

COURS DE PROBABILITES

Tous les résultats du cours d'ECS

ECE -MP – PC - PT - BCPST

Table des matières

Chapitre 1.	Introduction	3
Chapitre 2.	Légendes et abréviations	5
Chapitre 3.	Ensembles	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.	Lois de Morgan où $\overline{A} = \complement_E A$	8
7.	Inclusion d'ensembles	8
8.	Différence d'ensembles	9
9.	Différence symétrique d'ensembles	9
10.	Produit cartésien d'ensembles	9
11.	Fonction indicatrice (ou fonction caractéristique) d'un ensemble	9
12.	Liminf et Limsup d'ensembles	9
Chapitre 4.	Formules combinatoires	11
1.	Coefficient binomial	11
2.	Coefficient multinomial	12
3.	Produit	12
Chapitre 5.	Dénombrement	15
Chapitre 6.	Espaces probabilisés	17
1.	Tribu	17
2.	Correspondance entre relations logiques et ensemblistes	17
3.	Probabilité	18
Chapitre 7.	Variables aléatoires réelles	21
1.	Généralités	21
2.	Variables aléatoires particulières	23
3.	Moments	24
Chapitre 8.	Vecteurs aléatoires discrets	29
1.	Famille sommables	29
2.	Couples discrets	30
3.	Vecteurs discrets de dimension $n \geq 2$	31
4.	Droites de régression	38
5.	Produit scalaire dans $\mathcal{L}^2$	38
Chapitre 9.	Inégalités probabilistes	39
1.	Inégalités de concentration	39
2.	Inégalités de moments	40
Chapitre 10.	Convergences	43
1.	Convergence en probabilités	43

2. Convergence en loi	44
3. Liens entre ces deux types de convergence	45
Chapitre 11. Approximations	47
Chapitre 12. Estimation inférentielle	49
1. Estimation ponctuelle	49
2. Estimation par intervalle de confiance	51
Chapitre 13. Zoologie des lois usuelles du programme	53
1. Loi de Bernoulli (S)	53
2. Variable indicatrice (S)	53
3. Loi binomiale (S)	53
4. Loi de Dirac (ou loi certaine) (S)	54
5. Loi exponentielle (C)	54
6. Loi gamma à un seul paramètre (C)	55
7. Loi géométrique (S)	55
8. Lois normales (C)	55
9. Loi de Poisson (S)	56
10. Lois uniformes	57
11. Bilan des lois stables par convolution	58
12. Bilan des lois sans mémoire	58
Chapitre 14. Compléments hors programme mais ...	59
1. Formule de Vandermonde	59
2. Relation min et max	59
3. Urne multicolore	59
4. Espérance et fonction de survie	59
5. Interprétation géométrique de la corrélation	60
6. Minimisation de la variance	60
7. Espérance conditionnelle	60
8. Loi exponentielle	60
9. Variables symétriques	60
10. Variables aléatoires extrêmes	61
11. Loi géométrique	61
12. Loi normale et loi du chi-deux	61
13. Produit scalaire et norme associée dans $\mathcal{L}^2$	61
Chapitre 15. Outils	63
Index	65

CHAPITRE 1

Introduction

Ce document fête les 10 ans de spicesagros, avec beaucoup de retard, énormément de retard ! Ce fut l'un des tous premiers sites créés en France pour les prépas EC, à une époque où il n'y avait quasiment rien sur le net pour cette filière. Il présente l'intégralité des résultats du programme de 2014 en Maths 2. Sa structure est différente des formulaires que l'on trouve en librairie puisqu'il présente l'avantage de rassembler les résultats communs (que j'aime qualifier d'universels) que l'on retrouve dans l'étude des variables discrètes et des variables continues. Cela rend sa lecture plus rapide et efficace. Pour cela il faudra bien s'imprégner des légendes et abréviations qui suivent cette introduction. Certaines sont hors programme mais terriblement pratiques pour synthétiser le contenu. Pour être complet ce vadémécum contient un nombre conséquent de résultats hors programme, clairement signalés par un "HP mais ..." en rouge, car ce poly est en couleur ! Ces résultats ne sont pas mis au hasard et tiennent compte des plus récentes évolutions des sujets d'écrits et d'oraux qui les introduisent plus que régulièrement. Ils ne sont pas présents pour vous paniquer, mais pour donner un contour le plus fin possible de ce qui est susceptible de tomber le 9 mai 2022, et encore je n'ai pas tout mis, j'en laisse un peu plus pour mes étudiants ... En vrac, vous y trouverez un chapitre conséquent sur les vecteurs aléatoires, un autre très important de compléments, d'outils mathématiques pour aborder les maths 2. Toutes les lois avec de nombreux compléments sont regroupés en un seul chapitre etc, etc ... Enfin vous découvrirez une table des matières ainsi qu'un index thématique de plus d'une centaine d'entrées. Il est inutile de vous rappeler qu'en aucun cas ce document ne doit se substituer au cours de votre professeur dont l'étude minutieuse doit être votre priorité absolue. Ce poly pourra aussi être utilisé à profit par des étudiants de MP, PC, PT, BCPST, L1, L2, mais aussi par ceux préparant le capes ainsi que l'agrégation interne.

Je vous souhaite de prendre autant de plaisir à l'utiliser que moi pour le créer. Pour terminer je vous demande la plus grande indulgence si vous trouvez quelques rares coquilles qui subsisteraient après de multiples reprises du document qui s'étalent sur plusieurs années.

Bon travail à tous.

Le 19/12/2021 à 3h10.

Légendes et abréviations

- HP : résultat hors programme mais ...
- $\textcircled{C}$: résultat valable en continu
- $\textcircled{S}$: résultat valable en discret
- $\sqcup$: union disjointe
- ssrcv : sous réserve de convergence
- $\text{ssr } |cv|$: sous réserve de convergence absolue
- i.i.d. : variables indépendantes et identiquement distribuées
- SCE : système complet d'événements
- v.a.r. : variable aléatoire
- v.a.r.d. : variable aléatoire discrète
- v.a.r.a.d. : variable aléatoire à densité
- $\sum_i u_i < \infty$: 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)
- $\forall k \geq 1$, $\mathcal{L}^k$ désigne l'espace vectoriel des variables aléatoires admettant un moment d'ordre k (notation de l'ensemble hors programme à ne surtout pas utiliser)
- $\forall k \geq 1$, $\mathcal{L}_d^k$ 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 X et Y suivent la même loi.
- $\mathcal{S}_X$ support de X .

Ensembles

1. Définition intuitive d'un ensemble

Définition 3.1. 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 $\forall x \in A, x \in B$ et $\forall x \in B, x \in A$. Par contraposée $A \neq B$ lorsque $\exists x \in A, x \notin B$ ou $\exists x \in B, x \notin A$.

Remarque 3.1. L'ordre de l'écriture des éléments d'un ensemble n'a aucune importance, ainsi que leurs répétitions éventuelles. Par exemple : $\{1, 2, 3\} = \{2, 3, 1\} = \{1, 1, 2, 3, 3, 3\}$.

Définition 3.2. Soit A et B deux ensembles, on dira que B est un sous-ensemble de A ou que B est une partie de A si, et seulement si, tout élément de B est aussi un élément de A . Ainsi : $B \subset A$ lorsque $\forall x \in B, x \in A$.

Remarque 3.2. $(A = B) \Leftrightarrow (A \subset B \text{ et } B \subset A)$.

2. Ensemble des parties d'un ensemble

Définition 3.3. 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 et l'ensemble vide.

Remarque 3.3. ► Ecrire $A \subset E$ se traduit par $A \in \mathcal{P}(E)$ ( on n'écrit jamais $A \in E$!).

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

3. Union d'ensembles

Définition 3.4. Soit A et B deux parties d'un ensemble E non vide,

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

Plus généralement

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

Propriétés 3.1. ► $A \cup \overline{A} = A$

► $A \cup B = B \cup A$ (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 $\cup$)

► $E \cup A = E$ (E est absorbant pour $\cup$)

► $A \cup A = A$

4. Intersection d'ensembles

Définition 3.5. Soit A et B deux parties d'un ensemble E non vide,

$$A \cap B = \{x \in E, x \in A \text{ et } x \in B\}$$

Plus généralement

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

Proposition 3.1. ► $A \cap \overline{A} = \emptyset$

- $A \cap B = B \cap A$ (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 $\cap$)
- $E \cap A = A$ (E est neutre pour $\cap$)
- $A \cap A = A$

5. Distributivités

Propriétés 3.2. ► $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (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. Lois de Morgan où $\overline{A} = \complement_E A$

Propriétés 3.3. ► $\overline{A \cap B} = \overline{A} \cup \overline{B}$

- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $\overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \overline{A_i}$
- $\overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \overline{A_i}$ (*Lois de Morgan généralisées*)

Propriétés 3.4. ► $(A \cup B = E \text{ et } A \cap B = \emptyset) \Leftrightarrow (A \sqcup B = E)$

- $\overline{\overline{\Omega}} = \Omega$
- $\overline{\emptyset} = E$
- $\overline{\overline{A}} = A$
- $(B = \overline{A}) \Leftrightarrow (A = \overline{B})$
- $(A \subset B) \Leftrightarrow (\overline{B} \subset \overline{A})$

7. Inclusion d'ensembles

Définition 3.6. Soit A et B deux parties d'un ensemble E non vide, alors $A \subset B$ si

$$\forall x \in E, (x \in A \Rightarrow x \in B)$$

Propriétés 3.5. ► $(A \subset B) \Leftrightarrow (A \cap B = A)$

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

8. Différence d'ensembles

Définition 3.7. Soit A et B deux parties d'un ensemble E non vide,

$$A - B = \{x \in E, x \in A \text{ et } x \notin B\}$$

Propriétés 3.6. ► $A - B = A \cap \overline{B}$

► $A - B = A - (A \cap B)$

► $\overline{A} = \Omega - A$

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

9. Différence symétrique d'ensembles

Définition 3.8. (**HP mais ...**) $A \Delta B = (A \cup B) - (A \cap B) = (A - B) \cup (B - A)$

10. Produit cartésien d'ensembles

Définition 3.9. Soit A et B deux parties d'un ensemble E non vide,

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

Plus généralement,

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

11. Fonction indicatrice (ou fonction caractéristique) d'un ensemble

Définition 3.10. Soit E un ensemble. Pour toute partie A de E , on définit la fonction indicatrice (ou fonction caractéristique) de A notée $\mathbf{1}_A$ par

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

Propriétés 3.7. ► $\forall x \in E, \mathbf{1}_E(x) = 1, \mathbf{1}_\emptyset(x) = 0$

► $\mathbf{1}_{A \cap B} = \mathbf{1}_A \times \mathbf{1}_B$, le résultat se généralise par récurrence

► $\mathbf{1}_{A \cap B} = \inf(\mathbf{1}_A, \mathbf{1}_B)$, le résultat se généralise par récurrence

► $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A \cap B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \times \mathbf{1}_B$

► $\mathbf{1}_{A \cup B} = \sup(\mathbf{1}_A, \mathbf{1}_B)$, le résultat se généralise par récurrence

► Si $B \subset A \subset E$ alors $\mathbf{1}_{A-B} = \mathbf{1}_A - \mathbf{1}_B$

► $\forall n \in \mathbb{N}^*, (\mathbf{1}_A)^n = \mathbf{1}_A$

► (**HP mais ...**) Soit $(A_n)_{n \geq 0}$ une suite de sous-ensembles de E deux à deux disjoints, alors

$$\sum_{n=0}^{+\infty} \mathbf{1}_{A_n} = \mathbf{1}_{\bigcup_{n=0}^{+\infty} A_n}$$

Si les ensembles A_n constituent une partition de E alors

$$\sum_{n=0}^{+\infty} \mathbf{1}_{A_n} = 1$$

12. Liminf et Limsup d'ensembles

Définition 3.11. (**HP mais ...**) 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 3.8. $\bigcap_{n=0}^{+\infty} \bigcup_{k=n}^{+\infty} A_k$ signifie "les événements A_n sont réalisés une infinité de fois".

Propriétés 3.9. $\bigcup_{n=0}^{+\infty} \bigcap_{k=n}^{+\infty} A_k$ signifie "tous les événements A_n sont réalisés à partir d'un certain rang".

Propriétés 3.10. ► $\overline{\limsup_{n \rightarrow +\infty} (A_n)} = \liminf_{n \rightarrow +\infty} (\overline{A_n})$

► $\liminf_{n \rightarrow +\infty} (A_n) \subset \limsup_{n \rightarrow +\infty} (A_n)$

► $\limsup_{n \rightarrow +\infty} (A_n) = \left\{ \sum_{n \geq 0} \mathbf{1}_{A_n} = \infty \right\}$

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

Formules combinatoires

Proposition 4.1. Soit n un entier naturel non nul,

$$\begin{aligned} \blacktriangleright \sum_{k=1}^n k &= \frac{n(n+1)}{2} \\ \blacktriangleright \sum_{k=1}^n k^2 &= \frac{n(n+1)(2n+1)}{6} \\ \blacktriangleright \sum_{k=1}^n k^3 &= \left(\frac{n(n+1)}{2} \right)^2 \end{aligned}$$

1. Coefficient binomial

Définition 4.1. $\binom{n}{p} = \begin{cases} \frac{n!}{p!(n-p)!} & \text{si } 0 \leq p \leq n \\ 0 & \text{sinon} \end{cases}$

Propriétés 4.1. $\blacktriangleright \forall (n, p) \in \mathbb{N}^2, p \leq n, \binom{n}{p} = \binom{n}{n-p}$ (symétrie)

$\blacktriangleright \forall (n, p) \in \mathbb{N}^2, 1 \leq p \leq n, \binom{n}{p} = \frac{n}{p} \binom{n-1}{p-1}$ (la formule sans nom !)

$\blacktriangleright \forall (n, p) \in \mathbb{N}^2, 1 \leq p \leq n, \binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}$ (*triangle de Pascal*)

$\blacktriangleright \forall (p, d) \in \mathbb{N}^2, p \leq d, \sum_{k=p}^d \binom{k}{p} = \binom{d+1}{p+1}$ (*triangle de Pascal généralisée, HP mais ... et doit être redémontrée*)

Théorème 4.1. *Formule du binôme de Newton*

$$\forall n \in \mathbb{N}, \forall (a, b) \in \mathbb{C}^2, (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^n \binom{n}{k} b^k a^{n-k}$$

Corollaire 4.1. $\forall n \in \mathbb{N}, \sum_{k=0}^n \binom{n}{k} = 2^n$

Propriétés 4.2. $\forall n \in \mathbb{N}^*, \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} = 2^{n-1} = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1}$ (*HP mais ... et doit être redémontré*)

Propriétés 4.3. $\forall n \in \mathbb{N}^*, \sum_{k=1}^n k \binom{n}{k} = n 2^{n-1}$

Théorème 4.2. *Vandermonde (HP mais ... et doit être redémontré)*

$$\forall (n_1, n_2, n) \in \mathbb{N}^3, n \leq n_1 + n_2, \sum_{k=0}^n \binom{n_1}{k} \binom{n_2}{n-k} = \binom{n_1 + n_2}{n}$$

Corollaire 4.2. $\forall n \in \mathbb{N}, \sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}$

Théorème 4.3. Vandermonde généralisé (HP mais ...)

$$\forall (N_1, \dots, N_k) \in \mathbb{N}^k, \sum_{\substack{(n_1, \dots, n_k) \in \llbracket 0, n \rrbracket^k \\ n_1 + \dots + n_k = n}} \binom{N_1}{n_1} \times \dots \times \binom{N_k}{n_k} = \binom{N_1 + \dots + N_k}{n}$$

avec $n \leq N_1 + \dots + N_k$.

