

Le 29/04/2021 à 7H55

Tous les résultats du cours en un clin d'oeil

Cookbook

spicesagros.fr

05 décembre 2021

Table des matières

1	Légendes et abréviations	4
2	Ensembles	5
2.1	Définition intuitive d'un ensemble	5
2.2	Ensemble des parties d'un ensemble	5
2.3	Union d'ensembles	5
2.4	Intersection d'ensembles	5
2.5	Distributivités	5
2.6	Lois de Morgan où $\overline{A} = \complement_E A$	5
2.7	Inclusion d'ensembles	6
2.8	Différence d'ensembles	6
2.9	Différence symétrique d'ensembles (HP)	6
2.10	Produit cartésien d'ensembles	6
2.11	Fonction indicatrice (ou fonction caractéristique) d'un ensemble	6
3	Formules combinatoires	7
3.1	Coefficient binomial	7
3.2	Produit	7
4	Dénombrement	9

5	Espaces probabilisés	10
5.1	Tribu	10
5.2	Correspondance entre relations logiques et ensemblistes	10
5.3	Probabilité	10
5.3.1	Définition et propriétés	10
5.3.2	Indépendance d'événements	11
5.3.3	Probabilité conditionnelle	11
6	Variables aléatoires réelles	12
6.1	Généralités	12
6.2	Moments	13
6.3	Inégalités	14
6.4	Espérance conditionnelle	14
7	Vecteurs aléatoires discrets	15
7.1	Famille sommables	15
7.2	Couples discrets	15
7.3	Vecteurs discrets de dimension $n \geq 2$	16
7.4	Covariance	16
7.4.1	Droites de régression	17
7.4.2	Produit scalaire dans \mathcal{L}^2 (HP)	17
7.4.3	Variance d'une somme	17
7.4.4	Coefficient de corrélation linéaire	17
7.4.5	Matrice de covariance-variance	17
7.5	Indépendance	18
7.5.1	Définition et proposition	18
7.5.2	Le lemme des coalition	18
7.5.3	Convolution	18
8	Inégalités	19
8.1	Inégalités de concentration	19
8.2	Inégalités de moments	19
8.2.1	Inégalité de Cauchy-Schwarz (ICS)	19
8.2.2	Inégalité triangulaire	19
8.2.3	Inégalité de Jensen (HP)	19
9	Convergences	20
9.1	Convergence en probabilités	20
9.2	Convergence en loi	20
9.3	Approximations	21

10 Estimation inférentielle	22
10.1 Estimation ponctuelle	22
10.1.1 Généralités	22
10.1.2 Exemples d'estimateurs de paramètres usuels et sans biais et convergents	22
10.2 Estimation par intervalle de confiance	23
10.2.1 Estimation par intervalle de confiance d'une proportion	23
10.2.2 Estimation par intervalle de confiance asymptotique d'une moyenne de loi normale de variance donnée	23
11 Zoologie des lois usuelles du programme	24
11.1 Loi de Bernoulli \textcircled{S}	24
11.2 Variable indicatrice \textcircled{S}	24
11.3 Loi binomiale \textcircled{S}	24
11.4 Loi de Dirac (ou loi certaine) \textcircled{S}	24
11.5 Loi exponentielle \textcircled{C}	25
11.6 Loi gamma \textcircled{C}	25
11.7 Loi géométrique \textcircled{S}	25
11.8 Lois normales \textcircled{C}	25
11.8.1 Loi normale centrée et réduite ou de Laplace-Gauss centrée et réduite	25
11.8.2 Loi normale quelconque ou de Laplace-Gauss	26
11.9 Loi de Poisson \textcircled{S}	26
11.10 Lois uniformes	26
11.10.1 Loi uniforme discrète \textcircled{S}	26
11.10.2 Loi uniforme continue \textcircled{C}	26
11.11 Bilan des lois du programme stables par convolution	27
11.12 Bilan des lois du programme sans mémoire	27

1 Légendes et abréviations

- **D** Définition
- **T** Théorème
- **P** Proposition
- **Pr** Propriété (s)
- **C** Corollaire
- **R** Remarque
- **E** Exemple
- HP : résultat hors programme mais ...
- © : résultat valable en continu
- Ⓢ : résultat valable en discret
- \sqcup : union disjointe
- ssrcv : sous réserve de convergence
- 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 .

Pour l'index situé en fin de document, vous devrez ajouter 2 à la valeur de la page indiquée.

2 Ensembles

2.1 Définition intuitive d'un ensemble

D C'est le tout formant une collection d'objets *distincts* 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 $\exists x \in B, x \notin A$.

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

D 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 $\forall x \in B, x \in A$.

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

2.2 Ensemble des parties d'un ensemble

D 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.

R ▶ 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 $a \subset A!$).

2.3 Union d'ensembles

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

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

Pr ▶ $A \cup \bar{A} = A$

▶ $A \cup B = B \cup A$ (*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)

▶ $\Omega \cup A = \Omega$ (Ω est absorbant pour \cup)

▶ $A \cup A = A$

2.4 Intersection d'ensembles

D ▶ $A \cap B = \{x \in \Omega, x \in A \text{ et } x \in B\}$.

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

Pr ▶ $A \cap \bar{A} = \emptyset$

▶ $A \cap B = B \cap A$ (*commutativité*)

▶ $A \cap (B \cap C) = (A \cap B) \cap C$ (*associativité*)

▶ $A \cap B \subset A, A \cap B \subset B$

▶ $(A \subset B) \Leftrightarrow (A \cap B = A)$

▶ $\emptyset \cap A = \emptyset$ (\emptyset est absorbant pour \cap)

▶ $\Omega \cap A = A$ (Ω est neutre pour \cap)

▶ $A \cap A = A$

2.5 Distributivités

Pr ▶ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (*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*)

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

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

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

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

2.6 Lois de Morgan où $\bar{A} = \complement_E A$

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

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

▶ $\overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \bar{A}_i$

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

Pr ▶ $(A \cup B = E \text{ et } A \cap B = \emptyset) \Leftrightarrow (A \sqcup B = E)$

Pr ▶ $\overline{\bar{\Omega}} = \emptyset$

▶ $\overline{\bar{\emptyset}} = \Omega$

▶ $\overline{\bar{A}} = A$

- ▶ $(B = \overline{A}) \Leftrightarrow (A = \overline{B})$
- ▶ $(A \subset B) \Leftrightarrow (\overline{B} \subset \overline{A})$

2.7 Inclusion d'ensembles

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

Pr ▶ $(A \subset B) \Leftrightarrow (A \cap B = A)$

▶ $(A \subset B \text{ et } B \subset C) \Rightarrow (A \subset C)$

▶ $(A = B) \Leftrightarrow (A \subset B \text{ et } B \subset A)$

▶ $(A \subsetneq B) \Leftrightarrow (A \subset B \text{ et } A \neq B)$

▶ $(A \subset B) \Leftrightarrow (A \in \mathcal{P}(B))$

2.8 Différence d'ensembles

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

Pr ▶ $A - B = A \cap \overline{B}$

▶ $A - B = A - (A \cap B)$

▶ $\overline{A} = \Omega - A$

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

2.9 Différence symétrique d'ensembles (HP)

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

2.10 Produit cartésien d'ensembles

D ▶ $A \times B = \{(a, b) \mid a \in A \text{ et } b \in B\}$.

