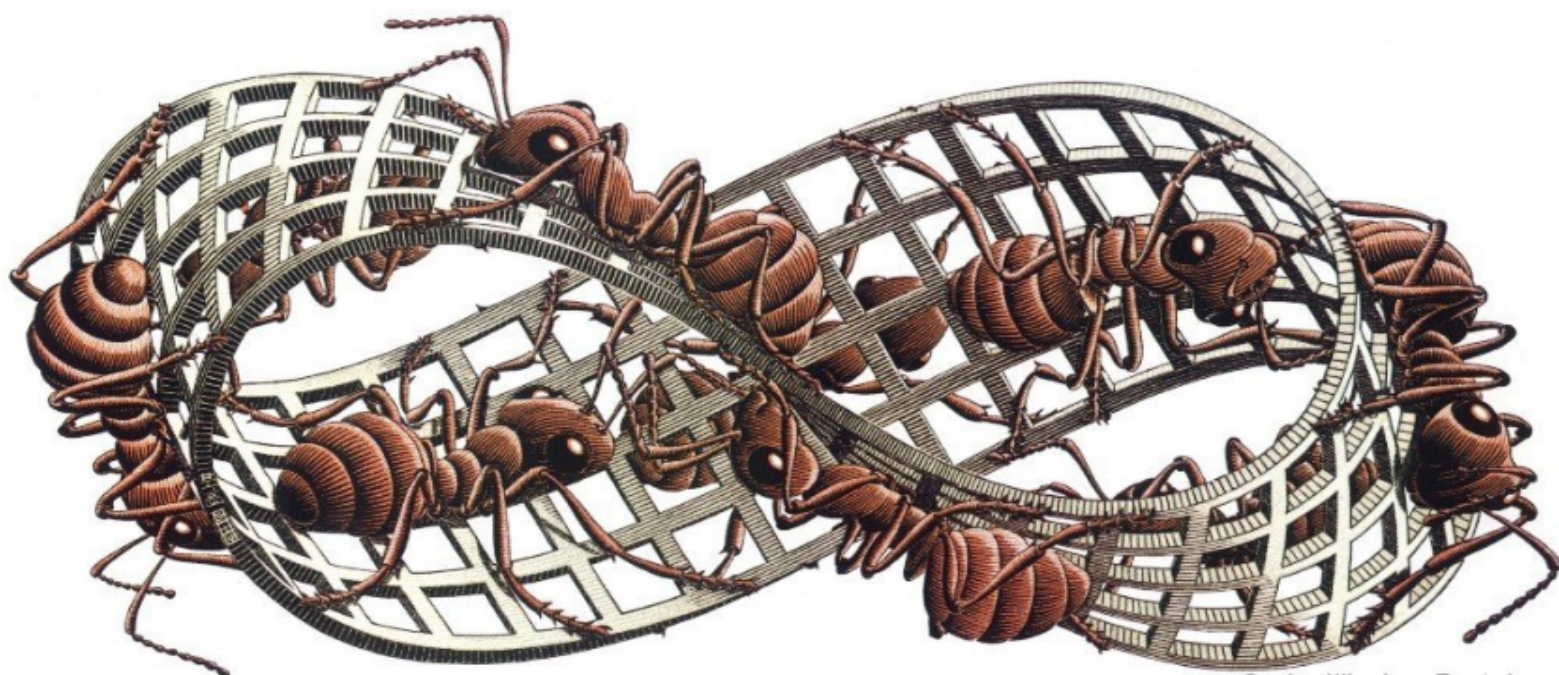


# Cours de BCPST 2

Mathématiques

BCPST 2A (*Marcelin Berthelot*)



# Table des matières

<b>1</b>	<b>Séries</b>	<b>4</b>
1.1	Définitions. vocabulaire. notation.	4
1.2	Des premiers résultats	5
1.3	Séries usuelles.	8
1.4	Théorèmes.	11
1.5	Absolue convergence.	13
<b>2</b>	<b>Polynômes</b>	<b>14</b>
2.1	Généralités.	14
2.2	Racines.	17
2.3	Dérivation (complément).	20
<b>3</b>	<b>Espaces vectoriels</b>	<b>22</b>
3.1	Structure vectorielle.	23
3.2	Sous-espace vectoriel	24
3.3	$\text{Vect}(u_1, \dots, u_n)$	25
3.4	Familles libres.	28
3.5	Bases.	29
3.6	Coordonnées dans une base.	29
3.7	Dimension d'un espace vectoriel	31
3.8	Famille libre, famille génératrice et dimension.	33
3.9	Systèmes linéaires homogènes et dimension.	34
3.10	Rang.	35
<b>4</b>	<b>Probabilités</b>	<b>38</b>
4.1	Ensemble des événements et probabilités.	38
4.2	Conditionnement et indépendance.	42
4.3	Variables aléatoires réelles.	45
4.4	Espérance, variance et moments d'une variable aléatoire réelle.	46
<b>5</b>	<b>Variables aléatoires discrètes</b>	<b>50</b>
5.1	Variables aléatoires discrètes.	50
5.2	Indépendance.	55
5.3	Lois finies usuelles.	57
5.4	Lois discrètes infinies usuelles.	61
<b>6</b>	<b>Applications linéaires.</b>	<b>67</b>
6.1	Applications linéaires.	67
6.2	Opérations et applications linéaires.	70
6.3	Noyau et Image.	72
6.4	Image d'une base.	74
6.5	Représentation matricielle.	77
6.6	Isomorphe à $\mathbb{K}^n$ .	80
6.7	Rang d'une application linéaire.	81
6.8	Une matrice vue comme une application linéaire.	82
6.9	Changement de base.	84
6.10	Matrices semblables	86
<b>7</b>	<b>Intégrales généralisées</b>	<b>88</b>
7.1	Introduction	88
7.2	Définitions. vocabulaire. notation.	89
7.3	Théorèmes de convergence.	93

7.4	Propriétés . . . . .	95
7.5	Absolue convergence. . . . .	99
7.6	Calculs. . . . .	100
7.7	Intégrales impropres usuelles. . . . .	102
<b>8</b>	<b>Diagonalisation</b> . . . . .	<b>103</b>
8.1	Introduction . . . . .	103
8.2	Éléments propres. . . . .	104
8.3	Propriétés des familles de vecteurs propres. . . . .	108
8.4	Diagonalisation. . . . .	110
<b>9</b>	<b>Variables aléatoires réelles à densité.</b> . . . .	<b>115</b>
9.1	Densités de probabilité et fonction de répartition. . . . .	115
9.2	Méthodes pratiques. . . . .	117
9.3	Calculs de probabilité. . . . .	119
9.4	Exemples de recherche de la loi de $u(X)$ . . . . .	121
9.5	Espérance. Propriétés. . . . .	122
9.6	Lois à densité usuelles. . . . .	124
9.7	Somme de variables aléatoires à densité indépendantes. . . . .	139
<b>10</b>	<b>Statistique.</b> . . . .	<b>141</b>
10.1	Statistique univariée. . . . .	141
10.2	Statistique bivariée. . . . .	144
10.3	Loi faible des grands nombres. . . . .	145
10.4	Echantillonnage. Estimation. . . . .	149
10.5	Convergence en loi . . . . .	152
10.6	Théorème central limite . . . . .	153
10.7	Intervalle de confiance de la moyenne. . . . .	156
10.8	Test de conformité à la moyenne. . . . .	158
<b>11</b>	<b>Produit scalaire dans <math>\mathbb{R}^n</math>.</b> . . . .	<b>161</b>
11.1	Généralités . . . . .	161
11.2	Bases orthonormales. . . . .	166
11.3	Matrice symétrique réelle. . . . .	170
11.4	Orthogonal d'un sous-espace vectoriel . . . . .	171
11.5	Projection orthogonale. . . . .	174
11.6	Distance. . . . .	176
<b>12</b>	<b>Couples de variables aléatoires réelles.</b> . . . .	<b>178</b>
12.1	Couples de variables aléatoires discrètes. . . . .	178
12.2	Covariance. Variance d'une somme. . . . .	182
<b>Annexes</b>		<b>187</b>
<b>A</b>	<b>Fonctions indicatrices.</b> . . . .	<b>187</b>
<b>B</b>	<b>Minimum. Maximum. Somme</b> . . . . .	<b>189</b>
B.1	Variables aléatoires à valeurs dans $\mathbb{N}$ . . . . .	189
B.2	Variables aléatoires à densité . . . . .	190
<b>C</b>	<b>Régression linéaire</b> . . . . .	<b>192</b>
<b>D</b>	<b>Annexe — Équations différentielles autonomes d'ordre 1</b> . . . . .	<b>195</b>
D.1	Modèle malthusien (croissance exponentielle). . . . .	195
D.2	Modèle logistique (croissance avec saturation). . . . .	196
D.3	Modèle de Gompertz. . . . .	197
D.4	Modèle de Lotka–Volterra (proie–prédateur). . . . .	197
D.5	Lien avec l'informatique (méthode d'Euler). . . . .	198
<b>E</b>	<b>Statistiques transversales : test de conformité, variabilité biologique et validation d'un modèle</b> . . . . .	<b>201</b>
E.1	Variabilité biologique (SVT) . . . . .	201

---

E.2	Test de conformité à la moyenne (mathématiques) . . . . .	202
E.3	Validation d'un modèle (physique) . . . . .	202
E.4	Incertitudes de type A et de type B . . . . .	203
E.5	Méthode de Monte-Carlo . . . . .	204
E.6	Lien avec les TIPE . . . . .	204
E.7	Vocabulaire et présentation des résultats . . . . .	204
	Tableau de synthèse . . . . .	205

# Séries

## Plan du chapitre

1.1	Définitions. vocabulaire. notation. . . . .	4
1.2	Des premiers résultats . . . . .	5
1.2.1	Indice de départ. . . . .	5
1.2.2	Une condition nécessaire de convergence. . . . .	6
1.2.3	Séries télescopiques. . . . .	6
1.2.4	Combinaisons linéaires de séries convergentes . . . . .	6
1.2.5	Séries $\sum u_n$ où $(u_n)$ est à support fini. . . . .	7
1.2.6	Séries à termes positifs. . . . .	7
1.3	Séries usuelles. . . . .	8
1.3.1	Séries géométriques et les dérivées. . . . .	8
1.3.2	Série exponentielle. . . . .	9
1.3.3	Séries de Riemann. . . . .	10
1.4	Théorèmes. . . . .	11
1.4.1	Théorème de convergence par comparaison des termes généraux positifs. . . . .	11
1.4.2	Théorème de convergence par équivalence des termes généraux positifs. . . . .	12
1.5	Absolute convergence. . . . .	13
1.5.1	Théorème. . . . .	13
1.5.2	Convergence commutative. . . . .	13

## 1.1 Définitions. vocabulaire. notation.

### Définition

Soit  $m$  un entier naturel non nul et  $(u_n)_{n \geq m}$  une suite de réels.

Étudier la série de terme général  $u_n$  signifie étudier la suite  $\left( \sum_{k=m}^n u_k \right)_{n \geq m}$  (on note  $(S_n)_{n \geq m}$  cette suite)

La série de terme général  $u_n$  est notée :  $\sum_{n \geq m} u_n$ .

- Les  $S_n = \sum_{k=m}^n u_k$  sont les **sommes partielles de la série**  $\sum_{n \geq m} u_n$ .
- Dire que la série  $\sum_{n \geq m} u_n$  est **convergente** signifie que  $(S_n)$  est convergente.
- Si  $(S_n)$  est convergente vers le réel  $\ell$  alors  $\ell$  est appelé **somme de la série**  $\sum_{n \geq m} u_n$ .

et dans ce cas on note :  $\sum_{k=m}^{+\infty} u_k = \lim_{n \rightarrow +\infty} S_n (= \ell)$

- Dire que la série  $\sum_{n \geq m} u_n$  est **divergente** signifie que  $(S_n)$  est divergente.

