

Maths Appro

# Vademecum

## Maths 2

Programme 2021

[spicesagros.fr](http://spicesagros.fr)

01 septembre 2025

# Table des matières

Chapitre 1. <b>Légendes et abréviations</b>	5
Chapitre 2. <b>Quelques notions sur les ensembles</b>	7
1. Définition intuitive d'un ensemble	7
2. Ensemble des parties d'un ensemble	7
3. Union d'ensembles	7
4. Intersection d'ensembles	7
5. Distributivités	8
6. Complémentation et lois de DE MORGAN	8
7. Inclusion d'ensembles	8
8. Différence d'ensembles	9
9. Produit cartésien d'ensembles	9
10. Image directe et image réciproque d'une partie par une application	9
11. Complément hors programme	10
Chapitre 3. <b>Éléments de dénombrement</b>	11
1. Propriétés du cardinal d'un ensemble fini	11
2. Deux problèmes	12
3. Des cardinaux de listes particulières particuliers	13
4. Exemples fondamentaux de tirages usuels	13
5. Formules combinatoires	13
6. Compléments hors programme	14
Chapitre 4. <b>Espaces probabilisés</b>	15
1. Ensemble des événements	15
2. Correspondance entre relations logiques et ensemblistes	15
3. Correspondance entre le langage des événements et celui des probabilités	16
4. Probabilité	16
Chapitre 5. <b>Variables aléatoires réelles</b>	21
1. Généralités	21
2. Variables aléatoires particulières	24
3. Moments	25
Chapitre 6. <b>Vecteurs aléatoires discrets</b>	29
1. Famille sommables	29
2. Couples discrets	30
3. Vecteurs discrets de dimension $n \geq 2$	31
4. Espérance conditionnelle	32
5. Covariance	33
6. Indépendance de variables aléatoires	36
7. Produit de convolution	37
8. Droites de régression	38
9. Complément hors programme	39
Chapitre 7. <b>Variables aléatoires extrêmes</b>	41
1. Panorama en dimension 2	41

2. Généralisation à la dimension $n$	42
<b>Chapitre 8. Inégalités probabilistes</b>	45
1. Inégalités de concentration	45
2. Inégalités de moments	46
<b>Chapitre 9. Convergences</b>	49
1. Convergence en probabilité	49
2. Convergence en loi	50
3. Liens entre ces deux types de convergence	51
<b>Chapitre 10. Approximations</b>	53
<b>Chapitre 11. Estimation inférentielle</b>	55
1. Estimation ponctuelle	55
2. Compléments hors programmes	57
3. Estimation par intervalle de confiance	57
<b>Chapitre 12. Zoologie des lois usuelles du programme</b>	61
1. Loi de BERNOULLI	61
2. Variable aléatoire indicatrice	61
3. Loi binomiale	62
4. Loi de DIRAC (ou loi certaine)	62
5. Loi exponentielle	63
6. Loi gamma à un seul paramètre	63
7. Loi géométrique	63
8. Lois normales	64
9. Loi de POISSON	65
10. Lois uniformes	65
11. Bilan des lois du programme stables par convolution	66
12. Bilan des lois sans mémoire	66
<b>Chapitre 13. Outils</b>	67
1. Les identités basiques	67
2. Quelques sommes que l'on peut retrouver dans les sujets	68
3. Calcul intégral	68
4. Les produits finis	68
5. Deux équivalents simples	70
6. Résultats divers	70
<b>Chapitre 14. Compléments hors programme mais ...</b>	71
1. Formule de VANDERMONDE	71
2. Le coefficient multinomial	71
3. Relation min et max	71
4. Urne multicolore	71
5. Espérance et fonction de survie	72
6. Variables aléatoires de signe fixé et indicatrices	72
7. Interprétation géométrique de la corrélation	72
8. Minimisation de la variance	72
9. Espérance conditionnelle	73
10. La loi gamma à deux paramètres	73
11. Loi exponentielle	74
12. Variables aléatoires symétriques	74
13. Loi géométrique	74
14. Loi normale et loi du chi-deux	75
15. Produit scalaire et norme associée dans $L^2(\Omega, \mathcal{A}, \mathbf{P})$	75





## Légendes et abréviations

- ▶  $\textcircled{C}$  : résultat valable en continu
- ▶  $\textcircled{S}$  : résultat valable en discret
- ▶  $\textcircled{S}\textcircled{C}$  : résultat valable en continu et en discret
- ▶  $\sqcup$  : union disjointe
- ▶ *i.i.d.* : variables aléatoires indépendantes et identiquement distribuées
- ▶  $\sum_i u_i < \infty$  : série simple convergente (symbole **hors programme** à ne surtout pas utiliser)
- ▶  $\sum_i u_i = \infty$  : série simple divergente (symbole **hors programme** à ne surtout pas utiliser)
- ▶  $\sum_{i,j} u_{i,j} < \infty$  : : série double convergente (symbole **hors programme** à ne surtout pas utiliser)
- ▶  $\sum_{i,j} u_{i,j} = \infty$  : série double divergente (symbole **hors programme** à ne surtout pas utiliser)
- ▶  $\int_a^b f < \infty$  : intégrale convergente (symbole **hors programme** à ne surtout pas utiliser)
- ▶  $\int_a^b f = \infty$  : intégrale divergente (symbole **hors programme** à ne surtout pas utiliser)
- ▶  $A \perp B$  :  $A$  et  $B$  indépendants (le symbole **hors programme** à ne surtout pas utiliser)
- ▶  $X \perp Y$  :  $X$  et  $Y$  indépendantes (le symbole **hors programme** à ne surtout pas utiliser)
- ▶  $X_1 \perp \dots \perp X_n$  :  $X_1, \dots, X_n$  mutuellement indépendantes (le symbole **hors programme** à ne surtout pas utiliser)
- ▶  $\forall k \geq 1, L^k(\Omega, \mathcal{A}, \mathbf{P})$  désigne l'espace vectoriel des variables aléatoires admettant un moment d'ordre  $k$  (notation de l'ensemble **hors programme** à ne surtout pas utiliser)
- ▶  $\forall k \geq 1, L_d^k(\Omega, \mathcal{A}, \mathbf{P})$  désigne l'espace vectoriel des variables aléatoires discrète admettant un moment d'ordre  $k$  (notation de l'ensemble **hors programme** à ne surtout pas utiliser)
- ▶  $P_X$  : loi de  $X$
- ▶  $X \sim Y$  : les variables aléatoires  $X$  et  $Y$  suivent la même loi
- ▶  $\mathcal{S}_X$  support de  $X$



## Quelques notions sur les ensembles

### 1. Définition intuitive d'un ensemble

**Définition 2.1.** C'est le tout formant une collection d'objets distincts, non rangés, et ayant une propriété commune. Chaque objet d'un ensemble est appelé élément, et écrire que  $x$  est un élément de  $A$  se note  $x \in A$ . Dans le cas contraire, on note  $x \notin A$ .

On dit que  $A = B$  lorsque pour tout  $x \in A$ ,  $x \in B$  et pour tout  $x \in B$ ,  $x \in A$ . Par contraposée  $A \neq B$  lorsque il existe  $x \in A$ ,  $x \notin B$  ou il existe  $x \in B$ ,  $x \notin A$ .

**Remarque 2.1.** L'ordre de l'écriture des éléments d'un ensemble n'a aucune importance, ainsi que leurs répétitions éventuelles. Par exemple :  $\{1, 2, 3\} = \{2, 3, 1\} = \{1, 1, 2, 3, 3, 3\}$ .

### 2. Ensemble des parties d'un ensemble

**Définition 2.2.** On appelle ensemble des parties de  $A$ , l'ensemble de tous les sous-ensembles de  $A$ , on le note  $\mathcal{P}(A)$ , il contient évidemment  $A$  lui-même ainsi que l'ensemble vide.

**Remarque 2.2.** ► Ecrire  $A \subset \Omega$  se traduit par  $A \in \mathcal{P}(\Omega)$  (**Attention** on n'écrit jamais  $A \in \Omega$ !).

► Ecrire  $a \in A$  se traduit par  $\{a\} \subset A$  (**Attention** on n'écrit jamais  $a \subset A$ !).

### 3. Union d'ensembles

**Définition 2.3.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide,

$$A \cup B = \{x \in \Omega, x \in A \text{ ou } x \in B\}$$

Plus généralement

$$\bigcup_{i \in I} A_i = \{x \in \Omega \mid \exists i \in I, x \in A_i\}$$

**Définition 2.4. Union disjointe**

Si  $A \cap B = \emptyset$  on dit que les ensembles  $A$  et  $B$  sont disjoints et  $A \cup B$  se note  $A \sqcup B$ .

**Propriétés 2.1.** Soit  $A$ ,  $B$  et  $C$  deux parties d'un ensemble  $\Omega$  non vide,

- $A \cap \overline{A} = \emptyset$
- $A \cup \overline{A} = \Omega$
- $A \cup B = B \cup A$  (commutativité)
- $A \cup (B \cup C) = (A \cup B) \cup C$  (associativité)
- $A \subset A \cup B$ ,  $B \subset A \cup B$
- $(A \subset B) \Leftrightarrow (A \cup B = B)$
- $\emptyset \cup A = A$  ( $\emptyset$  est neutre pour  $\cup$ )
- $\Omega \cup A = \Omega$  ( $\Omega$  est absorbant pour  $\cup$ )
- $A \cup A = A$  (réflexivité)

### 4. Intersection d'ensembles

**Définition 2.5.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide,

$$A \cap B = \{x \in \Omega \mid x \in A \text{ et } x \in B\}$$

Plus généralement

$$\bigcap_{i \in I} A_i = \{x \in \Omega \mid \forall i \in I, x \in A_i\}$$

**Proposition 2.1.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide.

- ▶  $A \cap \bar{A} = \emptyset$
- ▶  $A \cap B = B \cap A$  (commutativité)
- ▶  $A \cap (B \cap C) = (A \cap B) \cap C$  (associativité)
- ▶  $A \cap B \subset A, A \cap B \subset B$
- ▶  $(A \subset B) \Leftrightarrow (A \cap B = A)$
- ▶  $\emptyset \cap A = \emptyset$  ( $\emptyset$  est absorbant pour  $\cap$ )
- ▶  $\Omega \cap A = A$  ( $\Omega$  est neutre pour  $\cap$ )
- ▶  $A \cap A = A$  (réflexivité)

### 5. Distributivités

**Propriétés 2.2.** ▶  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$  (distributivité de  $\cap$  par rapport à  $\cup$ )

▶  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$  (distributivité de  $\cup$  par rapport à  $\cap$ )

▶  $\left( \bigcap_{i \in I} A_i \right) \cup B = \bigcap_{i \in I} (A_i \cup B)$

▶  $\left( \bigcup_{i \in I} A_i \right) \cap B = \bigcup_{i \in I} (A_i \cap B)$

▶  $\left( \bigcap_{i \in I} A_i \right) \cup \left( \bigcap_{j \in J} B_j \right) = \bigcap_{(i,j) \in I \times J} (A_i \cup B_j)$

▶  $\left( \bigcup_{i \in I} A_i \right) \cap \left( \bigcup_{j \in J} B_j \right) = \bigcup_{(i,j) \in I \times J} (A_i \cap B_j)$

### 6. Complémentation et lois de De Morgan

On notera  $\bar{A} = \complement_{\Omega} A$  (complémentaire de  $A$  dans  $\Omega$ ).

**Propriétés 2.3.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide.

- ▶  $\overline{\bar{\Omega}} = \emptyset$
- ▶  $\overline{\bar{\emptyset}} = \Omega$
- ▶  $\overline{\bar{A}} = A$
- ▶  $(B = \bar{A}) \Leftrightarrow (A = \bar{B})$

**Propriétés 2.4.** ▶  $\overline{A \cap B} = \bar{A} \cup \bar{B}$

▶  $\overline{A \cup B} = \bar{A} \cap \bar{B}$

▶  $\bigcup_{i \in I} A_i = \bigcap_{i \in I} \bar{A}_i$  et  $\bigcap_{i \in I} A_i = \overline{\bigcup_{i \in I} \bar{A}_i}$  (Lois de De Morgan généralisées)

### 7. Inclusion d'ensembles

#### Définition 2.6. Inclusion

Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$ , on dit que  $B$  est un sous-ensemble de  $A$  ou que  $B$  est une partie de  $A$  si tout élément de  $B$  est aussi un élément de  $A$ . On dit aussi que  $B$  est inclus dans  $A$  ce que l'on note  $B \subset A$ . Ainsi  $B \subset A$  lorsque

$$\forall x \in B, \quad x \in A$$

Autrement dit,  $B \subset A$  si

$$\forall x \in \Omega, \quad (x \in B \Rightarrow x \in A)$$

**Notation 2.1.** Le symbole  $\subset$  désigne une inclusion au sens large. Il est parfois noté  $\subseteq$ .

**Propriétés 2.5.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide.

- ▶  $(A \subset B) \Leftrightarrow (A \cap B = A)$
- ▶  $(A \subset B) \Leftrightarrow (A \cup B = B)$
- ▶  $A \cap B \subset A$  et  $A \cap B \subset B$
- ▶  $A \subset A \cup B$  et  $B \subset A \cup B$

- ▶  $(A \subset B \text{ et } B \subset C) \Rightarrow (A \subset C)$  (transitivité de  $\subset$ )
- ▶  $(A = B) \Leftrightarrow (A \subset B \text{ et } B \subset A)$
- ▶  $(A \subsetneq B) \Leftrightarrow (A \subset B \text{ et } A \neq B)$
- ▶  $(A \subset B) \Leftrightarrow (A \in \mathcal{P}(B))$
- ▶  $(A \subset B) \Leftrightarrow (\overline{B} \subset \overline{A})$

### 8. Différence d'ensembles

**Définition 2.7.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide, la différence de  $A$  et  $B$  est un ensemble noté  $A - B$  défini par

$$A - B = \{x \in \Omega \mid x \in A \text{ et } x \notin B\}$$

L'ensemble  $A - B$  est donc celui des éléments de  $A$  privé de ceux de  $B$ .

**Notation 2.2.** La différence de  $A$  et  $B$  peut être aussi notée  $A \setminus B$ .

**Propriétés 2.6.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide.

- ▶  $A - B = A \cap \overline{B}$

**Définition 2.8.**

**Propriétés 2.7.** ▶  $A - B = A - (A \cap B)$

- ▶  $\overline{A} = \Omega - A$

- ▶  $((A - B) = A) \Leftrightarrow (A \cap B = \emptyset)$

### 9. Produit cartésien d'ensembles

**Définition 2.9.** Soit  $A$  et  $B$  deux parties d'un ensemble  $\Omega$  non vide, on appelle

$$A \times B = \{(a, b) \in \Omega^2 \mid a \in A \text{ et } b \in B\}$$

Plus généralement,

$$A_1 \times \cdots \times A_n = \{(x_1, \dots, x_n) \in \Omega^n \mid x_1 \in A_1, \dots, x_n \in A_n\}$$

**Remarque 2.3.** ▶  $A \times B$  se lit "A croix B".

- ▶ Le produit cartésien  $\underbrace{A \times \cdots \times A}_{n \text{ fois}}$  se note  $A^n$ .

## 10. Image directe et image réciproque d'une partie par une application

**Définition 2.10. Image directe**

Soit  $f : E \rightarrow F$  une application et soit  $A$  une partie de  $E$ .

L'image directe de  $A$  par  $f$  est l'ensemble noté  $f(A)$  constitué des images par  $f$  des éléments de  $A$ , soit

$$f(A) = \{y \in F \mid \exists a \in A, y = f(a)\}$$

**Notation 2.3.** Un élément de  $f(A)$  s'écrit  $f(a)$  où  $a \in A$ .

On pourra prouver la proposition suivante à titre d'exercice :

**Proposition 2.2.** Soit  $f : E \rightarrow F$  une application et soit  $A$  et  $B$  deux parties de  $E$ , alors on a

- ▶  $(A \subset B) \Rightarrow (f(A) \subset f(B))$
- ▶  $f(A \cup B) = f(A) \cup f(B)$
- ▶  $f(A \cap B) \subset f(A) \cap f(B)$

**Définition 2.11. Image réciproque**

Soit  $f : E \rightarrow F$  une application et soit  $B$  une partie de  $F$ .

L'image réciproque de  $B$  par  $f$  est l'ensemble noté  $f^{-1}(B)$  constitué des éléments de  $E$  dont l'image appartient à  $B$ , soit

$$f^{-1}(B) = \{x \in E \mid f(x) \in B\}$$

C'est donc l'ensemble des antécédents des éléments de  $B$  par l'application  $f$ .

**Notation 2.4.** La notation de l'image réciproque d'une partie utilisant  $f^{-1}$  peut prêter à confusion et faire croire que  $f$  est supposée bijective. Or il n'en est rien et c'est pourquoi il n'est pas rare de voir utiliser les notations  $f^{-1}$  voire  $f^{(-1)}$  ou même  $f^{\leftarrow}$  que nous n'utiliserons pas dans cet ouvrage. Au passage on pourra remarquer que la notation  $f^{-1}$  privée de la partie concernée n'a aucun sens.

**Remarque 2.4.** Écrire qu'un élément  $x$  appartient à  $f^{-1}(B)$  revient à écrire que  $f(x) \in B$ .

On pourra prouver la proposition suivante à titre d'exercice :

**Proposition 2.3.** Soit  $f : E \rightarrow F$  une application et soit  $A$  et  $B$  deux parties de  $E$ , alors on a

- ▶  $(A \subset B) \Rightarrow (f^{-1}(A) \subset f^{-1}(B))$
- ▶  $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$
- ▶  $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$

## 11. Complément hors programme

**Définition 2.12. Différence symétrique d'ensembles**

$$A \Delta B = (A \cup B) - (A \cap B) = (A - B) \cup (B - A)$$

**Définition 2.13. Liminf et Limsup d'ensembles**

Soit une suite  $(A_n)_{n \geq 0}$  de parties d'un ensemble  $E$ . On pose

$$\liminf_{n \rightarrow +\infty} (A_n) = \bigcup_{n=0}^{+\infty} \bigcap_{k=n}^{+\infty} A_k \quad \text{et} \quad \limsup_{n \rightarrow +\infty} (A_n) = \bigcap_{n=0}^{+\infty} \bigcup_{k=n}^{+\infty} A_k$$

**Propriétés 2.8.**  $\bigcap_{n=0}^{+\infty} \bigcup_{k=n}^{+\infty} A_k$  signifie "les événements  $A_n$  sont réalisés une infinité de fois".

**Propriétés 2.9.**  $\bigcup_{n=0}^{+\infty} \bigcap_{k=n}^{+\infty} A_k$  signifie "tous les événements  $A_n$  sont réalisés à partir d'un certain rang".

**Propriétés 2.10.** ▶  $\overline{\limsup_{n \rightarrow +\infty} (A_n)} = \liminf_{n \rightarrow +\infty} (\overline{A_n})$

- ▶  $\liminf_{n \rightarrow +\infty} (A_n) \subset \limsup_{n \rightarrow +\infty} (A_n)$
- ▶  $\limsup_{n \rightarrow +\infty} (A_n) = \left\{ \sum_{n \geq 0} \mathbf{1}_{A_n} = \infty \right\}$
- ▶  $\liminf_{n \rightarrow +\infty} (A_n) = \left\{ \sum_{n \geq 0} \mathbf{1}_{A_n} < \infty \right\}$

## Éléments de dénombrement

**Définition 3.1. (du Larousse)** Faire le compte des unités composant un ensemble.

**Définition 3.2.** Soit  $E$  un ensemble.

- ▶ On dit que  $E$  est de cardinal  $n \in \mathbf{N}$  (ou bien que  $E$  possède  $n$  éléments) lorsqu'il existe une bijection de  $\llbracket 1, n \rrbracket$  dans  $E$ .
- ▶ On dit que  $E$  est fini lorsqu'il existe un entier naturel  $n$  tel que  $E$  soit de cardinal  $n$ .
- ▶ On dit que  $E$  est un ensemble dénombrable lorsqu'il existe une bijection  $\varphi : \mathbf{N} \rightarrow E$ . En d'autres termes on peut numéroter tous les éléments de  $E$  par des entiers, c'est à dire que l'on peut écrire  $E = \{i_n, n \in \mathbf{N}\}$  (où  $i_n = \varphi(n)$ ).
- ▶ On dit d'un ensemble qu'il est au plus dénombrable pour dire qu'il est fini ou dénombrable.

**Exemple 3.1. (à connaître par coeur)**

Les ensembles  $\mathbf{N}$ ,  $\mathbf{Z}$ ,  $\mathbf{Q}$  et  $\mathbf{N}^2$  sont dénombrables (il y en a d'autres), en revanche  $\mathbf{R}$ ,  $\mathbf{C}$ ,  $\mathcal{P}(\mathbf{N})$ ,  $[0, 1]$ , plus généralement aucun intervalle de  $\mathbf{R}$  d'intérieur non vide ne le sont pas etc ...

### 1. Propriétés du cardinal d'un ensemble fini

**Propriétés 3.1.** ▶ Soit  $A$  et  $B$  deux ensembles tels que  $A$  est fini et  $B$  est fini et  $B \subset A$ , alors

$$\text{Card}(B) \leq \text{Card}(A) \quad \text{et} \quad \text{Card}(A - B) = \text{Card}(A) - \text{Card}(B)$$

- ▶ Si  $A \subset B$  et  $\text{Card}(A) = \text{Card}(B)$  alors  $A = B$ .
- ▶  $\text{Card}(A \sqcup B) = \text{Card}(A) + \text{Card}(B)$
- ▶ Plus généralement, pour tout  $n \in \mathbf{N}^*$ ,

$$\text{Card}\left(\bigsqcup_{k=1}^n A_k\right) = \sum_{k=1}^n \text{Card}(A_k) \quad (\text{principe additif})$$

- ▶  $\text{Card}(A \times B) = \text{Card}(A) \text{Card}(B)$
- ▶ Plus généralement, pour tout  $n \in \mathbf{N}^*$ ,

$$\text{Card}(A_1 \times \cdots \times A_n) = \prod_{k=1}^n \text{Card}(A_k) \quad (\text{principe multiplicatif})$$

**Remarque 3.1.** On utilise le principe multiplicatif pour dénombrer le nombre d'issues que peut fournir une expérience aléatoire  $\mathcal{E}$  se décomposant à l'aide d'une succession de  $n$  sous-expériences  $\mathcal{E}_k$  avec  $k \in \llbracket 1, n \rrbracket$  amenant chacune  $n_k$  issues. Alors le nombre total des issues de  $\mathcal{E}$  est égal à  $\prod_{k=1}^n n_k$ .

**Théorème 3.1. Formule de POINCARÉ ou du crible pour deux ensembles**

$$\forall n \in \mathbf{N}^*, \quad \text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$$

**Remarque 3.2.** La généralisation de cette formule a disparu du programme.

**Définition 3.3. Partition d'un ensemble**

On dit que la famille d'événements  $(A_i)_{i \in I}$  est une partition de  $E$  lorsque :

- ▶  $\forall i \in I, A_i = \emptyset$ ,
- ▶  $(\forall (i, j) \in I^2, i \neq j) \Rightarrow (A_i \cap A_j = \emptyset)$  et
- ▶  $\bigsqcup_{i \in I} A_i = E$ .

**Remarque 3.3.** C'est la même définition pour un système complet d'événements avec  $E = \Omega$ .

**Proposition 3.1. Lemme des bergers**

Soit  $E$  et  $F$  deux ensembles finis et  $f$  une application surjective de  $E$  vers  $F$ . On suppose qu'il existe  $p \in \mathbb{N}^*$  tel que chaque élément  $y \in F$  possède  $p$  antécédents par  $f$ , alors

$$\text{Card}(E) = p \text{Card}(F)$$

**Proposition 3.2. Nombre d'arrangements de  $E_p$  vers  $F_n$**

Soit  $E_p$  et  $F_n$  deux ensembles avec  $\text{Card}(E_p) = p$ , et  $\text{Card}(F_n) = n$  et  $p \leq n$ , alors le nombre d'arrangements entre  $E_p$  et  $F_n$  est égal à

$$n(n-1) \times \cdots \times (n-p+1)$$

**Proposition 3.3.** Le nombre de permutations de  $E_n$  avec  $\text{Card}(E_n) = n \in \mathbb{N}^*$  est égal à  $n!$

**Proposition 3.4.** Le nombre de  $p$ -listes de  $E_n$  est égal à  $n^p$ .

**Proposition 3.5.** Le nombre de combinaisons de  $p$  éléments de  $E$  de cardinal  $n$  avec  $p \leq n$  est égal à

$$\binom{n}{p} = \frac{n(n-1) \times \cdots \times (n-p+1)}{p!}$$

**Remarque 3.4.** Voir les propriétés des coefficients binomiaux à la section 6 de ce chapitre.

**Proposition 3.6. Nombre de parties d'un ensemble  $E$**

$$\text{Card}(\mathcal{P}(E)) = 2^{\text{Card}(E)}$$

## 2. Deux problèmes

**Proposition 3.7. Le problème des anagrammes**

Soit un mot de  $n$  lettres constitué de  $n_1$  lettres  $L_1, \dots, n_p$  lettres  $L_p$  avec  $n_1 + \cdots + n_p = n$  alors le nombre d'anagrammes de ce mot est égal à

$$\binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \cdots \times \binom{n-n_1-\cdots-n_{p-1}}{n_p} = \frac{n!}{n_1! \times \cdots \times n_p!}$$

(nombre de permutations avec répétition)

**Remarque 3.5.** Ce résultat, hors programme, devra être redémontré à chaque emploi.

**Proposition 3.8. Le problème des chemins monotones dans  $\mathbb{N}^2$**

Un point mobile se déplace sur un quadrillage en reliant son origine  $O$  à un point  $M$  de ce quadrillage, de coordonnées entières  $(x, y)$ . Le point mobile est contraint de se déplacer à pas de longueur 1, soit vers la droite, soit vers le haut. Un tel parcours est appelé un chemin monotone. Le nombre de tels chemins reliant le point  $O$  au point  $M$  est égal à  $\binom{x+y}{x}$ .

**Remarque 3.6.** Ce problème montre qu'un chemin monotone peut être assimilé à un anagramme d'un mot de deux lettres.

### 3. Des cardinaux de listes particulières particuliers

Les résultats suivants ne sont pas au programme et sont à traiter à titre d'exercice.

**Proposition 3.9.** *Nombre de listes d'entiers sous contraintes d'égalité linéaire*

►  $\text{Card} \{(i, j) \in \mathbf{N}^2 \mid i + j = n\} = n + 1$

►  $\text{Card} \{(i, j) \in (\mathbf{N}^*)^2 \mid i + j = n\} = n - 1$

►  $\text{Card} \{(x_1, \dots, x_p) \in \mathbf{N}^p \mid x_1 + \dots + x_p = n\} = \binom{n+p-1}{p-1}$

►  $\text{Card} \{(x_1, \dots, x_p) \in (\mathbf{N}^*)^p \mid x_1 + \dots + x_p = n\} = \binom{n-1}{p-1}$  avec  $n \geq p$

**Remarque 3.7.** *Les deux premiers points de la proposition précédente interviennent dans le théorème de sommation par paquets d'une série double, lorsqu'on effectue une sommation selon des droites obliques.*

**Proposition 3.10.** *Nombre de  $p$ -listes strictement croissantes d'entiers (à savoir retrouver)*

$$\text{Card} \{(x_1, \dots, x_p) \in (\mathbf{N}^*)^p \mid 1 \leq x_1 < \dots < x_p \leq n\} = \binom{n}{p} \text{ (voir exercice ***)}$$

**Proposition 3.11.** *Nombre de  $p$ -listes croissantes d'entiers (à savoir retrouver)*

$$\text{Card} \{(x_1, \dots, x_p) \in (\mathbf{N}^*)^p \mid 1 \leq x_1 \leq \dots \leq x_p \leq n\} = \binom{n+p-1}{p} \text{ (voir exercice ***)}$$

### 4. Exemples fondamentaux de tirages usuels

Situation : une urne contient  $n$  boules numérotées de 1 à  $n$ , on y extrait  $p$  boules alors si

Type de tirage	Ordre	Répétitions d'éléments	Dénombrement
Successif avec remise	Oui	Oui	$n^p$ $p$ -listes
Successif sans remise	Oui	Non	$n(n-1) \times \dots \times (n-p+1)$ arrangements
Simultané	Non	Non	$\binom{n}{p}$ combinaisons

### 5. Formules combinatoires

**Définition 3.4.** 
$$\binom{n}{p} = \begin{cases} \frac{n!}{p!(n-p)!} & \text{si } 0 \leq p \leq n \\ 0 & \text{sinon} \end{cases}$$

**Propriétés 3.2.** ►  $\forall (n, p) \in \mathbf{N}^2, p \leq n, \binom{n}{p} = \binom{n}{n-p}$  (symétrie)

►  $\forall (n, p) \in \mathbf{N}^2, 1 \leq p \leq n, \binom{n}{p} = \frac{n}{p} \binom{n-1}{p-1}$

►  $\forall (n, p) \in \mathbf{N}^2, 1 \leq p \leq n, \binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}$  (triangle de PASCAL)

►  $\forall (p, d) \in \mathbf{N}^2, p \leq d, \sum_{k=p}^d \binom{k}{p} = \binom{d+1}{p+1}$  (triangle de PASCAL généralisée, hors programme)

►  $\forall n \in \mathbf{N}, \sum_{k=0}^n \binom{n}{k} = 2^n$

►  $\forall n \in \mathbf{N}^*, \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} = 2^{n-1} = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1}$  (hors programme)

►  $\forall n \in \mathbf{N}^*, \sum_{k=1}^n k \binom{n}{k} = n2^{n-1}$

**6. Compléments hors programme****Proposition 3.12. Nombre d'applications de  $E$  vers  $F$** 

Soit  $E$  et  $F$  deux ensembles et  $\mathcal{A}(E, F)$  l'ensemble des applications entre  $E$  et  $F$ , alors

$$\text{Card}(\mathcal{A}(E, F)) = (\text{Card}(F))^{\text{Card}(E)}$$

**Proposition 3.13. Nombre d'injections de  $E_p$  vers  $F_n$** 

Soit  $E_p$  et  $F_n$  deux ensembles et  $\mathcal{I}(E_p, F_n)$  l'ensemble des applications injectives entre  $E_p$  et  $F_n$ , avec  $\text{Card}(E_p) = p$ , et  $\text{Card}(F_n) = n$  et  $p \leq n$ , alors

$$\text{Card}(\mathcal{I}(E_p, F_n)) = n(n-1) \times \cdots \times (n-p+1)$$

## Espaces probabilisés

### 1. Ensemble des événements

Pour permettre l'analyse d'une expérience aléatoire on doit considérer, a priori, un ensemble  $\mathcal{A}$  d'événements qui seront identifiés avec la partie de l'univers  $\Omega$  dont les éléments (issues élémentaires de l'expérience) permettent de les réaliser. On pourrait naturellement penser à considérer  $\mathcal{P}(\Omega)$ , l'ensemble de toutes les parties de  $\Omega$  comme un ensemble d'événements, mais  $\mathcal{P}(\Omega)$  est souvent bien trop vaste comme par exemple lorsque  $\Omega = \mathbb{R}$  pour pouvoir l'affirmer! On imposera logiquement à cet ensemble  $\mathcal{A}$  des conditions de stabilité de façon à ce que les opérations ensemblistes usuelles ne fassent pas "sortir" de l'ensemble  $\mathcal{A}$  en question. Cet ensemble  $\mathcal{A}$  sera un ensemble de parties de  $\Omega$  contenant  $\Omega$ , stable par réunion finie et complémentation (on ne dit pas complémentarité!) ce qui induit aussi la stabilité pour l'intersection au plus dénombrable, la différence ainsi que la différence symétrique en utilisant les lois de Morgan. A savoir

**Définition 4.1. Ensemble des événements**

Soit  $\Omega \neq \emptyset$  et soit  $\mathcal{A} \subset \mathcal{P}(\Omega)$ . On dit que  $\mathcal{A}$  est un **ensemble d'événements** d'une expérience aléatoire<sup>1</sup> si  $\mathcal{A}$  contient  $\Omega$ , stable par réunion finie et complémentation. Autrement dit

- ▶  $\Omega \in \mathcal{A}$
- ▶  $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$
- ▶ Pour toute suite  $(A_n)_{n \in I}$  d'événements de  $\mathcal{A}$ ,  $\bigcup_{n \in I} A_n \in \mathcal{A}$ .