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

2.11 Fonction indicatrice (ou fonction caractéristique) d'un ensemble

D 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

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

Pr ▶ $\forall x \in E, \mathbf{1}_\Omega(x) = 1, \mathbf{1}_\emptyset(x) = 0$

▶ $\mathbf{1}_{A \cap B} = \mathbf{1}_A \times \mathbf{1}_B$

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

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

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

3 Formules combinatoires

P Soit n un entier naturel,

$$\blacktriangleright \sum_{k=1}^n k = \frac{n(n+1)}{2}$$

$$\blacktriangleright \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\blacktriangleright \sum_{k=1}^n k^3 = \left(\frac{n(n+1)}{2}\right)^2$$

3.1 Coefficient binomial

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

$$\mathbf{Pr} \blacktriangleright \forall (n, p) \in \mathbf{N}^2, p \leq n, \binom{n}{p} = \binom{n}{n-p} \text{ (symétrie)}$$

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

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

T Formule du binôme de Newton

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

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

$$\mathbf{Pr} \forall n \in \mathbf{N}^*, \sum_{k=1}^n k \binom{n}{k} = n2^{n-1} \times \dots \times a_p^{n_p}$$

$$\mathbf{Pr} \blacktriangleright \left(\sum_{k=1}^n x_k\right)^2 = \sum_{k=1}^n x_k^2 + 2 \sum_{1 \leq i < j \leq n} x_i x_j$$

Pr Identité de Bernoulli

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

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

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

3.2 Produit

Pr Soit $a \in \mathbb{R}$ et $(p, n, m) \in \mathbf{N}^3$, $p \leq n$ et $(x_k)_{1 \leq k \leq n}$ $(y_k)_{1 \leq k \leq n}$ deux familles de complexes

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

$$\blacktriangleright \prod_{k=p}^n a^k = a^{\sum_{k=p}^n k}$$

$$\blacktriangleright \prod_{k=p}^n a x_k = a^{n-p+1} \prod_{k=p}^n x_k \text{ où}$$

\blacktriangleright Pour toute permutation σ de $\llbracket p, n \rrbracket$ on a $\prod_{k=p}^n x_{\sigma(k)} = \prod_{k=p}^n x_k$

$$\blacktriangleright \prod_{k=1}^{n+p} x_k = \left(\prod_{k=1}^n x_k\right) \left(\prod_{k=n+1}^{n+p} x_k\right) \text{ (Associativité)}$$

$$\blacktriangleright \left(\prod_{k=p}^n x_k\right) \times \left(\prod_{k=p}^n y_k\right) = \prod_{k=p}^n (x_k y_k)$$

$$\blacktriangleright \left(\prod_{k=p}^n x_k\right)^m = \prod_{k=p}^n (x_k)^m$$

$$\blacktriangleright \frac{\prod_{k=1}^n x_k}{\prod_{k=1}^n y_k} = \prod_{k=1}^n \left(\frac{x_k}{y_k}\right) \text{ (où tous les complexes } y_k \neq 0)$$

$$\blacktriangleright \prod_{k=p}^n \frac{x_{k+1}}{x_k} = \frac{x_{n+1}}{x_p} \text{ (où tous les complexes } x_k \neq 0) \text{ (Télescopage)}$$

Pr Permutation des produits

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

$$\blacktriangleright \prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{i,j} = \prod_{i=1}^n \prod_{j=1}^n a_{i,j} = \prod_{j=1}^n \prod_{i=1}^n a_{i,j}$$

$$\blacktriangleright \prod_{1 \leq i \leq j \leq n} a_{i,j} = \prod_{j=1}^n \prod_{i=1}^j a_{i,j} = \prod_{i=1}^n \prod_{j=i}^n a_{i,j}$$

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

Pr Liens entre produit et somme

$$\blacktriangleright \forall n \geq 1, \forall k \in \{1, \dots, n\}, x_k > 0, \ln \left(\prod_{k=1}^n x_k \right) = \sum_{k=1}^n \ln(x_k)$$

$$\blacktriangleright \exp \left(\sum_{k=1}^n x_k \right) = \prod_{k=1}^n \exp(x_k)$$

$$\blacktriangleright \prod_i \sum_j a_{ij} \neq \sum_j \prod_i a_{ij}$$

$$\text{Pr} \prod_{k=1}^n k = n!$$

Pr Produits “à trous”

$$\blacktriangleright \prod_{k=1}^n (2k) = 2^n n!$$

$$\blacktriangleright \prod_{k=1}^n (2k+1) = \frac{(2n+1)!}{2^n n!}$$

4 Dénombrement

Pr ► Soit A et B deux ensembles tels que A est fini et B est fini et $B \subset A$, alors $\text{Card}(B) \leq \text{Card}(A)$ et $\text{Card}(A - B) = \text{Card}(A) - \text{Card}(B)$

► Si $A \subset B$ et $\text{Card}(A) = \text{Card}(B)$ alors $A = B$

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

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

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

T **Formule de Poincaré ou du crible pour deux ensembles**

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

R La généralisation de cette formule a disparu du programme.

D **Partition d'un ensemble**

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

▷ $\forall i \in I, A_i = \emptyset$,

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

▷ $\bigsqcup_{i \in I} A_i = E$.

Remarque : c'est la même déf. pour un SCE avec $E = \Omega$.

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

P **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 ,

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

P **Nombre d'injections de E_p vers F_n** avec $\text{Card}(E_p) = p$, $\text{Card}(F_n) = n$ et $p \leq n : A_n^p$

P **Nombre d'arrangements de E_p vers F_n** avec $\text{Card}(E_p) = p$, $\text{Card}(F_n) = n$ et $p \leq n : n(n-1) \times \cdots \times (n-p+1)$

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

P **Nombre de p -listes de E_n** : n^p

P **Nombre de combinaisons de p éléments de E de cardinal n** avec $p \leq n :$

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

P **Nombre de parties d'un ensemble E** : $\text{Card}(\mathcal{P}(E)) = 2^{\text{Card}(E)}$

P **Nombre de couples d'entiers :**

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

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

P **Nombre de suites d'entiers :**

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

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

E **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

5 Espaces probabilisés

5.1 Tribu

D 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 :

$$\triangleright \Omega \in \mathcal{A},$$

$$\triangleright \forall A \in \mathcal{A}, \bar{A} \in \mathcal{A},$$

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

$$\mathbf{P} \quad \emptyset \in \mathcal{A}, \mathcal{A} \text{ est stable pour } \cap, -, \Delta$$

$$\mathbf{E} \quad \triangleright \mathcal{A} = \{\emptyset, \Omega\} \text{ (tribu grossière, la plus triviale mais inexploitable)}$$

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

D 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 $\sigma(\mathcal{F})$.

$$\mathbf{E} \quad \mathbf{P} \quad \{\emptyset, A, \bar{A}, \Omega\} \text{ est une tribu notée } \sigma(A), \text{ la plus petite engendrée par } A \in \mathcal{A}.$$

$\mathbf{E} \quad \mathbf{P} \quad \triangleright \left\{ \bigsqcup_{\ell \in L} A_\ell \mid L \in \mathcal{P}(K) \right\}$ où $(A_k)_{k \in K}$ est un SCE est la tribu engendrée par le SCE et notée $\sigma((A_k)_{k \in K})$.

D Tribu borélienne de \mathbb{R}

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

P Quelle tribu utiliser le jour du concours ?

$$\triangleright \Omega \text{ est fini : } \mathcal{A} = \mathcal{P}(\Omega)$$

$$\triangleright \Omega \text{ est infini et dénombrable : } \mathcal{A} = \mathcal{P}(\Omega)$$

$\triangleright \Omega$ est infini et indénombrable : dans le cadre du programme, on se cherche pas à déterminer \mathcal{A} .

5.2 Correspondance entre relations logiques et ensemblistes

Relation logique	Relation ensembliste
négation non	complémentation
conjonction et	intersection \cap
disjonction ou	réunion \cup
implication \Rightarrow	inclusion \subset
équivalence \Leftrightarrow	égalité $=$

5.3 Probabilité

5.3.1 Définition et propriétés

D Axiomatique d'une probabilité

Soit (Ω, \mathcal{A}) un espace probabilisable, on appelle probabilité sur (Ω, \mathcal{A}) toute application $\mathbf{P} : \mathcal{A} \rightarrow [0, 1]$ vérifiant : $\mathbf{P}(\Omega) = 1, \forall (A_k)_{k \in K}$ famille d'événements deux à deux disjoints $\sum_k \mathbf{P}(A_k) < \infty$ et $\mathbf{P}\left(\bigsqcup_{k \in K} A_k\right) = \sum_{k \in K} \mathbf{P}(A_k)$ (σ -additivité de \mathbf{P}).