**Exemples :**

$$1) \sum_{n \geq 0} (-2)^{-n} \quad 2) \sum_{n \geq 0} n \quad 3) \sum_{n \geq 1} \frac{1}{2^n} \quad 4) \sum_{n \geq 1} \left( \frac{1}{n} - \frac{1}{n+1} \right) \quad 5) \sum_{n \geq 1} (-0.1)^n \quad 6) \sum_{n \geq 1} 2^n$$

**En pratique :** Pour étudier la nature d'une série :

- on étudie la convergence de la suite  $(S_n)$  définie par :  $S_n = \sum_{k=m}^n u_k$ .

$$\begin{aligned} S_n &= \sum_{k=m}^n u_k \\ &= \dots \\ &= \dots \\ &\xrightarrow{n \rightarrow +\infty} \ell \end{aligned}$$

- on conclut en utilisant le vocabulaire des séries :

Lorsque  $(S_n)$  converge vers  $\ell$ , on dit que la série  $\sum_{n \geq m} u_n$  converge et que sa somme  $\sum_{k=m}^{+\infty} u_k$  est égale à  $\ell$ .

Lorsque  $(S_n)$  diverge, on dit que la série  $\sum_{n \geq m} u_n$  diverge.

**Remarques :**

La suite  $(S_n)$  est la suite des sommes partielles de la série  $\sum_{n \geq m} u_n$

Lorsque  $(S_n)$  converge vers  $\ell$ , on dit que la somme " $\sum_{k=m}^{+\infty} u_k$  existe" et vaut  $\ell$ .

## 1.2 Des premiers résultats

### 1.2.1 Indice de départ.

**Proposition**

Soit  $(u_n)_{n \geq m}$  une suite de réels et  $n_1$  et  $n_2$  deux entiers naturels supérieurs à  $m$ ,  
 La série  $\sum_{n \geq n_1} u_n$  est convergente si, et seulement si, la série  $\sum_{n \geq n_2} u_n$  est convergente.

**En effet :** On suppose que  $n_1 < n_2$

$$\forall n \geq n_2, \quad \sum_{k=n_1}^n u_k = \underbrace{\sum_{k=n_1}^{n_2-1} u_k}_{\text{ne dépend pas de } n} + \sum_{k=n_2}^n u_k$$

donc les deux suites  $\left( \sum_{k=n_1}^n u_k \right)$  et  $\left( \sum_{k=n_2}^n u_k \right)$  convergent simultanément.

**Remarque :** en cas de convergence, toujours avec  $n_1 < n_2$ , on a

$$\sum_{k=n_1}^{+\infty} u_k = \sum_{k=n_1}^{n_2-1} u_k + \sum_{k=n_2}^{+\infty} u_k$$

**Remarques :**

- On dit aussi que les séries  $\sum_{n \geq n_1} u_n$  et  $\sum_{n \geq n_2} u_n$  ont même nature.

- La nature d'une série ne dépend pas des premiers termes.
- Quand on demande la nature de la série, il est inutile de donner le premier indice.
- On peut ne rien mettre sous la somme. " Déterminer la nature de  $\sum u_n$  ".

### 1.2.2 Une condition nécessaire de convergence.

**Théorème.**

Soit  $(u_n)_{n \geq m}$  une suite de réels,

Si la série  $\sum_{n \geq m} u_n$  est convergente alors la suite  $(u_n)$  tend vers 0.

**En effet :**  $u_n = S_n - S_{n-1}$ , donc si  $(S_n)$  converge vers  $\ell$ ,  $(S_{n-1})$  aussi et ainsi  $(u_n)$  converge vers 0.

**En pratique** on utilise souvent la forme *contraposée* de cette implication :

Si la suite  $(u_n)$  ne tend pas vers 0 alors la série  $\sum_{n \geq m} u_n$  est **divergente**

**Attention** la *réciproque* est fautive. Contre-exemple :  $\sum_{n \geq 1} \frac{1}{n}$  est divergente et pourtant  $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$

*Trop nombreux sont ceux qui font l'erreur de raisonnement de dire que la série converge lorsque  $u_n \xrightarrow[n \rightarrow +\infty]{} 0$ .*

### 1.2.3 Séries télescopiques.

**Proposition.** (Complément)

Soit  $(a_n)$  une suite réelle quelconque, on note  $(u_n)$  la suite définie par  $u_n = a_{n+1} - a_n$ ,

La série de terme général  $u_n$  converge si, et seulement si, la suite  $(a_n)$  est convergente.

et en cas de convergence :  $\sum_{k=m}^{+\infty} u_k = \lim_{n \rightarrow +\infty} a_n - a_m$

*En pratique on le redémontre, car il y a plusieurs situations possibles :  $u_n = a_{n-1} - a_n$ ,  $u_n = a_n - a_{n+1}$  ...*

**En effet :** C'est une conséquence directe de la somme télescopique :  $S_n = \sum_{k=m}^n (a_{k+1} - a_k) = a_{n+1} - a_m$

### 1.2.4 Combinaisons linéaires de séries convergentes

**Théorème**

Soient  $(u_n)_{n \geq m}$  et  $(v_n)_{n \geq m}$  deux suites de réels,  $\alpha$  et  $\beta$  deux réels.

Si les séries  $\sum_{n \geq m} u_n$  et  $\sum_{n \geq m} v_n$  sont convergentes alors la série  $\sum_{n \geq m} (\alpha u_n + \beta v_n)$  est convergente.

et alors :  $\sum_{k=m}^{+\infty} (\alpha u_k + \beta v_k) = \alpha \sum_{k=m}^{+\infty} u_k + \beta \sum_{k=m}^{+\infty} v_k$

**Démonstration :**

On suppose que les séries  $\sum u_n$  et  $\sum v_n$  convergent,

Pour  $n \geq m$

$$\sum_{k=m}^n (\alpha u_k + \beta v_k) = \alpha \sum_{k=m}^n u_k + \beta \sum_{k=m}^n v_k \quad (\text{Propriétés des sommes})$$

or les deux séries convergent donc  $\left(\sum_{k=m}^n (\alpha u_k + \beta v_k)\right)_{n \geq m}$  converge.

On a aussi l'égalité :  $\sum_{n=m}^{+\infty} (\alpha u_n + \beta v_n) = \alpha \sum_{n=m}^{+\infty} u_n + \beta \sum_{n=m}^{+\infty} v_n$

### Corollaire

Pour  $\alpha \neq 0$ ,

$$\sum_{n \geq m} u_n \text{ converge si, et seulement si, } \sum_{n \geq m} \alpha u_n \text{ converge.}$$

**En effet :** Quand  $\alpha \neq 0$   $\sum_{n \geq m} u_n = \frac{1}{\alpha} \sum_{n \geq m} \alpha u_n$

### 1.2.5 Séries $\sum u_n$ où $(u_n)$ est à support fini.

#### Proposition.

Soit  $(u_n)_{n \geq 0}$  une suite de réels,

S'il existe un rang  $N$  à partir duquel  $u_n = 0$  alors la série  $\sum u_n$  est convergente.

$$\text{et alors : } \sum_{k=0}^{+\infty} u_k = \sum_{k=0}^{N-1} u_k$$

**En effet :** La suite  $(S_n)$  est constante à partir de  $N$ . (La suite est stationnaire)

Phrase de rédaction : "La somme  $\sum_{k=0}^{+\infty} u_k$  existe car  $(u_n)$  est à support fini."

**Exemples :**  $\sum_{k=0}^{+\infty} \binom{10}{k}$  existe et vaut  $\sum_{k=0}^{10} \binom{10}{k} = 2^{10}$

### Lien avec les fonctions polynômiales réelles :

Dire qu'une fonction  $P$  est polynomiale signifie qu'il existe une suite  $(a_n)$  nulle à partir d'un certain rang telle

$$\text{que : } P : x \mapsto \sum_{k=0}^{+\infty} a_k x^k$$

### 1.2.6 Séries à termes positifs.

Deux résultats à redémontrer dans une copie.

#### Proposition. (Complément)

Soit  $(u_n)_{n \geq m}$  une suite de réels,

Si  $(u_n)$  est à **termes positifs ou nuls** alors la suite  $\left(\sum_{k=m}^n u_k\right)_{n \geq m}$  est croissante.

**En effet :** pour tout  $n \geq m$ ,  $S_{n+1} - S_n = u_{n+1} \geq 0$

#### Proposition. (Complément)

Soit  $(u_n)_{n \geq m}$  une suite de réels,

Si  $(u_n)$  est à **termes positifs ou nuls** alors on a l'équivalence suivante :

La série  $\sum_{n \geq m} u_n$  est convergente si, et seulement si, la suite  $\left(\sum_{k=m}^n u_k\right)_{n \geq m}$  est majorée.

$$\text{et en cas de convergence on a : } \forall n \geq m, \sum_{k=m}^n u_k \leq \sum_{k=m}^{+\infty} u_k$$

**En effet :** C'est une conséquence du théorème de convergence monotone appliqué à la suite  $(S_n)$

## 1.3 Séries usuelles.

### 1.3.1 Séries géométriques et les dérivées.

**Théorème.**

Soit  $q$  un nombre réel,

❶ Série géométrique de raison  $q$ .

La série  $\sum_{n \geq 0} q^n$  est convergente si, et seulement si,  $-1 < q < 1$ ; et alors  $\sum_{k=0}^{+\infty} q^k = \frac{1}{1-q}$ .

❷ Série géométrique dérivée de raison  $q$ .

La série  $\sum_{n \geq 1} nq^{n-1}$  est convergente si, et seulement si,  $-1 < q < 1$ ; et alors  $\sum_{k=1}^{+\infty} kq^{k-1} = \frac{1}{(1-q)^2}$ .

❸ Série géométrique dérivée d'ordre 2 de raison  $q$ .

La série  $\sum_{n \geq 2} n(n-1)q^{n-2}$  est convergente si, et seulement si,  $-1 < q < 1$

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

**Exemples.**

$$\sum_{k=0}^{+\infty} 2^{-k} = \frac{1}{1-\frac{1}{2}} = 2 \quad ; \quad \sum_{n=1}^{+\infty} n \left(-\frac{1}{3}\right)^{n-1} = \frac{1}{\left(1+\frac{1}{3}\right)^2} = \frac{9}{16} \quad ; \quad \sum_{k=2}^{+\infty} k(k-1) \left(\frac{2}{5}\right)^{k-2} = \frac{2}{\left(1-\frac{2}{5}\right)^3} = \frac{250}{27}$$

**Remarques :**

- pour retrouver l'expression des séries dérivées on dérive deux fois la fonction  $x \mapsto \frac{1}{1-x}$
- Pour  $q \in ]-1, 1[$ , on peut changer l'indice de départ pour les séries dérivées :

$$\sum_{k=0}^{+\infty} kq^{k-1} = \frac{1}{(1-q)^2} \quad \text{et} \quad \sum_{k=0}^{+\infty} k(k-1)q^{k-2} = \frac{2}{(1-q)^3}.$$

**Démonstrations.**

Si  $q \notin ]-1, 1[$  alors les termes généraux  $(q^n)$ ,  $(nq^{n-1})$  et  $(n(n-1)q^{n-2})$  ne convergent pas vers 0 donc les séries divergent dans les cas ❶, ❷ et ❸.