Propriétés 4.4. $\forall (n, p) \in \mathbb{N}^2, 1 \leq p \leq n, \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} 1 = \binom{n}{p}$ (HP mais ... et doit être redémontré)

- $\forall n \in \mathbb{N}^*, \left(\sum_{k=1}^n x_k \right)^2 = \sum_{k=1}^n x_k^2 + 2 \sum_{1 \leq i < j \leq n} x_i x_j$
- $\forall n \in \mathbb{N}^*, \left(\sum_{k=1}^n x_k \right)^3 = \sum_{k=1}^n x_k^3 + 3 \sum_{j \neq k} x_j^2 x_k + \sum_{j \neq k \neq \ell} x_j x_k x_\ell$ (très utile en stats HP mais ...)
- $\forall n \in \mathbb{N}^*, \left(\sum_{k=1}^n x_k \right)^4 = \sum_{k=1}^n x_k^4 + 4 \sum_{j \neq k} x_j x_k^3 + 3 \sum_{j \neq k} x_j^2 x_k^2 + 6 \sum_{i \neq j \neq k} x_i^2 x_j x_k + \sum_{i \neq j \neq k \neq \ell} x_i x_j x_k x_\ell$ (très utile en stats ! HP mais ...)

Propriétés 4.5. Identité de Bernoulli

$$\forall (a, b) \in \mathbb{C}^2, \forall n \in \mathbb{N}, a^{n+1} - b^{n+1} = (a - b) \sum_{k=0}^n a^{n-k} b^k = (a - b) \sum_{k=0}^n a^k b^{n-k}$$

Définition 4.2. ► $\forall (n, p) \in \mathbb{N}^2, p \leq n, \forall (a_p, \dots, a_n) \in \mathbb{R}^{n-p+1}, \prod_{k=p}^n a_k = a_p \times \dots \times a_n$.

- On pose $\prod_{k=p}^n a_k = 1$ lorsque $p > n$

2. Coefficient multinomial

Définition 4.3. (HP mais ...) $\binom{n}{n_1, \dots, n_p} = \frac{n!}{n_1! \times \dots \times n_p!}$ avec $n_1 + \dots + n_p = n$.

Propriétés 4.6. $\binom{n}{n_1, \dots, n_p} = \binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \binom{n-n_1-n_2}{n_3} \times \dots \times \binom{n-n_1-\dots-n_{p-1}}{n_p}$
avec $n_1 + \dots + n_p = n$.

Théorème 4.4. Formule multinomiale (HP mais ...)

$$\forall (a_1, \dots, a_p) \in \mathbb{R}^p, \forall n \in \mathbb{N}, (a_1 + \dots + a_p)^n = \sum_{\substack{(n_1, \dots, n_p) \in \mathbb{N}^p \\ n_1 + \dots + n_p = n}} \underbrace{\frac{n!}{n_1! \times \dots \times n_p!}}_{\text{coef}^{\text{e}} \text{ multinomial}} \times a_1^{n_1} \times \dots \times a_p^{n_p}$$

3. Produit

Propriétés 4.7. 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 complexes

- $\prod_{k=p}^n a = a^{n-p+1}$

- $\prod_{k=p}^n a^k = a^{\sum_{k=p}^n k}$
- $\prod_{k=p}^n ax_k = a^{n-p+1} \prod_{k=p}^n x_k$
- Pour toute permutation σ 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 complexes $y_k \neq 0$)
- $\prod_{k=p}^n \frac{x_{k+1}}{x_k} = \frac{x_{n+1}}{x_p}$ (où tous les complexes $x_k \neq 0$) (téléscopage)

Propriétés 4.8. *Permutation des produits*

Soit $(a_{i,j})_{1 \leq i,j \leq n}$ une famille de nombres complexes

- $\prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{i,j} = \prod_{i=1}^n \prod_{j=1}^n a_{i,j} = \prod_{j=1}^n \prod_{i=1}^n a_{i,j}$
- $\prod_{1 \leq i \leq j \leq n} a_{i,j} = \prod_{j=1}^n \prod_{i=1}^j a_{i,j} = \prod_{i=1}^n \prod_{j=i}^n a_{i,j}$
- $\prod_{1 \leq i < j \leq n} a_{i,j} = \prod_{j=2}^n \prod_{i=1}^{j-1} a_{i,j} = \prod_{i=1}^{n-1} \prod_{j=i+1}^n a_{i,j}$

Propriétés 4.9. *Liens entre produit et somme*

- $\forall n \in \mathbb{N}^*, \forall k \in \{1, \dots, n\}, x_k > 0, \ln \left(\prod_{k=1}^n x_k \right) = \sum_{k=1}^n \ln(x_k)$
- $\forall n \in \mathbb{N}^*, \exp \left(\sum_{k=1}^n x_k \right) = \prod_{k=1}^n \exp(x_k)$
-  $\prod_i \sum_j a_{ij} \neq \sum_j \prod_i a_{ij}$

Propriétés 4.10. $\prod_{k=1}^n k = n!$

Propriétés 4.11. *Produits "à trous"*

- $\forall n \in \mathbb{N}^*, \prod_{k=1}^n (2k) = 2^n n!$
- $\forall n \in \mathbb{N}^*, \prod_{k=1}^n (2k+1) = \frac{(2n+1)!}{2^n n!}$

Dénombrement

Propriétés 5.1. ► 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 $A = B$.

► $\text{Card}(A \sqcup B) = \text{Card}(A) + \text{Card}(B)$

► $\forall n \in \mathbb{N}^*, \text{Card}\left(\bigsqcup_{k=1}^n A_k\right) = \sum_{k=1}^n \text{Card}(A_k)$

► $\forall n \in \mathbb{N}^*, \text{Card}(A_1 \times \cdots \times A_n) = \prod_{k=1}^n \text{Card}(A_k)$

Théorème 5.1. *Formule de Poincaré ou du crible pour deux ensembles*

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

Remarque 5.1. La généralisation de cette formule a disparu du programme.

Définition 5.1. *Partition d'un ensemble*

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

► $\forall i \in I, A_i \cap A_j = \emptyset$,

► $(\forall (i, j) \in I^2, i \neq j) \Rightarrow (A_i \cap A_j = \emptyset)$ et

► $\bigsqcup_{i \in I} A_i = E$.

Remarque 5.2. C'est la même définition pour un système complet d'événements avec $E = \Omega$.

Proposition 5.1. *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 pour tout $y \in F$, $\text{Card}(f^{-1}(\{y\})) = p$, alors

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

Proposition 5.2. *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 5.3. *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

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

noté A_n^p (**HP mais ...**).

Proposition 5.4. *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

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

noté A_n^p (**HP mais ...**).

Proposition 5.5. *Nombre de permutations de E_n avec $\text{Card}(E_n) = n \in \mathbb{N}^*$:*

$$n!$$

Proposition 5.6. *Nombre de p -listes de E_n :*

$$n^p$$

Proposition 5.7. *Nombre de combinaisons de p éléments de E de cardinal n avec $p \leq n$:*

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

(la dernière égalité est **HP mais ...**)

Proposition 5.8. *Nombre de parties d'un ensemble E*

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

Proposition 5.9. *Problème des anagrammes (**HP mais ...**)*

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'annagrammes 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 d'arrangements avec répétition)

Proposition 5.10. *Nombre de couples d'entiers*

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

Remarque 5.3. La proposition précédente intervient dans le théorème de sommation par paquets d'une série double, lorsqu'on effectue une sommation selon des droites obliques.

Proposition 5.11. *Nombre de suites d'entiers*

- $\text{Card} \{(x_1, \dots, x_p) \in (\mathbb{N}^*)^p \mid 1 \leq x_1 < \cdots < x_p \leq n\} = \binom{n}{p}$
- $\text{Card} \{(x_1, \dots, x_p) \in (\mathbb{N}^*)^p \mid 1 \leq x_1 \leq \cdots \leq x_p \leq n\} = \binom{n+p-1}{p}$ (**HP mais ...**)

Exemple 5.1. *Exemples fondamentaux de tirages usuels*

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

Espaces probabilisés

1. Tribu

Définition 6.1. Tribu ou σ -algèbre sur Ω

Tout ensemble de parties d'un ensemble Ω , contenant Ω , stable par réunion au plus dénombrable et par passage au complémentaire, s'appelle une tribu ou σ -algèbre sur Ω , souvent notée $\mathcal{A}$. Autrement dit :

- $\Omega \in \mathcal{A}$,
- $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$,
- pour toute suite $(A_n)_{n \geq n_0}$ d'événements de $\mathcal{A}$, $\bigcup_{n=n_0}^{+\infty} A_n \in \mathcal{A}$.

Proposition 6.1. $\emptyset \in \mathcal{A}$, $\mathcal{A}$ est stable pour $\cap$, $-$, Δ

Exemple 6.1. ► $\mathcal{A} = \{\emptyset, \Omega\}$ (tribu grossière, la plus triviale mais inexploitable)

► Si Ω est au plus dénombrable $\mathcal{A} = \mathcal{P}(\Omega)$ (tribu complète, la plus fine)

Définition 6.2. Tribu engendrée par une famille

Soit $\mathcal{F} = \{A_i, i \in I\}$ une famille de sous-ensembles de Ω . Alors la tribu engendrée par $\mathcal{F}$ est la plus petite tribu sur Ω contenant tous les sous-ensembles $A_i, i \in I$. Elle est notée $\sigma(\mathcal{F})$.

Exemple 6.2.

Proposition 6.2. $\{\emptyset, A, \bar{A}, \Omega\}$ est une tribu notée $\sigma(A)$, c'est la plus petite engendrée par $A \in \mathcal{A}$.

Proposition 6.3. L'ensemble $\left\{ \bigcup_{\ell \in L} A_\ell \mid L \in \mathcal{P}(K) \right\}$ où $(A_k)_{k \in K}$ est un système complet d'événements est la tribu engendrée par celui-ci est notée $\sigma((A_k)_{k \in K})$.

Définition 6.3. Tribu borélienne de $\mathbb{R}$

$$\mathcal{B}(\mathbb{R}) = \sigma(\{]-\infty, x[\mid x \in \mathbb{R}\}) = \sigma(\{[x, +\infty[\mid x \in \mathbb{R}\}) = \dots$$

Proposition 6.4. Quelle tribu utiliser le jour du concours ?

- Si Ω est fini alors $\mathcal{A} = \mathcal{P}(\Omega)$.
- Si Ω est infini et dénombrable alors $\mathcal{A} = \mathcal{P}(\Omega)$.
- Si Ω est infini et indénombrable : dans le cadre du programme, on se cherche 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 $\Rightarrow$	inclusion $\subset$
équivalence $\Leftrightarrow$	égalité $=$

3. Probabilité

3.1. Définition et propriétés.

Définition 6.4. *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

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

(σ -additivité de $\mathbf{P}$).

Définition 6.5. *Espace probabilisable*

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

Définition 6.6. *Espace probabilisé*

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

Définition 6.7. *Événement négligeable*

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

Remarque 6.1. Il n'y a aucune raison que $A = \emptyset$.

Définition 6.8. *Événement quasi-certain*

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

Remarque 6.2. Il n'y a aucune raison que $A = \Omega$.

Définition 6.9. *Système complet d'événements*

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

Propriétés 6.1. Soit A et B deux événements de $(\Omega, \mathcal{A}, \mathbf{P})$,

- $\mathbf{P}(\emptyset) = 0$
- $\mathbf{P}(A \sqcup B) = \mathbf{P}(A) + \mathbf{P}(B)$
- $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$
- $\mathbf{P}(A - B) = \mathbf{P}(A) - \mathbf{P}(A \cap B)$
- $(B \subset A) \Rightarrow (\mathbf{P}(A - B) = \mathbf{P}(A) - \mathbf{P}(B))$
- $(B \subset A) \Rightarrow (\mathbf{P}(B) \leq \mathbf{P}(A))$ (croissance de $\mathbf{P}$)
- $(B \subset A) \Rightarrow (\mathbf{P}(\overline{A}) \leq \mathbf{P}(\overline{B}))$
- $\mathbf{P}(A \cap B) \leq \min\{\mathbf{P}(A), \mathbf{P}(B)\}$
- $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$ (*formule du crible ou de Poincaré pour deux ensembles dont la généralisation a disparu du programme*)
- $\mathbf{P}(A \cup B) = \mathbf{P}(\overline{A} \cap B) + \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \overline{B})$

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

Théorème 6.1. *Limite monotone*

- $\left((A_n)_{n \geq 0} \nearrow\right) \Rightarrow \left(\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}(A_n)\right)$
- $\left((A_n)_{n \geq 0} \searrow\right) \Rightarrow \left(\mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}(A_n)\right)$

Corollaire 6.1. Soit $(A_n)_{n \geq 0}$ une suite quelconque d'événements,

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

Proposition 6.5. Soit A et B deux événements de $(\Omega, \mathcal{A}, \mathbf{P})$,

$$\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$$

Théorème 6.2. *Inégalité de Boole ou sous-additivité (HP mais ...)*

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

Définition 6.10.

Théorème 6.3. *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 $n \in \mathbb{N}^*$. 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 6.3. On est dans une situation d'équiprobabilité dans le cas de tirages effectués au hasard.

Remarque 6.3.  Il ne peut y avoir d'équiprobabilité lorsque l'univers est infini.

3.2. Indépendance d'événements.

Définition 6.11. *Deux événements*

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

Proposition 6.6. $(A \perp A) \Leftrightarrow (\mathbf{P}(A) = 0 \text{ ou } \mathbf{P}(A) = 1)$ (HP mais ...)

Définition 6.12. *n événements deux à deux indépendants*

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

Définition 6.13. *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 6.4. 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 !

Remarque 6.5.  On n'écrit JAMAIS de produit infini. On se ramènera TOUJOURS à un produit fini à l'aide du théorème de la limite monotone.

Définition 6.14. *Suites d'événements*

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

Remarque 6.6. ► Pour $n \geq 2$, $\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1) \times \mathbf{P}(A_2) \times \dots \times \mathbf{P}(A_n)$.

► L'indépendance mutuelle est plus forte que l'indépendance deux à deux.

Proposition 6.7. Les assertions suivantes sont équivalentes pour deux événements A et B :

► $A \perp B$

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

Proposition 6.8. Soit $(A_n)_{n \geq 0}$ une suite d'événements indépendants alors $(B_n)_{n \geq 0}$ où pour tout entier n , $B_n = A_n$ ou $\overline{A_n}$, reste une suite d'événements indépendants.

Proposition 6.9. $(A \perp A) \Leftrightarrow (\mathbf{P}(A) = 0 \text{ ou } \mathbf{P}(A) = 1)$

3.3. Probabilité conditionnelle.

Définition 6.15. Soit $A \in \mathcal{A}$ tel que $\mathbf{P}(A) \neq 0$, $\mathbf{P}_A(B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}$.

Proposition 6.10. Toutes les propriétés vues sur les probabilités inconditionnelles sont encore valables et s'appliquent aux probabilités conditionnelles.

Proposition 6.11. 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 6.12. 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 6.13. 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 6.4. 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

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

Variables aléatoires réelles

1. Généralités

Définition 7.1. $\textcircled{S}\textcircled{C}$: $X : \Omega \rightarrow \mathbb{R}$ tel que pour tout $B \in \mathcal{B}(\mathbb{R})$ on a

$$X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\} \in \mathcal{A}$$

Remarque 7.1. La notation $X(\Omega)$ est abusive¹ car il n'est pas dit que $X(\Omega)$ coïncide avec l'image de l'application X car il se pourrait que Ω comporte des événements de probabilité nulle et dans ce cas la probabilité que X prenne certaines valeurs de $X(\Omega)$ soit nulle. Dans de nombreuses sources on voit que $X(\Omega)$ regroupe les valeurs qui peuvent effectivement être prises par X (en tenant compte des probabilités et on introduira ci-dessous la notion de support) et dans d'autres $X(\Omega)$ regroupe l'intégralité des valeurs prises par X . Donc on voit que le choix de $X(\Omega)$ n'est pas unique !