D Espace probabilisable

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

D Espace probabilisé

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

D Événement négligeable

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

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

D Événement quasi-certain

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

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

D Système complet d'événements

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

Pr Soit A et B deux évs de $(\Omega, \mathcal{A}, \mathbf{P})$,

$$\triangleright \mathbf{P}(\emptyset) = 0$$

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

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

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

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

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

- ▶ $(B \subset A) \Rightarrow (\mathbf{P}(\overline{A}) \leq \mathbf{P}(\overline{B}))$
- ▶ $\mathbf{P}(A \cap B) \leq \min\{\mathbf{P}(A), \mathbf{P}(B)\}$
- ▶ $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$ (formule du crible pour deux ensembles

dont la généralisation a disparu des programme)

- ▶ $\mathbf{P}(A \cup B) = \mathbf{P}(\overline{A} \cap B) + \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \overline{B})$
- ▶ $\mathbf{P}\left(\bigsqcup_{k=1}^n A_k\right) = \sum_{k=1}^n \mathbf{P}(A_k)$ (σ -additivité de \mathbf{P})

T Limite monotone

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

C Soit $(A_n)_{n \geq 0}$ une suite quelconque d'événements :

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

P Soit A et B deux évts de $(\Omega, \mathcal{A}, \mathbf{P})$, $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$.

T Relation de Laplace (cas d'équiprobabilité)

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

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

5.3.2 Indépendance d'événements

D ▶ Deux événements


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

▶ n événements deux à deux indépendants

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

▶ n événements mutuellement indépendants ou indépendants

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

R  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 !

▶ Suites d'événements

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

R ▶ 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.

P Les assertions suivantes sont équivalentes pour deux évts A et B :

$$\triangleright A \perp B$$

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

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

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

P 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.

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

5.3.3 Probabilité conditionnelle

D $\forall A \in \mathcal{A}$ tq $\mathbf{P}(A) \neq 0$, $\mathbf{P}_A(B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}$

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

P 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 :

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

$$\triangleright \mathbf{P}_A(B) = \mathbf{P}(B)$$

$$\triangleright \mathbf{P}_B(A) = \mathbf{P}(A)$$

$$\triangleright \mathbf{P}_B(A) \geq \mathbf{P}(A \cap B)$$

$$\triangleright \mathbf{P}_A(B) \geq \mathbf{P}(A \cap B)$$

P 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 :

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

P Formule des probabilités totales

Soit $(A_k)_{k \in K}$ un SCE tel que $\forall k \in K$, $\mathbf{P}(A_k) \neq 0$, alors

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

P Formule de Bayes

Soit $(A_k)_{k \in K}$ un SCE tel que $\forall k \in K$, $\mathbf{P}(A_k) \neq 0$, alors

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

6 Variables aléatoires réelles

6.1 Généralités


D \textcircled{C} : $X : \Omega \rightarrow \mathbb{R}$ tq $\forall B \in \mathcal{B}(\mathbb{R})$ on a $X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\} \in \mathcal{A}$.

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

D \textcircled{S} : $X : \Omega \rightarrow \mathbb{R}$ est une v.a.r.d. si $X(\Omega)$ est *au plus dénombrable* et $\sum_{x \in X(\Omega)} \mathbf{P}([X = x]) = 1$.

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

D \textcircled{C} : $X = Y$ lorsque $\forall \omega \in \Omega, X(\omega) = Y(\omega)$.

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

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

D **Variable aléatoire bornée**

\textcircled{C} : On dit qu'une variable X est bornée s'il existe une constante $M \geq 0$ telle que $\forall \omega \in \Omega, |X(\omega)| \leq M$.

D **Loi de probabilité**

\textcircled{S} : $P_X : X(\Omega) \rightarrow [0, 1]$, définie par $\forall x \in X(\Omega), P_X(x) = \mathbf{P}([X = x])$.

D **Fonction de répartition F**

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

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

P \textcircled{C} : $F' = f$ là où F est dérivable i.e. là où f est continue.

Pr ► \textcircled{C} : Dans le cas général $\lim_{-\infty} F = 0, \lim_{+\infty} F = 1, F$ est \mathcal{C}^0 à droite en tout point de \mathbb{R}, F est croissante au sens large.

► \textcircled{C} : On rajoute F est \mathcal{C}^0 sur \mathbb{R} et $\mathcal{C}^1 - I$ (I ensemble fini éventuellement vide) i.e. F est \mathcal{C}^1 presque partout.

P **Caractérisation d'une loi**


► \textcircled{S} : On donne $X(\Omega)$ ainsi que toutes les probabilités ponctuelles ou bien la fonction de répartition F de X .

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

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

P \textcircled{S} : Soit X une variable aléatoire telle que $X(\Omega) \subset \mathbb{Z}$ alors on a $\forall k \in \mathbb{Z}, \mathbf{P}([X = k]) = F(k) - F(k-1)$ où F est la fonction de répartition de X .

P \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).

P \textcircled{C} : $(X = Y) \Rightarrow (X \sim Y)$  la réciproque est fausse!

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

P **Caractérisation d'une densité**

\textcircled{C} : (f densité) \Leftrightarrow

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

Pr **Calcul de probabilité à partir de la fonction de répartition**

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

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

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

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

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

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

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

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

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

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

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

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

Pr **Calcul de probabilité à partir de la fonction de survie**

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

6.2 Moments

\mathbf{E} Espérance

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

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

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

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

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

\mathbf{P} Espérance d'une variable bornée

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

\mathbf{P} \textcircled{S} : $\forall a \in \mathbf{R}$, soit X une variable certaine égale à a alors $\mathbf{E}(X) = a$, i.e. $\mathbf{E}(a\mathbf{1}_{\Omega}) = a$.

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

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

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

\mathbf{T} Théorème de domination

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

\mathbf{D} Moment d'ordre r où $r \geq 0$

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

\mathbf{P} $\textcircled{S}\textcircled{C}$: $(m_2(X) = 0) \Leftrightarrow (X = 0 \text{ p.s.})$

\mathbf{R} \textcircled{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!

\mathbf{P} $\textcircled{S}\textcircled{C}$: Toute v.a.r. bornée admet des moments de tous ordres.

\mathbf{D} Moment absolu d'ordre r où $r \geq 0$

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

\mathbf{R} \textcircled{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!

\mathbf{P} $\textcircled{S}\textcircled{C}$: $(m_r(X) < \infty) \Leftrightarrow (m_r(|X|) < \infty)$

\mathbf{D} Moment centré d'ordre r où $r \geq 0$

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

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

\mathbf{D} $\textcircled{S}\textcircled{C}$: $\mathbf{V}(X) = \mu_2(X)$

\mathbf{D} $\textcircled{S}\textcircled{C}$: $\sigma(X) = \sqrt{\mathbf{V}(X)}$

\mathbf{P} $\textcircled{S}\textcircled{C}$: $(\mathbf{V}(X) < \infty) \Leftrightarrow (\mathbf{E}(X^2) < \infty)$

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

\mathbf{P} $\textcircled{S}\textcircled{C}$: $(\mathbf{E}(X^2) = 0) \Rightarrow (\mathbf{V}(X) = 0)$

\mathbf{P} $\textcircled{S}\textcircled{C}$: $(X \text{ bornée}) \Rightarrow (\mathbf{V}(X) < \infty)$

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

\mathbf{T} Théorème de Koenig-Huygens

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

\mathbf{T} Théorème de transfert

$$\mathbf{E}(\varphi(X)) = \begin{cases} \textcircled{S} : \sum_{x \in X(\Omega)} \varphi(x) \mathbf{P}([X = x]) \text{ ssr } |\text{cv}| \text{ où } \varphi \text{ déf sur } I \supset X(\Omega) \\ \textcircled{C} : \int_a^b \varphi(x) f(x) dx \text{ ssr } |\text{cv}| \text{ où } X(\Omega) =]a, b[\text{ p.s., } a \text{ et } b \text{ dans } \overline{\mathbf{R}} \\ \text{et } \varphi :]a, b[\rightarrow \mathbf{R} \text{ est } \mathcal{C}^0 \text{ presque partout dans }]a, b[\end{cases}$$

\mathbf{P} $\textcircled{S}\textcircled{C}$: $\forall (a, b) \in \mathbf{R}^* \times \mathbf{R}$, $(\mathbf{E}(X) < \infty) \Leftrightarrow (\mathbf{E}(aX + b) < \infty)$ et $\mathbf{E}(aX + b) = a\mathbf{E}(X) + b$.

\mathbf{P} $\textcircled{S}\textcircled{C}$: $\forall (a, b) \in \mathbf{R}^* \times \mathbf{R}$, $(\mathbf{V}(X) < \infty) \Leftrightarrow (\mathbf{V}(aX + b) < \infty)$ et $\mathbf{V}(aX + b) = a^2\mathbf{V}(X)$.