Il suffit donc d'étudier la convergence dans les trois cas avec  $q \in ]-1, 1[$  :

❶

$$\begin{aligned} \sum_{k=0}^n q^k &= \frac{1-q^{n+1}}{1-q} \\ &\xrightarrow{n \rightarrow +\infty} \frac{1}{1-q} \end{aligned}$$

❷

$$\begin{aligned} (1-q) \sum_{k=1}^n kq^{k-1} &= \sum_{k=1}^n kq^{k-1} - \sum_{k=1}^n kq^k \\ &= \sum_{k=0}^{n-1} (k+1)q^k - \sum_{k=1}^n kq^k \end{aligned}$$

$$\begin{aligned}
 &= \sum_{k=0}^{n-1} kq^k - \sum_{k=1}^n kq^k + \sum_{k=0}^{n-1} q^k \\
 &= -nq^n + \sum_{k=0}^{n-1} q^k \\
 &\xrightarrow{n \rightarrow +\infty} 0 + \frac{1}{1-q} \quad (\text{On utilise } \textcircled{1} \text{ et une croiss. comparée})
 \end{aligned}$$

$$\sum_{k=1}^n kq^{k-1} \xrightarrow{n \rightarrow +\infty} \frac{1}{(1-q)^2}$$

③

$$\begin{aligned}
 (1-q) \sum_{k=2}^n k(k-1)q^{k-2} &= \sum_{k=2}^n k(k-1)q^{k-2} - \sum_{k=2}^n k(k-1)q^{k-1} \\
 &= \sum_{k=1}^{n-1} (k+1)kq^{k-1} - \sum_{k=1}^n k(k-1)q^{k-1} \\
 &= \sum_{k=1}^n (k^2 + k - k^2 + k)q^{k-1} - (n+1)nq^{n-1} \\
 &= 2 \sum_{k=1}^n kq^{k-1} - (n+1)nq^{n-1} \\
 &\xrightarrow{n \rightarrow +\infty} \frac{2}{(1-q)^2} - 0 \quad (\text{On utilise } \textcircled{2})
 \end{aligned}$$

En effet :  $(n+1)nq^{n-1} \sim \frac{n^2q^n}{q}$  et  $\lim_{n \rightarrow +\infty} n^2q^n = 0$  (limite du cours appelé "croissances comparées")

$$\sum_{k=1}^n kq^{k-1} \xrightarrow{n \rightarrow +\infty} \frac{1}{(1-q)^2}$$

Quelques sommes autour de celles-ci (Savoir les retrouver rapidement, ne pas les apprendre)

$\sum_{k=1}^{+\infty} q^k = \frac{q}{1-q}$	$\sum_{n=1}^{+\infty} q^{n-1} = \frac{1}{1-q}$	$\sum_{k=m}^{+\infty} q^k = \frac{q^m}{1-q}$	$\sum_{n=m}^{+\infty} q^{n-m} = \frac{1}{1-q}$
$\sum_{k=1}^{+\infty} kq^k = \frac{q}{(1-q)^2}$	$\sum_{n=0}^{+\infty} (n+1)q^n = \frac{1}{(1-q)^2}$	$\sum_{k=0}^{+\infty} k(k-1)q^k = \frac{2q^2}{(1-q)^3}$	$\dots$

### 1.3.2 Série exponentielle.

**Théorème.**

Quel que soit le nombre réel  $x$ ,

la série  $\sum_{n \geq 0} \frac{x^n}{n!}$  est convergente et  $\sum_{k=0}^{+\infty} \frac{x^k}{k!} = e^x$

**Démonstration.** (Admis) Voir la feuille\_cours\_1 Ex 9

**Exemples.**

$$\sum_{n=0}^{+\infty} \frac{1}{n!} = e \qquad \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!} = e^{-1} \qquad \sum_{n=0}^{+\infty} \frac{(\ln 2)^n}{n!} = e^{\ln 2} = 2$$

Quelques sommes autour de celle-ci (Savoir les retrouver rapidement, ne pas les apprendre)

$$\sum_{n=1}^{+\infty} \frac{x^n}{n!} = e^x - 1 \quad \sum_{n=1}^{+\infty} \frac{x^n}{(n-1)!} = xe^x \quad \sum_{n=m}^{+\infty} \frac{x^n}{(n-m)!} = x^m e^x \quad \dots$$

### 1.3.3 Séries de Riemann.

#### Théorème.

- ❶ La série  $\sum_{n \geq 1} \frac{1}{n}$  est divergente.
- ❷ La série  $\sum_{n \geq 1} \frac{1}{n^2}$  est convergente. ( Complément : sa somme vaut  $\frac{\pi^2}{6}$  )
- ❸ Complément :  
 Pour  $\alpha \in \mathbb{R}$ ,  
 la série  $\sum_{n \geq 1} \frac{1}{n^\alpha}$  est convergente si, et seulement si,  $\alpha > 1$

**Démonstration.** (Rappel rapide, sinon revoir la feuille\_cours\_1)

Pour ❶ on utilise pour  $k \geq 1$  :  $\ln(k+1) - \ln(k) \leq \frac{1}{k}$  et pour ❷ on utilise pour  $k \geq 2$  :  $\frac{1}{k^2} \leq \frac{1}{k-1} - \frac{1}{k}$

Pour ❸ dans le cas  $\alpha > 1$ , on utilise pour  $k \geq 2$  :  $\frac{1}{k^\alpha} \leq \frac{1}{\alpha-1} ((k-1)^{1-\alpha} - k^{1-\alpha})$

La démonstration de ❸ faite en classe (on utilise un théorème de convergence pour la première étape) :

- On suppose  $\boxed{\alpha \leq 1}$ ,

$1 - \alpha \geq 0$  donc pour tout  $n \in \mathbb{N}^*$ ,  $\exp((1-\alpha)\ln(n)) \geq 1$  ou encore  $n^{1-\alpha} \geq 1$ ,

On peut alors affirmer que :  $\left\{ \begin{array}{l} \forall n \in \mathbb{N}^*, \quad 0 \leq \frac{1}{n} \leq \frac{1}{n^\alpha} \\ \sum_{n \geq 1} \frac{1}{n} \quad \text{diverge} \end{array} \right.$  ce qui entraîne (théorème de convergence) que :

la série  $\sum_{n \geq 1} \frac{1}{n^\alpha}$  est divergente

- On suppose  $\boxed{\alpha > 1}$ ,

Soit  $k$  un entier  $\geq 2$ ,

la fonction  $x \mapsto \frac{1}{x^\alpha}$  est décroissante sur  $[k-1, k]$  donc  $\forall x \in [k-1, k]$ ,  $\frac{1}{k^\alpha} \leq \frac{1}{x^\alpha}$

et en intégrant sur  $[k-1, k]$  (les fonctions sont continues) il vient :

$$\frac{1}{k^\alpha} \leq \int_{k-1}^k \frac{1}{x^\alpha} dx$$

En sommant pour  $k$  allant de 2 à  $n$  on obtient avec la relation de Chasles :

$$S_n = \sum_{k=2}^n \frac{1}{k^\alpha} \leq \int_1^n \frac{1}{x^\alpha} dx = \left[ \frac{x^{1-\alpha}}{1-\alpha} \right]_1^n = \frac{1}{\alpha-1} - \frac{x^{1-\alpha}}{\alpha-1} \leq \frac{1}{\alpha-1}$$

Donc la suite  $(S_n)$  est majorée.

de plus  $S_{n+1} - S_n = \frac{1}{(n+1)^\alpha} \geq 0$  donc  $(S_n)$  est croissante.

La suite  $(S_n)$  est majorée et croissante donc elle converge et ainsi :

la série  $\sum_{n \geq 1} \frac{1}{n^\alpha}$  est convergente

**Remarques :**

- La série  $\sum_{n \geq 1} \frac{1}{n}$  est appelée *série harmonique*.

- $\sum_{k=1}^n \frac{1}{k} \xrightarrow{n \rightarrow +\infty} +\infty$

*Non seulement on sait que la série diverge mais aussi que la suite des sommes partielles tend vers  $+\infty$ .*

- Complément :  $\sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$

- Ne pas confondre les sommes de Riemann, les séries de Riemann et les intégrales de Riemann.

## 1.4 Théorèmes.

$m$  désigne ici un entier naturel quelconque.

### 1.4.1 Théorème de convergence par comparaison des termes généraux positifs.

**Théorème**

Soient  $(u_n)_{n \geq m}$  et  $(v_n)_{n \geq m}$  deux suites de réels,

❶ Si  $\forall n \geq m, 0 \leq u_n \leq v_n$  et  $\sum v_n$  est convergente alors la série  $\sum u_n$  est convergente

❷ Si  $\forall n \geq m, 0 \leq u_n \leq v_n$  et  $\sum u_n$  est divergente alors la série  $\sum v_n$  est divergente

*Attention (erreur courante) : Trop nombreux sont ceux qui passent à la somme sur l'encadrement  $0 \leq u_n \leq v_n$ .*

**Démonstration :** (Faites au tableau)

Il suffit de faire l'une des deux démonstrations :

En effet, lorsque :  $\forall n \geq N, 0 \leq u_n \leq v_n$ , les deux implications suivantes sont la contraposée l'une de l'autre :

❶ Si  $\sum v_n$  CV alors  $\sum u_n$  CV      et      ❷ Si  $\sum u_n$  DV alors  $\sum v_n$  DV.

Montrons le ❶ de ce théorème :

On suppose que  $\forall n \geq m, 0 \leq u_n \leq v_n$  et que  $\sum_{n \geq m} v_n$  converge.

On note  $S_n = \sum_{k=m}^n u_k$  et  $S'_n = \sum_{k=m}^n v_k$ ,

- Pour tout  $n \in \mathbb{N} : S_n - S_{n-1} = u_n \geq 0$  et  $S'_n - S'_{n-1} = v_n \geq 0$  donc les suites  $(S_n)$  et  $(S'_n)$  sont croissantes.

- $\sum_{n \geq m} v_n$  converge, donc  $(S'_n)$  converge et elle est donc majorée. On note  $M$  un réel vérifiant  $\forall n \geq m, S'_n \leq M$ .

- $\forall k \geq m, u_k \leq v_k$  donc (en sommant pour  $k$  allant de  $m$  à un entier  $n$ ) pour tout  $n, S_n \leq S'_n$

et ainsi  $\forall n \geq m, S_n \leq M$

- On a montré que la suite  $(S_n)$  est croissante et majorée (par  $M$ ), donc elle converge.

*(Théorème de convergence monotone)*

En conclusion :  $(S_n)$  est convergente,      autrement dit la série  $\sum u_n$  converge.

**Corollaire**

Soient  $(u_n)_{n \geq m}$  et  $(v_n)_{n \geq m}$  deux suites de réels,  
 Si  $\forall n \geq m \quad 0 \leq u_n \leq v_n$  et  $\sum v_n$  est convergente  
 alors la série  $\sum u_n$  est convergente et  $\sum_{k=m}^{+\infty} u_k \leq \sum_{k=m}^{+\infty} v_k$

**En effet :** Comme les deux séries convergent, on peut passer à la limite sur les sommes partielles.

### 1.4.2 Théorème de convergence par équivalence des termes généraux positifs.

#### Théorème

Soient  $(u_n)_{n \geq m}$  et  $(v_n)_{n \geq m}$  deux suites de réels,

- ❶ Si  $u_n \sim v_n$  et si  $\forall n \geq m, v_n \geq 0$  et si  $\sum v_n$  est convergente,  
 alors la série  $\sum u_n$  est convergente
- ❷ Si  $u_n \sim v_n$  et si  $\forall n \geq m, v_n \geq 0$  et si  $\sum v_n$  est divergente,  
 alors la série  $\sum u_n$  est divergente

*Attention : Trop nombreux sont ceux qui oublient  $v_n \geq 0$ .*

#### Démonstration :

- $u_n \sim v_n$  donc il existe  $(t_n)$  telle que  $\forall n \geq m, v_n = u_n t_n$  et  $\lim_{n \rightarrow +\infty} t_n = 1$
- $\lim_{n \rightarrow +\infty} t_n = 1$  donc il existe  $N \geq m$  tel que  $\forall n \geq N, \frac{1}{2} \leq t_n \leq \frac{3}{2}$  et en multipliant par  $u_n \geq 0$  on obtient :  

$$\forall n \geq N, \quad \frac{u_n}{2} \leq v_n \leq \frac{3u_n}{2}$$

D'une part : Si  $\sum u_n$  converge

on a alors  $\forall n \geq N, \quad 0 \leq v_n \leq \frac{3u_n}{2}$  et la série  $\sum \frac{3u_n}{2}$  converge

donc (d'après le théorème de 1.4.1) :  $\sum v_n$  converge

D'autre part : Si  $\sum u_n$  diverge

on a alors  $\forall n \geq N, \quad 0 \leq \frac{u_n}{2} \leq v_n$  et la série  $\sum \frac{u_n}{2}$  diverge

donc (d'après le théorème de 1.4.1) :  $\sum v_n$  diverge

En conclusion :  $\sum_{n \geq m} u_n$  converge si, et seulement si,  $\sum_{n \geq m} v_n$  converge.