**Proposition 4.1.**  $\emptyset \in \mathcal{A}$ ,  $\mathcal{A}$  est stable pour  $\cap, -, \Delta$

**Exemple 4.1.** ▶  $\mathcal{A} = \{\emptyset, \Omega\}$  est un ensemble d'événement que l'on peut qualifier de grossier car inexploitable.

- ▶ Si  $\Omega$  est au plus dénombrable  $\mathcal{A} = \mathcal{P}(\Omega)$  est un ensemble d'événements, le plus fin possible.
- ▶ L'ensemble  $\{\emptyset, A, \bar{A}, \Omega\}$  constitue un ensemble d'événements qui est le plus petit possible engendré par  $A \in \mathcal{A}$ .

**Proposition 4.2. Quelle ensemble d'événements utiliser le jour du concours ?**

- ▶ Si  $\Omega$  est fini alors on pourra prendre  $\mathcal{A} = \mathcal{P}(\Omega)$ .
- ▶ Si  $\Omega$  est infini et dénombrable alors on pourra prendre  $\mathcal{A} = \mathcal{P}(\Omega)$ .
- ▶ Si  $\Omega$  est infini et indénombrable la situation se complique et sort du cadre de ce programme donc on ne cherchera pas à déterminer  $\mathcal{A}$ .

### 2. Correspondance entre relations logiques et ensemblistes

Relation logique	Relation ensembliste
négation "non"	complémentation $\bar{\cdot}$
conjonction "et"	intersection $\cap$
disjonction "ou"	réunion $\cup$
implication $\implies$	inclusion $\subset$
équivalence $\iff$	égalité $=$

<sup>1</sup>Comprendre : associé à une expérience aléatoire.

## 3. Correspondance entre le langage des événements et celui des probabilités

Langage des ensembles	Langage des probabilités
$\Omega$	<b>Événement certain</b> : il est toujours réalisé quelle que soit l'issue de l'épreuve.
$\emptyset$	<b>Événement impossible</b> : il n'est jamais réalisé quelle que soit l'issue de l'épreuve.
$\{\omega\}$	<b>Événement élémentaire</b> : c'est un singleton constitué d'un seul élément de $\Omega$ , il est réalisé si et seulement si le résultat de l'épreuve est exclusivement $\omega$ .
$A \cap B$	<b>Événement "A et B"</b> : il est réalisé si et seulement si $A \wedge B$ sont réalisés.
$A \cup B$	<b>Événement "A ou B"</b> : il est réalisé si et seulement si $A \vee B$ est réalisé.
$\bar{A}$	<b>Événement contraire de A</b> : il est réalisé si et seulement si $A$ ne l'est pas. Ainsi $\bar{A} = \complement_{\Omega} A$ .
$A \subset B$	<b>Événement "A implique B"</b> : ce qui signifie que quand $A$ est réalisé, $B$ l'est aussi.
$A = B$	<b>Événement "A égal B"</b> : ce qui signifie que les deux événements se réalisent simultanément quelle que soit l'issue $\omega$ de l'expérience. On dit aussi que les événements sont <b>équivalents</b> .
$A \Delta B$ (hors programme)	<b>Événement soit "A soit B"</b> : il est réalisé si et seulement si $A \cap B$ ne l'est pas.
$\bigcap_{i \in I} A_i$	<b>Événement "tous les <math>A_i</math> sont réalisés quand <math>i \in I</math>"</b> : il est réalisé si, et seulement si, tous les $A_i$ sont réalisés quand $i$ parcourt $I$ .
$\bigcap_{i \in I} A_i$	<b>Événement "tous les <math>A_i</math> sont réalisés quand <math>i \in I</math>"</b> : il est réalisé si, et seulement si, tous les $A_i$ sont réalisés quand $i$ parcourt $I$ .
$\bigcup_{i \in I} A_i$	<b>Événement "au moins un des <math>A_i</math> est réalisé quand <math>i \in I</math>"</b> : il est réalisé si, et seulement si, il existe $i \in I$ tel que $A_i$ soit réalisé.
$\bigcap_{n=1}^{+\infty} \bigcup_{k=n}^{+\infty} A_k$	<b>Événement "une infinité d'événements se réalise"</b> : il est réalisé si, et seulement si, pour tout $n \geq 1$ il existe $k \geq n$ tel que $A_k$ soit réalisé.
$\bigcup_{n=1}^{+\infty} \bigcap_{k=n}^{+\infty} A_k$	<b>Événement "tous les événements se réalisent à partir d'un certain rang"</b> : il est réalisé si, et seulement si, il existe $n \geq 1$ tel que pour tout $k \geq n$ tel que $A_k$ est réalisé.

## 4. Probabilité

## 4.1. Définition et propriétés.

**Définition 4.2. Axiomatique d'une probabilité**

Soit  $(\Omega, \mathcal{A})$  un espace probabilisable, on appelle probabilité sur  $(\Omega, \mathcal{A})$  toute application  $\mathbf{P} : \mathcal{A} \rightarrow [0, 1]$  vérifiant :  $\mathbf{P}(\Omega) = 1$ , pour toute  $(A_k)_{k \in K}$  famille d'événements deux à deux disjoints la série  $\sum_k \mathbf{P}(A_k) < \infty$  et

$$\mathbf{P}\left(\bigsqcup_{k \in K} A_k\right) = \sum_{k \in K} \mathbf{P}(A_k)$$

( $\sigma$ -additivité de  $\mathbf{P}$ ).

**Définition 4.3. Espace probabilisable**

On appelle espace probabilisable tout couple  $(\Omega, \mathcal{A})$ .

**Définition 4.4. Espace probabilisé**

On appelle espace probabilisé tout triplet  $(\Omega, \mathcal{A}, \mathbf{P})$ .

**Définition 4.5. Événement négligeable**

Soit  $A \in \mathcal{A}$ ,  $A$  est dit négligeable si  $\mathbf{P}(A) = 0$ .

**Remarque 4.1.** Il n'y a aucune raison que  $A = \emptyset$  bien que  $\mathbf{P}(\emptyset) = 0$ .

**Définition 4.6. Événement quasi-certain**

Soit  $A \in \mathcal{A}$ ,  $A$  est dit quasi-certain si  $\mathbf{P}(A) = 1$ .

**Remarque 4.2.** Il n'y a aucune raison que  $A = \Omega$  bien que  $\mathbf{P}(\Omega) = 1$ .

**Définition 4.7. Système complet d'événements**

On appelle système complet d'événements toute partition de l'univers.

**Propriétés 4.1.** Soit  $A$  et  $B$  deux événements de  $(\Omega, \mathcal{A}, \mathbf{P})$ ,

$$\blacktriangleright \mathbf{P}(A \sqcup B) = \mathbf{P}(A) + \mathbf{P}(B)$$

$$\blacktriangleright \mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$$

$$\blacktriangleright \mathbf{P}(A - B) = \mathbf{P}(A) - \mathbf{P}(A \cap B)$$

$$\blacktriangleright (B \subset A) \Rightarrow (\mathbf{P}(A - B) = \mathbf{P}(A) - \mathbf{P}(B))$$

$$\blacktriangleright (B \subset A) \Rightarrow (\mathbf{P}(B) \leq \mathbf{P}(A)) \text{ (croissance de } \mathbf{P})$$

$$\blacktriangleright (B \subset A) \Rightarrow (\mathbf{P}(\bar{A}) \leq \mathbf{P}(\bar{B}))$$

$$\blacktriangleright \mathbf{P}(A \cap B) \leq \min\{\mathbf{P}(A), \mathbf{P}(B)\}$$

$$\blacktriangleright \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) \text{ (formule du crible ou de POINCARÉ pour deux ensembles dont la généralisation a disparu du programme)}$$

$$\blacktriangleright \mathbf{P}(A \cup B) = \mathbf{P}(\bar{A} \cap B) + \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \bar{B})$$

$$\blacktriangleright \mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B) \text{ (la généralisation de ce résultat s'appelle l'inégalité de BOOLE mais est hors programme)}$$

$$\blacktriangleright \forall n \in \mathbf{N}^*, \mathbf{P}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \mathbf{P}(A_k) \text{ (\sigma-additivité de } \mathbf{P})$$

**Théorème 4.1. Limite monotone**

$$\blacktriangleright \left( (A_n)_{n \geq 0} \text{ est une suite croissante d'événements} \right) \Rightarrow \left( \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}(A_n) \right)$$

$$\blacktriangleright \left( (A_n)_{n \geq 0} \text{ est une suite décroissante d'événements} \right) \Rightarrow \left( \mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}(A_n) \right)$$

**Exemple 4.2. de suites monotones**

**Exemple 4.3.**  $\blacktriangleright \left(\bigcup_{k=1}^n A_k\right)_{n \geq 1}$  et  $\left(\bigcap_{k=n}^{+\infty} A_k\right)_{n \geq 1}$  sont deux suites croissantes.

$\blacktriangleright \left(\bigcap_{k=1}^n A_k\right)_{n \geq 1}$  et  $\left(\bigcup_{k=n}^{+\infty} A_k\right)_{n \geq 1}$  sont deux suites décroissantes.

**Corollaire 4.1.** Soit  $(A_n)_{n \geq 0}$  une suite quelconque d'événements,

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}\left(\bigcup_{k=0}^n A_k\right) \text{ et } \mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}\left(\bigcap_{k=0}^n A_k\right)$$

**Théorème 4.2. Inégalité de BOOLE ou sous-additivité (hors programme)**

$$\mathbf{P}\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \mathbf{P}(A_k) \text{ et } \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbf{P}(A_n) \text{ (où } \sum_{n \geq 0} \mathbf{P}(A_n) < +\infty \text{ ou } \sum_{n \geq 0} \mathbf{P}(A_n) = \infty)$$

**Proposition 4.3. Relation de LAPLACE (cas d'équiprobabilité)**

Soit  $(\Omega, \mathcal{A}, \mathbf{P})$  un espace probabilisé sur un univers fini  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$  avec  $n \in \mathbf{N}^*$ . Notons pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $p_k = \mathbf{P}(\omega_k)$ . On parle d'équiprobabilité lorsque

$$p_1 = p_2 = \dots = p_n$$

et dans ce cas

$$\forall A \in \mathcal{A}, \quad \mathbf{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$$

**Exemple 4.4.** On est dans une situation d'équiprobabilité dans le cas de tirages effectués au hasard.

**Remarque 4.3. Attention** il ne peut y avoir d'équiprobabilité lorsque l'univers est infini.

**4.2. Indépendance d'événements.****Définition 4.8. Indépendance de deux événements**

Deux événements  $A, B$  sont indépendants si la réalisation de l'un n'influe pas sur la probabilité de la réalisation de l'autre

$$\mathbf{P}_A(B) = \mathbf{P}(B) \text{ si } \mathbf{P}(A) \neq 0 \text{ ou } \mathbf{P}_B(A) = \mathbf{P}(A) \text{ si } \mathbf{P}(B) \neq 0$$

**Remarque 4.4.** On parle d'indépendance stochastique. L'indépendance dépend donc de la probabilité utilisée et d'ailleurs nous devrions écrire que les événements  $A$  et  $B$  sont indépendants pour la probabilité  $\mathbf{P}$ . On voit donc que l'indépendance de deux événements est relative à la probabilité utilisée sur l'univers  $\Omega$  et seulement à elle.

**Proposition 4.4. Caractérisation de l'indépendance de deux événements**

$$(X \perp Y) \Leftrightarrow (\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B))$$

On voit donc que l'indépendance mutuelle est plus forte que l'indépendance deux à deux.

**Remarque 4.5.** Ne pas confondre l'indépendance de deux événements avec le fait qu'ils soient incompatibles. En effet deux événements  $A$  et  $B$  incompatibles  $A$  et  $B$  et de probabilités non nulles sont nécessairement dépendants car

$$\mathbf{P}(A \cap B) = 0 \neq \mathbf{P}(A)\mathbf{P}(B)$$

On peut aussi se rendre compte que la notion d'indépendance est probabiliste alors que celle d'incompatibilité est ensembliste.

**Remarque 4.6.** On pourra prouver à titre d'exercice que

$$(A \perp A) \Leftrightarrow (\mathbf{P}(A) = 0 \text{ ou } \mathbf{P}(A) = 1)$$

**Remarque 4.7.** L'indépendance n'est pas transitive, à savoir

$$(A \perp B \text{ et } B \perp C) \not\Rightarrow (A \perp C)$$

**Proposition 4.5.** Les assertions suivantes sont équivalentes pour deux événements  $A$  et  $B$  :

- ▶  $A \perp \overline{B}$
- ▶  $\overline{A} \perp B$
- ▶  $\overline{\overline{A}} \perp B$
- ▶  $\overline{A} \perp \overline{\overline{B}}$

On pourra démontrer, à titre d'exercice, la proposition suivante :

**Proposition 4.6.** Soit  $A$  et  $B$  deux événements, alors on a

- ▶ Si  $\mathbf{P}(A) = 0$ , alors  $A$  et  $B$  sont indépendants.
- ▶ Si  $\mathbf{P}(B) = 0$ , alors  $A$  et  $B$  sont indépendants.
- ▶ Si  $\mathbf{P}(A) = 1$ , alors  $A$  et  $B$  sont indépendants.
- ▶ Si  $\mathbf{P}(B) = 1$ , alors  $A$  et  $B$  sont indépendants.

**Définition 4.9.  $n$  événements deux à deux indépendants**

$$\forall n \in \mathbf{N}^* - \{1\}, \quad \forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad (i \neq j) \Rightarrow \mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j)$$

**Définition 4.10.**  *$n$  événements mutuellement indépendants ou indépendants*

$$\forall I \in \mathcal{P}(\llbracket 1, n \rrbracket), \quad I \neq \emptyset, \quad \mathbf{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbf{P}(A_i)$$

**Remarque 4.8.** *Prouver que  $\mathbf{P}\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathbf{P}(A_i)$  ne suffit pas pour conclure que les événements  $A_1, \dots, A_n$  sont indépendants! Sur la copie on écrira que les événements  $A_1, \dots, A_n$  sont indépendants, si pour tout entier  $k$  de  $\llbracket 2, n \rrbracket$  et pour tout  $k$ -liste  $(i_1, i_2, \dots, i_k)$  tel que  $i_1 < i_2 < \dots < i_k$  on a*

$$\mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k}) = \prod_{\ell=1}^k \mathbf{P}(A_{i_\ell})$$

**Exemple 4.5.** *Si on a  $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$ ,  $\mathbf{P}(A \cap C) = \mathbf{P}(A)\mathbf{P}(C)$ ,  $\mathbf{P}(B \cap C) = \mathbf{P}(B)\mathbf{P}(C)$  et  $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$  alors les événements  $A, B$  et  $C$  sont indépendants.*

**Proposition 4.7.** *Indépendance mutuelle versus indépendance deux à deux*

*L'indépendance mutuelle d'événements entraîne l'indépendance deux à deux de ceux-ci, mais la réciproque est FAUSSE. L'indépendance mutuelle est donc plus forte que l'indépendance deux à deux.*

**Définition 4.11.** *Suites d'événements*

*On se ramène au cas fini pour toutes sous-suites finies.*

**Proposition 4.8.** *Soit  $(A_n)_{n \geq 0}$  une suite d'événements indépendants alors la suite  $(B_n)_{n \geq 0}$  où pour tout entier  $n$ ,  $B_n = A_n$  ou  $\overline{A_n}$ , reste une suite d'événements indépendants.*

**Remarque 4.9.** *On n'écrit JAMAIS de produit infini. On se ramènera TOUJOURS à un produit fini à l'aide du théorème de la limite monotone.*

### 4.3. Probabilité conditionnelle.

**Définition 4.12.** *Probabilité conditionnelle* Soit  $A \in \mathcal{A}$  tel que  $\mathbf{P}(A) \neq 0$ ,  $\mathbf{P}_A(B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}$ .

**Remarque 4.10.** *On lit "probabilité conditionnelle de  $B$  sachant que l'événement  $A$  est réalisé". Il n'est pas rare d'observer dans certaines sources que cette probabilité est notée  $\mathbf{P}(B/A)$ . Cette notation a disparu en ECG, et c'est tant mieux car elle faisait croire au lecteur que  $B/A$  était un événement, or il n'en est rien!*

**Proposition 4.9.** *Toutes les propriétés vues sur les probabilités inconditionnelles restent encore valables et s'appliquent aux probabilités conditionnelles.*

**Proposition 4.10.** *On dit que deux événements  $A$  et  $B$  où  $\mathbf{P}(A) \neq 0$  et  $\mathbf{P}(B) \neq 0$ , sont indépendants si l'une des conditions équivalentes est vérifiée :*

- ▶  $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$
- ▶  $\mathbf{P}_A(B) = \mathbf{P}(B)$
- ▶  $\mathbf{P}_B(A) = \mathbf{P}(A)$
- ▶  $\mathbf{P}_B(A) \geq \mathbf{P}(A \cap B)$
- ▶  $\mathbf{P}_A(B) \geq \mathbf{P}(A \cap B)$

**Proposition 4.11.** *Formule des probabilités composées*

*Soit  $A_1, \dots, A_n$ ,  $n$  événements tel que pour tout  $n \geq 2$ ,  $\mathbf{P}(A_1 \cap \dots \cap A_n) \neq 0$ , alors*

$$\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1) \mathbf{P}_{A_1}(A_2) \mathbf{P}_{A_1 \cap A_2}(A_3) \times \dots \times \mathbf{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n)$$

**Proposition 4.12.** *Formule des probabilités totales*

*Soit  $(A_k)_{k \in K}$  un système complet d'événements tel que pour tout  $k \in K$ ,  $\mathbf{P}(A_k) \neq 0$ , alors*

$$\forall B \in \mathcal{A}, \quad \sum_k \mathbf{P}_{A_k}(B) \mathbf{P}(A_k) < \infty \quad \text{et} \quad \mathbf{P}(B) = \sum_{k \in K} \mathbf{P}_{A_k}(B) \mathbf{P}(A_k)$$

**Théorème 4.3. Formule de BAYES**

Soit  $(A_k)_{k \in K}$  un système complet d'événements tel que pour tout  $k \in K$ ,  $\mathbf{P}(A_k) \neq 0$ , alors

$$\forall B \in \mathcal{A}, \quad \mathbf{P}(B) \neq 0, \quad \forall i \in K, \quad \mathbf{P}_B(A_i) = \frac{\mathbf{P}_{A_i}(B) \mathbf{P}(A_i)}{\sum_{k \in K} \mathbf{P}_{A_k}(B) \mathbf{P}(A_k)}$$

**Remarque 4.11.** Cette formule est surnommée la formule de la probabilité d'une cause. Elle nous permet de faire un retour vers le passé, en nous intéressant à l'événement  $A_i$  qui est l'une des causes produisant l'événement  $B$  qui en est la conséquence.

## Variabes aléatoires réelles

### 1. Généralités

**Définition 5.1.**  $\textcircled{S}\textcircled{C}$  :  $X : \Omega \rightarrow \mathbb{R}$  tel que pour tout réel  $x$ ,

$$\{X \leq x\} \in \mathcal{A}$$

**Définition 5.2. Univers image**

L'ensemble  $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$  est appelé univers image de  $X$ .

**Définition 5.3. Variable aléatoire discrète**

On dit qu'une variable aléatoire  $X$  est discrète si son univers image est au plus dénombrable et qu'elle est finie si celui-ci est fini. Si  $X(\Omega)$  est un ensemble infini et dénombrable, on dit que  $X$  est une variable aléatoire discrète infinie.

**Notation 5.1.** Dans la pratique et dans le cas d'une variable aléatoire discrète finie, il sera courant de noter  $X(\Omega) = \{x_1, x_2, \dots, x_n\}$  l'univers image d'une variable aléatoire  $X$  discrète finie et dans le cas infini

$$X(\Omega) = \{x_k \mid k \in K\}$$

où  $K$  est une partie non vide de  $\mathbb{N}$ .

**Remarque 5.1.**  $\blacktriangleright$  **Attention** si l'ensemble  $X(\Omega)$  est un ensemble infini et indénombrable, comme par exemple tout intervalle de  $\mathbb{R}$  on peut dire uniquement que  $X$  n'est pas discrète et rien d'autre, en particulier il ne faut pas annoncer de manière péremptoire qu'elle est à densité! Il ne faut surtout pas croire que le monde des variables aléatoires possède uniquement deux catégories. C'est à méditer ...

$\blacktriangleright$  Dans le cas discret, l'ensemble  $X(\Omega)$  pourrait regrouper des réels  $x_k$  (que l'on appelle des réalisations de  $X$ ) qui soit de probabilité nulle, à savoir que  $\mathbf{P}([X = x_k])$ . C'est la raison pour laquelle on introduit la notion de support d'une variable aléatoire, l'ensemble noté  $\mathcal{S}_X$  regroupant uniquement les réalisations de la variable aléatoire  $X$  qui sont de probabilité non nulle. Maintenant il faut dire que dans de la plupart des cas la différence entre ces ensembles  $X(\Omega)$  et  $\mathcal{S}_X$  n'est pas faite et on a tendance à les confondre.

Dans le cas des variables aléatoires à densité, on appellera le support de  $X$  l'ensemble pour lequel sa densité est non identiquement nulle. Faisons donc le bilan :

**Définition 5.4. Support d'une variable aléatoire**

$\blacktriangleright$   $\textcircled{S}$  : On appelle support d'une variable aléatoire discrète  $X$  l'ensemble noté  $\mathcal{S}_X$  défini par

$$\mathcal{S}_X = \{x \in X(\Omega) \mid \mathbf{P}([X = x]) \neq 0\}$$

$\blacktriangleright$   $\textcircled{C}$  : On appelle support d'une variable aléatoire continue  $X$  l'ensemble noté  $\mathcal{S}_X$  défini par

$$\mathcal{S}_X = \{x \in X(\Omega) \mid f(x) \neq 0\}$$

**Définition 5.5. Variable aléatoire à densité**

On dit qu'une variable aléatoire  $X$  est à densité si sa fonction de répartition  $F$  est continue sur  $\mathbb{R}$  et de classe  $\mathcal{C}^1$  sur  $\mathbb{R}$  éventuellement privé d'un ensemble fini de points.

**Notation 5.2. Événements**

Soit  $x$  un réel, on notera  $[X = x]$  l'événement  $\{\omega \in \Omega, X(\omega) = x\} = X^{-1}(\{x\})$  pour alléger les notations. De la même manière l'événement  $\{\omega \in \Omega, X(\omega) \leq x\} = X^{-1}([-\infty, x])$  sera noté  $[X \leq x]$  et ainsi de suite pour tous les autres cas de figure. Finalement reprenez bien que

$$[X = x] = \{\omega \in \Omega, X(\omega) = x\}$$

**Proposition 5.1.** (admis) L'ensemble  $\mathcal{V}(\Omega, \mathcal{A}, \mathbf{P})$  des variables aléatoires réelles définies sur  $(\Omega, \mathcal{A}, \mathbf{P})$  est un  $\mathbb{R}$ -espace vectoriel stable par multiplication.

**Proposition 5.2. Variables aléatoires extrêmes**

Ⓢⓐ : La borne inférieure<sup>1</sup> ainsi que la borne supérieure d'un vecteur aléatoire restent une variable aléatoire.

**Remarque 5.2.** Ces derniers temps la proposition précédente était admise en concours que ce soit à l'oral ou à l'écrit.

**Proposition 5.3. Normalisation**

Ⓢ :  $X : \Omega \rightarrow \mathbb{R}$  est une variable aléatoire discrète si  $X(\Omega)$  est au plus dénombrable et

$$\sum_{x \in X(\Omega)} \mathbf{P}([X = x]) = 1$$

**Définition 5.6. Variables aléatoires égales**

Ⓢⓐ : Nous avons  $X = Y$  lorsque pour tout  $\omega \in \Omega$ ,  $X(\omega) = Y(\omega)$ .

**Remarque 5.3. importante** Ne confondez surtout pas  $X = 7$  avec  $[X = 7]$ . En effet  $X = 7$  signifie que la fonction  $\omega \in \Omega \mapsto X(\omega) \in \mathbb{R}$  est constante, on parle de variable déterministe (i.e. variable certaine) alors que  $[X = 7]$  est un événement, soit donc un ensemble ! Voyez donc que l'on ne manipule pas le même type d'objet ...

**Définition 5.7. Variables aléatoires presque sûrement égales**

Ⓢⓐ : Nous avons  $X = Y$  p.s. lorsque  $\mathbf{P}([X = Y]) = 1$ .

**Définition 5.8. Loi de probabilité d'une variable aléatoire quelconque**

Ⓢ : On appelle loi de probabilité d'une variable aléatoire  $X : \Omega \rightarrow \mathbb{R}$  l'application  $P_X : X(\Omega) \rightarrow [0, 1]$ , définie par

$$\forall I \text{ intervalle de } \mathbb{R}, \quad P_X(I) = \mathbf{P}([X \in I])$$

**Définition 5.9. Loi de probabilité d'une variable discrète**

Ⓢ : On appelle loi de probabilité d'une variable aléatoire discrète  $X$  l'application  $P_X : X(\Omega) \rightarrow [0, 1]$ , définie par

$$\forall x \in X(\Omega), \quad P_X(x) = \mathbf{P}([X = x])$$

**Remarque 5.4. ► Attention** une notation comme  $\mathbf{P}(X)$  n'a aucun sens : les seuls objets mathématiques dont la probabilité est définie, ce sont les événements, or  $X$  n'est pas une partie de  $\Omega$  mais une application !

► On vérifie très facilement qu'une loi de probabilité d'une variable aléatoire  $X : \Omega \rightarrow E$  est une probabilité sur l'ensemble  $E$  où  $X$  prend ses valeurs.

► La loi de probabilité d'une variable aléatoire  $X$  joue un rôle capital car elle décrit les valeurs prises par  $X$  et avec quelle probabilité. C'est pourquoi je dis sans arrêt en cours : "pas de loi, pas de chocolat!".

**Définition 5.10. Densité de probabilité**

Toute fonction  $f$  à valeurs positives, qui éventuellement ne diffère de  $F$  qu'en un nombre ni de points, est une densité d'une variable aléatoire  $X$ .

**Proposition 5.4. Caractérisation d'une densité**

Ⓢ : ( $f$  densité)  $\Leftrightarrow \left( f \geq 0 \text{ sur son domaine, } \mathcal{C}^0 \text{ presque partout, } \int_{-\infty}^{+\infty} f < +\infty \text{ et vaut } 1 \text{ (normalisation)} \right)$

**Définition 5.11. Fonction de répartition d'une variable aléatoire notée  $F$**

► Ⓢ :  $\forall x \in \mathbb{R}, F(x) = \sum_{\substack{t \in X(\Omega) \\ t \leq x}} \mathbf{P}([X = t])$ .

► ⓐ :  $\forall x \in \mathbb{R}, F(x) = \int_{-\infty}^x f(u) du$ .

**Propriétés 5.1. ► Ⓢⓐ :** Dans le cas général  $\lim_{-\infty} F = 0, \lim_{+\infty} F = 1$ ,  $F$  est continue à droite en tout point de  $\mathbb{R}$ ,  $F$  est croissante au sens large.

► ⓐ : Dans le cas des variables aléatoires à densité, on rajoute que  $F$  est continue sur  $\mathbb{R}$  et de classe  $\mathcal{C}^1$  sur  $\mathbb{R} - I$  ( $I$  ensemble fini éventuellement vide, i.e.  $F$  est de classe  $\mathcal{C}^1$  presque partout).

<sup>1</sup>  $\inf_{k \in K} (X_k)$  et  $\sup_{k \in K} (X_k)$  où  $K$  est une partie finie non vide de  $\mathbf{N}$ .

**Proposition 5.5.** © : Soit  $X$  une variable aléatoire à densité, de densité  $f$  et de fonction de répartition  $F$  alors  $F' = f$  là où  $F$  est dérivable i.e. là où  $f$  est continue.

**Propriétés 5.2.** Calcul de probabilité à partir de la fonction de répartition d'une variable aléatoire quelconque

On notera  $F_X(a^-) = \lim_{x \rightarrow a^-} F_X(x)$ .

- ▶ © :  $\forall a \in \mathbb{R}, \mathbf{P}([X = a]) = F_X(a) - F_X(a^-)$ .
- ▶ © :  $\forall a \in \mathbb{R}, \mathbf{P}([X = a]) = 0$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X \leq b]) = F_X(b) - F_X(a^-)$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X \leq b]) = F_X(b) - F_X(a)$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X < b]) = F_X(b^-) - F_X(a^-)$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X < b]) = F_X(b^-) - F_X(a)$ .

**Propriétés 5.3.** Calcul de probabilité à partir de la fonction de répartition d'une variable aléatoire quelconque

- ▶ © : Pour tout intervalle  $I$  de  $\mathbb{R}$  on a,  $\mathbf{P}([X \in I]) = \int_I f(x) dx$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X \leq b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X \leq b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X < b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx$ .
- ▶ © :  $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X < b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx$ .

**Remarque 5.5.** © : La densité est définie arbitrairement aux points où  $F$  n'est pas dérivable. Donc il n'y a pas d'unicité de la densité.

**Proposition 5.6.** Caractérisation d'une loi

- ▶ © : On donne  $X(\Omega)$  ainsi que toutes les probabilités ponctuelles  $\mathbf{P}([X = x])$  pour tout  $x \in X(\Omega)$  ou bien la fonction de répartition  $F$  de  $X$ .
- ▶ © : On donne  $X(\Omega)$  puis une densité de  $X$  ou bien la fonction de répartition  $F$  de  $X$ .

**Proposition 5.7.** ©© : Pour toute variable aléatoire  $X$ , la fonction de répartition de  $X$  est continue en  $x \in \mathbb{R}$  si, et seulement si,  $\mathbf{P}([X = x]) = 0$ .

**Proposition 5.8.** © : Soit  $X$  une variable aléatoire telle que  $X(\Omega) \subset \mathbb{Z}$  alors on a

$$\forall k \in \mathbb{Z}, \mathbf{P}([X = k]) = F(k) - F(k-1)$$

où  $F$  est la fonction de répartition de  $X$ .