Proposition 7.1. (**HP mais ...**) $\textcircled{S}\textcircled{C}$: Soit $X : \Omega \rightarrow \mathbb{R}$ et $g : I \rightarrow \mathbb{R}$, $g \in \mathcal{C}^0(I, \mathbb{R})$ où I est un intervalle de $\mathbb{R}$ contenant $X(\Omega)$, alors $g \circ X$ reste une v.a.r.

Proposition 7.2. (**admis**) L'ensemble $\mathcal{V}(\Omega, \mathcal{A}, \mathbf{P})$ des variables réelles définies sur $(\Omega, \mathcal{A}, \mathbf{P})$ est un $\mathbb{R}$ -espace vectoriel stable par multiplication.

Proposition 7.3. (**HP mais ...**) $\textcircled{S}\textcircled{C}$: La borne inférieure ainsi que la borne supérieure d'un vecteur aléatoire restent une variable aléatoire.

Définition 7.2. **Support d'une variable aléatoire**

► $\textcircled{S}$: On appelle support d'une variable discrète X l'ensemble noté S_X défini par

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

► $\textcircled{C}$: On appelle support d'une variable continue X l'ensemble noté S_X défini par

$$S_X = \{x \in X(\Omega) \mid f(x) \neq 0\}$$

Remarque 7.2. On voit souvent dans de nombreuses sources que les ensembles $X(\Omega)$ et S_X soient confondus, mais il n'en est rien car Ω peut comporter des événements qui soient de probabilité nulle et il se peut alors que la probabilité que X prenne certaines valeurs de $X(\Omega)$ soit nulle.

Définition 7.3.

Proposition 7.4. (**Normalisation**) $\textcircled{S}$: $X : \Omega \rightarrow \mathbb{R}$ est une v.a.r.d. si $X(\Omega)$ est au plus dénombrable et

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

Définition 7.4. Si $X(\Omega)$ est ∞ et indénombrable alors X n'est pas discrète ( on ne peut rien dire d'autre sans étude supplémentaire !).

Définition 7.5. $\textcircled{S}\textcircled{C}$: $X = Y$ lorsque pour tout $\omega \in \Omega$, $X(\omega) = Y(\omega)$.

Définition 7.6. $\textcircled{S}\textcircled{C}$: $X = Y$ p.s. lorsque $\mathbf{P}([X = Y]) = 1$.

Définition 7.7. $\textcircled{S}$: $\mathcal{A}_X = \sigma([X = x]_{x \in X(\Omega)}) = \left\{ \bigsqcup_{x \in A} [X = x] \mid A \subset X(\Omega) \right\}$.

¹mais cela ne pose pas de problème ...

Définition 7.8. Loi de probabilité

⑤ : $P_X : X(\Omega) \rightarrow [0, 1]$, définie par

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

Remarque 7.3.  Une notation comme $\mathbf{P}(X)$ n'a aucun sens : les seules choses dont la probabilité est définie, ce sont les événements, or X n'est pas une partie de Ω mais une application !

Remarque 7.4. On vérifie très facilement qu'une loi de probabilité d'une variable $X : \Omega \rightarrow E$ est une probabilité sur l'ensemble E où X prend ses valeurs.

Remarque 7.5. La loi de probabilité d'une variable 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 7.9. 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.

Définition 7.10. 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 X .

Proposition 7.5. Caractérisation d'une densité

⑥ : (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 7.11. Fonction de répartition notée F

► ⑤ : $\forall x \in \mathbb{R}, F(x) = \sum_{\substack{t \in X(\Omega) \\ t \leq x}} \mathbf{P}([X = t]).$

► ⑥ : $\forall x \in \mathbb{R}, F(x) = \int_{-\infty}^x f(u) du.$

Propriétés 7.1. ► ⑤⑥ : 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.

► ⑥ : Dans le cas des variables 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).

Proposition 7.6. ⑥ : Soit X une v.a.r.a.d. 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 7.2. Calcul de probabilité à partir de la fonction de répartition

► ⑤ : $\forall a \in \mathbb{R}, \mathbf{P}([X = a]) = F_X(a) - F_X(a^-).$

► ⑥ : $\forall a \in \mathbb{R}, \mathbf{P}([X = a]) = 0.$

► ⑤ : $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X \leq b]) = F_X(b) - F_X(a^-).$

► ⑤ : $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X \leq b]) = F_X(b) - F_X(a).$

► ⑤ : $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X < b]) = F_X(b^-) - F_X(a^-).$

► ⑤ : $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X < b]) = F_X(b^-) - F_X(a).$

► ⑥ : Pour tout intervalle I de $\mathbb{R}$ on a, $\mathbf{P}([X \in I]) = \int_I f(x) dx.$

► ⑥ : $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X \leq b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx.$

► ⑥ : $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X \leq b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx.$

► ⑥ : $\forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a \leq X < b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx.$

$$\blacktriangleright \textcircled{C} : \forall (a, b) \in \mathbb{R}^2, a \leq b, \mathbf{P}([a < X < b]) = F_X(b) - F_X(a) = \int_a^b f(x) dx.$$

Remarque 7.6. $\textcircled{C}$: La densité est définie arbitrairement aux points où F n'est pas dérivable. Donc il n'y a pas d'unicité de la densité.

Définition 7.12. *Fonction de survie* (**HP mais ...**)

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

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

Proposition 7.7. *Caractérisation d'une loi*

- $\blacktriangleright \textcircled{S}$: 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 .
- $\blacktriangleright \textcircled{C}$: On donne $X(\Omega)$ puis une densité de X ou bien la fonction de répartition F de X .

Proposition 7.8. $\textcircled{S}\textcircled{C}$: 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 7.9. $\textcircled{S}$: Soit X une variable aléatoire telle que $X(\Omega) \subset \mathbb{Z}$ alors on a

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

où F est la fonction de répartition de X .

Proposition 7.10. $\textcircled{S}\textcircled{C}$: $(X \sim Y) \Leftrightarrow (F_X = F_Y)$ où F_X (resp. F_Y) est la fonction de répartition de X (resp. Y).

Remarque 7.7. Puisque la notation $X \sim Y$ est hors programme, on utilisera l'écriture $X \stackrel{L}{=} Y$.

Proposition 7.11. $\textcircled{S}\textcircled{C}$: $(X = Y) \Rightarrow (X \sim Y)$  la réciproque est fausse !

Proposition 7.12. $\textcircled{S}\textcircled{C}$: $(X \sim Y) \Rightarrow (\forall k \geq 0, \mathbf{E}(X^k) = \mathbf{E}(Y^k))$

$$\textbf{Proposition 7.13. } \forall A \subset X(\Omega), \mathbf{P}([X \in A]) = \begin{cases} \textcircled{S} : \sum_{x \in A} \mathbf{P}([X = x]) \\ \textcircled{C} : \int_{x \in A} f(x) dx \end{cases}$$

Propriétés 7.3. *Calcul de probabilité à partir de la fonction de survie*

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

2. Variables aléatoires particulières

Définition 7.13. *Variable aléatoire bornée*

$\textcircled{S}\textcircled{C}$: 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 7.1. Les variables de Dirac, uniformes, de Bernoulli, binomiales sont bornées.

Définition 7.14. *Variable aléatoire étagée*

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

Exemple 7.2. Toute variable de Bernoulli est une variable étagée.

Définition 7.15. *Variable aléatoire sans mémoire*

$\textcircled{S}\textcircled{C}$: 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 7.8. $\textcircled{S}\textcircled{C}$: 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 7.3. Les variables géométriques sont les seules variables discrètes sans mémoire et dans le cas continu, ce sont les variables exponentielles.

Définition 7.16. *Variable aléatoire symétrique* (**HP mais ...**)

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

$$X \sim -X$$

Définition 7.17. *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 7.4. $X \hookrightarrow \mathcal{U}(\{-1, 0, 1\})$.

Définition 7.18. *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 7.5. $X \hookrightarrow \mathcal{N}(0, 1)$.

3. Moments

Définition 7.19. *Espérance*

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

Proposition 7.14. $\textcircled{S}$: (**HP mais ...**) Si Ω est au plus dénombrable, alors

$$\mathbf{E}(X) = \sum_{\omega \in \Omega} X(\omega) \mathbf{P}(\{\omega\}) \text{ ssr } |\text{cv}|$$

Proposition 7.15. $\textcircled{S}\textcircled{C}$: $(X \text{ bornée}) \Rightarrow (\mathbf{E}(X) < \infty)$

Proposition 7.16. $\textcircled{S}\textcircled{C}$: $\forall (a, b) \in \mathbb{R}, (a \leq X \leq b) \Rightarrow (a \leq \mathbf{E}(X) \leq b)$

Proposition 7.17. $\textcircled{S}$: $\forall c \in \mathbb{R}, (\mathbf{P}([X = c]) = 1) \Rightarrow (\mathbf{E}(X) = c)$

Proposition 7.18. $\textcircled{S}\textcircled{C}$: $(\mathbf{E}(X) < \infty) \Leftrightarrow (\mathbf{E}(|X|) < \infty)$

Proposition 7.19. $\textcircled{S}\textcircled{C}$: (**HP mais ...**) Pour tout X variable positive nous avons

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

Proposition 7.20. *Espérance et fonction de survie*

► $\textcircled{S}$: (**HP mais ...**) Soit X une v.a.r. discrète telle que $X(\Omega) \subset \mathbb{N}$ 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,

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

► **©** : (**HP mais ...**) Soit X une v.a.r. 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 7.9.  Attention l'emploi de l'axiome de σ -additivité est inapplicable dans le cas discret, puisque les événements $([X > n])_{n \geq 0}$ ne sont pas disjoints. À méditer ...

Remarque 7.10. 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 $X \hookrightarrow \mathcal{E}(\lambda)$, $\int_0^{+\infty} \mathbf{P}([X > t]) dt = \int_0^{+\infty} e^{-\lambda t} dt = \frac{1}{\lambda}$.

Proposition 7.21. **Espérance d'une variable bornée**

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

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

Proposition 7.22. **©** : Soit $a \in \mathbb{R}$, et soit X une variable certaine égale à a alors $\mathbf{E}(X) = a$, i.e.

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

Proposition 7.23. **©©** : Si $X \geq 0$ alors $\mathbf{E}(X) \geq 0$ (positivité de l'espérance)

Proposition 7.24. **©©** : $(X \geq 0 \text{ et } \mathbf{E}(X) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

Corollaire 7.1. **©©** : $(\mathbf{E}(|X|) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

Corollaire 7.2. **©©** : $(\mathbf{E}(X^2) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

Corollaire 7.3. **©©** : $(\mathbf{E}((X - \mathbf{E}(X))^2) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

Proposition 7.25. **©©** : $\left. \begin{array}{l} \mathbf{E}(X) < \infty \\ Y \text{ est bornée} \end{array} \right\} \Rightarrow \mathbf{E}(XY) < \infty$.

Proposition 7.26. **©©** : $\forall (X, Y) \in (\mathcal{L}^1)^2$, $(X \leq Y) \Rightarrow (\mathbf{E}(X) \leq \mathbf{E}(Y))$ (croissance de l'espérance **admis dans le cas continu**)

Proposition 7.27. **©©** : $\forall (X, Y) \in (\mathcal{L}^1)^2$, $(X = Y) \Rightarrow (\mathbf{E}(X) = \mathbf{E}(Y))$

Théorème 7.1. **Théorème de domination (admis)**

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

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

Définition 7.20. **Moment d'ordre r où $r \geq 0$**

$$m_r(X) = \mathbf{E}(X^r) = \begin{cases} \text{©} : \sum_{x \in X(\Omega)} x^r \mathbf{P}([X = x]) \text{ ssr } |cv| \\ \text{©} : \int_{-\infty}^{+\infty} x^r f(x) dx \text{ ssr } |cv| \end{cases}$$

On note $\mathcal{L}^r(\Omega)$ voire $\mathcal{L}^r$ tout simplement, l'espace vectoriel des v.a.r. admettant un moment d'ordre r (**HP**).

Proposition 7.28. $\mathbb{S}\mathbb{C}$: $(m_2(X) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

Remarque 7.11. $\mathbb{C}$: Dans le cas continu, on pourrait se contenter de la convergence simple puisque l'intégrande ne change de signe qu'en 0, mais c'est une histoire de programme !

Proposition 7.29. $\mathbb{S}\mathbb{C}$: Toute v.a.r. bornée admet des moments de tous ordres.

Définition 7.21. *Moment absolu d'ordre r où $r \geq 0$*

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

Remarque 7.12. $\mathbb{C}$: Dans le cas continu, on pourrait se contenter de la convergence simple puisque l'intégrande ne change de signe qu'en 0, mais c'est une histoire de programme !

Proposition 7.30. $\mathbb{S}\mathbb{C}$: $(m_r(X) < \infty) \Leftrightarrow (m_r(|X|) < \infty)$

Définition 7.22. *Moment centré d'ordre r où $r \geq 0$*

Soit X une v.a.r. tel que $\mathbf{E}(X) < \infty$,

$$\mu_r(X) = \mathbf{E}((X - \mathbf{E}(X))^r) = \begin{cases} \mathbb{S} : \sum_{x \in X(\Omega)} (x - \mathbf{E}(X))^r \mathbf{P}([X = x]) \text{ ssr } |cv| \\ \mathbb{C} : \int_{-\infty}^{+\infty} (x - \mathbf{E}(X))^r f(x) dx \text{ ssr } |cv| \end{cases}$$

Remarque 7.13. 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 σ_X . Ainsi :

Définition 7.23. $\mathbb{S}\mathbb{C}$: On appelle écart-type de X le réel noté $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

Remarque 7.14. L'écart-type peut être considéré comme la distance euclidienne de X avec $\mathbf{E}(X)$. (**HP mais ...**)

Définition 7.24. $\mathbb{S}\mathbb{C}$: $\mathbf{V}(X) = \mu_2(X)$

Proposition 7.31. $\mathbb{S}\mathbb{C}$: $(\mathbf{V}(X) < \infty) \Leftrightarrow (\mathbf{E}(X^2) < \infty)$

Proposition 7.32. $\mathbb{S}\mathbb{C}$: $(\mathbf{V}(X) = 0) \Leftrightarrow (X = \mathbf{E}(X) \text{ p.s.})$

Proposition 7.33. $\mathbb{S}\mathbb{C}$: $(\mathbf{E}(X^2) = 0) \Rightarrow (\mathbf{V}(X) = 0)$

Proposition 7.34. $\mathbb{S}\mathbb{C}$: $(X \text{ bornée}) \Rightarrow (\mathbf{V}(X) < \infty)$

Théorème 7.2. *Théorème de Koenig-Huygens*

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

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

Définition 7.25. *Moment factoriel d'ordre r où $r \geq 1$ (**HP mais ...**)*

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

Proposition 7.35. $\mathbb{S}\mathbb{C}$: Soit $r \in \mathbb{N}$ alors $(m_{r+1}(X) < \infty) \Rightarrow (\forall k \in \llbracket 0, r \rrbracket, m_k(X) < \infty)$.