#### Une autre version de ce théorème

#### Théorème

Soient  $(u_n)_{n \geq m}$  et  $(v_n)_{n \geq m}$  deux suites de nombres réels,  
 Si  $u_n \sim v_n$  et si  $\forall n \geq m, v_n \geq 0$  alors  $\sum u_n$  et  $\sum v_n$  sont de même nature.

Remarque :  $\sum u_n$  et  $\sum v_n$  sont de même nature signifie :

$\sum u_n$  converge si, et seulement si,  $\sum v_n$  converge.

## 1.5 Absolue convergence.

### 1.5.1 Théorème.

#### Définition

Soit  $(u_n)_{n \geq m}$  une suite de réels,  
 Dire que la série  $\sum_{n \geq m} u_n$  est **absolument convergente** signifie que la série  $\sum_{n \geq m} |u_n|$  est convergente.

**Théorème.** (La convergence absolue entraîne la convergence)

Soit  $(u_n)_{n \geq m}$  une suite de réels,  
 Si la série  $\sum u_n$  est absolument convergente  
 alors elle est convergente et  $\left| \sum_{k=m}^{+\infty} u_k \right| \leq \sum_{k=m}^{+\infty} |u_k|$

#### Démonstration :

On suppose que  $\sum u_n$  est absolument convergente,

On définit les deux suites  $(u_n^+)$  et  $(u_n^-)$  par :

$$u_n^+ = \max(u_n, 0) \quad \text{et} \quad u_n^- = \min(u_n, 0)$$

on remarque que :  $\forall n \in \mathbb{N}, \quad u_n = u_n^+ + u_n^-$

$$\begin{array}{l} \text{or} \quad \left\{ \begin{array}{l} \text{Pour tout } n \geq m, \quad 0 \leq u_n^+ \leq |u_n| \\ \text{et } \sum |u_n| \text{ converge} \end{array} \right. \quad \text{donc } \sum u_n^+ \text{ converge} \\ \text{et} \quad \left\{ \begin{array}{l} \text{Pour tout } n \geq m, \quad 0 \leq -u_n^- \leq |u_n| \\ \text{et } \sum |u_n| \text{ converge} \end{array} \right. \quad \text{donc } \sum -u_n^- \text{ converge et ainsi } \sum u_n^- \text{ converge} \end{array}$$

donc (linéarité des séries convergentes)  $\sum u_n$  converge.

On a bien démontré que :

si  $\sum u_n$  est absolument convergente alors  $\sum u_n$  converge.

**Attention** la réciproque est fautive.

Contre-exemple : La série  $\sum_{n \geq 1} \frac{(-1)^n}{n}$  est convergente, mais pas absolument convergente.

### 1.5.2 Convergence commutative.

#### Théorème.

Soit  $(u_n)_{n \in \mathbb{N}}$  une suite de réels et  $\sigma$  une bijection de  $\mathbb{N}$  dans  $\mathbb{N}$ ,  
 Si la série  $\sum u_n$  est absolument convergente alors  $\sum u_{\sigma(n)}$  est convergente  
 et  $\sum_{k=0}^{+\infty} u_{\sigma(k)} = \sum_{k=0}^{+\infty} u_k$ .

Ce résultat est admis et nous servira dans le chapitre "Probabilité".

#### Autrement dit :

- La valeur de la somme d'une série absolument convergente ne dépend pas de l'ordre d'énumération de ses termes.
- Si la série est absolument convergente alors elle est commutativement convergente.

# Polynômes

## Plan du chapitre

---

2.1	Généralités. . . . .	14
2.1.1	Introduction et notation. . . . .	14
2.1.2	Vocabulaire. . . . .	15
2.1.3	Opérations. . . . .	15
2.1.4	Degré. . . . .	16
2.1.5	Intégrité . . . . .	17
2.2	Racines. . . . .	17
2.2.1	Racine et factorisation . . . . .	17
2.2.2	Racine multiple. Multiplicité. . . . .	18
2.2.3	Racine complexe d'un polynôme à coefficients réels . . . . .	19
2.2.4	Factorisation dans $\mathbb{C}[X]$ . . . . .	19
2.3	Dérivation (complément). . . . .	20
2.3.1	polynôme dérivé. . . . .	20
2.3.2	Dérivées successives. . . . .	20

---

## 2.1 Généralités.

Voir aussi le cours : *Fonctions polynomiales réelles.*

### 2.1.1 Introduction et notation.

Un polynôme est entièrement déterminé par la liste de ses coefficients (des nombres complexes). Seul un nombre fini de ses coefficients sont non nuls.

*La définition la plus simple : Un polynôme est une suite de nombres complexes nuls à partir d'un certain rang.*

- On note  $\mathbb{C}[X]$  l'ensemble des polynômes et  $\mathbb{R}[X]$  l'ensemble des polynômes à coefficients réels.
- On a l'inclusion :  $\mathbb{R}[X] \subset \mathbb{C}[X]$ .
- Un polynôme est nul si, et seulement si, tous ses coefficients sont nuls.
- Deux polynômes sont égaux si, et seulement s'ils ont les mêmes coefficients.

On écrira les polynômes comme une somme de monômes :  $P = \sum_{k=0}^n a_k X^k$

Pour comprendre les opérations on confondra les polynômes et les fonctions polynomiales de  $\mathbb{C}$  dans  $\mathbb{C}$ .

$$\text{on aura alors } X^k : \mathbb{C} \longrightarrow \mathbb{C} \quad \text{et} \quad \sum_{k=0}^n a_k X^k : \mathbb{C} \longrightarrow \mathbb{C}$$

$$z \longmapsto z^k \qquad \qquad \qquad z \longmapsto \sum_{k=0}^n a_k z^k$$

Les règles de calcul sur les polynômes sont les mêmes que sur les fonctions polynomiales.

On écrira indifféremment :  $P = \sum_{k=0}^n a_k X^k$  et  $P(X) = \sum_{k=0}^n a_k X^k$ .

## 2.1.2 Vocabulaire.

Pour  $P$  le polynôme  $\sum_{k=0}^n a_k X^k$  avec  $a_n \neq 0$  (un polynôme non nul)

On dit que le polynôme  $P$  est sous sa **forme développée réduite**.

- $a_n X^n$  est le monôme dominant de  $P$ .
- $n$  le degré de  $P$ .
- $a_n$  le coefficient dominant.
- $a_k X^k$  est le monôme d'ordre  $k$  de  $P$ . (ou de degré  $k$ )
- $a_k$  le coefficient d'ordre  $k$  de  $P$ . (ou de degré  $k$ )
- $a_0$  est le terme constant.

Un polynôme est une somme finie de monômes.

Dire qu'un polynôme est **unitaire**, signifie que son coefficient dominant vaut 1.

Dire qu'un polynôme est **constant** signifie qu'il est nul ou que son degré vaut 0.

## 2.1.3 Opérations.

**Propriétés :** (règles de calcul)

Soient  $P, Q$  et  $R$  trois polynômes, on a :

1.  $0 + P = P, \quad P + Q = Q + P, \quad (P + Q) + R = P + (Q + R)$
2.  $P \times 0 = 0, \quad P \times 1 = P, \quad P \times Q = Q \times P, \quad P \times (Q \times R) = (P \times Q) \times R$
3.  $P \times (Q + R) = P \times Q + P \times R$

- **Multiplication par un scalaire.** ( $\alpha \in \mathbb{C}$ )

$$\alpha \sum_{k=0}^n a_k X^k = \sum_{k=0}^n (\alpha a_k) X^k$$

- **Somme de polynômes.**

Pour  $m \leq n$ ,

$$\sum_{k=0}^n a_k X^k + \sum_{k=0}^m b_k X^k = \sum_{k=0}^m (a_k + b_k) X^k + \sum_{k=m+1}^n a_k X^k$$

**Remarque :**

Avec les deux opérations précédentes, **l'ensemble des polynômes est un espace vectoriel.**

- **Produit de polynômes.**

• **Produit de deux monômes :**  $aX^n \times bX^m = abX^{n+m}$

•  $\sum_{k=0}^n a_k X^k \times \sum_{k=0}^m b_k X^k = \sum_{k=0}^{n+m} c_k X^k$  avec  $c_k = \sum_{i=0}^k a_i b_{k-i}$  ou  $c_k = \sum_{i+j=k} a_i b_j$

**Remarques :**

- Lorsque  $a_n \neq 0$  et  $b_m \neq 0$ , le monôme dominant de  $\sum_{k=0}^n a_k X^k \times \sum_{k=0}^m b_k X^k$  est  $a_n b_m X^{n+m}$ .

- Le terme constant de  $\sum_{k=0}^n a_k X^k \times \sum_{k=0}^m b_k X^k$  est  $a_0 b_0$ .

- Dans la somme  $\sum_{i=0}^{n+m} a_i b_{n+m-i}$  le seul terme non nul est celui d'indice  $i = n$  donc  $c_{n+m} = a_n b_m$

• **Composée de polynômes.**

Pour  $P(X) = \sum_{k=0}^n a_k X^k$  et  $Q(X)$  deux polynômes,  $P \circ Q$  est le polynôme  $P \circ Q(X) = \sum_{k=0}^n a_k Q(X)^k$

**Remarques :**

- En général  $P \circ Q \neq Q \circ P$ .
- Si  $Q$  n'est pas constant et  $P(X) = \sum_{k=0}^n a_k X^k$  avec  $a_n \neq 0$  alors

le monôme dominant de  $P \circ Q$  est celui de  $a_n Q(X)^n$

Et si  $Q(X) = b_0 + \dots + b_m X^m$  alors le monôme dominant de  $Q(X)^n = b_m^n X^{nm}$

**A retenir :**

L'ensemble des polynômes est stable par combinaison linéaire, par produit et par composée.

**Attention :** Dans ce cours on ne fait pas de quotient de polynômes.

**2.1.4 Degré.**

**Définition :**

Soit  $P \in \mathbb{C}[X]$ ,

- ❶ si  $P = 0$  alors  $\deg(P) = -\infty$
- ❷ si  $P = a_0 + a_1 X + \dots + a_n X^n$  avec  $a_n \neq 0$  alors  $\deg(P) = n$

**Remarques :**

- $\deg(P) = \max\{i \mid a_i \neq 0\}$  (on prend comme convention  $\max(\emptyset) = -\infty$ )
- Pour  $P$  non nul définie par la suite de ses coefficients :  
le degré de  $P$  est l'indice de son dernier coefficient non nul.

**Théorème :**

Soient  $P$  et  $Q$  deux polynômes.