**Proposition 5.9.** ©© :  $(X \sim Y) \Leftrightarrow (F_X = F_Y)$  où  $F_X$  (resp.  $F_Y$ ) est la fonction de répartition de  $X$  (resp.  $Y$ ).

**Remarque 5.6.** Puisque la notation  $X \sim Y$  est hors programme, on utilisera l'écriture  $X \stackrel{\mathcal{L}}{=} Y$ .

**Proposition 5.10.** ©© :  $(X = Y) \Rightarrow (X \sim Y)$

**Attention** la réciproque est fautive !

**Proposition 5.11.** ©© :  $(X \sim Y) \Rightarrow (\forall k \geq 0, \mathbf{E}(X^k) = \mathbf{E}(Y^k))$

**Proposition 5.12.**  $\forall A \subset X(\Omega), \mathbf{P}([X \in A]) = \begin{cases} \text{©} : \sum_{x \in A} \mathbf{P}([X = x]) \\ \text{©} : \int_{x \in A} f(x) dx \end{cases}$

**Définition 5.12.** *Fonction de survie (hors programme)*

Ⓢⓐ : On appelle fonction de survie d'une variable aléatoire  $X$  la fonction  $\mathbb{S}_X$  définie par

$$\mathbb{S}_X : t \in \mathbb{R} \rightarrow \mathbf{P}([X > t]) = 1 - F(t)$$

**Propriétés 5.4.** *Calcul de probabilité à partir de la fonction de survie*

$$\text{Ⓢ} : \forall k \in \mathbb{N}^*, \quad \mathbf{P}([X = k]) = \mathbb{S}_X(k-1) - \mathbb{S}_X(k) = \mathbf{P}([X > k-1]) - \mathbf{P}([X > k])$$

**Proposition 5.13.** *Fonction d'une variable aléatoire*

Soit  $X$  une variable aléatoire définie sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$  et soit  $g$  une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$  telle que pour tout intervalle  $I$  de  $\mathbb{R}$ ,

$$\{x \in \mathbb{R} \mid g(x) \in I\} \in \mathcal{A}$$

L'application composée  $Y = g \circ X$  est une variable aléatoire, que l'on note  $g(X)$ ,

$$g(X) : \Omega \xrightarrow{X} \mathbb{R} \xrightarrow{g} \mathbb{R}$$

On dit que  $Y$  est fonction de  $X$ .

Le calcul de la loi s'effectue dans le cas discret de la manière suivante

$$\forall y \in g(X)(\Omega), \quad \mathbf{P}([g(X) = y]) = \sum_{\substack{x \in X(\Omega) \\ g(x) = y}} \mathbf{P}([X = x])$$

et dans le cas continu par la recherche de la fonction de répartition de  $g(X)$  notée  $F_{g(X)}$  définie sur  $\mathbb{R}$  par

$$F_{g(X)}(x) = \mathbf{P}([g(X) \leq x])$$

et si, par exemple,  $g$  réalise une bijection croissante sur  $I$  contenant  $X(\Omega)$ ,

$$\forall x \in \mathbb{R}, \quad \mathbf{P}([g(X) \leq x]) = \mathbf{P}([X \leq g^{-1}(x)])$$

**Remarque 5.7.** Dans le cas où  $X$  est discrète alors  $g(X)$  reste une variable aléatoire discrète, en revanche dans le cas continu, on ne pourra conclure sur la nature de  $g(X)$  sans effectuer une étude de régularité de sa fonction de répartition.

**Exemple 5.1.** On pourra méditer sur le fait que la variable aléatoire  $X - |X|$  où  $X \hookrightarrow \mathcal{N}(0, 1)$  n'est ni discrète ni continue.

## 2. Variables aléatoires particulières

**Définition 5.13.** *Variable aléatoire certaine ou déterministe*

Ⓢ : On dit qu'une variable aléatoire  $X$  est certaine (ou déterministe) lorsque son support est un singleton, c'est-à-dire que  $X$  ne prend qu'une seule valeur.

**Définition 5.14.** *Variable aléatoire quasi-certaine*

Ⓢ : On dit qu'une variable aléatoire  $X$  est quasi-certaine lorsque son support n'est pas nécessairement un singleton voire est infini dénombrable mais dont une seule réalisation est de probabilité maximale égale à 1, donc l'intégralité des autres réalisations sont de probabilités nulles.

**Définition 5.15.** *Variable aléatoire bornée*

Ⓢⓐ : On dit qu'une variable aléatoire  $X$  est bornée s'il existe une constante  $M \geq 0$  telle que pour tout  $\omega \in \Omega$ ,

$$|X(\omega)| \leq M$$

**Exemple 5.2.** Les variables aléatoires de DIRAC (variables aléatoires certaines), uniformes, de BERNOULLI, binomiales sont bornées.

**Définition 5.16.** *Variable aléatoire étagée*

Ⓢ : On dit qu'une variable aléatoire  $X$  est étagée, lorsqu'elle est discrète et prenant un nombre fini de valeurs.

**Exemple 5.3.** Toute variable aléatoire de BERNOULLI est une variable étagée.

**Définition 5.17.** *Variable aléatoire sans mémoire*

Ⓢⓐ : On dit qu'une variable aléatoire  $X$  est à valeurs dans  $\mathbb{R}_+^*$  est sans mémoire, lorsqu'on a

$$\forall (x, y) \in \mathbb{R}_+^2, \quad \mathbf{P}_{[X > y]}([X > x + y]) = \mathbf{P}([X > x])$$

**Remarque 5.8.**  $\textcircled{S}\textcircled{C}$  : La définition précédente est équivalente à écrire que

$$\forall (x, y) \in \mathbb{R}_+^2, \quad \mathbf{P}([X > x + y]) = \mathbf{P}([X > x]) \mathbf{P}([X > y])$$

**Exemple 5.4.** Les variables aléatoires géométriques sont les seules variables discrètes sans mémoire et dans le cas continu, ce sont les variables exponentielles.

**Définition 5.18. Variable aléatoire symétrique (hors programme)**

$\textcircled{S}\textcircled{C}$  : On dit qu'une variable aléatoire  $X$  est symétrique lorsque

$$X \sim -X$$

**Exemple 5.5.**  $\blacktriangleright X \hookrightarrow \mathcal{U}(\{-1, 0, 1\})$

$\blacktriangleright X \hookrightarrow \mathcal{N}(0, 1)$

**Définition 5.19. Variable aléatoire centrée**

$\textcircled{S}\textcircled{C}$  : On dit qu'une variable aléatoire  $X$  admettant une espérance est centrée lorsque

$$\mathbf{E}(X) = 0$$

**Exemple 5.6.**  $\blacktriangleright X \hookrightarrow \mathcal{U}(\{-1, 0, 1\})$

$\blacktriangleright X \hookrightarrow \mathcal{N}(0, 1)$

$\blacktriangleright$  Pour toute variable aléatoire  $X$  admettant une espérance,  $X - \mathbf{E}(X)$  est une variable centrée.

**Définition 5.20. Variable aléatoire réduite**

$\textcircled{S}\textcircled{C}$  : On dit qu'une variable aléatoire  $X$  admettant une variance est réduite lorsque

$$\mathbf{V}(X) = 1$$

**Exemple 5.7.**  $\blacktriangleright X \hookrightarrow \mathcal{N}(0, 1)$

$\blacktriangleright$  Pour toute variable aléatoire  $X$  admettant une variance non nulle,  $\frac{X}{\sqrt{\mathbf{V}(X)}}$  est une variable réduite.

### 3. Moments

**Définition 5.21. Espérance**

$$\mathbf{E}(X) = \begin{cases} \textcircled{S} : \sum_{x \in X(\Omega)} x \mathbf{P}([X = x]) \text{ sous réserve de convergence absolue} \\ \textcircled{C} : \int_{-\infty}^{+\infty} x f(x) dx \text{ sous réserve de convergence absolue} \end{cases}$$

**Proposition 5.14.**  $\textcircled{S}$  : (hors programme) Si  $\Omega$  est au plus dénombrable, alors

$$\mathbf{E}(X) = \sum_{\omega \in \Omega} X(\omega) \mathbf{P}(\{\omega\}) \text{ sous réserve de convergence absolue}$$

**Proposition 5.15.**  $\textcircled{S}\textcircled{C}$  :  $(\mathbf{E}(X) < \infty) \Leftrightarrow (\mathbf{E}(|X|) < \infty)$

**Proposition 5.16.**  $\textcircled{S}\textcircled{C}$  : (hors programme) Pour tout  $X$  variable aléatoire positive nous avons

$$(\mathbf{E}(X) < \infty) \Leftrightarrow (\mathbf{E}(\lfloor X \rfloor) < \infty)$$

**Proposition 5.17. Espérance d'une variable aléatoire bornée**

$\textcircled{S}\textcircled{C}$  : Si  $X(\Omega)$  possède un maximum  $x_{max}$  et un minimum  $x_{min}$ , alors  $\mathbf{E}(X) < \infty$  et l'on a,

$$x_{min} \leq \mathbf{E}(X) \leq x_{max}$$

**Corollaire 5.1.**  $\textcircled{S}\textcircled{C}$  :  $\forall (a, b) \in \mathbb{R}, (a \leq X \leq b) \Rightarrow (a \leq \mathbf{E}(X) \leq b)$

**Remarque 5.9.** Ainsi toute variable bornée admet une espérance et même des moments de tous ordres.

**Proposition 5.18.**  $\textcircled{S}$  :  $\forall c \in \mathbb{R}, (\mathbf{P}([X = c]) = 1) \Rightarrow (\mathbf{E}(X) = c)$

**Proposition 5.19. Espérance et fonction de survie (hors programme)**

►  $\textcircled{S}$  : Soit  $X$  une variable aléatoire discrète telle que  $X(\Omega) \subset \mathbf{N}$  alors

$$(\mathbf{E}(X) < \infty) \Leftrightarrow \left( \sum_{n \geq 0} \mathbf{P}([X > n]) < \infty \right)$$

et si l'une des deux conditions est vérifiée,

$$\mathbf{E}(X) = \sum_{n=0}^{+\infty} \mathbf{P}([X > n])$$

►  $\textcircled{C}$  : Soit  $X$  une variable aléatoire à densité positive, alors

$$(\mathbf{E}(X) < \infty) \Leftrightarrow \left( \int_0^{+\infty} \mathbf{P}([X > t]) dt < \infty \right)$$

et si l'une des deux conditions est vérifiée,

$$\mathbf{E}(X) = \int_0^{+\infty} \mathbf{P}([X > t]) dt$$

**Remarque 5.10.** ► **Attention** l'emploi de l'axiome de  $\sigma$ -additivité est inapplicable dans le cas discret pour assurer l'existence de l'espérance, puisque les événements  $([X > n])_{n \geq 0}$  ne sont pas disjoints. À méditer ...

► L'emploi de cette proposition hors programme, n'a d'intérêt, en discret, que si  $X$  est un temps d'attente.

Par exemple si  $X \hookrightarrow \mathcal{G}(p)$  alors  $\mathbf{E}(X) = \sum_{n=0}^{+\infty} q^n = \frac{1}{1-q} = \frac{1}{p}$ . Le pendant de la loi géométrique en continu étant la loi exponentielle (le lien se faisant à l'aide de la fonction partie entière), la formule de l'espérance sous forme intégrale lui sied très bien : en effet pour  $X \hookrightarrow \mathcal{E}(\lambda)$ ,

$$\int_0^{+\infty} \mathbf{P}([X > t]) dt = \int_0^{+\infty} e^{-\lambda t} dt = \frac{1}{\lambda}$$

► L'inégalité précédente fournit un moyen rapide et efficace de vérifier la cohérence de la valeur d'une espérance.

**Proposition 5.20.**  $\textcircled{S}$  : Soit  $a \in \mathbb{R}$ , et soit  $X$  une variable aléatoire certaine égale à  $a$  alors  $\mathbf{E}(X) = a$ , i.e.

$$\mathbf{E}(a1_{\Omega}) = a$$

**Proposition 5.21. Positivité de l'espérance**

$\textcircled{S}\textcircled{C}$  : Si  $X \geq 0$  alors  $\mathbf{E}(X) \geq 0$ .

**Proposition 5.22. Espérance nulle d'une variable positive**

$\textcircled{S}\textcircled{C}$  :  $(X \geq 0 \text{ et } \mathbf{E}(X) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

**Corollaire 5.2.**  $\textcircled{S}\textcircled{C}$

►  $(\mathbf{E}(|X|) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

►  $(\mathbf{E}(X^2) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

►  $(\mathbf{E}((X - \mathbf{E}(X))^2) = 0) \Leftrightarrow (X = \mathbf{E}(X) \text{ p.s.})$

**Proposition 5.23. Croissance de l'espérance**

$\textcircled{S}\textcircled{C}$  :  $\forall (X, Y) \in (L^1(\Omega, \mathcal{A}, \mathbf{P}))^2, (X \leq Y) \Rightarrow (\mathbf{E}(X) \leq \mathbf{E}(Y))$  (admis dans le cas continu)

**Attention** la réciproque est fautive.

**Proposition 5.24.**  $\textcircled{S}\textcircled{C}$  :  $\forall (X, Y) \in (L^1(\Omega, \mathcal{A}, \mathbf{P}))^2, (X = Y) \Rightarrow (\mathbf{E}(X) = \mathbf{E}(Y))$

**Attention** la réciproque est fautive.

**Théorème 5.1. Théorème de domination** (admis)

$\textcircled{S}\textcircled{C}$  : Soit  $X$  et  $Y$  deux variables aléatoires, si  $|X| \leq Y$  et si  $\mathbf{E}(Y) < \infty$  alors

$$\mathbf{E}(X) < \infty \quad \text{et} \quad |\mathbf{E}(X)| \leq \mathbf{E}(|X|) \leq \mathbf{E}(Y)$$

**Corollaire 5.3.**  $\textcircled{S}\textcircled{C}$  :  $\left. \begin{array}{l} \mathbf{E}(X) < \infty \\ Y \text{ est bornée} \end{array} \right\} \Rightarrow \mathbf{E}(XY) < \infty$

**Définition 5.22. Moment d'ordre  $r$  où  $r \geq 0$**

$$m_r(X) = \mathbf{E}(X^r) = \begin{cases} \textcircled{S} : \sum_{x \in X(\Omega)} x^r \mathbf{P}([X = x]) \text{ sous réserve de convergence absolue} \\ \textcircled{C} : \int_{-\infty}^{+\infty} x^r f(x) dx \text{ sous réserve de convergence absolue} \end{cases}$$

On rappelle que l'on note  $L^r(\Omega, \mathcal{A}, \mathbf{P})$  voire  $L^r$  tout simplement, l'espace vectoriel des variables aléatoires admettant un moment d'ordre  $r$  (hors programme).

**Proposition 5.25.**  $\textcircled{S}\textcircled{C} : \forall k \in \mathbf{N}^*, (m_{2k}(X) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

**Remarque 5.11.**  $\textcircled{C} : \text{Dans le cas continu, on pourrait se contenter de la convergence simple puisque l'intégrande ne change de signe qu'en } 0, \text{ mais c'est une histoire de programme!}$

**Proposition 5.26.**  $\textcircled{S}\textcircled{C} : \text{Toute variable aléatoire bornée admet des moments de tous ordres.}$

**Remarque 5.12.** *C'est le type de constat qui peut faire gagner beaucoup de temps en concours.*

**Définition 5.23. Moment absolu d'ordre  $r$  où  $r \geq 0$**

$$m_r(|X|) = \mathbf{E}(|X|^r) = \begin{cases} \textcircled{S} : \sum_{x \in X(\Omega)} |x|^r \mathbf{P}([X = x]) \text{ sous réserve de convergence} \\ \textcircled{C} : \int_{-\infty}^{+\infty} |x|^r f(x) dx \text{ sous réserve de convergence} \end{cases}$$

**Remarque 5.13.**  $\textcircled{C} : \text{Dans le cas continu, on pourrait se contenter de la convergence simple puisque l'intégrande ne change de signe qu'en } 0, \text{ mais c'est une histoire de programme!}$

**Proposition 5.27.**  $\textcircled{S}\textcircled{C} : \forall r \in \mathbf{N}^*, (m_r(X) < \infty) \Leftrightarrow (m_r(|X|) < \infty)$

**Définition 5.24. Moment centré d'ordre  $r$  où  $r \geq 0$**

Soit  $X$  une variable aléatoire tel que  $\mathbf{E}(X) < \infty$ ,

$$\mu_r(X) = \mathbf{E}((X - \mathbf{E}(X))^r) = \begin{cases} \textcircled{S} : \sum_{x \in X(\Omega)} (x - \mathbf{E}(X))^r \mathbf{P}([X = x]) \text{ sous réserve de convergence absolue} \\ \textcircled{C} : \int_{-\infty}^{+\infty} (x - \mathbf{E}(X))^r f(x) dx \text{ sous réserve de convergence absolue} \end{cases}$$

**Remarque 5.14.** *La variance est nombre positif représentant l'écart quadratique moyen de  $X$  à son espérance  $\mathbf{E}(X)$ . Il a la dimension d'un carré, c'est pourquoi on introduit sa racine carrée appelée écart-type noté  $\sigma(X)$  voire  $\sigma_X$ . Ainsi :*

**Définition 5.25. Écart-type  $\textcircled{S}\textcircled{C} :$**  On appelle écart-type de  $X$  le réel noté  $\sigma(X) = \sqrt{\mathbf{V}(X)}$ .

**Remarque 5.15.** *L'écart-type peut être considéré comme la distance euclidienne de  $X$  avec  $\mathbf{E}(X)$ . (hors programme)*

**Définition 5.26. Variance  $\textcircled{S}\textcircled{C} :$**   $\mathbf{V}(X) = \mu_2(X)$

**Proposition 5.28.**  $\textcircled{S}\textcircled{C} : (\mathbf{V}(X) < \infty) \Leftrightarrow (\mathbf{E}(X^2) < \infty)$

**Proposition 5.29.**  $\textcircled{S}\textcircled{C} : (\mathbf{V}(X) = 0) \Leftrightarrow (X = \mathbf{E}(X) \text{ p.s.})$

**Proposition 5.30.**  $\textcircled{S}\textcircled{C} : (\mathbf{E}(X^2) = 0) \Rightarrow (\mathbf{V}(X) = 0)$

**Proposition 5.31.**  $\textcircled{S}\textcircled{C} : (X \text{ bornée}) \Rightarrow (\mathbf{V}(X) < \infty)$

**Théorème 5.2. Théorème de KOENIG-HUYGENS**

$\textcircled{S}\textcircled{C} : \text{Soit } X \text{ tel que } \mathbf{E}(X^2) < \infty \text{ alors}$

$$\mathbf{V}(X) < \infty \quad \text{et} \quad \mathbf{V}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2$$

**Remarque 5.16. Attention** à ne pas utiliser ce théorème en guise de définition comme on le voit souvent en concours ...

**Définition 5.27.** *Moment factoriel d'ordre  $r$  où  $r \geq 1$  (hors programme)*

$$\mathbf{E}(X(X-1) \times \cdots \times (X-r+1)) = \begin{cases} \textcircled{S} : \sum_{x \in X(\Omega)} x(x-1) \times \cdots \times (x-r+1) \mathbf{P}([X=x]) \\ \textcircled{C} : \int_{-\infty}^{+\infty} x(x-1) \times \cdots \times (x-r+1) f(x) dx \end{cases}$$

sous réserve de convergence absolue dans chacun des cas

**Proposition 5.32.**  $\textcircled{S}\textcircled{C}$  : Soit  $r \in \mathbf{N}$  alors  $(m_{r+1}(X) < \infty) \Rightarrow (\forall k \in \llbracket 0, r \rrbracket, m_k(X) < \infty)$ .

**Remarque 5.17.**  $\textcircled{S}\textcircled{C}$  : Nous avons donc pour tout  $r \in \mathbf{N}^*$ ,  $L^r(\Omega, \mathcal{A}, \mathbf{P}) \subset L^{r-1}(\Omega, \mathcal{A}, \mathbf{P}) \subset \cdots \subset L^1(\Omega, \mathcal{A}, \mathbf{P})$ .

**Proposition 5.33.**  $\textcircled{S}\textcircled{C}$  : Soit  $r \in \mathbf{N}$  alors  $(\mu_{r+1}(X) < \infty) \Rightarrow (\forall k \in \llbracket 0, r \rrbracket, \mu_k(X) < \infty)$ .

**Théorème 5.3.** *Théorème de transfert (admis)*

$$\mathbf{E}(\varphi(X)) = \begin{cases} \textcircled{S} : \sum_{x \in X(\Omega)} \varphi(x) \mathbf{P}([X=x]) \text{ où } \varphi \text{ définie sur } I \text{ contenant } X(\Omega) \\ \textcircled{C} : \int_a^b \varphi(x) f(x) dx \text{ où } X(\Omega) \subset ]a, b[, a \text{ et } b \text{ dans } \overline{\mathbf{R}} \\ \text{et } \varphi : ]a, b[ \rightarrow \mathbf{R} \text{ est continue presque partout dans } ]a, b[ \end{cases}$$

sous réserve de convergence absolue dans chacun des cas

Voyons l'intérêt capital du théorème de transfert à travers un exemple.

**Exemple 5.8.** Soit  $X$  une variable aléatoire uniformément distribuée sur  $[0, 1]$ . L'espérance de la variable aléatoire  $Y = X^5 + 2X^3 - X^2 - 3X$  est égale à

$$\mathbf{E}(Y) = \int_0^1 (x^5 + 2x^3 - x^2 - 3x) dx = -\frac{7}{6}$$

La question qui se pose est : pouvons-nous obtenir ce résultat sans le théorème de transfert seulement à partir de la détermination de la loi de  $Y$  ? La réponse est clairement non ! Tout simplement car il n'existe pas de formule explicite pour résoudre une inéquation polynomiale comme

$$x^5 + 2x^3 - x^2 - 3x \leq y$$

de degré supérieur ou égal à 5 comme la démontré Evariste GALOIS (1811 – 1832) durant sa courte vie. Donc la densité de  $Y$  existe bien mais nous sommes bien incapables de l'obtenir. Il faut nous munir d'un ordinateur et employer des méthodes d'analyse numérique qui dépassent le cadre de notre programme.

**Proposition 5.34.** *Espérance et transformation affine*

$\textcircled{S}\textcircled{C}$  :  $\forall (a, b) \in \mathbf{R}^* \times \mathbf{R}$ ,

$$(\mathbf{E}(X) < \infty) \Leftrightarrow (\mathbf{E}(aX + b) < \infty) \quad \text{et} \quad \mathbf{E}(aX + b) = a\mathbf{E}(X) + b$$

**Proposition 5.35.** *Variance et transformation affine*

$\textcircled{S}\textcircled{C}$  :  $\forall (a, b) \in \mathbf{R}^* \times \mathbf{R}$ ,

$$(\mathbf{V}(X) < \infty) \Leftrightarrow (\mathbf{V}(aX + b) < \infty) \quad \text{et} \quad \mathbf{V}(aX + b) = a^2\mathbf{V}(X)$$

**Remarque 5.18.** *Attention* notez bien que le réel  $b$  a disparu du résultat précédent !

**Exemple 5.9.**  $\mathbf{V}(9X + 21) = 81\mathbf{V}(X)$ .

**Proposition 5.36.**  $\textcircled{S}\textcircled{C}$  :  $\inf_{a \in \mathbf{R}} \mathbf{E}((X - a\mathbf{1}_\Omega)^2) = \mathbf{V}(X)$  (hors programme)

**Définition 5.28.** *Variable aléatoire centrée réduite*

$\textcircled{S}\textcircled{C}$  : Soit  $X$  tel que  $\mathbf{E}(X) < \infty$  et  $\mathbf{V}(X) < \infty$ ,  $\mathbf{V}(X) \neq 0$  alors  $X^* = \frac{X - \mathbf{E}(X)}{\sigma(X)}$  est appelée variable aléatoire centrée réduite associée à  $X$ .

**Remarque 5.19.** L'étape faisant passer de  $X$  à  $X^*$  s'appelle la standardisation, consistant à centrer et réduire la variable aléatoire de départ.

## Vecteurs aléatoires discrets

### 1. Famille sommables

Conformément au programme, tout ce qui suit n'est pas exigible de la part des étudiants et devra être rappelé dans tout sujet de concours.

**Définition 6.1.** Soit  $I$  un ensemble dénombrable indexé par  $\mathbf{N}$  sous la forme  $I = \{\varphi(n), n \in \mathbf{N}\}$  où  $\varphi$  est une bijection de  $\mathbf{N}$  dans  $I$ . Si la série  $\sum u_{\varphi(n)}$  converge absolument, alors sa somme est indépendante de l'indexation  $\varphi$  et pourra également être notée  $\sum_{i \in I} u_i$ . On dira alors que la série est absolument convergente ou que la famille  $(u_i)_{i \in I}$  est sommable.

**Théorème 6.1. Le théorème de sommation par paquets (admis)**

Soit  $I = \bigsqcup_{j \in J} I_j$  un ensemble dénombrable et pour tout  $j$  de  $J$ , les  $I_j$  sont des ensembles dénombrables. Les assertions suivantes sont équivalentes :

- (1) la famille  $(u_i)_{i \in I}$  est sommable,
- (2) pour tout  $j \in J$ , la série  $\sum_{k \in I_j} u_k$  converge et la série  $\sum_j \sum_{k \in I_j} |u_k|$  converge. Dans ce cas nous avons

$$\sum_{i \in I} u_i = \sum_{j \in J} \sum_{k \in I_j} u_k$$

**Théorème 6.2. Cas particulier fondamental :  $I = \mathbf{N}^2$  (admis)**

Soit  $(I_n)_{n \in \mathbf{N}}$  une partition de  $\mathbf{N}^2$ . Les assertions suivantes sont équivalentes :

- (1) La famille  $(u_{i,j})_{(i,j) \in \mathbf{N}^2}$  est sommable.
- (2) Pour tout  $n \in \mathbf{N}$ , la série  $\sum_{i,j} |u_{i,j}|$  où  $(i,j) \in I_n$  converge et la série simple  $\sum_{n \geq 0} \sum_{(i,j) \in I_n} |u_{i,j}|$  converge aussi.

Dans ce cas nous avons

$$\sum_{(i,j) \in \mathbf{N}^2} u_{i,j} = \sum_{n=0}^{+\infty} \left( \sum_{(i,j) \in I_n} u_{i,j} \right)$$

**Proposition 6.1. Les trois grandes partitions de  $\mathbf{N}^2$**

- ▶  $\mathbf{N}^2 = \bigsqcup_{n \in \mathbf{N}} (\mathbf{N} \times \{n\})$
- ▶  $\mathbf{N}^2 = \bigsqcup_{n \in \mathbf{N}} (\{n\} \times \mathbf{N})$
- ▶  $\mathbf{N}^2 = \bigsqcup_{n \in \mathbf{N}} \{(i,j) \in \mathbf{N}^2 \mid i+j = n\}$

**Proposition 6.2. Produit de deux familles sommables (admis)**

Soit  $I$  et  $J$  deux ensembles dénombrables et soit  $(u_i)_{i \in I}$  et  $(v_j)_{j \in J}$  deux familles sommables, alors la série double  $\sum_{i,j} u_{i,j}$  est convergente et l'on a

$$\sum_{(i,j) \in I \times J} u_{i,j} = \left( \sum_{i \in I} u_i \right) \left( \sum_{j \in J} v_j \right)$$

## 2. Couples discrets

Toutes les variables aléatoires en jeu dans cette sous-section sont définies dans un même espace probabilisé.

### Définition 6.2. Loi de probabilité d'un couple

Soit  $C = (X, Y)$ . On appelle loi de  $C$  l'application  $P_C : X(\Omega) \times Y(\Omega) \rightarrow [0, 1]$  définie par

$$\forall (x_i, y_j) \in X(\Omega) \times Y(\Omega), \quad P_C(x_i, y_j) = \mathbf{P}([X = x_i] \cap [Y = y_j]) = p_{i,j}$$

### Définition 6.3. Loïs marginales

► **Loi de  $X$**  : c'est l'application  $P_X : X(\Omega) \rightarrow [0, 1]$  définie par

$$\forall x_i \in X(\Omega), \quad P_X(x_i) = \mathbf{P}([X = x_i]) = p_{i,\bullet}$$

► **Loi de  $Y$**  : c'est l'application  $P_Y : Y(\Omega) \rightarrow [0, 1]$  définie par

$$\forall y_j \in Y(\Omega), \quad P_Y(y_j) = \mathbf{P}([Y = y_j]) = p_{\bullet,j}$$

### Proposition 6.3. Loïs marginales

► **Loi de  $X$**  :  $\forall x_i \in X(\Omega), \mathbf{P}([X = x_i]) = \sum_{y_j \in Y(\Omega)} p_{i,j}$

► **Loi de  $Y$**  :  $\forall y_j \in Y(\Omega), \mathbf{P}([Y = y_j]) = \sum_{x_i \in X(\Omega)} p_{i,j}$ .

**Remarque 6.1.** On peut toujours obtenir les loïs marginales à partir de la loi d'un couple, en revanche on ne peut pas, en général, obtenir la loi d'un couple à partir de ses loïs marginales, SAUF si les deux variables en jeu sont indépendantes.

### Définition 6.4. Loïs conditionnelles

► On appelle loi conditionnelle de  $X$  sachant  $[Y = y_j]$  avec  $\mathbf{P}([Y = y_j]) \neq 0$  l'application

$$\begin{aligned} \mathbf{P}_{[Y=y_j]} : X(\Omega) &\rightarrow [0, 1] \\ x_i &\mapsto \mathbf{P}_{[Y=y_j]}([X = x_i]) = \frac{p_{i,j}}{p_{\bullet,j}} \end{aligned}$$

► On appelle loi conditionnelle de  $Y$  sachant  $[X = x_i]$  avec  $\mathbf{P}([X = x_i]) \neq 0$  l'application

$$\begin{aligned} \mathbf{P}_{[X=x_i]} : Y(\Omega) &\rightarrow [0, 1] \\ y_j &\mapsto \mathbf{P}_{[X=x_i]}([Y = y_j]) = \frac{p_{i,j}}{p_{i,\bullet}} \end{aligned}$$

### Proposition 6.4. Caractérisation de la loi d'une fonction d'un couple

Notons  $Z = g(X, Y)$ . Caractériser la loi de la variable aléatoire  $Z$ , c'est donner :

►  $Z(\Omega) = g(X, Y)(\Omega)$

► et pour tout  $z$  de  $Z(\Omega)$ ,

$$\mathbf{P}([Z = z]) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y)=z}} \mathbf{P}([X = x] \cap [Y = y])$$

$$\text{tel que } \sum_{z \in Z(\Omega)} \mathbf{P}([Z = z]) = 1.$$

**Proposition 6.5.**  $\forall I \in \mathcal{P}(Z(\Omega)), I \neq \emptyset, \mathbf{P}([Z \in I]) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y) \in I}} \mathbf{P}([X = x] \cap [Y = y]).$

### Théorème 6.3. ⑤ Théorème de transfert (admis)

Soit une application  $\varphi$  définie sur  $\mathcal{D}$  contenant  $X(\Omega) \times Y(\Omega)$ ,

$$\mathbf{E}(\varphi(X, Y)) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \varphi(x, y) \mathbf{P}([X = x] \cap [Y = y])$$

sous réserve de convergence absolue de la série.