\mathbf{D} \mathbf{P} Variable centrée réduite

③③ : Soit X tel que $\mathbf{E}(X) < \infty$ et $\mathbf{V}(X) < \infty$, $\mathbf{V}(X) \neq 0$ alors $X^* = \frac{X - \mathbf{E}(X)}{\sigma(X)}$ est appelée variable centrée réduite associée à X .

6.3 Inégalités

Pr ③③ Inégalités de moments

- ▶ $\forall X \in \mathcal{L}^1, (X \geq 0) \Rightarrow (\mathbf{E}(X) \geq 0)$ (positivité de l'espérance)
- ▶ $\forall (X, Y) \in (\mathcal{L}^1)^2, (X \leq Y) \Rightarrow (\mathbf{E}(X) \leq \mathbf{E}(Y))$ (croissance de l'espérance)
- ▶ $\forall (X, Y) \in (\mathcal{L}^1)^2, (X = Y) \Rightarrow (\mathbf{E}(X) = \mathbf{E}(Y))$
- ▶ $(\forall Y \in \mathcal{L}^1 \text{ et } X \text{ tel que } |X| \leq Y) \Rightarrow (\mathbf{E}(X) < \infty \text{ et } \mathbf{E}(|X|) \leq \mathbf{E}(Y))$
- ▶ $\forall X \in \mathcal{L}^1, |\mathbf{E}(X)| \leq \mathbf{E}(|X|)$ (inégalité triangulaire)
- ▶ $\forall X \in \mathcal{L}^2, \mathbf{E}(X^2) \geq (\mathbf{E}(X))^2$

6.4 Espérance conditionnelle

D ③ Espérance conditionnelle

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

Pr ③ : $\forall A \in \mathcal{A}, \mathbf{P}(A) \neq 0, (\mathbf{E}(X) < \infty) \Rightarrow (\mathbf{E}(X | A) < \infty)$.

Pr ③ 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.

- ▶ $\forall x \in \mathcal{S}_X, \forall c \in \mathbb{R}, \mathbf{E}(c \mathbf{1}_\Omega | [X = x]) = c$.
- ▶ $\forall x \in \mathcal{S}_X, \mathbf{E}(aZ + bT | [X = x]) = a\mathbf{E}(Z | [X = x]) + b\mathbf{E}(T | [X = x])$ (linéarité).
- ▶ Si $X \perp\!\!\!\perp Y$ alors $\forall x \in \mathcal{S}_X, \mathbf{E}(Y | [X = x]) = \mathbf{E}(Y)$ et $\forall y \in \mathcal{S}_Y, \mathbf{E}(X | [Y = y]) = \mathbf{E}(X)$.

R. Ces propriétés restent valables en continu mais sont hors programme.

P ③ Formule de l'espérance totale

Soit $(A_i)_{i \in I}$ un SCE tq $\forall i \in I, \mathbf{P}(A_i) \neq 0$ alors $\mathbf{E}(X) = \sum_{i \in I} \mathbf{E}(X | A_i) \mathbf{P}(A_i)$ ssi

$|\text{cv}|$.

D ③ 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 $\mathbf{E}(X) < \infty$.

7 Vecteurs aléatoires discrets

7.1 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.

D 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**.

T Le théorème de sommation par paquets

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

T Cas particulier fondamental : $I = \mathbf{N}^2$

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

P Les trois grandes partitions de \mathbf{N}^2

$$\triangleright \mathbf{N}^2 = \bigsqcup_{n \in \mathbf{N}} (\mathbf{N} \times \{n\})$$

$$\triangleright \mathbf{N}^2 = \bigsqcup_{n \in \mathbf{N}} (\{n\} \times \mathbf{N})$$

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

P Produit de deux familles sommables

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

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

7.2 Couples discrets

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

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

D Lois marginales

► **Loi de X** : c'est l'application $P_X : X(\Omega) \rightarrow [0, 1]$ définie par $\forall x_i \in X(\Omega), P_X(x_i) = \mathbf{P}([X = x_i]) = p_{i,\bullet}$.

► **Loi de Y** : c'est l'application $P_Y : Y(\Omega) \rightarrow [0, 1]$ définie par $\forall y_j \in Y(\Omega), P_Y(y_j) = \mathbf{P}([Y = y_j]) = p_{\bullet,j}$.

P Lois marginales

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

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

D Lois conditionnelles

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

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

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

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

P 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 de $Z(\Omega)$:

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

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

T \textcircled{S} Théorème de transfert

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

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

en jeu.

7.3 Vecteurs discrets de dimension $n \geq 2$


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

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

D **Lois marginales de dimension un**

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

R  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!)

P $\forall k \in [1, n], \forall x_k \in X_k(\Omega)$,

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

T **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

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

P **Caractérisation de la loi d'une fonction d'un vecteur**

Soit $V = (X_1, \dots, X_n)$ 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

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

$Z(\Omega)$.

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

7.4 Covariance


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


R \textcircled{SC} : Cette définition ne sert quasiment jamais au calcul d'une covariance car elle n'est pas très pratique à manipuler.

R \textcircled{SC} : La covariance entre deux variables aléatoires X et Y est une mesure de liaison entre elles.

D **Variables aléatoires non corrélées ou décorrélées**

\textcircled{SC} : 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!

R \textcircled{SC} :  Faites toujours la distinction entre indépendant et non-corrélé. Ne concluez jamais à l'indépendance après avoir obtenu un coefficient de corrélation nul!

T **Théorème de Koenig**

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

Pr \blacktriangleright \textcircled{SC} : $\text{Cov}(X, X) = \mathbf{V}(X) \geq 0$ (positivité)

\blacktriangleright \textcircled{SC} : $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ (symétrie)

\blacktriangleright \textcircled{SC} : 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)

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


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

\blacktriangleright \textcircled{SC} : Soit $(X, Y) \in (\mathcal{L}^2)^2$ et $\forall (a, b, c, d) \in \mathbb{R}^4$,

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

\blacktriangleright \textcircled{SC} : $\forall k \in [1, n], \forall \ell \in [1, m], X_k$ et $Y_\ell \in \mathcal{L}^2, \lambda_k$ et $\mu_\ell \in \mathbb{R}$,

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

R \textcircled{SC}  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.

7.4.1 Droites de régression

D P Droites de régression

► **Ⓢ** : Droite de régression de Y en X .

Elle est notée Δ et a pour équation :

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

► **Ⓢ** : Droite de régression de X en Y .

Elle est notée Δ' et a pour équation :

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

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

7.4.2 Produit scalaire dans \mathcal{L}^2 (HP)

P Produit scalaire et norme associée dans \mathcal{L}^2 (HP mais déjà tombé!)

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

7.4.3 Variance d'une somme

Pr **ⓈⓈ** : $\forall (X, Y) \in (\mathcal{L}^2)^2$ et $\forall (a, b) \in \mathbb{R}^2$,
 $\mathbf{V}(aX + bY) = a^2\mathbf{V}(X) + b^2\mathbf{V}(Y) + 2ab\text{Cov}(X, Y)$.

P Identité du parallélogramme

ⓈⓈ : Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors $\mathbf{V}(X) + \mathbf{V}(Y) = \frac{1}{2}(\mathbf{V}(X + Y) + \mathbf{V}(X - Y))$

P Identités de polarisation

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

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

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

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

P **ⓈⓈ** : 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).

 La réciproque est fautive.

7.4.4 Coefficient de corrélation linéaire

D Coefficient de corrélation linéaire

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

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

R **ⓈⓈ** : Le coefficient de corrélation mesure la qualité d'une relation linéaire entre deux variables X et Y .

Pr **ⓈⓈ** : Le coefficient de corrélation est *invariant* par changement d'échelle $\forall (a, b, c) \in \mathbb{R}^* \times \mathbb{R}^2, \rho(aX + b, aY + c) = \rho(X, Y)$.

Si de plus $d \neq 0, |\rho(aX + b, dY + c)| = |\rho(X, Y)|$.

Pr L'inégalité de Cauchy-Schwarz (ICS)

► **ⓈⓈ** : Soit $(X, Y) \in (\mathcal{L}^2)^2$ tq $\mathbf{V}(X) \neq 0$ et $\mathbf{V}(Y) \neq 0$ alors $|\rho(X, Y)| \leq 1$ soit $|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$.

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

R **ⓈⓈ** : Plus $|\rho(X, Y)|$ est proche de 1, plus l'ajustement linéaire entre X et Y est pertinent.