- ❶  $\deg(P + Q) \leq \max\{\deg(P); \deg(Q)\}$
- ❷ Si  $\deg(P) < \deg(Q)$  alors  $\deg(P + Q) = \deg(Q)$
- ❸  $\deg(P \times Q) = \deg(P) + \deg(Q)$
- ❹ Pour tout  $m \in \mathbb{N}^*$ ,  $\deg(P^m) = m \deg(P)$
- ❺ (complément)  
si  $Q$  n'est pas constant alors  $\deg(P \circ Q) = \deg(P) \times \deg(Q)$

Remarque : On accepte ici des calculs avec  $-\infty$  lorsque un polynôme nul.

**En effet :** (lorsque  $P$  et  $Q \neq 0$ ) on pose  $P(X) = \sum_{k=0}^n a_k X^k$  et  $Q(X) = \sum_{k=0}^m b_k X^k \dots$

**Notation :** pour  $n$  un entier naturel non nul,

$\mathbb{R}_n[X]$  l'ensemble des polynômes à coefficients réels de degré inférieur ou égal à  $n$ .  
 $\mathbb{C}_n[X]$  l'ensemble des polynômes à coefficients complexes de degré inférieur ou égal à  $n$ .

**Remarque.**

$\mathbb{R}_n[X]$  est un espace vectoriel sur  $\mathbb{R}$  de dimension  $n + 1$ .  
 $\mathbb{C}_n[X]$  est un espace vectoriel sur  $\mathbb{C}$  de dimension  $n + 1$ .

### 2.1.5 Intégrité

**Théorème** : *Intégrité.*

$$\begin{aligned} &\text{Soient } P \text{ et } Q \text{ deux polynômes de } \mathbb{C}[X], \\ &P \times Q = 0 \iff P = 0 \quad \text{ou} \quad Q = 0 \end{aligned}$$

**Démonstration** :

- $\boxed{\Leftarrow}$  Propriété du produit : si  $P$  ou  $Q$  est nul alors  $PQ = 0$ .
- $\boxed{\Leftarrow}$  (*Montrons sa contraposée*) On suppose que  $P \neq 0$  et  $Q \neq 0$ .  
on peut alors écrire  $P = \sum_{k=0}^n a_k X^k$  avec  $a_n \neq 0$  et  $Q = \sum_{k=0}^m b_k X^k$  avec  $b_m \neq 0$   
on a alors  $PQ = \sum_{k=0}^{n+m} c_k X^k$  avec  $c_{n+m} = a_n b_m \neq 0$  donc  $PQ \neq 0$ .

**Corollaires** :

- ❶ Soient  $P$  et  $Q$  deux polynômes de  $\mathbb{C}[X]$ ,  
 $P \times Q \neq 0 \iff P \neq 0 \quad \text{et} \quad Q \neq 0$
- ❷ Soient  $P, Q$  et  $R$  trois polynômes de  $\mathbb{C}[X]$ ,  
Lorsque  $Q \neq 0$ , on a l'équivalence :  $QP = QR \iff P = R$ .
- ❸ Soient  $P_1, P_2, \dots, P_n$  des polynômes :  
 $P_1 P_2 P_3 \cdots P_n = 0 \iff \exists i \in \llbracket 1; n \rrbracket, P_i = 0$

**Autrement dit** : Un produit de polynômes est nul si, et seulement si, un de ses facteurs est nul.

**En effet** : pour ❷  $QP = QR \iff QP - QR = 0 \iff Q(P - R) = 0 \underbrace{\iff}_{Q \neq 0} P - R = 0 \iff P = R$ .

## 2.2 Racines.

### 2.2.1 Racine et factorisation

**Définition** (*Racine de  $P$* ) :

$$\begin{aligned} &\text{Soient } P \in \mathbb{C}[X] \text{ et } \alpha \in \mathbb{C}, \\ &\text{Dire que } \alpha \text{ est une racine de } P \text{ signifie que : } P(\alpha) = 0 \end{aligned}$$

**Théorème** :

$$\begin{aligned} &\text{Soient } P \in \mathbb{C}[X] \text{ et } \alpha \in \mathbb{C}, \\ &P(\alpha) = 0 \quad \text{si, et seulement si,} \quad \text{il existe un polynôme } Q \text{ tel que } P = (X - \alpha)Q \end{aligned}$$

**Démonstration** :

- ❶ Pour chaque entier  $k$ , en posant  $Q_k(X) = \sum_{i=0}^{k-1} \alpha^{n-1-k} X^i$  on a :  $X^k - \alpha^k = (X - \alpha)Q_k(X)$   
(*Formule de Bernoulli*)
- ❷ Si  $\alpha$  est une racine de  $P = \sum_{k=0}^n a_k X^k$ ,

$$\begin{aligned} P(X) &= P(X) - P(\alpha) \\ &= \sum_{k=0}^n a_k X^k - \sum_{k=0}^n a_k \alpha^k \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=0}^n a_k (X^k - \alpha^k) \\
&= \sum_{k=0}^n a_k (X - \alpha) Q_k(X) \quad (\text{définis en } \bullet) \\
&= (X - \alpha) \sum_{k=0}^n a_k Q_k(X)
\end{aligned}$$

Finalement, en posant  $Q(X) = \sum_{k=0}^n a_k Q_k(X)$  on a bien  $Q(X) \in \mathbb{C}[X]$  et  $P(X) = (X - \alpha)Q(X)$

**Théorème :**

Soient  $P$  un polynôme et  $k$  un entier naturel non nul,

Si  $\alpha_1, \alpha_2, \dots, \alpha_k$  sont  $k$  **racines distinctes** de  $P$  alors  $P$  est factorisable par :

$$(X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_k)$$

ou encore  $(X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_k)$  divise  $P$

**Démonstration :** (*non faites*)

**Conséquences :**

- Un polynôme de degré  $n \in \mathbb{N}$  a au plus  $n$  racines distinctes.
- Le nombre de racines distinctes d'un polynôme non nul est majoré par son degré.
- Un polynôme de  $\mathbb{C}_n[X]$  ayant plus de  $n + 1$  racines est nul.
- Si un polynôme a une infinité de racines alors c'est le polynôme nul.

### 2.2.2 Racine multiple. Multiplicité.

**Définition :** *Racine multiple.*

Soient  $a \in \mathbb{C}$  et  $P$  un polynôme de  $\mathbb{C}[X]$

Dire que  $\alpha$  est une racine multiple de  $P$ ,

signifie qu'il existe un polynôme  $Q$  tel que :  $P(X) = (X - \alpha)^2 Q(X)$ .

**Définition :** *Multiplicité.*

Soient  $a \in \mathbb{C}$ ,  $P$  un polynôme de  $\mathbb{C}[X]$  et  $m$  un entier naturel non nul.

Dire que  $\alpha$  est une racine de  $P$  d'ordre de multiplicité  $m$

signifie qu'il existe un polynôme  $Q$  tel que :  $P(X) = (X - \alpha)^m Q(X)$  et  $Q(\alpha) \neq 0$ .

**Théorème :**

Soit  $P$  un polynôme de  $\mathbb{C}[X]$ ,  $r$  un entier naturel non nul.

Si  $\alpha_1, \alpha_2, \dots, \alpha_r$  sont des racines distincts de  $P$  d'ordre respectif  $m_1, m_2, \dots, m_r$  alors

il existe  $Q \in \mathbb{C}[X]$  tel que  $P(X) = \left( \prod_{i=1}^r (X - \alpha_i)^{m_i} \right) Q(X)$  et  $\forall i \in \llbracket 1; r \rrbracket, Q(\alpha_i) \neq 0$ .

**Démonstration :** (*non faites*)

**Conséquence :**

Un polynôme de degré  $n$  ( $n \geq 0$ ) possède au plus  $n$  racines comptées avec leur ordre de multiplicité.

### 2.2.3 Racine complexe d'un polynôme à coefficients réels

**Théorème :**

Soient  $P(X)$  un polynôme et  $\alpha$  un nombre complexe.  
Si  $P \in \mathbb{R}[X]$  et si  $\alpha$  est une racine de  $P$  alors  $\bar{\alpha}$  est une racine de  $P$ .

**Démonstration :**

Soient  $P = \sum_{k=0}^n a_k x^k$  un polynôme et  $\alpha$  un nombre complexe.

On suppose que  $P$  est à coefficients réels et que  $P(\alpha) = 0$ ,

$$\begin{aligned}
 P(\bar{\alpha}) &= \sum_{k=0}^n a_k (\bar{\alpha})^k \\
 &= \sum_{k=0}^n a_k \overline{\alpha^k} && \overline{z^n} = \bar{z}^n \\
 &= \sum_{k=0}^n \overline{a_k \alpha^k} && \text{car } a_k \in \mathbb{R} \text{ et } \overline{zz'} = \bar{z} \bar{z}' \\
 &= \overline{\sum_{k=0}^n a_k \alpha^k} && \overline{z + z'} = \bar{z} + \bar{z}' \\
 &= \overline{P(\alpha)} \\
 &= 0 && \text{car } P(\alpha) = 0
 \end{aligned}$$

si  $P \in \mathbb{R}[X]$  et si  $\alpha$  est une racine de  $P$  alors  $\bar{\alpha}$  est une racine de  $P$

### 2.2.4 Factorisation dans $\mathbb{C}[X]$ .

**Théorème : (admis) (d'Alembert-Gauss)**

Tout polynôme non constant admet une racine dans  $\mathbb{C}$ .

**Théorème :**

Soit  $P$  un polynôme de  $\mathbb{C}[X]$  tel que  $\deg(P) = n$  avec  $n \in \mathbb{N}^*$ ,

❶ Il existe des complexes  $\lambda, \alpha_1, \alpha_2, \dots, \alpha_n$  tels que :

$$P(X) = \lambda(X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_n)$$

❷ Il existe un entier  $p \leq n$ , des complexes  $\lambda, r_1, r_2, \dots, r_p$  et des entiers non nuls  $m_1, \dots, m_p$  tels que :

$$r_1, r_2, \dots, r_p \text{ sont distinctes et } P(X) = \lambda(X - r_1)^{m_1}(X - r_2)^{m_2} \cdots (X - r_p)^{m_p}$$

**Remarques :**

- Ces écritures sont les formes factorisées de  $P$ .
- On dit aussi que  $P$  est décomposé en produit de facteurs irréductibles.
- Dans  $\mathbb{C}[X]$  les polynômes irréductibles sont les polynômes de degré 1.  
(Ils admettent comme diviseurs que les constantes non nulles et eux-mêmes.)
- Tout polynôme de  $\mathbb{C}[X]$  de degré  $n$ , possède exactement  $n$  racines comptées avec leur ordre de multiplicité.
- Attention : ces théorèmes sont faux quand on se restreint à des polynômes de  $\mathbb{R}[X]$ .

## 2.3 Dérivation (complément).

### 2.3.1 polynôme dérivé.

Définition :

Soit  $P(X) = \sum_{k=0}^n a_k X^k$  un polynôme de  $\mathbb{C}[X]$ ,

on appelle polynôme dérivé de  $P(X)$  le polynôme :

$$P'(X) = \sum_{k=1}^n k a_k X^{k-1} \quad \text{ou} \quad P'(X) = \sum_{k=0}^{n-1} (k+1) a_{k+1} X^k$$

Remarques :

- Ici pas besoin de justifier une dérivabilité comme pour les fonctions.
- La définition est cohérente avec le cours sur la dérivation des fonctions de  $\mathbb{R} \rightarrow \mathbb{R}$ .
- $P(X)$  est constant équivaut à  $P'(X) = 0$   
(Attention avec les fonctions on a en plus besoin de **dérivable** sur un **intervalle**).

Propriétés :

$(\lambda, \mu) \in \mathbb{C}^2$  et  $P(X), Q(X)$  deux polynômes de  $\mathbb{C}[X]$ ,

- $(\lambda P + \mu Q)'(X) = \lambda P'(X) + \mu Q'(X)$
- $(PQ)'(X) = P'(X)Q(X) + P(X)Q'(X)$
- $(P^m)'(X) = mP'(X)P^{m-1}(X)$
- $(Q \circ P)'(X) = P'(X).Q'(P(X))$

Dérivation et degré.