**Exemple 6.1.** Soit  $X$  et  $Y$  deux variables aléatoires discrètes à valeurs dans  $\mathbf{N}^*$ , alors en cas de convergence absolue de la série en jeu,

$$\mathbf{E}\left(\frac{X}{Y}\right) = \sum_{(i,j) \in (\mathbf{N}^*)^2} \frac{i}{j} \times \mathbf{P}([X = i] \cap [Y = j])$$

**Conséquence 6.1.** © *Linéarité de l'espérance dans le cas de deux variables aléatoires (admis)*  
Soit  $X$  et  $Y$  deux variables aléatoires admettant une espérance et soit  $a$  et  $b$  deux réels, alors

$$\mathbf{E}(aX + bY) < \infty \quad \text{et} \quad \mathbf{E}(aX + bY) = a\mathbf{E}(X) + b\mathbf{E}(Y)$$

**Remarque 6.2.** ► © Le résultat reste valable pour les variables aléatoires à densité et même aux variables aléatoires quelconques. Il se généralise au cas de  $n$  variables aléatoires par récurrence.

► L'espérance est une forme linéaire positive sur le sous-espace vectoriel des variables aléatoires (soit discrètes, soit continues) admettant une espérance.

**Théorème 6.4.** © *Théorème de domination (admis)*

Soit  $X$  et  $Y$  deux variables aléatoires tel que  $0 \leq |X| \leq Y$  p.s. et  $\mathbf{E}(Y) < \infty$  alors  $\mathbf{E}(X) < \infty$  et dans ce cas

$$|\mathbf{E}(X)| \leq \mathbf{E}(Y)$$

**Remarque 6.3.** ► © Le résultat reste valable pour les variables aléatoires à densité.

► L'inégalité  $X \leq Y$  ne suffit pas pour conclure car  $X$  pourrait être négative et ne pas admettre d'espérance.

### 3. Vecteurs discrets de dimension $n \geq 2$

Toutes les variables aléatoires en jeu dans cette sous-section sont définies dans un même espace probabilisé.

#### 3.1. Généralités.

**Définition 6.5.** *Loi d'un vecteur discret de dimension  $n \geq 2$*

Soit  $V = (X_1, \dots, X_n)$  alors  $P_V : X_1(\Omega) \times \dots \times X_n(\Omega) \rightarrow [0, 1]$ , définie par

$$\forall (x_1, \dots, x_n) \in \prod_{i=1}^n X_i(\Omega), \quad P_V(x_1, \dots, x_n) = \mathbf{P}\left(\bigcap_{i=1}^n [X_i = x_i]\right)$$

**Remarque 6.4.** On a toujours  $V(\Omega) \subset X_1(\Omega) \times \dots \times X_n(\Omega)$ .

**Définition 6.6.** *Lois marginales de dimension un*

$$\forall k \in \llbracket 1, n \rrbracket, \quad \forall x_k \in X_k(\Omega), \quad P_{X_k}(x_k) = \mathbf{P}([X_k = x_k])$$

**Remarque 6.5.** *Attention* contrairement à ce que l'on pourrait croire, les lois marginales d'un vecteur de dimension  $n$  ne sont pas uniquement de dimension un. Elles peuvent être de dimension deux, trois etc ... (à méditer!)

**Proposition 6.6.** *Calcul de probabilités*

$\forall k \in \llbracket 1, n \rrbracket, \forall x_k \in X_k(\Omega),$

$$\mathbf{P}([X_k = x_k]) = \sum_{(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \in \prod_{m \in \llbracket 1, n \rrbracket - \{k\}} X_m(\Omega)} \mathbf{P}\left(\bigcap_{j=1}^n [X_j = x_j]\right)$$

**Théorème 6.5.** *Théorème de transfert (admis)*

Soit  $C = (X_1, X_2)$  et  $\varphi$  une fonction définie sur  $\mathbf{R}^n$  (il est suffisant d'avoir  $\varphi$  définie sur  $C(\Omega)$ ) alors

$$\mathbf{E}(\varphi(X_1, X_2)) = \sum_{(x_1, x_2) \in X_1(\Omega) \times X_2(\Omega)} \varphi(x_1, x_2) \mathbf{P}([X_1 = x_1] \cap [X_2 = x_2])$$

sous réserve de convergence absolue de la série.

**Remarque 6.6.** La généralisation de ce théorème au cas de  $n$  variables aléatoires  $X_1, \dots, X_n$  où  $n$  est supérieur à 2 est hors programme.

**Conséquence 6.2. Linéarité de l'espérance dans le cas d'au moins deux variables aléatoires**

Soit  $n$  un entier supérieur ou égal à 2 et  $X_1, \dots, X_n$   $n$  variables discrètes définies sur un même espace probabilisé admettant une espérance. Alors pour tout  $a_1, \dots, a_n$  réels,

$$\mathbf{E} \left( \sum_{k=1}^n a_k X_k \right) < \infty \quad \text{et} \quad \mathbf{E} \left( \sum_{k=1}^n a_k X_k \right) = \sum_{k=1}^n \mathbf{E}(a_k X_k)$$

**Remarque 6.7. Attention** la linéarité de l'espérance n'exige pas l'indépendance des variables aléatoires en jeu ! D'autre part on rappelle que le caractère de linéarité de l'espérance n'est pas un argument d'existence. Autrement dit on n'écrit pas une phrase du style "l'espérance existe car elle linéaire". La linéarité ne s'utilise que pour le calcul.

**Proposition 6.7. Caractérisation de la loi d'une fonction d'un vecteur**

Soit  $n \geq 2$  et  $V = (X_1, \dots, X_n)$  et  $\varphi$  une fonction définie sur une partie de  $\mathbf{R}^n$  contenant  $\prod_{i=1}^n X_i(\Omega)$ , alors la loi de  $Z = \varphi(V)$  est caractérisée par la donnée de  $Z(\Omega) \subset \text{Im}(\varphi)$  et par celles de  $s$  probabilités

$$\forall z \in Z(\Omega), \quad \mathbf{P}([Z = z]) = \sum_{\substack{(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega) \\ g(x_1, \dots, x_n) = z}} \mathbf{P} \left( \bigcap_{j=1}^n [X_j = x_j] \right)$$

avec

$$\sum_{z \in Z(\Omega)} \mathbf{P}([Z = z]) = 1$$

**Proposition 6.8. (Identité en loi - admis)** Soit  $V_1 = (X_1, \dots, X_n)$  et  $V_2 = (Y_1, \dots, Y_n)$  et  $g$  définie sur au moins  $V_1(\Omega)$  et  $V_2(\Omega)$  alors

$$(V_1 \sim V_2) \Rightarrow (g(V_1) \sim g(V_2))$$

**Proposition 6.9. (Identité en loi - admis)** Soit  $V_1 = (X_1, \dots, X_n)$  et  $V_2 = (Y_1, \dots, Y_n)$ . On suppose que les variables  $X_1, \dots, X_n$  sont indépendantes ainsi que les variables  $Y_1, \dots, Y_n$ . Enfin pour tout  $i$  de  $\{1, \dots, n\}$ ,  $X_i \sim Y_i$ . Alors

$$V_1 \sim V_2$$

**Remarque 6.8. (très importante)** Sans l'hypothèse d'indépendance des variables  $X_i$  d'un côté et des variables  $Y_j$  de l'autre, la coalition n'est pas validée. Autrement dit si l'on reprend les deux vecteurs  $V_1 = (X_1, \dots, X_n)$  et  $V_2 = (Y_1, \dots, Y_n)$  en supposant seulement que pour tout  $i$  de  $\{1, \dots, n\}$ ,  $X_i \sim Y_i$ , alors  $g(V_1)$  et  $g(V_2)$  ne suivent pas forcément la même loi. En revanche si l'on rajoute l'indépendance des variables  $X_i$  et  $Y_i$  alors  $g(V_1) \sim g(V_2)$ . À méditer ...

**4. Espérance conditionnelle****Définition 6.7. ⑤ Espérance conditionnelle**

$$\forall A \in \mathcal{A}, \quad \mathbf{P}(A) \neq 0, \quad \mathbf{E}(X | A) = \sum_{x \in X(\Omega)} x \mathbf{P}_A([X = x]) \text{ sous réserve de convergence absolue}$$

**Proposition 6.10. ⑤ :**  $\forall A \in \mathcal{A}, \mathbf{P}(A) \neq 0, (\mathbf{E}(X) < \infty) \Rightarrow (\mathbf{E}(X | A) < \infty)$ .

**Propriétés 6.1. ⑤** Soit  $X, Y, Z$  et  $T$  quatre variables aléatoires et  $g$  une fonction de deux variables réelles. Enfin  $a, b, c$  trois constantes réelles.

- ▶  $\forall x \in \mathcal{S}_X, \forall c \in \mathbf{R}, \mathbf{E}(c \mathbf{1}_\Omega | [X = x]) = c$ .
- ▶  $\forall y \in \mathcal{S}_Y, \mathbf{E}(X | [Y = y]) = \frac{\mathbf{E}(X \mathbf{1}_{[Y=y]})}{\mathbf{P}([Y = y])}$  (hors programme).
- ▶  $\forall x \in \mathcal{S}_X, \mathbf{E}(aZ + bT | [X = x]) = a\mathbf{E}(Z | [X = x]) + b\mathbf{E}(T | [X = x])$  (linéarité).
- ▶ Si  $X \perp Y$  alors  $\forall x \in \mathcal{S}_X, \mathbf{E}(Y | [X = x]) = \mathbf{E}(Y)$  et  $\forall y \in \mathcal{S}_Y, \mathbf{E}(X | [Y = y]) = \mathbf{E}(X)$ .
- ▶  $\mathbf{E}(g(X, Y) | [X = x]) = \mathbf{E}(g(x, Y) | [X = x])$
- ▶  $|\mathbf{E}(Y | [X = x])| \leq \mathbf{E}(|Y| | [X = x])$  (inégalité triangulaire)
- ▶ Si  $Z \leq Y$  alors  $\mathbf{E}(Z | [X = x]) \leq \mathbf{E}(Y | [X = x])$  (croissance de l'espérance)

**Remarque 6.9.** ► Ces propriétés restent valables en continu mais sont hors programme.

► Les espérances conditionnelles peuvent servir à construire des fonctions de régression qui étudient la valeur moyenne d'une variable aléatoire conditionnée par une autre.

**Définition 6.8.** ⑤ **Fonctions de régression**

Soit  $X$  et  $Y$  deux variables aléatoires telles que  $\mathbf{E}(Y) < \infty$ , alors on appelle fonction de régression de  $Y$  sur  $X$ , la fonction  $g_1 : \mathcal{S}_X \rightarrow \mathbb{R}$  définie par  $g_1(x) = \mathbf{E}(Y \mid [X = x])$ . De même on appelle fonction de régression de  $X$  sur  $Y$ , la fonction  $g_2 : \mathcal{S}_Y \rightarrow \mathbb{R}$  définie par  $g_2(y) = \mathbf{E}(X \mid [Y = y])$  à condition que  $\mathbf{E}(X) < \infty$ .

**Proposition 6.11.** ⑤ **Formule de l'espérance totale**

Soit  $(A_i)_{i \in I}$  un système complet d'événements tel que pour tout  $i \in I$ ,  $\mathbf{P}(A_i) \neq 0$  alors

$$(\mathbf{E}(X) < \infty) \Leftrightarrow \begin{cases} \forall i \in I, \mathbf{E}(|X| \mid A_i) < \infty \\ \sum_i \mathbf{E}(|X| \mid A_i) \mathbf{P}(A_i) < \infty \end{cases}$$

En cas d'existence

$$\mathbf{E}(X) = \sum_{i \in I} \mathbf{E}(X \mid A_i) \mathbf{P}(A_i)$$

## 5. Covariance

**Définition 6.9. Covariance** ⑤⑥ : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$ ,  $\text{Cov}(X, Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$ .

**Remarque 6.10.** ► ⑤⑥ : Cette définition ne sert quasiment jamais au calcul d'une covariance car elle n'est pas très pratique à manipuler.

► ⑤⑥ : La covariance entre deux variables aléatoires  $X$  et  $Y$  est une mesure de liaison entre elles. Si le signe de la covariance est positif, cela s'interprète en disant que si en moyenne l'une des deux variables aléatoires augmente, l'autre diminue. Et un signe négatif s'interprète en disant qu'en moyenne si une variable aléatoire tend à augmenter, l'autre diminue.

**Définition 6.10. Variables aléatoires non corrélées ou décorréelées**

⑤⑥ : On dit que deux variables aléatoires  $X$  et  $Y$  sont non corrélées ou décorréelées lorsque  $\text{Cov}(X, Y) = 0$ .

**Attention** cela ne veut pas dire qu'elles sont indépendantes !

**Remarque 6.11.** ⑤⑥ : **Attention** faites toujours la distinction entre indépendant et non-corrélation. Ne concluez jamais à l'indépendance après avoir obtenu un coefficient de corrélation nul !

**Théorème 6.6. Théorème de KOENIG-HUYGENS**

⑤⑥ : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$ ,  $\text{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$ .

**Proposition 6.12.** ⑤⑥ :  $\left. \begin{array}{l} \mathbf{E}(X^2) < \infty \\ \mathbf{E}(Y^2) < \infty \end{array} \right\} \Rightarrow (\mathbf{E}(XY) < \infty)$

**Remarque 6.12.** La proposition précédente donne une condition suffisante d'existence de la covariance du couple  $(X, Y)$ .

**Propriétés 6.2.** ► ⑤⑥ :  $\text{Cov}(X, X) = \mathbf{V}(X) \geq 0$  (positivité)

► ⑤⑥ :  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$  (symétrie)

► ⑤⑥ : Soit  $(X, Y, Z) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^3$  et  $(a, b) \in \mathbb{R}^2$ ,

$$\text{Cov}(aX + bY, Z) = a \text{Cov}(X, Z) + b \text{Cov}(Y, Z)$$

(linéarité à gauche)

► ⑤⑥ :  $\forall X \in L^2(\Omega, \mathcal{A}, \mathbf{P}), \forall a \in \mathbb{R}, \text{Cov}(X, a\mathbf{1}_\Omega) = 0$

► ⑤⑥ :  $\forall (a, b) \in \mathbb{R}^2, \text{Cov}(a\mathbf{1}_\Omega, b\mathbf{1}_\Omega) = 0$

► ⑤⑥ : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  et  $(a, b, c, d) \in \mathbb{R}^4$ ,

$$\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y)$$

► ⑤⑥ :  $\forall k \in \llbracket 1, n \rrbracket, \forall \ell \in \llbracket 1, m \rrbracket, (X_k, Y_\ell) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2, \lambda_k$  et  $\mu_\ell \in \mathbb{R}$ ,

$$\text{Cov}\left(\sum_{k=1}^n \lambda_k X_k, \sum_{\ell=1}^m \mu_\ell X_\ell\right) = \sum_{(k, \ell) \in \llbracket 1, n \rrbracket \times \llbracket 1, m \rrbracket} \lambda_k \mu_\ell (X_k, X_\ell) \quad (\text{bilinéarité})$$

**Exemple 6.2.**  $\text{Cov}(2X + 3Y, Z - W) = 2 \text{Cov}(X, Z) - 2 \text{Cov}(X, W) + 3 \text{Cov}(Y, Z) - 3 \text{Cov}(Y, W)$

**Remarque 6.13.**  $\textcircled{S}\textcircled{C}$  *Attention* puisque  $\text{Cov}(X, X) = 0$  n'entraîne pas nécessairement que  $X$  est la variable aléatoire nulle, la covariance ne définit pas en général un produit scalaire, sauf si l'on se réduit aux variables aléatoires de  $L^2(\Omega, \mathcal{A}, \mathbf{P})$  qui soient centrées et en assimilant variable aléatoire nulle p.s. et variable aléatoire nulle.

### 5.1. Matrice de covariance-variance.

**Définition 6.11.** *Matrice de covariance-variance d'un vecteur aléatoire*

$\textcircled{S}$  Soit pour tout  $k$  de  $\llbracket 1, n \rrbracket$ ,  $X_k \in L^2_d(\Omega, \mathcal{A}, \mathbf{P})$ , on appelle matrice de covariance-variance du vecteur aléatoire  $V = (X_1, \dots, X_n)$  la matrice symétrique réelle de  $S_n(\mathbb{R})$  notée  $\Gamma_V$  définie par  $\Gamma_V = (\text{Cov}(X_i, X_j))_{(i,j) \in \llbracket 1, n \rrbracket^2}$ .

**Remarque 6.14.**  $\blacktriangleright$   $\textcircled{S}$  La matrice de covariance-variance associée à un vecteur nous donne une idée de la dispersion du vecteur.

$\blacktriangleright$   $\textcircled{C}$  La définition reste valable en continu, mais elle n'est pas au programme.

**Proposition 6.13.**  $\blacktriangleright$   $\textcircled{S}$  Si les variables aléatoires  $X_1, \dots, X_n$  sont indépendantes,  $\Gamma_V$  est diagonale car toutes les covariances  $\text{Cov}(X_i, X_j)$  (avec  $i \neq j$ ) sont nulles.

**Proposition 6.14.** *Forme quadratique associée à  $\Gamma_V$*

$$\forall (a_1, \dots, a_n) \in \mathbb{R}^n, \quad {}^t [ a_1 \quad \dots \quad a_n ] \Gamma_V \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \mathbf{V} \left( \sum_{k=1}^n X_k \right) \geq 0$$

**Propriétés 6.3.** *Spectre d'une matrice de covariance-variance*

La forme quadratique associée à  $\Gamma_V$  montre que la matrice  $\Gamma_V$  est positive et donc ses valeurs propres sont positives ou nulles. Ainsi la matrice de covariance  $\Gamma_V$  est symétrique positive.

### 5.2. Variance d'une somme.

**Propriétés 6.4.**  $\textcircled{S}\textcircled{C}$  :  $\forall (X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  et  $\forall (a, b) \in \mathbb{R}^2$ , alors

$$\mathbf{V}(aX + bY) = a^2 \mathbf{V}(X) + b^2 \mathbf{V}(Y) + 2ab \text{Cov}(X, Y)$$

**Exemple 6.3.**  $\mathbf{V}(2X - Y) = 4\mathbf{V}(X) + \mathbf{V}(Y) - 4 \text{Cov}(X, Y)$

**Propriétés 6.5.**  $\textcircled{S}\textcircled{C}$  :  $\forall (X_1, \dots, X_n) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^n$  et  $\forall (a_1, \dots, a_n) \in \mathbb{R}^2$ ,

$$\mathbf{V} \left( \sum_{k=1}^n a_k X_k \right) = \sum_{k=1}^n a_k^2 \mathbf{V}(X_k) + 2 \sum_{1 \leq i < j \leq n} a_i a_j \text{Cov}(X_i, X_j) \quad (\text{hors programme})$$

**Corollaire 6.1.** Si les variables aléatoires  $X_1, \dots, X_n$  sont de même loi alors

$$\mathbf{V} \left( \sum_{k=1}^n a_k X_k \right) = n \mathbf{V}(X_1) + 2 \binom{n}{2} \text{Cov}(X_1, X_2) \quad (\text{hors programme})$$

**Proposition 6.15.** *Identité du parallélogramme*

$\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  alors

$$\mathbf{V}(X) + \mathbf{V}(Y) = \frac{1}{2} (\mathbf{V}(X + Y) + \mathbf{V}(X - Y))$$

**Proposition 6.16.** *Identités de polarisation<sup>1</sup>*

$\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  alors,

- $\blacktriangleright$   $\text{Cov}(X, Y) = \frac{1}{4} (\mathbf{V}(X + Y) - \mathbf{V}(X - Y))$
- $\blacktriangleright$   $\text{Cov}(X, Y) = \frac{1}{2} ((\mathbf{V}(X + Y) - \mathbf{V}(X) - \mathbf{V}(Y)))$
- $\blacktriangleright$   $\text{Cov}(X, Y) = \frac{1}{2} ((\mathbf{V}(X) + \mathbf{V}(Y) - \mathbf{V}(X - Y)))$

<sup>1</sup>Une identité de polarisation permet d'exprimer une forme bilinéaire symétrique en fonction de la forme quadratique associée.

**Remarque 6.15.** *On retrouve les identités remarquables précédentes pour les nombres réels et les produits scalaires. En effet*

$$\begin{aligned}(a+b)^2 &= a^2 + 2ab + b^2 \\ (a-b)^2 &= a^2 - 2ab + b^2 \\ (a+b)^2 - (a-b)^2 &= 4ab \\ \|u+v\|^2 &= \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 \\ \|u-v\|^2 &= \|u\|^2 - 2\langle u, v \rangle + \|v\|^2 \\ \|u+v\|^2 - \|u-v\|^2 &= 4\langle u, v \rangle\end{aligned}$$

**Proposition 6.17.**  $\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  alors,

$$\text{Cov}(X+Y, X-Y) = \mathbf{V}(X) - \mathbf{V}(Y)$$

**Remarque 6.16.** *On retrouve l'identité remarquable précédente pour les nombres réels et les produits scalaires. En effet*

$$\begin{aligned}a^2 - b^2 &= (a+b)(a-b) \\ \|u\|^2 - \|v\|^2 &= \langle u+v, u-v \rangle\end{aligned}$$

**Proposition 6.18.**  $\textcircled{S}\textcircled{C}$  : Si  $(X_1, \dots, X_n) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^n$  indépendantes, alors

$$\mathbf{V}\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n \mathbf{V}(X_k)$$

(additivité de la variance).

**Remarque 6.17.** *Attention la réciproque est fausse.*

### 5.3. Coefficient de corrélation linéaire.

**Définition 6.12.** *Coefficient de corrélation linéaire*

$\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  tel que  $\mathbf{V}(X) \neq 0$  et  $\mathbf{V}(Y) \neq 0$  alors on appelle coefficient de corrélation linéaire le nombre réel noté  $\rho(X, Y)$  défini par

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbf{V}(X)\mathbf{V}(Y)}} = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

**Remarque 6.18.**  $\textcircled{S}\textcircled{C}$  : Le coefficient de corrélation mesure la qualité d'une relation linéaire entre deux variables aléatoires  $X$  et  $Y$ . C'est un indicateur de dépendance linéaire entre les variables aléatoires  $X$  et  $Y$ , c'est-à-dire à quel point  $Y$  peut être approchée par des fonctions affines de  $X$ .

**Propriétés 6.6.**  $\textcircled{S}\textcircled{C}$  : Le coefficient de corrélation est invariant par changement d'échelle

$$\forall (a, b, c) \in \mathbf{R}^* \times \mathbf{R}^2, \quad \rho(aX + b, aY + c) = \rho(X, Y)$$

Si de plus  $d \neq 0$ ,

$$|\rho(aX + b, dY + c)| = |\rho(X, Y)|$$

**Propriétés 6.7.** *L'inégalité de CAUCHY-SCHWARZ*

►  $\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  tel que  $\mathbf{V}(X) \neq 0$  et  $\mathbf{V}(Y) \neq 0$  alors  $|\rho(X, Y)| \leq 1$  soit

$$|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$$

►  $\textcircled{S}\textcircled{C}$  :  $(|\rho(X, Y)| = 1) \Leftrightarrow (\exists (a, b) \in \mathbf{R}^* \times \mathbf{R} \mid Y = aX + b \text{ p.s.})$

**Remarque 6.19.**  $\textcircled{S}\textcircled{C}$  : Plus  $|\rho(X, Y)|$  est proche de 1, plus l'ajustement linéaire entre  $X$  et  $Y$  est pertinent.

## 6. Indépendance de variables aléatoires

### 6.1. Définition et proposition

#### Définition 6.13. Indépendance de $n \geq 2$ variables

ⓈⓈ : Soit  $n \geq 2$  on dit que les variables aléatoires  $X_1, \dots, X_n$  sont indépendantes ou mutuellement indépendantes lorsque

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^n, \quad \mathbf{P} \left( \bigcap_{i=1}^n [X_i \leq x_i] \right) = \prod_{i=1}^n \mathbf{P}([X_i \leq x_i])$$

Autrement dit

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

**Proposition 6.19.** Une caractérisation de l'indépendance de  $n \geq 2$  variables aléatoires (admis)

$$\text{ⓈⓈ} : (X_1 \perp \dots \perp X_n) \Leftrightarrow \left( \text{Soit } I_1, \dots, I_n, n \text{ intervalles de } \mathbb{R}, \mathbf{P} \left( \bigcap_{i=1}^n [X_i \in I_i] \right) = \prod_{i=1}^n \mathbf{P}([X_i \in I_i]) \right)$$

**Proposition 6.20.** Une caractérisation de l'indépendance de  $n \geq 2$  variables aléatoires discrètes (admis)

$$\text{Ⓢ} : (X_1 \perp \dots \perp X_n) \Leftrightarrow \left( \forall (x_1, \dots, x_n) \in \prod_{i=1}^n X_i(\Omega), \mathbf{P} \left( \bigcap_{i=1}^n [X_i = x_i] \right) = \prod_{i=1}^n \mathbf{P}([X_i = x_i]) \right)$$

**Proposition 6.21.** ⓈⓈ Soit  $X$  une variable aléatoire certaine, alors pour toute variable aléatoire  $Y$  on a  $X$  et  $Y$  indépendantes. Ceci reste valable si  $X = a$  p.s. ( $a \in \mathbb{R}$ ). Autrement dit toute variable aléatoire certaine ou quasi-certaine est indépendante (au sens du calcul des probabilités) de toute autre variable aléatoire  $y$  compris d'elle-même. Autrement dit pour tout  $X$  variable aléatoire et pour tout  $a$  réel,  $X \perp a \mathbf{1}_\Omega$ .

**Remarque 6.20. Attention** il n'y a aucune relation entre la notion d'indépendance et celle d'indépendance conditionnelle.

Prenez l'exemple de deux variables aléatoires  $X_1$  et  $X_2$  i.i.d. tel que  $X_i \hookrightarrow \mathcal{B}(p)$ . Soit  $S_2 = X_1 + X_2$  leur somme. Alors d'évidence  $S_2$  prend la valeur 0 avec une probabilité non nulle et  $\mathbf{P}_{[S_2=0]}([X_1 = 1]) > 0$  de même  $\mathbf{P}_{[S_2=0]}([X_2 = 1]) > 0$  alors que  $\mathbf{P}_{[S_2=0]}([X_1 = 1] \cap [X_2 = 1]) = 0$ .

Ainsi  $\mathbf{P}_{[S_2=0]}([X_1 = 1] \cap [X_2 = 1]) \neq \mathbf{P}_{[S_2=0]}([X_1 = 1]) \mathbf{P}_{[S_2=0]}([X_2 = 1])$ .

### 6.2. Le lemme des coalitions ou la stabilité de l'indépendance par composition

**Proposition 6.22.** (Cas de deux variables aléatoires admis) ⓈⓈ : Soit  $X, Y$  deux variables aléatoires et soit  $(\varphi, \psi) \in \mathcal{F}(I, \mathbb{R}) \times \mathcal{F}(J, \mathbb{R})$  tel que  $I$  contient  $X(\Omega)$  et  $J$  contient  $Y(\Omega)$  alors  $\varphi(X), \psi(Y)$  restent des variables aléatoires, et

$$(X \perp Y) \Rightarrow (\varphi(X) \perp \psi(Y))$$

**Remarque 6.21. Attention** la réciproque est fautive. En effet considérer les variables aléatoires  $X$  et  $Y$  tel que  $Y = X^2$  et  $\varphi : t \mapsto 6$  et  $\psi : t \mapsto 7$ , alors  $\varphi(X) \perp \psi(Y)$  bien que les variables aléatoires  $X$  et  $Y$  ne le soient pas.

**Proposition 6.23.** (Cas de  $n$  variables aléatoires admis) ⓈⓈ : Soit  $n \geq 2$  et  $X_1, \dots, X_n$   $n$  variables aléatoires et pour tout  $(\varphi, \psi) \in \mathcal{F}(D, \mathbb{R}) \times \mathcal{F}(D', \mathbb{R})$  tel que  $D$  contenant  $X_1(\Omega) \times \dots \times X_k(\Omega)$  et  $D'$  contenant  $X_{k+1}(\Omega) \times \dots \times X_n(\Omega)$  où pour tout  $k \in \llbracket 1, n-1 \rrbracket$ ,  $\varphi(X_1, \dots, X_k)$  et  $\psi(X_{k+1}, \dots, X_n)$  restent des variables aléatoires alors

$$\forall k \in \llbracket 1, n-1 \rrbracket, \quad (X_1 \perp \dots \perp X_n) \Rightarrow (\varphi(X_1, \dots, X_k) \perp \psi(X_{k+1}, \dots, X_n))$$

**Remarque 6.22. Attention** cette propriété peut s'étendre à un nombre fini de fonctions s'appliquant à une partition des variables aléatoires et en particulier au cas de  $(\varphi_1(X_1), \varphi_2(X_2), \dots, \varphi_n(X_n))$ .

**Exemple 6.4.** Pour  $n$  et  $m$  deux entiers naturels non nuls et des variables aléatoires indépendantes  $X_1, \dots, X_n, X_{n+1}, \dots, X_{n+m}$  les sommes de paquets disjoints  $\sum_{i=1}^n X_i$  et  $\sum_{j=n+1}^{n+m} X_j$  sont des variables aléatoires indépendantes.

**Proposition 6.24.** (admis)  $\textcircled{S}\textcircled{C}$  :  $\forall n \geq 2, \forall (X_1, \dots, X_n) \in (L^1(\Omega, \mathcal{A}, \mathbf{P}))^n$ ,

$$(X_1 \perp \dots \perp X_n) \Rightarrow \left( \mathbf{E} \left( \prod_{i=1}^n X_i \right) < \infty \text{ et } \mathbf{E} \left( \prod_{i=1}^n X_i \right) = \prod_{i=1}^n \mathbf{E}(X_i) \right)$$

**Proposition 6.25.**  $\textcircled{S}\textcircled{C}$  :  $\forall (X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2, (X \perp Y) \Rightarrow (\text{Cov}(X, Y) = 0)$

**Remarque 6.23.** *Attention* la réciproque de l'implication précédente est fautive en général. En effet considérer  $X \hookrightarrow \mathcal{U}(\{-1, 0, 1\})$  et  $Y = X^2$ . Maintenant elle est vérifiée si les deux variables aléatoires sont au plus binaires entre autres mais ceci est hors programme (voir exo\*\*\*).

**Proposition 6.26.**  $\textcircled{S}\textcircled{C}$  :  $\forall (X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2, (\text{Cov}(X, Y) \neq 0) \Rightarrow \text{non}(X \perp Y)$  (par contraposée)

**Proposition 6.27.**  $\textcircled{S}\textcircled{C}$  :  $\forall (X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2, (X \perp Y) \Rightarrow (\rho(X, Y) = 0)$

**Remarque 6.24.** *Attention* la réciproque de l'implication précédente est fautive en général (sauf si les deux variables aléatoires sont au plus binaires mais ceci est hors programme).

