.d :** *Loi de la somme de variables aléatoires absolument continues*

Etant données  $m$  v.a.r absolument continues  $X_1, \dots, X_m$ , la loi de  $Y = X_1 + \dots + X_m$  est définie par :

1.  $Y(\Omega) = X_1(\Omega) + \dots + X_m(\Omega)$  ;
2.  $\forall y \in Y(\Omega), \mathbf{P}(Y \leq y) = \iint \dots \int_{D_y} f(t_1, \dots, t_m) dt_1 \dots dt_m,$

où  $f$  désigne la densité du vecteur  $(X_1, \dots, X_m)$ ,

et  $D_y = \{(x_1, \dots, x_m) \in \mathbb{R}^m \text{ tels que } x_1 + \dots + x_m \leq y\}$ .

**Proposition 4.2.2.e :** *Loi de la somme de deux variables aléatoires absolument continues*

Etant données deux v.a.r absolument continues  $X$  et  $Y$ , la loi de  $Z = X + Y$  est définie par :

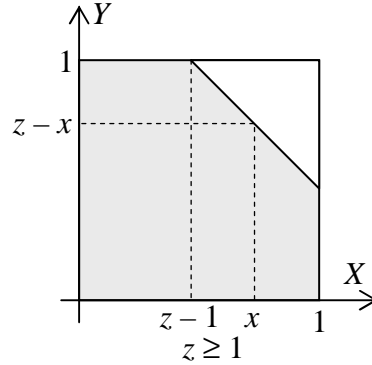
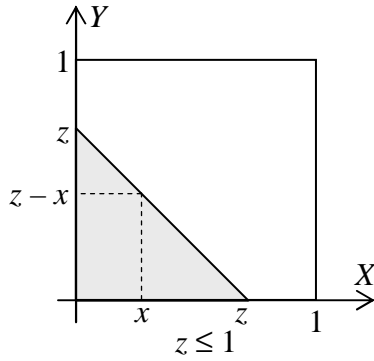
1.  $Z(\Omega) = X(\Omega) + Y(\Omega)$  ;
2.  $\forall z \in Z(\Omega), \mathbf{P}(Z \leq z) = \iint_{D_z} f(x, y) dx dy$  avec  $D_z = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } x + y \leq z\}$ ,

$$\text{soit : } \mathbf{P}(Z \leq z) = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{z-y} f(x, y) dx \right) dy = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{z-x} f(x, y) dy \right) dx.$$

où  $f$  désigne la densité du vecteur  $(X, Y)$ .

**Exemples 4.2.2.f** (reprenant les exemples 4.2.1.g) :

Pour  $m = 2$  et  $V(\Omega) = [0; 1]^2$ , en posant  $V = (X, Y)$  et  $Z = X + Y$ , on a  $Z(\Omega) = [0; 2]$ , et :



$$\text{si } z \leq 1, F_Z(z) = \mathbf{P}(Z \leq z) = \int_0^z \int_0^{z-x} f(x,y) dy dx ;$$

$$\text{si } z \geq 1, F_Z(z) = \mathbf{P}(Z \leq z) = \int_0^{z-1} \int_0^1 f(x,y) dy dx + \int_{z-1}^1 \int_0^{z-x} f(x,y) dy dx$$

On obtient donc pour les trois premiers exemples :

$$1^\circ) \text{ si } z \leq 1 : F_Z(z) = \int_0^z \int_0^{z-x} dy dx = \int_0^z (z-x) dx = \left[ zx - \frac{x^2}{2} \right]_0^z = \frac{z^2}{2} \Rightarrow f_Z(z) = z ;$$

$$\text{si } z \geq 1 : F_Z(z) = z - 1 + \int_{z-1}^1 \int_0^{z-x} dy dx = z - 1 + \left[ zx - \frac{x^2}{2} \right]_{z-1}^1 = -\frac{z^2}{2} + 2z - 1 \Rightarrow f_Z(z) = 2 - z ;$$

$$2^\circ) \text{ si } z \leq 1 : F_Z(z) = \int_0^z \int_0^{z-x} 4xy dy dx = 4 \int_0^z x \left[ \frac{y^2}{2} \right]_0^{z-x} dx = 2 \int_0^z (xz^2 - 2x^2z + x^3) dx = \frac{z^4}{6} \\ \Rightarrow f_Z(z) = \frac{2z^3}{3} ;$$

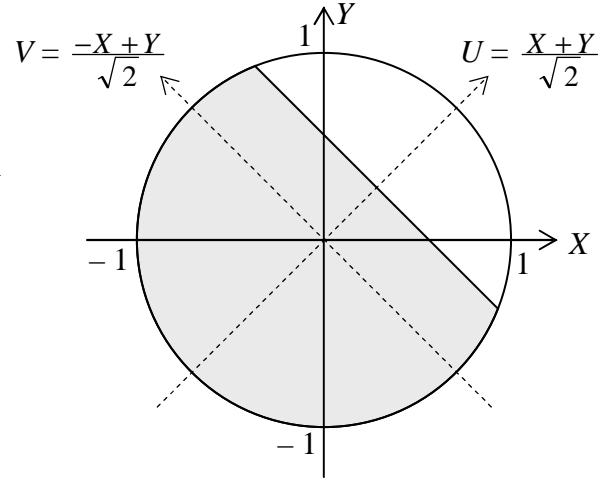
$$\text{si } z \geq 1 : F_Z(z) = 4 \left( \int_0^{z-1} \int_0^1 xy dy dx + \int_{z-1}^1 \int_0^{z-x} xy dy dx \right) = 4 \left( \int_0^{z-1} x \left[ \frac{y^2}{2} \right]_0^1 dx + \int_{z-1}^1 x \left[ \frac{y^2}{2} \right]_0^{z-x} dx \right) \\ = 2 \left( \int_0^{z-1} x dx + \int_{z-1}^1 (xz^2 - 2x^2z + x^3) dx \right) = \frac{1}{6} (-z^4 + 12z^2 - 16z + 6) \Rightarrow f_Z(z) = -\frac{2z^3}{3} + 4z - \frac{8}{3} ;$$

$$3^\circ) \text{ si } z \leq 1 : F_Z(z) = \int_0^z \int_0^{z-x} (x+y) dy dx = \int_0^z \left[ xy + \frac{y^2}{2} \right]_0^{z-x} dx = \frac{1}{2} \int_0^z (z^2 - x^2) dx = \frac{z^3}{3} \\ \Rightarrow f_Z(z) = z^2 ;$$

$$\text{si } z \geq 1 : F_Z(z) = \int_0^{z-1} \int_0^1 (x+y) dy dx + \int_{z-1}^1 \int_0^{z-x} (x+y) dy dx \\ = \int_0^{z-1} \left[ xy + \frac{y^2}{2} \right]_0^1 dx + \int_{z-1}^1 \left[ xy + \frac{y^2}{2} \right]_0^{z-x} dx = \int_0^{z-1} \left( x + \frac{1}{2} \right) dx + \frac{1}{2} \int_{z-1}^1 (z^2 - x^2) dx = -\frac{z^3}{3} + z^2 + \frac{1}{3} \\ \Rightarrow f_Z(z) = 2z - z^2 ;$$

4°) Pour  $m = 2$ , et  $V$  vecteur aléatoire uniforme sur  $V(\Omega) = \{ (x, y) \in \mathbb{R}^2 \text{ t.q. } x^2 + y^2 \leq 1 \}$ , en posant  $V = (X, Y)$  et  $Z = X + Y$ , on a  $Z(\Omega) = [-\sqrt{2}; \sqrt{2}]$  et  $F_Z(z) = \mathbf{P}(Z \leq z)$  est égale à la surface de la zone grisée sur la figure ci-dessous :

$$\begin{aligned} \text{Donc } F_Z(z) &= \frac{1}{\pi} \int_{-1}^{z/\sqrt{2}} \int_{-\sqrt{1-u^2}}^{\sqrt{1-u^2}} dv du \\ &= \frac{2}{\pi} \int_{-1}^{z/\sqrt{2}} \sqrt{1-u^2} du = \frac{1}{\pi} \left[ \arcsin u + u \sqrt{1-u^2} \right]_{-1}^{z/\sqrt{2}} \\ &= \frac{1}{\pi} \left( \arcsin \frac{z}{\sqrt{2}} + \frac{\pi}{2} + \frac{z}{\sqrt{2}} \sqrt{1 - \frac{z^2}{2}} \right) \\ \Rightarrow f_Z(z) &= \frac{\sqrt{2-z^2}}{\pi} \end{aligned}$$



**Proposition 4.2.2.g :** *Loi du quotient de deux variables aléatoires absolument continues*

Si  $X$  et  $Y$  sont deux v.a.r. absolument continues dont la loi conjointe est définie par une densité de probabilité  $f$  sur  $\mathbb{R}^2$ , leur quotient  $\frac{Y}{X}$  est une v.a.r. absolument continue admettant pour densité la fonction  $q$  définie par :

$$q(t) = \int_0^{+\infty} x (f(x, tx) + f(-x, -tx)) dx .$$

Démonstration :

Notons  $f_X$  et  $f_Y$  les densités respectives de  $X$  et  $Y$ .  $X$  étant une v.a.r. absolument continue, la probabilité  $\mathbf{P}(X = 0)$  est nulle, et la variable aléatoire  $Q = \frac{Y}{X}$  peut donc être définie.

Montrons que cette variable aléatoire admet une densité.

Pour tout réel  $z$ , notons :

$$-D_z^+ = \left\{ (x, y) \in \mathbb{R}^2 \mid x > 0 \text{ et } \frac{y}{x} \leq z \right\} ;$$

$$-D_z^- = \left\{ (x, y) \in \mathbb{R}^2 \mid x < 0 \text{ et } \frac{y}{x} \leq z \right\}$$

$$\begin{aligned} \text{On a alors } \mathbf{P}(Q \leq z) &= \iint_{D_z^+ \cup D_z^-} f(x, y) dx dy = \iint_{D_z^+} f(x, y) dx dy + \iint_{D_z^-} f(x, y) dx dy \\ &= \int_0^{+\infty} \left( \int_{-\infty}^{zx} f(x, y) dy \right) dx + \int_{-\infty}^0 \left( \int_{zx}^{+\infty} f(x, y) dy \right) dx . \end{aligned}$$

En effectuant le changement de variable  $y = tx$ , on a :



$$\begin{aligned}
\mathbf{P}(Q \leq z) &= \int_0^{+\infty} \left( \int_{-\infty}^z x f(x, tx) dt \right) dx - \int_{-\infty}^0 \left( \int_{-\infty}^z x f(x, tx) dt \right) dx \\
&= \int_{-\infty}^z \left( \int_0^{+\infty} x f(x, tx) dx \right) dt - \int_{-\infty}^0 \left( \int_{-\infty}^0 x f(x, tx) dx \right) dt \\
&= \int_{-\infty}^z \left( \int_0^{+\infty} x f(x, tx) dx \right) dt + \int_{-\infty}^0 \left( \int_0^{+\infty} x f(-x, -tx) dx \right) dt \\
&= \int_{-\infty}^z \left( \int_0^{+\infty} x [f(x, tx) + f(-x, -tx)] dx \right) dt = \int_{-\infty}^z q(t) dt .
\end{aligned}$$

**Proposition 4.2.2.h :** *Espérance de  $Y = \varphi(X_1, \dots, X_n)$*

Etant donné un vecteur aléatoire continu  $V = (X_1, \dots, X_m)$  de densité  $f$ , et une fonction  $\varphi$  à valeurs réelles définie en tout point de  $V(\Omega)$ , soit la variable aléatoire  $Y = \varphi(X_1, \dots, X_m)$ .

On a alors :

$$\mathbf{E}[Y] = \int_{\mathbb{R}^m} \varphi(x_1, \dots, x_m) f(x_1, \dots, x_m) dx_1 \dots dx_m ,$$

sous réserve que cette intégrale existe.

**Proposition 4.2.2.i :** *Linéarité de l'espérance*

Soient  $X$  et  $Y$  deux v.a.r continues admettant chacune une espérance, et soient  $a$  et  $b$  deux nombres réels. Alors la variable aléatoire  $aX + bY$  admet une espérance, et on a :

$$\mathbf{E}[aX + bY] = a \mathbf{E}[X] + b \mathbf{E}[Y] .$$

Démonstration :

Notons  $f$  la densité du vecteur  $(X, Y)$  et respectivement  $f_X$  et  $f_Y$  les densités des variables  $X$  et  $Y$ .

On a alors :

$$\begin{aligned}
\mathbf{E}[aX + bY] &= \int_{\mathbb{R}^2} (ax + by) f(x, y) dx dy \\
&= a \int_{\mathbb{R}^2} x f(x, y) dx dy + b \int_{\mathbb{R}^2} y f(x, y) dx dy \\
&= a \int_{\mathbb{R}} \left( x \int_{\mathbb{R}} f(x, y) dy \right) dx + b \int_{\mathbb{R}} \left( y \int_{\mathbb{R}} f(x, y) dx \right) dy \\
&= a \int_{\mathbb{R}} x f_X(x) dx + b \int_{\mathbb{R}} y f_Y(y) dy \\
&= a \mathbf{E}[X] + b \mathbf{E}[Y] .
\end{aligned}$$

### 4.3. Vecteurs aléatoires mixtes

#### Définition 4.3.a : Vecteurs aléatoires mixtes

Soient  $n, m \in \mathbb{N}^*$ ,  $X = (X_1, \dots, X_n)$  un vecteur aléatoire discret et  $Y = (Y_1, \dots, Y_m)$  un vecteur aléatoire continu. Alors, le vecteur  $V = [X \ Y] = (X_1, \dots, X_n, Y_1, \dots, Y_m)$  est appelé un *vecteur aléatoire mixte*.

#### Définition 4.3.b : Densité de probabilité mixte discrète/continue

Soit un vecteur aléatoire mixte  $V = [X \ Y]$ , avec  $X = (X_1, \dots, X_n)$  discret et  $Y = (Y_1, \dots, Y_m)$  continu. La *fonction de densité mixte discrète/continue* est une fonction  $f_V : X(\Omega) \times Y(\Omega) \rightarrow \mathbb{R}$  telle que pour tout  $x = (x_1, \dots, x_n) \in X(\Omega)$  et pour toute partie  $B \subset Y(\Omega)$  on a :

$$\mathbf{P}(X = x \cap Y \in B) = \int_B f_V(x, y) dy$$

Remarque :

$$\sum_{x \in X(\Omega)} \int_{Y(\Omega)} f_V(x, y) dy = 1 \text{ et } \forall A \subset X(\Omega) \text{ on a : } \mathbf{P}(V \in A \times B) = \sum_{x \in A} \int_B f_V(x, y) dy.$$

#### Exemple 4.3.c : avec $n = m = 1$

Soit  $V = (X, Y)$  avec  $X(\Omega) = \mathbb{N}^*$  et  $Y(\Omega) = ]0; 1[$ , dont la fonction de densité mixte est donnée par  $f_V(x, y) = y(1 - y)^{x-1}$  sur le support  $X(\Omega) \times Y(\Omega)$ .

#### Définition 4.3.d : Lois marginales

Soit un vecteur aléatoire mixte  $V = [X \ Y]$ , avec  $X = (X_1, \dots, X_n)$  discret et  $Y = (Y_1, \dots, Y_m)$  continu. Les *lois marginales* des vecteurs  $X$  et  $Y$  sont définies de la façon suivante :

$$\forall x = (x_1, \dots, x_n) \in X(\Omega), \mathbf{P}(X = x) = \int_{Y(\Omega)} f_V(x, y) dy$$

$$\text{et } \forall y = (y_1, \dots, y_m) \in Y(\Omega) \text{ par } f_Y(y) = \sum_{x \in X(\Omega)} f_V(x, y).$$

#### Exemple 4.3.c (suite) :

$$f_Y(y) = \sum_{n=1}^{\infty} y(1 - y)^{n-1} = 1, \text{ donc } Y \sim \mathcal{U}_{]0; 1[},$$

$$\text{et } \mathbf{P}(X = n) = \int_0^1 (1 - y)^{n-1} y dy \underset{u=1-y}{=} \int_0^1 (1 - u) u^{n-1} du = \left[ \frac{u^n}{n} - \frac{u^{n+1}}{n+1} \right]_0^1 = \frac{1}{n(n+1)}.$$

## 5. Indépendance et corrélation

### 5.1. Rappels : variables aléatoires indépendantes

**Définition 5.1.a** (rappel) : *Variables aléatoires indépendantes*

Deux variables aléatoires  $X_1$  et  $X_2$  sont dites *indépendantes* si et seulement si pour tout couple  $(B_1, B_2)$  de  $X_1(\Omega) \times X_2(\Omega)$ ,  $X_1^{-1}(B_1)$  et  $X_2^{-1}(B_2)$  sont deux événements indépendants.

On a donc  $\mathbf{P}((X_1 \in B_1) \cap (X_2 \in B_2)) = \mathbf{P}(X_1^{-1}(B_1) \cap X_2^{-1}(B_2))$   
 $= \mathbf{P}(X_1^{-1}(B_1)) \cdot \mathbf{P}(X_2^{-1}(B_2)) = \mathbf{P}(X_1 \in B_1) \cdot \mathbf{P}(X_2 \in B_2)$

**Définition 5.1.b** : *Variables aléatoires mutuellement indépendantes*

$m$  variables aléatoires  $X_1, \dots, X_m$  sont dites *mutuellement indépendantes* (ou *indépendantes entre elles*) si et seulement si pour tout  $m$ -uplet  $(B_1, \dots, B_m)$  de  $X_1(\Omega) \times \dots \times X_m(\Omega)$ , les événements  $X_1^{-1}(B_1), \dots, X_m^{-1}(B_m)$  sont mutuellement indépendants, c'est-à-dire si et seulement si :

$$\forall I \in \llbracket 1, m \rrbracket, \mathbf{P}\left(\bigcap_{i \in I} (X_i \in B_i)\right) = \prod_{i \in I} \mathbf{P}(X_i \in B_i).$$

Remarque : si des variables aléatoires sont mutuellement indépendantes, alors elles sont indépendantes deux à deux, mais la réciproque est fausse.

### 5.2. Vecteurs aléatoires à composantes indépendantes

#### 5.2.1 Vecteurs aléatoires discrets

**Proposition 5.2.1.a** :

Etant données  $m$  v.a.r discrètes  $X_1, \dots, X_m$ , le vecteur aléatoire discret  $V = (X_1, \dots, X_m)$  a ses  $m$  composantes mutuellement indépendantes si et seulement si :

$$\forall (x_1, \dots, x_m) \in X_1(\Omega) \times \dots \times X_m(\Omega), \mathbf{P}[V = (x_1, \dots, x_m)] = \prod_{i=1}^m \mathbf{P}(X_i = x_i)$$

**Exemples 5.2.1.b** (reprenant les exemples 4.1.1.b) :

1°)  $\forall (n, m) \in \llbracket 1, 6 \rrbracket^2$ ,  $\mathbf{P}((X, Y) = (n, m)) = \frac{1}{36}$  et  $\mathbf{P}(X = n) = \mathbf{P}(Y = m) = \frac{1}{6}$ , donc les

composantes  $X$  et  $Y$  sont indépendantes.

2°)  $\forall (n, m) \in \mathbb{N}^2$ ,  $\mathbf{P}((X, Y) = (n, m)) = \frac{1}{2^{n+1}} \frac{e^{-1}}{m!}$ ,  $\mathbf{P}(X = n) = \frac{1}{2^{n+1}}$  et  $\mathbf{P}(Y = m) = \frac{e^{-1}}{m!}$ ,

donc les composantes  $X$  et  $Y$  sont indépendantes.

3°)  $\forall (n, m) \in \mathbb{N}^2$ ,  $\mathbf{P}((X, Y) = (n, m)) = \frac{e^{-n} n^m}{m! 2^{n+1}}$ ,  $\mathbf{P}(X = n) = \frac{1}{2^{n+1}}$

et  $\mathbf{P}(Y = m) = \sum_{n=0}^{+\infty} \frac{e^{-n} n^m}{m! 2^{n+1}}$ , donc les composantes  $X$  et  $Y$  ne sont pas indépendantes.

4°)  $\forall (n, m) \in \{(n, m) \in \mathbb{N}^2 \text{ t. q. } m \leq n\}$ ,  $\mathbf{P}((X, Y) = (n, m)) = \frac{1}{(n+1) 2^{n+1}}$ ,

$\mathbf{P}(X = n) = \frac{1}{2^{n+1}}$  et  $\mathbf{P}(Y = m) = \sum_{n=m}^{+\infty} \frac{1}{(n+1) 2^{n+1}}$ , donc les composantes  $X$  et  $Y$  ne sont pas

indépendantes.

**Proposition 5.2.1.c :**

Si  $m$  v.a.r discrètes  $X_1, \dots, X_m$  sont indépendantes, alors la fonction de répartition du vecteur aléatoire discret  $V = (X_1, \dots, X_m)$  est égale au produit des fonctions de répartitions de ses composantes, et réciproquement.

Démonstration :

$$\begin{aligned}
 F_V(x_1, \dots, x_m) &= \mathbf{P}(X_1 \leq x_1 \cap \dots \cap X_m \leq x_m) \\
 &= \sum_{k_1 \leq x_1} \dots \sum_{k_m \leq x_m} \mathbf{P}(X_1 = k_1 \cap \dots \cap X_m = k_m) \\
 &= \sum_{k_1 \leq x_1} \dots \sum_{k_m \leq x_m} \mathbf{P}(X_1 = k_1) \times \dots \times \mathbf{P}(X_m = k_m) \\
 &= \sum_{k_1 \leq x_1} \mathbf{P}(X_1 = k_1) \times \dots \times \sum_{k_m \leq x_m} \mathbf{P}(X_m = k_m) \\
 &= F_{X_1}(x_1) \times \dots \times F_{X_m}(x_m)
 \end{aligned}$$

### 5.2.2 Vecteurs aléatoires continus

#### Proposition 5.2.2.a :

Etant données  $m$  v.a.r absolument continues  $X_1, \dots, X_m$  de densités respectives  $f_1, \dots, f_m$ , le vecteur aléatoire continu  $V = (X_1, \dots, X_m)$  a ses  $m$  composantes mutuellement indépendantes si et seulement si il admet pour densité la fonction  $f$  définie sur  $\mathbb{R}^m$  par :

$$\forall (x_1, \dots, x_m) \in \mathbb{R}^m, f(x_1, \dots, x_m) = \prod_{i=1}^m f_i(x_i).$$

#### Démonstration :

Supposons les variables  $X_1, \dots, X_m$  mutuellement indépendantes. Alors,  $\forall (x_1, \dots, x_m) \in \mathbb{R}^m$ , en notant  $D = ]-\infty; x_1] \times \dots \times ]-\infty; x_m]$ , on a :

$$\mathbf{P}(V \in D) = \prod_{i=1}^m \mathbf{P}(X_i \leq x_i) = \prod_{i=1}^m \int_{-\infty}^{x_i} f_i(t_i) dt_i = \int_D \prod_{i=1}^m f_i(t_i) dt_1 \dots dt_m,$$

et donc la densité du vecteur  $V$  est  $f(x_1, \dots, x_m) = \prod_{i=1}^m f_i(x_i)$ .

Réciproquement, si la densité du vecteur  $V$  est  $f(x_1, \dots, x_m) = \prod_{i=1}^m f_i(x_i)$ , alors pour tout  $m$ -uplet  $(B_1, \dots, B_m)$  de  $X_1(\Omega) \times \dots \times X_m(\Omega)$ , on a :

$$\begin{aligned} \mathbf{P}((X_1 \in B_1) \cap \dots \cap (X_m \in B_m)) &= \int_{B_1} \int_{B_2} \dots \int_{B_m} \prod_{i=1}^m f_i(t_i) dt_1 \dots dt_m \\ &= \prod_{i=1}^m \int_{B_i} f_i(t_i) dt_i = \prod_{i=1}^m \mathbf{P}(X_i \in B_i). \end{aligned}$$

Les variables  $X_1, \dots, X_m$  sont donc mutuellement indépendantes.

#### Exemples 5.2.2.b (reprenant les exemples 4.2.1.g) :

1°)  $\forall x, y \in [0; 1], f(x, y) = 1, f_X(x) = 1$  et  $f_Y(y) = 1$ , donc les composantes  $X$  et  $Y$  sont indépendantes ;

2°)  $\forall x, y \in [0; 1], f(x, y) = 4xy, f_X(x) = 2x$  et  $f_Y(y) = 2y$ , donc les composantes  $X$  et  $Y$  sont indépendantes ;

3°)  $\forall x, y \in [0; 1], f(x, y) = x + y, f_X(x) = x + \frac{1}{2}$  et  $f_Y(y) = y + \frac{1}{2}$ , donc les composantes  $X$  et  $Y$  ne sont pas indépendantes ;

4°)  $\forall x, y \in [0; 1], f(x, y) = 1, f_X(x) = 2\sqrt{1-x^2}$  et  $f_Y(y) = 2\sqrt{1-y^2}$ , donc les composantes  $X$  et  $Y$  ne sont pas indépendantes.

**Proposition 5.2.2.c :**

Si  $m$  v.a.r absolument continues  $X_1, \dots, X_m$  sont indépendantes, alors la fonction de répartition du vecteur aléatoire continu  $V = (X_1, \dots, X_m)$  est égale au produit des fonctions de répartition de ses composantes, et réciproquement.

Démonstration :

$$\begin{aligned}
 F_V(x_1, \dots, x_m) &= \mathbf{P}(X_1 \leq x_1 \cap \dots \cap X_m \leq x_m) \\
 &= \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_m} f_V(t_1, \dots, t_m) dt_m \dots dt_1 \\
 &= \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_m} f_{X_1}(t_1) \times \dots \times f_{X_m}(t_m) dt_m \dots dt_1 \\
 &= \int_{-\infty}^{x_1} f_{X_1}(t_1) dt_1 \times \dots \times \int_{-\infty}^{x_m} f_{X_m}(t_m) dt_m \\
 &= F_{X_1}(x_1) \times \dots \times F_{X_m}(x_m)
 \end{aligned}$$

**5.3. Corrélation linéaire de deux variables aléatoires****5.3.1 Espérance du produit de deux lois indépendantes**

**Proposition 5.3.1.a :** *Espérance du produit de deux lois indépendantes*

Etant données deux v.a.r indépendantes,  $X$  et  $Y$ , admettant chacune une espérance, leur produit  $XY$  admet une espérance, et :

$$\mathbf{E}[XY] = \mathbf{E}[X] \cdot \mathbf{E}[Y].$$

Démonstration :

Si  $X$  et  $Y$  sont deux variables aléatoires discrètes, alors :

$$\begin{aligned}
 \mathbf{E}[XY] &= \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j \mathbf{P}((X = x_i) \cap (Y = y_j)) \\
 &= \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} (x_i \mathbf{P}(X = x_i)) \cdot (y_j \mathbf{P}(Y = y_j)) \\
 &= \sum_{x_i \in X(\Omega)} x_i \mathbf{P}(X = x_i) \cdot \sum_{y_j \in Y(\Omega)} y_j \mathbf{P}(Y = y_j) = \mathbf{E}[X] \cdot \mathbf{E}[Y].
 \end{aligned}$$

Si  $X$  et  $Y$  sont deux variables aléatoires absolument continues, de densités respectives  $f_X$  et  $f_Y$ , alors le vecteur  $(X, Y)$  a une densité  $f$  telle que :

$$\forall (x, y) \in \mathbb{R}^2, f(x, y) = f_X(x) \cdot f_Y(y).$$

On a donc :

$$\mathbf{E} [ XY ] = \iint_{\mathbb{R}^2} xy f_X(x) f_Y(y) dx dy = \left( \int_{-\infty}^{+\infty} x f_X(x) dx \right) \cdot \left( \int_{-\infty}^{+\infty} y f_Y(y) dy \right) = \mathbf{E} [ X ] \cdot \mathbf{E} [ Y ]$$

Remarque :

La réciproque est fautive, c'est-à-dire que si  $X$  et  $Y$  sont deux v.a.r telles que

$$\mathbf{E} [ XY ] = \mathbf{E} [ X ] \cdot \mathbf{E} [ Y ],$$

alors elles ne sont pas forcément indépendantes.

En prenant, par exemple  $X$  telle que  $\mathbf{E} [ X ] = \mathbf{E} [ X^3 ] = 0$ , et  $Y = X^2$ ,

on a bien  $\mathbf{E} [ XY ] = \mathbf{E} [ X^3 ] = 0$ , et  $\mathbf{E} [ X ] \cdot \mathbf{E} [ Y ] = 0$ , alors que  $X$  et  $Y$  ne sont évidemment pas indépendantes.