Soit  $P(X)$  un polynôme non constant ( $\deg P \geq 1$ ),

$$\deg(P') = \deg(P) - 1$$

### 2.3.2 Dérivées successives.

Définition :

On définit, par récurrence sur  $k$ , les polynômes dérivés successifs de  $P$  :

$$\begin{cases} P^{(0)}(X) = P(X) \\ \forall k \in \mathbb{N}, \quad P^{(k+1)}(X) = (P^{(k)})'(X) \end{cases}$$

Propriété :

$$(\lambda, \mu) \in \mathbb{C}^2, m \text{ un entier et } P(X), Q(X) \text{ deux polynômes de } \mathbb{C}[X],$$

$$(\lambda P + \mu Q)^{(m)}(X) = \lambda P^{(m)}(X) + \mu Q^{(m)}(X)$$

Proposition :

Le dérivé  $i$ -ième du monôme  $a_k X^k$  est égale :

$$\begin{aligned} &\text{\`a} \quad 0 \quad \text{si } i > k \\ &\text{\`a} \quad a_k \times \frac{k!}{(k-i)!} X^{k-i} \quad \text{si } 0 \leq i \leq k \end{aligned}$$

**Remarques :**

- la dérivée  $k$ -ième de  $a_k X^k$  est le polynôme constant  $a_k \times k!$

- le monôme dominant de  $P^{(k)}(X)$  est :  $a_n \times \frac{n!}{(n-k)!} X^{n-k}$

- le terme constant de  $P^{(k)}(X)$  est égal à :  $a_k \times k!$ , ou encore

$$a_k = \frac{P^{(k)}(0)}{k!}$$

- si  $k > n$  et  $\deg(P(X)) = n$  alors  $P^{(k)}(X) = 0$

---

# Espaces vectoriels

## Plan du chapitre

---

3.1	Structure vectorielle. . . . .	<b>23</b>
3.1.1	Définition . . . . .	23
3.1.2	Sommes de $n$ vecteurs. . . . .	23
3.1.3	Exemples de références. . . . .	24
3.2	Sous-espace vectoriel . . . . .	<b>24</b>
3.2.1	Définition. . . . .	24
3.2.2	Intersection de sous-espaces vectoriels . . . . .	25
3.3	$\text{Vect}(u_1, \dots, u_n)$ . . . . .	<b>25</b>
3.3.1	Combinaisons linéaires. . . . .	25
3.3.2	Définitions - notations. . . . .	26
3.3.3	C'est un sous-espace vectoriel. . . . .	26
3.3.4	Opérations élémentaires. . . . .	26
3.3.5	Familles génératrices. . . . .	27
3.3.6	Espace vectoriel de dimension finie. . . . .	27
3.3.7	Bases canoniques. . . . .	27
3.4	Familles libres. . . . .	<b>28</b>
3.4.1	Définition. . . . .	28
3.4.2	Identification. . . . .	28
3.4.3	Familles de polynômes. . . . .	28
3.5	Bases. . . . .	<b>29</b>
3.5.1	Définition d'une base. . . . .	29
3.5.2	Caractérisation d'une base. . . . .	29
3.6	Coordonnées dans une base. . . . .	<b>29</b>
3.6.1	Définition. . . . .	29
3.6.2	Application $v \mapsto \text{Coord}_{\mathcal{B}}(v)$ . . . . .	30
3.6.3	Matrice d'une famille de vecteurs dans une base. . . . .	30
3.7	Dimension d'un espace vectoriel . . . . .	<b>31</b>
3.7.1	Définition. . . . .	31
3.7.2	Dans une famille libre chaque vecteur donne une nouvelle direction. . . . .	32
3.7.3	Compléter une famille libre ou extraire d'une famille génératrice. . . . .	33
3.8	Famille libre, famille génératrice et dimension. . . . .	<b>33</b>
3.8.1	Famille génératrice en dimension $n$ . . . . .	33
3.8.2	Famille libre en dimension $n$ . . . . .	34
3.8.3	Dimension et sous-espace vectoriel. . . . .	34
3.9	Systèmes linéaires homogènes et dimension. . . . .	<b>34</b>
3.10	Rang. . . . .	<b>35</b>
3.10.1	Rang d'une famille de vecteurs. . . . .	35
3.10.2	Rang d'une matrice. . . . .	36
3.10.3	Matrices et familles de vecteurs. . . . .	36
3.10.4	Rang d'un système. . . . .	37

---

### 3.1 Structure vectorielle.

On travaille avec des nombres réels ou des nombres complexes ; pour simplifier on écrira :  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ .  
Les lettres  $n, m, p$  et  $r$  désignent dans tout ce chapitre des entiers naturels non nuls.

#### 3.1.1 Définition

**Définition** (*Compléments*)

On note  $E$  un ensemble muni de deux opérations : une loi de composition interne  $+$  et une loi de composition externe  $\cdot$  (définie sur  $\mathbb{K} \times E$ ).

Dire que  $(E, +, \cdot)$  est un  $\mathbb{K}$ -espace vectoriel signifie que les deux lois ont les propriétés suivantes :

- a<sub>1</sub>)** pour tout  $u$  et  $v$  de  $E$ ,  $u + v = v + u$
- a<sub>2</sub>)** pour tout  $u, v, w$  des éléments de  $E$ ,  $(u + v) + w = u + (v + w)$
- a<sub>3</sub>)** il existe  $0_E$  un élément de  $E$  tel que pour tout  $u \in E$ ,  $u + 0_E = u$
- a<sub>4</sub>)** pour tout  $u$  de  $E$ , il existe un  $u'$  dans  $E$  tel que  $u + u' = 0_E$
- b<sub>1</sub>)** pour tout  $u$  de  $E$   $1.u = u$
- b<sub>2</sub>)** pour tout  $u$  de  $E$  pour tout  $\alpha$  et  $\beta$  de  $\mathbb{K}$ ,  $\alpha.(\beta.u) = (\alpha\beta).u$
- b<sub>3</sub>)** pour tout  $u$  de  $E$  pour tout  $\alpha$  et  $\beta$  de  $\mathbb{K}$ ,  $(\alpha + \beta).u = \alpha.u + \beta.u$
- b<sub>4</sub>)** pour tout  $u$  et  $v$  de  $E$ , pour tout  $\alpha$  de  $\mathbb{K}$ ,  $\alpha.(u + v) = \alpha.u + \alpha.v$

**Remarques :** ① Les éléments d'un espace vectoriel sont appelés **vecteurs**.

② Les éléments de  $\mathbb{K}$ , (*les nombres*), sont appelés **scalaires**.

**Propositions**

- ①  $0_E$  est unique et est nommé : vecteur nul de  $E$ .
- ② Pour tout  $v \in E$  et  $k \in \mathbb{K}$ ,  $kv = 0_E$  si, et seulement si,  $k = 0$  ou  $v = 0_E$ .
- ③  $u'$  est unique et on le note  $-u$ .
- ④  $\forall v \in E$ ,  $-v = (-1)v$ .
- ⑤  $\forall v \in E$  et  $\forall k \in \mathbb{K}$ ,  $(-k)v = k(-v) = -(kv)$ .

**Démonstration.**

(On peut démontrer ①, ②, ③, ④ et ⑤ avec les  $a_i$ ) et  $b_i$ ), mais on rencontre rarement ce type de raisonnement en BCPST)

**Théorème.**

Si  $E$  est un espace vectoriel alors :

- ①  $\forall (\alpha, \beta) \in \mathbb{K}^2$ ,  $\forall (u, v) \in E^2$ ,  $\alpha u + \beta v \in E$
- ②  $\forall (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ ,  $\forall (u_1, \dots, u_n) \in E^n$ ,  $\sum_{k=1}^n \lambda_k u_k \in E$

**En effet :** Cela vient de l'existence des deux lois.

#### 3.1.2 Sommes de $n$ vecteurs.

**Définition.**

Soit  $(u_n)$  une suite de vecteurs de  $E$ , on définit

$$\sum_{k=0}^0 u_k = u_0 \quad \text{et} \quad \forall n \in \mathbb{N}, \quad \sum_{k=0}^{n+1} u_k = \left( \sum_{k=0}^n u_k \right) + u_{n+1}$$

Plus simplement on pourra écrire :  $\sum_{k=0}^n u_k = \underbrace{u_0 + \dots + u_n}_{n+1 \text{ vecteurs}}$

**Remarque :** on ne fait aucune somme d'un nombre infini de vecteurs.

**Proposition.**

Soient  $n$  un entier naturel non nul et  $(u_1, \dots, u_n) \in E^n$ ,

- ❶ On ne change pas la somme des vecteurs en modifiant l'ordre des vecteurs.

Autrement dit : Si  $\sigma$  est une bijection de  $\llbracket 1, n \rrbracket$  dans  $\llbracket 1, n \rrbracket$  alors 
$$\sum_{k=1}^n u_{\sigma(k)} = \sum_{k=1}^n u_k$$

- ❷ Pour  $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$  et  $a \in \mathbb{K}$ , 
$$a \left( \sum_{k=1}^n \alpha_k u_k \right) = \sum_{k=1}^n (a\alpha_k) u_k$$
- ❸ Pour  $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$  et  $(\beta_1, \dots, \beta_n) \in \mathbb{K}^n$ , 
$$\left( \sum_{k=1}^n \alpha_k u_k \right) + \left( \sum_{k=1}^n \beta_k u_k \right) = \sum_{k=1}^n (\alpha_k + \beta_k) u_k$$

**Remarques :**

- Les points ❷ et ❸ permettront de montrer que  $\text{Vect}(u_1, \dots, u_n)$  est stable par combinaison linéaire.
- On peut faire des changements indices comme sur les sommes de nombres.

### 3.1.3 Exemples de références.

Des  $\mathbb{R}$ -espaces vectoriels :

- $E$  : l'ensemble des vecteurs du plan (resp. de l'espace) ( avec  $0_E = \vec{0}$  )
- $E = \mathbb{R}^n$ , pour  $n$  un entier naturel non nul ( avec  $0_E = (0, 0, \dots, 0)$  )
- $E = \mathcal{M}_{n,m}(\mathbb{R})$ , pour  $n$  et  $m$  deux entiers naturels non nuls ( avec  $0_E$  : la matrice nulle )
- (complément)  $E$  : l'ensemble des suites réelles ( avec  $0_E$  : la suite nulle )
- $E$  : l'ensemble des fonctions de  $I$  dans  $\mathbb{R}$ , pour  $I$  un intervalle de  $\mathbb{R}$  ( avec  $0_E$  : fonction nulle sur  $I$  )
- $E$  : l'ensemble des fonctions de classe  $C^n$  sur  $I$  un intervalle de  $\mathbb{R}$  ( avec  $0_E$  : fonction nulle sur  $I$  )
- $E$  : l'ensemble des fonctions de classe  $C^\infty$  sur  $I$  un intervalle de  $\mathbb{R}$  ( avec  $0_E$  : fonction nulle sur  $I$  )
- $E = \mathbb{R}[X]$  ( avec  $0_E$  : le polynôme nul )
- $E$  : l'ensemble des variables aléatoires à valeurs réelles. ( avec  $0_E$  : la variable certaine égale à 0. )

Des  $\mathbb{C}$ -espaces vectoriels :

- $E = \mathbb{C}^n$ , pour  $n$  un entier naturel non nul ( avec  $0_E = (0, 0, \dots, 0)$  )
- $E = \mathcal{M}_{n,m}(\mathbb{C})$ , pour  $n$  et  $m$  deux entiers naturels non nuls ( avec  $0_E$  : la matrice nulle )
- (complément)  $E$  : l'ensemble des suites à valeurs complexes ( avec  $0_E$  : la suite nulle )
- $E = \mathbb{C}[X]$  ( avec  $0_E$  : le polynôme nul )

On définit les deux opérations suivantes pour  $E$  et  $F$  deux espaces vectoriels :

- Pour  $f$  une application  $E$  dans  $F$  et  $\alpha$  un scalaire on définit l'application : 
$$\alpha f : E \longrightarrow F$$
  
$$u \longmapsto \alpha f(u)$$
- Pour deux applications  $f$  et  $g$  de  $E$  dans  $F$  on définit l'application : 
$$f + g : E \longrightarrow F$$
  
$$u \longmapsto f(u) + g(u)$$

L'ensemble des applications de  $E$  dans  $F$  (noté  $F^E$ ) muni de ces deux lois est un espace vectoriel.