**Proposition 6.28.**  $\textcircled{S}\textcircled{C}$  :  $\forall (X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2, (\rho(X, Y) \neq 0) \Rightarrow \text{non}(X \perp Y)$  (par contraposée)

**Proposition 6.29.**  $\textcircled{S}$  :  $(\forall (X, Y) \in (L^1(\Omega, \mathcal{A}, \mathbf{P}))^2 \text{ et } X \perp Y) \Rightarrow \begin{cases} \mathbf{E}(X \mid [Y = y_j]) = \mathbf{E}(X) \\ \mathbf{E}(Y \mid [X = x_i]) = \mathbf{E}(Y) \end{cases}$   
où  $\forall y_j \in Y(\Omega), \mathbf{P}([Y = y_j]) \neq 0$  et  $\forall x_i \in X(\Omega), \mathbf{P}([X = x_i]) \neq 0$ .

**Proposition 6.30.** (admis)  $\textcircled{S}\textcircled{C}$  :  $\forall n \geq 2, \forall (X_1, \dots, X_n) \in \mathbf{R}^n, \forall (X_1, \dots, X_n) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^n$ ,

$$(X_1 \perp \dots \perp X_n) \Rightarrow \left( \mathbf{V} \left( \sum_{i=1}^n a_i X_i \right) < \infty \text{ et } \mathbf{V} \left( \sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i^2 \mathbf{V}(X_i) \right)$$

(quadraticité de la variance)

## 7. Produit de convolution

$\textcircled{S}\textcircled{C}$  **Attention** le théorème qui va suivre ne s'emploie qu'en cas de recherche de la loi de la somme de deux variables aléatoires  $X$  et  $Y$  indépendantes.

**Définition 6.14.** *Produit de convolution*

►  $\textcircled{S}$  On appelle produit de convolution de deux suites  $u$  et  $v$  (également appelé produit de CAUCHY), la suite  $w$  définie par

$$\forall n \in \mathbf{N}, w_n = \sum_{\substack{(i,j) \in \mathbf{N}^2 \\ i+j=n}} u_i v_j = \sum_{k=0}^n u_k v_{n-k}$$

►  $\textcircled{C}$  On appelle produit de convolution de deux fonctions  $f$  et  $g$ , la fonction notée  $f * g$  définie en tout réel  $x$  où elle existe, par

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(t) g(x-t) dt$$

**Remarque 6.25.**  $\textcircled{C}$  L'intégrale est convergente pour tout réel  $x$  si au moins une des deux fonctions  $f$  ou  $g$  est à support compact.

**Théorème 6.7.** *Le théorème de convolution* (admis)

- (S) Soit  $X$  et  $Y$  avec  $X \perp Y$  alors

$$\forall z \in (X + Y)(\Omega), \quad \mathbf{P}([X + Y = z]) = \sum_{x \in \{ \substack{x \in X(\Omega) \\ z-x \in Y(\Omega) \}} } \mathbf{P}([X = x]) \mathbf{P}([Y = z - x])$$

- (C) Soit  $X$  et  $Y$  avec  $X \perp Y$  alors la fonction  $f_{X+Y}$  définie par

$$\forall z \in \mathbf{R}, \quad f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(t) f_Y(z-t) dt = \int_{-\infty}^{+\infty} f_X(z-t) f_Y(t) dt$$

est une densité, à condition que  $f_{X+Y}$  soit définie et continue presque partout, ce qui est le cas, condition suffisante, si au moins l'une des deux densités est bornée (c'est-à-dire que l'une des deux densités est nulle en dehors d'un compact  $[-R, R]$  avec  $R \geq 0$ ), et on note

$$f_{X+Y} = f_X * f_Y$$

**Remarque 6.26.** ► On dit que la loi de la somme est la convolée des lois de  $X$  et de  $Y$ .

- Un excellent moyen de s'entraîner à "convoler", que ce soit en discret ou en continu, consiste à prendre deux variables aléatoires uniformes indépendantes.
- (C) En continu il est très fortement conseillé de déterminer les valeurs de  $t$  pour lesquelles soit  $f_X(t) f_Y(z-t)$  soit  $f_X(z-t) f_Y(t)$  n'est pas nul. Par exemple lorsque les deux densités sont non nulles uniquement sur  $\mathbf{R}_+$ , alors  $f_X(z-t) f_Y(t) \neq 0$  si, et seulement si,  $0 \leq t \leq z$  et les bornes d'intégration sont donc 0 et  $z$ .

### Propriétés 6.8. (C)

- Le produit de convolution est commutatif autrement dit avec les notations en continu

$$f_{X+Y} = f_X * f_Y = f_Y * f_X$$

#### Propriétés complémentaires (hors programme)

- Le produit de convolution est associatif. Autrement dit avec les notations en continu

$$(f_X * f_Y) * f_Z = f_X * (f_Y * f_Z)$$

- D'autre part  $(f_X, f_Y) \mapsto f_X * f_Y$  est une forme bilinéaire.
- Il n'existe pas de fonction  $e$  tel que  $f * e = e * f = f$ . Autrement dit, le produit de convolution n'admet pas d'élément neutre.

**Conséquence 6.3.** Le théorème de convolution nous permet de conclure que les lois Binomiale, de POISSON, gamma et normale sont stables par convolution. Il y en a beaucoup d'autres ...

## 8. Droites de régression

### Définition 6.15.

#### Proposition 6.31. Droites de régression

- (S) : Droite de régression de  $Y$  en  $X$ .  
Elle est notée  $\Delta$  et a pour équation :

$$\hat{Y} = \left( \frac{\text{Cov}(X, Y)}{\mathbf{V}(X)} \right) (X - \mathbf{E}(X)) + \mathbf{E}(Y)$$

- (S) : Droite de régression de  $X$  en  $Y$ .  
Elle est notée  $\Delta'$  et a pour équation :

$$\hat{X} = \left( \frac{\text{Cov}(X, Y)}{\mathbf{V}(Y)} \right) (Y - \mathbf{E}(Y)) + \mathbf{E}(X)$$

**Remarque 6.27.** Ces équations de droite restent valables en continu mais ne sont pas au programme.

### 9. Complément hors programme

**Proposition 6.32.** *Produit scalaire et norme associée dans  $L^2(\Omega, \mathcal{A}, \mathbf{P})$  (déjà tombé!)*

©© : L'application

$$\varphi : (X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2 \mapsto \varphi(X, Y) = \mathbf{E}(XY)$$

définit un produit scalaire dans  $L^2(\Omega, \mathcal{A}, \mathbf{P})$  en assimilant variable aléatoire nulle p.s. et variable aléatoire nulle. La norme associée est définie par

$$\forall X \in L^2(\Omega, \mathcal{A}, \mathbf{P}), \quad \|X\| = \varphi(X, X) = \sqrt{\mathbf{E}(X^2)}$$

**Exemple 6.5.** *Pour tout  $X \in L^2(\Omega, \mathcal{A}, \mathbf{P})$ , on peut écrire que  $\mathbf{V}(X) = \varphi(X - \mathbf{E}(X), X - \mathbf{E}(X))$ . D'où l'interprétation de la variance comme l'écart quadratique moyen de  $X$  à son espérance, et la distance de  $X$  à son espérance est appelée écart-type de  $X$ .*



## Variables aléatoires extrêmes

Ces variables aléatoires que l'on retrouve quasiment tout le temps en concours sont extrêmement importantes. Elles sont principalement utilisées dans des situations de crises pour modéliser des résultats aléatoires. On les rencontre dans l'analyse des impacts des changements climatiques, dans la mesure du risque d'un marché, dans la mesure de la vitesse d'évasion des étoiles, dans l'étude des séismes etc ... Elles méritent donc qu'on leur consacre un chapitre à part entière.

### 1. Panorama en dimension 2

**Définition 7.1. Variables extrêmes**

Soit 2 variables aléatoires  $X$  et  $Y$  définies sur un même espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$ . On appelle variables aléatoires extrêmes, les variables aléatoires

$$\inf(X, Y) \quad \text{et} \quad \sup(X, Y)$$

définies par

$$\begin{aligned} \inf(X, Y) &: \omega \in \Omega \mapsto \min(X(\omega), Y(\omega)) \in \mathbb{R} \\ \sup(X, Y) &: \omega \in \Omega \mapsto \max(X(\omega), Y(\omega)) \in \mathbb{R} \end{aligned}$$

**Proposition 7.1. (Des identités)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\text{Sup}(X, -X) = |X|$
- ▶  $\text{Inf}(X, -X) = -|X|$
- ▶  $\text{Sup}(X, Y) + \text{Inf}(X, Y) = X + Y$
- ▶  $\text{Sup}(X, Y) - \text{Inf}(X, Y) = |X - Y|$
- ▶  $\text{Sup}(X, Y) = \frac{X + Y + |X - Y|}{2}$
- ▶  $\text{Inf}(X, Y) = \frac{X + Y - |X - Y|}{2}$

**Remarque 7.1.** Les deux dernières identités prouvent que  $\text{Sup}(X, Y)$  et  $\text{Inf}(X, Y)$  restent des variables aléatoires par combinaison linéaire.

**Proposition 7.2. (Des inégalités)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\text{Inf}(X, Y) \leq \text{Sup}(X, Y)$
- ▶  $|\text{Sup}(X, Y)| \leq |X| + |Y|$
- ▶  $|\text{Inf}(X, Y)| \leq |X| + |Y|$

**Proposition 7.3. (Monotonie)** On considère quatre variables aléatoires  $X, X'$  et  $Y, Y'$  tel que  $X \leq X'$  et  $Y \leq Y'$  alors

- ▶  $\text{Sup}(X, Y) \leq \text{Sup}(X', Y')$
- ▶  $\text{Inf}(X, Y) \leq \text{Inf}(X', Y')$

**Proposition 7.4. (Composition avec des fonctions monotones)** On considère deux variables aléatoires  $X$  et  $Y$ .

- (1) Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  croissante, alors
  - (a)  $f(\text{Sup}(X, Y)) = \text{Sup}(f(X), f(Y))$
  - (b)  $f(\text{Inf}(X, Y)) = \text{Inf}(f(X), f(Y))$
- (2) Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  décroissante, alors
  - (a)  $f(\text{Sup}(X, Y)) = \text{Inf}(f(X), f(Y))$

$$(b) f(\text{Inf}(X, Y)) = \text{Sup}(f(X), f(Y))$$

**Proposition 7.5. (Liens entre Sup et Inf)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\text{Sup}(X, Y) = -\text{Inf}(-X, -Y)$
- ▶  $\text{Inf}(X, Y) = -\text{Sup}(-X, -Y)$

**Proposition 7.6. (Idempotence)** On considère  $X$  une variable aléatoire.

- ▶  $\text{Sup}(X, X) = X$
- ▶  $\text{Inf}(X, X) = X$

**Proposition 7.7. (Commutativité)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\text{Sup}(X, Y) = \text{Sup}(Y, X)$
- ▶  $\text{Inf}(X, Y) = \text{Inf}(Y, X)$

**Proposition 7.8. (Associativité)** On considère trois variables aléatoires  $X, Y$ , et  $Z$ .

- ▶  $\text{Sup}(\text{Sup}(X, Y), Z) = \text{Sup}(X, \text{Sup}(Y, Z))$
- ▶  $\text{Inf}(\text{Inf}(X, Y), Z) = \text{Inf}(X, \text{Inf}(Y, Z))$

**Proposition 7.9. (Distributivité)** On considère trois variables aléatoires  $X, Y$ , et  $Z$ .

- ▶  $\text{Sup}(X, \text{Inf}(Y, Z)) = \text{Inf}(\text{Sup}(X, Y), \text{Sup}(X, Z))$
- ▶  $\text{Inf}(X, \text{Sup}(Y, Z)) = \text{Sup}(\text{Inf}(X, Y), \text{Inf}(X, Z))$

**Proposition 7.10. (Invariance par translation ou stabilité par translation)** On considère trois variables aléatoires  $X, Y$ , et  $Z$ .

- ▶  $\text{Sup}(X + Z, Y + Z) = \text{Sup}(X, Y) + Z$
- ▶  $\text{Inf}(X + Z, Y + Z) = \text{Inf}(X, Y) + Z$

**Proposition 7.11. (Les lois)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\forall x \in \mathbb{R}, F_{\text{Sup}(X, Y)}(x) = \mathbf{P}([X \leq x] \cap [Y \leq x])$
- ▶  $\forall x \in \mathbb{R}, F_{\text{Inf}(X, Y)}(x) = 1 - \mathbf{P}([X > x] \cap [Y > x])$

**Proposition 7.12. (Condition suffisante de l'existence de l'espérance des variables extrêmes)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\left. \begin{array}{l} \mathbf{E}(X) < +\infty \\ \mathbf{E}(Y) < +\infty \end{array} \right\} \Rightarrow \mathbf{E}(\text{Sup}(X, Y)) < +\infty$
- ▶  $\left. \begin{array}{l} \mathbf{E}(X) < +\infty \\ \mathbf{E}(Y) < +\infty \end{array} \right\} \Rightarrow \mathbf{E}(\text{Inf}(X, Y)) < +\infty$

**Proposition 7.13. (Condition suffisante de l'existence du moment d'ordre 2 des variables extrêmes)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\left. \begin{array}{l} \mathbf{E}(X^2) < +\infty \\ \mathbf{E}(Y^2) < +\infty \end{array} \right\} \Rightarrow \mathbf{V}(\text{Sup}(X, Y)) < +\infty$
- ▶  $\left. \begin{array}{l} \mathbf{E}(X^2) < +\infty \\ \mathbf{E}(Y^2) < +\infty \end{array} \right\} \Rightarrow \mathbf{V}(\text{Inf}(X, Y)) < +\infty$

**Proposition 7.14. (Des inégalités de moments)** On considère deux variables aléatoires  $X$  et  $Y$ .

- ▶  $\mathbf{E}(\text{Sup}(X, Y)) \geq \max(\mathbf{E}(X), \mathbf{E}(Y))$
- ▶  $\mathbf{E}(\text{Inf}(X, Y)) \leq \min(\mathbf{E}(X), \mathbf{E}(Y))$

## 2. Généralisation à la dimension $n$

**Définition 7.2. Variables extrêmes**

Soit  $n$  ( $n \geq 2$ ) variables aléatoires  $X_1, \dots, X_n$  définies sur un même espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$ . On appelle variables aléatoires extrêmes, les variables aléatoires

$$\inf(X_1, \dots, X_n) \quad \text{et} \quad \sup(X_1, \dots, X_n)$$

définies par

$$\begin{aligned} \inf(X_1, \dots, X_n) &: \omega \in \Omega \mapsto \min(X_1(\omega), \dots, X_n(\omega)) \in \mathbb{R} \\ \sup(X_1, \dots, X_n) &: \omega \in \Omega \mapsto \max(X_1(\omega), \dots, X_n(\omega)) \in \mathbb{R} \end{aligned}$$

**Remarque 7.2.** ► Ces deux applications sont des bien des variables aléatoires ce qu'il est facile de voir, en particulier en dimension 2 puisque

$$\sup(X, Y) = \frac{X + Y + |X - Y|}{2} \quad \text{et} \quad \inf(X, Y) = \frac{X + Y - |X - Y|}{2}$$

► Comme  $\max$  et  $\min$  sont des lois de compositions internes commutatives et associatives<sup>1</sup> nous avons, par exemple

$$\sup(X_1, \dots, X_n, X_{n+1}) = \sup(\sup(X_1, \dots, X_n), X_{n+1}) \quad \text{et} \quad \inf(X_1, \dots, X_n, X_{n+1}) = \inf(\inf(X_1, \dots, X_n), X_{n+1})$$

La propriété d'associativité est très pratique à utiliser au cours d'une démonstration par récurrence.

### Propriétés 7.1. Événements

Soit  $x$  un réel,

$$\text{► } \{\inf(X_1, \dots, X_n) \geq x\} = \bigcap_{k=1}^n \{X_k \geq x\}$$

$$\text{► } \{\sup(X_1, \dots, X_n) \leq x\} = \bigcap_{k=1}^n \{X_k \leq x\}$$

**Remarque 7.3.** Il est très vivement conseillé d'éviter de travailler "ponctuellement" avec des variables aléatoires extrêmes (cas 3) et 4).

### Proposition 7.15. Loi des variables extrêmes

Soit  $n$  ( $n \geq 2$ ) variables aléatoires indépendantes  $X_1, \dots, X_n$  définies sur un même espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$ . On appelle définit les variables aléatoires

$$I_n = \inf(X_1, \dots, X_n) \quad \text{et} \quad S_n = \sup(X_1, \dots, X_n)$$

Alors pour tout  $x$  réel,

$$F_{I_n}(x) = \prod_{k=1}^n F_{X_k}(x) \quad \text{et} \quad F_{S_n}(x) = 1 - \prod_{k=1}^n (1 - F_{X_k}(x))$$

### Proposition 7.16. Relations reliant $\inf$ et $\sup$ en dimension $n$

$$\forall n \in \mathbf{N}^* - \{1\}, \quad \sup_{1 \leq k \leq n} (-X_k) = - \inf_{1 \leq k \leq n} (X_k) \quad \text{et} \quad \sup_{1 \leq k \leq n} (X_k) = - \inf_{1 \leq k \leq n} (-X_k)$$

**Exemple 7.1.** ► Comme la fonction  $t \mapsto \alpha t$  est bijective croissante sur  $\mathbf{R}$  pour  $\alpha$  réel strictement positif alors

$$\sup_{1 \leq k \leq n} (\alpha X_k) = \alpha \sup_{1 \leq k \leq n} (X_k) \quad \text{et} \quad \inf_{1 \leq k \leq n} (\alpha X_k) = \alpha \inf_{1 \leq k \leq n} (X_k)$$

► Comme la fonction  $t \mapsto \frac{1}{t}$  est bijective décroissante sur  $\mathbf{R}_+^*$  alors

$$\sup_{1 \leq k \leq n} \left( \frac{1}{X_k} \right) = \frac{1}{\inf_{1 \leq k \leq n} (X_k)} \quad \text{et} \quad \inf_{1 \leq k \leq n} \left( \frac{1}{X_k} \right) = \frac{1}{\sup_{1 \leq k \leq n} (X_k)}$$

► Comme la fonction  $t \mapsto \exp(t)$  est bijective sur  $\mathbf{R}$  alors

$$\exp \left( \sup_{1 \leq k \leq n} (X_k) \right) = \sup_{1 \leq k \leq n} (\exp(X_k)) \quad \text{et} \quad \exp \left( \inf_{1 \leq k \leq n} (X_k) \right) = \inf_{1 \leq k \leq n} (\exp(X_k))$$

### Proposition 7.17. La formule d'inclusion-exclusion (hors programme).

$$\sup(X_1, \dots, X_n) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_n \leq n} \inf(X_{i_1}, \dots, X_{i_k})$$

### Proposition 7.18. Monotonie des suites de variables aléatoires extrêmes

► La suite  $\left( \sup_{1 \leq k \leq n} (X_k) \right)_{n \geq 1}$  est croissante.

<sup>1</sup> $\max(x, y) = \max(y, x)$  et  $\max(x, y, z) = \max((x, y), z) = \max((y, x), z) = \dots$  Idem pour  $\min$ . nous avons, par exemple  $\sup(X_1, \dots, X_n, X_{n+1}) = \sup((X_1, \dots, X_n), X_{n+1})$  idem pour  $\inf$ .

- La suite  $\left(\inf_{1 \leq k \leq n} (X_k)\right)_{n \geq 1}$  est décroissante.

**Proposition 7.19. Comparaison des variables aléatoires extrêmes**

Soit  $n$  ( $n \geq 2$ ) variables aléatoires  $X_1, \dots, X_n$  définies sur un même espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$

- $\forall n \geq 1, \inf_{1 \leq k \leq n} (X_k) \leq \sup_{1 \leq k \leq n} (X_k)$
- $\forall k \in \llbracket 1, n \rrbracket, 0 \leq \left| \inf_{1 \leq k \leq n} (X_k) \right| \leq \sum_{k=1}^n |X_k|$
- $\forall k \in \llbracket 1, n \rrbracket, |X_k| \leq \left| \sup_{1 \leq k \leq n} (X_k) \right| \leq \sum_{k=1}^n |X_k|$

**Théorème 7.1. Théorème de Fisher -Tippett -Gnedenko : convergence en loi d'une suite de variables aléatoires extrêmes (hors programme)**

Puisque les variables aléatoires extrêmes ne sont pas fonction de sommes de variables aléatoires, le théorème central limite leur est inapplicable dans une étude de convergence en loi. On utilise le théorème de FISHER–TIPPETT–GNEDENKO (de la fin des années 20 et 40 hors programme) qui tombe régulièrement sans le dire ... Il est alors l'occasion de découvrir les lois de WEIBULL, de FRÉCHET et de GUMBEL qui sont des lois à densité.

Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires i.i.d. de fonction de répartition commune  $F$ . S'il existe deux suites réelles normalisantes  $(a_n)_{n \geq 1}$  de termes strictement positifs et  $(b_n)_{n \geq 1}$  de réels quelconque et une loi non dégénérée (non certaine) tel que telles que la suite  $\left(\frac{1}{a_n} \left(\sup_{1 \leq k \leq n} (X_k) - b_n\right)\right)_{n \geq 1}$  converge en loi, converge en loi vers une loi limite non dégénérée (non de DIRAC), c'est-à-dire que

$$\lim_{n \rightarrow +\infty} \mathbf{P} \left( \left[ \frac{\sup_{1 \leq k \leq n} (X_k) - b_n}{a_n} \leq x \right] \right) = \lim_{n \rightarrow +\infty} F^n(a_n x + b_n) = G_\gamma(x)$$

alors

$$G_\gamma(x) = \begin{cases} \exp\left(- (1 + \gamma x)_+^{-1/\gamma}\right) & \text{si } \gamma \neq 0 \\ \exp(-\exp(-x)) & \text{si } \gamma = 0 \end{cases}$$

où  $\gamma \in \mathbb{R}$  et  $z_+ = \max(0, z)$ .

**Remarque 7.4.** On dit que l'on normalise  $\sup_{1 \leq k \leq n} (X_k)$  de sorte qu'elle converge en loi. On trouve des résultats similaires avec une suite de  $\inf_{1 \leq k \leq n} (X_k)$ .

On distingue trois formes différentes pour la loi  $G_\gamma$ , selon le signe de  $\gamma$ . On parle de **domaines d'attraction**, et des recherches récentes ont amené à affiner ces domaines comme le prouvent les travaux d'EMBRECHTS et de ses collaborateurs on proposé en 2013 en proposant une liste non exhaustive de lois associées aux différents domaines d'attraction selon les valeurs du paramètre  $\gamma$ . Il faut savoir qu'une grande partie de ces lois constituent ce que j'appelle en cours le fameux "Top 30" dont vous comprenez maintenant l'importance que ce soit en sujet de concours soit en cours en Ecole ... Voici un tableau récapitulatif

Signe de $\gamma$ et domaines d'attraction	$\gamma < 0$ WEIBULL	$\gamma = 0$ GUMBEL	$\gamma > 0$ FRÉCHET
	Uniforme Beta	Normale Exponentielle Log-normale Gamma WEIBULL Logistique GUMBEL	PARETO STUDENT BURR Khi-deux FRÉCHET CAUCHY

**Remarque 7.5.** L'application de ce théorème est très régulièrement donnée en concours.

## Inégalités probabilistes

### 1. Inégalités de concentration

**Définition 8.1.**  $\textcircled{S}\textcircled{C}$  : On appelle *inégalité de concentration*, toute inégalité donnant une majoration du type ( $f$  étant une fonction explicite)

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq x) \leq f(x)$$

**Théorème 8.1. Inégalité de MARKOV**

►  $\textcircled{S}\textcircled{C}$  : Soit  $X \in L^1(\Omega, \mathcal{A}, \mathbf{P})$  à valeurs positives, alors

$$\forall \varepsilon > 0, \quad \mathbf{P}([X \geq \varepsilon]) \leq \frac{\mathbf{E}(X)}{\varepsilon}$$

►  $\textcircled{S}\textcircled{C}$  : Ce qui se généralise à l'ordre  $r$  pour  $r \in \mathbf{N}^*$ ,

$$\forall \varepsilon > 0, \quad \mathbf{P}([X \geq \varepsilon]) \leq \frac{\mathbf{E}(X^r)}{\varepsilon^r}$$

(Inégalité de MARKOV à l'ordre  $r$ )

►  $\textcircled{S}\textcircled{C}$  : (hors programme) Si  $\varphi : I \mapsto \mathbf{R}_+^*$  est une fonction bijective et strictement croissante sur  $I$  contenant  $X(\Omega)$  alors

$$\forall \varepsilon > 0, \quad \mathbf{P}([X \geq \varepsilon]) = \mathbf{P}([\varphi(X) \geq \varphi(\varepsilon)]) \leq \frac{\mathbf{E}(\varphi(X))}{\varphi(\varepsilon)}$$

**Remarque 8.1.**  $\textcircled{S}\textcircled{C}$  : Dans le cas d'une variable aléatoire  $X$  de signe quelconque on a

$$\forall \varepsilon > 0, \quad \mathbf{P}(|X| \geq \varepsilon) \leq \frac{\mathbf{E}(|X|)}{\varepsilon}$$

et en généralisant

$$\forall r \in \mathbf{N}^*, \quad \forall \varepsilon > 0, \quad \mathbf{P}(|X|^r \geq \varepsilon) \leq \frac{\mathbf{E}(|X|^r)}{\varepsilon}$$

**Théorème 8.2. Inégalité de BIENAYMÉ-TSCHEBYCHEV**

►  $\textcircled{S}\textcircled{C}$  : Soit  $X \in L^2(\Omega, \mathcal{A}, \mathbf{P})$  admettant une variance non nulle, alors

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

►  $\textcircled{S}\textcircled{C}$  : (hors programme) Ce qui se généralise à l'ordre  $r$  par

$$\forall r \in \mathbf{N}^*, \quad \forall \varepsilon > 0, \quad \mathbf{P}(|X - \mathbf{E}(X)| \geq \varepsilon) \leq \frac{\mathbf{V}(X^r)}{\varepsilon^r}$$

**Remarque 8.2.** Dans le cas où  $\varepsilon > 1$ , la "morale" est la suivante : plus  $X$  admet un moment d'ordre élevé (on dit aussi plus elle est intégrable), plus elle est concentrée autour de sa moyenne.

**Théorème 8.3. Inégalité de CHERNOFF (hors programme)**

$\textcircled{S}\textcircled{C}$  : Soit  $X$  une variable aléatoire alors pour tout  $t$  réel et tout  $a > 0$

$$\mathbf{P}([X \geq t]) \leq e^{-at} \mathbf{E}(e^{aX}) \quad \text{et} \quad \mathbf{P}([X \leq t]) \leq e^{-at} \mathbf{E}(e^{-aX})$$

sous réserve d'existence des espérances en jeu.

## 2. Inégalités de moments

### 2.1. Inégalité de Cauchy-Schwarz.

**Théorème 8.4.**  $\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  alors  $\text{Cov}(X, Y) < \infty$  et

$$(\text{Cov}(X, Y))^2 \leq \mathbf{V}(X) \mathbf{V}(Y)$$

*D'autres versions*

$\blacktriangleright$   $\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  alors  $\mathbf{E}(XY) < \infty$  et

$$(\mathbf{E}(XY))^2 \leq \mathbf{E}(X^2) \mathbf{E}(Y^2)$$

Le cas d'égalité est atteint lorsqu'il existe un réel  $a$  non nul tel que  $Y = aX$  p.s.

$\blacktriangleright$   $\textcircled{h}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  alors  $\mathbf{E}(XY) < \infty$  et

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

**Remarque 8.3.**  $\blacktriangleright$  On retrouve le résultat  $\mathbf{E}(|X|) \leq \sqrt{\mathbf{E}(X^2)}$  en posant  $Y$  la variable constante égale à 1 dans l'inégalité précédente.

$\blacktriangleright$  On retrouve le premier résultat du théorème en posant  $X' = X - \mathbf{E}(X)$  et  $Y' = Y - \mathbf{E}(Y)$ .

**Remarque 8.4.**  $\blacktriangleright$  On peut utiliser l'inégalité de CAUCHY-SCHWARZ à partir d'une variable aléatoire  $X$  positive, en remarquant que l'on peut écrire  $X = X\mathbf{1}_{[X>0]}$ .

$\blacktriangleright$  (hors programme) Avec le produit scalaire défini en proposition 8.32, l'inégalité de CAUCHY-SCHWARZ peut s'écrire

$$(\langle X, Y \rangle)^2 \leq \|X\|^2 \|Y\|^2$$

### 2.2. Inégalité triangulaire.

**Théorème 8.5.**  $\textcircled{S}\textcircled{C}$  : Soit  $X$  et  $Y$  telles que  $\mathbf{E}(X) < \infty$  et  $\mathbf{E}(Y) < \infty$  alors

$$\mathbf{E}(|X + Y|) < \infty \quad \text{et} \quad \mathbf{E}(|X + Y|) \leq \mathbf{E}(|X|) + \mathbf{E}(|Y|)$$

**Remarque 8.5.**  $\textcircled{S}\textcircled{C}$  : Cette inégalité se généralise au cas de  $n$  variables aléatoires aléatoires.

### 2.3. Inégalité de Jensen (hors programme).

**Théorème 8.6. Inégalité de JENSEN**  $\textcircled{S}\textcircled{C}$  : Soit  $X$  une variable aléatoire à valeurs dans un intervalle  $I$  et  $\varphi$  une fonction convexe sur un intervalle  $I$  contenant  $X(\Omega)$ , alors si  $X$  et  $\varphi(X)$  admettent une espérance on a,

$$\varphi(\mathbf{E}(X)) \leq \mathbf{E}(\varphi(X))$$

**Remarque 8.6.**  $\textcircled{S}\textcircled{C}$  : **Attention** en général  $\varphi(\mathbf{E}(X)) \neq \mathbf{E}(\varphi(X))$ .

**Exemple 8.1.**  $\blacktriangleright$   $\exp(\mathbf{E}(X)) \leq \mathbf{E}(\exp(X))$

- $\blacktriangleright$   $\frac{1}{\mathbf{E}(X)} \leq \mathbf{E}\left(\frac{1}{X}\right)$
- $\blacktriangleright$   $|\mathbf{E}(X)| \leq \mathbf{E}(|X|)$  (inégalité triangulaire)
- $\blacktriangleright$   $(\mathbf{E}(X))^2 \leq \mathbf{E}(X^2)$
- $\blacktriangleright$   $\ln(\mathbf{E}(X)) \geq \mathbf{E}(\ln(X))$
- $\blacktriangleright$   $\sqrt{\mathbf{E}(X)} \geq \mathbf{E}(\sqrt{X})$

### 2.4. Condition suffisante d'existence de l'espérance du produit de deux variables aléatoires.

**Proposition 8.1.**  $\textcircled{S}\textcircled{C}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$  alors  $\mathbf{E}(|XY|) < \infty$  et

$$\mathbf{E}(|XY|) \leq \frac{\mathbf{E}(X^2) + \mathbf{E}(Y^2)}{2}$$

Voir la preuve dans l'exo \*\*\*

**Corollaire 8.1.**  $\textcircled{S}\textcircled{C}$  : Soit  $X \in L^2(\Omega, \mathcal{A}, \mathbf{P})$  alors  $\mathbf{E}(|X|) < \infty$  et

$$\mathbf{E}(|X|) \leq \frac{1 + \mathbf{E}(X^2)}{2}$$

**Remarque 8.7.** *On retrouve avec ce dernier résultat par domination*

$$(\mathbf{E}(X^2) < \infty) \Rightarrow (\mathbf{E}(X) < \infty)$$

**2.5. Une inégalité avec une variable aléatoire indicatrice.**

**Proposition 8.2.**  $\textcircled{S}\textcircled{C}$  :  $\forall \varepsilon > 0, \mathbf{E}(X\mathbf{1}_{[X \leq \varepsilon]}) \leq \varepsilon.$

**Remarque 8.8.** *Ce résultat n'est pas dans le cours, il faut donc savoir le démontrer.*



## Convergences

### 1. Convergence en probabilité

**Définition 9.1.** *Convergence en probabilité*  $\textcircled{S}\textcircled{C}$  :  $(X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X$  lorsque

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbf{P}(|X_n - X| \geq \varepsilon) = 0$$

soit encore

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbf{P}(|X_n - X| < \varepsilon) = 1$$

**Remarque 9.1.**  $\blacktriangleright$  Toutes les variables aléatoires en jeu étant définies sur le même espace probabilisé, c'est ce que l'on appelle une convergence spatiale.