### 5.3.2 Covariance

**Définition 5.3.2.a :** *Covariance*

Etant données deux v.a.r admettant chacune une espérance, on définit la *covariance* de  $X$  et  $Y$  par :

$$\mathbf{cov} (X, Y) = \mathbf{E} [ (X - \mathbf{E} [ X ]) (Y - \mathbf{E} [ Y ]) ]$$

**Propositions 5.3.2.b :**

1.  $\mathbf{cov} (X, Y) = \mathbf{E} [ XY ] - \mathbf{E} [ X ] \mathbf{E} [ Y ]$  ;
2. Si  $X$  et  $Y$  sont indépendantes, alors  $\mathbf{cov} (X, Y) = \mathbf{E} [ XY ] - \mathbf{E} [ X ] \mathbf{E} [ Y ] = 0$  ;
3. Etant données deux v.a.r  $X$  et  $Y$ , on a  $\mathbf{V} [ X + Y ] = \mathbf{V} [ X ] + \mathbf{V} [ Y ] + 2 \mathbf{cov} (X, Y)$ .

Dans le cas où  $X$  et  $Y$  sont indépendantes, on a donc  $\mathbf{V} [ X + Y ] = \mathbf{V} [ X ] + \mathbf{V} [ Y ]$ .

Démonstrations :

$$\begin{aligned} 1. \mathbf{cov} (X, Y) &= \mathbf{E} [ (X - \mathbf{E} [ X ]) (Y - \mathbf{E} [ Y ]) ] \\ &= \mathbf{E} [ XY - \mathbf{E} [ Y ] X - \mathbf{E} [ X ] Y + \mathbf{E} [ X ] \mathbf{E} [ Y ] ] \\ &= \mathbf{E} [ XY ] - \mathbf{E} [ \mathbf{E} [ Y ] X ] - \mathbf{E} [ \mathbf{E} [ X ] Y ] + \mathbf{E} [ \mathbf{E} [ X ] \mathbf{E} [ Y ] ] \\ &= \mathbf{E} [ XY ] - \mathbf{E} [ Y ] \mathbf{E} [ X ] - \mathbf{E} [ X ] \mathbf{E} [ Y ] + \mathbf{E} [ X ] \mathbf{E} [ Y ] \\ &= \mathbf{E} [ XY ] - \mathbf{E} [ X ] \mathbf{E} [ Y ]. \end{aligned}$$

2. Ce point découle immédiatement la proposition 7.5.a.

$$\begin{aligned} 3. \mathbf{V} [ X + Y ] &= \mathbf{E} [ (X + Y)^2 ] - (\mathbf{E} [ X + Y ])^2 \\ &= \mathbf{E} [ X^2 + 2XY + Y^2 ] - (\mathbf{E} [ X ] + \mathbf{E} [ Y ])^2 \\ &= \mathbf{E} [ X^2 ] + 2 \mathbf{E} [ XY ] + \mathbf{E} [ Y^2 ] - \mathbf{E} [ X ]^2 - 2 \mathbf{E} [ X ] \mathbf{E} [ Y ] - \mathbf{E} [ Y ]^2 \\ &= \mathbf{V} [ X ] + \mathbf{V} [ Y ] + 2 \mathbf{cov} (X, Y). \end{aligned}$$

### 5.3.3 Corrélation linéaire

#### Définition 5.3.3.a : Coefficient de corrélation linéaire

Si  $X$  et  $Y$  sont deux v.a.r non certaines (i.e. de variance non nulle) et admettant une covariance, on peut définir le *coefficient de corrélation linéaire* entre  $X$  et  $Y$  par :

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\mathbf{V}[X]\mathbf{V}[Y]}} = \frac{\text{cov}(X, Y)}{\sigma[X]\sigma[Y]}.$$

Si  $\rho(X, Y) \neq 0$ , les variables  $X$  et  $Y$  sont dites *corrélées*.

#### Propositions 5.3.3.b : Inégalité de Cauchy-Schwartz

Soient  $X$  et  $Y$  deux v.a.r non certaines (i.e. de variance non nulle) et admettant une covariance.

On a :  $\text{cov}(X, Y)^2 \leq \mathbf{V}[X]\mathbf{V}[Y]$ .

Il en découle immédiatement que  $|\rho(X, Y)| \leq 1$ .

Démonstration :

Comme  $\forall \lambda \in \mathbb{R}, \mathbf{V}[Y - \lambda X] = \lambda^2 \mathbf{V}[X] - 2\lambda \text{cov}(X, Y) + \mathbf{V}[Y] \geq 0$ , on a :

$\Delta' = \text{cov}(X, Y)^2 - \mathbf{V}[X]\mathbf{V}[Y] \leq 0$ , donc  $\rho(X, Y)^2 \leq 1$ , et finalement :

$$|\rho(X, Y)| \leq 1.$$

#### Propositions 5.3.3.c : Liaison fonctionnelle linéaire entre deux v.a.r

Soient  $X$  et  $Y$  deux v.a.r non certaines (i.e. de variance non nulle) et admettant une covariance.

Si  $\rho(X, Y) = 1$ , alors  $Y$  est une fonction affine de  $X$ , et on peut donc écrire  $Y = aX + b$ , avec :

$$a = \frac{\text{cov}(X, Y)}{\mathbf{V}[X]} \text{ et } b = \mathbf{E}[Y] - a \mathbf{E}[X].$$

Démonstration :

Si  $\rho(X, Y) = 1$ , alors  $\Delta' = 0$ , donc pour  $a = \frac{\text{cov}(X, Y)}{\mathbf{V}[X]}$  on a  $\mathbf{V}[Y - aX] = 0$ .

$Y - aX$  est donc une variable certaine, donc il existe  $b$  tel que :

$$\mathbf{P}(Y - aX = b) = \mathbf{P}(Y = aX + b) = 1,$$

ce que l'on peut écrire  $Y = aX + b$ , avec  $b = \mathbf{E}[Y - aX] = \mathbf{E}[Y] - a \mathbf{E}[X]$ .



### 5.3.4 Deux exemples de v.a.r non corrélées linéairement, mais dépendantes

#### Exemple 5.3.4.a :

Soit un vecteur aléatoire discret  $V = (X, Y)$  défini par  $V(\Omega) = \{1, 2, 3\}^2$  et  $\mathbf{P}[V = (x, y)]$

donné par le tableau suivant :

$Y \downarrow X \rightarrow$	1	2	3
1	0,04	0,14	0,02
2	0,18	0,20	0,22
3	0,08	0,06	0,06

La loi marginale de  $X$  est donnée par  $X(\Omega) = \{1, 2, 3\}$  et :

$$\mathbf{P}(X=1) = 0,04 + 0,18 + 0,08 = 0,3 ; \mathbf{P}(X=2) = 0,4 \text{ et } \mathbf{P}(X=3) = 0,3.$$

La loi marginale de  $Y$  est donnée par  $Y(\Omega) = \{1, 2, 3\}$  et :

$$\mathbf{P}(Y=1) = 0,04 + 0,14 + 0,02 = 0,2 ; \mathbf{P}(Y=2) = 0,6 \text{ et } \mathbf{P}(Y=3) = 0,2.$$

On a donc  $\mathbf{E}[X] = 0,3 + 2 \times 0,4 + 3 \times 0,3 = 2$  et  $\mathbf{E}[Y] = 0,2 + 2 \times 0,6 + 3 \times 0,2 = 2$ .

La loi de  $Z = XY$  est donnée par  $Z(\Omega) = \{1, 2, 3, 4, 6, 9\}$  et :

$$\mathbf{P}(Z=1) = 0,04 ; \mathbf{P}(Z=2) = 0,18 + 0,14 = 0,32 ; \mathbf{P}(Z=3) = 0,08 + 0,02 = 0,10 ;$$

$$\mathbf{P}(Z=4) = 0,20 ; \mathbf{P}(Z=6) = 0,06 + 0,22 = 0,28 ; \mathbf{P}(Z=9) = 0,06.$$

On a donc  $\mathbf{E}[XY] = 0,04 + 2 \times 0,32 + 3 \times 0,10 + 4 \times 0,20 + 6 \times 0,28 + 9 \times 0,06 = 4$ .

Finalement  $\text{cov}(X, Y) = \mathbf{E}[XY] - \mathbf{E}[X] \times \mathbf{E}[Y] = 0$ , donc  $X$  et  $Y$  ne sont pas corrélées linéairement.

Elles ne sont pourtant pas indépendantes, puisque :

$$\mathbf{P}(X=1 \cap Y=1) = 0,04 \neq 0,06 = \mathbf{P}(X=1) \times \mathbf{P}(Y=1).$$

#### Exemple 5.3.4.b :

On peut prendre comme exemple deux variables aléatoires  $X$  et  $Y$ , avec  $Y$  qui est une fonction quadratique de  $X$  :  $X$  uniforme sur  $\{-1, 0, 1\}$  et  $Y = X^2$ .

On a alors :  $\mathbf{P}(X=-1) = \mathbf{P}(X=0) = \mathbf{P}(X=1) = 1/3$  ;  $\mathbf{P}(Y=0) = 1/3$  ;  $\mathbf{P}(Y=1) = 2/3$ .

Donc,  $\mathbf{E}[X] = 0$  et  $\mathbf{E}[Y] = 2/3$ . La loi  $XY = X^3$  est, elle aussi, uniforme sur  $\{-1, 0, 1\}$ , donc  $\mathbf{E}[XY] = 0$  et en conséquence  $\text{cov}(X, Y) = \mathbf{E}[XY] - \mathbf{E}[X] \times \mathbf{E}[Y] = 0$ .

## 5.4. Caractéristiques des vecteurs aléatoires

### Définition 5.4.a : *Espérance mathématique*

Etant donné un vecteur aléatoire  $V = (X_1, \dots, X_m)$ , l'*espérance mathématique* de  $V$  est le vecteur :

$$\mathbf{E} [ V ] = ( \mathbf{E} [ X_1 ], \dots, \mathbf{E} [ X_m ] ).$$

### Définition 5.4.b : *Matrice des variances et covariances*

Etant donné un vecteur aléatoire  $V = (X_1, \dots, X_m)$ , la *matrice des variances et covariances* de  $V$  est :

$$\Sigma_V = \begin{pmatrix} \mathbf{V}(X_1) & & & & \\ & \ddots & & & \\ & & \mathbf{V}(X_i) & \cdots & \mathbf{cov}(X_j, X_i) \\ & & \vdots & \ddots & \vdots \\ & & \mathbf{cov}(X_i, X_j) & \cdots & \mathbf{V}(X_j) \\ & & & & \ddots & \\ & & & & & \mathbf{V}(X_m) \end{pmatrix}.$$

### Propriété 5.4.c :

Etant donné un vecteur (colonne) aléatoire  $V = \begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix}$ , on a :

$$\Sigma_V = \mathbf{E} [ (V - \mathbf{E} [ V ]) \cdot {}^t(V - \mathbf{E} [ V ]) ] = \mathbf{E} [ V {}^tV - \mathbf{E} [ V ] {}^t\mathbf{E} [ V ] ].$$

où, pour un vecteur  $X$ , la notation  ${}^tX$  désigne le vecteur transposé de  $X$  et où la définition de l'espérance d'un vecteur aléatoire a été généralisée à la matrice aléatoire  $(V - \mathbf{E} [ V ]) \cdot {}^t(V - \mathbf{E} [ V ])$ .

### Démonstration :

Le terme en  $i^{\text{ème}}$  ligne et  $j^{\text{ème}}$  colonne de  $\Sigma_V$  vaut :

$$\mathbf{E} [ (X_i - \mathbf{E} [ X_i ]) \cdot (X_j - \mathbf{E} [ X_j ]) ] = \mathbf{E} [ X_i X_j ] - \mathbf{E} [ X_i ] \mathbf{E} [ X_j ] = \mathbf{cov} (X_i, X_j).$$

**Proposition 5.4.d :**

Etant donné un vecteur (colonne) aléatoire  $V = \begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix}$ , une matrice  $M \in M_m(\mathbb{R})$  et un vecteur

$$C \in \mathbb{R}^m, \text{ on a : } \mathbf{E} [ MV + C ] = M \cdot \mathbf{E} [ V ] + C,$$

$$\text{et : } \Sigma_{MV+C} = M \cdot \Sigma_V \cdot {}^tM.$$

Démonstration :

La première propriété découle immédiatement de la linéarité de l'espérance mathématique d'une variable aléatoire.

$$\text{On a : } (MV + C) - \mathbf{E} [ MV + C ] = M ( V - \mathbf{E} [ V ] ),$$

$$\text{donc : } \Sigma_{MV+C} = \mathbf{E} [ M ( V - \mathbf{E} [ V ] ) ( V - \mathbf{E} [ V ] ) {}^tM ] = M \cdot \Sigma_V \cdot {}^tM.$$

**Théorème 5.4.e** (admis) :

Une condition nécessaire et suffisante pour qu'une matrice symétrique  $\Sigma$  soit la matrice de variance-covariance d'un vecteur aléatoire est qu'elle soit positive, *i.e.* :  $\forall X \in \mathbb{R}^m$ , vecteur colonne, on a  ${}^tX \Sigma X \geq 0$ .

**Proposition 5.4.f** (admise) :

Toute matrice symétrique positive  $\Sigma \in S_m^+(\mathbb{R})$  est diagonalisable et ses valeurs propres sont positives. Il existe donc une matrice inversible  $P$  orthogonale (*i.e.* telle que  ${}^tP = P^{-1}$ ) et une matrice diagonale  $\Lambda = \text{Diag} (\lambda_i)_{1 \leq i \leq m}$  avec  $\lambda_i \geq 0$  pour tout  $i$ , telles que  $\Sigma = P \Lambda {}^tP$ .

**Proposition 5.4.g** : *Transformation de Mahalanobis*

Soit  $V$  un vecteur aléatoire dont l'espérance mathématique est  $\mathbf{E} [ V ] = \mu$  et la matrice de variance-covariance est  $\Sigma$ . Alors, en notant  $\Sigma^{-1/2} = P \Lambda^{-1/2} {}^tP$ , avec  $\Lambda^{-1/2} = \text{Diag} (\sqrt{\lambda_i})$ , le vecteur aléatoire  $V^* = \Sigma^{-1/2} (V - \mu)$  est un vecteur aléatoire centré réduit, à composantes non corrélées (*i.e.* sa matrice de variance-covariance est la matrice identité).

Démonstration :

D'après la proposition 5.4.d, pour toute matrice  $M$  et tout vecteur  $C$ ,

$$\mathbf{E} [ MV + C ] = M \cdot \mathbf{E} [ V ] + C \text{ et } \Sigma_{MV+C} = M \cdot \Sigma_V \cdot {}^tM.$$

Avec  $M = \Sigma^{-1/2}$  et  $C = -\Sigma^{-1/2} \mu$ , on obtient donc :

$$\mathbf{E} [ V^* ] = \Sigma^{-1/2} \mu - \Sigma^{-1/2} \mu = \mathbf{0}, \text{ et } \Sigma_{V^*} = \Sigma^{-1/2} \Sigma \Sigma^{-1/2} = P \Lambda^{-1/2} {}^tP P \Lambda {}^tP P \Lambda^{-1/2} {}^tP^T = \mathbf{I}_m.$$

## 5.5. Vecteurs gaussiens

### Définition 5.5.a : Vecteur gaussien

Etant donné un vecteur aléatoire  $U = (U_1, \dots, U_m)$  dont les composantes sont des variables aléatoires normales centrées réduites indépendantes, on appelle *vecteur gaussien* de dimension  $m$  toute transformation linéaire du vecteur  $U$  :

$$V = AU + \mu, \text{ avec } A \in M_m(\mathbb{R}) \text{ et } \mu \in \mathbb{R}^m.$$

Comme on a  $\mathbf{E}[U] = 0$  et  $\Sigma_U = \mathbf{I}_m$ , alors :

$$\mathbf{E}[V] = A \mathbf{E}[U] + \mu = \mu \text{ et } \Sigma_V = A \Sigma_U A^T = A A^T.$$

On note alors  $V \sim \mathcal{N}_m(\mu, \Sigma_V)$ .

### Remarque :

Comme  $\Sigma_V = A A^T = P \Lambda P^T = P \Lambda^{1/2} V \Lambda^{1/2} P^T = P \Lambda^{1/2} P P^T \Lambda^{1/2} P^T$ , on a  $A = P \Lambda^{1/2} P = \Sigma_V^{1/2}$ .

En appliquant la transformation de Mahalanobis pour centrer le vecteur  $V$ , on retrouve donc naturellement  $V^* = \Sigma_V^{-1/2}(V - \mu) = A^{-1}(V - \mu) = U$ .

### Proposition 5.5.2.b (admise) : Densité de probabilité d'un vecteur gaussien

La densité de probabilité  $f_V$  d'un vecteur gaussien  $V \sim \mathcal{N}_m(\mu, \Sigma)$  est définie par :

$$\forall X \in \mathbb{R}^m, f_V(X) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(\Sigma)}} e^{-\frac{(X-\mu)^T \Sigma^{-1} (X-\mu)}{2}}.$$

### Cas $m = 2$ :

Avec un vecteur  $V = (X, Y)$ , en notant  $\mathbf{E}[X] = \mu_X$ ,  $\mathbf{E}[Y] = \mu_Y$ ,  $\mathbf{V}[X] = \sigma_X^2$ ,  $\mathbf{V}[Y] = \sigma_Y^2$ ,

et  $\rho = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$  le coefficient de corrélation de  $X$  et  $Y$ , on peut poser :

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{pmatrix}.$$

On a alors :  $\det(\Sigma) = (\sigma_X \sigma_Y)^2 (1 - \rho^2)$  et

$$\Sigma^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} \frac{1}{\sigma_X^2} & -\frac{\rho}{\sigma_X \sigma_Y} \\ -\frac{\rho}{\sigma_X \sigma_Y} & \frac{1}{\sigma_Y^2} \end{pmatrix} = \frac{(\sigma_X^2 \sigma_Y^2)^{-1}}{1 - \rho^2} \begin{pmatrix} \sigma_Y^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_X^2 \end{pmatrix}.$$

On a finalement  $f_V(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x,y)}$  avec :

$$Q(x, y) = \frac{1}{1-\rho^2} \left( \frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X \sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} \right).$$

Cas particuliers :

1°) Vecteur réduit :  $\Sigma = \mathbf{Id}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . On a alors  $\sigma_X = \sigma_Y = 1$  et  $\rho = 0$ , d'où

$$Q(x, y) = (x - \mu_X)^2 + (y - \mu_Y)^2 \text{ et } f_V(x, y) = \frac{1}{2\pi} e^{-\frac{(x-\mu_X)^2 + (y-\mu_Y)^2}{2}}.$$

2°) Vecteur centré réduit :  $\Sigma = \mathbf{Id}_2$  et  $\mu_X = \mu_Y = 0$ , d'où  $Q(x, y) = x^2 + y^2$  et

$$f_V(x, y) = \frac{1}{2\pi} e^{-\frac{x^2 + y^2}{2}}.$$

L'allure de la représentation graphique de  $f_V$  est montrée par la figure ci-dessous.

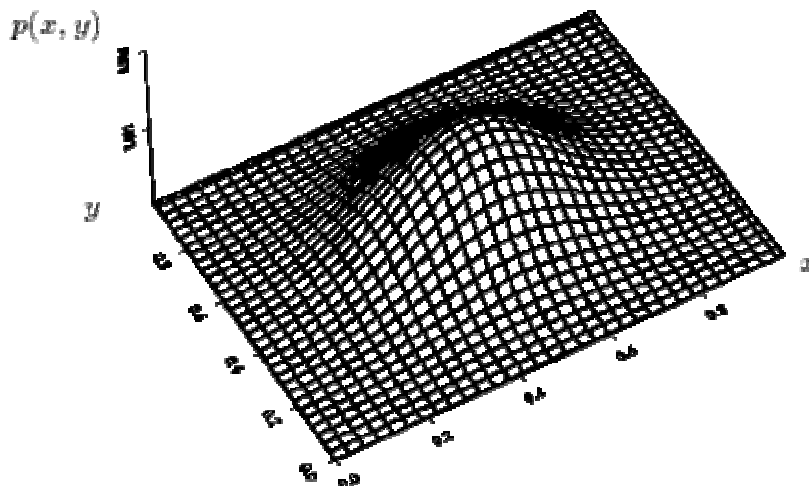


Figure 5.5.2.a : Densité de probabilité d'un vecteur gaussien de dimension 2

## 5.6. Somme de variables aléatoires indépendantes

### 5.6.1. Somme de deux variables aléatoires discrètes indépendantes

#### Proposition 5.6.1.a :

Etant données deux v.a.r discrètes indépendantes  $X$  et  $Y$ , la loi de  $Z = X + Y$  est définie par :

$$\begin{aligned}
 1. \quad Z(\Omega) &= X(\Omega) + Y(\Omega) ; \\
 2. \quad \forall z \in Z(\Omega), \mathbf{P}(Z = z) &= \sum_{\substack{x \in X(\Omega), y \in Y(\Omega) \\ x + y = z}} \mathbf{P}((X = x) \cap (Y = y)) \\
 &= \sum_{\substack{x \in X(\Omega), y \in Y(\Omega) \\ x + y = z}} \mathbf{P}(X = x) \cdot \mathbf{P}(Y = y) \\
 &= \sum_{x \in X(\Omega)} \mathbf{P}(X = x) \cdot \mathbf{P}(Y = z - x) \\
 &= \sum_{y \in Y(\Omega)} \mathbf{P}(X = z - y) \cdot \mathbf{P}(Y = y).
 \end{aligned}$$

#### Exemples 5.6.1.b :

1°) Somme de deux lois binomiales.

Soient  $X$  et  $Y$  deux v.a.r dont les lois binomiales de paramètres respectifs  $(n_1, p)$  et  $(n_2, p)$  sont définies par :

$$\begin{aligned}
 1. \quad X(\Omega) &= \llbracket 0, n_1 \rrbracket \text{ et } \mathbf{P}(X = k) = C_{n_1}^k p^k (1-p)^{n_1-k}, p \in ]0; 1[, n_1 \in \mathbb{N}^*; \\
 2. \quad Y(\Omega) &= \llbracket 0, n_2 \rrbracket \text{ et } \mathbf{P}(Y = k) = C_{n_2}^k p^k (1-p)^{n_2-k}, p \in ]0; 1[, n_2 \in \mathbb{N}^*.
 \end{aligned}$$

La loi de  $Z = X + Y$  est définie par :  $Z(\Omega) = \llbracket 0, n_1 + n_2 \rrbracket$  et

$$\begin{aligned}
 \mathbf{P}(Z = k) &= \sum_{i=0}^k \mathbf{P}(X = i) \mathbf{P}(Y = k - i) = \sum_{i=0}^k C_{n_1}^i p^i (1-p)^{n_1-i} C_{n_2}^{k-i} p^{k-i} (1-p)^{n_2-(k-i)} \\
 &= p^k (1-p)^{n_1+n_2-k} \sum_{i=0}^k C_{n_1}^i C_{n_2}^{k-i}.
 \end{aligned}$$

Or comme d'après la formule du binôme de Newton,  $C_{n_1}^i$  est le coefficient de  $\mathbf{X}^i$  du polynôme  $(1 + \mathbf{X})^{n_1}$ ,  $C_{n_2}^{k-i}$  est le coefficient de  $\mathbf{X}^{k-i}$  du polynôme  $(1 + \mathbf{X})^{n_2}$ , et comme

$(1 + \mathbf{X})^{n_1 + n_2} = (1 + \mathbf{X})^{n_1} (1 + \mathbf{X})^{n_2}$ ,  $\sum_{i=0}^k C_{n_1}^i C_{n_2}^{k-i}$  est alors le coefficient de  $\mathbf{X}^k$  du polynôme

$(1 + \mathbf{X})^{n_1 + n_2}$  et on a donc :  $\sum_{i=0}^k C_{n_1}^i C_{n_2}^{k-i} = C_{n_1 + n_2}^k$  (formule de Vandermonde, cf. §7.6.1.f).

Finalement  $\mathbf{P}(Z = k) = C_{n_1 + n_2}^k p^k (1-p)^{n_1 + n_2 - k}$ , et  $Z$  est donc distribuée suivant une loi binomiale de paramètres  $n_1 + n_2$  et  $p$ .

2°) Somme de deux lois de Poisson.

Soient  $X$  et  $Y$  deux v.a.r dont les lois de Poisson de paramètres respectifs  $\lambda$  et  $\mu$  sont définies par :

$$1. X(\Omega) = \mathbb{N}, \text{ et } \mathbf{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \lambda \in \mathbb{R}_+^* ;$$

$$2. Y(\Omega) = \mathbb{N}, \text{ et } \mathbf{P}(Y = k) = e^{-\mu} \frac{\mu^k}{k!}, \mu \in \mathbb{R}_+^* ;$$

La loi de  $Z = X + Y$  est définie par :  $Z(\Omega) = \mathbb{N}$ , et

$$\begin{aligned} \mathbf{P}(Z = k) &= \sum_{i=0}^k \mathbf{P}(X = i) \mathbf{P}(Y = k - i) = \sum_{i=0}^k e^{-\lambda} \frac{\lambda^i}{i!} e^{-\mu} \frac{\mu^{k-i}}{(k-i)!} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^k C_k^i \lambda^i \mu^{k-i} = e^{-(\lambda+\mu)} \frac{(\lambda + \mu)^k}{k!}, \text{ donc } Z \text{ est distribuée suivant une loi} \end{aligned}$$

de Poisson de paramètre  $\lambda + \mu$ .

## 5.6.2. Somme de deux v.a.r absolument continues indépendantes

### Proposition-définition 5.6.2.a : Produit de convolution

Etant données deux v.a.r absolument continues indépendantes  $X$  et  $Y$  de densités respectives  $f_X$  et  $f_Y$ , la loi de  $Z = X + Y$  est définie par :

$$1. Z(\Omega) = X(\Omega) + Y(\Omega) ;$$

$$2. \forall z \in Z(\Omega), f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) \cdot f_Y(y) dy = \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z-x) dx .$$

La densité  $f_Z$  de la v.a.r absolument continue  $Z$  est le *produit de convolution* des densités  $f_X$  et  $f_Y$  des v.a.r.  $X$  et  $Y$ , et est notée :  $f_Z = f_X * f_Y$ .

Démonstration :

Notons  $f$  la densité du vecteur aléatoire continu  $(X, Y)$ , et  $f_X$  et  $f_Y$  les densités respectives de  $X$  et de  $Y$ . On a déjà vu que :

$$\forall z \in Z(\Omega), \mathbf{P}(Z \leq z) = \iint_{D_z} f(x, y) dx dy \text{ avec } D_z = \{ (x, y) \in \mathbb{R}^2 \text{ t.q. } x + y \leq z \}, \text{ soit :}$$

$$\mathbf{P}(Z \leq z) = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{z-y} f(x, y) dx \right) dy$$

En effectuant le changement de variable  $t = x + y \Leftrightarrow x = z - y$ , on obtient :

$$\mathbf{P}(Z \leq z) = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^z f(t - y, y) dt \right) dy = \int_{-\infty}^z \left( \int_{-\infty}^{+\infty} f(t - y, y) dy \right) dt.$$

Or  $\mathbf{P}(Z \leq z) = \int_{-\infty}^z f_Z(t) dt$ , donc la densité de  $Z$  est  $f_Z(t) = \int_{-\infty}^{+\infty} f(t - y, y) dy$

$X$  et  $Y$  étant indépendantes, on a :  $\forall (t, y) \in \mathbb{R}^2, f(t - y, y) = f_X(t - y) \cdot f_Y(y)$ .

D'où, finalement, en remplaçant la variable muette  $t$  par  $z$  :

$$\forall z \in Z(\Omega), f_Z(z) = \int_{-\infty}^{+\infty} f_X(z - y) \cdot f_Y(y) dy.$$

En intervertissant  $X$  et  $Y$ , on obtient également  $f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z - x) dx$ .

### **Proposition 5.6.2.b :**

Si les supports  $X(\Omega)$  et  $Y(\Omega)$  des densités  $f_X$  et  $f_Y$  ne sont que des parties de  $\mathbb{R}$ , la densité de  $Z = X + Y$  peut s'écrire :

1.  $f_Z(z) = \int_{Y(\Omega)} f_X(z - y) \cdot f_Y(y) dy$  ;
2.  $f_Z(z) = \int_{D_z} f_X(z - y) \cdot f_Y(y) dy$  avec  $D_z = \{ z - x, x \in X(\Omega) \}$ .

Dans le premier cas on est assuré que  $y \in Y(\Omega)$ , mais l'on a pas forcément  $z - y \in X(\Omega)$ .

Dans le second cas on est assuré que  $z - y \in X(\Omega)$ , mais l'on a pas forcément  $y \in Y(\Omega)$ .

### Démonstration :

Le support de  $f_Y$  est  $Y(\Omega)$ , donc :  $f_Z(z) = \int_{Y(\Omega)} f_X(z - y) \cdot f_Y(y) dy$ .