Remarque 7.15. $\mathbb{S}\mathbb{C}$: Nous avons donc pour tout $r \in \mathbb{N}^*$, $\mathcal{L}^r \subset \mathcal{L}^{r-1} \subset \cdots \subset \mathcal{L}^1$.

Proposition 7.36. $\mathbb{S}\mathbb{C}$: Soit $r \in \mathbb{N}$ alors $(\mu_{r+1}(X) < \infty) \Rightarrow (\forall k \in \llbracket 0, r \rrbracket, \mu_k(X) < \infty)$.

Théorème 7.3. Théorème de transfert (admis)

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

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

Exemple 7.6. Soit X une variable aléatoire uniformément distribuée sur $[0, 1]$. L'espérance de la variable $Y = X^5 + 2X^3 - X^2 - 3X$ 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 : pouvions-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 7.37. Translation $\textcircled{S}\textcircled{C} : \forall (a, b) \in \mathbb{R}^* \times \mathbb{R}, (\mathbf{E}(X) < \infty) \Leftrightarrow (\mathbf{E}(aX + b) < \infty)$ et

$$\mathbf{E}(aX + b) = a\mathbf{E}(X) + b$$

Proposition 7.38. Translation $\textcircled{S}\textcircled{C} : \forall (a, b) \in \mathbb{R}^* \times \mathbb{R}, (\mathbf{V}(X) < \infty) \Leftrightarrow (\mathbf{V}(aX + b) < \infty)$ et

$$\mathbf{V}(aX + b) = a^2\mathbf{V}(X)$$

Remarque 7.16.  Notez bien que le réel b a disparu du résultat précédent !

Exemple 7.7. $\mathbf{V}(9X + 21) = 81\mathbf{V}(X)$.

Proposition 7.39. $\textcircled{S}\textcircled{C} : \inf_{a \in \mathbb{R}} \mathbf{E}((X - a)^2) = \mathbf{V}(X)$ (**HP mais ...**)

Définition 7.26.

Proposition 7.40. Variable centrée réduite

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

Remarque 7.17. L'étape faisant passer de X à X^* s'appelle la standardisation, consistant à centrer et réduire la variable 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 8.1. Soit I un ensemble dénombrable indexé par $\mathbf{N}$ sous la forme $I = \{\varphi(n), n \in \mathbf{N}\}$ où φ 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 φ 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 8.1. *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 8.2. *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 8.1. *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 8.2. *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 en jeu dans cette sous-section sont définies dans un même espace probabilisé.

Définition 8.2. 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 8.3. Lois 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 8.3. Lois marginales

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

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

Définition 8.4. Lois conditionnelles

► **Loi conditionnelle de X sachant $[Y = y_j]$** avec $\mathbf{P}([Y = y_j]) \neq 0$

$$\mathbf{P}_{[Y=y_j]} : X(\Omega) \rightarrow [0, 1] \text{ définie par } \forall x_i \in X(\Omega), \quad \mathbf{P}_{[Y=y_j]}([X = x_i]) = \frac{p_{i,j}}{p_{\bullet,j}}.$$

► **Loi conditionnelle de Y sachant $[X = x_i]$** avec $\mathbf{P}([X = x_i]) \neq 0$

$$\mathbf{P}_{[X=x_i]} : Y(\Omega) \rightarrow [0, 1] \text{ définie par } \forall y_j \in Y(\Omega), \quad \mathbf{P}_{[X=x_i]}([Y = y_j]) = \frac{p_{i,j}}{p_{i,\bullet}}.$$

Proposition 8.4. Caractérisation de la loi d'une fonction d'un couple

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

► $Z(\Omega) = g(X, Y)(\Omega)$ et pour tout $z \in Z(\Omega)$,

► $\mathbf{P}([Z = z]) = \sum_{\left\{ \begin{smallmatrix} (x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y) = z \end{smallmatrix} \right\}} \mathbf{P}([X = x] \cap [Y = y])$ tel que $\sum_{z \in Z(\Omega)} \mathbf{P}([Z = z]) = 1$.

Proposition 8.5. $\forall I \in \mathcal{P}(Z(\Omega)), I \neq \emptyset, \mathbf{P}([Z \in I]) = \sum_{\left\{ \begin{smallmatrix} (x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y) \in I \end{smallmatrix} \right\}} \mathbf{P}([X = x] \cap [Y = y]).$

Théorème 8.3. ⑤ Théorème de transfert (admis)

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

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

Exemple 8.1. Soit X et Y deux variables discrètes à valeurs dans $\mathbf{N}^*$, alors en cas de convergence absolue de la série en jeu, $\mathbf{E}\left(\frac{X}{Y}\right) = \sum_{(i,j) \in (\mathbf{N}^*)^2} \frac{i}{j} \mathbf{P}([X = i] \cap [Y = j]).$

Conséquence 8.1. ⑤ **Linéarité de l'espérance** (*admis*)

Soit X et Y deux v.a.r. 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 8.1. ③ Le résultat reste valable pour les variables à densité et même aux variables aléatoires quelconques. Il se généralise au cas de n variables par récurrence.

Remarque 8.2. L'espérance est une forme linéaire positive sur le sous-espace vectoriel des variables aléatoires (soit discrètes, soit continues) admettant une espérance.

Théorème 8.4. ⑤ **Théorème de domination** (*admis*)

Soit X et Y deux v.a.r. 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 8.3. ③ Le résultat reste valable pour les variables à densité.

Remarque 8.4. 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 en jeu dans cette sous-section sont définies dans un même espace probabilisé.

3.1. Généralités.**Définition 8.5.** **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 8.5. $V(\Omega) \subset X_1(\Omega) \times \dots \times X_n(\Omega)$

Définition 8.6. **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 8.6.  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 8.6. $\forall k \in \llbracket 1, n \rrbracket, \quad \forall x_k \in X_k(\Omega),$

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

Théorème 8.5. **Théorème de transfert**

Soit $V = (X_1, \dots, X_n)$ et φ une fonction définie sur $\mathbb{R}^n$ (il est suffisant d'avoir φ définie sur $V(\Omega)$) alors

$$\mathbf{E}(\varphi(X_1, \dots, X_n)) = \sum_{(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)} \varphi(x_1, \dots, x_n) \mathbf{P}\left(\bigcap_{j=1}^n [X_j = x_j]\right)$$

ssr | cv |.

Conséquence 8.2. 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 8.7.  La linéarité de l'espérance n'exige pas l'indépendance des variables en jeu !

Proposition 8.7. *Caractérisation de la loi d'une fonction d'un vecteur*

Soit $n \geq 2$ et $V = (X_1, \dots, X_n)$ et φ une fonction définie sur une partie de $\mathbb{R}^n$ 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

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

pour tout z de $Z(\Omega)$ avec

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

Proposition 8.8. (*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

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

3.2. Espérance conditionnelle.

Définition 8.7.  *Espérance conditionnelle*

$$\forall A \in \mathcal{A}, \quad \mathbf{P}(A) \neq 0, \quad \mathbf{E}(X | A) = \sum_{x \in X(\Omega)} x \mathbf{P}_A([X = x]) \text{ srr } |cv|$$

Proposition 8.9.  : $\forall A \in \mathcal{A}, \mathbf{P}(A) \neq 0, (\mathbf{E}(X) < \infty) \Rightarrow (\mathbf{E}(X | A) < \infty)$.

Propriétés 8.1.  Soit X, Y, Z et T quatre v.a.r.d. et g une fonction de deux variables. Enfin a, b, c trois constantes réelles.

- ▶ $\forall x \in \mathcal{S}_X, \forall c \in \mathbb{R}, \mathbf{E}(c \mathbf{1}_\Omega | [X = x]) = c$.
- ▶ $\forall y \in \mathcal{S}_Y, \mathbf{E}(X | [Y = y]) = \frac{\mathbf{E}(X \mathbf{1}_{[Y=y]})}{\mathbf{P}([Y = y])}$ (*HP mais ...*).
- ▶ $\forall x \in \mathcal{S}_X, \mathbf{E}(aZ + bT | [X = x]) = a\mathbf{E}(Z | [X = x]) + b\mathbf{E}(T | [X = x])$ (linéarité).
- ▶ Si $X \perp\!\!\!\perp Y$ 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 8.8. Ces propriétés restent valables en continu mais sont hors programme.

Remarque 8.9. Les espérances conditionnelles peuvent servir à construire des fonctions de régression qui étudient la valeur moyenne d'une variable conditionnée par une autre.

Proposition 8.10.  *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

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

En cas d'existence

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

Définition 8.8.  *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 $g_1(x) = \mathbf{E}(Y | [X = x])$. De même on appelle fonction de régression de X sur Y , la fonction $g_2 : \mathcal{S}_Y \rightarrow \mathbb{R}$ définie par $g_2(y) = \mathbf{E}(X | [Y = y])$ à condition que $\mathbf{E}(X) < \infty$.

3.3. Covariance.

Définition 8.9. $\textcircled{S}\textcircled{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$, $\text{Cov}(X, Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$.

Remarque 8.10. $\textcircled{S}\textcircled{C}$: Cette définition ne sert quasiment jamais au calcul d'une covariance car elle n'est pas très pratique à manipuler.

Remarque 8.11. $\textcircled{S}\textcircled{C}$: 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 augmente, l'autre diminue. Et un signe négatif s'interprète en disant qu'en moyenne si une variable tend à augmenter, l'autre diminue.

Définition 8.10. *Variables aléatoires non corrélées ou décorrélées*

$\textcircled{S}\textcircled{C}$: On dit que deux variables X et Y sont non corrélées ou décorrélées lorsque $\text{Cov}(X, Y) = 0$.  Cela ne veut pas dire qu'elles sont indépendantes !

Remarque 8.12. $\textcircled{S}\textcircled{C}$:  Faites toujours la distinction entre indépendant et non-corrélation. Ne concluez jamais à l'indépendance après avoir obtenu un coefficient de corrélation nul !

Théorème 8.6. *Théorème de Koenig-Huygens*

$\textcircled{S}\textcircled{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$, $\text{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$.

Proposition 8.11. $\textcircled{S}\textcircled{C}$: $\left. \begin{array}{l} \mathbf{E}(X^2) < \infty \\ \mathbf{E}(Y^2) < \infty \end{array} \right\} \Rightarrow (\mathbf{E}(XY) < \infty)$

Remarque 8.13. La proposition précédente donne une condition suffisante d'existence de la covariance du couple (X, Y) .

Propriétés 8.2. $\textcircled{S}\textcircled{C}$: $\text{Cov}(X, X) = \mathbf{V}(X) \geq 0$ (positivité)

$\textcircled{S}\textcircled{C}$: $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ (symétrie)

$\textcircled{S}\textcircled{C}$: Soit $(X, Y, Z) \in (\mathcal{L}^2)^3$ et $\forall (a, b) \in \mathbb{R}^2$, $\text{Cov}(aX + bY, Z) = a \text{Cov}(X, Z) + b \text{Cov}(Y, Z)$ (linéarité à gauche)

$\textcircled{S}\textcircled{C}$: $\forall X \in \mathcal{L}^2$, $\forall a \in \mathbb{R}$, $\text{Cov}(X, a\mathbf{1}_\Omega) = 0$

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

$\textcircled{S}\textcircled{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ et $\forall (a, b, c, d) \in \mathbb{R}^4$, $\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y)$.

$\textcircled{S}\textcircled{C}$: $\forall k \in \llbracket 1, n \rrbracket$, $\forall \ell \in \llbracket 1, m \rrbracket$, X_k et $Y_\ell \in \mathcal{L}^2$, λ_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 8.2. $\text{Cov}(2X + 3Y, Z - W) = 2 \text{Cov}(X, Z) - 2 \text{Cov}(X, W) + 3 \text{Cov}(Y, Z) - 3 \text{Cov}(Y, W)$

Remarque 8.14. $\textcircled{S}\textcircled{C}$  La covariance ne définit pas en général un produit scalaire, sauf si l'on se réduit aux variables de $\mathcal{L}^2$ qui soient centrées et en assimilant variable nulle p.s. et variable nulle.

3.4. Matrice de covariance-variance.

Définition 8.11. *Matrice de covariance-variance d'un vecteur aléatoire*

$\textcircled{S}$ Soit pour tout k de $\llbracket 1, n \rrbracket$, $X_k \in \mathcal{L}_d^2$, 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 Γ_V définie par $\Gamma_V = (\text{Cov}(X_i, X_j))_{(i, j) \in \llbracket 1, n \rrbracket^2}$.

Remarque 8.15. $\textcircled{S}$ La matrice de covariance-variance associée à un vecteur nous donne une idée de la dispersion du vecteur.

Proposition 8.12. $\textcircled{S}$ Si les variables $X_1, \dots, X_n$ sont indépendantes, Γ_V est diagonale car toutes les covariances $\text{Cov}(X_i, X_j)$ (avec $i \neq j$) sont nulles.

Proposition 8.13. *Forme quadratique associée à Γ_V*

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

Propriétés 8.3. La forme quadratique associée à Γ_V montre que la matrice Γ_V est positive et donc ses valeurs propres sont positives ou nulles. Ainsi la matrice de covariance Γ_V est symétrique positive.

3.5. Variance d'une somme.

Propriétés 8.4. $\mathbb{S}\mathbb{C}$: $\forall (X, Y) \in (\mathcal{L}^2)^2$ 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 8.3. $\mathbf{V}(2X - Y) = 4\mathbf{V}(X) + \mathbf{V}(Y) - 4\text{Cov}(X, Y)$

Propriétés 8.5. $\mathbb{S}\mathbb{C}$: $\forall (X_1, \dots, X_n) \in (\mathcal{L}^2)^n$ et $\forall (a_1, \dots, a_n) \in \mathbb{R}^n$,

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

Corollaire 8.1. Si de plus les variables $X_1, \dots, X_n$ sont de même loi alors

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

Proposition 8.14. *Identité du parallélogramme*

$\mathbb{S}\mathbb{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors

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

Proposition 8.15. *Identités de polarisation¹*

$\mathbb{S}\mathbb{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors,

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

Remarque 8.16. On retrouve les identités remarquables précédentes pour les nombres 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 8.16. $\mathbb{S}\mathbb{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors,

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

Remarque 8.17. On retrouve l'identité remarquable précédente pour les nombres et les produits scalaires. En effet

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

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

Proposition 8.17. $\textcircled{S}\textcircled{C}$: Si $(X_1, \dots, X_n) \in (\mathcal{L}^2)^n$ 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 8.18.  La réciproque est fausse.

3.6. Coefficient de corrélation linéaire.

Définition 8.12. *Coefficient de corrélation linéaire*

$\textcircled{S}\textcircled{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ 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 8.19. $\textcircled{S}\textcircled{C}$: Le coefficient de corrélation mesure la qualité d'une relation linéaire entre deux variables X et Y . C'est un indicateur de dépendance linéaire entre les variables X et Y , c'est-à-dire d'à quel point Y peut être approchée par des fonctions affines de X .

Propriétés 8.6. $\textcircled{S}\textcircled{C}$: Le coefficient de corrélation est invariant par changement d'échelle

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

Si de plus $d \neq 0$,

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

Propriétés 8.7. *L'inégalité de Cauchy-Schwarz (ICS)*

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

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

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

Remarque 8.20. $\textcircled{S}\textcircled{C}$: Plus $|\rho(X, Y)|$ est proche de 1, plus l'ajustement linéaire entre X et Y est pertinent.