7.4.5 Matrice de covariance-variance

D Matrice de covariance-variance d'un vecteur aléatoire

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

R **Ⓢ** La matrice de covariance-variance associée à un vecteur nous donne une idée de la *dispersion* du vecteur.

P ► **Ⓢ** 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.

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

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*.

7.5 Indépendance

7.5.1 Définition et proposition

D \textcircled{C} : $X_1 \perp \dots \perp X_n$ (on dit que les variables sont indépendantes ou mutuellement indépendantes) lorsque pour tous B_1, \dots, B_n , boréliens de \mathbf{R} on a

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

P Deux caractérisations de l'indépendance de v.a.r.

\textcircled{C} : $(X_1 \perp \dots \perp X_n) \Leftrightarrow$

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

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

P Une caractérisation de l'indépendance de v.a.r. discrètes

\textcircled{C} : $(X_1 \perp \dots \perp X_n) \Leftrightarrow$

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

P \textcircled{C} Soit X une variable certaine, alors $\forall Y$ v.a.r. on a $X \perp Y$. Ceci reste valable si $X = a$ p.s. ($a \in \mathbf{R}$). Autrement dit toute variable aléatoire certaine ou quasi-certaine est indépendante (au sens du calcul des probabilités) de toute autre variable.

P \textcircled{C} : $\forall X$ v.a.r. $\forall a \in \mathbf{R}, X \perp a \mathbf{1}_\Omega$.

7.5.2 Le lemme des coalition

P \textcircled{C} : $\forall (X, Y)$ deux variables et $\forall (\varphi, \psi) \in \mathcal{F}(I, \mathbf{R}) \times \mathcal{F}(J, \mathbf{R}) \mid I \supset X(\Omega)$ et $J \supset Y(\Omega)$ alors $\varphi(X), \psi(Y)$ restent des variables, et $(X \perp Y) \Rightarrow (\varphi(X) \perp \psi(Y))$.

! La réciproque est fautive.

P \textcircled{C} : $\forall (X_1, \dots, X_n)$ n variables et $\forall (\varphi, \psi) \in \mathcal{F}(D, \mathbf{R}) \times \mathcal{F}(D', \mathbf{R})$ tel que $D \supset X_1(\Omega) \times \dots \times X_k(\Omega)$ et $D' \supset X_{k+1}(\Omega) \times \dots \times X_n(\Omega)$ tq $\forall k \in \llbracket 1, n-1 \rrbracket, \varphi(X_1, \dots, X_k)$ et $\psi(X_{k+1}, \dots, X_n)$ restent des variables alors $\forall k \in \llbracket 1, n-1 \rrbracket,$

$$(X_1 \perp \dots \perp X_n) \Rightarrow (\varphi(X_1, \dots, X_k) \perp \psi(X_{k+1}, \dots, X_n)).$$

! On peut "bordéliser" les v.a.r. comme on veut!

P \textcircled{C} : $\forall (X_1, \dots, X_n) \in (\mathcal{L}^1)^n,$

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

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

! La réciproque est fautive.

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

P \textcircled{C} : $\forall (X, Y) \in (\mathcal{L}^2)^2, (X \perp Y) \Rightarrow (\rho(X, Y) = 0)$

! La réciproque est fautive en général (sauf si les deux variables sont au plus binaires ou gaussiennes (HP)).

P \textcircled{C} : $\forall (X, Y) \in (\mathcal{L}^2)^2, (\rho(X, Y) \neq 0) \Rightarrow \text{non}(X \perp Y)$ (par contraposée)

P \textcircled{C} : $(\forall (X, Y) \in (\mathcal{L}^1)^n \text{ et } X \perp Y) \Rightarrow \left(\begin{array}{l} \mathbf{E}(X \mid [Y = y_j]) = \mathbf{E}(X) \\ \mathbf{E}(Y \mid [X = x_i]) = \mathbf{E}(Y) \end{array} \right)$

où $\forall y_j \in Y(\Omega), \mathbf{P}([Y = y_j]) \neq 0$ et $\forall x_i \in X(\Omega), \mathbf{P}([X = x_i]) \neq 0$.

P \textcircled{C} : $\forall (X_1, \dots, X_n) \in \mathbf{R}^n, \forall (X_1, \dots, X_n) \in (\mathcal{L}^2)^n,$

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

7.5.3 Convolution

T \textcircled{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.

\textcircled{S} Soit X et Y avec $X \perp Y$ alors $\forall z \in (X + Y)(\Omega),$

$$\mathbf{P}([X + Y = z]) = \sum_{x \mid \begin{array}{l} x \in X(\Omega) \\ z-x \in Y(\Omega) \end{array}} \mathbf{P}([X = x]) \mathbf{P}([Y = z - x]).$$

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

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

que f_{X+Y} soit définie et \mathcal{C}^0 presque partout, ce qui est le cas si au moins l'une des deux densités est bornée et on note $f_{X+Y} = f_X * f_Y$.

Pr \textcircled{C} Le produit de convolution est commutatif et associatif.

Autrement dit $f_{X+Y} = f_X * f_Y = f_Y * f_X$ et $(f_X * f_Y) * f_Z = f_X * (f_Y * f_Z)$.

8 Inégalités

8.1 Inégalités de concentration

D \textcircled{C} : On appelle inégalité de concentration, toute inégalité donnant une majoration du type (f étant une fonction explicite) $\mathbf{P}(|X - \mathbf{E}(X)| \geq x) \leq f(x)$.

T **Inégalité de Markov (IM)**

► \textcircled{C} : Soit $X \in \mathcal{L}^1$ à valeurs positives, alors $\forall \varepsilon > 0$, $\mathbf{P}([X \geq \varepsilon]) \leq \frac{\mathbf{E}(X)}{\varepsilon}$ (IM à l'ordre 1)

► \textcircled{C} : Ce qui se généralise à l'ordre r pour $r \in \mathbf{N}^*$, $\forall \varepsilon > 0$, $\mathbf{P}([X \geq \varepsilon]) \leq \frac{\mathbf{E}(X^r)}{\varepsilon^r}$.

R \textcircled{C} : Dans le cas d'une variable X de signe quelconque on a

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

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

T **Inégalité de Bienaymé-Tchebychev (IBT)**

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

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

R 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.

8.2 Inégalités de moments

8.2.1 Inégalité de Cauchy-Schwarz (ICS)

T ► \textcircled{C} : 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{C} : 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 $Y = aX$.

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

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

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

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

8.2.2 Inégalité triangulaire

T \textcircled{C} : Soit X et Y telles que $\mathbf{E}(X) < \infty$ et $\mathbf{E}(Y) < \infty$ alors $\mathbf{E}(|X + Y|) < \infty$ et

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

R \textcircled{C} : Cette inégalité se généralise au cas de n variables aléatoires.

8.2.3 Inégalité de Jensen (HP)

T \textcircled{C} : Soit X une variable aléatoire à valeurs dans un intervalle I et φ une fonction convexe de I dans \mathbf{R} alors si X et $\varphi(X)$ admettent une espérance on a,

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

R \textcircled{C} : En général $\varphi(\mathbf{E}(X)) \neq \mathbf{E}(\varphi(X))$.

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

P ► \textcircled{C} : Soit $(X, Y) \in (\mathcal{L}^2)^2$ alors $\mathbf{E}(|XY|) < \infty$ et $\mathbf{E}(|XY|) \leq \frac{\mathbf{E}(X^2) + \mathbf{E}(Y^2)}{2}$.

► \textcircled{C} : Soit $X \in \mathcal{L}^2$ alors $\mathbf{E}(|X|) \leq \frac{1 + \mathbf{E}(X^2)}{2}$.

9 Convergences

9.1 Convergence en probabilités

D $\textcircled{S}\textcircled{C}$: $(X_n)_{n \geq 1} \xrightarrow{\mathbf{P}} X$ lorsque $\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbf{P}(|X_n - X| \geq \varepsilon) = 0$ soit encore $\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbf{P}(|X_n - X| < \varepsilon) = 1$.

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

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

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

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

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

T **Théorème de composition**

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

R Il est suffisant d'avoir $f \in \mathcal{C}^0(X(\Omega), \mathbb{R})$ mais ce n'est pas dit dans le programme.

T **Loi faible des grands nombres**

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

et de variance commune σ^2 . Soit $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ alors $(\bar{X}_n)_{n \geq 1} \xrightarrow{\mathbf{P}} m$.