Le support de  $f_X$  est  $X(\Omega)$ , et l'on a :  $z - y \in X(\Omega) \Leftrightarrow \exists x \in X(\Omega)$  tel que  $z - y = x$

$$\Leftrightarrow \exists x \in X(\Omega) \text{ tel que } y = z - x$$

$$\Leftrightarrow y \in D_z = \{ z - x, x \in X(\Omega) \}.$$

On a donc aussi :  $f_Z(z) = \int_{D_z} f_X(z - y) \cdot f_Y(y) dy$ .

### Remarque :

En intervertissant  $X$  et  $Y$ , on a encore :

$$f_Z(z) = \int_{X(\Omega)} f_X(x) \cdot f_Y(z - x) dx = \int_{D_z} f_X(x) \cdot f_Y(z - x) dx, \text{ avec } D_z = \{ z - y, y \in Y(\Omega) \}.$$



### Exemples 5.6.2.c :

1°) Soient  $X$  et  $Y$  deux variables aléatoires uniformes sur  $[0 ; 1]$  et  $Z = X + Y$ .

Donc  $Z(\Omega) = [0 ; 2]$  et :

$$f_Z(z) = \int_{D_t} f_X(x) \cdot f_Y(t-x) dx, \text{ avec } D_t = \{ t-y, y \in [0 ; 1] \} = [t-1 ; t].$$

$$\text{Comme } t-x \in [0 ; 1], f_Z(t) = \int_{t-1}^t f_X(x) dx.$$

$$\text{Donc, } \forall t \leq 1, f_Z(t) = \int_0^t dx = t, \text{ et } \forall t \geq 1, f_Z(t) = \int_{t-1}^1 dx = 2-t.$$

2°) Soient  $X$  et  $Y$  deux variables aléatoires telles que :

$$X(\Omega) = [0 ; 1] \text{ et } \forall x \in X(\Omega), f_X(x) = 2x; Y(\Omega) = [0 ; 1] \text{ et } \forall y \in Y(\Omega), f_Y(y) = 2y.$$

Soit  $Z = X + Y$ . Alors  $Z(\Omega) = [0 ; 2]$  et :

$$f_Z(z) = \int_{D_t} f_X(x) \cdot f_Y(t-x) dx, \text{ avec } D_t = \{ t-y, y \in [0 ; 1] \} = [t-1 ; t].$$

$$\text{Comme } t-x \in [0 ; 1], f_Z(t) = 2 \int_{t-1}^t (t-x) f_X(x) dx.$$

$$\text{Donc, } \forall t \leq 1, f_Z(t) = 4 \int_0^t x(t-x) dx = 4 \left[ t \frac{x^2}{2} - \frac{x^3}{3} \right]_0^t = \frac{2t^3}{3},$$

$$\text{et } \forall t \geq 1, f_Z(t) = 4 \int_{t-1}^1 x(t-x) dx = 4 \left[ t \frac{x^2}{2} - \frac{x^3}{3} \right]_{t-1}^1 = -\frac{2t^3}{3} + 4t - \frac{8}{3}.$$

3°) Soient  $X$  et  $Y$  indépendantes telles que :

$$X(\Omega) = [a ; b] \text{ et } \forall x \in X(\Omega), f_X(x) = \frac{1}{b-a}; Y(\Omega) = [c ; d] \text{ et } \forall y \in Y(\Omega), f_Y(y) = \frac{1}{d-c}.$$

Soit  $Z = X + Y$ , on a donc  $Z(\Omega) = [a+c ; b+d]$  et :

$$f_Z(t) = \int_{D_t} f_X(x) \cdot f_Y(t-x) dx, \text{ avec } D_t = \{ t-y, y \in [c ; d] \} = [t-d ; t-c],$$

$$\text{donc } f_Z(t) = \frac{1}{d-c} \int_{t-d}^{t-c} f_X(x) dx.$$

1<sup>er</sup> cas :  $b-a < d-c$

si  $t-d < a < t-c < b$  :  $t \in [a+c ; b+c]$



$$f_Z(t) = \frac{1}{d-c} \int_a^{t-c} \frac{dx}{b-a} = \frac{1}{d-c} \times \frac{t-c-a}{b-a} \text{ varie linéairement de 0 à } \frac{1}{d-c} \text{ quand } t$$

varie de  $a+c$  à  $b+c$  ;

si  $t-d < a < b < t-c$  :  $t \in [b+c ; a+d]$



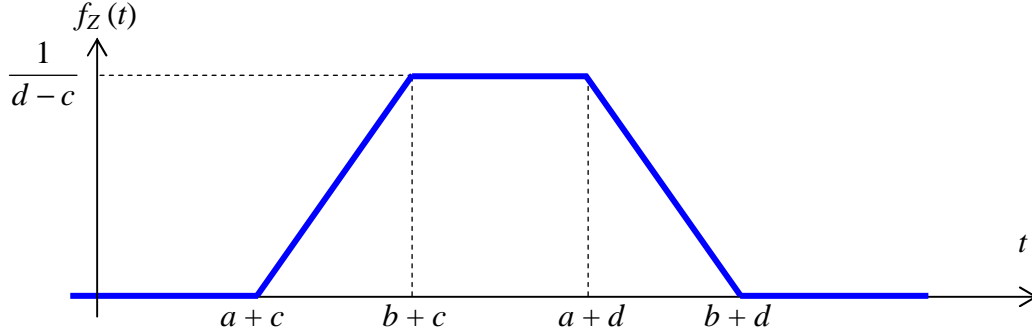
$$f_Z(t) = \frac{1}{d-c} \int_a^b \frac{dx}{b-a} = \frac{1}{d-c} ;$$

si  $a < t-d < b < t-c : t \in [a+d; b+d]$



$$f_Z(t) = \frac{1}{d-c} \int_{t-d}^b \frac{dx}{b-a} = \frac{1}{d-c} \times \frac{b-t+d}{b-a} \text{ varie linéairement de } \frac{1}{d-c} \text{ à } 0 \text{ quand } t$$

varie de  $a+d$  à  $b+d$  ;

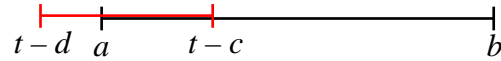


2<sup>ème</sup> cas :  $b-a > d-c$  ; on peut bien sûr se ramener au premier cas en prenant la formule

$$f_Z(t) = \int_{D_t} f_X(t-y) \cdot f_Y(y) dy, \text{ avec } D_t = \{ t-x, x \in X(\Omega) \}, \text{ cependant on peut être amené à}$$

conserver la première formule pour des raisons de facilité de calcul :

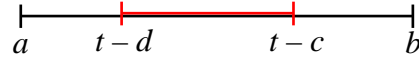
- si  $t-d < a < t-c < b : t \in [a+c; a+d]$



$$f_Z(t) = \frac{1}{d-c} \int_a^{t-c} \frac{dx}{b-a} = \frac{1}{d-c} \times \frac{t-c-a}{b-a} \text{ varie linéairement de } 0 \text{ à } \frac{1}{b-a} \text{ quand } t$$

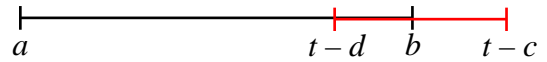
varie de  $a+c$  à  $a+d$  ;

- si  $a < t-d < t-c < b : t \in [a+d; b+c]$



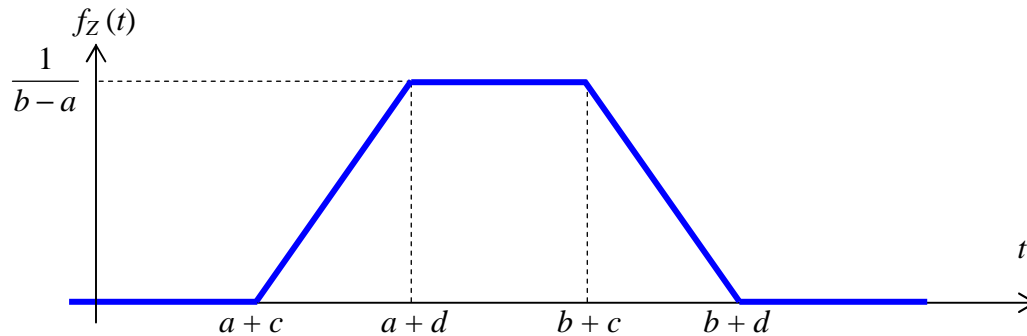
$$f_Z(t) = \frac{1}{d-c} \int_{t-d}^{t-c} \frac{dx}{b-a} = \frac{1}{b-a} ;$$

- si  $a < t-d < b < t-c : t \in [b+c; b+d]$



$$f_Z(t) = \frac{1}{d-c} \int_{t-d}^b \frac{dx}{b-a} = \frac{1}{d-c} \times \frac{b-t+d}{b-a} \text{ varie linéairement de } \frac{1}{b-a} \text{ à } 0 \text{ quand } t$$

varie de  $b+c$  à  $b+d$  ;



3<sup>ème</sup> cas :  $b-a = d-c$  ;



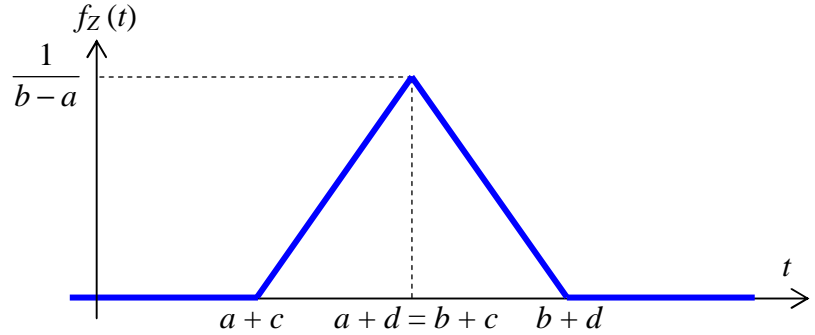
- si  $t-d < a < t-c < b : t \in [a+c; a+d]$

$f_Z(t) = \frac{1}{d-c} \int_a^{t-c} \frac{dx}{b-a} = \frac{1}{d-c} \times \frac{t-c-a}{b-a}$  varie linéairement de 0 à  $\frac{1}{b-a}$  quand  $t$  varie de  $a+c$  à  $a+d$  ;

- si  $a < t-d < b < t-c : t \in [b+c ; b+d]$



$f_Z(t) = \frac{1}{d-c} \int_{t-d}^b \frac{dx}{b-a} = \frac{1}{d-c} \times \frac{b-t+d}{b-a}$  varie linéairement de  $\frac{1}{b-a}$  à 0 quand  $t$  varie de  $b+c$  à  $b+d$  ;



4°) Soient  $X$  et  $Y$  indépendantes telles que :

$$X(\Omega) = \mathbb{R}_+ \text{ et } \forall x \in X(\Omega), f_X(t) = \frac{\lambda^{\alpha_1} t^{\alpha_1-1} e^{-\lambda t}}{\Gamma(\alpha_1)}, \lambda > 0, \alpha_1 > 0 ;$$

$$Y(\Omega) = \mathbb{R}_+ \text{ et } \forall y \in Y(\Omega), f_Y(t) = \frac{\lambda^{\alpha_2} t^{\alpha_2-1} e^{-\lambda t}}{\Gamma(\alpha_2)}, \lambda > 0, \alpha_2 > 0 ,$$

qui sont des lois gamma de paramètres respectifs  $\alpha_1, \lambda$  et  $\alpha_2, \lambda$  (cf. § 12.1.) .

Soit  $Z = X + Y$ . On a alors  $Z(\Omega) = \mathbb{R}_+$  et :

$$f_Z(t) = \int_{D_t} f_X(x) \cdot f_Y(t-x) dx, \text{ avec } D_t = \{ t-y, y \in \mathbb{R}_+ \} = ]-\infty ; t],$$

$$\text{Comme } t-x \in \mathbb{R}_+, f_Z(t) = \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} \int_{-\infty}^t (t-x)^{\alpha_2-1} e^{-\lambda(t-x)} f_X(x) dx ,$$

$$\begin{aligned} \text{Donc pour } t > 0, f_Z(t) &= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1) \Gamma(\alpha_2)} \int_0^t (t-x)^{\alpha_2-1} e^{-\lambda(t-x)} x^{\alpha_1-1} e^{-\lambda x} dx \\ &= \frac{\lambda^{\alpha_1+\alpha_2} e^{-\lambda t}}{\Gamma(\alpha_1) \Gamma(\alpha_2)} \int_0^t (t-x)^{\alpha_2-1} x^{\alpha_1-1} dx \end{aligned}$$

$$\text{Le changement de variable } x = vt \text{ donne } f_Z(t) = \frac{\int_0^1 (1-v)^{\alpha_2-1} v^{\alpha_1-1} dv}{\Gamma(\alpha_1) \Gamma(\alpha_2)} \lambda^{\alpha_1+\alpha_2} t^{\alpha_1+\alpha_2-1} e^{-\lambda t} .$$

Or la densité  $\phi$  d'une loi gamma de paramètres  $\alpha_1 + \alpha_2$  et  $\lambda$  est :

$$\phi(t) = \frac{\lambda^{\alpha_1+\alpha_2} t^{\alpha_1+\alpha_2-1} e^{-\lambda t}}{\Gamma(\alpha_1 + \alpha_2)} .$$

Il existe donc  $K \in \mathbb{R}_+$  telle que  $f_Z = K\phi$ , mais comme  $\int_0^{+\infty} f_Z(t) dt = \int_0^{+\infty} \phi(t) dt = 1$ , alors  $K = 1$ , et donc finalement  $Z$  est distribuée suivant une loi gamma de paramètres  $\alpha_1 + \alpha_2$  et  $\lambda$ .

5°) Soient  $X$  et  $Y$  deux lois gaussiennes centrées réduites.

Soit  $Z = X + Y$ . On a alors  $Z(\Omega) = \mathbb{R}$  et :

$$f_Z(t) = \int_{D_t} f_X(x) \cdot f_Y(t-x) dx, \text{ avec } D_t = \{ t-y, y \in \mathbb{R} \} = \mathbb{R}.$$

$$\text{On a donc } f_Z(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} \cdot e^{-\frac{(t-x)^2}{2}} dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\left(x^2 - tx + \frac{t^2}{2}\right)} dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\left(x - \frac{t}{2}\right)^2 - \frac{t^2}{4}} dx,$$

$$\text{soit } f_Z(t) = \frac{1}{2\pi} e^{-\frac{t^2}{4}} \int_{-\infty}^{+\infty} e^{-u^2} du, \text{ en posant } u = x - \frac{t}{2}.$$

L'intégrale qui figure dans l'expression ci-dessus est l'intégrale de Gauss, et elle vaut  $\sqrt{\pi}$ .

$$\text{On a donc } f_Z(t) = \frac{1}{2\sqrt{\pi}} e^{-\frac{t^2}{4}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}, \text{ avec } \sigma = \sqrt{2}, \text{ et finalement } Z \sim \mathcal{N}(0, \sqrt{2}).$$

Dans le cas, un peu plus général, où  $X \sim \mathcal{N}(\mu_1, \sigma)$  et  $Y \sim \mathcal{N}(\mu_2, \sigma)$  suivent deux lois normales d'espérances respectives  $\mu_1$  et  $\mu_2$ , et de variance commune  $\sigma^2$ , on a, en posant

$$X^* = \frac{X - \mu_1}{\sigma} \text{ et } Y^* = \frac{Y - \mu_2}{\sigma} :$$

$$X^* + Y^* = \frac{X + Y - (\mu_1 + \mu_2)}{\sigma} \sim \mathcal{N}(0, \sqrt{2}), \text{ d'où } X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma).$$

Dans le cas général où  $X \sim \mathcal{N}(\mu_1, \sigma_1)$  et  $Y \sim \mathcal{N}(\mu_2, \sigma_2)$  avec  $\sigma_1 \neq \sigma_2$ , il est préférable d'utiliser une autre méthode pour déterminer la distribution de  $X + Y$ .

## 6. Distributions conditionnelles

### 6.1. Distribution conditionnelle de $X$ sachant $X \in A$

#### 6.1.1. Cas d'une variable aléatoire

##### Définition 6.1.1.a : Loi conditionnelle

Etant donnée une variable aléatoire réelle  $X$  et une partie non vide  $A$  de  $X(\Omega)$ , on définit la loi conditionnelle de  $X$  sachant  $X \in A$ , notée  $X_{|X \in A}$ , par :

$$\forall B \subset X(\Omega), \mathbf{P}(X_{|X \in A} \in B) = \mathbf{P}(X \in B | X \in A) = \frac{\mathbf{P}(X \in A \cap B)}{\mathbf{P}(X \in A)}.$$

On a naturellement  $X_{|X \in A}(\Omega) = A$ .

Cas particulier :

Si  $A$  est un singleton  $\{a\}$ , alors  $X_{|X \in \{a\}} \sim C(a)$ .

##### Exemples 6.1.1.b :

1°)  $X(\Omega) = \llbracket 1, 4 \rrbracket$  et  $\forall k \in X(\Omega), \mathbf{P}(X = k) = \frac{k}{10}$ .

Avec  $A = \{3, 4\}$ , comme  $\mathbf{P}(X \in A) = \frac{7}{10}$ ,  $\forall k \in A, \mathbf{P}(X_{|X \in A} = k) = \frac{k/10}{7/10} = \frac{k}{7}$ .

2°)  $X \sim \mathcal{P}(\lambda)$  et  $A = \mathbb{N}^*$ . On a  $X(\Omega) = \mathbb{N}^*$ , et :

$$\forall k > 0, \mathbf{P}(X_{|X \in A} = k) = \mathbf{P}(X = k | X \neq 0) = \frac{e^{-\lambda} \frac{\lambda^k}{k!}}{1 - e^{-\lambda}}.$$

##### Définition 6.1.1.c : Fonction de répartition conditionnelle

Etant donnée une partie non vide  $A$  de  $X(\Omega)$ , la *fonction de répartition conditionnelle* de  $X$  sachant  $X \in A$  est la fonction de répartition de la variable conditionnelle  $X_{|X \in A}$ , notée  $F_{X|X \in A}$  :

$$\forall x \in A, F_{X|X \in A}(x) = \mathbf{P}((X|A) \in ]-\infty; x]) = \mathbf{P}(X \leq x | X \in A).$$

Remarques :

Cette fonction a toutes les propriétés d'une fonction de répartition, en particulier :

$$\lim_{x \rightarrow +\infty} F_{X|X \in A}(x) = 1.$$

**Définition 6.1.1.d : Densité conditionnelle d'une v.a.r absolument continue**

Si, étant donnée une partie non vide  $A$  de  $X(\Omega)$ , la fonction de répartition conditionnelle de  $X$  sachant  $X \in A$ ,  $F_{X|X \in A}$ , est dérivable par morceaux, alors la variable conditionnelle  $X|_{X \in A}$  est absolument continue, et sa fonction de densité,  $f_{X|X \in A}$ , est la dérivée de  $F_{X|X \in A}$ .

**Exemples 6.1.1.e :**

1°) Si  $X \sim \mathcal{U}_{[0;1]}$ , et  $A = [a; b] \subset [0; 1]$  on a  $\mathbf{P}(X \in A) = b - a$ , d'où

$$\forall x \in [a; b], F_{X|X \in A}(x) = \frac{\mathbf{P}(a \leq X \leq x)}{b - a} = \frac{x - a}{b - a}, \text{ d'où } f_{X|X \in A}(x) = \frac{1}{b - a}.$$

Finalement  $X|_{X \in A} \sim \mathcal{U}_{[a; b]}$ .

2°) Si  $X \sim \mathcal{N}(0, 1)$ , et  $A = \mathbb{R}_+$ , on a  $\mathbf{P}(X \in A) = 1/2$

$$\forall x \in \mathbb{R}_+, F_{X|X \in A}(x) = \frac{\mathbf{P}(0 \leq X \leq x)}{\mathbf{P}(X \geq 0)} = 2(F_X(x) - 1), \text{ d'où } f_{X|X \in A}(x) = 2f_X(x).$$

**Définition 6.1.1.f : Espérance conditionnelle**

Etant donnée une partie non vide  $A$  de  $X(\Omega)$ , l'espérance conditionnelle de  $X$  sachant  $X \in A$ , notée  $\mathbf{E}[X|_{X \in A}]$ , est l'espérance mathématique de la variable conditionnelle  $X|_{X \in A}$  :

$$\text{- si } X|_{X \in A} \text{ est une v.a.r discrète, } \mathbf{E}[X|_{X \in A}] = \sum_{k \in A} k \times \mathbf{P}((X|_{X \in A}) = k) = \frac{\sum_{k \in A} k \times \mathbf{P}(X = k)}{\mathbf{P}(X \in A)} ;$$

$$\text{- si } X|_{X \in A} \text{ est une v.a.r absolument continue, } \mathbf{E}[X|_{X \in A}] = \int_A x f_{X|X \in A}(x) dx.$$

Remarque : On définit de même tous les autres moments conditionnels.

**Proposition 6.1.1.g :**

Etant donné une partition  $(A_i)_{i \in I}$  de  $\Omega$ , et une variable aléatoire réelle  $X$ , la fonction de répartition de  $X$  s'exprime comme combinaison linéaire des fonctions de répartition conditionnelles de  $X$  sachant  $X \in A_i$  :  $F_X(x) = \sum_{i \in I} F_{X|X \in A_i}(x) \times \mathbf{P}(X \in A_i)$ .

**Démonstration :**

Il suffit d'appliquer le théorème des probabilités totales aux définitions des fonctions  $F_X$  et  $F_{X|X \in A_i}$ ,  $i \in I$ .

**Proposition 6.1.1.h :**

Etant donné une partition  $(A_i)_{i \in I}$  de  $\Omega$ , et une variable aléatoire réelle  $X$ , l'espérance mathématique de  $X$  s'exprime comme combinaison linéaire des espérances mathématiques conditionnelles de  $X$  sachant  $X \in A_i$  :  $\mathbf{E}[X] = \sum_{i \in I} \mathbf{E}[X|_{X \in A_i}] \times \mathbf{P}(X \in A_i)$ .

Démonstration :

Il suffit d'appliquer le théorème des probabilités totales aux définitions des espérances  $\mathbf{E}[X]$  et  $\mathbf{E}[X|_{X \in A_i}]$ ,  $i \in I$ .

**6.1.2. Cas d'un vecteur aléatoire****Définitions 6.1.2.a :**

Etant donné un vecteur aléatoire  $V = (X_1, \dots, X_m)$ , et une partie  $A$  de  $V(\Omega)$  telle que  $\mathbf{P}(V \in A) > 0$ , on définit, comme pour une variable aléatoire, la *loi conditionnelle* de  $V$  sachant  $V \in A$ , notée  $V|_{V \in A}$ , par :

$$V|_{V \in A}(\Omega) = A \text{ et } \forall B \subset V(\Omega), \mathbf{P}(V|_{V \in A} \in B) = \mathbf{P}(V \in B | V \in A) = \frac{\mathbf{P}(V \in A \cap B)}{\mathbf{P}(V \in A)}.$$

La *fonction de répartition conditionnelle* est alors :

$$F_{V|_{V \in A}}(x_1, \dots, x_m) = \mathbf{P}(X_1 \leq x_1 \cap \dots \cap X_m \leq x_m | V \in A).$$

Dans le cas où cette fonction de répartition admet une dérivée d'ordre  $m$  par rapport à  $x_1, \dots,$

$$x_m, \text{ la fonction de densité conditionnelle est alors } f_{V|_{V \in A}} = \frac{\partial^m F_{(V|A)}}{\partial x_1 \dots \partial x_m}.$$

L'*espérance conditionnelle* se définit suivant le même principe que pour une variable aléatoire à partir de la définition correspondante pour les vecteurs aléatoires.

On peut, de même, définir des *distributions conditionnelles marginales* :

- dans le cas discret :

$$\mathbf{P}[(X_i|_{V \in A}) = x_i] = \sum_{j \neq i} \sum_{k_j \in X_j(\Omega) \cap A} \mathbf{P}((X_1, \dots, X_m) = (k_1, \dots, k_{i-1}, x_i, k_{i+1}, \dots, k_m) | V \in A)$$

- dans le cas continu :

$$f_{X_i|_{V \in A}}(x_i) = \int \int \dots \int_{\mathbb{R}^{m-1} \cap A} f_{V|_{V \in A}}(t_1, \dots, t_{i-1}, x_i, t_{i+1}, \dots, t_m) dt_1 \dots dt_{i-1} dt_{i+1} \dots dt_m$$

**Proposition 6.1.2.b :**

Si  $V$  est un vecteur discret tel que  $V(\Omega) = \{v_i\}_{i \in I}$ , la loi de  $V_{|V \in A}$  est déterminée par :

$$V_{|V \in A}(\Omega) = A \text{ et } \forall v \in A, \mathbf{P}(V_{|V \in A} = v) = \frac{\mathbf{P}(V = v)}{\mathbf{P}(V \in A)}.$$

Démonstration :

Il suffit de prendre  $B = \{v\}$  dans la formule de la définition **10.2.a** ci-dessus.

**Exemple 6.1.2.c** (reprenant l'exemple **5.3.4.a**) :

Soit un vecteur aléatoire discret  $V = (X, Y)$  défini par  $V(\Omega) = \{1, 2, 3\}^2$  et  $\mathbf{P}[V = (x, y)]$  donné par le tableau ci-contre :

$Y \downarrow X \rightarrow$	1	2	3
1	0,04	0,14	0,02
2	0,18	0,20	0,22
3	0,08	0,06	0,06

Déterminons par exemple la loi conditionnelle de  $V$  sachant que  $X + Y \leq 4$ . On a donc dans ce cas  $A = \{(x, y) \in \{1, 2, 3\}^2 / x + y \leq 4\}$ .

On a  $V_{|V \in A}(\Omega) = A = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)\}$  et  $\mathbf{P}(V \in A) = 0,66$ .

D'après la proposition **6.1.2.b** la loi de  $V_{|V \in A}$  est donc simplement obtenue en divisant les probabilités de ces couples par 0,66.

On obtient donc le tableau ci-contre :

$Y \downarrow X \rightarrow$	1	2	3
1	$0,06$	$0,21$	$0,03$
2	$0,27$	$0,30$	
3	$0,12$		

Les lois marginales conditionnelles de  $X$  et de  $Y$  sachant que  $V \in A$  sont respectivement :

-  $X_{|V \in A}(\Omega) = \{1, 2, 3\}$  et  $\mathbf{P}(X_{|V \in A} = 1) = 0,45$ ,  $\mathbf{P}(X_{|V \in A} = 2) = 0,51$  et  $\mathbf{P}(X_{|V \in A} = 3) = 0,03$ .

-  $Y_{|V \in A}(\Omega) = \{1, 2, 3\}$  et  $\mathbf{P}(Y_{|V \in A} = 1) = 0,30$ ,  $\mathbf{P}(Y_{|V \in A} = 2) = 0,57$  et  $\mathbf{P}(Y_{|V \in A} = 3) = 0,12$ .

Les espérances conditionnelles sont  $\mathbf{E}[X_{|V \in A}] = 1,57$  et  $\mathbf{E}[Y_{|V \in A}] = 1,81$ .

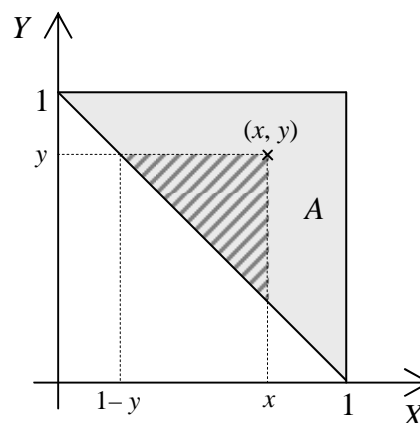
**Exemple 6.1.2.d :**

Soit  $V = (X, Y)$ , uniforme sur  $[0; 1]^2$ .

Considérons une condition sur  $V$ , par exemple  $X + Y \geq 1$ .

La partie  $A$  correspondant à la condition de type  $V \in A$  est donc le triangle grisé sur la figure ci-contre :

On a  $\mathbf{P}(V \in A) = \frac{1}{2}$ .