3.7. Indépendance.

3.7.1. Définition et proposition.

Définition 8.13. $\textcircled{S}\textcircled{C}$: Soit $n \geq 2$ et $X_1 \perp \dots \perp X_n$ (on dit que les variables sont indépendantes ou mutuellement indépendantes) lorsque pour tous $B_1, \dots, B_n$, boréliens de $\mathbb{R}$ on a

$$\mathbf{P} \left(\bigcap_{i=1}^n [X_i \in B_i] \right) = \prod_{i=1}^n \mathbf{P}([X_i \in B_i])$$

Proposition 8.18. *Deux caractérisations de l'indépendance de $n \geq 2$ v.a.r. (admis)*

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

► $(X_1 \perp \dots \perp X_n) \Leftrightarrow (\forall (x_1, \dots, x_n) \in \mathbb{R}^n, \mathbf{P} \left(\bigcap_{i=1}^n [X_i \leq x_i] \right) = \prod_{i=1}^n \mathbf{P}([X_i \leq x_i]))$

Proposition 8.19. *Une caractérisation de l'indépendance de $n \geq 2$ v.a.r. discrètes (admis)*

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

Proposition 8.20. $\textcircled{S}\textcircled{C}$ Soit X une variable certaine, alors pour toute variable aléatoire Y on a $X \perp\!\!\!\perp Y$. 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.

Proposition 8.21. $\textcircled{S}\textcircled{C}$: Pour tout X variable aléatoire et pour tout a réel, $X \perp\!\!\!\perp a \mathbf{1}_\Omega$.

Remarque 8.21. Autrement dit, toute variable constante est indépendante de toute variable y compris d'elle-même.

Remarque 8.22.  Il n'y a aucune relation entre la notion d'indépendance et celle d'indépendance conditionnelle.

Prenez l'exemple de deux variables 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 $\mathbf{P}_{[S_2=0]}([X_1=1] \cap [X_2=1]) \neq \mathbf{P}_{[S_2=0]}([X_1=1]) \mathbf{P}_{[S_2=0]}([X_2=1])$.

3.7.2. **Le lemme des coalition ou la stabilité de l'indépendance par composition.**

Proposition 8.22. (Cas de deux variables **admis**) $\textcircled{S}\textcircled{C}$: Soit X, Y deux variables et soit $(\varphi, \psi) \in \mathcal{F}(I, \mathbb{R}) \times \mathcal{F}(J, \mathbb{R})$ | $I \supset X(\Omega)$ et $J \supset Y(\Omega)$ alors $\varphi(X)$, $\psi(Y)$ restent des variables, et

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

Remarque 8.23.  La réciproque est fausse. En effet considérer les variables X et Y tel que $Y = X^2$ et $\varphi : t \mapsto 6$ et $\psi : t \mapsto 7$, alors $\varphi(X) \perp\!\!\!\perp \psi(Y)$ bien que les variables X et Y ne soient pas indépendantes.

Proposition 8.23. (Cas de n variables **admis**) $\textcircled{S}\textcircled{C}$: Soit $n \geq 2$ et $X_1, \dots, X_n$ n variables et pour tout $(\varphi, \psi) \in \mathcal{F}(D, \mathbb{R}) \times \mathcal{F}(D', \mathbb{R})$ tel que $D \supset X_1(\Omega) \times \dots \times X_k(\Omega)$ et $D' \supset X_{k+1}(\Omega) \times \dots \times X_n(\Omega)$ 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 alors

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

Remarque 8.24.  Cette propriété peut s'étendre à un nombre fini de fonctions s'appliquant à une partition des variables et en particulier au cas de $(\varphi_1(X_1), \varphi_2(X_2), \dots, \varphi_n(X_n))$.

Exemple 8.4. Pour n et m deux entiers naturels non nuls et des v.a.r. 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 indépendantes.

Proposition 8.24. (**admis**) $\textcircled{S}\textcircled{C}$: $\forall n \geq 2, \forall (X_1, \dots, X_n) \in (\mathcal{L}^1)^n$,

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

Proposition 8.25. $\textcircled{S}\textcircled{C}$: $\forall (X, Y) \in (\mathcal{L}^2)^2$, $(X \perp\!\!\!\perp Y) \Rightarrow (\text{Cov}(X, Y) = 0)$

Remarque 8.25.  La réciproque est fausse en général (sauf si les deux variables sont au plus binaires ou gaussiennes (**HP mais ...**))

Remarque 8.26.  La réciproque est fausse. En effet considérer $X \hookrightarrow \mathcal{U}(\{-1, 0, 1\})$ et $Y = X^2$.

Proposition 8.26. $\textcircled{S}\textcircled{C}$: $\forall (X, Y) \in (\mathcal{L}^2)^2$, $(\text{Cov}(X, Y) \neq 0) \Rightarrow \text{non}(X \perp\!\!\!\perp Y)$ (par contraposée)

Proposition 8.27. $\textcircled{S}\textcircled{C}$: $\forall (X, Y) \in (\mathcal{L}^2)^2$, $(X \perp\!\!\!\perp Y) \Rightarrow (\rho(X, Y) = 0)$

Remarque 8.27.  La réciproque est fausse en général (sauf si les deux variables sont au plus binaires ou gaussiennes (**HP mais ...**))

Proposition 8.28. $\mathbb{S}\mathbb{C} : \forall (X, Y) \in (\mathcal{L}^2)^2, (\rho(X, Y) \neq 0) \Rightarrow \text{non}(X \perp Y)$ (par contraposée)

Proposition 8.29. $\mathbb{S} : \left(\forall (X, Y) \in (\mathcal{L}^1)^2 \text{ et } X \perp Y \right) \Rightarrow \left(\begin{cases} \mathbf{E}(X \mid [Y = y_j]) = \mathbf{E}(X) \\ \mathbf{E}(Y \mid [X = x_i]) = \mathbf{E}(Y) \end{cases} \right)$
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 8.30. (*admis*) $\mathbb{S}\mathbb{C} : \forall n \geq 2, \forall (X_1, \dots, X_n) \in \mathbb{R}^n, \forall (X_1, \dots, X_n) \in (\mathcal{L}^2)^n,$

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

(quadraticité de la variance)

3.7.3. Convolution.

Définition 8.14. *Produit de convolution*

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

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

► $\mathbb{C}$ 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 8.28. $\mathbb{C}$ L'intégrale est convergente, par exemple, si au moins une des deux fonctions f ou g est à support compact.

Remarque 8.29. $\mathbb{S}\mathbb{C}$  Le théorème qui va suivre ne s'emploie qu'en cas de recherche de la loi de la somme de deux v.a.r. X et Y indépendantes. Il existe des exceptions, par exemple avec des variables de Cauchy dépendantes, mais ce n'est pas au programme.

Théorème 8.7. (*admis*)

► $\mathbb{S}$ Soit X et Y avec $X \perp Y$ alors

$$\forall z \in (X + Y)(\Omega), \quad \mathbf{P}([X + Y = z]) = \sum_{x \in \left\{ \begin{smallmatrix} x \in X(\Omega) \\ z-x \in Y(\Omega) \end{smallmatrix} \right\}} \mathbf{P}([X = x]) \mathbf{P}([Y = z - x])$$

► $\mathbb{C}$ Soit X et Y avec $X \perp Y$ alors la fonction f_{X+Y} définie par

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

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 8.30. On dit que la loi de la somme est la convolée des lois de X et de Y .

Remarque 8.31. Un excellent moyen de s'entraîner à "convoler", que ce soit en discret ou en continu, consiste à prendre deux variables uniformes indépendantes.

Remarque 8.32. En continu il est très fortement conseillé de déterminer les valeurs de t pour lesquelles le produit $f_X(t) f_Y(z-t)$ soit $f_X(z-t) f_Y(t)$ n'est pas nul.

Propriétés 8.8. $\mathbb{C}$

► Le produit de convolution est commutatif et associatif. Autrement dit avec les notations en continu

$$f_{X+Y} = f_X * f_Y = f_Y * f_X \quad \text{et} \quad (f_X * f_Y) * f_Z = f_X * (f_Y * f_Z)$$

- 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 8.3. Le théorème de convolution nous permet de conclure les lois Binomiale, de Poisson, gamma et normale sont stables par convolution. Il y en a beaucoup d'autres ...

4. Droites de régression

Définition 8.15.

Proposition 8.31. Droites de régression

- $\textcircled{S}$: Droite de régression de Y en X .
Elle est notée Δ et a pour équation :

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

- $\textcircled{S}$: Droite de régression de X en Y .
Elle est notée Δ' et a pour équation :

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

Remarque 8.33. Ces équations de droite restent valables en continu mais ne sont pas au programme.

5. Produit scalaire dans $\mathcal{L}^2$

Proposition 8.32. Produit scalaire et norme associée dans $\mathcal{L}^2$ (**HP mais ...** déjà tombé !)

$\textcircled{S}\textcircled{C}$: L'application

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

définit un produit scalaire dans $\mathcal{L}^2$ en assimilant variable nulle p.s. et variable nulle. La norme associée est définie par

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

Exemple 8.5. $\forall X \in \mathcal{L}^2$, 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 .

Inégalités probabilistes

1. Inégalités de concentration

Définition 9.1. $\textcircled{S}\textcircled{C}$: 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 9.1. *Inégalité de Markov (IM)*

► $\textcircled{S}\textcircled{C}$: Soit $X \in \mathcal{L}^1$ à valeurs positives, alors

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

(IM à l'ordre 1)

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

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

(IM à l'ordre r)

► $\textcircled{S}\textcircled{C}$: (**HP mais ...**) Si $\varphi : I \mapsto \mathbf{R}_+^*$ est une fonction bijective et strictement croissante sur $I \supset X(\Omega)$ 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 9.1. $\textcircled{S}\textcircled{C}$: Dans le cas d'une variable 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 9.2. *Inégalité de Bienaymé-Tchebychev (IBT)*

► $\textcircled{S}\textcircled{C}$: Soit $X \in \mathcal{L}^2$ admettant une variance non nulle, alors

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

► $\textcircled{S}\textcircled{C}$: (**HP mais ...**) 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 9.2. 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 9.3. *Inégalité de Chernoff (HP mais ...)*

$\textcircled{S}\textcircled{C}$: Soit X une v.a.r. 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 (ICS).

Théorème 9.4. ► $\textcircled{SC}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors $\text{Cov}(X, Y) < \infty$ et

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

D'autres versions

► $\textcircled{SC}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ 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.

► $\textcircled{SC}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors $\mathbf{E}(XY) < \infty$ et

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

Exemple 9.1. $\mathbf{E}(|X|) \leq \sqrt{\mathbf{E}(X^2)}$

Remarque 9.3. On peut utiliser l'ICS à partir d'une variable X positive, en remarquant que l'on peut écrire $X = X \mathbf{1}_{[X>0]}$.

Remarque 9.4. Avec le produit scalaire défini en proposition 8.32, l'ICS peut s'écrire $(\langle X, Y \rangle)^2 \leq \|X\|^2 \|Y\|^2$. (**HP mais ...**)

2.2. Inégalité triangulaire.

Théorème 9.5. $\textcircled{SC}$: 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 9.5. $\textcircled{SC}$: Cette inégalité se généralise au cas de n variables aléatoires.

2.3. Inégalité de Jensen (**HP mais ...**)

Théorème 9.6. $\textcircled{SC}$: Soit X une variable aléatoire à valeurs dans un intervalle I et φ une fonction convexe sur un intervalle $I \supset X(\Omega)$, alors si X et $\varphi(X)$ admettent une espérance on a,

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

Remarque 9.6. $\textcircled{SC}$:  En général $\varphi(\mathbf{E}(X)) \neq \mathbf{E}(\varphi(X))$.

Exemple 9.2. ► $\exp(\mathbf{E}(X)) \leq \mathbf{E}(\exp(X))$

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

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

Proposition 9.1. $\textcircled{SC}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors $\mathbf{E}(|XY|) < \infty$ et

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

Corollaire 9.1. $\textcircled{SC}$: Soit $X \in \mathcal{L}^2$ alors $\mathbf{E}(|X|) < \infty$ et

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

Remarque 9.7. 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 indicatrice.

Proposition 9.2. $\textcircled{\text{S}}\textcircled{\text{C}}$: $\forall \varepsilon > 0, \mathbf{E} \left(X \mathbf{1}_{[X \leq \varepsilon]} \right) \leq \varepsilon$.

Remarque 9.8. *Ce résultat n'est pas dans le cours, il faut donc savoir le démontrer.*

Convergences

1. Convergence en probabilités

Définition 10.1. $\textcircled{S}\textcircled{C}$: $(X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X$ lorsque

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

soit encore

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

Remarque 10.1. Toutes les variables en jeu étant définies sur le même espace probabilisé, c'est ce que l'on appelle une convergence spciale.

Remarque 10.2. $\textcircled{S}\textcircled{C}$: $\forall \varepsilon > 0, [|X_n - X| \geq \varepsilon] \subset [|X_n - X| > \frac{\varepsilon}{2}]$.

Propriétés 10.1. $\textcircled{S}\textcircled{C}$: $\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X\right) \Leftrightarrow \left((X_n - X)_{n \geq 1} \xrightarrow{\mathbf{P}} 0\right)$.

Propriétés 10.2. $\textcircled{S}\textcircled{C}$: $\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X\right) \not\Leftrightarrow \left(\lim_{n \rightarrow +\infty} \mathbf{E}(X_n) = \mathbf{E}(X) \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}(X_n - X) = 0\right)$

Propriétés 10.3. $\textcircled{S}\textcircled{C}$: $\left(\lim_{n \rightarrow +\infty} \mathbf{E}(X_n - X) = 0 \in \mathbb{R} \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}(X_n - X) = 0\right) \Rightarrow \left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X\right)$

Corollaire 10.1. $\textcircled{S}\textcircled{C}$: $\left(\lim_{n \rightarrow +\infty} \mathbf{E}(X_n) = a \in \mathbb{R} \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}(X_n) = 0\right) \Rightarrow \left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} a\right)$

Propriétés 10.4. (**HP mais ...** à savoir démontrer ...)

► $\textcircled{S}\textcircled{C}$: $\left(|X_n|_{n \geq 1} \xrightarrow{\mathbf{P}} 0\right) \Leftrightarrow \left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} 0\right)$

► $\textcircled{S}\textcircled{C}$: $\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y\right) \Rightarrow (X = Y \text{ p.s.})$

► $\textcircled{S}\textcircled{C}$: $\forall (a, b) \in \mathbb{R}^2, \left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y\right) \Rightarrow \left((aX_n + bY_n)_{n \geq 1} \xrightarrow{\mathbf{P}} aX + bY\right)$

► $\textcircled{S}\textcircled{C}$: $\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y\right) \Rightarrow \left((X_n Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} XY\right)$

► $\textcircled{S}\textcircled{C}$: $\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X \text{ et } (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} Y\right) \Rightarrow \left(\left(\frac{X_n}{Y_n}\right)_{n \geq 1} \xrightarrow{\mathbf{P}} \frac{X}{Y} \text{ avec } \mathbf{P}([Y = 0]) = 0\right)$

Théorème 10.1. **Théorème de composition** (**admis**)

$\textcircled{S}\textcircled{C}$: Si $(X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X$ et si $f \in C^0(\mathbb{R}, \mathbb{R})$ alors

$$(f(X_n))_{n \geq 1} \xrightarrow{\mathbf{P}} f(X)$$

Remarque 10.3. 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.

Remarque 10.4. $\textcircled{S}\textcircled{C}$: Ce résultat s'applique essentiellement en concours aux transformations affines, exponentielles, à la valeur absolue, à l'élevation à une puissance entière.

Théorème 10.2. **Loi faible des grands nombres**

$\textcircled{S}\textcircled{C}$: Soit $(X_n)_{n \geq 1}$ une suite de variables indépendantes, d'espérance commune m et de variance commune

σ^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 10.5. $\textcircled{S} : \text{Si } \forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{B}(p)$ 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 10.2. $\textcircled{S} \textcircled{C} : (X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X$ lorsque $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$ en tout point x où F_X est continue.