$\blacktriangleright$  En français : La convergence en probabilité de la suite  $(X_n)_{n \geq 1}$  vers  $X$  garantit que la probabilité que les variables aléatoires  $X_n$  et  $X$  restent "éloignées" tend vers 0.

**Remarque 9.2.**  $\textcircled{S}\textcircled{C}$  :  $\forall \varepsilon > 0, [|X_n - X| \geq \varepsilon] \subset [|X_n - X| > \frac{\varepsilon}{2}]$ .

**Propriétés 9.1.**  $\blacktriangleright \textcircled{S}\textcircled{C}$  :  $\left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \right) \Leftrightarrow \left( (X_n - X)_{n \geq 1} \xrightarrow{\mathbf{P}} 0 \right)$ .

$\blacktriangleright \textcircled{S}\textcircled{C}$  :  $\left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \right) \not\Leftarrow \left( \lim_{n \rightarrow +\infty} \mathbf{E}(X_n) = \mathbf{E}(X) \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}(X_n - X) = 0 \right)$

$\blacktriangleright \textcircled{S}\textcircled{C}$  :  $\left( \lim_{n \rightarrow +\infty} \mathbf{E}(X_n - X) = 0 \in \mathbf{R} \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}(X_n - X) = 0 \right) \Rightarrow \left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \right)$

**Corollaire 9.1.**  $\textcircled{S}\textcircled{C}$  :  $\left( \lim_{n \rightarrow +\infty} \mathbf{E}(X_n) = a \in \mathbf{R} \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}(X_n) = 0 \right) \Rightarrow \left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} a \right)$

**Propriétés 9.2.**  $\textcircled{S}\textcircled{C}$  :  $\left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y \right) \Rightarrow \left( (X_n + Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X + Y \right)$

**Propriétés 9.3.**  $\textcircled{S}\textcircled{C}$  (hors programme mais ... à savoir démontrer ... voir exo \*\*\*)

$\blacktriangleright \left( (|X_n|)_{n \geq 1} \xrightarrow{\mathbf{P}} 0 \right) \Leftrightarrow \left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} 0 \right)$

$\blacktriangleright \left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y \right) \Rightarrow (X = Y \text{ p.s.})$

$\blacktriangleright \forall (a, b) \in \mathbf{R}^2, \left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y \right) \Rightarrow \left( (aX_n + bY_n)_{n \geq 1} \xrightarrow{\mathbf{P}} aX + bY \right)$

$\blacktriangleright \left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y \right) \Rightarrow \left( (X_n Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} XY \right)$

$\blacktriangleright \left( (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y \right) \Rightarrow \left( \left( \frac{X_n}{Y_n} \right)_{n \geq 1} \xrightarrow{\mathbf{P}} \frac{X}{Y} \text{ avec } \mathbf{P}([Y = 0]) = 0 \right)$

**Théorème 9.1.** *Stabilité de la convergence en probabilité par composition* (admis)

$\textcircled{S}\textcircled{C}$  : Si  $(X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X$  et si  $f \in \mathcal{C}^0(\mathbf{R}, \mathbf{R})$  alors

$$\left( f(X_n) \right)_{n \geq 1} \xrightarrow{\mathbf{P}} f(X)$$

**Remarque 9.3.**  $\blacktriangleright$  Il est suffisant d'avoir  $f$  continue sur un intervalle  $I$  tel que  $\mathbf{P}([X \in I]) = 1$ , mais ce n'est pas dit dans le programme.

$\blacktriangleright \textcircled{S}\textcircled{C}$  : Ce résultat s'applique essentiellement en concours aux transformations affines, exponentielles, à la valeur absolue, à l'élevation à une puissance entière.

**Théorème 9.2. Loi faible des grands nombres**

ⓈⓈ : Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires indépendantes, d'espérance commune  $m$  et de variance commune  $\sigma^2$ . Soit  $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$  alors

$$(\overline{X}_n)_{n \geq 1} \xrightarrow{\mathbf{P}} m$$

**Propriétés 9.4.** Ⓢ : Si pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $X_k \hookrightarrow \mathcal{B}(p)$  alors en posant  $F_n = \frac{1}{n} \sum_{k=1}^n X_k$  on a,

$$\forall \varepsilon > 0, \quad \mathbf{P}(|F_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$$

**2. Convergence en loi**

**Définition 9.2.** ⓈⓈ :  $(X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X$  lorsque  $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$  en tout point  $x$  où  $F_X$  est continue.

**Remarque 9.4.** ► (*culturelle hors programme*) Les fonctions  $F_{X_n}$  et  $F_X$  représentent la fonction de répartition de  $X_n$  (resp. de  $X$ ), les variables aléatoires en jeu n'ont pas l'obligation d'être définies sur le même espace. La convergence en loi n'est donc pas une convergence spatiale.

► On évitera une écriture abusive du type  $(X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$  alors qu'en toute rigueur on doit écrire  $(X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X$  où  $X \hookrightarrow \mathcal{N}(0, 1)$ .

**Proposition 9.1.** ⓈⓈ : Soit  $a$  et  $b$  deux réels tel que  $a < b$  alors on a

$$\lim_{n \rightarrow +\infty} \mathbf{P}([a < X_n \leq b]) = \mathbf{P}([a < X \leq b])$$

**Proposition 9.2. Caractérisation de la convergence en loi dans le cas discret**

Ⓢ :  $((X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X) \Leftrightarrow \left( \forall k \in \mathbf{N}, \lim_{n \rightarrow +\infty} \mathbf{P}([X_n = k]) = \mathbf{P}([X = k]) \right)$  où les variables aléatoires  $X_n$  et  $X$  sont à valeurs dans  $\mathbf{N}$ .

**Remarque 9.5.** ► **Attention** n'employer cette proposition qu'en cas où vous soyez sûrs de connaître la nature discrète de TOUTES les variables aléatoires en jeu.

► **Attention** par rapport à la précédente réforme, la proposition 11.2 est réduite uniquement aux variables aléatoires à valeurs dans  $\mathbf{N}$ .

**Théorème 9.3. Stabilité de la convergence en loi par composition (admis)**

ⓈⓈ : Si  $(X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X$  et si  $f \in \mathcal{C}^0(\mathbf{R}, \mathbf{R})$  alors on a

$$(f(X_n))_{n \geq 1} \xrightarrow{\mathcal{L}} f(X)$$

**Remarque 9.6.** ► Il est suffisant d'avoir  $f$  continue sur un intervalle  $I$  tel que  $\mathbf{P}([X \in I]) = 1$ , mais ce n'est pas dit dans le programme.

► ⓈⓈ : Ce résultat s'applique essentiellement en concours aux transformations affines, exponentielles, à la valeur absolue, à l'élevation à une puissance entière.

**Théorème 9.4. Théorème de SLUTSKY (hors programme)**

► ⓈⓈ :  $\left. \begin{array}{l} (X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X \\ (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} c \in \mathbf{R} \end{array} \right\} \Rightarrow (X_n + Y_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X + c$  (SLUTSKY **additif**)

► ⓈⓈ :  $\left. \begin{array}{l} (X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X \\ (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} c \in \mathbf{R} \end{array} \right\} \Rightarrow (X_n Y_n)_{n \geq 1} \xrightarrow{\mathcal{L}} cX$  (SLUTSKY **multiplicatif**)

**Remarque 9.7. Attention** le théorème de SLUTSKY mélange les types de convergence.

**Théorème 9.5. Théorème central limite (admis)**

►  $\textcircled{S}\textcircled{C}$  : Soit  $(X_n)_{n \geq 1}$  i.i.d. alors  $(S_n^*)_{n \geq 1} \xrightarrow{\mathcal{L}} N$  où  $N \hookrightarrow \mathcal{N}(0, 1)$  où pour tout  $n \in \mathbf{N}^*$ ,  $S_n = \sum_{k=1}^n X_k$  et

$$S_n^* = \frac{S_n - \mathbf{E}(S_n)}{\sigma(S_n)} \text{ où } \sigma(S_n) > 0.$$

Autrement dit pour tout  $x \in \mathbb{R}$ ,

$$\lim_{n \rightarrow +\infty} \mathbf{P}([S_n^* \leq x]) = \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

En posant  $\forall n \geq 1$ ,  $\mathbf{E}(X_n) = \mu$  et  $\mathbf{V}(X_n) = \sigma^2$ , alors  $\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}\right)_{n \geq 1} \xrightarrow{\mathcal{L}} N$  où  $N \hookrightarrow \mathcal{N}(0, 1)$ .

►  $\textcircled{S}\textcircled{C}$  : Le théorème central limite existe aussi en “version moyenne”, c’est-à-dire qu’en utilisant les mêmes notations et en posant  $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$  alors

$$\left(\overline{X}_n^*\right)_{n \geq 1} = \left(\sqrt{n} \left(\frac{\overline{X}_n - \mu}{\sigma}\right)\right)_{n \geq 1} \xrightarrow{\mathcal{L}} N$$

où  $N \hookrightarrow \mathcal{N}(0, 1)$ .

**Remarque 9.8. Attention** on ne dit pas théorème de la limite centrée !

**Corollaire 9.2.** (admis)  $\textcircled{S}\textcircled{C}$  :  $\forall (a, b) \in \mathbb{R}^2$ ,  $a < b$ ,

$$\lim_{n \rightarrow +\infty} \mathbf{P} \left( \left[ a \leq \frac{S_n - n\mathbf{E}(X_1)}{\sqrt{n\mathbf{V}(X_1)}} \leq b \right] \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

### 3. Liens entre ces deux types de convergence

**Proposition 9.3.** (hors programme)  $\textcircled{S}\textcircled{C}$  : Soit  $(X_n)_{n \geq 1}$  et  $X$  des variables aléatoires définies sur le même espace probabilisé, alors

$$\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X\right) \Rightarrow (X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X$$

**Remarque 9.9.** La réciproque de cette implication est fautive, il suffit de considérer  $X_n = (-1)^n X$  où  $X \hookrightarrow \mathcal{N}(0, 1)$ .

**Remarque 9.10.** Cette implication montre que la convergence en loi est plus faible que la convergence en loi.

**Proposition 9.4.** (hors programme)  $\textcircled{S}\textcircled{C}$  :  $\forall a \in \mathbb{R}$ ,  $\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} a\right) \Leftrightarrow \left((X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} a\right)$ .



## Approximations

**Proposition 10.1. Binomiale par POISSON**

Pour tout  $n \in \mathbf{N}^*$ ,  $X_n \hookrightarrow \mathcal{B}(n, p_n)$  avec  $\lim_{n \rightarrow +\infty} np_n = \lambda > 0$ , alors  $(X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X$  où  $X \hookrightarrow \mathcal{P}(\lambda)$  (conditions :  $n \geq 30$  et  $p \leq 0,1$ ).

**Remarque 10.1.** C'est cette approximation qui donne le surnom de loi des phénomènes rares à la loi de POISSON.

**Proposition 10.2. Binomiale par Normale (admis)**

Soit pour tout  $n \in \mathbf{N}^*$ ,  $X_n \hookrightarrow \mathcal{B}(p)$  des variables aléatoires indépendantes, ainsi  $S_n = \sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p)$ . Alors

$$(S_n^*)_{n \geq 1} \xrightarrow{\mathcal{L}} N$$

où  $N \hookrightarrow \mathcal{N}(0, 1)$  avec pour tout  $n \in \mathbf{N}^*$ ,

$$S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}}$$

(conditions  $n \geq 30$ ,  $np \geq 5$  et  $nq \geq 5$ ).

**Proposition 10.3. POISSON par Normale (admis)**

Soit pour tout  $X_n \hookrightarrow \mathcal{P}(\lambda)$  des variables aléatoires indépendantes, ainsi  $S_n = \sum_{k=1}^n X_k \hookrightarrow \mathcal{P}(n\lambda)$ . Alors

$$(S_n^*)_{n \geq 1} \xrightarrow{\mathcal{L}} N$$

où  $N \hookrightarrow \mathcal{N}(0, 1)$  avec pour tout  $n \in \mathbf{N}^*$ ,

$$S_n^* = \frac{S_n - n\lambda}{\sqrt{n\lambda}}$$

(condition  $\lambda \geq 15$ ).

**Remarque 10.2.** Le programme stipule que les conditions d'approximation devront être systématiquement rappelées dans l'énoncé, mais malgré tout il serait bon d'avoir un ordre de grandeur en tête pour les oraux ...



## Estimation inférentielle

### 1. Estimation ponctuelle

La statistique inférentielle est une branche des mathématiques permettant d'estimer des caractéristiques d'une population à partir des données fournies par un échantillon représentatif issue de cette population.

#### 1.1. Généralités

• Soit  $n \in \mathbf{N}^*$  et  $\theta \in \Theta$  où  $\Theta$  est une partie de  $\mathbf{R}$  et  $g : \Theta \rightarrow \mathbf{R}$ . Le but de l'**estimation inférentielle** est celle de l'estimation d'un paramètre  $g(\theta)$  d'une loi suivie par une variable aléatoire  $X$  dont on connaît la loi générique (mais non les paramètres) appartenant à une famille de loi notée  $\mu_\theta$ .

#### Définition 11.1. $n$ -échantillon d'une variable aléatoire

Ⓢⓐ : On appelle un  $n$ -échantillon  $(X_1, \dots, X_n)$  d'une variable aléatoire  $X$  dont on cherche à estimer un ou des paramètres de sa loi (connue), un vecteur aléatoire de dimension  $n$  constitué de variables aléatoires  $X_k$  i.i.d. suivant toutes la loi de  $X$ .

#### Définition 11.2. Statistique ou estimateur d'un paramètre

Ⓢⓐ : Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon de  $X$ , et  $\varphi : \mathbf{R}^n \rightarrow \mathbf{R}$ . On appelle statistique ou estimateur de  $g(\theta)$  toute suite de variables aléatoires  $(T_n)_{n \geq 1}$  où pour tout  $n \in \mathbf{N}^*$ ,  $T_n = \varphi(X_1, \dots, X_n)$  est indépendante de  $\theta$  et dont les valeurs appartiennent à  $g(\Theta)$ .

#### Définition 11.3. Estimation

Ⓢⓐ : On appelle estimation de  $g(\theta)$  toute réalisation de la variable aléatoire  $T_n$  soit

$$T_n(\omega) = \varphi(X_1(\omega), \dots, X_n(\omega)) = \varphi(x_1, \dots, x_n) \text{ où } \omega \in \Omega$$

#### Définition 11.4. Estimateur convergent ou consistant d'un paramètre

Ⓢⓐ : On dit qu'un estimateur est convergent ou consistant lorsque

$$(T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta)$$

**Remarque 11.1.** Dans le langage courant on commet un abus de langage en parlant d'estimateur est convergent alors qu'on considère une suite d'estimateurs.

#### Définition 11.5. Biais d'un estimateur $T_n$ par rapport à $g(\theta)$

Ⓢⓐ : C'est un réel noté  $b_\theta(T_n)$  défini par  $b_\theta(T_n) = \mathbf{E}_\theta(T_n - g(\theta))$ .

▶ Ⓢⓐ : Si  $\forall \theta \in \Theta$ ,  $b_\theta(T_n) = 0$  on dit que l'estimateur est sans biais.

▶ Ⓢⓐ : Si  $\forall \theta \in \Theta$ ,  $\lim_{n \rightarrow +\infty} b_\theta(T_n) = 0$  on dit que l'estimateur est asymptotiquement sans biais.

#### Proposition 11.1. Conditions suffisantes qu'un estimateur soit convergent ou consistant

▶ Ⓢⓐ :  $\left( b_\theta(T_n) = 0 \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}_\theta(T_n) = 0 \right) \Rightarrow \left( (T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta) \right)$

▶ Ⓢⓐ :  $\left( \lim_{n \rightarrow +\infty} b_\theta(T_n) = 0 \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}_\theta(T_n) = 0 \right) \Rightarrow \left( (T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta) \right)$

#### 1.2. Méthodes de construction d'un estimateur ponctuel

• Les principales sont au nombre de trois et apparaissent très régulièrement dans les concours.

- ▶ La méthode du maximum de vraisemblance (HEC II 2020 – EDHEC 2016).
- ▶ La méthode des moments (ECRICOME 2021 – 2022).
- ▶ La méthode des moindres carrés que l'on doit à GAUSS.

**Remarque 11.2.** ► Vous en déduirez que, si je ne me trompe, la dernière méthode n'est pas tombée aux écrits ... mais encore une fois je peux me tromper ...

► Comme ces méthodes ne sont pas aux programmes elles vous seront entièrement guidées. Cependant vous disposez d'une miscule latitude de création d'un estimateur sans biais, obtenu à partir d'un autre biaisé, en profitant de la linéarité de l'espérance.

### 1.3. Exemples d'estimateurs de paramètres usuels et sans biais et convergents.

#### 1.3.1. Moyenne empirique

. Situation : Soit une variable aléatoire  $X$  suivant une loi d'espérance inconnue  $\mu \in \mathbb{R}$  mais de variance  $\sigma^2 \in \mathbb{R}_+^*$  connue.

#### Définition 11.6. Moyenne empirique

Ⓢⓐ : Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon de  $X$  d'espérance inconnue  $X$ , on appelle moyenne empirique la variable aléatoire  $\overline{X}_n$  définie pour tout  $n \in \mathbb{N}^*$  par  $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$  où pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $\mathbf{E}_\mu(X_k) = \mu$  (inconnue).

#### Proposition 11.2. Moments

Ⓢⓐ :  $\forall \mu \in \mathbb{R}$ ,  $\mathbf{E}_\mu(\overline{X}_n) = \mu$  et  $\mathbf{V}_\mu(\overline{X}_n) = \frac{\sigma^2}{n}$ .

#### Proposition 11.3. Loi dans un cas particulier

ⓐ :  $(\forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{N}(\mu, \sigma^2)) \Rightarrow \left( \overline{X}_n \hookrightarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \right)$  où  $\mu \in \mathbb{R}$  et  $\sigma^2 \in \mathbb{R}_+^*$ .

#### Proposition 11.4. Approximation

Ⓢⓐ :  $(\forall k \in \llbracket 1, n \rrbracket, X_k \not\hookrightarrow \mathcal{N}(\mu, \sigma^2)) \Rightarrow \left( \overline{X}_n \underset{\simeq}{\hookrightarrow} \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \right)$  par le théorème central limite pour  $n$  grand où  $\mu \in \mathbb{R}$  et  $\sigma^2 \in \mathbb{R}_+^*$ .

**Proposition 11.5.** Ⓢⓐ :  $(\overline{X}_n)_{n \geq 1} \xrightarrow{\mathbf{P}} \mu$ .

#### 1.3.2. Fréquence empirique

. Situation : Soit une variable aléatoire  $X$  tel que  $X \hookrightarrow \mathcal{B}(p)$  où  $p$  est un réel de  $]0, 1[$  inconnu donc à estimer.

#### Définition 11.7. Fréquence empirique

Ⓢ : Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon de  $X \hookrightarrow \mathcal{B}(p)$  d'espérance inconnue  $p$ . On appelle fréquence empirique la variable aléatoire  $F_n$  définie par

$$\forall n \in \mathbb{N}^*, \quad F_n = \frac{1}{n} \sum_{k=1}^n X_k$$

avec pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $X_k \hookrightarrow \mathcal{B}(p)$  où  $p$  est inconnu.

#### Proposition 11.6. Moments

$\forall n \in \mathbb{N}^*, \forall p \in ]0, 1[, \mathbf{E}_p(F_n) = p$  et  $\mathbf{V}_p(F_n) = \frac{p(1-p)}{n}$ .

#### Proposition 11.7. Approximation

$F_n \underset{\simeq}{\hookrightarrow} \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$  par le théorème central limite pour  $n$  grand,  $np \geq 5$  et  $np(1-p) \geq 5$ .

**Proposition 11.8.**  $(F_n)_{n \geq 1} \xrightarrow{\mathbf{P}} p$ .

#### 1.3.3. Variance empirique

. Situation : Soit une variable aléatoire  $X$  suivant une loi d'espérance connue  $\mu \in \mathbb{R}$  mais de variance inconnue  $\sigma^2 \in \mathbb{R}_+^*$ .

#### Définition 11.8. Variance empirique

Ⓢ : Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon tel que pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $\mathbf{E}_\sigma(X_k) = \mu$  et  $\mathbf{V}_\sigma(X_k) = \sigma^2$  inconnu. On

appelle variance empirique la variable aléatoire  $S_n^2$  définie par

$$\forall n \in \mathbf{N}^*, \quad S_n^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X}_n)^2$$

**Définition 11.9. Variance empirique rectifiée (hors programme)**

Ⓢ : Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon tel que pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $\mathbf{E}_\sigma(X_k) = \mu$  et  $\mathbf{V}_\sigma(X_k) = \sigma^2$  inconnu. On appelle variance empirique rectifiée la variable aléatoire notée  $\widehat{S}_n^2$  définie par

$$\forall n \in \mathbf{N}^*, \quad \widehat{S}_n^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2$$

**Proposition 11.9.** ⓈⓈ :  $\forall n \in \mathbf{N}^*$ ,  $S_n^2 = \frac{1}{n} \sum_{k=1}^n X_k^2 - (\bar{X}_n)^2$  (à savoir retrouver).

**Proposition 11.10.** ⓈⓈ :  $\forall n \in \mathbf{N}^*$ ,  $\forall \sigma \in \mathbf{R}_+^*$ ,  $\mathbf{E}_\sigma(S_n^2) = \left(\frac{n-1}{n}\right) \sigma^2$  (à savoir retrouver).

**Remarque 11.3.** On voit par le résultat précédent que l'estimateur  $S_n^2$  sous-estime en moyenne la variance.

**Proposition 11.11.** ⓈⓈ :  $\left(\left(\frac{n-1}{n}\right) S_n^2\right) \xrightarrow{\mathbf{P}} \sigma^2$  (à savoir retrouver).

**Proposition 11.12.** ⓈⓈ :  $\forall n \in \mathbf{N}^*$ ,  $\forall \sigma \in \mathbf{R}_+^*$ ,  $\mathbf{E}_\sigma(\widehat{S}_n^2) = \sigma^2$  (hors programme mais à savoir retrouver voir exo\*\*\*).

**Proposition 11.13.** ⓈⓈ :  $(\widehat{S}_n^2) \xrightarrow{\mathbf{P}} \sigma^2$  (hors programme mais à savoir retrouver voir exo\*\*\*).

## 2. Compléments hors programmes

**Définition 11.10. Risque quadratique moyen**

ⓈⓈ : C'est le nombre réel noté  $r_\theta(T_n)$  défini par

$$\forall \theta \in \Theta, \quad r_\theta(T_n) = \mathbf{E}_\theta \left( (T_n - g(\theta))^2 \right)$$

**Remarque 11.4.** ⓈⓈ :  $\forall \theta \in \Theta$ ,  $r_\theta(T_n) = \|T_n - g(\theta)\|^2$

**Proposition 11.14.** ⓈⓈ

- ▶  $\forall \theta \in \Theta$ ,  $r_\theta(T_n) = \mathbf{E}_\theta \left( (T_n - g(\theta))^2 \right)$
- ▶  $\forall \theta \in \Theta$ ,  $\left( \lim_{n \rightarrow +\infty} r_\theta(T_n) = 0 \right) \Rightarrow \left( (T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta) \right)$
- ▶  $\forall \theta \in \Theta$ ,  $r_\theta(T_n) = \mathbf{V}_\theta(T_n) + b_\theta^2(T_n)$
- ▶  $\forall \theta \in \Theta$ ,  $(r_\theta(T_n) = \mathbf{V}_\theta(T_n)) \Leftrightarrow (b_\theta(T_n) = 0)$
- ▶  $\left( \lim_{n \rightarrow +\infty} r_\theta(T_n) = 0 \right) \Leftrightarrow \left( \lim_{n \rightarrow +\infty} \mathbf{V}_\theta(T_n) = 0 \text{ et } \lim_{n \rightarrow +\infty} b_\theta(T_n) = 0 \right)$

**Remarque 11.5.** ⓈⓈ : Le meilleur moyen de comparer deux estimateurs sera de prendre celui dont le risque quadratique est le plus petit.

## 3. Estimation par intervalle de confiance

### 3.1. Généralités.

**Définition 11.11. Intervalle de confiance d'un paramètre**

ⓈⓈ : L'intervalle aléatoire  $[U_n, V_n]$  est intervalle de confiance (par excès) de  $g(\theta)$  au niveau de confiance  $1 - \alpha$  (ou de risque  $\alpha$ ) si

$$\mathbf{P}_\theta([U_n, V_n] \ni g(\theta)) \geq 1 - \alpha$$

où  $U_n$  et  $V_n$  sont deux estimateurs  $g(\theta)$ . Sa réalisation est l'estimation de cet intervalle de confiance.

**Remarque 11.6.** ► (*essentielle*) **Attention** par abus d'écriture<sup>1</sup> (de langage !) on voit dans notre programme

$$\mathbf{P}_\theta ([U_n \leq g(\theta) \leq V_n]) \geq 1 - \alpha$$

ce qui est équivalent à la véritable définition ci-dessus, certes, mais elle engendre des confusions lorsque l'on s'exprime. En effet, il est incorrect de dire que : "g(θ) a 95% de chances d'être compris entre 58 et 74". En effet, dans cette dernière écriture, il n'y a rien d'aléatoire ! g(θ) est ou n'est pas dans l'intervalle [58, 74]. La probabilité que g(θ) soit compris entre 60 et 68 est donc 0 ou 1, mais pas 95% ! Alors faites très attention lorsque vous parlez ! C'est pourquoi on parle de confiance qui est la probabilité que l'intervalle contienne effectivement à 100% la vérité.

- Il n'y a pas de notation spécifique d'un intervalle de confiance dans le programme, mais il est d'usage de le noter  $IC_\alpha(g(\theta))$  voire  $IC_{1-\alpha}(g(\theta))$ <sup>2</sup>.
- Si  $[U_n, V_n]$  vérifie  $\mathbf{P}_\theta ([U_n, V_n] \ni g(\theta)) = 1 - \alpha$  alors on parle d'intervalle de confiance exact.

**Exemple 11.1.** Par exemple un intervalle de confiance à 95% est une fourchette autour d'une valeur estimée qui a 95 chances sur 100 d'intercepter la vraie valeur (la vérité).

**Remarque 11.7.** ► L'information apportée par cet intervalle est bien plus riche que celle fournie par l'estimation ponctuelle de g(θ) : plus l'intervalle est étroit, plus précise est l'estimation.

- (*essentielle*) Un intervalle de confiance se construit toujours à partir d'un estimateur ponctuel  $T_n$  à l'aide de plusieurs méthodes : soit à partir de la loi exacte de  $T_n$ , soit à partir de l'inégalité de MARKOV, soit à partir de l'inégalité de BIENAYMÉ-TCHEBYCHEV, soit à partir de la méthode analytique (polynôme), soit à l'aide d'autres inégalités améliorant les deux premières citées (comme les inégalités de CHERNOFF, d'HOEFFDING et bien d'autres ...). Toutes ces méthodes sont déjà tombées au cours des années.
- Si la loi de l'estimateur  $T_n$  est bilatérale, il est raisonnable de choisir un intervalle de confiance symétrique et centré autour d'un certain point. Dans le cas contraire, l'intervalle de confiance doit être construit autour des zones de forte probabilité.
- Si  $\alpha$  augmente (ou si  $n$  augmente), l'amplitude de l'intervalle de confiance diminue et donc la précision augmente.

**Définition 11.12. Intervalle de confiance asymptotique d'un paramètre**

©© : L'intervalle aléatoire  $[U_n, V_n]$  est intervalle de confiance asymptotique de g(θ) au niveau de confiance  $1 - \alpha$  (ou de risque  $\alpha$ ) si il existe une suite  $(\alpha_n)_{n \geq 1}$  telle que  $(\alpha_n)_{n \geq 1}$  soit convergente vers  $\alpha$  et

$$\mathbf{P}_\theta ([U_n, V_n] \ni g(\theta)) \geq 1 - \alpha_n$$

donc

$$\lim_{n \rightarrow +\infty} \mathbf{P}_\theta ([U_n, V_n] \ni g(\theta)) \geq 1 - \alpha$$

où  $U_n$  et  $V_n$  sont deux estimateurs g(θ).

**Remarque 11.8.** ► Un intervalle asymptotique se construit à partir d'une limite via le théorème central limite, par exemple.

- On pourra parfaitement présenter dans une copie ou à l'oral un intervalle de confiance asymptotique sous la forme suivante

$$IC_\alpha(g(\theta)) \underset{n \rightarrow +\infty}{\simeq} [U_n, V_n]$$

### 3.2. Méthodes de construction d'un intervalle de confiance

• Avant toute chose chose, il faut savoir que les différents moyens présentés ci-dessous partent systématiquement de la connaissance de  $T_n$  un estimateur ponctuel de g(θ)

Elles sont au nombre de 5 (la quatrième sera entièrement guidée par l'énoncé).

- Utilisation de l'inégalité de BIENAYMÉ-TCHEBYCHEV. Elle livrera un intervalle par excès et centré dont la définition est

$$\mathbf{P}_\theta ([U_n \leq g(\theta) \leq V_n]) \geq 1 - \alpha$$

- Utilisation de l'inégalité de MARKOV. Elle livrera un intervalle par excès et centré dont la définition est

$$\mathbf{P}_\theta ([U_n \leq g(\theta) \leq V_n]) \geq 1 - \alpha$$

- Utilisation d'autres inégalités. Elle livrera le plus souvent un intervalle par excès et centré.

<sup>1</sup>Car g(θ) est une grandeur inconnue mais non aléatoire, on dit qu'elle est déterministe.

<sup>2</sup>Il n'y a aucune ambiguïté entre  $\alpha$  et  $1 - \alpha$ .