$F_{V|V \in A}(x, y) = \frac{\mathbf{P}((X \leq x) \cap (Y \leq y) \cap (V \in A))}{\mathbf{P}(V \in A)}$  est la probabilité que  $V$  appartienne à la

surface hachurée sachant qu'il appartient à  $A$  :

$$\begin{aligned} F_{V|V \in A}(x, y) &= 2 \int_{1-y}^x \int_{1-u}^y dv du = 2 \int_{1-y}^x (y + u - 1) du \\ &= 2 \left[ uy + \frac{u^2}{2} - u \right]_{1-y}^x = (x + y - 1)^2. \end{aligned}$$

Cette fonction admet une dérivée d'ordre 2 en  $x$  et  $y$ . On peut donc déterminer  $f_{V|V \in A}$ .

$$\frac{\partial F_{V|V \in A}}{\partial y}(x, y) = 2(x + y - 1), \text{ d'où } f_{V|V \in A}(x, y) = \frac{\partial^2 F_{V|V \in A}}{\partial y \partial x}(x, y) = 2 = \frac{f_V(x, y)}{\mathbf{P}(V \in A)}.$$

Les lois marginales conditionnelles de  $X$  et de  $Y$  sachant  $V \in A$  sont :

$$f_{X|V \in A}(x) = 2 \int_{1-x}^1 dy = 2x \text{ et par symétrie } f_{Y|V \in A}(y) = 2y.$$

Enfin, les espérances conditionnelles de  $X$  et de  $Y$  sachant  $V \in A$  sont :

$$\mathbf{E}[X|V \in A] = \int_0^1 2x^2 dx = \frac{2}{3} = \mathbf{E}[Y|V \in A].$$

## 6.2. Distribution conditionnelle de $X$ sachant $Y$

### 6.2.1. Loi conditionnelle de $X$ sachant que $Y \leq y$

**Définition 6.2.1.a :** Loi conditionnelle de  $X$  sachant que  $Y \leq y$

Etant donné un vecteur aléatoire  $V = (X, Y)$ , la loi conditionnelle de  $X$  sachant que  $Y \leq y$ , notée  $X|Y \leq y$ , est donnée par sa fonction de répartition :

$$F_{X|Y \leq y}(x) = \mathbf{P}(X \leq x | Y \leq y) = \frac{\mathbf{P}((X \leq x) \cap (Y \leq y))}{\mathbf{P}(Y \leq y)} = \frac{F_V(x, y)}{F_Y(y)},$$

Remarques :

$$1^\circ) F_Y(y) = \lim_{x \rightarrow +\infty} F_V(x, y).$$

2°) Cette définition particulière correspond à la définition générale **10.2.a**, en prenant dans le cas particulier  $A = (Y \leq y)$  et  $B = (X \leq x)$ .

**Proposition 6.2.1.b :**

Si  $V = (X, Y)$  est un vecteur discret, on a, pour tout  $x_i \in X(\Omega)$  :

$$\mathbf{P}(X|_{Y \leq y_j} = x_i) = \mathbf{P}(X = x_i | Y \leq y_j) = \frac{\sum_{y_k \leq y_j} \mathbf{P}(V = (x_i, y_k))}{\mathbf{P}(Y \leq y_j)}.$$

**Exemple 6.2.1.c** (reprenant l'exemple 5.3.4.a) :

Soit un vecteur aléatoire discret  $V = (X, Y)$  défini par  $V(\Omega) = \{1, 2, 3\}^2$  et  $\mathbf{P}[V = (x, y)]$  donné par le tableau ci-contre :

$Y \downarrow X \rightarrow$	1	2	3
1	0,04	0,14	0,02
2	0,18	0,20	0,22
3	0,08	0,06	0,06

Déterminons par exemple la loi conditionnelle de  $X$  sachant que  $Y \leq 2$ .

On a  $X|_{Y \leq 2}(\Omega) = \{1, 2, 3\}$ .

$$\forall n \in \{1, 2, 3\}, \text{ on a } \mathbf{P}(X|_{Y \leq 2} = n) = \frac{\sum_{y \leq 2} \mathbf{P}(V = (n, y))}{\mathbf{P}(Y \leq 2)} = \frac{\mathbf{P}(V = (n, 1)) + \mathbf{P}(V = (n, 2))}{\mathbf{P}(Y \leq 2)}.$$

La distribution marginale de  $Y$  montre que  $\mathbf{P}(Y \leq 2) = 0,8$  (le résultat est obtenu en sommant les probabilités des deux premières lignes du tableau).

On a donc :  $\mathbf{P}(X|_{Y \leq 2} = 1) = \frac{0,04 + 0,18}{0,8} = 0,275$  ;  $\mathbf{P}(X|_{Y \leq 2} = 2) = \frac{0,14 + 0,20}{0,8} = 0,425$  et

$$\mathbf{P}(X|_{Y \leq 2} = 3) = \frac{0,02 + 0,22}{0,8} = 0,3.$$

L'espérance conditionnelle de  $X$  sachant que  $Y \leq 2$  est  $\mathbf{E}[X|_{Y \leq 2}] = 2,025$ .

**Exemple 6.2.1.d :**

Soit un vecteur aléatoire continu  $V = (X, Y)$  défini par :

$$V(\Omega) = (\mathbb{R}_+^*)^2 \text{ et } \forall (x, y) \in V(\Omega), f_V(x, y) = (x + y + xy) e^{-(x + y + xy)}.$$

Déterminons la loi conditionnelle de  $X$  sachant que  $Y \leq 2$  :

$$X|_{Y \leq 2}(\Omega) = \mathbb{R}_+^*, \text{ et } F_{X|Y \leq 2}(x) = \frac{F_V(x, 2)}{F_V(x, \infty)} \text{ avec } F_V(y) = \lim_{x \rightarrow +\infty} F_V(x, y).$$

$$\text{On a } F_V(x, y) = \int_0^x \int_0^y (u + v + uv) e^{-(u+v+uv)} dv du$$

$$\int_0^y (u + v + uv) e^{-(u+v+uv)} dv = e^{-u} \left( u \int_0^y e^{-(1+u)v} dv + \int_0^y v(1+u) e^{-(1+u)v} dv \right)$$

En intégrant par parties la deuxième intégrale de la parenthèse, on obtient :

$$\begin{aligned}
\int_0^y (u+v+uv) e^{-(u+v+uv)} dv &= e^{-u} \left( u \int_0^y e^{-(1+u)v} dv - \left[ v e^{-(1+u)v} \right]_0^y + \int_0^y e^{-(1+u)v} dv \right) \\
&= e^{-u} \left( \int_0^y (1+u) e^{-(1+u)v} dv - \left[ v e^{-(1+u)v} \right]_0^y \right) \\
&= -e^{-u} \left[ (1+v) e^{-(1+u)v} \right]_0^y = e^{-u} - (1+y) e^{-(u+y+uy)}
\end{aligned}$$

$$\begin{aligned}
\text{On a donc } F_V(x, y) &= \int_0^x \left( e^{-u} - (1+y) e^{-(u+y+uy)} \right) du \\
&= \left[ -e^{-u} + e^{-(u+y+uy)} \right]_0^x = 1 - e^{-x} - e^{-y} + e^{-x-y-xy}.
\end{aligned}$$

$$\text{Alors } F_Y(y) = \lim_{x \rightarrow +\infty} F_V(x, y) = 1 - e^{-y}.$$

$$\text{Finalement, } F_{X|Y \leq 2}(x) = \frac{F_V(x, 2)}{F_Y(2)} = \frac{1 - e^{-x} - e^{-2} + e^{-(3x+2)}}{1 - e^{-2}} = 1 + \frac{e^{-(3x+2)} - e^{-x}}{1 - e^{-2}},$$

$$\text{et } f_{X|Y \leq 2}(x) = \frac{e^{-x} - 3e^{-(3x+2)}}{1 - e^{-2}}.$$

### 6.2.2. Loi conditionnelle de $X$ sachant que $Y = y$

**Remarque préliminaire 6.2.2.a :** *Loi conditionnelle de  $X$  sachant que  $Y = y$*

La définition générale **6.1.2.a**, ne peut plus s'appliquer dans ce cas. En effet, on obtiendrait  $F_{X|Y=y}(x) = \mathbf{P}(X \leq x | Y = y)$ , or, dans le cas où  $Y$  est une variable aléatoire absolument continue, l'événement  $A = (Y = y)$  est un événement presque impossible.

Cette remarque nous amènera à distinguer les cas discrets et continus.

**Définition 6.2.2.b :** *Loi conditionnelle de  $X$  sachant que  $Y = y$  dans le cas discret*

Etant donné un vecteur aléatoire discret  $V = (X, Y)$ , la loi conditionnelle de  $X$  sachant que  $Y = y_j$ , notée  $X_{|Y=y_j}$ , est donnée par :

$$\forall x_i \in X(\Omega), \mathbf{P}(X_{|Y=y_j} = x_i) = \mathbf{P}(X = x_i | Y = y_j) = \frac{\mathbf{P}(V = (x_i, y_j))}{\mathbf{P}(Y = y_j)} = \frac{p_{ij}}{\sum_i p_{ij}}.$$

**Exemple 6.2.2.c** (reprenant l'exemple 5.3.4.a) :

Soit un vecteur aléatoire discret  $V = (X, Y)$  défini par  $V(\Omega) = \{1, 2, 3\}^2$  et  $\mathbf{P}[V = (x, y)]$  donné par le tableau ci-contre :

$Y \downarrow X \rightarrow$	1	2	3
1	0,04	0,14	0,02
2	0,18	0,20	0,22
3	0,08	0,06	0,06

Déterminons par exemple la loi conditionnelle de  $X$  sachant que  $Y = 2$ .

On a  $X_{|Y=2}(\Omega) = \{1, 2, 3\}$ .

$$\forall n \in \{1, 2, 3\}, \text{ on a } \mathbf{P}(X_{|Y=2} = n) = \frac{\mathbf{P}(V = (n, 2))}{\mathbf{P}(Y = 2)}.$$

La distribution marginale de  $Y$  montre que  $\mathbf{P}(Y = 2) = 0,6$  (le résultat est obtenu en sommant les probabilités de la deuxième ligne du tableau).

$$\text{On a donc : } \mathbf{P}(X_{|Y=2} = 1) = \frac{0,18}{0,6} = 0,3 ; \mathbf{P}(X_{|Y=2} = 2) = \frac{0,2}{0,6} = 0,3\bar{3}$$

$$\text{et } \mathbf{P}(X_{|Y=2} = 3) = \frac{0,22}{0,6} = 0,3\bar{6}.$$

L'espérance conditionnelle de  $X$  sachant que  $Y \leq 2$  est  $\mathbf{E}[X_{|Y \leq 2}] = 2,0\bar{6}$ .

**Définition 6.2.2.d** : Loi conditionnelle de  $X$  sachant que  $Y = y$  dans le cas continu

Soit  $V = (X, Y)$  un vecteur continu, de densité  $f_V$ . Notons respectivement  $f_X$  et  $f_Y$  les densités marginales de  $X$  et  $Y$ . La loi conditionnelle de  $X$  sachant  $Y = y$ , est définie, pour chaque

$$y \in Y(\Omega) \text{ tel que } f_Y(y) > 0, \text{ par : } f_{X|Y=y}(x) = \frac{f_V(x, y)}{f_Y(y)}.$$

Remarque :

Cette définition se justifie en posant  $F_{X|Y=y}(x) = \lim_{h \rightarrow 0^+} \mathbf{P}(X \leq x | y - h \leq Y \leq y + h)$ .

$$\begin{aligned} \mathbf{P}(X \leq x | y - h \leq Y \leq y + h) &= \frac{\mathbf{P}((X \leq x) \cap (y - h \leq Y \leq y + h))}{\mathbf{P}(y - h \leq Y \leq y + h)} \\ &= \frac{\int_{y-h}^{y+h} \left( \int_{-\infty}^x f_V(u, v) du \right) dv}{\int_{y-h}^{y+h} f_Y(v) dv} = \frac{\frac{1}{2h} \int_{y-h}^{y+h} \left( \int_{-\infty}^x f_V(u, v) du \right) dv}{\frac{1}{2h} \int_{y-h}^{y+h} f_Y(v) dv}. \end{aligned}$$

Comme pour tout fonction  $g$ , continue en  $y$ , on a  $\lim_{h \rightarrow 0^+} \frac{1}{2h} \int_{y-h}^{y+h} g(v) dv = g(y)$ , on en déduit que

$$\text{Alors } F_{X|Y=y}(x) = \lim_{h \rightarrow 0^+} \mathbf{P}(X \leq x | y - h \leq Y \leq y + h) = \frac{\int_{-\infty}^x f_V(u, y) du}{f_Y(y)} = \int_{-\infty}^x \frac{f_V(u, y)}{f_Y(y)} du,$$

d'où le résultat, obtenu par dérivation de  $F_{X|Y=y}$ .

**Exemple 6.2.2.e** (reprenant l'exemple 6.2.1.d) :

Soit un vecteur aléatoire continu  $V = (X, Y)$  défini par :

$$V(\Omega) = (\mathbb{R}_+^*)^2 \text{ et } \forall (x, y) \in V(\Omega), f_V(x, y) = (x + y + xy) e^{-(x+y+xy)}.$$

Déterminons la loi conditionnelle de  $X$  sachant que  $Y = y$  :

$$X_{|Y=y}(\Omega) = \mathbb{R}_+^*, \text{ et } f_{X|Y=y}(x) = \frac{f_V(x, y)}{f_Y(y)}.$$

$$\begin{aligned} \text{On a } f_Y(y) &= \int_0^{+\infty} (x + y + xy) e^{-(x+y+xy)} dx \\ &= e^{-y} \left( y \int_0^{+\infty} e^{-(1+y)x} dx + \int_0^{+\infty} x(1+y) e^{-(1+y)x} dx \right) \end{aligned}$$

En intégrant par parties la deuxième intégrale de la parenthèse, on obtient :

$$\begin{aligned} f_Y(y) &= e^{-y} \left( y \int_0^{+\infty} e^{-(1+y)x} dx - \left[ x e^{-(1+y)x} \right]_0^{+\infty} + \int_0^{+\infty} e^{-(1+y)x} dx \right) \\ &= e^{-y} \int_0^{+\infty} (1+y) e^{-(1+y)x} dx = -e^{-y} \left[ (1+y) e^{-(1+y)x} \right]_0^{+\infty} = e^{-y}. \end{aligned}$$

$$\text{On a donc, } f_{X|Y=y}(x) = \frac{f_V(x, y)}{f_Y(y)} = \frac{(x + y + xy) e^{-(x+y+xy)}}{e^{-y}} = (x + y + xy) e^{-(x+xy)}.$$

$$\text{Finalement } f_{X|Y=2}(x) = (3x + 2) e^{-(2x+1)}.$$

**Définition 6.2.2.f** : Loi conditionnelle de  $X$  sachant que  $Y = y$  dans le cas mixte

Premier cas :  $X$  discret et  $Y$  continu ;

Soient  $X$  un vecteur aléatoire discret,  $Y$  un vecteur aléatoire continu de densité  $f_Y$  et  $V = [X Y]$  un vecteur aléatoire mixte, de densité mixte  $f_V$ .

La loi conditionnelle de  $X$  sachant  $Y = y$ , est définie, pour chaque  $y \in Y(\Omega)$  tel que  $f_Y(y) > 0$ ,

$$\text{par : } \mathbf{P}(X = x | Y = y) = \frac{f_V(x, y)}{f_Y(y)}.$$

Deuxième cas :  $X$  continu et  $Y$  discret ;

Soient  $X$  un vecteur aléatoire continu de densité  $f_X$ ,  $Y$  un vecteur aléatoire discret et  $V = [X Y]$  un vecteur aléatoire mixte, de densité mixte  $f_V$ .

La loi conditionnelle de  $X$  sachant  $Y = y$ , est définie, pour chaque  $y \in Y(\Omega)$  tel que

$$\mathbf{P}(Y = y) > 0, \text{ par : } f_{X|Y=y}(x) = \frac{f_V(x, y)}{\mathbf{P}(Y = y)}.$$

**Exemple 6.2.2.g :**

Soit  $V = (X, Y)$  avec  $X(\Omega) = \mathbb{N}^*$  et  $Y(\Omega) = ]0; 1[$ , dont la fonction de densité mixte est donnée par  $f_V(x, y) = y(1-y)^{x-1}$  sur le support  $X(\Omega) \times Y(\Omega)$ .

Les lois marginales de  $X$  et  $Y$  sont données par :

$$f_Y(y) = \sum_{n=1}^{\infty} y(1-y)^{n-1} = 1, \text{ donc } Y \sim \mathcal{U}_{]0; 1[},$$

$$\text{et } \mathbf{P}(X = n) = \int_0^1 (1-y)^{n-1} y \, dy \underset{u=1-y}{=} \int_0^1 (1-u) u^{n-1} \, du = \left[ \frac{u^n}{n} - \frac{u^{n+1}}{n+1} \right]_0^1 = \frac{1}{n(n+1)}.$$

La loi conditionnelle de  $X$  sachant que  $Y = p$  est :

$$\mathbf{P}(X = n | Y = p) = \frac{f_V(n, p)}{f_Y(p)} = p(1-p)^{n-1}.$$

$X_{|Y=p}$  est donc distribuée suivant une loi géométrique de paramètre  $p$ .

La loi conditionnelle de  $Y$  sachant que  $X = n$  est :

$$f_{Y|X=n}(p) = \frac{f_V(n, p)}{\mathbf{P}(X = n)} = n(n+1)p(1-p)^{n-1}.$$

**6.2.3. Cas de deux variables indépendantes****Proposition 6.2.3.a :**

Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes, alors les variables conditionnelles  $X_{|Y \leq y}$  et  $X_{|Y=y}$  ont la même distribution que la variable  $X$ .

Démonstration :

$\forall x \in X(\Omega)$ , on a :

$$F_{X|Y \leq y}(x) = \mathbf{P}(X \leq x | Y \leq y) = \frac{\mathbf{P}((X \leq x) \cap (Y \leq y))}{\mathbf{P}(Y \leq y)} = \mathbf{P}(X \leq x) = F_X(x).$$

Si le vecteur aléatoire  $(X, Y)$  est discret, alors  $\mathbf{P}(X_{|Y=y} = x) = \mathbf{P}(X = x | Y = y) = \mathbf{P}(X = x)$ .

Si le vecteur aléatoire  $(X, Y)$  est continu, alors  $f_{X|Y=y}(x) = \frac{f_V(x, y)}{f_Y(y)} = \frac{f_X(x) \times f_Y(y)}{f_Y(y)} = f_X(x)$ .

#### 6.2.4. Espérance et variances conditionnelles de $X$ sachant $Y$

**Définition 6.2.4.a :** *Espérance conditionnelle de  $X$  sachant que  $Y = y$*

Soient  $X$  et  $Y$  deux variables aléatoires. Pour tout  $y \in Y(\Omega)$  on définit l'espérance conditionnelle de  $X$  sachant que  $Y = y$  par :

- Si  $X$  est discrète :  $\mathbf{E} [ X | Y = y ] = \sum_{x \in X(\Omega)} x \times \mathbf{P} (X = x | Y = y) ;$
- Si  $X$  est continue :  $\mathbf{E} [ X | Y = y ] = \int_{X(\Omega)} x \times f_{X|Y=y}(x) dx ;$

Remarque :

Dans le cas où le vecteur  $V = (X, Y)$  est mixte de densité mixte  $f_V$ :

1°) Si  $X$  est discrète et  $Y$  continue,  $\mathbf{P} (X = x | Y = y) = \frac{f_V(x, y)}{f_Y(y)} ;$

2°) Si  $X$  est continue et  $Y$  discrète,  $f_{X|Y=y}(x) = \frac{f_V(x, y)}{\mathbf{P}(Y = y)} ;$

**Exemple 6.2.4.b :**

Soit  $V = (X, Y)$  avec  $X(\Omega) = \mathbb{N}^*$  et  $Y(\Omega) = ] 0 ; 1 [$ , dont la fonction de densité mixte est donnée par  $f_V(x, y) = y(1 - y)^{x-1}$  sur le support  $X(\Omega) \times Y(\Omega)$ .

On a vérifié précédemment que quand  $Y = p$ ,  $X|_{Y=p}$  est distribuée suivant une loi géométrique de paramètre  $p$  et que  $f_{Y|X=n}(p) = \frac{f_V(n, p)}{\mathbf{P}(X = n)} = n(n + 1)p(1 - p)^{n-1}$ .

On a donc naturellement  $\mathbf{E} [ X | Y = p ] = \sum_{n=1}^{\infty} n p(1 - p)^{n-1} = \frac{1}{p}$ , espérance d'une loi géométrique de paramètre  $p$ .

$$\begin{aligned}
 \text{De plus, } \mathbf{E} [ Y | X = n ] &= \int_{Y(\Omega)} y \times f_{Y|X=n}(y) dy \\
 &= \int_0^1 p n(n + 1)(1 - p)^{n-1} p dp \\
 &= n(n + 1) \int_0^1 (1 - u)^2 u^{n-1} du \\
 &= n(n + 1) \int_0^1 (u^{n-1} - 2u^n + u^{n+1}) du \\
 &= n(n + 1) \left[ \frac{1}{n} - \frac{2}{n+1} + \frac{1}{n+2} \right]_0^1 = \frac{2}{n+2}.
 \end{aligned}$$

**Définition 6.2.4.c : Espérance conditionnelle de  $X$  sachant  $Y$** 

L'espérance conditionnelle de  $X$  sachant  $Y$ , notée  $\mathbf{E} [ X | Y ]$  est une fonction  $\varphi(Y)$ , la fonction  $\varphi$  étant définie par  $\varphi(y) = \mathbf{E} [ X | Y = y ]$  si  $y \in Y(\Omega)$  et  $\varphi(y) = 0$  sinon.

**Proposition 6.2.4.d : (admise)**

L'espérance conditionnelle de  $X$  sachant  $Y$  est la fonction de  $Y$ ,  $Z = \varphi(Y)$ , qui minimise  $\mathbf{E} [ (X - Z)^2 ]$ . On peut considérer que c'est la meilleure "estimation" de  $X$  par une fonction de  $Y$ .

**Exemple 6.2.4.e :**

Soit  $V = (X, Y)$  avec  $X(\Omega) = \mathbb{N}^*$  et  $Y(\Omega) = ] 0 ; 1 [$ , dont la fonction de densité mixte est donnée par  $f_V(x, y) = y(1 - y)^{x-1}$  sur le support  $X(\Omega) \times Y(\Omega)$ .

La loi marginale de  $Y$  est uniforme sur  $] 0 ; 1 [$ , et la loi de  $X_{|Y=p}$  est une loi géométrique de paramètre  $p$ .

Comme  $\mathbf{E} [ X | Y = p ] = \frac{1}{p}$ , on a :  $\mathbf{E} [ X | Y ] = \frac{1}{Y}$ . Ceci signifie que la meilleure fonction de

$Y$  qui "estime" une loi géométrique dont le paramètre est la valeur prise par  $Y$  est  $\frac{1}{Y}$ .

De même, comme  $\mathbf{E} [ Y | X = n ] = \frac{2}{n+2}$ , on a  $\mathbf{E} [ Y | X ] = \frac{2}{X+2}$ , ce qui signifie que la meilleure "estimation" de  $Y$  (c'est-à-dire de la loi qui donne le paramètre de  $X$ ) en fonction de  $X$  est  $\frac{2}{X+2}$ .

**Définition 6.2.4.f : Variance conditionnelle de  $X$  sachant  $Y$** 

$$\mathbf{V} [ X | Y ] = \mathbf{E} [ (X - \mathbf{E} [ X | Y ])^2 | Y ] = \mathbf{E} [ X^2 | Y ] - (\mathbf{E} [ X | Y ])^2.$$

Interprétation :  $\mathbf{E} [ X | Y ]$  représente une variable aléatoire permettant d'estimer  $X$  quand on connaît  $Y$  et  $\mathbf{V} [ X | Y ]$  représente l'erreur quadratique moyenne de cette estimation.

**Proposition 6.2.4.g : (admise)**

$$\mathbf{V} [ X ] = \mathbf{E} [ \mathbf{V} [ X | Y ] ] + \mathbf{V} [ \mathbf{E} [ X | Y ] ].$$



## 7. Fonctions génératrices et caractéristiques

### 7.1. Fonctions génératrices des probabilités d'une variable entière

**Définition 7.1.a :** *Fonction génératrice des probabilités d'une variable entière*

Etant donnée une variable aléatoire entière  $X$  (c'est-à-dire telle que  $X(\Omega) \subset \mathbb{N}$ ), on appelle *fonction génératrice des probabilités* de  $X$  (ou, plus simplement, *fonction génératrice* de  $X$ , quand il n'y a pas de risque de confusion), et on note  $G_X$ , la fonction définie sur  $\{ t \in \mathbb{R} / \mathbf{E} [ t^X ] \text{ existe} \}$  par  $G_X : t \mapsto G_X(t) = \mathbf{E} [ t^X ]$ .

**Proposition 7.1.b :** *Existence de la fonction génératrice des probabilités*

Soit  $X$  une variable aléatoire entière. Si l'on note  $p_k = \mathbf{P}(X = k)$ ,  $G_X$  est donnée par la série entière  $\sum_{k=0} p_k t^k$ , et est donc définie sur le disque de convergence de cette série entière.

En particulier, si  $X$  est bornée son rayon de convergence est infini.

**Exemples 7.1.c :**

1°) Loi binomiale.

Soit  $X \sim \mathcal{B}(n, p)$ . Pour tout  $t \in \mathbb{R}$ , on a :

$$G_X(t) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} t^k = \sum_{k=0}^n \binom{n}{k} (pt)^k q^{n-k} = (pt + q)^n.$$

2°) Loi géométrique.

Soit  $X \sim \mathcal{P}(1, p)$ . Pour tout  $t \in \mathbb{R}$ , tel que  $|t| < \frac{1}{q}$  on a :

$$G_X(t) = \sum_{k=1}^{+\infty} p q^{k-1} t^k = pt \sum_{k=1}^{+\infty} (qt)^{k-1} = pt \sum_{k=0}^{+\infty} (qt)^k = \frac{pt}{1 - qt}.$$

3°) Loi de Poisson.

Soit  $X \sim \mathcal{P}(\lambda)$ . Pour tout  $t \in \mathbb{R}$ , on a :

$$G_X(t) = \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} t^k = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda t)^k}{k!} = e^{\lambda(t-1)}.$$

**Propositions 7.1.d :**

1°)  $G_X$  est toujours définie en 1, et  $G_X(1) = 1$ . Le rayon de convergence est donc supérieur ou égal à 1.

2°)  $\forall a, b \in \mathbb{R}, G_{aX+b}(t) = \mathbf{E}[t^{(aX+b)}] = t^b \mathbf{E}[t^{aX}] = t^b G_X(t^a)$ .

**Proposition 7.1.e :**

Si  $X$  est une variable aléatoire entière dont la fonction génératrice  $G_X$  a un rayon de convergence supérieur à 1, alors  $G_X$  est infiniment dérivable, et  $\forall n \in \mathbb{N}$ , on a:

3.  $G_X^{(n)}(0) = n! p_n$  (d'où le nom de fonction génératrice des probabilités) ;
4.  $G_X^{(n)}(1) = \mathbf{E}[X(X-1) \dots (X-n+1)]$ .

On en déduit :  $\mathbf{E}[X] = G_X'(1)$  et  $\mathbf{V}[X] = G_X''(1) + G_X'(1) - (G_X'(1))^2$ .

Démonstration :

$$G_X^{(n)}(t) = \left( \sum_{k \geq 0} p_k t^k \right)^{(n)} = \sum_{k \geq n} \left( p_k t^k \right)^{(n)} = \sum_{k \geq n} p_k k(k-1) \dots (k-n+1) t^{k-n}, \text{ d'où les résultats}$$

en appliquant à  $t = 0$  et  $t = 1$ .

**Proposition 7.1.f :**

Si  $X$  et  $Y$  ont même fonction génératrice, alors elles sont identiquement distribuées.

Démonstration :

Si  $G_X$  et  $G_Y$  sont définies pour  $t \in \mathbb{R}$ , alors :  $G_X(t) = \sum_{k=0}^{+\infty} \mathbf{P}(X=k) t^k$  et  $G_Y(t) = \sum_{k=0}^{+\infty} \mathbf{P}(Y=k) t^k$ .

Pour tout  $n \in \mathbb{N}$ , d'après la formule de Taylor, le coefficient de  $t^n$  dans la série entière  $G_X(t)$

(respectivement  $G_Y(t)$ ) est  $\frac{G_X^{(n)}(0)}{n!}$  (respectivement  $\frac{G_Y^{(n)}(0)}{n!}$ ).

Donc  $\frac{G_X^{(n)}(0)}{n!} = \mathbf{P}(X=n)$  et  $\frac{G_Y^{(n)}(0)}{n!} = \mathbf{P}(Y=n)$ .