Remarque 10.5. Les fonctions F_{X_n} et F_X représentent la fonction de répartition de X_n (resp. de X), les variables 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.

Remarque 10.6. 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 10.1. $\textcircled{S} \textcircled{C} : \text{Soit } a \text{ et } b \text{ deux réels tel que } a < b \text{ alors on a}$

$$\lim_{n \rightarrow +\infty} \mathbf{P}([a < X_n \leq b]) = \mathbf{P}([a < X \leq b])$$

Proposition 10.2. $\textcircled{S} : \left((X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X \right) \Leftrightarrow \left(\forall k \in \mathbb{Z}, \lim_{n \rightarrow +\infty} \mathbf{P}([X_n = k]) = \mathbf{P}([X = k]) \right)$ où les variables X_n et X sont à valeurs dans $\mathbb{Z}$.

Remarque 10.7.  N'employer cette proposition qu'en cas où vous soyez sûrs de connaître la nature discrète de TOUTES les variables en jeu.

Théorème 10.3. *Théorème de composition (admis)*

$\textcircled{S} \textcircled{C} : \text{Si } (X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X \text{ et si } f \in \mathcal{C}^0(\mathbb{R}, \mathbb{R}) \text{ alors on a}$

$$(f(X_n))_{n \geq 1} \xrightarrow{\mathcal{L}} f(X)$$

Remarque 10.8. 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.

Remarque 10.9. $\textcircled{S} \textcircled{C} : \text{Ce résultat s'applique essentiellement en concours aux transformations affines, exponentielles, à la valeur absolue, à l'élévation à une puissance entière.}$

Théorème 10.4. *Théorème de Slutsky (admis)*

Propriétés 10.6. $\textcircled{S} \textcircled{C} : \left. \begin{array}{l} (X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X \\ (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} c \in \mathbb{R} \end{array} \right\} \Rightarrow (X_n + Y_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X + c$

Propriétés 10.7. $\textcircled{S} \textcircled{C} : \left. \begin{array}{l} (X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X \\ (Y_n)_{n \geq 1} \xrightarrow{\mathbf{P}} c \in \mathbb{R} \end{array} \right\} \Rightarrow (X_n Y_n)_{n \geq 1} \xrightarrow{\mathcal{L}} cX$

Remarque 10.10.  Slutsky mélange les types de convergence.

Théorème 10.5. *Théorème central limite (TCL) (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 $n \in \mathbb{N}^*$, $S_n = \sum_{k=1}^n X_k$

et $S_n^* = \frac{S_n - \mathbf{E}(S_n)}{\sigma(S_n)}$ 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 TCL 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

$$\left(\overline{X}_n^* \right)_{n \geq 1} = \left(\sqrt{n} \left(\frac{\overline{X}_n - \mu}{\sigma} \right) \right) \xrightarrow{\mathcal{L}} N$$

où $N \hookrightarrow \mathcal{N}(0, 1)$.

Remarque 10.11.  On ne dit pas théorème de la limite centrée !

Corollaire 10.2. (*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 10.3. ► (*HP mais ...*) $\textcircled{S}\textcircled{C}$: Soit $(X_n)_{n \geq 1}$ et X des variables 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 10.12. La réciproque de cette implication est fausse, il suffit de considérer $X_n = (-1)^n X$ où $X \hookrightarrow \mathcal{N}(0, 1)$.

Proposition 10.4. ► $\textcircled{S}\textcircled{C}$: $\forall a \in \mathbb{R}$, $\left((X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} a \right) \Leftrightarrow \left((X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} a \right)$.

Approximations

Proposition 11.1. *Binomiale par Poisson*

$\forall n \in \mathbf{N}^*$, $X_n \hookrightarrow \mathcal{B}(n, p_n)$ 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 $p \leq 0,1$).

Proposition 11.2. *Binomiale par Normale (admis)*

Soit pour tout $n \in \mathbf{N}^*$, $X_n \hookrightarrow \mathcal{B}(p)$ des variables 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 11.3. *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 11.1. 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 $n \in \mathbb{N}^*$ et $\theta \in \Theta$ où Θ est une partie de $\mathbb{R}$ et $g : \Theta \rightarrow \mathbb{R}$. Le but de l'estimation inférentielle est celle de l'estimation d'un paramètre $g(\theta)$ d'une loi suivie par une variable X dont on connaît la loi générique (mais non les paramètres) appartenant à une famille de loi notée μ_θ .

Définition 12.1. *n-échantillon d'une variable aléatoire*

ⓈⓈ : On appelle un n -échantillon $(X_1, \dots, X_n)$ d'une variable X dont on cherche à estimer un ou des paramètres de sa loi (connue), un vecteur aléatoire de dimension n constitué de variables X_k i.i.d. suivant toutes la loi de X .

Définition 12.2. *Statistique ou estimateur d'un paramètre*

ⓈⓈ : Soit $(X_1, \dots, X_n)$ un n -échantillon de X , et $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$. On appelle statistique ou estimateur de $g(\theta)$ toute suite de variables aléatoires $(T_n)_{n \geq 1}$ où pour tout $n \in \mathbb{N}^*$, $T_n = \varphi(X_1, \dots, X_n)$ est indépendante de θ et dont les valeurs appartiennent à $g(\Theta)$.

Définition 12.3. *Estimation*

ⓈⓈ : On appelle estimation de $g(\theta)$ toute réalisation de la variable T_n soit $\varphi(X_1(\omega), \dots, X_n(\omega)) = \varphi(x_1, \dots, x_n)$ où $\omega \in \Omega$.

Définition 12.4. *Estimateur convergent ou consistant d'un paramètre*

ⓈⓈ : On dit qu'un estimateur est convergent ou consistant lorsque

$$(T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta)$$

Définition 12.5. *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 12.1. *Conditions suffisantes qu'un estimateur soit convergent ou consistant*

► ⓈⓈ : $\left(b_\theta(T_n) = 0 \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}_\theta(T_n) = 0 \right) \Rightarrow \left((T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta) \right)$

► ⓈⓈ : $\left(\lim_{n \rightarrow +\infty} b_\theta(T_n) = 0 \text{ et } \lim_{n \rightarrow +\infty} \mathbf{V}_\theta(T_n) = 0 \right) \Rightarrow \left((T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta) \right)$

Définition 12.6. *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 12.1. ⓈⓈ : $\forall \theta \in \Theta$, $r_\theta(T_n) = \|T_n - g(\theta)\|^2$ (**HP mais ...**)

Proposition 12.2. $\textcircled{S}\textcircled{C} : \forall \theta \in \Theta, r_\theta(T_n) = \mathbf{E}_\theta \left((T_n - g(\theta))^2 \right)$

Proposition 12.3. $\textcircled{S}\textcircled{C} : \forall \theta \in \Theta, \left(\lim_{n \rightarrow +\infty} r_\theta(T_n) = 0 \right) \Rightarrow \left((T_n)_{n \geq 1} \xrightarrow{\mathbf{P}} g(\theta) \right)$

Proposition 12.4. $\textcircled{S}\textcircled{C} : \forall \theta \in \Theta, r_\theta(T_n) = \mathbf{V}_\theta(T_n) + b_\theta^2(T_n)$

Proposition 12.5. $\textcircled{S}\textcircled{C} : \forall \theta \in \Theta, (r_\theta(T_n) = \mathbf{V}_\theta(T_n)) \Leftrightarrow (b_\theta(T_n) = 0)$

Proposition 12.6. $\textcircled{S}\textcircled{C} : \forall \theta \in \Theta, \left(\lim_{n \rightarrow +\infty} r_\theta(T_n) = 0 \right) \Leftrightarrow \left(\lim_{n \rightarrow +\infty} \mathbf{V}_\theta(T_n) = 0 \text{ et } \lim_{n \rightarrow +\infty} b_\theta(T_n) = 0 \right)$

Remarque 12.2. $\textcircled{S}\textcircled{C} : \text{Le meilleur moyen de comparer deux estimateurs sera de prendre celui dont le risque quadratique est le plus petit.}$

1.2. Exemples d'estimateurs de paramètres usuels et sans biais et convergents.

1.2.1. Moyenne empirique

. *Situation* : Soit une variable X suivant une loi d'espérance inconnue $\mu \in \mathbb{R}$ mais de variance $\sigma^2 \in \mathbb{R}_+^*$ connue.

Définition 12.7. Moyenne empirique

$\textcircled{S}\textcircled{C} : \text{Soit } (X_1, \dots, X_n) \text{ un } n\text{-échantillon de } X \text{ d'espérance inconnue } X, \text{ on appelle moyenne empirique la variable aléatoire } \overline{X}_n \text{ définie pour tout } n \in \mathbb{N}^* \text{ par } \overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k \text{ où pour tout } k \in \llbracket 1, n \rrbracket, \mathbf{E}_\mu(X_k) = \mu \text{ (inconnue).}$

Proposition 12.7. $\textcircled{S}\textcircled{C} : \forall \mu \in \mathbb{R}, \mathbf{E}_\mu(\overline{X}_n) = \mu \text{ et } \mathbf{V}_\mu(\overline{X}_n) = \frac{\sigma^2}{n}.$

Proposition 12.8. $\blacktriangleright \textcircled{C} : (\forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{N}(\mu, \sigma^2)) \Rightarrow \left(\overline{X}_n \hookrightarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \right) \text{ où } \mu \in \mathbb{R} \text{ et } \sigma^2 \in \mathbb{R}_+^*.$

$\blacktriangleright \textcircled{S}\textcircled{C} : (\forall k \in \llbracket 1, n \rrbracket, X_k \not\hookrightarrow \mathcal{N}(\mu, \sigma^2)) \Rightarrow \left(\overline{X}_n \not\hookrightarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \right) \text{ par le TCL pour } n \text{ grand où } \mu \in \mathbb{R} \text{ et } \sigma^2 \in \mathbb{R}_+^*.$

Proposition 12.9. $\textcircled{S}\textcircled{C} : (\overline{X}_n)_{n \geq 1} \xrightarrow{\mathbf{P}} \mu.$

1.2.2. 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 12.8. Fréquence empirique

$\textcircled{S} : \text{Soit } (X_1, \dots, X_n) \text{ un } n\text{-échantillon de } X \hookrightarrow \mathcal{B}(p) \text{ d'espérance inconnue } p. \text{ On appelle fréquence empirique la variable aléatoire } F_n \text{ définie par}$

$$\forall n \in \mathbb{N}^*, F_n = \frac{1}{n} \sum_{k=1}^n X_k$$

avec pour tout $k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{B}(p)$ où p est inconnu.

Proposition 12.10. $\forall n \in \mathbb{N}^*, \forall p \in]0, 1[, \mathbf{E}_p(F_n) = p \text{ et } \mathbf{V}_p(F_n) = \frac{p(1-p)}{n}.$

Proposition 12.11. $F_n \hookrightarrow \mathcal{N}\left(p, \frac{p(1-p)}{n}\right) \text{ par le TCL pour } n \text{ grand, } np \geq 5 \text{ et } np(1-p) \geq 5.$

Proposition 12.12. $(F_n)_{n \geq 1} \xrightarrow{\mathbf{P}} p.$

1.2.3. Variance empirique

. Situation : Soit une variable X suivant une loi d'espérance connue $\mu \in \mathbb{R}$ mais de variance inconnue $\sigma^2 \in \mathbb{R}_+^*$.

Définition 12.9. Variance empirique

⑤ : 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

$$\forall n \in \mathbb{N}^*, \quad S_n^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \overline{X}_n)^2$$

Définition 12.10. Variance empirique rectifiée (HP mais ...)

⑤ : 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 $\widehat{S}_n^2$ définie par

$$\forall n \in \mathbb{N}^*, \quad \widehat{S}_n^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \overline{X}_n)^2$$

Proposition 12.13. ⑤⑥ : $\forall n \in \mathbb{N}^*, S_n^2 = \frac{1}{n} \sum_{k=1}^n X_k^2 - (\overline{X}_n)^2$ (à savoir retrouver).

Proposition 12.14. ⑤⑥ : $\forall n \in \mathbb{N}^*, \forall \sigma \in \mathbb{R}_+^*, \mathbf{E}_\sigma(S_n^2) = \left(\frac{n-1}{n}\right) \sigma^2$ (à savoir retrouver).

Proposition 12.15. ⑤⑥ : $\left(\left(\frac{n-1}{n}\right) S_n^2\right) \xrightarrow{\mathbf{P}} \sigma^2$ (à savoir retrouver).

Proposition 12.16. ⑤⑥ : $\forall n \in \mathbb{N}^*, \forall \sigma \in \mathbb{R}_+^*, \mathbf{E}_\sigma(\widehat{S}_n^2) = \sigma^2$ (à savoir retrouver) (HP mais ...).

Proposition 12.17. ⑤⑥ : $(\widehat{S}_n^2) \xrightarrow{\mathbf{P}} \sigma^2$ (à savoir retrouver) (HP mais ...).

2. Estimation par intervalle de confiance

Définition 12.11. Intervalle de confiance d'un paramètre

⑤⑥ : L'intervalle aléatoire $[U_n, V_n]$ est intervalle de confiance de $g(\theta)$ au niveau de confiance $1 - \alpha$ (ou de risque α) 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 12.3. Un intervalle de confiance peut se construire à l'aide de plusieurs outils : IM, IBT, méthode analytique (polynôme) ou à l'aide d'autres inégalités améliorant les deux premières citées. Toutes ces méthodes sont déjà tombées au cours des années.

Définition 12.12. Intervalle de confiance asymptotique d'un paramètre

⑤⑥ : L'intervalle aléatoire $[U_n, V_n]$ est intervalle de confiance asymptotique de $g(\theta)$ au niveau de confiance $1 - \alpha$ (ou de risque α) si il existe une suite $(\alpha_n)_{n \geq 1}$ telle que $(\alpha_n)_{n \geq 1}$ soit convergente vers α 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 $g(\theta)$.

Remarque 12.4. Un intervalle asymptotique s'obtient donc par une limite via le TCL ou bien par encadrement à partir de l'IM soit l'IBT voire d'autres inégalités présentées dans un sujet (comme Chernoff, Hoeffding et bien d'autres ...).

2.1. Estimation par intervalle de confiance d'une proportion.

2.1.1. *Par l'inégalité de Bienaymé-Tchebychev*

Proposition 12.18. *Nous avons*

$$IC_{\alpha}(p) = \left[F_n - \frac{1}{2\sqrt{n\alpha}} \quad ; \quad F_n + \frac{1}{2\sqrt{n\alpha}} \right]$$

2.1.2. *Par le théorème de la limite centrée*

Proposition 12.19. *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

$$u_{1-\alpha/2} = \Phi^{-1}(1 - \alpha/2)$$

et

$$IC_{\alpha}(p) \subset \left[F_n - \frac{u_{1-\alpha/2}}{2\sqrt{n}} \quad ; \quad F_n + \frac{u_{1-\alpha/2}}{2\sqrt{n}} \right]$$

à condition que

$$\min(nu_n, nv_n, n(1-u_n), n(1-v_n)) \geq 5$$

2.2. Estimation par intervalle de confiance asymptotique d'une moyenne de loi normale de variance donnée.

Proposition 12.20. *Nous avons*

$$IC_{\alpha}(\mu) = \left[\overline{X}_n - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \quad ; \quad \overline{X}_n + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

où

$$u_{1-\alpha/2} = \Phi^{-1}(1 - \alpha/2)$$

à condition que $n \geq 30$. Dans le cas où σ 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

$$IC_{\alpha}(\mu) = \left[\overline{X}_n - u_{1-\alpha/2} \frac{s}{\sqrt{n}} \quad ; \quad \overline{X}_n + u_{1-\alpha/2} \frac{s}{\sqrt{n}} \right]$$

Zoologie des lois usuelles du programme

Les lois sont classées par ordre alphabétique.