$\blacktriangleright \textcircled{S}$: Si $\forall k \in [1, n], X_k \hookrightarrow \mathcal{B}(p)$ alors en posant $F_n = \frac{1}{n} \sum_{k=1}^n X_k$ on a,

$$\forall \varepsilon > 0, \mathbf{P}(|F_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$

9.2 Convergence en loi


D $\textcircled{S}\textcircled{C}$: $(X_n)_{n \geq 1} \xrightarrow{\mathcal{L}} X$ lorsque $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$ en tout point où F_X est continue.

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.

P $\textcircled{S}\textcircled{C}$: Soit a et b deux réels tel que $a < b$ alors on a,

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

P \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} .

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

T **Théorème de composition**


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

R Il est suffisant d'avoir $f \in \mathcal{C}^0(X(\Omega), \mathbb{R})$ mais ce n'est pas dit dans le programme.

T **Théorème de Slutsky**

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

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

R  Slutsky mélange les types de convergence.

T **Théorème central limite (TCL)**

$\blacktriangleright \textcircled{S}\textcircled{C}$: 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

$$\forall n \in \mathbb{N}^*, S_n = \sum_{k=1}^n X_k \text{ et } S_n^* = \frac{S_n - \mathbf{E}(S_n)}{\sigma(S_n)} \text{ où } \sigma(S_n) > 0.$$

Autrement dit $\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbf{P}([S_n^* \leq x]) = \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$.

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

$\blacktriangleright \textcircled{S}\textcircled{C}$: Le TCL existe aussi en "version moyenne", c'est-à-dire qu'en utilisant les mêmes notations et en posant $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ alors

$$\left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \right)_{n \geq 1} \xrightarrow{\mathcal{L}} N \text{ où } N \hookrightarrow \mathcal{N}(0, 1).$$

C $\textcircled{S}\textcircled{C}$: $\forall (a, b) \in \mathbb{R}^2, a < b$,

$$\lim_{n \rightarrow +\infty} \mathbf{P} \left(\left[a \leq \frac{S_n - n\mathbf{E}(X_1)}{\sqrt{n\mathbf{V}(X_1)}} \leq b \right] \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

9.3 Approximations

P 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 $p \leq 0,1$).

P Binomiale par Normale

Soit $\forall n \in \mathbf{N}^*$, $X_n \hookrightarrow \mathcal{B}(p)$ 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 $\forall n \in \mathbf{N}^*$, $S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}}$ (conditions $n \geq 30$, $np \geq 5$ et $nq \geq 5$).

P Poisson par Normale

Soit $\forall n \in \mathbf{N}^*$, $X_n \hookrightarrow \mathcal{P}(\lambda)$ 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 $\forall n \in \mathbf{N}^*$, $S_n^* = \frac{S_n - n\lambda}{\sqrt{n\lambda}}$ (condition $\lambda \geq 15$).

R 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 ...

10 Estimation inférentielle

10.1 Estimation ponctuelle

10.1.1 Généralités

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

D n -échantillon d'une variable aléatoire

S**C** : 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 .

D Statistique ou estimateur d'un paramètre

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

D Estimation

S**C** : 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 $\omega \in \Omega$.

D Estimateur convergent ou consistant d'un paramètre

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

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

D Biais d'un estimateur T_n par rapport à $g(\theta)$

S**C** : C'est un réel noté $b_\theta(T_n)$ défini par $b_\theta(T_n) = \mathbf{E}(T_n - g(\theta))$.

► **S****C** : Si $\forall \theta \in \Theta$, $b_\theta(T_n) = 0$ on dit que l'**estimateur est sans biais**.

► **S****C** : Si $\forall \theta \in \Theta$, $\lim_{n \rightarrow +\infty} b_\theta(T_n) = 0$ on dit que l'**estimateur est asymptotiquement sans biais**.

P Conditions suffisantes qu'un estimateur soit convergent ou consistant

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

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

D Risque quadratique moyen

S**C** : C'est le nombre réel noté $r_\theta(T_n)$ défini par

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

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

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

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

P **S****C** : $\forall \theta \in \Theta, (r_\theta(T_n) = \mathbf{V}_\theta(T_n)) \Leftrightarrow (b_\theta(T_n) = 0)$

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

10.1.2 Exemples d'estimateurs de paramètres usuels et sans biais et convergents

10.1.2.1 Moyenne empirique

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

D Moyenne empirique

S**C** : Soit (X_1, \dots, X_n) un n -échantillon de X d'espérance inconnue μ , on appelle moyenne empirique la variable aléatoire \bar{X}_n définie par $\forall n \in \mathbf{N}^*$, $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ où

$\forall k \in \llbracket 1, n \rrbracket, \mathbf{E}_\mu(X_k) = \mu$ (inconnue).

P **S****C** : $\forall \mu \in \mathbf{R}, \mathbf{E}_\mu(\bar{X}_n) = \mu$ et $\mathbf{V}_\mu(\bar{X}_n) = \frac{\sigma^2}{n}$.

P ► **C** : $(\forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{N}(\mu, \sigma^2)) \Rightarrow \left(\bar{X}_n \hookrightarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \right)$ où $\mu \in \mathbf{R}$ et $\sigma^2 \in \mathbf{R}_+^*$.

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

P **S****C** : $(\bar{X}_n)_{n \geq 1} \xrightarrow{\mathbf{P}} \mu$.

10.1.2.2 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.

D Fréquence empirique

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

$$\forall n \in \mathbf{N}^*, F_n = \frac{1}{n} \sum_{k=1}^n X_k \text{ avec } \forall k \in \llbracket 1, n \rrbracket, X_k \hookrightarrow \mathcal{B}(p) \text{ où } p \text{ est inconnu.}$$

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

P $F_n \underset{\approx}{\simeq} \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$ par le **TCL** pour n grand, $np \geq 5$ et $np(1-p) \geq 5$.

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

10.1.2.3 Variance empirique

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

D **Variance empirique**

S : Soit (X_1, \dots, X_n) un n -échantillon tel que $\forall k \in \llbracket 1, n \rrbracket$, $\mathbf{E}(X_k) = \mu$ et $\mathbf{V}(X_k) = \sigma^2$ inconnu. On appelle variance empirique la variable aléatoire S_n^2 définie par $\forall n \in \mathbf{N}^*$,

$$S_n^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X}_n)^2.$$

P **S** : $\forall n \in \mathbf{N}^*$, $S_n^2 = \frac{1}{n} \sum_{k=1}^n X_k^2 - (\bar{X}_n)^2$ (à savoir retrouver).

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

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

10.2 Estimation par intervalle de confiance

D **Intervalle de confiance d'un paramètre**

S : 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 $\mathbf{P}([U_n, V_n] \ni g(\theta)) \geq 1 - \alpha$ où U_n et V_n sont deux estimateurs $g(\theta)$. Sa réalisation est l'estimation de cet intervalle de confiance.

R *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.*

D **Intervalle de confiance asymptotique d'un paramètre**

S : 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 $\mathbf{P}([U_n, V_n] \ni g(\theta)) \geq 1 - \alpha_n$ donc $\lim_{n \rightarrow +\infty} \mathbf{P}([U_n, V_n] \ni g(\theta)) \geq 1 - \alpha$ où U_n et V_n sont deux estimateurs $g(\theta)$.

R *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 le sujet.*

10.2.1 Estimation par intervalle de confiance d'une proportion

10.1.2.1 Par l'inégalité de Bienaymé-Tchebychev

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

10.1.2.2 Par le théorème de la limite centrée

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

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

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

à condition que $\min(nu_n, nv_n, n(1-u_n), n(1-v_n)) \geq 5$.

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

$$\mathbf{P} IC_\alpha(\mu) = \left[\bar{X}_n - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} ; \bar{X}_n + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \text{ où } u_{1-\alpha/2} = \Phi^{-1}(1 - \alpha/2)$$

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

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

11 Zoologie des lois usuelles du programme

Les lois sont classées par ordre alphabétique.