Finalement, si  $G_X = G_Y$ , alors  $G_X^{(n)}(0) = G_Y^{(n)}(0)$ , et donc pour tout  $n \in \mathbb{N}$ :

$$\mathbf{P}(X=n) = \mathbf{P}(Y=n).$$

**Proposition 7.1.g :**

Si  $X$  et  $Y$  sont deux variables entières indépendantes admettant chacune une fonction génératrice, alors la variable aléatoire  $X + Y$  admet aussi une fonction génératrice, et on a :

$$G_{X+Y} = G_X \times G_Y.$$

Démonstration :

Pour tout  $t \in \mathbb{R}$ ,  $G_{X+Y}(t) = \mathbf{E} [ t^{X+Y} ] = \mathbf{E} [ t^X t^Y ]$ . Or, comme  $X$  et  $Y$  sont indépendantes,  $t^X$  et  $t^Y$  le sont également, donc  $\mathbf{E} [ t^X t^Y ] = \mathbf{E} [ t^X ] \mathbf{E} [ t^Y ] = G_X(t) \times G_Y(t)$ .

**Corollaire 7.1.h** (admis) :

Si  $(X_k)_{k \in K}$  est une famille de variables entières mutuellement indépendantes admettant chacune une fonction génératrice  $G_{X_k}$ , alors la variable aléatoire  $Z = \sum_{k \in K} X_k$  admet aussi une

fonction génératrice  $G_Z$ , et on a  $G_Z = \prod_{k \in K} G_{X_k}$ .

**Proposition 7.1.i** : *Loi de la somme de variables aléatoires mutuellement indépendantes*

Si l'on considère  $n$  variables entières mutuellement indépendantes  $X_1, \dots, X_n$ , admettant chacune une fonction génératrice, alors la fonction génératrice de  $Z = \sum_{i=1}^n X_i$  se calcule facilement en appliquant la proposition 10.1.d.

Si la fonction génératrice de  $Z$  est déjà connue comme étant la fonction génératrice d'une loi  $\mathcal{L}$ , alors en vertu de la proposition 10.1.c,  $Z \sim \mathcal{L}$ .

**Exemples 7.1.j** :

1°) La fonction génératrice d'une loi de Bernoulli de paramètre  $p$  est  $t \mapsto q + pt$ . Toute loi binomiale de paramètres  $n$  et  $p$  étant la somme de  $n$  lois de Bernoulli indépendantes, sa fonction génératrice est donc  $t \mapsto (q + pt)^n$ .

2°) Toute loi de Pascal de paramètres  $r$  et  $p$  étant la somme de  $r$  lois géométriques, sa fonction génératrice est donc  $t \mapsto \left( \frac{pt}{1-qt} \right)^r$ .

3°) La somme de deux lois de Poisson indépendantes est une loi de Poisson.

En effet, si  $X \sim \mathcal{P}(\lambda)$  et  $Y \sim \mathcal{P}(\mu)$ , alors pour tout  $t \in \mathbb{R}$  :

$$G_{X+Y}(t) = G_X(t) \times G_Y(t) = e^{\lambda(t-1)} \times e^{\mu(t-1)} = e^{(\lambda+\mu)(t-1)}, \text{ donc en vertu de la proposition 10.1.c:}$$

$$(X + Y) \sim \mathcal{P}(\lambda + \mu).$$

## 7.2. Fonctions génératrices des moments

### Définition 7.2.a : Fonction génératrice des moments

Etant donnée une v.a.r  $X$ , on appelle *fonction génératrice des moments* de  $X$ , et on note  $L_X$ , la fonction définie sur  $\{ t \in \mathbb{R} / \mathbf{E} [ e^{tX} ] \text{ existe} \}$  par  $L_X : t \mapsto L_X(t) = \mathbf{E} [ e^{tX} ]$ .

### Remarques :

1°) Si  $X$  est une variable entière, alors pour tout  $t \in \mathbb{R}$  pour lequel  $G_X$  et  $L_X$  sont définies, on a  $L_X(t) = G_X(e^t)$ .

2°) Il existe des lois dont la fonction génératrice des moments n'est définie que pour  $t = 0$ , comme par exemple la loi de Cauchy, de densité  $x \mapsto \frac{1}{\pi(1+x^2)}$  sur  $\mathbb{R}$ .

### Propositions 7.2.b : Existence et calcul de la fonction génératrice des moments

1°) Si  $X$  est une v.a.r. discrète telle que  $X(\Omega) = \{ x_k ; k \in K \subset \mathbb{N} \}$ , alors  $L_X(t)$  existe si et seulement si la série  $\sum_{k \in K} \mathbf{P}(X = x_k) e^{tx_k}$  converge.

En particulier, si  $X$  est bornée cette série converge.

2°) Si  $X$  est une v.a.r. absolument continue, alors  $L_X(t)$  existe si et seulement si l'application  $x \mapsto e^{tx} f(x)$  est intégrable sur  $X(\Omega)$ . On a alors :

$$L_X(t) = \int_{X(\Omega)} e^{tx} f(x) dx$$

Si  $X$  est bornée, alors  $L_X$  est définie sur  $\mathbb{R}$  tout entier.

### Propositions 7.2.c :

1°)  $L_X$  est toujours définie en 0, et  $L_X(0) = 1$ .

2°)  $\forall a, b \in \mathbb{R}, L_{aX+b}(t) = \mathbf{E} [ e^{t(aX+b)} ] = e^{tb} \mathbf{E} [ e^{atX} ] = e^{tb} L_X(at)$ .

**Exemples 7.2.d :**

1°) Loi uniforme discrète sur  $\left\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\right\}$ .

$$\text{Pour tout } t \in \mathbb{R}, L_X(t) = \mathbf{E}[e^{tX}] = \sum_{i=0}^{n-1} \frac{1}{n} e^{\frac{it}{n}} = \frac{1}{n} \sum_{i=0}^{n-1} \left(e^{\frac{t}{n}}\right)^i = \frac{e^t - 1}{n \left(e^{\frac{t}{n}} - 1\right)}.$$

2°) Loi uniforme continue sur  $[a; b]$ .

$$\text{Pour tout } t \in \mathbb{R}, L_X(t) = \frac{1}{b-a} \int_a^b e^{tx} dx = \frac{1}{b-a} \left[ \frac{1}{t} e^{tx} \right]_a^b = \frac{e^{bt} - e^{at}}{(b-a)t}.$$

3°) Loi exponentielle de paramètre  $\lambda$ .

$$\begin{aligned} \text{Pour tout } t \in \mathbb{R}, \text{ tel que } |t| < \lambda, L_X(t) &= \int_0^{+\infty} e^{tx} \lambda e^{-\lambda x} dx = \int_0^{+\infty} \lambda e^{(t-\lambda)x} dx \\ &= \frac{\lambda}{t-\lambda} \left[ e^{(t-\lambda)x} \right]_0^{+\infty} = \frac{\lambda}{\lambda-t}. \end{aligned}$$

4°) Loi normale centrée réduite.

$$\text{Pour tout } t \in \mathbb{R}, L_X(t) = \int_{-\infty}^{+\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2} + tx} dx.$$

$$\text{Or } -\frac{x^2}{2} + tx = -\frac{1}{2}(x^2 - 2tx) = -\frac{1}{2}[(x-t)^2 - t^2]$$

$$= -\frac{1}{2}(x-t)^2 + \frac{t^2}{2} = -\frac{u^2}{2} + \frac{t^2}{2} \text{ pour } u = x-t.$$

$$\text{Alors, } L_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2} + \frac{t^2}{2}} du = e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du = e^{\frac{t^2}{2}}.$$

5°) Loi normale de paramètres  $\mu$  et  $\sigma$ .

Si  $X \sim \mathcal{N}(\mu, \sigma)$ , il existe  $Y \sim \mathcal{N}(0, 1)$  telle que  $X = \sigma Y + \mu$ .

$$\text{Alors } L_X(t) = L_{\sigma Y + \mu}(t) = e^{\mu t} L_Y(\sigma t) = e^{\mu t} e^{-\frac{t^2 \sigma^2}{2}}.$$

**Proposition 7.2.e** (admise) :

Si tous les moments  $m_X(n) = \mathbf{E} [ X^n ]$  d'une v.a.r.  $X$  sont définis et si la série  $\sum_{n \geq 0} \frac{m_X(n)}{n!} u^n$  converge, avec un rayon de convergence  $R$  non nul, alors pour tout  $t \in \mathbb{R}$  tel que  $|t| < R$ ,  $L_X(t)$  existe et est égale à la somme de cette série :

$$L_X(t) = \sum_{n=0}^{+\infty} \frac{m_X(n)}{n!} u^n .$$

Conséquence :

Pour tout  $n \in \mathbb{N}$ , d'après la formule de Taylor, le coefficient de  $t^n$  dans la série entière  $L_X(t)$  est  $\frac{L_X^{(n)}(0)}{n!}$ , et donc  $L_X^{(n)}(0) = m_X(n) = \mathbf{E} [ X^n ]$ , d'où le nom de fonction génératrice des moments pour  $L_X$ .

**Proposition 7.2.f** (admise) :

Si  $X$  et  $Y$  ont même fonction génératrice des moments, alors elles sont identiquement distribuées.

**Proposition 7.2.g** :

Si  $X$  et  $Y$  sont deux v.a.r indépendantes admettant chacune une fonction génératrice des moments, alors la variable aléatoire  $X + Y$  admet aussi une fonction génératrice des moments, et on a  $L_{X+Y} = L_X \times L_Y$ .

Démonstration :

Pour tout  $t \in \mathbb{R}$ ,  $L_{X+Y}(t) = \mathbf{E} [ e^{t(X+Y)} ] = \mathbf{E} [ e^{tX} e^{tY} ]$ .

Or, comme  $X$  et  $Y$  sont indépendantes,  $e^{tX}$  et  $e^{tY}$  le sont également, donc :

$$\mathbf{E} [ e^{tX} e^{tY} ] = \mathbf{E} [ e^{tX} ] \mathbf{E} [ e^{tY} ] = L_X(t) \times L_Y(t).$$

**Corollaire 7.2.h** (admis) :

Si  $(X_k)_{k \in K}$  est une famille de v.a.r mutuellement indépendantes admettant chacune une fonction génératrice des moments  $L_{X_k}$ , alors la variable aléatoire  $Z = \sum_{k \in K} X_k$  admet aussi une

fonction génératrice des moments  $L_Z$ , et on a  $L_Z = \prod_{k \in K} L_{X_k}$ .

**Proposition 7.2.i :** *Loi de la somme de variables aléatoires mutuellement indépendantes*

Si l'on considère  $n$  v.a.r mutuellement indépendantes  $X_1, \dots, X_n$ , admettant chacune une fonction génératrice des moments, alors la fonction génératrice des moments de  $Z = \sum_{i=1}^n X_i$  se calcule facilement en appliquant la proposition 10.2.d. Si la fonction génératrice des moments de  $Z$  est déjà connue comme étant la fonction génératrice des moments d'une loi  $\mathcal{L}$ , alors en vertu de la proposition 10.2.c,  $Z \sim \mathcal{L}$ .

**Exemples 7.2.j :**

La somme de deux lois normales indépendantes est une loi normale.

Soient  $X_1 \sim \mathcal{N}(\mu_1, \sigma_1)$  et  $X_2 \sim \mathcal{N}(\mu_2, \sigma_2)$ , indépendantes.

On a  $L_{X_1}(t) = e^{\mu_1 t} e^{-\frac{t^2 \sigma_1^2}{2}}$  et  $L_{X_2}(t) = e^{\mu_2 t} e^{-\frac{t^2 \sigma_2^2}{2}}$ , donc :

$$L_{X_1+X_2}(t) = e^{\mu_1 t} e^{-\frac{t^2 \sigma_1^2}{2}} e^{\mu_2 t} e^{-\frac{t^2 \sigma_2^2}{2}} = e^{(\mu_1 + \mu_2)t} e^{-\frac{t^2 (\sigma_1^2 + \sigma_2^2)}{2}}.$$

On en déduit donc de la proposition 7.2.h que  $(X_1 + X_2) \sim \mathcal{N}\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right)$ .

**7.3. Fonctions caractéristiques****Définition 7.3.a :**

Soit  $X$  une v.a.r. On appelle *fonction caractéristique* de  $X$  (ou de la loi de  $X$ ) la fonction de la variable réelle  $t$  définie par :

$$\varphi_X(t) = \mathbf{E} [ e^{itX} ].$$

**Proposition 7.3.b :** *Existence et calcul de la fonction caractéristique*

1°) Si  $X$  est une variable aléatoire discrète telle que  $X(\Omega) = \{x_k; k \in K \subset \mathbb{N}\}$ ,  $\varphi_X(t)$  existe car la série  $\sum_{k \in K} \mathbf{P}(X = x_k) e^{itx_k}$  converge. En effet :

$$\sum_{k \in K} \left| \mathbf{P}(X = x_k) e^{itx_k} \right| = \sum_{k \in K} \mathbf{P}(X = x_k) \left| e^{itx_k} \right| = \sum_{k \in K} \mathbf{P}(X = x_k) = 1.$$

On a alors  $\varphi_X(t) = \sum_{k \in K} e^{itx_k} \times \mathbf{P}(X = x_k)$ .

2°) Si  $X$  est une variable aléatoire absolument continue, alors  $\varphi_X(t)$  existe car l'application  $x \mapsto e^{itx} f(x)$  est intégrable sur  $\Omega$  (si  $f = g + i h$ , alors  $\int f = \int g + i \int h$ ). En effet :

$$\int_{X(\Omega)} |e^{itx} f(x)| dx = \int_{X(\Omega)} |e^{itx}| |f(x)| dx = \int_{X(\Omega)} |f(x)| dx = 1.$$

On a alors  $\varphi_X(t) = \int_{X(\Omega)} e^{itx} f(x) dx$ .

Conclusion :

Pour toute v.a.r  $X$ , la fonction caractéristique est définie sur  $\mathbb{R}$  tout entier.

**Proposition 7.3.c :**

1°)  $\varphi_X(0) = 1$ .

2°)  $\forall t \in \mathbb{R}, |\varphi_X(t)| \leq \sum_{k \in K} \mathbf{P}(X = x_k) |e^{itx_k}| = \sum_{k \in K} \mathbf{P}(X = x_k) = 1$  (cas discret),

ou  $|\varphi_X(t)| \leq \int_{X(\Omega)} |e^{itx}| |f(x)| dx = \int_{X(\Omega)} |f(x)| dx = 1$  (cas continu),

3°)  $\forall t \in \mathbb{R}, \varphi_X(-t) = \mathbf{E}[e^{i(-t)X}] = \mathbf{E}[\overline{e^{itX}}] = \overline{\mathbf{E}[e^{itX}]} = \overline{\varphi_X(t)}$ . Donc si une fonction caractéristique est réelle, alors elle est paire.

4°)  $\forall a, b \in \mathbb{R}, \forall t \in \mathbb{R}, \varphi_{aX+b}(t) = \mathbf{E}[e^{it(aX+b)}] = e^{itb} \mathbf{E}[e^{i(ta)X}] = e^{itb} \varphi_X(ta)$  ;

5°) Si la loi de  $X$  est symétrique (i. e. si  $X$  et  $-X$  ont la même loi), alors  $\varphi_X$  est une fonction réelle, donc paire :  $\varphi_X(t) = \mathbf{E}[e^{itX}] = \mathbf{E}[e^{it(-X)}] = \mathbf{E}[e^{i(-t)X}] = \varphi_X(-t) = \overline{\varphi_X(t)}$ .

**Proposition 7.3.d** (admise) : *Unicité de l'inversion de la fonction caractéristique*

La fonction caractéristique des moments d'une variable aléatoire détermine la loi de cette variable.

**Proposition 7.3.e :**

Si  $X$  et  $Y$  sont deux v.a.r indépendantes admettant chacune une fonction caractéristique, alors la variable aléatoire  $X + Y$  admet aussi une fonction caractéristique, et on a  $\varphi_{X+Y} = \varphi_X \times \varphi_Y$ .

Démonstration :

$$\varphi_Z(t) = \mathbf{E}[e^{itZ}] = \mathbf{E}[e^{it(X+Y)}] = \mathbf{E}[e^{itX} e^{itY}].$$

Or, comme  $X$  et  $Y$  sont indépendantes, les variables  $e^{itX}$  et  $e^{itY}$  le sont aussi.

$$\text{Donc : } \varphi_Z(t) = \mathbf{E}[e^{itX} e^{itY}] = \mathbf{E}[e^{itX}] \mathbf{E}[e^{itY}] = \varphi_X(t) \times \varphi_Y(t).$$



**Corollaire 7.3.f** (admis) :

Si  $(X_k)_{k \in K}$  est une famille de v.a.r mutuellement indépendantes admettant chacune une fonction caractéristique  $\varphi_{X_k}$ , alors la variable aléatoire  $Z = \sum_{k \in K} X_k$  admet aussi une fonction caractéristique  $\varphi_Z$ , et on a :

$$\varphi_Z = \prod_{k \in K} \varphi_{X_k}.$$

**Proposition 7.3.g** : *Loi de la somme de deux variables aléatoires indépendantes*

Si  $X$  et  $Y$  sont deux v.a.r indépendantes, alors la fonction caractéristique de  $Z = X + Y$  se calcule facilement en appliquant la proposition 10.3.c. Si cette fonction caractéristique est déjà connue comme étant la fonction caractéristique d'une loi  $\mathcal{L}$ , alors en vertu de la proposition 10.3.b,  $Z \sim \mathcal{L}$ .

**Proposition 7.3.h** :

Soit  $X$  une variable aléatoire de fonction caractéristique  $\varphi$ . Si  $\mathbf{E}[|X|]$  existe (est bornée), alors  $\varphi$  est continûment dérivable et  $\varphi'(0) = i \mathbf{E}[X]$ .

Démonstration (pour le cas continu) :

On a  $\varphi(t) = \mathbf{E}[e^{itX}] = \int_{\mathbb{R}} e^{itx} f(x) dx$ .

Soit  $F : \begin{cases} \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C} \\ (t, x) \mapsto e^{itx} f(x) \end{cases}$  :

- $F$  est continue par morceaux sur  $\mathbb{R} \times \mathbb{R}$ ;
- $F$  vérifie l'hypothèse de domination sur  $\mathbb{R} \times \mathbb{R}$  :

$\forall (t, x) \in \mathbb{R} \times \mathbb{R}, |F(t, x)| = f(x)$  et  $f$  est positive, continue par morceaux et intégrable sur  $\mathbb{R}$ ;

- $\frac{\partial F}{\partial t} : (t, x) \mapsto ix e^{itx} f(x)$  existe et est continue sur  $\mathbb{R} \times \mathbb{R}$ ;
- $\frac{\partial F}{\partial t}$  vérifie l'hypothèse de domination sur  $\mathbb{R} \times \mathbb{R}$  :

$\forall (t, x) \in \mathbb{R} \times \mathbb{R}, \left| \frac{\partial F}{\partial t}(t, x) \right| = |x| f(x)$  et l'application  $x \mapsto |x| f(x)$  est positive, continue et intégrable sur  $\mathbb{R}$  car  $\mathbf{E}[|X|]$  existe.

Donc, en vertu du théorème de dérivation sous le signe somme pour une intégrale dépendant d'un paramètre :

- $\forall t \in \mathbb{R}, F(t, \cdot)$  et  $\frac{\partial F}{\partial t}(t, \cdot)$  sont intégrables sur  $\mathbb{R}$ ;
- $\varphi$  est continûment dérivable sur  $\mathbb{R}$ , et  $\forall t \in \mathbb{R}, \varphi'(t) = \int_{\mathbb{R}} ix e^{itx} f(x) dx$ .

Donc, finalement  $\varphi'(0) = i \mathbf{E}[X]$ .

**Proposition 7.3.i :**

Soit  $X$  une variable aléatoire de fonction caractéristique  $\varphi$ . S'il existe un entier  $n$  tel que  $\mathbf{E}[|X|^n]$  existe (est bornée), alors :

- $\varphi$  est continûment dérivable jusqu'à l'ordre  $n$  inclus ;
- pour tout  $k = 0, 1, \dots, n$  on a :  $\varphi^{(k)}(0) = i^k \mathbf{E}[X^k]$  ;
- $\varphi$  peut être représenté par la formule de Taylor :

$$\varphi(t) = \sum_{k=0}^n \frac{(it)^k}{k!} \mathbf{E}[X^k] + o(|t|^n), \text{ lorsque } t \in V(0).$$

Démonstration :

En appliquant, comme pour le théorème précédent, le théorème de dérivation sous le signe somme pour une intégrale dépendant d'un paramètre aux dérivées successives de  $\varphi$  (les dérivées partielles existent, sont continues et intégrables car  $\forall k \leq n, \mathbf{E}[|X|^k]$  existe), on déduit que  $\varphi$  est continûment dérivable jusqu'à l'ordre  $n$  inclus et que :

$$\varphi^{(k)}(t) = \int_{\mathbb{R}} \frac{\partial^k}{\partial t^k} e^{itx} f(x) dx = \int_{\mathbb{R}} i^k x^k e^{itx} f(x) dx.$$

Donc  $\varphi^{(k)}(0) = \int_{\mathbb{R}} i^k x^k f(x) dx = i^k \mathbf{E}[X^k]$ , et au voisinage de 0 :

$$\varphi(t) = \sum_{k=0}^n \frac{t^k}{k!} \varphi^{(k)}(0) + o(|t|^n) = \sum_{k=0}^n \frac{(it)^k}{k!} \mathbf{E}[X^k] + o(|t|^n).$$

## 7.4. Fonctions caractéristiques des lois usuelles

### 7.4.1 Loi discrètes

**Proposition 7.4.1.a :** *Fonction caractéristique d'une loi discrète uniforme sur  $\llbracket 1, N \rrbracket$*

Si  $X \sim \mathcal{U}_{\llbracket 1, N \rrbracket}$ , alors sa fonction caractéristique est définie par  $\varphi(t) = \frac{e^{it} - e^{itN}}{1 - e^{it}}$

Démonstration :

$$\varphi(t) = \mathbf{E}[e^{itX}] = \frac{1}{N} \sum_{k=1}^N e^{itk} = \frac{e^{it}}{N} \sum_{k=0}^{N-1} e^{itk} = \frac{e^{it} - e^{itN}}{1 - e^{it}}$$

**Proposition 7.4.1.b :** *Fonction caractéristique d'une loi de Bernoulli*

Si  $X \sim \mathcal{B}(p)$  alors sa fonction caractéristique est définie par  $\varphi(t) = \mathbf{E}[e^{itX}] = pe^{it} + q$ .

**Proposition 7.4.1.c :** *Fonction caractéristique d'une loi binomiale*

Si  $X \sim \mathcal{B}(n, p)$  alors sa fonction caractéristique est définie par  $\varphi(t) = (pe^{it} + q)^n$ .

Démonstration :

$$\varphi(t) = \mathbf{E}[e^{itX}] = \sum_{k=0}^n e^{itk} C_n^k p^k (1-p)^{n-k} = \sum_{k=0}^n C_n^k (pe^{it})^k (1-p)^{n-k} = (pe^{it} + q)^n.$$

Remarque :

La fonction caractéristique d'une loi de Bernoulli de paramètre  $p$  est définie par  $\varphi_B(t) = pe^{it} + q$ . Comme  $X$  est la somme de  $n$  lois de Bernoulli indépendantes de paramètre  $p$ , on a donc  $\varphi(t) = (pe^{it} + q)^n$ .

**Proposition 7.4.1.d :** *Fonction caractéristique d'une loi de Poisson*

Si  $X \sim \mathcal{P}(\lambda)$  alors sa fonction caractéristique est définie par  $\varphi(t) = e^{\lambda(e^{it} - 1)}$ .

Démonstration :

$$\varphi(t) = \mathbf{E}[e^{itX}] = \sum_{k=0}^{+\infty} e^{itk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it} - 1)}.$$

### 7.4.2 Loi continues classiques

**Proposition 7.4.2.a** : *Fonction caractéristique d'une loi continue uniforme sur  $[a; b]$*

Si  $X \sim \mathcal{U}_{[a; b]}$ , alors sa fonction caractéristique est définie par  $\varphi(t) = \frac{1}{b-a} e^{i\left(\frac{a+b}{2}\right)t} 2 \frac{\sin\left(\frac{b-a}{2}t\right)}{t}$ .

Si  $X \sim \mathcal{U}_{[-\alpha; \alpha]}$ , alors sa fonction caractéristique est définie par  $\varphi(t) = \frac{\sin \alpha t}{\alpha t}$ .

Démonstration :

$$\begin{aligned} \varphi(t) &= \mathbf{E} [e^{itX}] = \frac{1}{b-a} \int_a^b e^{itx} dx = \frac{1}{b-a} \left[ \int_a^b \cos tx dx + i \int_a^b \sin tx dx \right] \\ &= \frac{1}{b-a} \left( \left[ \frac{\sin tx}{t} \right]_a^b + i \left[ -\frac{\cos tx}{t} \right]_a^b \right) \\ &= \frac{1}{b-a} \frac{1}{t} ( \sin bt - i \cos bt - \sin at + i \cos at ) \\ &= \frac{1}{b-a} \frac{1}{it} ( ( \cos bt + i \sin bt ) - ( \cos at + i \sin at ) ) \\ &= \frac{1}{b-a} \frac{e^{ibt} - e^{iat}}{it} = \frac{1}{b-a} \frac{e^{i\left(\frac{a+b}{2} + \frac{b-a}{2}\right)t} - e^{i\left(\frac{a+b}{2} - \frac{b-a}{2}\right)t}}{it} = \frac{1}{b-a} e^{i\left(\frac{a+b}{2}\right)t} 2 \frac{\sin\left(\frac{b-a}{2}t\right)}{t}. \end{aligned}$$

Dans le cas où  $[a; b] = [-\alpha; \alpha]$ ,  $\varphi$  est réelle et  $\varphi(t) = \frac{\sin \alpha t}{\alpha t}$ .

**Proposition 7.4.2.b** (admise) : *Fonction caractéristique d'une loi exponentielle*

Si  $X \sim \mathcal{E}(\lambda)$ , alors sa fonction caractéristique est définie par  $\varphi(t) = \frac{\lambda}{\lambda - it}$ .

**Proposition 7.4.2.c** (admise) : *Fonction caractéristique d'une loi normale centrée réduite*

Si  $X \sim \mathcal{N}(0, 1)$ , alors sa fonction caractéristique est définie par  $\varphi(t) = e^{-t^2/2}$ .

**Proposition 7.4.2.d** (admise) : *Fonction caractéristique d'une loi normale*

Si  $X \sim \mathcal{N}(\mu, \sigma^2)$ , alors sa fonction caractéristique est définie par  $\varphi(t) = e^{i\mu t} e^{-\frac{\sigma^2 t^2}{2}}$ .

Démonstration :

Comme il existe  $Y \in \mathcal{N}(0, 1)$  telle que  $X = \sigma Y + \mu$ , on a :

$$\varphi_X(t) = \varphi_{\sigma Y + \mu}(u) = e^{i\mu t} \varphi_Y(\sigma t) = e^{i\mu t} e^{-\frac{\sigma^2 t^2}{2}}.$$

## 7.5. Fonctions caractéristiques de vecteurs aléatoires

### 7.5.1 Définitions et propriétés

**Définition 7.4.2.a :** *Fonction caractéristique d'un vecteur aléatoire*

Soit  $V = (X_1, \dots, X_m)$  un vecteur aléatoire. Pour tout vecteur  $T = (t_1, \dots, t_m) \in \mathbb{R}^m$ , la fonction caractéristique,  $\Phi_V$ , de  $V$  est définie par :

$$\Phi_V(T) = \mathbf{E} \left[ e^{i(t_1 X_1 + \dots + t_m X_m)} \right].$$

**Proposition 7.4.2.b :** *Fonction caractéristique d'une variable marginale*

Etant donné un vecteur aléatoire  $V = (X_1, \dots, X_m)$ , la fonction caractéristique de la variable marginale  $X_k$ ,  $\varphi_{X_k}$ , se déduit de la fonction caractéristique de  $V$  de la façon suivante :

$$\forall t \in \mathbb{R}, \varphi_{X_k}(t) = \Phi_V(T_k), \text{ avec } T_k = (t_1, \dots, t_m) \text{ tel que } t_i = 1 \text{ si } i = k \text{ et } 0 \text{ sinon.}$$

**Théorème 7.4.2.c :**

Etant donné un vecteur aléatoire  $V = (X_1, \dots, X_m)$ , ses variables marginales sont indépendantes

si et seulement si  $\forall T = (t_1, \dots, t_m) \in \mathbb{R}^m, \Phi_V(T) = \prod_{k=1}^m \varphi_{X_k}(t_k)$ .

Démonstration (pour  $m = 2$ ) :

Si  $X_1$  et  $X_2$  sont indépendantes, pour tout  $t \in \mathbb{R}$ ,  $e^{itX_1}$  et  $e^{itX_2}$  le sont aussi, et donc :

$$\Phi_{(X_1, X_2)}(t_1, t_2) = \mathbf{E} \left[ e^{i(t_1 X_1 + t_2 X_2)} \right] = \mathbf{E} \left[ e^{i t_1 X_1} \right] \times \mathbf{E} \left[ e^{i t_2 X_2} \right] = \varphi_{X_1}(t_1) \times \varphi_{X_2}(t_2).$$

Réciproquement si  $\Phi_{(X_1, X_2)}(t_1, t_2) = \varphi_{X_1}(t_1) \times \varphi_{X_2}(t_2)$ , alors, en raison de l'unicité de l'inversion de la fonction caractéristique (proposition 7.3.d), les variables aléatoires  $X_1$  et  $X_2$  sont indépendantes.