- Utilisation du théorème central limite. Il livrera un intervalle approximé et centré pour  $n$  grand dont la définition est

$$\mathbf{P}_\theta ([U_n \leq g(\theta) \leq V_n]) \simeq 1 - \alpha \text{ pour } n \text{ grand}$$

- Utilisation de la loi de l'estimateur  $T_n$  sans aucune approximation. Elle livrera un intervalle exact centré ou non, au choix. Cependant si la loi de  $T_n$  est à densité paire on préférera un intervalle centré dont la définition est

$$\mathbf{P}_\theta ([U_n \leq g(\theta) \leq V_n]) = 1 - \alpha \text{ pour } n \text{ grand}$$

### 3.3. Estimation par intervalle de confiance d'une proportion.

#### 3.3.1. Par l'inégalité de BIENAYMÉ-TCHEBYCHEV

**Proposition 11.15.** *Nous avons*

$$IC_\alpha(p) = \left[ F_n - \frac{1}{2\sqrt{n\alpha}} \quad ; \quad F_n + \frac{1}{2\sqrt{n\alpha}} \right]$$

#### 3.3.2. Par le théorème central limite

**Proposition 11.16.** *Nous avons*

$$IC_\alpha(p) = \left[ F_n - u_{1-\alpha/2} \sqrt{\frac{F_n(1-F_n)}{n}} \quad ; \quad F_n + u_{1-\alpha/2} \sqrt{\frac{F_n(1-F_n)}{n}} \right]$$

où

$$u_{1-\alpha/2} = \inf \{x \in \mathbb{R} : \Phi(x) \geq 1 - \alpha/2\} = \Phi^{-1}(1 - \alpha/2)$$

et même

$$IC_\alpha(p) \subset \left[ F_n - \frac{u_{1-\alpha/2}}{2\sqrt{n}} \quad ; \quad F_n + \frac{u_{1-\alpha/2}}{2\sqrt{n}} \right]$$

à condition que

$$\min(nu_n, nv_n, n(1-u_n), n(1-v_n)) \geq 5$$

**Remarque 11.9.** *L'inclusion a été obtenue à l'aide de la majoration*

$$F_n(1-F_n) \leq \frac{1}{4}$$

mais celle-ci n'a d'intérêt que si le paramètre à estimer est proche de 1/2 sinon l'intervalle

$$\left[ F_n - \frac{u_{1-\alpha/2}}{2\sqrt{n}} \quad ; \quad F_n + \frac{u_{1-\alpha/2}}{2\sqrt{n}} \right]$$

est trop grossier.

**Définition 11.13.** *Quantile d'ordre  $1 - \alpha/2$*

Le réel  $u_{1-\alpha/2}$  est appelé quantile d'ordre  $1 - \alpha/2$  avec

$$u_{1-\alpha/2} = \inf \{x \in \mathbb{R} : \Phi(x) \geq 1 - \alpha/2\}$$

**Définition 11.14.** *Précision de l'intervalle*

La précision de l'intervalle est donnée par la demi-longueur de l'intervalle (ou rayon) égale à

$$u_{1-\alpha/2} \sqrt{\frac{F_n(1-F_n)}{n}}$$

Si l'on souhaite que cette longueur soit inférieure à une constante  $c$  (à risque  $\alpha$  fixé) cela impose que la taille  $n$  vérifie

$$n \geq u_{1-\alpha/2}^2 \frac{F_n(1-F_n)}{c^2}$$

**Exemple 11.2.** *Si l'on veut diviser par 2 le rayon de l'intervalle, il faut multiplier la taille de l'échantillon par 4.*

### 3.4. Estimation par intervalle de confiance asymptotique d'une moyenne de loi normale de variance donnée.

**Proposition 11.17.** *Nous avons*

$$IC_{\alpha}(\mu) = \left[ \bar{X}_n - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \quad ; \quad \bar{X}_n + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

où

$$u_{1-\alpha/2} = \Phi^{-1}(1 - \alpha/2)$$

à condition que  $n \geq 30$ . Dans le cas où  $\sigma$  est inconnu, on pourra l'estimer par la réalisation de  $\sqrt{\frac{n}{n-1}} S_n$  notée  $s$  dans l'échantillon, ce qui donne

$$IC_{\alpha}(\mu) = \left[ \bar{X}_n - u_{1-\alpha/2} \frac{s}{\sqrt{n}} \quad ; \quad \bar{X}_n + u_{1-\alpha/2} \frac{s}{\sqrt{n}} \right]$$

**Définition 11.15. Précision de l'intervalle**

La précision de l'intervalle est donnée par la demi-longueur de l'intervalle (ou rayon) égale à

$$u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Si l'on souhaite que cette longueur soit inférieure à une constante  $c$  (à risque  $\alpha$  fixé) cela impose que la taille  $n$  vérifie

$$n \geq u_{1-\alpha/2}^2 \frac{\sigma^2}{c^2}$$

**Exemple 11.3.** Si l'on veut diviser par 2 le rayon de l'intervalle, il faut multiplier la taille de l'échantillon par 4.

## Zoologie des lois usuelles du programme

Les lois sont classées par ordre alphabétique.

### 1. Loi de Bernoulli

- ▶ **Notation** :  $X \hookrightarrow \mathcal{B}(p)$  avec  $p \in ]0, 1[$ .
- ▶ **Paramètre** :  $p$
- ▶ **Epreuve type** : épreuve amenant seulement deux issues : soit un succès soit un échec.
- ▶  $X(\Omega) = \{0, 1\}$
- ▶  $\mathbf{P}([X = 0]) = 1 - p$  et  $\mathbf{P}([X = 1]) = p$
- ▶  $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } x \in [0, 1[ \\ 1 & \text{si } x \geq 1 \end{cases}$
- ▶  $\mathbf{E}(X) = p$
- ▶  $\mathbf{V}(X) = p(1 - p)$
- ▶  $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (1 - X \hookrightarrow \mathcal{B}(1 - p))$
- ▶  $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X \hookrightarrow \mathcal{B}(1, p))$
- ▶  $\forall n \geq 1, (X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X^n \hookrightarrow \mathcal{B}(p))$  (on a même  $X^n = X$ )
- ▶  $X_1, \dots, X_n$  i.i.d. ( $\forall k \in [1, n], X_k \hookrightarrow \mathcal{B}(p)$ )  $\Rightarrow \left( \sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p) \right)$ .

#### Définition 12.1. Schéma de Bernoulli d'ordre $n$

On appelle schéma de Bernoulli d'ordre  $n$ , la répétition de  $n$  épreuves de Bernoulli indépendantes et de même paramètre.

### 2. Variable aléatoire indicatrice

**Définition 12.2.**  $\forall A \in \mathcal{A}, \mathbf{1}_A = \begin{cases} 0 & \text{si } \bar{A} \\ 1 & \text{si } A \end{cases}$

**Remarque 12.1.** Par définition une variable aléatoire indicatrice est donc une variable aléatoire de BERNOULLI.

- ▶ **Notation** :  $\mathbf{1}_A$
- ▶ **Paramètre** :  $\mathbf{P}(A)$
- ▶ **Epreuve type** : épreuve de BERNOULLI.
- ▶  $\mathbf{1}_A(\Omega) = \{0, 1\}$
- ▶  $\mathbf{P}([\mathbf{1}_A = 1]) = \mathbf{P}(A)$  et  $\mathbf{P}([\mathbf{1}_A = 0]) = \mathbf{P}(\bar{A})$
- ▶  $\forall x \in \mathbb{R}, F_{\mathbf{1}_A}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \mathbf{P}(A) & \text{si } x \in [0, 1[ \\ 1 & \text{si } x \geq 1 \end{cases}$
- ▶  $\mathbf{E}(\mathbf{1}_A) = \mathbf{P}(A)$
- ▶  $\mathbf{V}(\mathbf{1}_A) = \mathbf{P}(A)(1 - \mathbf{P}(A))$
- ▶  $\mathbf{1}_\Omega = 1$
- ▶  $\forall a \in \mathbb{R}, a = a\mathbf{1}_\Omega$  (ce résultat est une conséquence du précédent)
- ▶  $\mathbf{1}_\emptyset = 0$
- ▶  $\mathbf{1}_{\bar{A}} = 1 - \mathbf{1}_A$
- ▶ Si  $B \subset A \subset E$  alors  $\mathbf{1}_A - \mathbf{1}_B$  et  $\mathbf{1}_{A-B} = \mathbf{1}_A - \mathbf{1}_B$

- ▶  $\mathbf{1}_{A \cap B} = \inf(\mathbf{1}_A, \mathbf{1}_B)$ , (le résultat se généralise par récurrence)
- ▶  $\mathbf{1}_{A \cap B} = \mathbf{1}_A \times \mathbf{1}_B$  (le résultat se généralise par récurrence)
- ▶  $\forall n \geq 1, (\mathbf{1}_A)^n = \mathbf{1}_A$  (ce résultat est une conséquence du précédent)
- ▶  $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A \cap B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \times \mathbf{1}_B$  (la généralisation est hors programme)
- ▶  $\mathbf{1}_{A \cup B} = \sup(\mathbf{1}_A, \mathbf{1}_B)$  (le résultat se généralise par récurrence)
- ▶ Soit  $(A_k)_{k \in [1, n]}$  une suite finie d'événements d'un espace probabilisé deux à deux disjoints, alors

$$\sum_{k=1}^n \mathbf{1}_{A_k} = \mathbf{1}_{\bigsqcup_{k=1}^n A_k}$$

Si les événements  $A_k$  constituent une partition de  $\Omega$  alors

$$\sum_{k=1}^n \mathbf{1}_{A_k} = 1$$

- ▶  $(A \subset B) \Rightarrow (\mathbf{1}_A \leq \mathbf{1}_B)$
- ▶  $\mathbf{1}_{A \Delta B} = |\mathbf{1}_A - \mathbf{1}_B|$  (complément hors programme)

### 3. Loi binomiale

- ▶ **Notation** :  $X \hookrightarrow \mathcal{B}(n, p)$  avec  $n \in \mathbf{N}^*$  et  $p \in ]0, 1[$
- ▶ **Paramètres** :  $n$  et  $p$
- ▶ **Epreuve type** : succession de  $n$  épreuves de BERNOULLI indépendantes et de même paramètre  $p$ .
- ▶  $X(\Omega) = \llbracket 0, n \rrbracket$
- ▶  $\forall k \in \llbracket 0, n \rrbracket, \mathbf{P}([X = k]) = \binom{n}{k} p^k (1-p)^{n-k}$
- ▶  $\mathbf{E}(X) = np$
- ▶  $\mathbf{V}(X) = np(1-p)$
- ▶  $(X \hookrightarrow \mathcal{B}(n, p)) \Leftrightarrow (n - X \hookrightarrow \mathcal{B}(n, 1-p))$
- ▶  $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X \hookrightarrow \mathcal{B}(1, p))$
- ▶  $X_1, \dots, X_n$  i.i.d. ( $\forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{B}(p)$ )  $\Rightarrow \left( \sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p) \right)$
- ▶ **Stabilité de la loi binomiale par convolution**

$$\mathcal{B}(n_1, p) * \dots * \mathcal{B}(n_\ell, p) = \mathcal{B}\left(\sum_{k=1}^{\ell} n_k, p\right)$$

#### ▶ Approximations

- ▷  $(X \hookrightarrow \mathcal{B}(n, p)$  et  $n \geq 30, p < 0, 1) \Rightarrow (X \underset{\simeq}{\hookrightarrow} \mathcal{P}(np))$
- ▷  $\left. \begin{array}{l} X \hookrightarrow \mathcal{B}(n, p) \\ n \geq 30, np \geq 5, np(1-p) \geq 5 \end{array} \right\} \Rightarrow (X \underset{\simeq}{\hookrightarrow} \mathcal{N}(np, np(1-p)))$

### 4. Loi de Dirac (ou loi certaine)

- ▶ **Notation** :  $X \hookrightarrow \delta_c$  où  $c \in \mathbb{R}$
- ▶ **Paramètre** :  $c$
- ▶ **Epreuve type** : numéro associée à une boule tirée d'une urne ne contenant que des boules portant le numéro  $c$ .
- ▶  $X(\Omega) = \{c\}$
- ▶  $\mathbf{P}([X = c]) = 1$
- ▶  $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < c \\ 1 & \text{si } x \geq c \end{cases}$
- ▶  $\mathbf{E}(X) = c$
- ▶  $\mathbf{V}(X) = 0$
- ▶  $(\mathbf{V}(X) = 0) \Leftrightarrow (X \hookrightarrow \delta_c \text{ p.s.})$
- ▶  $(\mathbf{E}(X^2) = 0) \Leftrightarrow (X \hookrightarrow \delta_0 \text{ p.s.})$
- ▶ Toute variable aléatoire certaine est indépendante de toute autre, y compris d'elle-même.

## 5. Loi exponentielle

- ▶ **Notation :**  $X \hookrightarrow \mathcal{E}(\lambda)$  où  $\lambda \in \mathbb{R}_+^*$ .
- ▶ **Paramètre :**  $\lambda$
- ▶ **Epreuve type :** temps d'attente entre deux phénomènes indépendants tels que des arrivées à un guichet, ou des appels téléphoniques ...
- ▶  $X(\Omega) \subset \mathbb{R}_+$
- ▶  $\forall x \in \mathbb{R}, f_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x)$
- ▶  $\forall x \in \mathbb{R}, F_X(x) = (1 - e^{-\lambda x}) \mathbf{1}_{\mathbb{R}_+}(x)$
- ▶  $\mathbf{E}(X) = \frac{1}{\lambda}$
- ▶  $\mathbf{V}(X) = \frac{1}{\lambda^2}$
- ▶  $\forall n \in \mathbb{N}, \mathbf{E}(X^n) = \frac{n!}{\lambda^n}$  (hors programme)
- ▶ **Absence de mémoire**  
 $(X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (\forall (x, y) \in \mathbb{R}_+^2, \mathbf{P}_{[X>x]}([X > x + y]) = \mathbf{P}([X > y]))$ , autrement dit ce qui s'est passé sur l'intervalle de  $]-\infty, x[$  n'affecte en rien ce qui se passera sur l'intervalle  $]x, x + y]$ .
- ▶  $\forall a > 0, (X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (aX \hookrightarrow \mathcal{E}\left(\frac{\lambda}{a}\right))$
- ▶  $\forall \lambda > 0, (X \hookrightarrow \mathcal{E}(1)) \Leftrightarrow \left(\frac{X}{\lambda} \hookrightarrow \mathcal{E}(\lambda)\right)$
- ▶  $\forall \lambda > 0, (X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (\lambda X \hookrightarrow \mathcal{E}(1))$
- ▶  $\mathcal{E}(1) = \gamma(1)$
- ▶  $\underbrace{\mathcal{E}(1) * \dots * \mathcal{E}(1)}_{n \text{ fois}} = \gamma(n)$
- ▶  $\forall \lambda > 0, \underbrace{\mathcal{E}(\lambda) * \dots * \mathcal{E}(\lambda)}_{n \text{ fois}} = \Gamma\left(\frac{1}{\lambda}, n\right)$  (loi d'ERLANG - hors programme)
- ▶ **Attention** contrairement à une légende urbaine, la loi exponentielle n'est pas stable par convolution !

## 6. Loi gamma à un seul paramètre

- ▶ **Notation :**  $X \hookrightarrow \gamma(v)$  avec  $v \in \mathbb{R}_+^*$ .
- ▶ **Paramètre :**  $v$
- ▶  $X(\Omega) \subset \mathbb{R}_+^*$
- ▶  $\forall x \in \mathbb{R}, f_X(x) = \frac{x^{v-1} e^{-x}}{\Gamma(v)} \mathbf{1}_{\mathbb{R}_+^*}(x)$
- ▶  $\mathbf{E}(X) = v$
- ▶  $\mathbf{V}(X) = v$
- ▶  $\gamma(1) = \mathcal{E}(1)$
- ▶  $\underbrace{\mathcal{E}(1) * \dots * \mathcal{E}(1)}_{n \text{ fois}} = \gamma(n)$
- ▶  $\gamma(v_1) * \dots * \gamma(v_n) = \gamma\left(\sum_{k=1}^n v_k\right)$
- ▶  $\forall b \in \mathbb{R}_+^*, (X \hookrightarrow \gamma(v)) \Leftrightarrow (bX \hookrightarrow \Gamma(b, v))$  (hors programme)
- ▶  $\forall b \in \mathbb{R}_+^*, (X \hookrightarrow \Gamma(b, v)) \Leftrightarrow \left(\frac{1}{b}X \hookrightarrow \gamma(v)\right)$  (hors programme)

## 7. Loi géométrique

- ▶ **Notation :**  $X \hookrightarrow \mathcal{G}(p)$  où  $p \in ]0, 1[$ .
- ▶ **Paramètre :**  $p$
- ▶ **Epreuve type :** c'est le rang d'appartition du premier succès lors d'une succession illimitée d'épreuves de BERNOULLI indépendantes et de même paramètre  $p$ .
- ▶  $X(\Omega) = \mathbb{N}^*$

- ▶  $\forall k \in \mathbf{N}^*$ ,  $\mathbf{P}([X = k]) = q^{k-1}p$  avec  $q = 1 - p$ .
- ▶  $\forall k \in \mathbf{N}$ ,  $\mathbf{P}([X > k]) = q^k$
- ▶  $\forall x \in \mathbf{R}$ ,  $F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1 - q^k & \text{si } x \in [k, k + 1[, k \in \mathbf{N}^* \end{cases}$
- ▶  $\mathbf{E}(X) = \frac{1}{p}$
- ▶  $\mathbf{V}(X) = \frac{q}{p^2}$
- ▶ **Absence de mémoire**  
 $(X \hookrightarrow \mathcal{G}(p)) \Leftrightarrow (\forall (n, m) \in (\mathbf{N})^2, \mathbf{P}_{[X > n]}([X > n + m]) = \mathbf{P}([X > m]))$  ce qui s'est passé sur l'intervalle  $]-\infty, n]$  n'affecte en rien ce qui se passera sur l'intervalle  $]n, n + m]$ .

## 8. Lois normales

### 8.1. Loi normale centrée et réduite ou de Laplace-Gauss centrée et réduite.

- ▶ **Notation** :  $X \hookrightarrow \mathcal{N}(0, 1)$
- ▶ **Paramètres** : 0 et 1
- ▶  $X(\Omega) \subset \mathbf{R}$
- ▶  $\forall x \in \mathbf{R}$ ,  $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
- ▶ **Intégrale de GAUSS**
  - ▷  $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1$
  - ▷  $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$
- ▶ **Fonction de répartition**
  - ▷  $\forall x \in \mathbf{R}$ ,  $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$
  - ▷  $\forall x \in \mathbf{R}$ ,  $\Phi(-x) = 1 - \Phi(x)$
  - ▷  $\Phi(0) = 1/2$
- ▶ **Mode** : 0
- ▶ **Médiane** : 0
- ▶  $\mathbf{E}(X) = 0$
- ▶  $\mathbf{V}(X) = 1$
- ▶  $\forall n \in \mathbf{N}$ ,  $\mathbf{E}(X^n) = \begin{cases} 0 & \text{si } n \in 2\mathbf{N} + 1 \\ \frac{(2m)!}{2^m m!} & \text{si } n = 2m \in 2\mathbf{N} \end{cases}$  (hors programme)
- ▶ **Stabilité de la loi normale centrée réduite par convolution**

$$\underbrace{\mathcal{N}(0, 1) * \cdots * \mathcal{N}(0, 1)}_{n \text{ fois}} = \mathcal{N}(0, n)$$

- ▶ **Approximation**

$$\left( (X_n)_{n \geq 1} \text{ i.i.d.} \right) \Rightarrow \left( \frac{\sum_{k=1}^n X_k - \mathbf{E}(\sum_{k=1}^n X_k)}{\sqrt{\mathbf{V}(\sum_{k=1}^n X_k)}} \right)_{n \geq 1} \xrightarrow{\mathcal{L}} N \quad \text{où } N \hookrightarrow \mathcal{N}(0, 1)$$

via le théorème central limite.

### 8.2. Loi normale quelconque ou de Laplace-Gauss.

- ▶ **Notation** :  $X \hookrightarrow \mathcal{N}(m, \sigma^2)$  où  $m \in \mathbf{R}$  et  $\sigma \in \mathbf{R}_+^*$ .
- ▶ **Paramètres** :  $m$  et  $\sigma^2$
- ▶  $X(\Omega) \subset \mathbf{R}$
- ▶  $\forall x \in \mathbf{R}$ ,  $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$
- ▶ **Mode** :  $m$
- ▶ **Médiane** :  $m$

- ▶  $\mathbf{E}(X) = m$
- ▶  $\mathbf{V}(X) = \sigma^2$
- ▶ **Théorème fondamental de la loi normale**

$$(X \hookrightarrow \mathcal{N}(m, \sigma^2)) \Leftrightarrow \left( \frac{X - m}{\sigma} \hookrightarrow \mathcal{N}(0, 1) \right)$$

- ▶ **Stabilité de la loi normale par convolution**

$$\mathcal{N}(m_1, \sigma_1^2) * \dots * \mathcal{N}(m_n, \sigma_n^2) = \mathcal{N}\left(\sum_{k=1}^n m_k, \sum_{k=1}^n \sigma_k^2\right)$$

## 9. Loi de Poisson

- ▶ **Notation** :  $X \hookrightarrow \mathcal{P}(\lambda)$  où  $\lambda \in \mathbb{R}_+^*$ .
- ▶ **Paramètre** :  $\lambda$
- ▶ **Epreuve type** : nombre d'apparitions d'un phénomène rare durant un intervalle de temps donné.
- ▶  $X(\Omega) = \mathbb{N}$
- ▶  $\forall k \in \mathbb{N}, \mathbf{P}([X = k]) = \frac{e^{-\lambda} \lambda^k}{k!}$
- ▶  $\mathbf{E}(X) = \lambda$
- ▶  $\mathbf{V}(X) = \lambda$
- ▶ **Stabilité de la loi de POISSON par convolution**

$$\mathcal{P}(\lambda_1) * \dots * \mathcal{P}(\lambda_n) = \mathcal{P}\left(\sum_{k=1}^n \lambda_k\right)$$

- ▶ Raikov a démontré que si  $X_1, \dots, X_n$  sont  $n$  variables aléatoires indépendantes pour lesquelles

$$S_n = \sum_{k=1}^n X_k \text{ suit une loi de POISSON, alors chacune des variables aléatoires } X_k \text{ suit une loi de POISSON}$$

(hors programme). On dit que la loi de POISSON est indéfiniment divisible, c'est la seule à valeurs dans  $\mathbb{N}$ .

- ▶ Si  $X \hookrightarrow \mathcal{P}(\lambda)$  alors  $\mathbf{P}([X = 0]) = 1 - F_Y(1)$  où  $Y \hookrightarrow \mathcal{E}(\lambda)$  (hors programme à savoir retrouver).
- ▶ Soit  $X \hookrightarrow \mathcal{P}(\lambda x)$  et  $Y \hookrightarrow \mathcal{E}(\lambda)$  alors  $\mathbf{P}([X > x]) = \mathbf{P}([Y = 0])$  (hors programme à savoir retrouver).
- ▶ **Approximation**

$$\left. \begin{array}{l} X \hookrightarrow \mathcal{P}(\lambda) \\ \lambda \geq 15 \end{array} \right\} \Rightarrow \left( X \underset{\simeq}{\hookrightarrow} \mathcal{N}(\lambda, \lambda) \right)$$

## 10. Loïs uniformes

### 10.1. Loi uniforme discrète.

- ▶ **Notation** :  $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$
- ▶ **Paramètre** :  $\llbracket 1, n \rrbracket$
- ▶ **Epreuve type** : numéro d'une boule tirée à partir d'une urne constituée de boules numérotées de 1 à  $n$ .
- ▶  $X(\Omega) = \llbracket 1, n \rrbracket$
- ▶  $\forall k \in \llbracket 1, n \rrbracket, \mathbf{P}([X = k]) = \frac{1}{n}$
- ▶  $\mathbf{E}(X) = \frac{n+1}{2}$
- ▶  $\mathbf{V}(X) = \frac{n^2-1}{12}$
- ▶ Soit  $(a, b) \in \mathbb{R}^2$  avec  $a < b$  alors

$$(X \hookrightarrow \mathcal{U}(\llbracket a, b \rrbracket)) \Leftrightarrow (X - a + 1 \hookrightarrow \mathcal{U}(\llbracket 1, b - a + 1 \rrbracket))$$

$$\text{et } \mathbf{E}(X) = \frac{a+b}{2} \text{ et } \mathbf{V}(X) = \frac{(b-a+1)^2-1}{12}.$$

- ▶ **Attention** contrairement à une légende urbaine, la loi uniforme n'est pas stable par convolution !

**10.2. Loi uniforme continue.**

- ▶ **Notation :**  $X \hookrightarrow \mathcal{U}(I)$  où  $I$  est un intervalle d'extrémités  $a$  et  $b$  deux réels tel que  $a < b$ .
- ▶ **Paramètre :**  $I$
- ▶  $X(\Omega) \subset I$
- ▶  $\forall x \in \mathbb{R}, f_X(x) = \left(\frac{1}{b-a}\right) \mathbf{1}_I(x)$
- ▶  $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}$
- ▶  $\mathbf{E}(X) = \frac{a+b}{2}$
- ▶  $\mathbf{V}(X) = \frac{(b-a)^2}{12}$
- ▶ Si  $U \hookrightarrow \mathcal{U}([0, 1])$  alors  $Y = -\frac{\ln(1-U)}{\lambda} \hookrightarrow \mathcal{E}(\lambda)$  (*hors programme mais à connaître ...*)
- ▶  $(X \hookrightarrow \mathcal{U}([0, 1])) \Leftrightarrow (\forall (a, b) \in (\mathbb{R}_+^*)^2, a + (b-a)X \hookrightarrow \mathcal{U}([a, b]))$
- ▶  $(X \hookrightarrow \mathcal{U}([a, b])) \Leftrightarrow \left(\frac{X-a}{b-a} \hookrightarrow \mathcal{U}([0, 1])\right)$
- ▶  $(X \hookrightarrow \mathcal{U}([0, 1])) \Leftrightarrow \left(-\frac{1}{\lambda} \ln(X) \hookrightarrow \mathcal{E}(\lambda)\right)$  (*hors programme*)
- ▶ **Attention** contrairement à une légende urbaine, la loi uniforme n'est pas stable par convolution !

**11. Bilan des lois du programme stables par convolution**

- ▶  $\mathcal{B}(n_1, p) * \mathcal{B}(n_2, p) = \mathcal{B}(n_1 + n_2, p)$
- ▶  $\mathcal{P}(\lambda_1) * \mathcal{P}(\lambda_2) = \mathcal{P}(\lambda_1 + \lambda_2)$
- ▶  $\gamma(v_1) * \gamma(v_2) = \gamma(v_1 + v_2)$
- ▶  $\mathcal{N}(m_1, \sigma_1^2) * \mathcal{N}(m_2, \sigma_2^2) = \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

**12. Bilan des lois sans mémoire**

- ▶ **Ⓢ :** Les seules variables aléatoires à valeurs dans  $\mathbb{N}^*$  qui ont la propriété d'être sans mémoire, c'est-à-dire vérifiant

$$\forall (n, m) \in (\mathbb{N})^2, \mathbf{P}([X > n + m]) = \mathbf{P}([X > m]) \mathbf{P}([X > n])$$

sont les variables aléatoires géométriques.

- ▶ **Ⓢ :** Les seules variables aléatoires à valeurs dans  $\mathbb{R}_+$  qui ont la propriété d'être sans mémoire, c'est-à-dire vérifiant

$$\forall (s, t) \in (\mathbb{R}_+)^2, \mathbf{P}([X > s + t]) = \mathbf{P}([X > t]) \mathbf{P}([X > s])$$

sont les variables aléatoires exponentielles.

## Outils

### 1. Les identités basiques

Soit  $n$  un entier naturel non nul,

- ▶  $\sum_{k=1}^n k = \frac{n(n+1)}{2}$
- ▶  $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$
- ▶  $\sum_{k=1}^n k^3 = \left(\frac{n(n+1)}{2}\right)^2$

**Théorème 13.1.** *Formule du binôme de NEWTON*

$$\forall n \in \mathbf{N}, \quad \forall (a, b) \in \mathbf{R}^2, \quad (a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^n \binom{n}{k} b^k a^{n-k}$$

**Propriétés 13.1.** ▶  $\forall (n, p) \in \mathbf{N}^2, 1 \leq p \leq n,$   $\sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} 1 = \binom{n}{p}$  (*hors programme*)

- ▶  $\forall n \in \mathbf{N}^*, \left(\sum_{k=1}^n x_k\right)^2 = \sum_{k=1}^n x_k^2 + 2 \sum_{1 \leq i < j \leq n} x_i x_j$
- ▶  $\forall n \in \mathbf{N}^*, \left(\sum_{k=1}^n x_k\right)^3 = \sum_{k=1}^n x_k^3 + 3 \sum_{j \neq k} x_j^2 x_k + \sum_{j \neq k \neq \ell} x_j x_k x_\ell$  (*hors programme très utile en statistique*)
- ▶  $\forall n \in \mathbf{N}^*, \left(\sum_{k=1}^n x_k\right)^4 = \sum_{k=1}^n x_k^4 + 4 \sum_{j \neq k} x_j x_k^3 + 3 \sum_{j \neq k} x_j^2 x_k^2 + 6 \sum_{i \neq j \neq k} x_i^2 x_j x_k + \sum_{i \neq j \neq k \neq \ell} x_i x_j x_k x_\ell$  (*hors programme mais très utile en statistique*)

**Propriétés 13.2.** *Identité de BERNOULLI*

$$\forall (a, b) \in \mathbf{R}^2, \quad \forall n \in \mathbf{N}, \quad a^{n+1} - b^{n+1} = (a - b) \sum_{k=0}^n a^{n-k} b^k = (a - b) \sum_{k=0}^n a^k b^{n-k}$$

On note  $(X_i)_{1 \leq i \leq n}$  une suite de variables aléatoires d'un même espace probabilisé. On pose  $S_n = X_1 + \dots + X_n$ , alors on a

- ▶  $\mathbf{E}(S_n^2) = \sum_{i=1}^n \mathbf{E}(X_i^2) + 2 \sum_{1 \leq i < j \leq n} \mathbf{E}(X_i X_j)$
- ▶  $(\mathbf{E}(S_n))^2 = \sum_{i=1}^n (\mathbf{E}(X_i))^2 + 2 \sum_{1 \leq i < j \leq n} \mathbf{E}(X_i) \mathbf{E}(X_j)$
- ▶  $\mathbf{V}(S_n) = \mathbf{E}(S_n^2) - (\mathbf{E}(S_n))^2 = \sum_{i=1}^n \mathbf{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$

## 2. Quelques sommes que l'on peut retrouver dans les sujets

- ▶  $\forall x \in \mathbb{R}, \sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x$
- ▶  $\sum_{n=0}^{+\infty} t^n = \frac{1}{1-t}$  où  $|t| < 1$
- ▶  $\sum_{n=0}^{+\infty} (-t)^n = \frac{1}{1+t}$  où  $|t| < 1$
- ▶  $\sum_{n=0}^{+\infty} \frac{t^n}{n} = -\ln(1-t)$  où  $t \in [-1, 1[$  (*hors programme*)
- ▶  $\sum_{n=0}^{+\infty} \binom{n+k-1}{k-1} t^n = \frac{1}{(1-t)^k}$  où  $|t| < 1$
- ▶  $\forall x \in \mathbb{R}, e^{tx} + e^{-tx} = 2 \sum_{n=0}^{+\infty} \frac{(tx)^{2n}}{(2n)!}$  (*à savoir retrouver*)
- ▶  $\forall x \in \mathbb{R}, e^{tx} - e^{-tx} = 2 \sum_{n=0}^{+\infty} \frac{(tx)^{2n+1}}{(2n+1)!}$  (*à savoir retrouver*)

## 3. Calcul intégral

- ▶  $\Gamma : x \in \mathbb{R}_+^* \mapsto \int_0^{+\infty} t^{x-1} e^{-t} dt$
- ▶  $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} dx = 1$  (*intégrale de GAUSS - version 1*)
- ▶  $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$  (*intégrale de GAUSS - version 2*)
- ▶  $\frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-x^2/2} dx = \frac{1}{2}$  (*intégrale de GAUSS et parité de l'intégrande - version 1*)
- ▶  $\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$  (*intégrale de GAUSS et parité de l'intégrande - version 2*)
- ▶  $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-x^2/2} dx = 0$  (*espérance d'une variable aléatoire de loi  $\mathcal{N}(0, 1)$* )
- ▶  $\frac{1}{\sqrt{2\pi}} \int_0^{+\infty} x^2 e^{-x^2/2} dx = 1$  (*moment d'ordre deux d'une variable aléatoire de loi  $\mathcal{N}(0, 1)$* )
- ▶  $\forall n \in \mathbb{N}, \int_0^{+\infty} x^n e^{-x} dx = n!$  (*valeur de  $\Gamma(n+1)$* )
- ▶  $\int_0^{+\infty} x e^{-x} dx = 1$  (*espérance d'une variable aléatoire de loi  $\mathcal{E}(1)$* )

## 4. Les produits finis

Soit  $a \in \mathbb{R}$  et  $(p, n, m) \in \mathbb{N}^3$ ,  $p \leq n$  et  $(x_k)_{1 \leq k \leq n}$ ,  $(y_k)_{1 \leq k \leq n}$  deux familles de réels.