11.1 Loi de Bernoulli \textcircled{S}

- ▶ **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 $\mathbf{P}([X = 1]) = p$
- ▶ $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } x \in [0, 1[\\ 1 & \text{si } x \geq 1 \end{cases}$
- ▶ $\mathbf{E}(X) = p$
- ▶ $\mathbf{V}(X) = p(1 - p)$
- ▶ $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (1 - X \hookrightarrow \mathcal{B}(1 - p))$
- ▶ $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X \hookrightarrow \mathcal{B}(1, p))$
- ▶ X_1, \dots, X_n i.i.d. ($\forall k \in [1, n], X_k \hookrightarrow \mathcal{B}(p)$) $\Rightarrow \left(\sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p) \right)$

11.2 Variable indicatrice \textcircled{S}

$$\boxed{\mathbf{D}} \forall A \in \mathcal{A}, \mathbf{1}_A = \begin{cases} 0 & \text{si } \bar{A} \\ 1 & \text{si } A \end{cases}$$

$\boxed{\mathbf{R}}$ Par définition une variable 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 $\mathbf{P}(\mathbf{1}_A = 0) = \mathbf{P}(\bar{A})$
- ▶ $\forall x \in \mathbb{R}, F_{\mathbf{1}_A}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \mathbf{P}(A) & \text{si } x \in [0, 1[\\ 1 & \text{si } x \geq 1 \end{cases}$
- ▶ $\mathbf{E}(\mathbf{1}_A) = \mathbf{P}(A)$
- ▶ $\mathbf{V}(\mathbf{1}_A) = \mathbf{P}(A)(1 - \mathbf{P}(A))$
- ▶ $\mathbf{1}_{A \cap B} = \mathbf{1}_A \mathbf{1}_B$ (ce qui se généralise par récurrence)
- ▶ $\forall n \geq 1, (\mathbf{1}_A)^n = \mathbf{1}_A$
- ▶ $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \mathbf{1}_B$
- ▶ $\mathbf{1}_{A \Delta B} = |\mathbf{1}_A - \mathbf{1}_B|$

11.3 Loi binomiale \textcircled{S}

- ▶ **Notation** : $X \hookrightarrow \mathcal{B}(n, p)$ avec $n \in \mathbb{N}^*$ et $p \in]0, 1[$
- ▶ **Paramètre** : n et p
- ▶ **Epreuve type** : succession de n épreuves de Bernoulli indépendantes et de même paramètre p .
- ▶ $X(\Omega) = [0, n]$
- ▶ $\forall k \in [0, n], \mathbf{P}([X = k]) = \binom{n}{k} p^k (1 - p)^{n-k}$
- ▶ $\mathbf{E}(X) = np$
- ▶ $\mathbf{V}(X) = np(1 - p)$
- ▶ $(X \hookrightarrow \mathcal{B}(n, p)) \Leftrightarrow (n - X \hookrightarrow \mathcal{B}(n, 1 - p))$
- ▶ $(X \hookrightarrow \mathcal{B}(p)) \Leftrightarrow (X \hookrightarrow \mathcal{B}(1, p))$
- ▶ X_1, \dots, X_n i.i.d. ($\forall k \in [1, n], X_k \hookrightarrow \mathcal{B}(p)$) $\Rightarrow \left(\sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p) \right)$
- ▶ **Stabilité de la loi binomiale par convolution**
- ▶ $\mathcal{B}(n_1, p) * \dots * \mathcal{B}(n_\ell, p) \hookrightarrow \mathcal{B}\left(\sum_{k=1}^{\ell} n_k, p\right)$
- ▶ **Approximations**
- ▶ $(X \hookrightarrow \mathcal{B}(n, p)$ et $n \geq 30, p < 0, 1) \Rightarrow \left(X \underset{\approx}{\hookrightarrow} \mathcal{P}(np) \right)$
- ▶ $\left. \begin{array}{l} X \hookrightarrow \mathcal{B}(n, p) \\ n \geq 30, np \geq 5, np(1 - p) \geq 5 \end{array} \right\} \Rightarrow \left(X \underset{\approx}{\hookrightarrow} \mathcal{N}(np, np(1 - p)) \right)$

11.4 Loi de Dirac (ou loi certaine) \textcircled{S}

- ▶ **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.

11.5 Loi exponentielle ③

- ▶ **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 ...

- ▶ $X(\Omega) = \mathbb{R}_+$ p.s.
- ▶ $\forall x \in \mathbb{R}, f_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x)$
- ▶ $\forall x \in \mathbb{R}, F_X(x) = (1 - e^{-\lambda x}) \mathbf{1}_{\mathbb{R}_+}(x)$
- ▶ $\mathbf{E}(X) = \frac{1}{\lambda}$
- ▶ $\mathbf{V}(X) = \frac{1}{\lambda^2}$
- ▶ **Absence de mémoire**

$(X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (\forall (x, y) \in \mathbb{R}_+^2, \mathbf{P}_{[X > x]}([X > x + y]) = \mathbf{P}([X > y]))$, autrement dit ce qui s'est passé sur l'intervalle de $]-\infty, x]$ n'affecte en rien ce qui se passera sur l'intervalle $]x, x + y]$.

- ▶ $\forall a > 0, (X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (aX \hookrightarrow \mathcal{E}\left(\frac{\lambda}{a}\right))$
- ▶ $\forall \lambda > 0, (X \hookrightarrow \mathcal{E}(1)) \Leftrightarrow \left(\frac{X}{\lambda} \hookrightarrow \mathcal{E}(\lambda)\right)$
- ▶ $\forall \lambda > 0, (X \hookrightarrow \mathcal{E}(\lambda)) \Leftrightarrow (\lambda X \hookrightarrow \mathcal{E}(1))$
- ▶ $\mathcal{E}(1) = \gamma(1)$

11.6 Loi gamma ③

- ▶ **Notation** : $X \hookrightarrow \gamma(v)$ avec $v \in \mathbb{R}_+^*$
- ▶ **Paramètre** : v
- ▶ $X(\Omega) = \mathbb{R}_+^*$ p.s.
- ▶ $\forall x \in \mathbb{R}, f_X(x) = \frac{x^{v-1} e^{-x}}{\Gamma(v)} \mathbf{1}_{\mathbb{R}_+^*}(x)$
- ▶ $\mathbf{E}(X) = v$
- ▶ $\mathbf{V}(X) = v$
- ▶ $\gamma(1) = \mathcal{E}(1)$
- ▶ $\gamma(v_1) * \dots * \gamma(v_n) = \gamma\left(\sum_{k=1}^n v_k\right)$

11.7 Loi géométrique ③

- ▶ **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.

- ▶ $X(\Omega) = \mathbb{N}^*$
- ▶ $\forall k \in \mathbb{N}^*, \mathbf{P}([X = k]) = q^{k-1} p$ 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**

$(X \hookrightarrow \mathcal{G}(p)) \Leftrightarrow (\forall (n, m) \in (\mathbb{N})^2, \mathbf{P}_{[X > n]}([X > n + m]) = \mathbf{P}([X > m]))$ ce qui s'est passé sur l'intervalle $]-\infty, n]$ n'affecte en rien ce qui se passera sur l'intervalle $]n, n + m]$.

11.8 Lois normales ③

11.8.1 Loi normale centrée et réduite ou de Laplace-Gauss centrée et réduite

- ▶ **Notation** : $X \hookrightarrow \mathcal{N}(0, 1)$
- ▶ **Paramètre** : 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**
 - ▷ $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1$
 - ▷ $\int_{-\infty}^{+\infty} e^{-x^2} dx = 1$
- ▶ **Fonction de répartition**
 - ▷ $\forall x \in \mathbb{R}, \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$
 - ▷ $\forall x \in \mathbb{R}, \Phi(-x) = 1 - \Phi(x)$
 - ▷ $\Phi(0) = 1/2$
- ▶ **Mode** : 0
- ▶ **Médiane** : 0
- ▶ $\mathbf{E}(X) = 0$
- ▶ $\mathbf{V}(X) = 1$
- ▶ $\left(\frac{X - m}{\sigma} \hookrightarrow \mathcal{N}(0, 1)\right) \Leftrightarrow (X \hookrightarrow \mathcal{N}(m, \sigma^2))$
- ▶ **Stabilité de la loi normale centrée réduite par convolution**

$$\underbrace{\mathcal{N}(0,1) * \dots * \mathcal{N}(0,1)}_{n \text{ fois}} = \mathcal{N}(0,n)$$

► **Approximation**

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

11.8.2 Loi normale quelconque ou de Laplace-Gauss

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

► **Paramètre** : m et σ^2

► $X(\Omega) = \mathbb{R}$ p.s.