## 7.5.2 Applications aux vecteurs gaussiens

### Rappel 7.5.2.a : Vecteur gaussien

Etant donné un vecteur aléatoire  $U = (U_1, \dots, U_m)$  dont les composantes sont des variables aléatoires normales centrées réduites indépendantes, on appelle *vecteur gaussien* de dimension  $m$  toute transformation linéaire du vecteur  $U$  :

$$V = AU + \mu, \text{ avec } A \in M_m(\mathbb{R}) \text{ et } \mu \in \mathbb{R}^m.$$

### Théorème 7.5.2.b : Théorème de Cramer-Wold

La loi d'un vecteur aléatoire est entièrement déterminée par celles de toutes les combinaisons linéaires de ces composantes.

#### Démonstration :

Soit  $V = (X_1, \dots, X_m)$  un vecteur aléatoire. En vertu de l'unicité de l'inversion de la fonction caractéristique, la loi de  $V$  est entièrement déterminée si  $\Phi_V$  est connue, c'est-à-dire si, pour tout  $T \in \mathbb{R}^m$ , on peut déterminer  $\Phi_V(T) = \mathbf{E} \left[ e^{iT^TV} \right]$ .

Or,  $T^TV$  est une combinaison linéaire des composantes de  $V$ , donc, d'après l'hypothèse, c'est une variable aléatoire connue. Posons donc  $Y = T^TV$ .

Comme  $Y$  est connue,  $\phi_Y(t) = \mathbf{E} \left[ e^{itY} \right] = \mathbf{E} \left[ e^{iT^TV} \right]$  peut être déterminé pour tout  $t$ , donc en particulier pour  $t = 1$ . On a alors  $\phi_Y(1) = \mathbf{E} \left[ e^{iT^TV} \right] = \Phi_V(T)$ .

### Proposition 7.5.2.d (admise) :

Un vecteur aléatoire est gaussien si et seulement si toute combinaison linéaire de ces composantes est une variable aléatoire gaussienne.

#### Remarques :

- 1°) En vertu du théorème de Cramer-Wold, cette définition détermine bien une loi de probabilité.
- 2°) Si un vecteur est gaussien, ces composantes sont gaussiennes : il suffit de prendre une combinaison linéaire où tous les coefficients sont nuls sauf un.
- 3°) Si les composantes d'un vecteur sont gaussiennes et indépendantes, alors le vecteur est gaussien.

**Proposition 7.5.2.d :** *Fonction caractéristique d'un vecteur gaussien*

La fonction caractéristique d'un vecteur gaussien  $V \sim \mathcal{N}_m(\mu, \Sigma)$  est définie par :

$$\forall T \in \mathbb{R}^m, \Phi_V(T) = e^{iT^T \mu} e^{-\frac{1}{2} T^T \Sigma T}.$$

Démonstration :

Soit  $W = V - \mu$ , centré, et  $T \in \mathbb{R}^m$ .

La variable aléatoire  $X = T^T W$  étant une combinaison linéaire des composantes d'un vecteur gaussien est donc une variable gaussienne, d'après la proposition 7.5.2.d.

On a de plus,  $\mathbf{E}[X] = 0$  car  $W$  est centré, et  $\mathbf{V}[X] = \sigma^2 = T^T \Sigma_W T$ .

On a alors, pour tout  $t \in \mathbb{R}$ ,  $\varphi_X(t) = e^{-\frac{(\sigma t)^2}{2}} = e^{-\frac{t^2}{2} T^T \Sigma_W T}$ .

Toute combinaison linéaire du vecteur  $W$  étant déterminée, d'après le théorème de Cramer-Wold, la loi de  $W$  est donc, elle aussi, entièrement déterminée par :

$$\Phi_W(T) = \varphi_X(1) = e^{-\frac{1}{2} T^T \Sigma_W T}.$$

Comme  $V = W + \mu$ ,  $\Phi_V(T) = \mathbf{E}[e^{iT^T V}] = \mathbf{E}[e^{iT^T (W + \mu)}] = e^{iT^T \mu} \mathbf{E}[e^{iT^T W}] = e^{iT^T \mu} \Phi_W(T)$ .

**Théorème 7.5.2.e :**

Les composantes d'un vecteur gaussien sont indépendantes si et seulement si elles sont non corrélées.

Démonstration :

Soit  $V = (V_1, \dots, V_m) \sim \mathcal{N}_m(\mu, \Sigma)$  un vecteur gaussien.

Posons, pour tout  $k \in \llbracket 1, m \rrbracket$ ,  $\mathbf{E}[V_k] = \mu_k$  et  $\mathbf{V}[V_k] = \sigma_k^2$ .

Si les composantes de  $V$  sont indépendantes, alors elles sont non corrélées.

Si les composantes de  $V$  sont non corrélées, alors la matrice  $\Sigma$  est diagonale.

On a donc  $\Sigma = \text{Diag}(\sigma_k^2)$ . Soit  $T = (t_1, \dots, t_m)$  un vecteur colonne de  $\mathbb{R}^m$ .

$$\begin{aligned} \text{Alors } \Phi_V(T) &= \mathbf{E}[e^{iT^T V}] = e^{iT^T \mu} e^{-\frac{1}{2} T^T \Sigma T} = e^{i \sum_{k=1}^m t_k \mu_k} e^{-\frac{1}{2} \sum_{k=1}^m \sigma_k^2 t_k^2} = \prod_{k=1}^m e^{i t_k \mu_k} \prod_{k=1}^m e^{-\frac{(\sigma_k t_k)^2}{2}} \\ &= \prod_{k=1}^m \left( e^{i t_k \mu_k} e^{-\frac{(\sigma_k t_k)^2}{2}} \right) = \prod_{k=1}^m \varphi_{V_k}(t_k). \end{aligned}$$

Donc, d'après le théorème 7.4.2.c les composantes de  $V$  sont indépendantes.





## 8. Suites de variables aléatoires

### 8.1. Convergences des suites de variables aléatoires

Soit  $(X_n)_{n \in \mathbb{N}}$ , une suite de v.a.r.. On définit les quatre types de convergence suivants :

**Définition 8.1.a :** Convergence *presque sûre* ou convergence *forte*

Une suite de v.a.r.  $(X_n)_{n \in \mathbb{N}}$  converge *presque sûrement* vers une v.a.r.  $X$  si et seulement si :

$$\mathbf{P} ( \{ \omega \in \Omega \text{ tels que } \lim_{n \rightarrow +\infty} X_n(\omega) \neq X(\omega) \} ) = 0 .$$

**Définition 8.1.b :** Convergence *en moyenne d'ordre  $p$*

Une suite de v.a.r.  $(X_n)_{n \in \mathbb{N}}$  converge *en moyenne d'ordre  $p$*  vers une v.a.r.  $X$  si et seulement si :

$$\lim_{n \rightarrow +\infty} \mathbf{E} [ |X_n - X|^p ] = 0 .$$

Remarque :

Pour  $p = 2$ , on parle de convergence en moyenne *quadratique*.

**Définition 8.1.c :** Convergence *en probabilité*

Une suite de v.a.r.  $(X_n)_{n \in \mathbb{N}}$  converge *en probabilités* vers une v.a.r.  $X$  si et seulement si :

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbf{P} ( |X_n - X| > \varepsilon ) = 0 .$$

**Définition 8.1.d :** Convergence *en loi*

Une suite de v.a.r.  $(X_n)_{n \in \mathbb{N}}$  converge *en loi* vers une v.a.r.  $X$  si et seulement si la suite des fonctions de répartition des variables aléatoires  $X_n$ ,  $(F_n)_{n \in \mathbb{N}}$  converge vers la fonction de répartition  $F$  de la loi  $X$ , en tout point de continuité de  $F$ .

Remarques :

1°) Si une suite de v.a.r. absolument continues  $(X_n)_{n \in \mathbb{N}}$  converge *en loi* vers une v.a.r.  $X$ , alors la suite des fonctions de densité des variables aléatoires  $X_n$ ,  $(f_n)_{n \in \mathbb{N}}$  ne converge pas forcément vers la fonction de densité  $f$  de la loi  $X$ .

2°) Comme une suite de fonctions en escalier peut converger vers une loi continue, une suite de v.a.r. discrètes peut converger en loi vers une v.a.r. absolument continue.

**Théorème 8.1.e** (admis) : *Théorème de Paul Lévy*

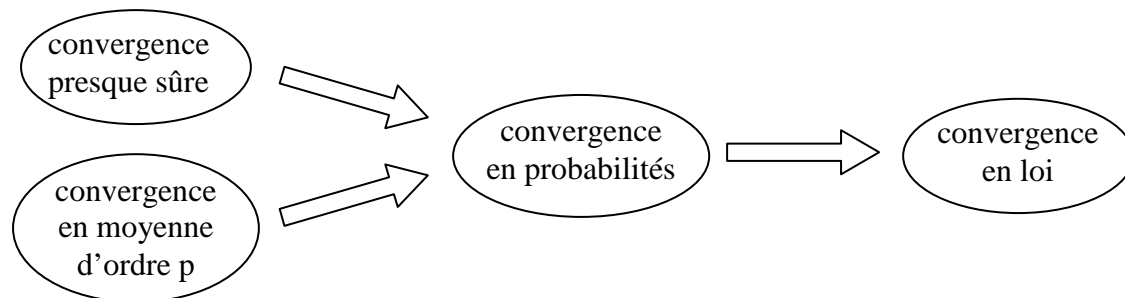
1°) Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires qui converge en loi vers une variable  $X$ . Alors la suite  $(\varphi_n)_{n \in \mathbb{N}}$  des fonctions caractéristiques des variables  $X_n$  converge vers la fonction caractéristique  $\varphi$  de  $X$ , et ceci uniformément dans tout intervalle fini.

2°) Soient  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires et  $(\varphi_n)_{n \in \mathbb{N}}$  la suite des fonctions caractéristiques des variables  $X_n$ . Supposons que la suite  $(\varphi_n)_{n \in \mathbb{N}}$  converge, au sens de la convergence simple, vers une fonction  $\varphi$ , dont la partie réelle est continue à l'origine. Alors :

- a)  $\varphi$  est une fonction caractéristique, c'est-à-dire qu'il existe une variable aléatoire  $X$  dont  $\varphi$  est la fonction caractéristique ;
- b) la suite  $(X_n)_{n \in \mathbb{N}}$  converge en loi vers  $X$ .

**Théorème 8.1.f** (admis) :

Pour une suite de v.a.r. la convergence presque sûre ou la convergence en moyenne d'ordre  $p$  impliquent la convergence en probabilités, qui implique elle-même la convergence en loi, les réciproques de ces implications étant fausses.



## 8.2. Lois des grands nombres

Soit une suite de v.a.r. identiquement distribuées  $(X_n)_{n \in \mathbb{N}^*}$ . Puisque les  $X_n$  sont identiquement distribuées on peut poser :  $\forall n \in \mathbb{N}^*, \mathbf{E}[X_n] = \mu, \mathbf{V}[X_n] = \sigma^2$ .

Soit la suite de v.a.r.  $(\bar{X}_n)_{n \in \mathbb{N}}$ , définie par :  $\forall n \in \mathbb{N}^*, \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

### **Théorème 8.2.a :** *Loi faible des grands nombres*

Si les  $X_n$  sont deux à deux non corrélées, alors la suite de v.a.r.  $(\bar{X}_n)_{n \in \mathbb{N}}$  converge en probabilités vers  $\mu$ .

#### Démonstration :

Comme les  $X_n$  sont deux à deux non corrélées, on a :

$$\begin{aligned} \forall i \neq j, \mathbf{E}[(X_i - \mu)(X_j - \mu)] &= \mathbf{E}[X_i X_j] - \mu \mathbf{E}[X_i] - \mu \mathbf{E}[X_j] + \mu^2 \\ &= \mathbf{E}[X_i X_j] - \mu^2 = \text{cov}(X_i, X_j) = 0. \end{aligned}$$

Posons :  $\forall n \in \mathbb{N}^*, Y_n = X_n - \mu$ .

Alors  $\forall i \neq j, \mathbf{E}[Y_i Y_j] = 0$ , et  $\forall i \in \mathbb{N}^*, \mathbf{E}[Y_i] = 0$  et  $\mathbf{V}[Y_i] = \mathbf{E}[Y_i^2] = \mathbf{V}[X_i] = \sigma^2$ .

En posant  $\forall n \in \mathbb{N}^*, \bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$ , on a  $\bar{Y}_n = \bar{X}_n - \mu$ .

$$\begin{aligned} \text{Alors : } \mathbf{E}[\bar{Y}_n^2] &= \frac{1}{n^2} \mathbf{E}\left[\left(\sum_{i=1}^n Y_i\right)^2\right] = \frac{1}{n^2} \mathbf{E}\left[\sum_{i=1}^n Y_i^2 + 2 \sum_{i>j} Y_i Y_j\right] \\ &= \frac{1}{n^2} \left(\sum_{i=1}^n \mathbf{E}[Y_i^2] + 2 \sum_{i>j} \mathbf{E}[Y_i Y_j]\right) = \frac{\sigma^2}{n}. \end{aligned}$$

Donc  $\lim_{n \rightarrow +\infty} \mathbf{E}[\bar{Y}_n^2] = \lim_{n \rightarrow +\infty} \mathbf{E}\left[\left|\bar{X}_n - \mu\right|^2\right] = 0$ , donc  $(\bar{X}_n)_{n \in \mathbb{N}}$  converge en moyenne quadratique vers  $\mu$ .

#### Remarque :

La convergence d'une suite de v.a.r. vers une constante réelle rentre dans le cadre des définitions données précédemment : il s'agit du cas particulier, mais fréquent, où la limite est une variable aléatoire certaine.

### **Théorème 8.2.b** (admis) : *Loi forte des grands nombres*

Si les  $X_n$  sont indépendantes deux à deux, alors la suite de v.a.r.  $(\bar{X}_n)_{n \in \mathbb{N}}$  converge presque sûrement vers  $\mu$ .

### 8.3. Théorème de la limite centrée

#### Proposition 8.3.a :

Soit une suite de v.a.r.  $(X_n)_{n \in \mathbb{N}^*}$ , indépendantes et identiquement distribuées, d'espérance commune  $\mu$ , et de variance commune  $\sigma^2$ .

Du fait de la linéarité de l'espérance et de l'indépendance des variables  $X_n$ , on a :

$$\mathbf{E} \left[ \sum_{i=1}^n X_i \right] = \sum_{i=1}^n \mathbf{E} [X_i] = n \mu ;$$

$$\mathbf{V} \left[ \sum_{i=1}^n X_i \right] = \sum_{i=1}^n \mathbf{V} [X_i] = n \sigma^2.$$

#### Théorème 8.3.b : Théorème de la limite centrée

Etant donnée une suite de v.a.r.  $(X_n)_{n \in \mathbb{N}^*}$ , indépendantes identiquement distribuées, d'espérance commune  $\mu$ , et de variance commune  $\sigma^2$ , la suite de variables aléatoires

$$\left( \frac{\sum_{i=1}^n X_i - n \mu}{\sigma \sqrt{n}} \right)_{n \in \mathbb{N}^*} \text{ converge en loi vers une variable aléatoire gaussienne centrée réduite.}$$

#### Démonstration :

On a  $\mathbf{E} \left[ \sum_{i=1}^n X_i \right] = n \mu$  et  $\mathbf{V} \left[ \sum_{i=1}^n X_i \right] = n \sigma^2$ , donc  $\frac{\sum_{i=1}^n X_i - n \mu}{\sigma \sqrt{n}} = \sum_{i=1}^n \frac{X_i - \mu}{\sigma \sqrt{n}}$  est une variable aléatoire centrée réduite.

Si les  $X_i$  sont identiquement distribuées, alors les  $\frac{X_i - \mu}{\sigma \sqrt{n}}$  sont donc identiquement distribuées,

et ont donc toutes une même fonction caractéristique notée  $\varphi_{\frac{X - \mu}{\sigma \sqrt{n}}}$ .

Si, pour tout  $k \in \mathbb{N}$ ,  $\mathbf{E} \left[ \left| \frac{X - \mu}{\sigma \sqrt{n}} \right|^k \right]$  existe, on a :

$$\varphi_{\frac{\sum_{i=1}^n X_i - n \mu}{\sigma \sqrt{n}}}(t) = \varphi_{\sum_{i=1}^n \frac{X_i - \mu}{\sigma \sqrt{n}}}(t) = \left[ \varphi_{\frac{X - \mu}{\sigma \sqrt{n}}}(t) \right]^n = \left[ \sum_{k=0}^{+\infty} \frac{t^k}{k!} \varphi_{\frac{X - \mu}{\sigma \sqrt{n}}}^{(k)}(0) \right]^n.$$

Or pour toutes variable aléatoire  $X$  pour laquelle le  $k^{\text{ième}}$  moment est défini, on a :

$$\varphi_X^{(k)}(0) = i^k \mathbf{E} [X^k].$$

Comme  $\mathbf{E}\left[\frac{X-\mu}{\sigma\sqrt{n}}\right] = 0$ , et  $\mathbf{E}\left[\left(\frac{X-\mu}{\sigma\sqrt{n}}\right)^2\right] = \mathbf{V}\left[\frac{X-\mu}{\sigma\sqrt{n}}\right] = \frac{1}{n}$ ,

on obtient  $\varphi_{\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}}(t) = \left[1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right]^n = e^{n \ln\left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)} = e^{n\left(-\frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)} = e^{-\frac{t^2}{2} + o(1)}$ .

Donc  $\lim_{n \rightarrow +\infty} \varphi_{\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}}(t) = e^{-\frac{t^2}{2}}$ .

Le théorème de Paul Lévy permet donc de conclure que la suite  $\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}\right)_{n \in \mathbb{N}^*}$  converge

en loi vers une loi normale centrée réduite.



## 9. Statistiques sur un échantillon aléatoire

### 9.1. Définitions

#### Définition 9.1.a : Echantillon aléatoire

On appelle *échantillon aléatoire* d'effectif  $n \in \mathbb{N}^*$ , tout vecteur aléatoire  $(X_1, \dots, X_n)$  dont les composantes  $X_i, i \in \llbracket 1, n \rrbracket$ , sont indépendantes entre elles et identiquement distribuées.

#### Définition 9.1.b : Statistique

Etant donné un échantillon aléatoire  $(X_1, \dots, X_n)$ , on appelle *statistique* sur cet échantillon toute variable aléatoire  $\varphi (X_1, \dots, X_n ; n)$ , fonction de  $X_1, \dots, X_n$ , et de  $n$ .

#### Définition 9.1.c : Moyenne arithmétique aléatoire

Etant donné un échantillon aléatoire  $(X_1, \dots, X_n)$ , on appelle *moyenne arithmétique aléatoire* sur cet échantillon la statistique  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ .

#### Définition 9.1.d : Variance estimée aléatoire

Etant donné un échantillon aléatoire  $(X_1, \dots, X_n)$ , on appelle *variance estimée aléatoire* sur cet échantillon la statistique  $S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \left( \frac{1}{n} \sum_{i=1}^n X_i^2 \right) - (\bar{X})^2$ .

#### Remarque :

Si l'espérance de la loi commune à toutes les variables aléatoires  $X_i, i \in \llbracket 1, n \rrbracket$ , que l'on peut noter  $\mathbf{E} [ X ]$ , est connue, la variance estimée aléatoire peut également être représentée par la

statistique  $T_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mathbf{E} [ X ])^2 = \left( \frac{1}{n} \sum_{i=1}^n X_i^2 \right) - \mathbf{E} [ X ]^2$ .

## 9.2. Distribution de la moyenne aléatoire d'un échantillon

**Proposition 9.2.a :** *Espérance et variance de la moyenne aléatoire d'un échantillon*

Soit un échantillon aléatoire  $(X_1, \dots, X_n)$ . Les variables  $X_n$  étant identiquement distribuées, on note respectivement  $\mu$  et  $\sigma^2$  leur espérance et leur variance commune.

Soit  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  la moyenne aléatoire de cet échantillon. L'espérance et la variance de  $\bar{X}_n$  sont données par :

$$\begin{aligned} \mathbf{E}[\bar{X}_n] &= \mathbf{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n \mathbf{E}[X_i] = \mu ; \\ \mathbf{V}[\bar{X}_n] &= \mathbf{V}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n^2} \sum_{i=1}^n \mathbf{V}[X_i] = \frac{\sigma^2}{n} . \end{aligned}$$

**Proposition 9.2.b :** *Distribution asymptotique de la moyenne aléatoire d'un échantillon*

La suite de variables aléatoires  $(\bar{X}_n)_{n \in \mathbb{N}^*}$  converge en loi vers une variable aléatoire gaussienne d'espérance  $\mu$  et de variance  $\frac{\sigma^2}{n}$ .

Grâce à cette propriété de convergence, on peut donc considérer que pour un échantillon suffisamment grand (en pratique, à partir d'une trentaine d'individus),  $\bar{X}_n$  est approximativement distribuée suivant une variable aléatoire gaussienne d'espérance  $\mu$  et de variance  $\frac{\sigma^2}{n}$ , ce que l'on peut noter :  $\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ .

Démonstration :

La distribution asymptotique de  $\bar{X}_n$  se déduit immédiatement du théorème de la limite

centrale :  $\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$ , donc la suite de variables aléatoires  $\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right)_{n \in \mathbb{N}^*}$  converge

en loi vers une variable aléatoire gaussienne centrée réduite, d'où le résultat.

**Proposition 9.2.c :** *Application à la distribution d'une proportion aléatoire*

On considère, dans une population donnée, un caractère binaire, valant 0 ou 1, tel que la proportion d'individus dans la population ayant une valeur de leur caractère égale à 1 est



$p \in [0; 1]$ . Ce caractère peut donc être représenté par une variable statistique distribuée suivant une loi de Bernoulli de paramètre  $p$ .

Etant donné un échantillon d'effectif  $n$ , le caractère de l'individu numéro  $i$ ,  $i \in \llbracket 1, n \rrbracket$ , est représenté par  $X_i \in \mathcal{B}(1, p)$ , et le nombre d'individus dans l'échantillon ayant une valeur de

leur caractère égale à 1 est la variable aléatoire  $S = \sum_{i=1}^n X_i \in \mathcal{B}(n, p)$ . On a donc :

$$\mathbf{E}[S] = np \text{ et } \mathbf{V}[S] = np(1-p).$$

La *proportion aléatoire* d'individus dans l'échantillon ayant une valeur de leur caractère égale à 1 est donc finalement donnée par la variable aléatoire  $F = \frac{S}{n} = \bar{X}_n$ . On a :

$$\mathbf{E}[F] = p \text{ et } \mathbf{V}[F] = \frac{p(1-p)}{n}.$$

D'après les résultats ci-dessus, on peut donc considérer que si l'échantillon est suffisamment grand  $F \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$ .

### 9.3. Distribution de la variance aléatoire d'un échantillon

**Proposition 9.3.a :** *Espérance et variance de la variance aléatoire d'un échantillon*

Soit un échantillon aléatoire  $(X_1, \dots, X_n)$ . Les variables  $X_n$  étant identiquement distribuées, on note respectivement  $\mu$  et  $\sigma^2$  leur espérance et leur variance commune.

Soit  $S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \left( \frac{1}{n} \sum_{i=1}^n X_i^2 \right) - (\bar{X})^2$  la variance aléatoire de cet échantillon.

L'espérance et la variance de  $S_X^2$  sont données par :

$$\mathbf{E}[S_X^2] = \frac{n-1}{n} \sigma^2;$$

$$\mathbf{V}[S_X^2] = \frac{n-1}{n^3} [(n-1)\mu_4 - (n-3)\sigma^4] \underset{n \rightarrow \infty}{\sim} \frac{\mu_4 - \sigma^4}{n}, \text{ avec } \mu_4 = \mathbf{E}[(X_i - \mu)^4], i \in \llbracket 1, n \rrbracket.$$

Démonstration (résultat admis pour la variance) :

$$\begin{aligned} \mathbf{E}[S_X^2] &= \mathbf{E}\left[\frac{1}{n} \sum_{i=1}^n X_i^2\right] - \mathbf{E}[(\bar{X})^2] = \frac{1}{n} \sum_{i=1}^n \mathbf{E}[X_i^2] - \mathbf{E}[(\bar{X})^2] \\ &= \left(\frac{1}{n} \sum_{i=1}^n \mathbf{E}[X_i^2]\right) - \mu^2 - \left(\mathbf{E}[(\bar{X})^2] - \mu^2\right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n} \left( \sum_{i=1}^n \mathbf{E}[X_i^2] - n \mu^2 \right) - \mathbf{V}[\bar{X}] \\
&= \frac{1}{n} \sum_{i=1}^n (\mathbf{E}[X_i^2] - \mu^2) - \mathbf{V}[\bar{X}] \\
&= \frac{1}{n} \sum_{i=1}^n \mathbf{V}[X_i] - \mathbf{V}[\bar{X}] = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2.
\end{aligned}$$

**Proposition 9.3.a :** *Distribution asymptotique de la variance aléatoire d'un échantillon*

Soit un échantillon aléatoire  $(X_1, \dots, X_n)$ . Les variables  $X_n$  étant identiquement distribuées, on note respectivement  $\mu$  et  $\sigma^2$  leur espérance et leur variance commune.

La variance aléatoire de l'échantillon,  $S_X^2$ , est approximativement distribuée suivant une loi gaussienne d'espérance  $\sigma^2$  et de variance  $\frac{\mu_4 - \sigma^4}{n}$

Démonstration :

La distribution de  $S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ , se déduit de l'application du théorème de la limite centrale aux variables aléatoires  $(X_i - \bar{X})^2$ .

D'après le théorème de la limite centrale, la suite de variables aléatoires

$\left( \sum_{i=1}^n (X_i - \bar{X})^2 \right)_{n \in \mathbb{N}^*}$  converge en loi vers une variable aléatoire gaussienne d'espérance

$\mathbf{E} \left[ \sum_{i=1}^n (X_i - \bar{X})^2 \right]$  et de variance  $\mathbf{V} \left[ \sum_{i=1}^n (X_i - \bar{X})^2 \right]$ .

Or  $\mathbf{E}[S_X^2] = \frac{1}{n} \mathbf{E} \left[ \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{n-1}{n} \sigma^2$ , donc  $\mathbf{E} \left[ \sum_{i=1}^n (X_i - \bar{X})^2 \right] = (n-1) \sigma^2$ ,

et  $\mathbf{V}[S_X^2] = \frac{1}{n^2} \mathbf{V} \left[ \sum_{i=1}^n (X_i - \bar{X})^2 \right] \underset{n \rightarrow \infty}{\sim} \frac{\mu_4 - \sigma^4}{n}$ , donc  $\mathbf{V} \left[ \sum_{i=1}^n (X_i - \bar{X})^2 \right] \underset{n \rightarrow \infty}{\sim} n (\mu_4 - \sigma^4)$ .

Donc, pour un échantillon suffisamment grand,  $\sum_{i=1}^n (X_i - \bar{X})^2 \sim \mathcal{N}((n-1) \sigma^2, n (\mu_4 - \sigma^4))$ ,

et donc finalement  $S_X^2 \sim \mathcal{N} \left( \sigma^2, \sqrt{\frac{\mu_4 - \sigma^4}{n}} \right)$ .

## 9.4. Cas des échantillons gaussiens

### Définition 9.4.a : Echantillon gaussien

Un échantillon aléatoire  $(X_1, \dots, X_n)$  est dit *gaussien*, s.s. si toute combinaison linéaire des  $X_i$  est gaussienne. En particulier, si les variables aléatoires  $X_1, \dots, X_n$  d'un échantillon aléatoire  $(X_1, \dots, X_n)$  sont distribuées suivant une loi gaussienne, alors cet échantillon est gaussien.