1. Loi de Bernoulli ⑤

- **Notation** : $X \hookrightarrow \mathcal{B}(p)$ avec $p \in]0, 1[$.
- **Paramètre** : p
- **Epreuve type** : épreuve amenant deux issues seulement : soit succès soit échec.
- $X(\Omega) = \{0, 1\}$
- $\mathbf{P}([X = 0]) = 1 - p$ et $\mathbf{P}([X = 1]) = p$
- $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } x \in [0, 1[\\ 1 & \text{si } x \geq 1 \end{cases}$
- $\mathbf{E}(X) = p$
- $\mathbf{V}(X) = p(1 - p)$
- $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (1 - X \hookrightarrow \mathcal{B}(1 - p))$
- $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X \hookrightarrow \mathcal{B}(1, p))$
- $\forall n \geq 1, (X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X^n \hookrightarrow \mathcal{B}(p))$
- $X_1, \dots, X_n$ i.i.d. ($\forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{B}(p)$) $\Rightarrow \left(\sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p) \right)$. Autrement dit

2. Variable indicatrice ⑤

Définition 13.1. $\forall A \in \mathcal{A}, \mathbf{1}_A = \begin{cases} 0 & \text{si } \overline{A} \\ 1 & \text{si } A \end{cases}$

Remarque 13.1. Par définition une variable indicatrice est donc une variable 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 $\mathbf{P}([\mathbf{1}_A = 0]) = \mathbf{P}(\overline{A})$
- $\forall x \in \mathbb{R}, F_{\mathbf{1}_A}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \mathbf{P}(A) & \text{si } x \in [0, 1[\\ 1 & \text{si } x \geq 1 \end{cases}$
- $\mathbf{E}(\mathbf{1}_A) = \mathbf{P}(A)$
- $\mathbf{V}(\mathbf{1}_A) = \mathbf{P}(A)(1 - \mathbf{P}(A))$
- $\mathbf{1}_{A \cap B} = \mathbf{1}_A \mathbf{1}_B$ (ce qui se généralise par récurrence)
- $\forall n \geq 1, (\mathbf{1}_A)^n = \mathbf{1}_A$
- $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \mathbf{1}_B$ (la généralisation est hors programme)
- $\mathbf{1}_{A \Delta B} = |\mathbf{1}_A - \mathbf{1}_B|$

3. Loi binomiale ⑤

- **Notation** : $X \hookrightarrow \mathcal{B}(n, p)$ avec $n \in \mathbb{N}^*$ et $p \in]0, 1[$
- **Paramètres** : n et p
- **Epreuve type** : succession de n épreuves de Bernoulli indépendantes et de même paramètre p .
- $X(\Omega) = \llbracket 0, n \rrbracket$

- $\forall k \in \llbracket 0, n \rrbracket, \mathbf{P}([X = k]) = \binom{n}{k} p^k (1-p)^{n-k}$
- $\mathbf{E}(X) = np$
- $\mathbf{V}(X) = np(1-p)$
- $(X \hookrightarrow \mathcal{B}(n, p)) \Leftrightarrow (n - X \hookrightarrow \mathcal{B}(n, 1-p))$
- $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X \hookrightarrow \mathcal{B}(1, p))$
- $X_1, \dots, X_n \text{ i.i.d. } (\forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{B}(p)) \Rightarrow \left(\sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p) \right)$
- **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**

- ▷ $(X \hookrightarrow \mathcal{B}(n, p) \text{ et } n \geq 30, p < 0, 1) \Rightarrow \left(X \underset{\sim}{\hookrightarrow} \mathcal{P}(np) \right)$
- ▷ $\left. \begin{array}{l} X \hookrightarrow \mathcal{B}(n, p) \\ n \geq 30, np \geq 5, np(1-p) \geq 5 \end{array} \right\} \Rightarrow \left(X \underset{\sim}{\hookrightarrow} \mathcal{N}(np, np(1-p)) \right)$

4. Loi de Dirac (ou loi certaine) ⑤

- **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 certaine est indépendante de toute autre, y compris d'elle-même.

5. Loi exponentielle ⑥

- **Notation** : $X \hookrightarrow \mathcal{E}(\lambda)$ où $\lambda \in \mathbb{R}_+^*$.
- **Paramètre** : λ
- **Epreuve type** : temps d'attente entre deux phénomènes indépendants tels que des arrivées à un guichet, ou des appels téléphoniques ...
- $X(\Omega) = \mathbb{R}_+ \text{ p.s.}$
- $\forall x \in \mathbb{R}, f_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x)$
- $\forall x \in \mathbb{R}, F_X(x) = (1 - e^{-\lambda x}) \mathbf{1}_{\mathbb{R}_+}(x)$
- $\mathbf{E}(X) = \frac{1}{\lambda}$
- $\mathbf{V}(X) = \frac{1}{\lambda^2}$
- $\forall n \in \mathbb{N}, \mathbf{E}(X^n) = \frac{n!}{\lambda^n}$ (**HP** mais bon à savoir ...)
- **Absence de mémoire**
 $(X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (\forall (x, y) \in \mathbb{R}_+^2, \mathbf{P}_{[X > x]}([X > x + y]) = \mathbf{P}([X > y])), \text{ autrement dit ce qui s'est passé sur l'intervalle de }]-\infty, x] \text{ n'affecte en rien ce qui se passera sur l'intervalle }]x, x + y].$
- $\forall a > 0, (X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow \left(aX \hookrightarrow \mathcal{E}\left(\frac{\lambda}{a}\right) \right)$
- $\forall \lambda > 0, (X \hookrightarrow \mathcal{E}(1)) \Leftrightarrow \left(\frac{X}{\lambda} \hookrightarrow \mathcal{E}(\lambda) \right)$

- ▶ $\forall \lambda > 0, (X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (\lambda X \hookrightarrow \mathcal{E}(1))$
- ▶ $\mathcal{E}(1) = \gamma(1)$
- ▶ $\underbrace{\mathcal{E}(1) * \dots * \mathcal{E}(1)}_{n \text{ fois}} = \gamma(n)$
- ▶ $\forall \lambda > 0, \underbrace{\mathcal{E}(\lambda) * \dots * \mathcal{E}(\lambda)}_{n \text{ fois}} = \Gamma\left(\frac{1}{\lambda}, n\right)$ (**loi d'Erlang HP mais ...**)
- ▶  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
- ▶ $X(\Omega) = \mathbb{R}_+^*$ p.s.
- ▶ $\forall x \in \mathbb{R}, f_X(x) = \frac{x^{v-1} e^{-x}}{\Gamma(v)} \mathbf{1}_{\mathbb{R}_+^*}(x)$
- ▶ $\mathbf{E}(X) = v$
- ▶ $\mathbf{V}(X) = v$
- ▶ $\gamma(1) = \mathcal{E}(1)$
- ▶ $\underbrace{\mathcal{E}(1) * \dots * \mathcal{E}(1)}_{n \text{ fois}} = \gamma(n)$
- ▶ $\gamma(v_1) * \dots * \gamma(v_n) = \gamma\left(\sum_{k=1}^n v_k\right)$
- ▶ $\forall b \in \mathbb{R}_+^*, (X \hookrightarrow \gamma(v)) \Leftrightarrow (bX \hookrightarrow \Gamma(b, v))$ (**HP mais ...**)
- ▶ $\forall b \in \mathbb{R}_+^*, (X \hookrightarrow \Gamma(b, v)) \Leftrightarrow \left(\frac{1}{b}X \hookrightarrow \gamma(v)\right)$ (**HP mais ...**)

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 .
- ▶ $X(\Omega) = \mathbb{N}^*$
- ▶ $\forall k \in \mathbb{N}^*, \mathbf{P}([X = k]) = q^{k-1}p$ avec $q = 1 - p$.
- ▶ $\forall k \in \mathbb{N}, \mathbf{P}([X > k]) = q^k$
- ▶ $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1 - q^k & \text{si } x \in [k, k+1[, k \in \mathbb{N}^* \end{cases}$
- ▶ $\mathbf{E}(X) = \frac{1}{p}$
- ▶ $\mathbf{V}(X) = \frac{q}{p^2}$
- ▶ **Absence de mémoire**
 $(X \hookrightarrow \mathcal{G}(p)) \Leftrightarrow \left(\forall (n, m) \in (\mathbb{N})^2, \mathbf{P}_{[X > n]}([X > n+m]) = \mathbf{P}([X > m]) \right)$ 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
- ▶ $X(\Omega) = \mathbb{R}$ p.s.
- ▶ $\forall x \in \mathbb{R}, f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
- ▶ **Intégrale de Gauss**

$$\triangleright \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1$$

$$\triangleright \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

► **Fonction de répartition**

$$\triangleright \forall x \in \mathbb{R}, \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

$$\triangleright \forall x \in \mathbb{R}, \Phi(-x) = 1 - \Phi(x)$$

$$\triangleright \Phi(0) = 1/2$$

► **Mode** : 0

► **Médiane** : 0

► **E**(X) = 0

► **V**(X) = 1

$$\triangleright \forall n \in \mathbb{N}, \mathbf{E}(X^n) = \begin{cases} 0 & \text{si } n \in 2\mathbb{N} + 1 \\ \frac{(2m)!}{2^m m!} & \text{si } n = 2m \in 2\mathbb{N} \end{cases} \quad (\text{HP mais ... à savoir retrouver})$$

$$\triangleright \left(\frac{X - m}{\sigma} \hookrightarrow \mathcal{N}(0, 1) \right) \Leftrightarrow (X \hookrightarrow \mathcal{N}(m, \sigma^2))$$

► **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**

$$\left((X_n)_{n \geq 1} \text{ i.i.d.} \right) \Rightarrow \left(\frac{\sum_{k=1}^n X_k - \mathbf{E}(\sum_{k=1}^n X_k)}{\sqrt{\mathbf{V}(\sum_{k=1}^n X_k)}} \right)_{n \geq 1} \xrightarrow{\mathcal{L}} N \quad \text{où } N \hookrightarrow \mathcal{N}(0, 1) \quad (TCL)$$

8.2. Loi normale quelconque ou de Laplace-Gauss.

► **Notation** : $X \hookrightarrow \mathcal{N}(m, \sigma^2)$ où $m \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$.

► **Paramètres** : m et σ^2

► $X(\Omega) = \mathbb{R}$ p.s.

$$\triangleright \forall x \in \mathbb{R}, f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$$

► **Mode** : m

► **Médiane** : m

► **E**(X) = m

► **V**(X) = σ^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**

$$\mathcal{N}(m_1, \sigma_1^2) * \cdots * \mathcal{N}(m_n, \sigma_n^2) = \mathcal{N}\left(\sum_{k=1}^n m_k, \sum_{k=1}^n \sigma_k^2\right)$$

9. Loi de Poisson ⑤

► **Notation** : $X \hookrightarrow \mathcal{P}(\lambda)$ où $\lambda \in \mathbb{R}_+^*$.

► **Paramètre** : λ

► **Epreuve type** : nombre d'apparitions d'un phénomène rare durant un intervalle de temps donné.

► $X(\Omega) = \mathbb{N}$

$$\triangleright \forall k \in \mathbb{N}, \mathbf{P}([X = k]) = \frac{e^{-\lambda} \lambda^k}{k!}$$

► **E**(X) = λ

► **V**(X) = λ

► **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 indépendantes pour lesquelles $S_n = \sum_{k=1}^n X_k$ suit une loi de Poisson, alors chacune des variables X_k suit une loi de Poisson (**HP mais ...**). On dit que la loi de Poisson est indéfiniment divisible, c'est la seule à valeurs dans $\mathbf{N}$.
- Si $X \hookrightarrow \mathcal{P}(\lambda)$ alors $\mathbf{P}([X=0]) = 1 - F_Y(1)$ où $Y \hookrightarrow \mathcal{E}(\lambda)$ (**HP mais ...** à savoir retrouver).
- Soit $X \hookrightarrow \mathcal{P}(\lambda x)$ et $Y \hookrightarrow \mathcal{E}(\lambda)$ alors $\mathbf{P}([X > x]) = \mathbf{P}([Y=0])$ (**HP mais ...** 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. Lois uniformes

10.1. Loi uniforme discrète ⑤.

- **Notation** : $X \hookrightarrow \mathcal{U}([1, n])$
- **Paramètre** : $[1, n]$
- **Epreuve type** : numéro d'une boule tirée à partir d'une urne constituée de boules numérotées de 1 à n .
- $X(\Omega) = [1, n]$
- $\forall k \in [1, n], \mathbf{P}([X=k]) = \frac{1}{n}$
- $\mathbf{E}(X) = \frac{n+1}{2}$
- $\mathbf{V}(X) = \frac{n^2-1}{12}$
- Soit $(a, b) \in \mathbb{R}^2$ avec $a < b$ alors

$$(X \hookrightarrow \mathcal{U}([a, b])) \Leftrightarrow (X - a + 1 \hookrightarrow \mathcal{U}([1, b - a + 1]))$$

$$\text{et } \mathbf{E}(X) = \frac{a+b}{2} \text{ et } \mathbf{V}(X) = \frac{(b-a+1)^2-1}{12}.$$



- Contrairement à une légende urbaine, la loi uniforme n'est pas stable par convolution !

10.2. Loi uniforme continue ⑥.

- **Notation** : $X \hookrightarrow \mathcal{U}(I)$ où I est un intervalle d'extrémités a et b deux réels tel que $a < b$.
- **Paramètre** : I
- $X(\Omega) = I$ p.s.
- $\forall x \in \mathbb{R}, f_X(x) = \left(\frac{1}{b-a}\right) \mathbf{1}_I(x)$
- $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}$
- $\mathbf{E}(X) = \frac{a+b}{2}$
- $\mathbf{V}(X) = \frac{(b-a)^2}{12}$
- Si $U \hookrightarrow \mathcal{U}([0, 1])$ alors $Y = -\frac{\ln(1-U)}{\lambda} \hookrightarrow \mathcal{E}(\lambda)$ (**HP mais** à connaître ...)
- $(X \hookrightarrow \mathcal{U}([0, 1])) \Leftrightarrow \left(\forall (a, b) \in (\mathbb{R}_+^*)^2, a + (b-a)X \hookrightarrow \mathcal{U}([a, b]) \right)$
- $(X \hookrightarrow \mathcal{U}([a, b])) \Leftrightarrow \left(\frac{X-a}{b-a} \hookrightarrow \mathcal{U}([0, 1]) \right)$

$$\blacktriangleright (X \hookrightarrow \mathcal{U}([0, 1]) \Leftrightarrow \left(-\frac{1}{\lambda} \ln(X) \hookrightarrow \mathcal{E}(\lambda)\right)$$



$\blacktriangleright$ Contrairement à une légende urbaine, la loi uniforme n'est pas stable par convolution !