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

► **Mode** : m

► **Médiane** : m

► $\mathbf{E}(X) = m$

► $\mathbf{V}(X) = \sigma^2$

► **Théorème fondamental de la loi normale**

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

► **Stabilité de la loi normale par convolution**

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

11.9 Loi de Poisson (S)

► **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.

► $X(\Omega) = \mathbb{N}$

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

► $\mathbf{E}(X) = \lambda$

► $\mathbf{V}(X) = \lambda$

► **Stabilité de la loi de Poisson par convolution**

$$\mathcal{P}(\lambda_1) * \dots * \mathcal{P}(\lambda_n) = \mathcal{P}\left(\sum_{k=1}^n \lambda_k\right)$$

► **Approximation**

$$\left. \begin{array}{l} X \hookrightarrow \mathcal{P}(\lambda) \\ \lambda \geq 15 \end{array} \right\} \Rightarrow \left(X \underset{\approx}{\hookrightarrow} \mathcal{N}(\lambda, \lambda) \right)$$

11.10 Loïs uniformes

11.10.1 Loi uniforme discrète (S)

► **Notation** : $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$

► **Paramètre** : $\llbracket 1, n \rrbracket$

► **Epreuve type** : numéro d'une tirée d'une constituée de boules numérotées de 1 à n .

► $X(\Omega) = \llbracket 1, n \rrbracket$

► $\forall k \in \llbracket 1, n \rrbracket, \mathbf{P}([X = k]) = \frac{1}{n}$

► $\mathbf{E}(X) = \frac{n+1}{2}$

► $\mathbf{V}(X) = \frac{n^2-1}{12}$

► Soit $(a, b) \in \mathbb{R}^2$ avec $a < b$ alors

$(X \hookrightarrow \mathcal{U}(\llbracket a, b \rrbracket)) \Leftrightarrow (X - a + 1 \hookrightarrow \mathcal{U}(\llbracket 1, b - a + 1 \rrbracket))$ et $\mathbf{E}(X) = \frac{a+b}{2}$ et

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

11.10.2 Loi uniforme continue (C)

► **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.

► $\forall x \in \mathbb{R}, f_X(x) = \left(\frac{1}{b-a}\right) \mathbf{1}_I(x)$

► $\forall x \in \mathbb{R}, F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}$

► $\mathbf{E}(X) = \frac{a+b}{2}$

► $\mathbf{V}(X) = \frac{(b-a)^2}{12}$

► $(X \hookrightarrow \mathcal{U}([0, 1])) \Leftrightarrow (\forall (a, b) \in (\mathbb{R}_+^*)^2, a + (b-a) X \hookrightarrow \mathcal{U}([a, b]))$

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

11.11 Bilan des lois du programme stables par convolution

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

11.12 Bilan des lois du programme sans mémoire

▶ \textcircled{S} : Les seules variables à valeurs dans \mathbf{N}^* qui ont la propriété d'être sans mémoire, c'est-à-dire vérifiant : $\forall (n, m) \in (\mathbf{N}^*)^2$, $\mathbf{P}([X > n + m]) = \mathbf{P}([X > m]) \mathbf{P}([X > n])$ sont les variables géométriques.

▶ \textcircled{C} : Les seules variables à valeurs dans \mathbf{R}_+ qui ont la propriété d'être sans mémoire, c'est-à-dire vérifiant : $\forall (s, t) \in (\mathbf{R}_+)^2$, $\mathbf{P}([X > n + m]) = \mathbf{P}([X > m]) \mathbf{P}([X > n])$ sont les variables exponentielles.



Index

Approximation

- binomiale par Normale, 19
- binomiale par Poisson, 19
- poisson par Normale, 19

Calcul de probabilité

- à partir de la fonction de répartition, 10
- à partir de la fonction de survie, 10

Coefficient

- de corrélation linéaire, 15

Coefficient

- binomial, 5

Convergence

- en loi, 18
- en probabilités, 18

Dénombrement

- exemples fondamentaux de tirages usuels, 7
- lemme des bergers, 7
- nombre d'applications de E vers F, 7
- nombre d'arrangements de E vers F, 7
- nombre d'injections de E vers F, 7
- nombre de combinaisons de p éléments de E de cardinal n, 7
- nombre de couples d'entiers, 7
- nombre de p-listes de E, 7
- nombre de parties d'un ensemble E, 7
- nombre de permutations de E, 7
- nombre de suites d'entiers, 7

Densité

- caractérisation, 10

Droites de régression, 15

Echantillon, 20

Ensemble

- définition intuitive d'un ensemble, 3
- différence d'ensembles, 4
- différence symétrique d'ensembles, 4
- distributivités, 3

ensemble des parties d'un ensemble, 3

fonction indicatrice ou fonction caractéristique, 4

inclusion d'ensembles, 4

intersection d'ensembles, 3

lois de Morgan, 3

partition, 7

produit cartésien, 4

union, 3

Espérance

conditionnelle, 12

définition, 11

et variable bornée, 11

Espace

probabilisé, 8

probabilisable, 8

Estimateur

biais, 20

consistant, 20

convergent, 20

d'un paramètre, 20

risque quadratique moyen, 20

Estimation

définition, 20

par intervalle de confiance asymptotique d'une moyenne de loi normale de variance donnée, 21

ponctuelle, 20

Evénement

négligeable, 8

quasi-certain, 8

système complet, 8

Fonction

de répartition, 10

Fonctions de régression, 12

Formule

de Bayes, 9

de l'espérance totale, 12

- de Poincaré ou du crible pour deux ensembles
 - dénombrement, 7
- des probabilités composées, 9
- des probabilités totales, 9
- du binôme de Newton, 5
- du crible pour deux ensembles
 - probabilité, 9
- du triangle de Pascal, 5
- Fréquence empirique, 20
- Identité
 - de Bernoulli, 5
 - du parallélogramme, 15
- Identités
 - de polarisation, 15
- Inégalité
 - de Bienaymé-Tchebychev, 17
 - de Cauchy-Schwarz, 15, 17
 - de concentration, 17
 - de Jensen, 17
 - de Markov, 17
 - de moments, 12, 17
 - triangulaire, 17
- Indépendance
 - deux événements, 9
 - mutuelle, 9
 - n événements, 9
 - suite d'événements, 9
 - variables aléatoires
 - caractérisation, 16
- Intervalle de confiance
 - asymptotique d'un paramètre, 21
 - d'un paramètre, 21
 - d'une proportion, 21
- Lemme des coalitions, 16
- Les trois grandes partitions de \mathbb{N}^2
 - de deux familles sommables, 13
- Loi
 - binomiale, 22
 - caractérisation, 10
 - caractérisation de la loi d'une fonction d'un couple, 13
 - caractérisation de la loi d'une fonction d'un vecteur, 14
 - certaine, 22
 - conditionnelle, 13
 - d'un couple, 13
 - de Bernoulli, 22
 - de Dirac, 22
 - de Poisson, 24
 - de probabilité, 10
 - exponentielle, 23
 - faible des grands nombres, 18
 - géométrique, 23
 - gamma, 23
 - marginale, 13
 - normale, 23
 - uniforme, 24
- Matrice de covariance-variance d'un vecteur aléatoire, 15
- Moyenne empirique, 20
- Norme dans L_2 , 15
- Produit
 - à trous, 6
 - de deux familles sommables, 13
 - lien entre produit et somme, 6
 - permutation des produits, 5
 - propriétés élémentaires, 5
 - scalaire dans L_2 , 15
- Relation de Laplace, 9
- Statistique, 20
- Théorème
 - de composition
 - convergence en loi, 18
 - convergence en probabilité, 18
 - de convolution, 16
 - de domination, 11
 - de Koenig-Huygens, 11

- de la limite centrée, 18
- de la limite monotone, 9
- de Slutsky, 18
- de sommation par paquets, 13
- de transfert en dimension un, 11
- de transfert fonction d'un couple, 13
- de transfert fonction d'un vecteur, 14
- Koenig, 14

Tribu, 8

- borélienne de \mathbb{R} , 8
- engendrée par une famille, 8

Variable aléatoire

- bornée, 10
- centrée réduite, 11
- indicatrice, 22
- support, 10

Variables aléatoires

- non corrélées ou décorréélées, 14

Variance

- additivité, 15
- empirique, 21