### Proposition 9.4.b : Distribution de la moyenne aléatoire d'un échantillon gaussien

Etant donné un échantillon aléatoire gaussien  $(X_1, \dots, X_n)$ , du fait de l'indépendance des variables aléatoires  $X_i, i \in \llbracket 1, n \rrbracket$ , la moyenne aléatoire  $\bar{X}_n$  est distribuée exactement suivant une loi gaussienne d'espérance  $\mu$  et de variance  $\sigma^2 / n$  :

$$\bar{X}_n \in \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right).$$

### Proposition 9.4.c (admise): Distribution de la variance aléatoire d'un échantillon gaussien

Etant donné un échantillon aléatoire gaussien  $(X_1, \dots, X_n)$ , la variable aléatoire  $\frac{n S_{\bar{X}}^2}{\sigma^2}$  est distribuée suivant une loi du chi-deux à  $n - 1$  degrés de liberté :

$$\frac{n S_{\bar{X}}^2}{\sigma^2} \in \chi_{n-1}^2.$$

### Définition 9.4.d : Loi de Student

Une variable aléatoire réelle  $X$  suit une loi de Student à  $n$  degrés de liberté, et on note  $X \in \mathcal{T}(n)$ , si elle admet pour densité la fonction :

$$f: \mathbb{R} \rightarrow \mathbb{R}_+ \\ t \mapsto \frac{1}{\sqrt{n} B\left(\frac{1}{2}, \frac{n}{2}\right) \left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}, \text{ avec } B(n, p) = \frac{\Gamma(n) \Gamma(p)}{\Gamma(n+p)}$$

### Propositions 9.4.e (admisses) :

1. Si  $U$  et  $X$  sont deux variables aléatoires indépendantes telles que  $U \in \mathcal{N}(0, 1)$  et  $X \in \chi_n^2$ ,

alors  $T = \frac{U}{\sqrt{\frac{X}{n}}} \in \mathcal{T}(n)$ .

2.  $E[X] = 0$  si  $n > 1$  ; pour  $n = 1, f(t) = \frac{1}{\pi(1+t^2)}$ ,  $\mathcal{T}(1)$  est la loi de Cauchy, qui n'admet

aucun moment fini ;

3.  $V[X] = \frac{n}{n-2}$  si  $n > 2$  ;

4. Si  $f_n$  désigne la fonction de densité d'une loi du chi-deux à  $n$  degrés de liberté, la suite de fonction  $(f_n)_{n \in \mathbb{N}}$  tend, lorsque  $n$  tend vers l'infini, vers la fonction de densité d'une loi normale centrée réduite : la suite de variables aléatoires  $(X_n)_{n \in \mathbb{N}}$ , telle que  $\forall n \in \mathbb{N}^*, X_n \in \mathcal{T}(n)$ , converge en loi vers une loi normale centrée réduite.

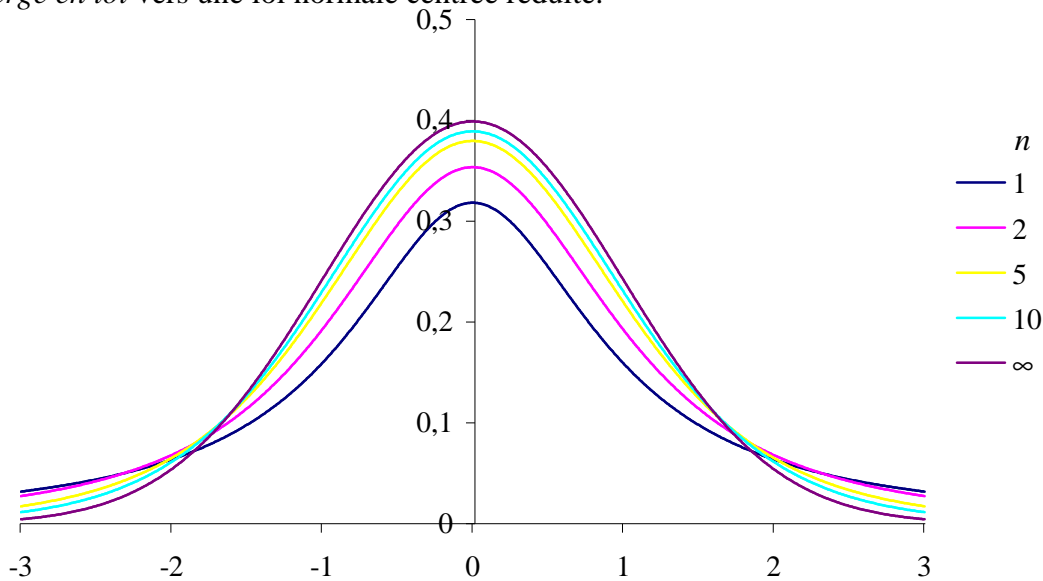


Figure 9.3.f : Fonctions de densité de lois de Student pour quelques valeurs du nombre de degrés de liberté  $n$ .

**Proposition 9.4.g** (admise): *Distribution de la moyenne aléatoire lorsque  $\sigma$  est inconnue*

On sait que  $\bar{X}_n \in \mathcal{N}(\mu, \sigma^2/n)$ , donc  $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \in \mathcal{N}(0, 1)$ , et que  $\frac{n S_X^2}{\sigma^2} \in \chi_{n-1}^2$ . Donc :

$$\frac{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{n S_X^2}{(n-1)\sigma^2}}} = \frac{\bar{X}_n - \mu}{S_X / \sqrt{n-1}} \in \mathcal{T}(n-1).$$

Remarque :

Les distributions qui sont données pour la moyenne aléatoire et la variance aléatoire d'un échantillon gaussien (ou pour des fonctions de celles-ci) sont exactes, et non asymptotiques, et sont donc valables quelle que soit la taille de l'échantillon considéré.

## 10. Notions d'estimation

### 10.1. Généralités

#### Définition 10.1.a : Estimation

Soit un caractère associé aux individus d'une population, représenté par une variable statistique  $X$  distribuée suivant une loi dont un paramètre est inconnu. L'estimation consiste à *estimer* ce paramètre inconnu à partir d'un échantillon donné, en proposant soit une valeur pour celui-ci (estimation *ponctuelle*), soit un intervalle dans lequel celui-ci a une grande probabilité de se trouver (estimation par *intervalle de confiance*).

#### Définition 10.1.b : Estimateur

On appelle *estimateur* d'un paramètre inconnu toute statistique permettant d'estimer ce paramètre.

#### Exemples 10.1.c :

Considérons une variable statistique  $X$  distribuée suivant une loi gaussienne d'espérance  $\mu$  et de variance  $\sigma^2$ .  $\mu$  et  $\sigma^2$  sont deux paramètres de la loi gaussienne de  $X$  que l'on peut chercher à estimer à partir d'un échantillon aléatoire  $(X_1, \dots, X_n)$ .

1°) un estimateur de l'espérance  $\mu$  est la moyenne aléatoire :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i ;$$

2°) un estimateur de la variance  $\sigma^2$  est la variance aléatoire :

$$S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \left( \frac{1}{n} \sum_{i=1}^n X_i^2 \right) - (\bar{X})^2 ;$$

3°) si l'espérance  $\mu$  est connue, un autre estimateur de la variance  $\sigma^2$  est :

$$T_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 = \left( \frac{1}{n} \sum_{i=1}^n X_i^2 \right) - \mu^2 .$$

**Définition 10.1.d : Biais d'un estimateur**

Etant donné un échantillon aléatoire  $(X_1, \dots, X_n)$ , un estimateur  $T(X_1, \dots, X_n; n)$  d'un paramètre  $\theta$  est *sans biais* si et seulement si :

$$\mathbf{E} [ T(X_1, \dots, X_n; n) ] = \theta .$$

Si tel n'est pas le cas, alors l'estimateur est dit *biaisé*, le *biais* étant égal à la différence  $\mathbf{E} [ T(X_1, \dots, X_n; n) ] - \theta$ .

L'estimateur est *asymptotiquement sans biais* si et seulement si :

$$\lim_{n \rightarrow \infty} \mathbf{E} [ T(X_1, \dots, X_n; n) ] = \theta .$$

**Exemples 10.1.e :**

Soit un échantillon  $(X_1, \dots, X_n)$ , avec :  $\forall i \in \llbracket 1, n \rrbracket, \mathbf{E} [ X_i ] = \mu$  et  $\mathbf{V} [ X_i ] = \sigma^2$ .

1°)  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  est un estimateur sans biais de  $\mu$  puisque  $\mathbf{E} [ \bar{X} ] = \mu$  ;

2°)  $S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  est un estimateur biaisé de  $\sigma^2$  puisque

$$\mathbf{E} [ S_X^2 ] = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2 .$$

3°)  $S_X'^2 = \frac{n}{n-1} S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  est donc un estimateur sans biais de  $\sigma^2$ .

**Définition 10.1.f : Estimateur convergent**

Etant donné un échantillon aléatoire  $(X_1, \dots, X_n)$ , un estimateur  $T(X_1, \dots, X_n; n)$  d'un paramètre  $\theta$  est *convergent* si et seulement si la suite de variables aléatoires :

$$( T(X_1, \dots, X_n; n) )_{n \in \mathbb{N}}$$

converge en probabilités vers  $\theta$ .

**Exemples 10.1.g :**

Soit un échantillon  $(X_1, \dots, X_n)$ , avec :  $\forall i \in \llbracket 1, n \rrbracket, \mathbf{E} [ X_i ] = \mu$  et  $\mathbf{V} [ X_i ] = \sigma^2$ .

1°) D'après la loi faible des grands nombres la suite  $(\bar{X}_n)_{n \in \mathbb{N}} = \left( \frac{1}{n} \sum_{i=1}^n X_i \right)_{n \in \mathbb{N}}$  converge en probabilités vers  $\mu$ , donc  $\bar{X}$  est un estimateur convergent.

2°) On démontre de même que les estimateurs  $S_X^2$ ,  $S_X'^2$  et  $T_X^2$  sont convergents (*i.e.* que les suites de variables aléatoires qui leur sont associées convergent en probabilité vers  $\sigma^2$ ).

## 10.2. Détermination d'un estimateur de maximum de vraisemblance

Soient un caractère associé aux individus d'une population, représenté par une variable statistique  $X$  et un échantillon aléatoire  $(X_1, \dots, X_n)$ . On considère les variables aléatoires  $X_i$  indépendantes et identiquement distribuées suivant une loi  $\mathcal{L}$  de paramètre  $\theta$ .

### Définition 10.2.a : Fonction de vraisemblance

On appelle *fonction de vraisemblance* du paramètre  $\theta$  sur l'échantillon  $(X_1, \dots, X_n)$  la fonction  $V(x_1, \dots, x_n; \theta)$ , définie par :

- si  $\mathcal{L}$  est une loi discrète,

$$V(x_1, \dots, x_n; \theta) = \mathbf{P}((X_1, \dots, X_n) = (x_1, \dots, x_n); \theta) = \prod_{i=1}^n \mathbf{P}(X_i = x_i; \theta),$$

où  $\mathbf{P}(A; \theta)$  représente la probabilité d'un événement  $A$  en fonction de  $\theta$  ;

- si  $\mathcal{L}$  est une loi continue,

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

où  $f_{(X_1, \dots, X_n)}(x_1, \dots, x_n; \theta)$  (respectivement  $f(x_i; \theta)$ ,  $1 \leq i \leq n$ ) représente la densité du vecteur aléatoire  $(X_1, \dots, X_n)$  au point  $(x_1, \dots, x_n)$  (respectivement la densité de chaque variable aléatoire  $X_i$  au point  $x_i$ ) exprimée comme une fonction de  $(x_1, \dots, x_n)$  et de  $\theta$  (respectivement des fonctions de  $x_i$  et de  $\theta$ ).

### Définition-Proposition 10.2.b : Estimateur de maximum de vraisemblance

$\hat{\theta}(X_1, \dots, X_n)$  est l'*estimateur de maximum de vraisemblance* du paramètre  $\theta$  si la fonction de vraisemblance  $V(x_1, \dots, x_n; \theta)$  présente un maximum en  $\hat{\theta}(x_1, \dots, x_n)$ .

En notant par simplification  $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ , et sous réserve d'existence des dérivées partielles,  $\hat{\theta}$  est solution du système d'équations :

$$\begin{cases} \frac{\partial V}{\partial \theta} \Big|_{(x_1, \dots, x_n; \hat{\theta})} = 0 \\ \frac{\partial^2 V}{\partial \theta^2} \Big|_{(x_1, \dots, x_n; \hat{\theta})} < 0 \end{cases}$$

### Remarques :

1°) Comme la fonction de vraisemblance est positive, et comme la fonction logarithme est croissante sur  $\mathbb{R}_+^*$ , ce système d'équation peut être remplacé par :

$$\begin{cases} \frac{\partial \ln V}{\partial \theta} \Big|_{(x_1, \dots, x_n; \hat{\theta})} = 0 \\ \frac{\partial^2 \ln V}{\partial \theta^2} \Big|_{(x_1, \dots, x_n; \hat{\theta})} < 0 \end{cases}.$$

Ce changement est avantageux du fait que  $V$  s'exprimant comme un produit,  $\ln V$  s'exprime comme une somme dont la dérivée est alors beaucoup plus facile à calculer.

2°) Cette méthode peut également être appliquée à une loi  $\mathcal{L}$  ayant  $k$  paramètres  $\theta_1, \dots, \theta_k$ .

$\hat{\theta}_1(X_1, \dots, X_n), \dots, \hat{\theta}_k(X_1, \dots, X_n)$  sont les *estimateurs de maximum de vraisemblance* des paramètres  $\theta_1, \dots, \theta_k$  si la fonction de vraisemblance  $V(x_1, \dots, x_n; \theta_1, \dots, \theta_k)$  présente un maximum en  $(\hat{\theta}_1(x_1, \dots, x_n), \dots, \hat{\theta}_k(x_1, \dots, x_n))$  (cf. exemple 4°)).

### **Exemples 10.2.c :**

1°) Estimation du paramètre  $\lambda$  d'une loi de Poisson

$$\text{On a } \mathbf{P}(X_i = x_i) = \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}.$$

$$\text{donc } V(x_1, \dots, x_n; \lambda) = \prod_{i=1}^n \mathbf{P}(X_i = x_i) = \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\lambda},$$

$$\text{donc } \ln V(x_1, \dots, x_n; \lambda) = \left( \sum_{i=1}^n x_i \right) \ln \lambda - \ln \left( \prod_{i=1}^n x_i! \right) - n\lambda$$

$$\text{donc } \frac{\partial \ln V}{\partial \lambda}(x_1, \dots, x_n; \lambda) = \frac{\sum_{i=1}^n x_i}{\lambda} - n.$$

$$\text{On a alors } \frac{\partial \ln V}{\partial \lambda}(x_1, \dots, x_n; \hat{\lambda}) = \frac{\sum_{i=1}^n x_i}{\hat{\lambda}} - n = 0 \Rightarrow \hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i.$$

$$\text{Par ailleurs } \frac{\partial^2 \ln V}{\partial \lambda^2}(x_1, \dots, x_n; \lambda) = -\frac{\sum_{i=1}^n x_i}{\lambda^2}, \text{ donc } \frac{\partial^2 \ln V}{\partial \lambda^2}(x_1, \dots, x_n; \hat{\lambda}) = -\frac{\sum_{i=1}^n x_i}{\hat{\lambda}^2} \leq 0,$$

car  $\forall i \in \llbracket 1, n \rrbracket, x_i \geq 0$



$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i$  est une réalisation de la variable aléatoire  $\hat{\lambda}(X_1, \dots, X_n)$  pour  $(X_1, \dots, X_n) = (x_1, \dots, x_n)$ , donc l'estimateur de maximum de vraisemblance du paramètre  $\lambda$  est :

$$\hat{\lambda}(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}.$$

2°) Estimation de l'espérance  $m$  d'une loi gaussienne de variance connue  $\sigma^2$  ;

On a  $f(x_i; m) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left( \frac{x_i - m}{\sigma} \right)^2}$ ,

donc  $V(x_1, \dots, x_n; m) = \prod_{i=1}^n f(x_i; m) = \left( \frac{1}{\sqrt{2\pi} \sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2}$ ,

donc  $\ln V(x_1, \dots, x_n; m) = -n \ln(\sqrt{2\pi} \sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2$ ,

donc  $\frac{\partial \ln V}{\partial m}(x_1, \dots, x_n; m) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - m)$ .

On a alors  $\frac{\partial \ln V}{\partial m}(x_1, \dots, x_n; \hat{m}) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \hat{m}) = 0 \Rightarrow \hat{m} = \frac{1}{n} \sum_{i=1}^n x_i$ .

Par ailleurs  $\frac{\partial^2 \ln V}{\partial m^2}(x_1, \dots, x_n; m) = -\frac{n}{\sigma^2} \Rightarrow \frac{\partial^2 \ln V}{\partial m^2}(x_1, \dots, x_n; \hat{m}) = -\frac{n}{\sigma^2} < 0$ .

$\hat{m} = \frac{1}{n} \sum_{i=1}^n x_i$  est une réalisation de la variable aléatoire  $\hat{m}(X_1, \dots, X_n)$  pour  $(X_1, \dots, X_n) = (x_1, \dots, x_n)$ , donc l'estimateur de maximum de vraisemblance du paramètre  $m$  est :

$$\hat{m}(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}.$$

3°) Estimation de la variance  $\sigma^2$  d'une loi gaussienne d'espérance connue  $m$  ;

On a  $f(x_i; \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left( \frac{x_i - m}{\sigma} \right)^2}$ ,

donc  $V(x_1, \dots, x_n; \sigma^2) = \prod_{i=1}^n f(x_i; \sigma^2) = \left( \frac{1}{\sqrt{2\pi} \sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2}$ .

Posons  $\theta = \sigma^2$ . On a alors  $V(x_1, \dots, x_n; \theta) = \left( \frac{1}{\sqrt{2\pi}} \right)^n \theta^{-\frac{n}{2}} e^{-\frac{1}{2\theta} \sum_{i=1}^n (x_i - m)^2}$ ,

donc  $\ln V(x_1, \dots, x_n; \theta) = -n \ln(\sqrt{2\pi}) - \frac{n}{2} \ln \theta - \frac{1}{2\theta} \sum_{i=1}^n (x_i - m)^2$ ,

$$\text{donc } \frac{\partial \ln V}{\partial \theta} (x_1, \dots, x_n; \theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^n (x_i - m)^2.$$

$$\text{On a alors } \frac{\partial \ln V}{\partial \theta} (x_1, \dots, x_n; \hat{\theta}) = -\frac{n}{2\hat{\theta}} + \frac{1}{2\hat{\theta}^2} \sum_{i=1}^n (x_i - m)^2 = 0 \Rightarrow \hat{\theta} = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2.$$

$$\text{Par ailleurs } \frac{\partial^2 \ln V}{\partial \theta^2} (x_1, \dots, x_n; \theta) = \frac{n}{2\theta^2} - \frac{1}{\theta^3} \sum_{i=1}^n (x_i - m)^2,$$

$$\text{donc } \frac{\partial^2 \ln V}{\partial \theta^2} (x_1, \dots, x_n; \hat{\theta}) = \frac{n}{2\hat{\theta}^2} - \frac{1}{\hat{\theta}^3} \sum_{i=1}^n (x_i - m)^2 = \frac{1}{\hat{\theta}^2} \left( \frac{n}{2} - \frac{1}{\hat{\theta}} \sum_{i=1}^n (x_i - m)^2 \right) = \frac{1}{\hat{\theta}^2} \left( \frac{n}{2} - n \right) < 0.$$

$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2$  est une réalisation de la variable aléatoire  $\hat{\theta}(X_1, \dots, X_n)$  pour  $(X_1, \dots, X_n) = (x_1, \dots, x_n)$ , donc l'estimateur de maximum de vraisemblance du paramètre  $\theta = \sigma^2$  est :

$$\hat{\theta}(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2.$$

4°) La densité d'une variable aléatoire gaussienne d'espérance mathématique  $m$  et de variance

$$\sigma^2 \text{ est } f(x_i; m, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left( \frac{x_i - m}{\sigma} \right)^2},$$

$$\text{donc } V(x_1, \dots, x_n; m, \sigma^2) = \prod_{i=1}^n f(x_i; m, \sigma^2) = \left( \frac{1}{\sqrt{2\pi} \sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2}.$$

$$\text{Posons } \theta = \sigma^2. \text{ On a alors } V(x_1, \dots, x_n; m, \theta) = \left( \frac{1}{\sqrt{2\pi}} \right)^n \theta^{-\frac{n}{2}} e^{-\frac{1}{2\theta} \sum_{i=1}^n (x_i - m)^2},$$

$$\text{donc } \ln V(x_1, \dots, x_n; m, \theta) = -n \ln(\sqrt{2\pi}) - \frac{n}{2} \ln \theta - \frac{1}{2\theta} \sum_{i=1}^n (x_i - m)^2,$$

$$\text{donc } \frac{\partial \ln V}{\partial m} (x_1, \dots, x_n; m, \theta) = \frac{1}{\theta} \sum_{i=1}^n (x_i - m), \text{ et } \frac{\partial \ln V}{\partial \theta} (x_1, \dots, x_n; m, \theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^n (x_i - m)^2.$$

$$\text{On a alors le système d'équations } \begin{cases} \frac{\partial \ln V}{\partial m} (x_1, \dots, x_n; \hat{m}, \hat{\theta}) = \frac{1}{\hat{\theta}} \sum_{i=1}^n (x_i - \hat{m}) = 0 \\ \frac{\partial \ln V}{\partial \theta} (x_1, \dots, x_n; \hat{m}, \hat{\theta}) = -\frac{n}{2\hat{\theta}} + \frac{1}{2\hat{\theta}^2} \sum_{i=1}^n (x_i - \hat{m})^2 = 0 \end{cases},$$

$$\text{dont la solution est } \begin{cases} \hat{m} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x} \\ \hat{\theta} = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{m})^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \end{cases}.$$

Soient  $r = \frac{\partial^2 \ln V}{\partial \hat{m}^2}(x_1, \dots, x_n; \hat{m}, \hat{\theta}) = -\frac{n}{\hat{\theta}}$ ,  $t = \frac{\partial^2 \ln V}{\partial \hat{\theta}^2}(x_1, \dots, x_n; \hat{m}, \hat{\theta}) = \frac{n}{2\hat{\theta}^2} - \frac{1}{\hat{\theta}^3} \sum_{i=1}^n (x_i - \hat{m})^2$

$$= \frac{n}{2\hat{\theta}^2} - \frac{n}{\hat{\theta}^2} = -\frac{n}{2\hat{\theta}^2} \quad (\text{car } \sum_{i=1}^n (x_i - \hat{m})^2 = n\hat{\theta}^2), \quad \text{et } s = \frac{\partial^2 \ln V}{\partial \hat{m} \partial \hat{\theta}}(x_1, \dots, x_n; \hat{m}, \hat{\theta}) =$$

$$-\frac{1}{\hat{\theta}^2} \sum_{i=1}^n (x_i - \hat{m}) = 0. \quad \text{On a } rt - s^2 = \frac{n^2}{2\hat{\theta}^3} > 0 \text{ car } \hat{\theta} > 0 \text{ et } r < 0, \text{ donc } (\hat{m}, \hat{\theta}) \text{ est bien un}$$

maximum de la fonction  $\ln V$ .

$\hat{m} = \bar{x}$  et  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$  sont donc les réalisations pour  $(X_1, \dots, X_n) = (x_1, \dots, x_n)$  des estimateurs de l'espérance et de la variance d'une variable aléatoire gaussienne par la méthode du maximum de vraisemblance.

Ces estimateurs sont donc respectivement  $\hat{m} = \bar{X}$  et  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ .

### 10.3. Estimation par intervalles de confiance

#### 10.3.1 Définitions et principes généraux

**Définition 10.3.1.a :** *Intervalle de confiance*

Soient un échantillon aléatoire  $(X_1, \dots, X_n)$  et un estimateur  $T(X_1, \dots, X_n; n)$  d'un paramètre  $\theta$ . On appelle intervalle de confiance sur  $\theta$ , au niveau de confiance  $1 - \alpha$  (valeur proche de 1), un intervalle  $[A; B]$  construit à partir de l'échantillon aléatoire (ses bornes  $A$  et  $B$  sont donc des variables aléatoires) telle que :  $\mathbf{P}(\theta \in [A; B]) = 1 - \alpha$ .

**Proposition 10.3.1.b :** *Principe de construction de l'intervalle de confiance*

La distribution de  $T$  dépend de la valeur de  $\theta$ , donc les fractiles d'ordre  $\alpha/2$  et  $1 - \alpha/2$  sont donc des fonctions de  $\theta$  que l'on peut noter respectivement  $t_{\alpha/2}$  et  $t_{1-\alpha/2}$ . On a donc :

$$\mathbf{P}(t_{\alpha/2}(\theta) \leq T(X_1, \dots, X_n; n) \leq t_{1-\alpha/2}(\theta)) = 1 - \alpha.$$

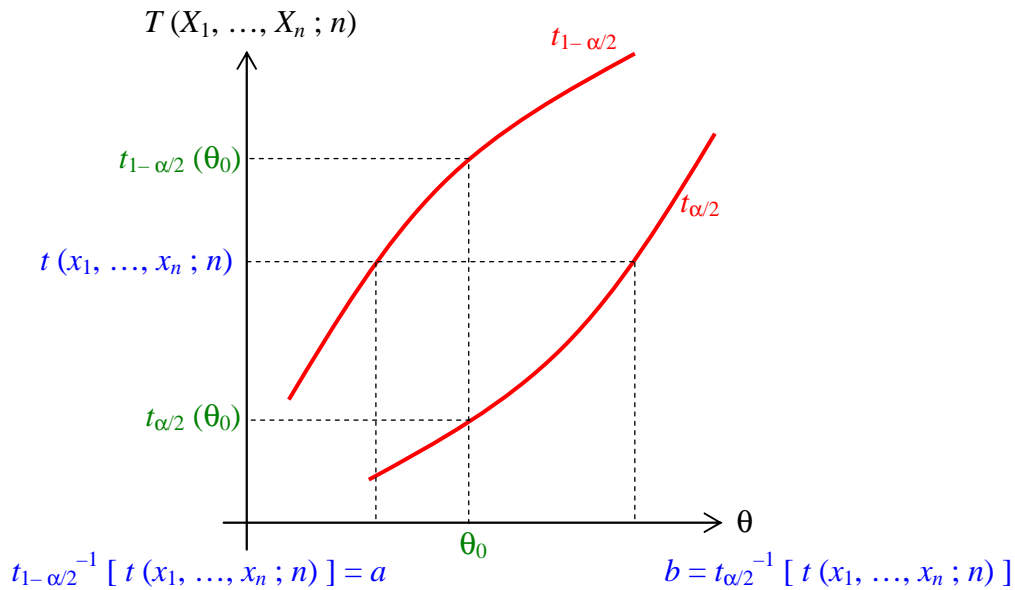
L'intervalle de confiance est construit en inversant, quand cela est possible, chacune des deux inégalités ci-dessus. Si par exemple les fonctions  $t_{\alpha/2}$  et  $t_{1-\alpha/2}$  sont croissantes, on a

$$t_{\alpha/2}(\theta) \leq T(X_1, \dots, X_n; n) \Leftrightarrow \theta \leq t_{\alpha/2}^{-1}[T(X_1, \dots, X_n; n)], \text{ et}$$

$$T(X_1, \dots, X_n; n) \leq t_{1-\alpha/2}(\theta) \Leftrightarrow \theta \geq t_{1-\alpha/2}^{-1}[T(X_1, \dots, X_n; n)].$$

On obtient donc l'intervalle de confiance :

$$\mathbf{P} \left( t_{1-\alpha/2}^{-1} [ T(X_1, \dots, X_n; n) ] \leq \theta \leq t_{\alpha/2}^{-1} [ T(X_1, \dots, X_n; n) ] \right) = 1 - \alpha .$$



**Figure 10.3.1.c** : construction de l'intervalle de confiance

### 10.3.2 Intervalle de confiance de l'espérance d'une loi gaussienne

Soit un échantillon aléatoire gaussien  $(X_1, \dots, X_n) : \forall i \in \llbracket 1, n \rrbracket, X_i \in \mathcal{N}(\mu, \sigma)$ .

**Proposition 10.3.2.a** : *Cas où  $\sigma$  est connue*

On sait que  $\bar{X} \in \mathcal{N}(\mu, \sigma^2/n)$ , donc  $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \in \mathcal{N}(0, 1)$ .

En notant  $u_\alpha$  le fractile d'ordre  $\alpha$  de la loi gaussienne, on a donc :

$$\mathbf{P}\left(u_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < u_{1-\alpha/2}\right) = 1 - \alpha,$$

et donc 
$$\mathbf{P}\left(\bar{X} + u_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha .$$

Remarque :

Une valeur de  $\alpha$  couramment utilisée est  $\alpha = 0.05$  (niveau de confiance de 95 %). On a alors les fractiles suivants :  $u_{\alpha/2} = u_{0.025} \approx -1.96$  et  $u_{1-\alpha/2} = u_{0.975} \approx 1.96$ .

**Proposition 10.3.2.b :** *Cas où  $\sigma$  est inconnue*

On sait que  $T = \frac{\bar{X} - \mu}{S'_X / \sqrt{n-1}} = \frac{\bar{X} - \mu}{S'_X / \sqrt{n}} \in \mathcal{T}(n-1)$ .