**Définition 13.1. Produit fini** ▶  $\forall (n, p) \in \mathbb{N}^2$ ,  $p \leq n$ ,  $\forall (a_p, \dots, a_n) \in \mathbb{R}^{n-p+1}$ ,  $\prod_{k=p}^n a_k = a_p \times \dots \times a_n$ .

▶ On pose  $\prod_{k=p}^n a_k = 1$  lorsque  $p > n$ .

### 4.1. Propriétés principales.

**Propriétés 13.3.** ▶  $\prod_{k=p}^n a = a^{n-p+1}$

- ▶  $\prod_{k=p}^n a^k = a^{\sum_{k=p}^n k}$
- ▶  $\prod_{k=p}^n ax_k = a^{n-p+1} \prod_{k=p}^n x_k$
- ▶ Pour toute permutation  $\sigma$  de  $\llbracket p, n \rrbracket$  on a  $\prod_{k=p}^n x_{\sigma(k)} = \prod_{k=p}^n x_k$ .
- ▶  $\prod_{k=1}^{n+p} x_k = \left( \prod_{k=1}^n x_k \right) \left( \prod_{k=n+1}^{n+p} x_k \right)$  (associativité)
- ▶  $\left( \prod_{k=p}^n x_k \right) \times \left( \prod_{k=p}^n y_k \right) = \prod_{k=p}^n (x_k y_k)$
- ▶  $\left( \prod_{k=p}^n x_k \right)^m = \prod_{k=p}^n (x_k)^m$
- ▶  $\frac{\prod_{k=1}^n x_k}{\prod_{k=1}^n y_k} = \prod_{k=1}^n \left( \frac{x_k}{y_k} \right)$  (où tous les réels  $y_k \neq 0$ )
- ▶  $\prod_{k=p}^n \frac{x_{k+1}}{x_k} = \frac{x_{n+1}}{x_p}$  (où tous les réels  $x_k \neq 0$ ) (téléscopage)

**4.2. Permutation des produits.** Soit  $(a_{i,j})_{1 \leq i, j \leq n}$  une famille de nombres complexes

**Propriétés 13.4.** ▶  $\prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{i,j} = \prod_{i=1}^n \prod_{j=1}^n a_{i,j} = \prod_{j=1}^n \prod_{i=1}^n a_{i,j}$

- ▶  $\prod_{1 \leq i < j \leq n} a_{i,j} = \prod_{j=1}^n \prod_{i=1}^j a_{i,j} = \prod_{i=1}^n \prod_{j=i}^n a_{i,j}$
- ▶  $\prod_{1 \leq i < j \leq n} a_{i,j} = \prod_{j=2}^n \prod_{i=1}^{j-1} a_{i,j} = \prod_{i=1}^{n-1} \prod_{j=i+1}^n a_{i,j}$

**4.3. Liens entre produit et somme.**

**Propriétés 13.5.** ▶  $\forall n \in \mathbb{N}^*, \forall k \in \{1, \dots, n\}, x_k > 0, \ln \left( \prod_{k=1}^n x_k \right) = \sum_{k=1}^n \ln(x_k)$

- ▶  $\forall n \in \mathbb{N}^*, \forall (x_1, \dots, x_n) \in \mathbb{R}^n, \exp \left( \sum_{k=1}^n x_k \right) = \prod_{k=1}^n \exp(x_k)$
- ▶ **Attention :**  $\prod_i \sum_j a_{ij} \neq \sum_j \prod_i a_{ij}$

**4.4. Quelques produits particuliers.**

- ▶  $\prod_{k=1}^n k = n!$
- ▶  $\forall n \in \mathbb{N}^*, \prod_{k=1}^n (2k) = 2^n n!$  (résultat à savoir retrouver)
- ▶  $\forall n \in \mathbb{N}^*, \prod_{k=1}^n (2k+1) = \frac{(2n+1)!}{2^n n!}$

### 5. Deux équivalents simples

- ▶  $\binom{n}{k} \underset{n \rightarrow +\infty}{\sim} \frac{n^k}{k!}$  où  $k$  est entier *fixé* (résultat à savoir retrouver)
- ▶  $e^{-n} \sum_{k=0}^n \frac{n^k}{k!} \underset{n \rightarrow +\infty}{\sim} \frac{1}{2}$

### 6. Résultats divers

- ▶  $\forall p \in \mathbf{N}^*, \forall k \in \llbracket 1, p \rrbracket, |x|^k \leq |x|^p + 1$
- ▶  $\forall (x, y) \in \mathbf{R}^2, |xy| \leq \frac{x^2 + y^2}{2}$
- ▶  $\forall x \in \mathbf{R}, |x| \leq \frac{x^2 + 1}{2}$
- ▶  $\forall x \in \mathbf{R}, \lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$
- ▶  $\forall (x, y) \in \mathbf{R}^2, \max(x, y) + \min(x, y) = x + y$   
(**Attention** ce résultat ne se généralise pas en  $\max(x_1, \dots, x_n) + \min(x_1, \dots, x_n) = x_1 + \dots + x_n$ )
- ▶  $\forall (x, y) \in \mathbf{R}^2, \max(x, y) - \min(x, y) = |x - y|$

## Compléments hors programme mais ...

### 1. Formule de Vandermonde

**Proposition 14.1.** Pour tout triplet  $(n_1, n_2, n) \in \mathbf{N}^3$  tel que  $n \leq n_1 + n_2$  alors

$$\sum_{k=0}^n \binom{n_1}{k} \binom{n_2}{n-k} = \binom{n_1 + n_2}{n}$$

**Corollaire 14.1.**  $\forall n \in \mathbf{N}, \sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}$

**Théorème 14.1. Formule de VANDERMONDE généralisée**

$$\forall (N_1, \dots, N_k) \in \mathbf{N}^k, \sum_{\substack{(n_1, \dots, n_k) \in [0, n]^k \\ n_1 + \dots + n_k = n}} \binom{N_1}{n_1} \times \dots \times \binom{N_k}{n_k} = \binom{N_1 + \dots + N_k}{n}$$

avec  $n \leq N_1 + \dots + N_k$ .

### 2. Le coefficient multinomial

**Définition 14.1.**  $\binom{n}{n_1, \dots, n_p} = \frac{n!}{n_1! \times \dots \times n_p!}$  avec  $n_1 + \dots + n_p = n$ .

**Propriétés 14.1.**  $\binom{n}{n_1, \dots, n_p} = \binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \binom{n-n_1-n_2}{n_3} \times \dots \times \binom{n-n_1-\dots-n_{p-1}}{n_p}$  avec  $n_1 + \dots + n_p = n$ .

**Théorème 14.2. Formule multinomiale**

$$\forall (a_1, \dots, a_p) \in \mathbf{R}^p, \quad \forall n \in \mathbf{N}, \quad (a_1 + \dots + a_p)^n = \sum_{\substack{(n_1, \dots, n_p) \in \mathbf{N}^p \\ n_1 + \dots + n_p = n}} \underbrace{\frac{n!}{n_1! \times \dots \times n_p!}}_{\text{coef}^{\text{t}} \text{ multinomial}} \times a_1^{n_1} \times \dots \times a_p^{n_p}$$

### 3. Relation min et max

**Proposition 14.2.** Pour toute liste  $(x_1, \dots, x_n) \in \mathbf{R}^n$ ,

$$\max(x_1, \dots, x_n) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \min(x_{i_1}, \dots, x_{i_k})$$

### 4. Urne muticolore

**Proposition 14.3.** Soit  $N$  le nombre de tirages de  $n$  boules dans une urne contenant  $n_1$  boules de couleurs  $C_1, \dots, n_k$  boules de couleurs  $C_k$ . On cherche le nombre de tirages avec  $c_1$  boules de couleur  $C_1, \dots, c_k$  boules de couleur  $C_k$ .

► Si les tirages sont non ordonnées et successifs sans remise ou simultanés, alors

$$N = \binom{n_1}{c_1} \binom{n_2}{c_2} \times \dots \times \binom{n_k}{c_k}$$

avec  $n = c_1 + \dots + c_k$ .

► Si les tirages sont non ordonnées et successifs avec remise, alors

$$N = \binom{n}{c_1} \binom{n-c_1}{c_2} \times \dots \times \binom{n-c_1-\dots-c_{k-1}}{c_k} \times n_1^{c_1} \times n_2^{c_2} \times \dots \times n_k^{c_k}$$

avec  $n = c_1 + \dots + c_k$ .

### 5. Espérance et fonction de survie

**Proposition 14.4.** ►  $\textcircled{S}$  : Soit  $X$  une variable aléatoire à valeur dans  $\mathbf{N}$  alors

$$(\mathbf{E}(X) < \infty) \Leftrightarrow \left( \sum_{n \geq 0} \mathbf{P}([X > n]) < \infty \right)$$

En cas de convergence,

$$\mathbf{E}(X) = \sum_{n=0}^{+\infty} \mathbf{P}([X > n])$$

►  $\textcircled{C}$  : Soit  $X$  une variable aléatoire à valeur dans  $\mathbf{R}_+$  de fonction de répartition  $F$  alors

$$(\mathbf{E}(X) < \infty) \Leftrightarrow \left( \int_0^{+\infty} \mathbf{P}([X > x]) < \infty \right)$$

En cas de convergence,

$$\mathbf{E}(X) = \int_0^{+\infty} \mathbf{P}([X > x]) = \int_0^{+\infty} (1 - F(x)) dx$$

**Proposition 14.5.** Pour toute variable aléatoire discrète  $X$ , on a

$$\lim_{x \rightarrow +\infty} \mathbf{P}([|X| > x]) = 0$$

### 6. Variables aléatoires de signe fixé et indicatrices

**Proposition 14.6.**  $\textcircled{S}\textcircled{C}$

► Pour tout  $X$  variable aléatoire positive on a

$$X = X \mathbf{1}_{[X > 0]}$$

► Pour tout  $X$  variable aléatoire négative on a

$$X = X \mathbf{1}_{[X < 0]}$$

**Remarque 14.1.** L'intérêt de la proposition précédente est de pouvoir appliquer l'inégalité de CAUCHY-SCHWARZ au produit  $X \mathbf{1}_{[X > 0]}$  et au produit  $X \mathbf{1}_{[X < 0]}$ , sous réserve que  $\mathbf{E}(X^2) < \infty$ .

### 7. Interprétation géométrique de la corrélation

**Définition 14.2.**  $\textcircled{S}\textcircled{C}$  : Soit  $\langle \bullet, \bullet \rangle$  le produit scalaire et  $\|\bullet\|$  la norme de  $L^2(\Omega, \mathcal{A}, \mathbf{P})$ , alors

►  $\text{Cov}(X, Y) = \langle X - \mathbf{E}(X), Y - \mathbf{E}(Y) \rangle$

►  $\rho(X, Y) = \frac{\langle X - \mathbf{E}(X), Y - \mathbf{E}(Y) \rangle}{\|X - \mathbf{E}(X)\| \|Y - \mathbf{E}(Y)\|}$

Autrement dit,  $\rho(X, Y)$  est le "cosinus de l'angle" entre  $X - \mathbf{E}(X)$  et  $Y - \mathbf{E}(Y)$ . Donc  $\rho(X, Y) = \pm 1$  signifie que  $X - \mathbf{E}(X)$  et  $Y - \mathbf{E}(Y)$  sont colinéaires : il existe un réel  $a$  non nul tel que

$$Y - \mathbf{E}(Y) = a(X - \mathbf{E}(X))$$

### 8. Minimisation de la variance

**Proposition 14.7.**  $\textcircled{S}\textcircled{C}$  :  $\mathbf{V}(X) = \inf_{a \in \mathbf{R}} \mathbf{E}((X - a)^2)$  (à savoir retrouver)

### 9. Espérance conditionnelle

**Proposition 14.8.**  $\textcircled{S} : \forall B \in \mathcal{A}, \mathbf{P}(B) \neq 0, \mathbf{E}(X | B) = \frac{\mathbf{E}(X1_B)}{\mathbf{P}(B)}$  (à savoir retrouver)

**Proposition 14.9.**  $\textcircled{S} : \forall y \in \mathcal{S}_Y, \mathbf{E}(X | [Y = y]) = \frac{\mathbf{E}(X1_{[Y=y]})}{\mathbf{P}([Y = y])}$  (à savoir retrouver)

### 10. La loi gamma à deux paramètres

**Définition 14.3.** Soit  $b$  et  $v$  deux réels strictement positifs. On dit que  $X$  une variable aléatoire à densité sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$  suit la loi gamma à deux paramètres de paramètres  $b$  et  $v$ , ce qu'on note  $X \hookrightarrow \Gamma(b, v)$ , si et seulement si une densité  $f$  associée à  $X$  est définie par

$$f_X(x) = \begin{cases} \frac{e^{-\frac{x}{b}} x^{v-1}}{\Gamma(v) b^v} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

**Remarque 14.2.**  $\blacktriangleright$  La distribution gamma est une généralisation de la loi exponentielle. En effet si la loi exponentielle correspond à la distribution de probabilité du temps séparant l'apparition de deux événements donnés, la loi gamma fournit la distribution du temps qui s'écoule entre la  $k^{\text{ème}}$  et la  $(k+r)^{\text{ème}}$  apparition de l'événement. La loi gamma est appliquée comme modèle de probabilité pour prévoir la durée de vie des appareils qui subissent une usure.

$\blacktriangleright$  La loi  $\Gamma(1, v)$  est dite **loi gamma à un paramètre  $v$** .

**Proposition 14.10.** La loi gamma et la loi exponentielle

$$\forall b \in \mathbf{R}_+^*, \Gamma(b, 1) = \mathcal{E}\left(\frac{1}{b}\right)$$

**Proposition 14.11.** Lien entre les lois  $\gamma$  et  $\Gamma$

Soit  $b, b', v$  trois réels strictement positifs alors on a les équivalences

$$\begin{aligned} (X \hookrightarrow \gamma(v)) &\Leftrightarrow (bX \hookrightarrow \Gamma(b, v)) \\ (X \hookrightarrow \Gamma(b, v)) &\Leftrightarrow \left(\frac{1}{b}X \hookrightarrow \gamma(v)\right) \\ (X \hookrightarrow \Gamma(b, v)) &\Leftrightarrow (b'X \hookrightarrow \Gamma(b'b, v)) \end{aligned}$$

**Remarque 14.3.** En général on ne sait pas expliciter la fonction de répartition sans intégrale. En revanche c'est possible quand  $t$  est un entier naturel, surtout petit, où l'on procède par intégration par parties.

**Proposition 14.12.** Si  $X \hookrightarrow \Gamma(b, v)$  alors

$$\mathbf{E}(X) = bv \quad \text{et} \quad \mathbf{V}(X) = b^2v$$

**Proposition 14.13.** Stabilité de la loi gamma pour la convolution

(1) (**Cas de deux variables aléatoires**) Soit  $X, Y$  deux variables aléatoires à densité indépendantes alors

$$(X \hookrightarrow \Gamma(b, v_1), Y \hookrightarrow \Gamma(b, v_2)) \Rightarrow (X + Y \hookrightarrow \Gamma(b, v_1 + v_2))$$

(2) (**Cas de  $n$  variables aléatoires**) On peut généraliser le résultat précédent par récurrence sur  $n \in \mathbf{N}^*$ . Soit  $X_1, \dots, X_n$ ,  $n$  variables aléatoires à densité indépendantes définies sur un même espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$  telles que pour tout entier  $k$  de  $\llbracket 1, n \rrbracket$ ,  $X_k \hookrightarrow \Gamma(b, v_k)$ . Alors

$$\sum_{k=1}^n X_k \hookrightarrow \Gamma\left(b, \sum_{k=1}^n v_k\right)$$

**Proposition 14.14.** Cas particulier important : somme de variables aléatoires exponentielles indépendantes

Soit  $X_1, \dots, X_n$ ,  $n$  variables aléatoires à densité **indépendantes** sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$  telles que pour tout entier  $k$  de  $\llbracket 1, n \rrbracket$ ,  $X_k \hookrightarrow \mathcal{E}(\lambda)$ . Alors

$$\sum_{k=1}^n X_k \hookrightarrow \Gamma\left(\frac{1}{\lambda}, n\right) \quad \text{loi d'ERLANG}$$

- Remarque 14.4.** (1) *Ne parlez surtout pas de stabilité de la loi exponentielle, c'est rédhibitoire!!!*  
 (2) *La loi d'ERLANG modélise les temps d'arrivée et de service dans les files d'attente de système de communication.*

### 11. Loi exponentielle

**Proposition 14.15.** © : Soit  $X \hookrightarrow \mathcal{E}(\lambda)$  et  $Y \hookrightarrow \mathcal{E}(\mu)$ , tel que  $X \perp\!\!\!\perp Y$  alors

- ▶  $\mathbf{P}([X \leq Y]) = \frac{\lambda}{\lambda + \mu}$
- ▶  $\mathbf{P}([Y \leq X]) = \frac{\mu}{\lambda + \mu}$
- ▶  $\mathbf{P}([X = Y]) = 0$
- ▶  $\inf(X, Y) \hookrightarrow \mathcal{E}(\lambda + \mu)$  (ce qui se généralise par récurrence, il n'y a rien de remarquable à dire concernant la loi de  $\sup(X, Y)$ )
- ▶  $\mathbf{P}([X \leq Y]) = \frac{\lambda}{\lambda + \mu}$
- ▶ © :  $\mathcal{E}(\lambda) = \Gamma\left(\frac{1}{\lambda}, 1\right)$
- ▶ © :  $\underbrace{\mathcal{E}(\lambda) * \dots * \mathcal{E}(\lambda)}_{n \text{ fois}} = \Gamma\left(\frac{1}{\lambda}, n\right)$  (Loi d'ERLANG)
- ▶ Soit  $X \hookrightarrow \mathcal{E}(\lambda)$  et  $Y \hookrightarrow \mathcal{P}(\lambda x)$  alors  $\mathbf{P}([X > x]) = \mathbf{P}([Y = 0])$
- ▶ Si  $X \hookrightarrow \mathcal{E}(\lambda)$  alors  $\lfloor X \rfloor + 1 \hookrightarrow \mathcal{G}(\lfloor \lambda \rfloor)$

**Proposition 14.16.** © :  $\gamma(v) = \Gamma(1, v)$

### 12. Variables aléatoires symétriques

**Définition 14.4. Variable aléatoires symétrique**

©© : On dit qu'une variable aléatoire est symétrique lorsque  $X \sim -X$ .

**Proposition 14.17. Distribution bilatérale**

© : Soit  $X$  une variable aléatoire à densité de densité  $f_X$  alors

$$(X \sim -X) \Leftrightarrow (f_X \text{ est paire})$$

On parle de distribution bilatérale.

**Proposition 14.18.** ©© :  $\forall x \in \mathbb{R}, F_X(-x) = 1 - F_X(x)$

**Proposition 14.19.** ©© : En cas d'existence, toute variable symétrique possède une espérance nulle.

**Proposition 14.20.** ©© : Soit  $X$  et  $Y$  deux variables aléatoires symétriques indépendantes alors  $X \pm Y$  reste une variable aléatoire symétrique.

### 13. Loi géométrique

**Proposition 14.21.** ▶ © : Soit  $X_1, \dots, X_n$   $n$  variables aléatoires indépendantes telles que pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $X_k \hookrightarrow \mathcal{G}(p_k)$  alors

$$\inf(X_1, \dots, X_n) \hookrightarrow \mathcal{G}\left(1 - \prod_{i=1}^n (1 - p_i)\right)$$

- ▶ © :  $\underbrace{\mathcal{G}(p) * \dots * \mathcal{G}(p)}_{n \text{ fois}} = \mathcal{P}(n, p)$  (loi de PASCAL de paramètres  $n$  et  $p$  généralisant la loi géométrique)

**Remarque 14.5.** Il n'y a rien de remarquable à dire concernant la loi de  $\sup(X_1, \dots, X_n)$ .

## 14. Loi normale et loi du chi-deux

**Proposition 14.22.** ►  $\textcircled{C}$  :  $(X \hookrightarrow \mathcal{N}(0, 1)) \Rightarrow \left( X^2 \hookrightarrow \Gamma\left(2, \frac{1}{2}\right) \right)$  (*loi du chi-deux à 1 degrés de liberté*)

►  $\textcircled{C}$  : Soit  $X_1, \dots, X_n$  des variables aléatoires i.i.d.

$$(\forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{N}(0, 1)) \Rightarrow \left( \sum_{k=1}^n X_k^2 \hookrightarrow \Gamma\left(2, \frac{n}{2}\right) \right)$$

(*loi du chi-deux à  $n$  degrés de liberté hors programme*).

15. Produit scalaire et norme associée dans  $L^2(\Omega, \mathcal{A}, \mathbf{P})$ 

**Proposition 14.23.** ►  $\textcircled{\text{S}}\textcircled{\text{C}}$  : Soit  $(X, Y) \in (L^2(\Omega, \mathcal{A}, \mathbf{P}))^2$ ,  $\langle X, Y \rangle = \mathbf{E}(XY)$

►  $\textcircled{\text{S}}\textcircled{\text{C}}$  : Soit  $X \in L^2(\Omega, \mathcal{A}, \mathbf{P})$ ,  $\|X\|^2 = \mathbf{E}(X^2)$



# Index

- Approximation
  - binomiale par Normale, 53
  - binomiale par Poisson, 53
  - poisson par Normale, 53
- Calcul de probabilité
  - à partir de la fonction de répartition à partir d'une varad, 23
  - à partir de la fonction de répartition d'une var quelconque, 23
  - à partir de la fonction de survie, 24
- Cardinal d'un ensemble
  - définition, 11
  - propriétés, 11
- Coefficient
  - de corrélation linéaire, 35
  - multinomial, 71
- Coefficient
  - binomial, 13
- Confiance, 58
- Convergence
  - en loi, 50
    - caractérisation dans le cas discret, 50
    - stabilité par composition, 50
  - en probabilité, 49
    - stabilité par composition, 49
  - spaciale, 49
- Convergences
  - lien, 51
- Convulée de lois, 38
- Dénombrement
  - de listes d'entiers sous contrainte d'égalité linéaire, 13
  - exemples fondamentaux de tirages usuels, 13
  - lemme des bergers, 12
  - nombre d'applications de E vers F, 14
  - nombre d'arrangements de E vers F, 12
  - nombre d'injections de E vers F, 14
  - nombre de combinaisons de p éléments de E de cardinal n, 12
  - nombre de p-listes croissantes, 13
  - nombre de p-listes de E, 12
  - nombre de p-listes strictement croissantes, 13
  - nombre de parties d'un ensemble E, 12
  - nombre de permutations de E, 12
  - problème des chemins monotones, 12
- Densité de probabilité
  - caractérisation, 22
  - définition, 22
- Dispersion d'un vecteur aléatoire, 34
- Distribution bilatérale, 74
- Droites de régression, 38
- Ecart quadratique moyen, 39
- Ecart-type, 27, 39
- Echantillon, 55
  - taille, 59, 60
- Egalité
  - en loi, 23
- Ensemble
  - complémentation, 8
  - définition intuitive d'un ensemble, 7
  - différence d'ensembles, 9
  - différence symétrique d'ensembles, 10
  - distributivités, 8
  - ensemble des parties d'un ensemble, 7
  - inclusion d'ensembles, 8
  - intersection d'ensembles, 7
  - lois de Morgan, 8
  - partition, 11
  - produit cartésien, 9
  - union, 7
    - disjointe, 7
- Equiprobabilité, 18
- Espérance
  - conditionnelle, 32
  - croissance, 26
  - définition, 25
  - domination, 26
  - et fonction de survie, 25
  - et variable bornée, 25
  - linéarité, 31
  - nulle d'une variable positive, 26
  - positivité, 26
- Espace
  - probabilisé, 17
  - probabilisable, 17
- Estimateur
  - biais, 55
  - consistant, 55
  - convergent, 55
  - d'un paramètre, 55
  - risque quadratique moyen, 57
- Estimation
  - définition, 55
  - inférentielle, 55
  - par intervalle de confiance asymptotique d'une moyenne de loi normale de variance donnée, 60
  - ponctuelle, 55
- Événement
  - négligeable, 17
  - quasi-certain, 17
  - système complet, 17

- Fonction
  - de régression, 33
  - de répartition, 22
  - de survie, 24
- Fonctions de régression, 33
- Forme quadratique associée à une matrice de covariance, 34
- Formule
  - de Bayes, 20
  - de l'espérance totale, 33
  - de la probabilité d'une cause, 20
  - de Poincaré ou du crible pour deux ensembles
    - dénombrement, 11
  - de Vandermonde, 71
  - des probabilités composées, 19
  - des probabilités totales, 19
  - du binôme de Newton, 67
  - du crible pour deux ensembles
    - probabilité, 17
  - du triangle de Pascal, 13
  - multinomiale, 71
- Fréquence empirique, 56
- Identité
  - de Bernoulli, 67
  - du parallélogramme, 34
- Identités
  - de polarisation, 34
- Inégalité
  - avec une variable indicatrice, 47
  - de Bienaymé-Tchebychev, 45
  - de Boole ou sous-additivité, 17
  - de Cauchy-Schwarz, 35, 46, 72
  - de Chernoff, 45
  - de concentration, 45
  - de Hoeffding, 58
  - de Jensen, 46
  - de Markov, 45
  - de moments, 46
  - triangulaire, 46
- Indépendance
  - deux événements
    - caractérisation, 18
  - mutuelle, 19
  - n événements, 18
  - suite d'événements, 19
  - variables aléatoires
    - caractérisation, 36
- Intervalle de confiance
  - asymptotique d'un paramètre, 58
  - d'un paramètre
    - exact, 58
    - par excès, 57
  - d'une proportion, 59
  - précision, 58-60
- Lemme des coalitions, 36
- Les trois grandes partitions de  $\mathbb{N}^2$ 
  - de deux familles sommables, 29
- Loi
  - beta, 44
  - binomiale, 62
  - caractérisation, 23
  - caractérisation de la loi d'une fonction d'un couple, 30
  - caractérisation de la loi d'une fonction d'un vecteur, 32
  - certaine, 62
  - conditionnelle, 30
  - d'Erlang, 63, 74
  - d'un couple, 30
  - de Bernoulli, 61
  - de Burr, 44
  - de Cauchy, 44
  - de Dirac, 62
  - de Fréchet, 44
  - de Gumbel, 44
  - de khi-deux, 44
  - de Pareto, 44
  - de Pascal, 74
  - de Poisson, 65
  - de probabilité
    - d'une variable discrète, 22
    - d'une variable queconque, 22
  - de Student, 44
  - de Weibull, 44
  - du chi-deux, 75
  - exponentielle, 63
  - faible des grands nombres, 50
  - géométrique, 63
  - gamma
    - à deux paramètres, 73
    - à un seul paramètre, 63
  - log-normale, 44
  - logistique, 44
  - marginale, 30
  - normale, 64
  - uniforme, 65
- Matrice de covariance-variance d'un vecteur aléatoire, 34
- Moment
  - factoriel d'ordre  $r$ , 28
- Moyenne empirique, 56
- Norme dans  $L_2$ , 39
- Principe additif, 11
- Principe multiplicatif, 11
- Produit
  - de Cauchy, 37
  - de convolution, 37
  - de deux familles sommables, 29
  - lien entre produit et somme, 69
  - permutation des produits, 68
  - propriétés élémentaires, 68
  - scalaire dans  $L_2$ , 39
- Quantile d'ordre  $1-\alpha/2$ , 59
- Relation
  - de Laplace, 18
  - min et max, 71
- Schéma de Bernoulli, 61
- Standardisation d'une variable aléatoire, 28
- Statistique, 55
  - inférentielle, 55
- Théorème
  - d'existence d'une espérance par domination, 31
  - de composition
    - convergence en loi, 50
    - convergence en probabilité, 49
  - de convolution, 37
  - de domination, 26

- de Fisher-Tippett-Gnedenko, 44
  - domaines d'attraction, 44
- de Koenig-Huygens, 27
- de la limite centrée, 50
- de la limite monotone, 17
- de Slutsky, 50
- de sommation par paquets, 29
- de transfert en dimension un, 28
- de transfert fonction d'un couple, 30
- de transfert fonction d'un vecteur, 31
- Koenig, 33
- Tirages
  - au hasard, 18
- Tribu
  - grossière, 15
  - la plus fine, 15
- Univers image d'une variable, 21
- Variable aléatoire
  - étagée, 24
  - à densité, 21
  - bornée, 24
  - centrée, 25
  - centrée réduite, 28
  - certaine, 22, 24
  - déterministe, 22, 24
  - discrète, 21
    - finie, 21
    - infinie, 21
  - indicatrice, 61
  - normalisation, 22
  - quasi-certaine, 24
  - sans mémoire, 24
  - support, 21
  - symétrique, 25, 74
- Variations aléatoires
  - de même loi, 23
  - extrêmes, 41, 42
    - événements, 43
    - comparaison, 44
    - formule d'inclusion-exclusion, 43
    - monotonie, 43
  - non corrélées ou décorréelées, 33
- Variance
  - additivité, 35
  - d'une somme, 34
  - empirique, 56
    - rectifiée, 57
  - quadraticité, 37
- Vecteur discret
  - loi, 31