## 3.2 Sous-espace vectoriel

### 3.2.1 Définition.

**Définition :**

Soient  $(E, +, \cdot)$  un  $\mathbb{K}$ -espace vectoriel et  $F$  une partie de  $E$ ,

Dire que  $F$  est un **sous-espace vectoriel** de  $E$  signifie que :

$$(F, +, \cdot)$$
 est un espace vectoriel.

**Caractérisation :**

Soient  $E$  un  $\mathbb{K}$ -espace vectoriel et  $F$  un ensemble

$F$  est un **sous-espace vectoriel** de  $E$  si, et seulement si, :

$$\textcircled{1} F \subset E. \quad \textcircled{2} 0_E \in F. \quad \textcircled{3} \forall (\alpha, \beta) \in \mathbb{K}^2, \quad \forall (u, v) \in F^2, \quad \alpha u + \beta v \in F$$

**Remarques :**

- Certains remplacent  $\textcircled{2}$  par  $F \neq \emptyset$ .
- Certains remplacent  $\textcircled{3}$  par  $\forall \lambda \in \mathbb{K}, \quad \forall (u, v) \in F^2, \quad u + \lambda v \in F$
- D'autres remplacent  $\textcircled{3}$  par  $\forall \alpha \in \mathbb{K}, \forall u \in F, \quad \alpha u \in F.$  et  $\forall (u, v) \in F^2, \quad u + v \in F.$
- Le singleton  $\{0_E\}$  est un sous-espace vectoriel de  $E$ , c'est d'ailleurs le seul qui contient un nombre fini d'éléments. (*Tous les autres contiennent un nombre infini de vecteurs*)

*Des exemples dans la feuille\_Cours\_3*

**3.2.2 Intersection de sous-espaces vectoriels****Théorème :**

Soient  $E$  un  $\mathbb{K}$ -espace vectoriel et  $F_1, F_2$  deux parties de  $E$ .

Si  $F_1$  et  $F_2$  sont deux sous-espaces vectoriels de  $E$

alors  $F_1 \cap F_2$  est un sous-espace vectoriel de  $E$ .

**Démonstration.**

- $F \cap G \subset E$
  - $0_E \in F$  et  $0_E \in G$  donc  $0_E \in F \cap G$ .
  - Soit  $(\alpha, \beta) \in \mathbb{K}^2, (u, v) \in (F \cap G)^2$ ,  
 $F \cap G \subset F$  donc  $u \in F$  et  $v \in F$  or  $F$  est un sous-espace vectoriel donc  $\alpha u + \beta v \in F$   
 $F \cap G \subset G$  donc  $u \in G$  et  $v \in G$  or  $G$  est un sous-espace vectoriel donc  $\alpha u + \beta v \in G$   
donc  $\alpha u + \beta v \in F \cap G$
- En conclusion :  $F \cap G$  est un sous-espace vectoriel de  $E$

**Théorème : (Généralisation)**

Soient  $E$  un  $\mathbb{K}$ -espace vectoriel,  $n \in \mathbb{N}^*$  et  $F_1, \dots, F_n$  des parties de  $E$ .

Si  $F_1, \dots, F_n$  sont des sous-espaces vectoriels de  $E$ ,

alors  $\bigcap_{i=1}^n F_i$  est un sous-espace vectoriel de  $E$ .

**En effet :** *On peut faire exactement le même raisonnement que pour le théorème précédent.*

**Attention :** En général, la réunion de deux sous-espaces vectoriels n'est pas un sous-espace vectoriel.

*Une condition nécessaire et suffisante : "l'un est inclus dans l'autre" (voir la feuille\_Cours\_3)*

**3.3 Vect( $u_1, \dots, u_n$ )****3.3.1 Combinaisons linéaires.**

Dire qu'un vecteur  $v$  de  $E$  est une **combinaison linéaire** des vecteurs  $u_1, \dots, u_n$  de  $E$

signifie qu'il existe  $\lambda_1, \dots, \lambda_n$  des scalaires tels que :  $v = \sum_{i=1}^n \lambda_i u_i$ .

**Remarque :** Un espace vectoriel est stable par combinaisons linéaires.

### 3.3.2 Définitions - notations.

#### Définition.

Soit  $(u_1, \dots, u_n)$  une famille de  $n$  vecteurs de  $E$ ,  
 On note  $\mathbf{Vect}(u_1, \dots, u_n)$  l'ensemble des éléments de  $E$  qui sont combinaisons linéaires des vecteurs  $u_1, \dots, u_n$ .

Notation ensembliste :

$$\mathbf{Vect}(u_1, \dots, u_n) = \left\{ \sum_{i=1}^n \lambda_i u_i \mid (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n \right\} \text{ ou encore } \left\{ x \in E \mid \exists (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n \ x = \sum_{i=1}^n \lambda_i u_i \right\}$$

Caractérisation des éléments :

$$\text{Pour } x \in E, \quad x \in \mathbf{Vect}(u_1, \dots, u_n) \iff \exists (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n : x = \sum_{i=1}^n \lambda_i u_i$$

### 3.3.3 C'est un sous-espace vectoriel.

#### Proposition :

Pour  $u_1, \dots, u_n$  des vecteurs de  $E$ ,  $\mathbf{Vect}(u_1, \dots, u_n)$  est un sous-espace vectoriel de  $E$ .

**Démonstration.** On note  $F = \mathbf{Vect}(u_1, \dots, u_n)$  où  $u_1, \dots, u_n$  sont des vecteurs de  $E$ .

- $E$  est un  $\mathbb{K}$ -espace vectoriel et les  $u_i$  sont dans  $E$  donc  $F \subset E$ .
- $0_E = \sum_{k=1}^n 0u_k$  donc  $0_E \in F$
- Soient  $v_1 = \sum_{k=1}^n \alpha_k u_k$  et  $v_2 = \sum_{k=1}^n \beta_k u_k$  deux vecteurs de  $F$  et  $\lambda_1, \lambda_2$  deux éléments de  $\mathbb{K}$ .

$$\begin{aligned} \lambda_1 v_1 + \lambda_2 v_2 &= \lambda_1 \sum_{k=1}^n \alpha_k u_k + \lambda_2 \sum_{k=1}^n \beta_k u_k \\ &= \sum_{k=1}^n \lambda_1 \alpha_k u_k + \sum_{k=1}^n \lambda_2 \beta_k u_k \\ &= \sum_{k=1}^n (\lambda_1 \alpha_k u_k + \lambda_2 \beta_k u_k) \\ &= \sum_{k=1}^n (\lambda_1 \alpha_k + \lambda_2 \beta_k) u_k \\ &\in F \end{aligned}$$

En conclusion :  $F$  est un sous-espace vectoriel.

#### Remarques :

- $\mathbf{Vect}(u_1, \dots, u_n)$  est le plus petit sous-espace vectoriel de  $E$  contenant les vecteurs  $(u_1, \dots, u_n)$
- $\mathbf{Vect}(u_1, \dots, u_n)$  est appelé **sous-espace vectoriel engendré** par les vecteurs  $u_1, \dots, u_n$ .
- C'est un autre moyen de montrer qu'une partie de  $E$  est un sous-espace vectoriel

### 3.3.4 Opérations élémentaires.

On appelle opérations élémentaires sur la famille  $(u_1, u_2, \dots, u_m)$  les transformations suivantes :

- Permuter deux vecteurs :  $u_i \longleftrightarrow u_j$
- Multiplier un vecteur par un scalaire  $\lambda$  non nul :  $u_i \longleftarrow \lambda u_i$
- Ajouter un vecteur à un autre vecteur :  $u_i \longleftarrow u_i + u_j$

On combine souvent ces opérations. Par exemple on fait souvent des transvections :  $u_i \longleftarrow u_i - \alpha u_j$  avec  $i \neq j$

**Propositions**

Soit  $(u_1, \dots, u_m)$  une famille de vecteurs de  $E$  on note :  $F = \text{Vect}(u_1, u_2, \dots, u_m)$ .

- ❶ On ne modifie pas  $F$  en changeant l'ordre des  $u_i$ .
- ❷ On ne modifie pas  $F$  par des opérations élémentaires sur les  $u_i$ .
- ❸ On ne modifie pas  $F$  en supprimant  $0_E$  s'il est dans la liste des  $u_i$ .
- ❹ On ne modifie pas  $F$  en ajoutant à un vecteur  $u_i$  une combinaison linéaire des autres vecteurs.

**Démonstrations.**

**3.3.5 Familles génératrices.**

On note  $E$  un espace vectoriel et  $F$  un sous-espace vectoriel de  $E$ .

**Définition :**

Soit  $(u_1, u_2, \dots, u_n)$  une famille de vecteurs de  $E$   
 Dire que  $(u_1, u_2, \dots, u_n)$  est une **famille génératrice** de  $F$  signifie que,  $F = \text{Vect}(u_1, \dots, u_n)$

**Remarques :**

- En pratique lorsque  $(u_1, u_2, \dots, u_n) \in F^n$ ,
- $(u_1, u_2, \dots, u_n)$  est une **famille génératrice** de  $F$  si et seulement si,  $\forall v \in F, \exists (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n : v = \sum_{k=1}^n \lambda_k u_k$
- Toute sur-famille d'une famille génératrice de  $F$  est génératrice de  $F$ .

**3.3.6 Espace vectoriel de dimension finie.**

**Définition**

Dire qu'un espace vectoriel  $E$  est de dimension finie signifie qu'il existe  $n \in \mathbb{N}^*$  et  $(u_1, \dots, u_n) \in E^n$  tels que :

$$E = \text{Vect}(u_1, u_2, \dots, u_n)$$

**Remarque :**

On définit ici cette notion "de dimension finie" avant de définir la notion de dimension d'un espace vectoriel.

**3.3.7 Bases canoniques.**

**Base canonique de  $\mathbb{K}^n$  ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ )**

La famille  $(e_1, \dots, e_n)$  définie par :

$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad e_n = (0, \dots, 0, 1)$$

**Base canonique des matrices colonnes  $\mathcal{M}_{n,1}(\mathbb{R})$  et  $\mathcal{M}_{n,1}(\mathbb{C})$ .**

La famille  $(E_1, \dots, E_n)$  définie par :

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad E_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

**Base canonique des ensembles de polynômes  $\mathbb{R}_n[X]$  ou  $\mathbb{C}_n[X]$**

La famille  $(1, X, X^2, \dots, X^n)$

**Remarque :** Ici on donne ces familles qui engendrent naturellement ces espaces vectoriels. Comme leur nom l'indique nous verrons que ces familles sont des bases.

### 3.4 Familles libres.

Soient  $E$  un espace vectoriel sur  $\mathbb{K}$  et  $n$  un entier naturel non nul.