11. Bilan des lois stables par convolution

- $\blacktriangleright \mathcal{B}(n_1, p) * \mathcal{B}(n_2, p) = \mathcal{B}(n_1 + n_2, p)$
- $\blacktriangleright \mathcal{P}(\lambda_1) * \mathcal{P}(\lambda_2) = \mathcal{P}(\lambda_1 + \lambda_2)$
- $\blacktriangleright \gamma(v_1) * \gamma(v_2) = \gamma(v_1 + v_2)$
- $\blacktriangleright \mathcal{N}(m_1, \sigma_1^2) * \mathcal{N}(m_2, \sigma_2^2) = \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

12. Bilan des lois sans mémoire

- $\blacktriangleright \textcircled{S}$: Les seules variables à valeurs dans $\mathbf{N}^*$ qui ont la propriété d'être sans mémoire, c'est-à-dire vérifiant :

$$\forall (n, m) \in (\mathbf{N})^2, \quad \mathbf{P}([X > n + m]) = \mathbf{P}([X > m]) \mathbf{P}([X > n])$$

sont les variables géométriques.

- $\blacktriangleright \textcircled{C}$: Les seules variables à valeurs dans $\mathbf{R}_+$ qui ont la propriété d'être sans mémoire, c'est-à-dire vérifiant :

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

sont les variables exponentielles.

Compléments hors programme mais ...

1. Formule de Vandermonde

Proposition 14.1. *Pour tout triplet $(n_1, n_2, n) \in \mathbb{N}^3$ tel que $n \leq n_1 + n_2$ alors*

$$\sum_{k=0}^n \binom{n_1}{k} \binom{n_2}{n-k} = \binom{n_1 + n_2}{n}$$

2. Relation min et max

Proposition 14.2. *Pour toute liste $(x_1, \dots, x_n) \in \mathbb{R}^n$,*

$$\max(x_1, \dots, x_n) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \min(x_{i_1}, \dots, x_{i_k})$$

3. 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 $n = c_1 + \dots + c_k$.

4. Espérance et fonction de survie

Proposition 14.4. ► ⑤ : *Soit X une variable à valeur dans $\mathbb{N}$ alors*

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

En cas de convergence,

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

► ⑥ : *Soit X une variable à valeur dans $\mathbb{R}_+$ 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

$$\lim_{x \rightarrow +\infty} \mathbf{P}(|X| > x) = 0$$

5. Interprétation géométrique de la corrélation

Définition 14.1. $\textcircled{S}\textcircled{C}$: Soit $\langle \bullet, \bullet \rangle$ le produit scalaire et $\|\bullet\|$ la norme de $\mathcal{L}^2$, 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

$$Y - \mathbf{E}(Y) = a(X - \mathbf{E}(X))$$

6. Minimisation de la variance

Proposition 14.6. $\textcircled{S}\textcircled{C}$: $\mathbf{V}(X) = \inf_{a \in \mathbf{R}} \mathbf{E}((X - a)^2)$ (à savoir retrouver)

7. Espérance conditionnelle

Proposition 14.7. $\textcircled{S}$: $\forall B \in \mathcal{A}, \mathbf{P}(B) \neq 0, \mathbf{E}(X | B) = \frac{\mathbf{E}(X \mathbf{1}_B)}{\mathbf{P}(B)}$ (à savoir retrouver)

Proposition 14.8. $\textcircled{S}$: $\forall y \in \mathcal{S}_Y, \mathbf{E}(X | [Y = y]) = \frac{\mathbf{E}(X \mathbf{1}_{[Y=y]})}{\mathbf{P}([Y = y])}$ (à savoir retrouver)

8. Loi exponentielle

Proposition 14.9. $\textcircled{C}$: 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 $\sup(X, Y)$)

► $\mathbf{P}([X \leq Y]) = \frac{\lambda}{\lambda + \mu}$

► $\textcircled{C} : \mathcal{E}(\lambda) = \Gamma\left(\frac{1}{\lambda}, 1\right)$

► $\textcircled{C} : \underbrace{\mathcal{E}(\lambda) * \dots * \mathcal{E}(\lambda)}_{n \text{ fois}} = \Gamma\left(\frac{1}{\lambda}, n\right)$ (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 $[X] + 1 \hookrightarrow \mathcal{G}([\lambda])$

Proposition 14.10. $\textcircled{C}$: $\gamma(v) = \Gamma(1, v)$

9. Variables symétriques

Définition 14.2. *Variable symétrique*

$\textcircled{S}\textcircled{C}$: On dit qu’une variable est symétrique lorsque $X \sim -X$

Proposition 14.11. *Distribution bilatérale*

$\textcircled{C}$: Soit X une v.a.r.a.d. de densité f_X alors

$$(X \sim -X) \Leftrightarrow (f_X \text{ est paire})$$

On parle de distribution bilatérale.

Proposition 14.12. $\textcircled{S}\textcircled{C}$: $\forall x \in \mathbf{R}, F_X(-x) = 1 - F_X(x)$

Proposition 14.13. $\textcircled{S}\textcircled{C}$: Soit X et Y deux variables symétriques indépendantes alors $X \pm Y$ reste une variable symétrique.

10. Variables aléatoires extrêmes

Définition 14.3. Soit n ($n \geq 2$) variables $X_1, \dots, X_n$ définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$. On appelle variables extrêmes, les variables aléatoires

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

définies par

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

Proposition 14.14. Soit n ($n \geq 2$) variables 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,

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

Remarque 14.1.  Puisque les variables extrêmes ne sont pas fonction de sommes de variables aléatoires, le TCL 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 **HP mais ...**) 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.

11. Loi géométrique

Proposition 14.15. $\blacktriangleright \textcircled{S}$: 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

$$\inf(X_1, \dots, X_n) \hookrightarrow \mathcal{G}\left(1 - \prod_{i=1}^n (1 - p_i)\right)$$

$\blacktriangleright \textcircled{S}$: $\underbrace{\mathcal{G}(p) * \dots * \mathcal{G}(p)}_{n \text{ fois}} = \mathcal{P}(n, p)$ (loi de Pascal de paramètres n et p généralisant la loi géométrique)

Remarque 14.2. Il n'y a rien de remarquable à dire concernant la loi de $\sup(X_1, \dots, X_n)$.

12. Loi normale et loi du chi-deux

Proposition 14.16. $\blacktriangleright \textcircled{C}$: $(X \hookrightarrow \mathcal{N}(0, 1)) \Rightarrow \left(X^2 \hookrightarrow \Gamma\left(2, \frac{1}{2}\right)\right)$ (loi du chi-deux à 1 degrés de liberté)

$\blacktriangleright \textcircled{C}$: Soit $X_1, \dots, X_n$ des variables 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 HP).

13. Produit scalaire et norme associée dans $\mathcal{L}^2$

Proposition 14.17. $\blacktriangleright \textcircled{S}\textcircled{C}$: Soit $(X, Y) \in (\mathcal{L}^2)^2$, $\langle X, Y \rangle = \mathbf{E}(XY)$

$\blacktriangleright \textcircled{S}\textcircled{C}$: Soit $X \in \mathcal{L}^2$, $\|X\|^2 = \mathbf{E}(X^2)$

Outils

- ▶ $\forall p \in \mathbb{N}^*, \forall k \in \llbracket 1, p \rrbracket, |x|^k \leq |x|^p + 1$
- ▶ $\forall (x, y) \in \mathbb{R}^2, |xy| \leq \frac{x^2 + y^2}{2}$
- ▶ $\forall x \in \mathbb{R}, |x| \leq \frac{x^2 + 1}{2}$
- ▶ $\forall x \in \mathbb{R}, \lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$
- ▶ $\forall (x, y) \in \mathbb{R}^2, \max(x, y) + \min(x, y) = x + y$
(ce résultat ne se généralise pas en $\max(x_1, \dots, x_n) + \min(x_1, \dots, x_n) = x_1 + \dots + x_n$)
- ▶ $\forall (x, y) \in \mathbb{R}^2, \max(x, y) - \min(x, y) = |x - y|$
- ▶ $\sum_{n=0}^{+\infty} t^n = \frac{1}{1-t}$ 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 $t \in [-1, 1[$ (**HP mais ...**)
- ▶ $\sum_{n=0}^{+\infty} \binom{n+k-1}{k-1} t^n = \frac{1}{(1-t)^k}$ où $|t| < 1$
- ▶ $\sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x$ où $x \in \mathbb{R}$
- ▶ $e^{tX} + e^{-tX} = 2 \sum_{n=0}^{+\infty} \frac{(tX)^{2n}}{(2n)!}$ (à savoir retrouver)
- ▶ $e^{tX} - e^{-tX} = 2 \sum_{n=0}^{+\infty} \frac{(tX)^{2n+1}}{(2n+1)!}$ (à savoir retrouver)
- ▶ $\Gamma : x \in \mathbb{R}_+^* \mapsto \int_0^{+\infty} t^{x-1} e^{-t} dt$
- ▶ $\binom{n}{k} \underset{n \rightarrow +\infty}{\sim} \frac{n^k}{k!}$ où k est entier fixé.
- ▶ $\lim_{n \rightarrow +\infty} e^{-n} \sum_{k=0}^n \frac{n^k}{k!} = \frac{1}{2}$ (**HP mais** à savoir retrouver)
- ▶ $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} dx = 1$ (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 $\mathcal{N}(0, 1)$)

- $\frac{1}{\sqrt{2\pi}} \int_0^{+\infty} x^2 e^{-x^2/2} dx = 1$ (*moment d'ordre deux d'une variable $\mathcal{N}(0,1)$*)
- $\forall n \in \mathbb{N}, \int_0^{+\infty} x^n e^{-x} dx = n!$ (*valeur de $\Gamma(n+1)$*)
- $\int_0^{+\infty} x e^{-x} dx = 1$ (*espérance d'une variable $\mathcal{E}(1)$*)

Index

- Approximation
 - binomiale par Normale, 47
 - binomiale par Poisson, 47
 - poisson par Normale, 47
- Calcul de probabilité
 - à partir de la fonction de répartition, 22
 - à partir de la fonction de survie, 23
- Coefficient
 - de corrélation linéaire, 35
- Coefficient
 - binomial, 11
 - multinomial, 12
- Convergence
 - en loi, 44
 - en probabilités, 43
 - spaciale, 43
- Convergences
 - lien, 45
- Dénombrement
 - exemples fondamentaux de tirages usuels, 16
 - lemme des bergers, 15
 - nombre d'applications de E vers F, 15
 - nombre d'arrangements de E vers F, 15
 - nombre d'injections de E vers F, 15
 - nombre de combinaisons de p éléments de E de cardinal n, 16
 - nombre de couples d'entiers, 16
 - nombre de p-listes de E, 16
 - nombre de parties d'un ensemble E, 16
 - nombre de permutations de E, 16
 - nombre de suites d'entiers, 16
 - problème des anagrammes, 16
- Densité de probabilité
 - caractérisation, 22
 - définition, 22
- Dispersion d'un vecteur aléatoire, 33
- Distribution bilatérale, 60
- Droites de régression, 38
- Ecart quadratique moyen, 38
- Ecart-type, 26, 38
- Echantillon, 49
- Egalité
 - en loi, 23
- Ensemble
 - définition intuitive d'un ensemble, 7
 - différence d'ensembles, 9
 - différence symétrique d'ensembles, 9
 - distributivités, 8
 - ensemble des parties d'un ensemble, 7
 - fonction indicatrice ou fonction caractéristique, 9
 - inclusion d'ensembles, 8
 - intersection d'ensembles, 7
 - lois de Morgan, 8
 - partition, 15
 - produit cartésien, 9
 - union, 7
- Equiprobabilité, 19
- Espérance
 - conditionnelle, 32
 - définition, 24
 - et fonction de survie, 24
 - et variable bornée, 25
 - linéarité, 31
- Espace
 - probabilisé, 18
 - probabilisable, 18
- Estimateur
 - biais, 49
 - consistant, 49
 - convergent, 49
 - d'un paramètre, 49
 - risque quadratique moyen, 49
- Estimation
 - définition, 49
 - par intervalle de confiance asymptotique d'une moyenne de loi normale de variance donnée, 52
 - ponctuelle, 49
- Événement
 - négligeable, 18
 - quasi-certain, 18
 - système complet, 18
- Fonction
 - de régression, 32
 - de répartition, 22
 - de survie, 23
- Fonctions de régression, 32
- Formule
 - de Bayes, 20
 - de l'espérance totale, 32
 - de Poincaré ou du crible pour deux ensembles dénombrement, 15
 - de Vandermonde, 11
 - des probabilités composées, 20
 - des probabilités totales, 20
 - du binôme de Newton, 11
 - du crible pour deux ensembles probabilité, 18
 - du triangle de Pascal, 11

- multinomiale, 12
- Fréquence empirique, 50
- Identité
 - de Bernoulli, 12
 - du parallélogramme, 34
- Identités
 - de polarisation, 34
- Inégalité
 - avec une variable indicatrice, 41
 - de Bienaymé-Tchebychev, 39
 - de Boole ou sous-additivité, 19
 - de Cauchy-Schwarz, 35, 40
 - de Chernoff, 39
 - de concentration, 39
 - de Jensen, 40
 - de Markov, 39
 - de moments, 40
 - triangulaire, 40
- Indépendance
 - deux événements, 19
 - mutuelle, 19
 - n événements, 19
 - suite d'événements, 19
 - variables aléatoires
 - caractérisation, 35
- Intervalle de confiance
 - asymptotique d'un paramètre, 51
 - d'un paramètre, 51
 - d'une proportion, 52
- Lemme des coalitions, 36
- Les trois grandes partitions de $\mathcal{I}N^2$
 - de deux familles sommables, 29
- Loi
 - binomiale, 53
 - 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, 54
 - conditionnelle, 30
 - d'Erlang, 55, 60
 - d'un couple, 30
 - de Bernoulli, 53
 - de Dirac, 54
 - de Fréchet, 61
 - de Gumbel, 61
 - de Poisson, 56
 - de probabilité, 22
 - de Weibull, 61
 - du chi-deux, 61
 - exponentielle, 54
 - faible des grands nombres, 43
 - géométrique, 55
 - gamma, 55
 - marginale, 30
 - normale, 55
 - uniforme, 57
- Matrice de covariance-variance d'un vecteur aléatoire, 33
- Moment
 - factoriel d'ordre r , 26
- Moyenne empirique, 50
- Norme dans L^2 , 38
- Produit
 - à trous, 13
 - de Cauchy, 37
 - de convolution, 37
 - de deux familles sommables, 29
 - lien entre produit et somme, 13
 - permutation des produits, 12
 - propriétés élémentaires, 12
 - scalaire dans L^2 , 38
- Relation
 - de Laplace, 19
 - min et max, 59
- Standardisation d'une variable aléatoire, 27
- Statistique, 49
 - inférentielle, 49
- Théorème
 - d'existence d'une espérance par domination, 31
 - de composition
 - convergence en loi, 44
 - convergence en probabilité, 43
 - de convolution, 37
 - de domination, 25
 - de Fisher-Tippett-Gnedenko, 61
 - de Koenig-Huygens, 26
 - de la limite centrée, 44
 - de la limite monotone, 18
 - de Slutsky, 44
 - de sommation par paquets, 29
 - de transfert en dimension un, 27
 - de transfert fonction d'un couple, 30
 - de transfert fonction d'un vecteur, 31
 - Koenig, 33
- Tirages
 - au hasard, 19
- Tribu, 17
 - borélienne de $\mathcal{I}R$, 17
 - engendrée par une famille, 17
 - grossière, 17
 - la plus fine, 17
- Variable aléatoire
 - étagée, 23
 - à densité, 22
 - bornée, 23
 - centrée, 24
 - centrée réduite, 27
 - indicatrice, 53
 - normalisation, 21, 22
 - sans mémoire, 23
 - support, 21
 - symétrique, 24, 60
- Variables aléatoires
 - de même loi, 23
 - extrêmes, 61
 - non corrélées ou décorrélées, 33
- Variance
 - additivité, 35
 - empirique, 51
 - rectifiée, 51
 - quadraticité, 37