En notant  $t_{n-1;\alpha}$  le fractile d'ordre  $\alpha$  de la loi de Student à  $n-1$  degrés de liberté, on a donc :

$$\mathbf{P}\left(t_{n-1;\alpha/2} < \frac{\bar{X} - \mu}{S'_X / \sqrt{n}} < t_{n-1;1-\alpha/2}\right) = 1 - \alpha,$$

et donc  $\mathbf{P}\left(\bar{X} + t_{n-1;\alpha/2} \frac{S'_X}{\sqrt{n}} < \mu < \bar{X} + t_{n-1;1-\alpha/2} \frac{S'_X}{\sqrt{n}}\right) = 1 - \alpha$ .

### 10.3.3 Intervalle de confiance de la variance d'une loi gaussienne

Soit un échantillon aléatoire gaussien  $(X_1, \dots, X_n) : \forall i \in \llbracket 1, n \rrbracket, X_i \in \mathcal{N}(\mu, \sigma)$ .

**Proposition 10.3.3.a :** *Cas où  $\mu$  est connue*

$T_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$  est le meilleur estimateur de  $\sigma^2$  et  $\frac{nT_X^2}{\sigma^2} = \sum_{i=1}^n \left( \frac{X_i - \mu}{\sigma} \right)^2 \in \chi_n^2$

puisque'il s'agit de la somme de  $n$  variables aléatoires gaussiennes centrées réduites indépendantes.

En notant  $\chi_{n;\alpha}^2$  le fractile d'ordre  $\alpha$  de la loi du chi-deux à  $n$  degrés de liberté, on a donc :

$$\mathbf{P}\left(\chi_{n;\alpha/2}^2 < \frac{nT_X^2}{\sigma^2} < \chi_{n;1-\alpha/2}^2\right) = 1 - \alpha,$$

et donc  $\mathbf{P}\left(\frac{nT_X^2}{\chi_{n;1-\alpha/2}^2} < \sigma^2 < \frac{nT_X^2}{\chi_{n;\alpha/2}^2}\right) = 1 - \alpha$ .

**Proposition 10.3.3.b :** *Cas où  $\mu$  est inconnue*

Les estimateurs de  $\sigma^2$  utilisés sont  $S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  ou bien  $S_X'^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ .

On sait que  $\frac{n S_X^2}{\sigma^2} = \frac{(n-1) S_X'^2}{\sigma^2} \in \chi_{n-1}^2$

En notant  $\chi_{n-1;\alpha}^2$  le fractile d'ordre  $\alpha$  de la loi du chi-deux à  $n - 1$  degrés de liberté, on a donc :

$$\mathbf{P}\left(\chi_{n-1;\alpha/2}^2 < \frac{n S_X^2}{\sigma^2} < \chi_{n-1;1-\alpha/2}^2\right) = 1 - \alpha,$$

et donc 
$$\mathbf{P}\left(\frac{\frac{n S_X^2}{2}}{\chi_{n-1;1-\alpha/2}^2} < \sigma^2 < \frac{\frac{n S_X^2}{2}}{\chi_{n-1;\alpha/2}^2}\right) = \mathbf{P}\left(\frac{\frac{(n-1) S_X'^2}{2}}{\chi_{n-1;1-\alpha/2}^2} < \sigma^2 < \frac{\frac{(n-1) S_X'^2}{2}}{\chi_{n-1;\alpha/2}^2}\right) = 1 - \alpha.$$

#### 10.3.4 Cas des échantillons non gaussiens

Soit un échantillon non gaussien,  $(X_1, \dots, X_n)$  tel que  $\forall i \in \llbracket 1, n \rrbracket$ ,  $\mathbf{E}[X_i] = \mu$  et  $\mathbf{V}[X_i] = \sigma^2$ .  $\mu$  n'est pas forcément un paramètre de la loi des  $X_i$ . Cependant, si l'échantillon est suffisamment grand, on sait, en vertu du théorème de la limite centrale, que  $\bar{X}$  est approximativement distribué suivant une loi gaussienne d'espérance  $\mu$  et de variance  $\sigma^2 / n$ . On peut donc, avec la construction présentée ci-dessus obtenir un intervalle de confiance approximatif sur  $\mu$  et sur  $\sigma^2$ .

Dans les cas où l'un des deux paramètres est inconnu, les correctifs appliqués dans le cas gaussien (utilisation de  $S_X^2$  et d'une loi de Student pour un intervalle de confiance sur  $\mu$ , et utilisation de  $\bar{X}$  et d'une loi du chi-deux à  $n$  degrés de liberté pour un intervalle de confiance sur  $\sigma^2$ ) restent approximativement valables et peuvent être utilisés.

**Proposition 10.3.4.a : Intervalle de confiance d'une proportion**

Soient une population dans laquelle la proportion d'individus pour lesquels la valeur d'un caractère binaire vaut 1 est  $p \in [0 ; 1]$ , et un échantillon issu de cette population d'effectif  $n$ , suffisamment grand ( $n \geq 30$ ).

L'intervalle de confiance pour la proportion  $p$ , construit au niveau de confiance  $1 - \alpha$  à partir d'un grand échantillon, est asymptotiquement donné par :

$$\mathbf{P} \left( F + u_{\alpha/2} \sqrt{\frac{F(1-F)}{n}} < p < F + u_{1-\alpha/2} \sqrt{\frac{F(1-F)}{n}} \right) = 1 - \alpha .$$

Démonstration :

On sait que la proportion aléatoire  $F$  d'individus dans cet échantillon qui ont leur caractère égal à 1 est approximativement distribuée suivant une loi gaussienne d'espérance  $p$  et de variance  $\frac{p(1-p)}{n}$ .

On a donc :

$$\mathbf{P} \left( p + u_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} < F < p + u_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}} \right) = 1 - \alpha .$$

On peut donc chercher un intervalle de confiance sur le paramètre  $p$  en inversant les inégalités ci-dessus. Comme  $u_{\alpha/2} = -u_{1-\alpha/2}$ , l'encadrement de  $F$  ci-dessus est équivalent à :

$$|F - p| < u_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}} \Leftrightarrow (F - p)^2 < (u_{1-\alpha/2})^2 \frac{p(1-p)}{n} .$$

La construction de l'intervalle de confiance revient donc à résoudre l'inéquation du second degré d'inconnue  $p$  ci-dessus. Finalement, les bornes de l'intervalle de confiance valent :

$$\frac{2F + \frac{(u_{1-\alpha/2})^2}{n} \pm \sqrt{4 \frac{(u_{1-\alpha/2})^2}{n} F(1-F) + \frac{(u_{1-\alpha/2})^4}{n^2}}}{2 \left( 1 + \frac{(u_{1-\alpha/2})^2}{n} \right)} \underset{n \rightarrow \infty}{\sim} F \pm u_{1-\alpha/2} \sqrt{\frac{F(1-F)}{n}}$$





# **TABLES NUMERIQUES**



**Table 1 : Fonction de répartition de la loi normale centrée réduite :**  
pour  $X \in \mathcal{N}(0, 1)$ , la table donne  $\mathbf{P}(X \leq x)$

$x$	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

(De 3,49 à 3,61 : 0,9998 ; de 3,62 à 3,89 : 0,9999 ; à partir de 3,90 : 1,0000 )

**Table 2 : Fractiles de la loi normale centrée réduite :**  
pour  $X \in \mathcal{N}(0, 1)$ , la table donne  $x = u_\alpha$  tel que  $\mathbf{P}(X \leq x) \approx \alpha$

$\alpha$	0	0,001	0,002	0,003	0,004	0,005	0,006	0,007	0,008	0,009	0,01	
0	$\infty$	3,0902	2,8782	2,7478	2,6521	2,5758	2,5121	2,4573	2,4089	2,3656	2,3263	0,99
0,01	2,3263	2,2904	2,2571	2,2262	2,1973	2,1701	2,1444	2,1201	2,0969	2,0748	2,0537	0,98
0,02	2,0537	2,0335	2,0141	1,9954	1,9774	1,9600	1,9431	1,9268	1,9110	1,8957	1,8808	0,97
0,03	1,8808	1,8663	1,8522	1,8384	1,8250	1,8119	1,7991	1,7866	1,7744	1,7624	1,7507	0,96
0,04	1,7507	1,7392	1,7279	1,7169	1,7060	1,6954	1,6849	1,6747	1,6646	1,6546	1,6449	0,95
0,05	1,6449	1,6352	1,6258	1,6164	1,6072	1,5982	1,5893	1,5805	1,5718	1,5632	1,5548	0,94
0,06	1,5548	1,5464	1,5382	1,5301	1,5220	1,5141	1,5063	1,4985	1,4909	1,4833	1,4758	0,93
0,07	1,4758	1,4684	1,4611	1,4538	1,4466	1,4395	1,4325	1,4255	1,4187	1,4118	1,4051	0,92
0,08	1,4051	1,3984	1,3917	1,3852	1,3787	1,3722	1,3658	1,3595	1,3532	1,3469	1,3408	0,91
0,09	1,3408	1,3346	1,3285	1,3225	1,3165	1,3106	1,3047	1,2988	1,2930	1,2873	1,2816	0,90
0,10	1,2816	1,2759	1,2702	1,2646	1,2591	1,2536	1,2481	1,2426	1,2372	1,2319	1,2265	0,89
0,11	1,2265	1,2212	1,2160	1,2107	1,2055	1,2004	1,1952	1,1901	1,1850	1,1800	1,1750	0,88
0,12	1,1750	1,1700	1,1650	1,1601	1,1552	1,1503	1,1455	1,1407	1,1359	1,1311	1,1264	0,87
0,13	1,1264	1,1217	1,1170	1,1123	1,1077	1,1031	1,0985	1,0939	1,0893	1,0848	1,0803	0,86
0,14	1,0803	1,0758	1,0714	1,0669	1,0625	1,0581	1,0537	1,0494	1,0451	1,0407	1,0364	0,85
0,15	1,0364	1,0322	1,0279	1,0237	1,0194	1,0152	1,0110	1,0069	1,0027	0,9986	0,9945	0,84
0,16	0,9945	0,9904	0,9863	0,9822	0,9782	0,9741	0,9701	0,9661	0,9621	0,9581	0,9542	0,83
0,17	0,9542	0,9502	0,9463	0,9424	0,9385	0,9346	0,9307	0,9269	0,9230	0,9192	0,9154	0,82
0,18	0,9154	0,9116	0,9078	0,9040	0,9002	0,8965	0,8927	0,8890	0,8853	0,8816	0,8779	0,81
0,19	0,8779	0,8742	0,8706	0,8669	0,8632	0,8596	0,8560	0,8524	0,8488	0,8452	0,8416	0,80
0,20	0,8416	0,8381	0,8345	0,8310	0,8274	0,8239	0,8204	0,8169	0,8134	0,8099	0,8064	0,79
0,21	0,8064	0,8030	0,7995	0,7961	0,7926	0,7892	0,7858	0,7824	0,7790	0,7756	0,7722	0,78
0,22	0,7722	0,7688	0,7655	0,7621	0,7588	0,7554	0,7521	0,7488	0,7454	0,7421	0,7388	0,77
0,23	0,7388	0,7356	0,7323	0,7290	0,7257	0,7225	0,7192	0,7160	0,7128	0,7095	0,7063	0,76
0,24	0,7063	0,7031	0,6999	0,6967	0,6935	0,6903	0,6871	0,6840	0,6808	0,6776	0,6745	0,75
0,25	0,6745	0,6713	0,6682	0,6651	0,6620	0,6588	0,6557	0,6526	0,6495	0,6464	0,6433	0,74
0,26	0,6433	0,6403	0,6372	0,6341	0,6311	0,6280	0,6250	0,6219	0,6189	0,6158	0,6128	0,73
0,27	0,6128	0,6098	0,6068	0,6038	0,6008	0,5978	0,5948	0,5918	0,5888	0,5858	0,5828	0,72
0,28	0,5828	0,5799	0,5769	0,5740	0,5710	0,5681	0,5651	0,5622	0,5592	0,5563	0,5534	0,71
0,29	0,5534	0,5505	0,5476	0,5446	0,5417	0,5388	0,5359	0,5330	0,5302	0,5273	0,5244	0,70
0,30	0,5244	0,5215	0,5187	0,5158	0,5129	0,5101	0,5072	0,5044	0,5015	0,4987	0,4958	0,69
0,31	0,4958	0,4930	0,4902	0,4874	0,4845	0,4817	0,4789	0,4761	0,4733	0,4705	0,4677	0,68
0,32	0,4677	0,4649	0,4621	0,4593	0,4565	0,4538	0,4510	0,4482	0,4454	0,4427	0,4399	0,67
0,33	0,4399	0,4372	0,4344	0,4316	0,4289	0,4261	0,4234	0,4207	0,4179	0,4152	0,4125	0,66
0,34	0,4125	0,4097	0,4070	0,4043	0,4016	0,3989	0,3961	0,3934	0,3907	0,3880	0,3853	0,65
0,35	0,3853	0,3826	0,3799	0,3772	0,3745	0,3719	0,3692	0,3665	0,3638	0,3611	0,3585	0,64
0,36	0,3585	0,3558	0,3531	0,3505	0,3478	0,3451	0,3425	0,3398	0,3372	0,3345	0,3319	0,63
0,37	0,3319	0,3292	0,3266	0,3239	0,3213	0,3186	0,3160	0,3134	0,3107	0,3081	0,3055	0,62
0,38	0,3055	0,3029	0,3002	0,2976	0,2950	0,2924	0,2898	0,2871	0,2845	0,2819	0,2793	0,61
0,39	0,2793	0,2767	0,2741	0,2715	0,2689	0,2663	0,2637	0,2611	0,2585	0,2559	0,2533	0,60
0,40	0,2533	0,2508	0,2482	0,2456	0,2430	0,2404	0,2378	0,2353	0,2327	0,2301	0,2275	0,59
0,41	0,2275	0,2250	0,2224	0,2198	0,2173	0,2147	0,2121	0,2096	0,2070	0,2045	0,2019	0,58
0,42	0,2019	0,1993	0,1968	0,1942	0,1917	0,1891	0,1866	0,1840	0,1815	0,1789	0,1764	0,57
0,43	0,1764	0,1738	0,1713	0,1687	0,1662	0,1637	0,1611	0,1586	0,1560	0,1535	0,1510	0,56
0,44	0,1510	0,1484	0,1459	0,1434	0,1408	0,1383	0,1358	0,1332	0,1307	0,1282	0,1257	0,55
0,45	0,1257	0,1231	0,1206	0,1181	0,1156	0,1130	0,1105	0,1080	0,1055	0,1030	0,1004	0,54
0,46	0,1004	0,0979	0,0954	0,0929	0,0904	0,0878	0,0853	0,0828	0,0803	0,0778	0,0753	0,53
0,47	0,0753	0,0728	0,0702	0,0677	0,0652	0,0627	0,0602	0,0577	0,0552	0,0527	0,0502	0,52
0,48	0,0502	0,0476	0,0451	0,0426	0,0401	0,0376	0,0351	0,0326	0,0301	0,0276	0,0251	0,51
0,49	0,0251	0,0226	0,0201	0,0175	0,0150	0,0125	0,0100	0,0075	0,0050	0,0025	0,0000	0,50
	0,01	0,009	0,008	0,007	0,006	0,005	0,004	0,003	0,002	0,001	0	

Fractiles de la loi normale centrée réduite.

**Table 3 : Fractiles de la loi du chi-deux à  $\nu$  degrés de liberté :**

pour  $X \in \chi^2_{\nu}$  la table donne  $x = \chi^2_{\nu; \alpha}$  tel que  $\mathbf{P} ( X \leq x ) \approx \alpha$

$\nu \setminus \alpha$	0,001	0,01	0,025	0,05	0,1	0,2	0,3	0,4	0,5
1	$1,6 \cdot 10^{-6}$	$1,6 \cdot 10^{-4}$	$9,8 \cdot 10^{-4}$	$3,9 \cdot 10^{-3}$	$1,6 \cdot 10^{-2}$	$6,4 \cdot 10^{-2}$	0,15	0,27	0,45
2	$2,0 \cdot 10^{-3}$	$2,0 \cdot 10^{-2}$	$5,1 \cdot 10^{-2}$	0,10	0,21	0,45	0,71	1,02	1,39
3	$2,4 \cdot 10^{-2}$	0,11	0,22	0,35	0,58	1,01	1,42	1,87	2,37
4	$9,1 \cdot 10^{-2}$	0,30	0,48	0,71	1,06	1,65	2,19	2,75	3,36
5	0,21	0,55	0,83	1,15	1,61	2,34	3,00	3,66	4,35
6	0,38	0,87	1,24	1,64	2,20	3,07	3,83	4,57	5,35
7	0,60	1,24	1,69	2,17	2,83	3,82	4,67	5,49	6,35
8	0,86	1,65	2,18	2,73	3,49	4,59	5,53	6,42	7,34
9	1,15	2,09	2,70	3,33	4,17	5,38	6,39	7,36	8,34
10	1,48	2,56	3,25	3,94	4,87	6,18	7,27	8,30	9,34
11	1,83	3,05	3,82	4,57	5,58	6,99	8,15	9,24	10,34
12	2,21	3,57	4,40	5,23	6,30	7,81	9,03	10,18	11,34
13	2,62	4,11	5,01	5,89	7,04	8,63	9,93	11,13	12,34
14	3,04	4,66	5,63	6,57	7,79	9,47	10,82	12,08	13,34
15	3,48	5,23	6,26	7,26	8,55	10,31	11,72	13,03	14,34
16	3,94	5,81	6,91	7,96	9,31	11,15	12,62	13,98	15,34
17	4,42	6,41	7,56	8,67	10,09	12,00	13,53	14,94	16,34
18	4,90	7,01	8,23	9,39	10,86	12,86	14,44	15,89	17,34
19	5,41	7,63	8,91	10,12	11,65	13,72	15,35	16,85	18,34
20	5,92	8,26	9,59	10,85	12,44	14,58	16,27	17,81	19,34
21	6,45	8,90	10,28	11,59	13,24	15,44	17,18	18,77	20,34
22	6,98	9,54	10,98	12,34	14,04	16,31	18,10	19,73	21,34
23	7,53	10,20	11,69	13,09	14,85	17,19	19,02	20,69	22,34
24	8,08	10,86	12,40	13,85	15,66	18,06	19,94	21,65	23,34
25	8,65	11,52	13,12	14,61	16,47	18,94	20,87	22,62	24,34
26	9,22	12,20	13,84	15,38	17,29	19,82	21,79	23,58	25,34
27	9,80	12,88	14,57	16,15	18,11	20,70	22,72	24,54	26,34
28	10,39	13,56	15,31	16,93	18,94	21,59	23,65	25,51	27,34
29	10,99	14,26	16,05	17,71	19,77	22,48	24,58	26,48	28,34
30	11,59	14,95	16,79	18,49	20,60	23,36	25,51	27,44	29,34
35	14,69	18,51	20,57	22,47	24,80	27,84	30,18	32,28	34,34
40	17,92	22,16	24,43	26,51	29,05	32,34	34,87	37,13	39,34
45	21,25	25,90	28,37	30,61	33,35	36,88	39,58	42,00	44,34
50	24,67	29,71	32,36	34,76	37,69	41,45	44,31	46,86	49,33
55	28,17	33,57	36,40	38,96	42,06	46,04	49,06	51,74	54,33
60	31,74	37,48	40,48	43,19	46,46	50,64	53,81	56,62	59,33
65	35,36	41,44	44,60	47,45	50,88	55,26	58,57	61,51	64,33
70	39,04	45,44	48,76	51,74	55,33	59,90	63,35	66,40	69,33
75	42,76	49,48	52,94	56,05	59,79	64,55	68,13	71,29	74,33
80	46,52	53,54	57,15	60,39	64,28	69,21	72,92	76,19	79,33
85	50,32	57,63	61,39	64,75	68,78	73,88	77,71	81,09	84,33
90	54,16	61,75	65,65	69,13	73,29	78,56	82,51	85,99	89,33
95	58,02	65,90	69,92	73,52	77,82	83,25	87,32	90,90	94,33
100	61,92	70,06	74,22	77,93	82,36	87,95	92,13	95,81	99,33

**Table 4 : Fractiles de la loi du chi-deux à v degrés de liberté**

P	0,5	0,6	0,7	0,8	0,9	0,95	0,975	0,99	0,999
v									
1	0,45	0,71	1,07	1,64	2,71	3,84	5,02	6,63	10,83
2	1,39	1,83	2,41	3,22	4,61	5,99	7,38	9,21	13,82
3	2,37	2,95	3,66	4,64	6,25	7,81	9,35	11,34	16,27
4	3,36	4,04	4,88	5,99	7,78	9,49	11,14	13,28	18,47
5	4,35	5,13	6,06	7,29	9,24	11,07	12,83	15,09	20,51
6	5,35	6,21	7,23	8,56	10,64	12,59	14,45	16,81	22,46
7	6,35	7,28	8,38	9,80	12,02	14,07	16,01	18,48	24,32
8	7,34	8,35	9,52	11,03	13,36	15,51	17,53	20,09	26,12
9	8,34	9,41	10,66	12,24	14,68	16,92	19,02	21,67	27,88
10	9,34	10,47	11,78	13,44	15,99	18,31	20,48	23,21	29,59
11	10,34	11,53	12,90	14,63	17,28	19,68	21,92	24,73	31,26
12	11,34	12,58	14,01	15,81	18,55	21,03	23,34	26,22	32,91
13	12,34	13,64	15,12	16,98	19,81	22,36	24,74	27,69	34,53
14	13,34	14,69	16,22	18,15	21,06	23,68	26,12	29,14	36,12
15	14,34	15,73	17,32	19,31	22,31	25,00	27,49	30,58	37,70
16	15,34	16,78	18,42	20,47	23,54	26,30	28,85	32,00	39,25
17	16,34	17,82	19,51	21,61	24,77	27,59	30,19	33,41	40,79
18	17,34	18,87	20,60	22,76	25,99	28,87	31,53	34,81	42,31
19	18,34	19,91	21,69	23,90	27,20	30,14	32,85	36,19	43,82
20	19,34	20,95	22,77	25,04	28,41	31,41	34,17	37,57	45,31
21	20,34	21,99	23,86	26,17	29,62	32,67	35,48	38,93	46,80
22	21,34	23,03	24,94	27,30	30,81	33,92	36,78	40,29	48,27
23	22,34	24,07	26,02	28,43	32,01	35,17	38,08	41,64	49,73
24	23,34	25,11	27,10	29,55	33,20	36,42	39,36	42,98	51,18
25	24,34	26,14	28,17	30,68	34,38	37,65	40,65	44,31	52,62
26	25,34	27,18	29,25	31,79	35,56	38,89	41,92	45,64	54,05
27	26,34	28,21	30,32	32,91	36,74	40,11	43,19	46,96	55,48
28	27,34	29,25	31,39	34,03	37,92	41,34	44,46	48,28	56,89
29	28,34	30,28	32,46	35,14	39,09	42,56	45,72	49,59	58,30
30	29,34	31,32	33,53	36,25	40,26	43,77	46,98	50,89	59,70
35	34,34	36,47	38,86	41,78	46,06	49,80	53,20	57,34	66,62
40	39,34	41,62	44,16	47,27	51,81	55,76	59,34	63,69	73,40
45	44,34	46,76	49,45	52,73	57,51	61,66	65,41	69,96	80,08
50	49,33	51,89	54,72	58,16	63,17	67,50	71,42	76,15	86,66
55	54,33	57,02	59,98	63,58	68,80	73,31	77,38	82,29	93,17
60	59,33	62,13	65,23	68,97	74,40	79,08	83,30	88,38	99,61
65	64,33	67,25	70,46	74,35	79,97	84,82	89,18	94,42	106,0
70	69,33	72,36	75,69	79,71	85,53	90,53	95,02	100,4	112,3
75	74,33	77,46	80,91	85,07	91,06	96,22	100,8	106,4	118,6
80	79,33	82,57	86,12	90,41	96,58	101,9	106,6	112,3	124,8
85	84,33	87,67	91,32	95,73	102,1	107,5	112,4	118,2	131,0
90	89,33	92,76	96,52	101,1	107,6	113,1	118,1	124,1	137,2
95	94,33	97,85	101,7	106,4	113,0	118,8	123,9	130,0	143,3
100	99,33	102,9	106,9	111,7	118,5	124,3	129,6	135,8	149,4

**Table 5 : Fractiles de la loi de Student  $T$  à  $n$  degrés de liberté**

pour  $X \in \mathcal{T}(n)$ , la table donne  $x = t_\alpha$  tel que  $\mathbf{P} ( X \leq x ) \approx \alpha$

$\mathbf{P} (  X  \geq t_\alpha )$	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
$\alpha$	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	0,95	0,975	0,99	0,995	0,9995
$n$													
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,656	636,57
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,600
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,768
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,689
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,660
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,526	0,678	0,846	1,043	1,292	1,664	1,990	2,374	2,639	3,416
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
$\infty$	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,290

**Table 6 : Fractiles d'ordre 0,95 de la loi de Fisher à  $v_1$  et  $v_2$  degrés de liberté :**  
pour  $X \in \mathcal{F}(v_1, v_2)$ , la table donne  $x = f_{v_1, v_2; 0,95}$  tel que  $P ( X \leq x ) \approx 0,95$

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	26	30	40	60	120	$\infty$
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.76	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.10	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.72	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.59	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.49	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.33	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.27	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.22	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.17	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.13	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.10	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.04	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.01	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	1.99	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.95	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.93	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.91	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.90	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.88	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.87	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.77	1.74	1.69	1.64	1.58	1.51
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.95	1.87	1.78	1.72	1.69	1.63	1.58	1.51	1.44
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.68	1.65	1.59	1.53	1.47	1.39
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97	1.89	1.81	1.72	1.65	1.62	1.57	1.50	1.44	1.35
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.88	1.79	1.70	1.63	1.60	1.54	1.48	1.41	1.33
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94	1.86	1.78	1.69	1.62	1.59	1.53	1.46	1.39	1.30
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.85	1.77	1.68	1.61	1.57	1.51	1.45	1.38	1.28
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.59	1.55	1.50	1.43	1.35	1.25
$\infty$	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.50	1.46	1.39	1.32	1.22	1.00



**Table 7 : Fractiles d'ordre 0.975 de la loi de Fisher à  $v_1$  et  $v_2$  degrés de liberté**  
pour  $X \in \mathcal{F}(v_1, v_2)$ , la table donne  $x = f_{v_1; v_2; 0,975}$  tel que  $\mathbf{P} ( X \leq x ) \approx 0,975$

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	26	30	40	60	120	$\infty$
1	648	799	864	900	922	937	948	957	963	969	977	985	993	999	<i>100</i>	<i>101</i>	<i>101</i>	<i>101</i>	<i>102</i>
2	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5	39.5	39.5
3	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.2	14.2	14.1	14.1	14.0	14.0	13.9	13.9
4	12.2	10.7	9.98	9.60	9.36	9.20	9.07	8.97	8.90	8.84	8.75	8.66	8.56	8.49	8.46	8.41	8.36	8.31	8.26
5	10.0	8.43	7.76	7.38	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.26	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.10	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.39	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.93	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.59	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.34	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.15	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.00	2.96	2.91	2.85	2.79	2.73
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.87	2.83	2.78	2.72	2.66	2.57
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.77	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.81	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.68	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.60	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.54	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.48	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.43	2.39	2.33	2.27	2.30	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.39	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.35	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.31	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.28	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.25	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.84	2.75	2.68	2.61	2.51	2.41	2.30	2.22	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.19	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.17	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.15	2.11	2.05	1.98	1.91	1.83
29	5.79	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.13	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.11	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	1.98	1.94	1.88	1.80	1.72	1.64
50	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38	2.32	2.22	2.11	1.99	1.91	1.87	1.80	1.72	1.64	1.55
60	5.29	3.93	3.34	3.01	2.77	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.86	1.82	1.74	1.67	1.58	1.48
70	5.25	3.89	3.31	2.97	2.75	2.59	2.47	2.38	2.30	2.24	2.14	2.03	1.91	1.82	1.78	1.71	1.63	1.54	1.44
80	5.22	3.86	3.28	2.95	2.73	2.57	2.45	2.35	2.28	2.21	2.11	2.00	1.88	1.79	1.75	1.68	1.60	1.51	1.40
90	5.20	3.84	3.26	2.93	2.71	2.55	2.43	2.34	2.26	2.19	2.09	1.98	1.86	1.77	1.73	1.66	1.58	1.48	1.37
100	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24	2.18	2.08	1.97	1.85	1.76	1.71	1.64	1.56	1.46	1.35
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.95	1.82	1.73	1.69	1.61	1.53	1.43	1.31
$\infty$	5.03	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.95	1.83	1.71	1.61	1.57	1.49	1.39	1.27	1.04

*Les nombres en italique doivent être multipliés par 10 .*