#### 3.4.1 Définition.

**Définition :**

Soit  $(u_1, u_2, \dots, u_n)$  une famille de vecteurs de  $E$ ,  
dire que  $(u_1, u_2, \dots, u_n)$  est une **famille libre** signifie que,  
quel que soit  $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ , si  $\sum_{k=1}^n \lambda_k u_k = 0_E$  alors  $\forall k \in \llbracket 1; n \rrbracket, \lambda_k = 0$

**Remarques :**

- La famille est **liée** lorsqu'il existe  $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$  tel que  $\sum_{k=1}^n \lambda_k u_k = 0_E$
- $(u_1, \dots, u_n)$  est libre si, et seulement si,  $\forall (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \left( \sum_{k=1}^n \lambda_k u_k = 0_E \iff \forall k \in \llbracket 1; n \rrbracket, \lambda_k = 0 \right)$
- Toute sous-famille d'une famille libre est libre.

#### 3.4.2 Identification.

**Théorème** (*Unicité de l'écriture sur une famille libre. Identification.*) :

Soient  $(u_1, u_2, \dots, u_n)$  une famille de vecteurs de  $E$ ,  $(a_1, a_2, \dots, a_n) \in \mathbb{K}^n$  et  $(b_1, b_2, \dots, b_n) \in \mathbb{K}^n$ .  
Si  $(u_1, u_2, \dots, u_n)$  est libre alors on a l'équivalence :  
$$a_1 u_1 + a_2 u_2 + \dots + a_n u_n = b_1 u_1 + b_2 u_2 + \dots + b_n u_n \iff (a_1, a_2, \dots, a_n) = (b_1, b_2, \dots, b_n)$$

**En effet :**  $\sum_{k=1}^n a_k u_k = \sum_{k=1}^n b_k u_k \iff \sum_{k=1}^n (a_k - b_k) u_k = 0_E \iff \forall k \in \llbracket 1, n \rrbracket, a_k - b_k = 0$   
 $(u_i)_i$  est libre

#### 3.4.3 Familles de polynômes.

**Théorème**

Toute famille finie de polynômes non nuls de degré deux à deux distincts est libre.

**Démonstration.**

(WLOG) on suppose que  $0 \leq \deg(P_1) < \dots < \deg(P_n)$

Soit  $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$  tel que  $\sum_{k=1}^n \lambda_k P_k = 0$

on raisonne par l'absurde en supposant  $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$

en notant :  $m = \max(\{k \in \llbracket 1, n \rrbracket \mid \lambda_k \neq 0\})$  on obtient :  $\sum_{k=1}^m \lambda_k P_k = 0$

comme  $\lambda_m \neq 0$ , on en déduit que :  $P_m = -\frac{1}{\lambda_m} \sum_{k=1}^{m-1} \lambda_k P_k$

mais l'hypothèse sur les degrés entraîne que  $-\frac{1}{\lambda_m} \sum_{k=1}^{m-1} \lambda_k P_k$  est de degré strictement inférieur à  $\deg(P_m)$

C'est impossible, donc nécessairement tous les  $\lambda_i$  sont nuls et ainsi la famille est libre. ■

### 3.5 Bases.

#### 3.5.1 Définition d'une base.

Définition :

Soient  $u_1, \dots, u_n$  des vecteurs de  $E$ .  
 Dire que  $(u_1, \dots, u_n)$  est une base de  $E$  signifie que,  
 $(u_1, \dots, u_n)$  est une **famille libre et génératrice** de  $E$ .

**Remarque :** Si  $(u_1, \dots, u_n)$  est une famille libre alors c'est une base de  $\text{Vect}(u_1, \dots, u_n)$ .

*En effet :*  $(u_1, \dots, u_n)$  est naturellement une famille génératrice de  $\text{Vect}(u_1, \dots, u_n)$ .

#### 3.5.2 Caractérisation d'une base.

Théorème :

Soient  $u_1, \dots, u_n$  des vecteurs de  $E$ ,  
 $(u_1, \dots, u_n)$  est une base de  $E$  si, et seulement si,  
 pour tout vecteur  $v$  de  $E$ , il existe une unique  $(x_1, \dots, x_n) \in \mathbb{K}^n$  tel que :  $v = \sum_{k=1}^n x_k u_k$

Démonstration.

Rapidement.

D'une part  $(u_1, \dots, u_n)$  est génératrice si, et seulement si,  $\forall v \in E : \underbrace{\exists(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n : v = \sum_{k=1}^n \lambda_k u_k}_{\text{existence}}$

D'autre part  $(u_1, \dots, u_n)$  est libre si, et seulement si,

$$\forall(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, (\lambda'_1, \dots, \lambda'_n) \in \mathbb{K}^n, \underbrace{\sum_{k=1}^n \lambda_k u_k = \sum_{k=1}^n \lambda'_k u_k \implies (\lambda_1, \dots, \lambda_n) = (\lambda'_1, \dots, \lambda'_n)}_{\text{unicité}}$$

On a bien

$$(u_1, \dots, u_n) \text{ est libre et génératrice si, et seulement si, } \forall v \in E : \underbrace{\exists!(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n : v = \sum_{k=1}^n \lambda_k u_k}_{\text{existence et unicité}}$$

### 3.6 Coordonnées dans une base.

#### 3.6.1 Définition.

Définition :

$\mathcal{B} = (e_1, \dots, e_n)$  est une base de  $E$ ,  
 à chaque vecteur  $v$  de  $E$  on peut associer l'unique matrice  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  telle que :  $v = x_1 e_1 + \dots + x_n e_n$   
 Cette matrice est appelée : **matrice colonne des coordonnées de  $v$  dans la base  $\mathcal{B}$**   
 On note :  $\text{Coord}_{\mathcal{B}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

**Remarque :**  $\forall v \in E, \quad \text{Coord}_{\mathcal{B}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \iff v = x_1 e_1 + \dots + x_n e_n$

### 3.6.2 Application $v \mapsto \text{Coord}_{\mathcal{B}}(v)$

**Théorème :**

Si  $\mathcal{B} = (e_1, e_2, \dots, e_n)$  est une base de  $E$  alors  
l'application  $\Phi : E \rightarrow \mathcal{M}_{n,1}(\mathbb{K})$  qui à  $v$  associe  $\text{Coord}_{\mathcal{B}}(v)$  est un isomorphisme.

**Rappels du cours application linéaire :**  $\Phi$  est isomorphisme signifie qu'elle est linéaire et bijective.

- Linéaire :  $\forall(\alpha, \beta) \in \mathbb{K}^2, \forall(u, v) \in E^2, \Phi(\alpha u + \beta v) = \alpha\Phi(u) + \beta\Phi(v)$
- Bijective :  $\forall X \in \mathcal{M}_{n,1}(\mathbb{K}), \exists! u \in E : \Phi(u) = X$

**Démonstration :**

En notant :  $\Phi : E \rightarrow \mathcal{M}_{n,1}(\mathbb{K})$   
 $u \mapsto \text{Coord}_{\mathcal{B}}(u)$

- Si  $u = \sum_{i=1}^n x_i e_i$  et  $v = \sum_{i=1}^n x'_i e_i$  alors  $\alpha u + \beta v = \sum_{i=1}^n (\alpha x_i + \beta x'_i) e_i$  donc  $\Phi(\alpha u + \beta v) = \alpha\Phi(u) + \beta\Phi(v)$
- Pour tout  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$  il existe un unique  $u \in E$  tel que  $u = \sum_{i=1}^n x_i e_i$  donc tel que  $\Phi(u) = X$

En conclusion :  $\Phi$  est un isomorphisme.

**Proposition :**

Soient  $\mathcal{B}$  une base de  $E$  et  $u_1, \dots, u_n$  des vecteurs de  $E$ .

- 1 Pour tout  $v \in E$ ,  $v \in \text{Vect}(u_1, \dots, u_n) \iff \text{Coord}_{\mathcal{B}}(v) \in \text{Vect}(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$
- 2  $(u_1, \dots, u_n)$  est libre si, et seulement si,  $(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$  est libre.

**Démonstration :**

- $\Phi$  est un isomorphisme donc :

$$\sum_{k=1}^n \lambda_k u_k = 0_E \quad \text{équivalent à} \quad \sum_{k=1}^n \lambda_k \text{Coord}_{\mathcal{B}}(u_k) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

ce qui entraîne que :

$(u_1, \dots, u_n)$  est libre si, et seulement si,  $(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$  est libre.

- De même on a l'équivalence

$$v = \sum_{k=1}^n \lambda_k u_k \quad \text{équivalent à} \quad \text{Coord}_{\mathcal{B}}(v) = \sum_{k=1}^n \lambda_k \text{Coord}_{\mathcal{B}}(u_k)$$

ce qui entraîne que :

$v \in \text{Vect}(u_1, \dots, u_n)$  si, et seulement si,  $\text{Coord}_{\mathcal{B}}(v) \in \text{Vect}(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$ .

### 3.6.3 Matrice d'une famille de vecteurs dans une base.

**Définition** (Matrice d'une famille de vecteurs dans une base) :

Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ , à chaque famille de vecteurs  $(v_1, \dots, v_m)$  de  $E$  on peut associer l'unique matrice  $M$  de  $\mathcal{M}_{n,m}(\mathbb{K})$  telle que : pour tout  $j \in \llbracket 1; m \rrbracket$ , la  $j$ -ième colonne de  $M$  est  $\text{Coord}_{\mathcal{B}}(v_j)$ .

Cette matrice est la **matrice de la famille de vecteurs** de  $(v_1, \dots, v_m)$  dans la base  $\mathcal{B}$ .

On la note :  $\text{Mat}_{\mathcal{B}}(v_1, \dots, v_m)$

**Remarques :**

- On peut résumer avec la relation.

$$\text{Mat}_{\mathcal{B}}(v_1, \dots, v_m) = \left( \text{Coord}_{\mathcal{B}}(v_1) \mid \dots \mid \text{Coord}_{\mathcal{B}}(v_m) \right)$$

- Multiplication par une matrice colonne :

$$\text{Mat}_{\mathcal{B}}(v_1, \dots, v_m) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \sum_{k=1}^m x_k \text{Coord}_{\mathcal{B}}(v_k)$$

$$\text{Mat}_{\mathcal{B}}(v_1, \dots, v_m) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \text{Coord}_{\mathcal{B}} \left( \sum_{k=1}^m x_k v_k \right)$$

Extrait de la feuille cours\_3\_2.

Soient  $A \in \mathcal{M}_{n,p}(\mathbb{R})$  et  $X \in \mathcal{M}_{p,1}(\mathbb{R})$ .

$$\begin{aligned} AX &= \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^p a_{1,j} x_j \\ \vdots \\ \sum_{j=1}^p a_{n,j} x_j \end{pmatrix} \\ &= \sum_{j=1}^p \begin{pmatrix} a_{1,j} x_j \\ \vdots \\ a_{n,j} x_j \end{pmatrix} \\ &= \sum_{j=1}^p x_j \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix} \end{aligned}$$

$AX$  est une combinaison linéaire des colonnes de  $A$

## 3.7 Dimension d'un espace vectoriel

### 3.7.1 Définition.

#### Théorème et définition

Soit  $E$  un espace vectoriel de dimension finie différent de  $\{0_E\}$ .

- ❶  $E$  possède au moins une base.
- ❷ toutes les bases de  $E$  ont le même nombre de vecteurs.
- ❸ Le nombre de vecteurs d'une base de  $E$  est appelée dimension de  $E$ . On note  $\dim(E)$  la dimension de  $E$ .

**Démonstration.** On admet ce théorème.

**Remarques :** L'espace vectoriel  $\{0_E\}$  a par convention une dimension nulle.  $\dim(\{0_E\}) = 0$ .  
(Sa base est la famille vide)

**Rappel sur les systèmes homogènes :**

- S'il y a strictement plus d'inconnues que d'équations alors le système homogène a une infinité de solutions.
- S'il y a strictement plus d'inconnues que d'équations alors le système homogène a d'autres solutions que  $(0, \dots, 0)$ .
- Si le système homogène a une unique solution alors le nombre d'équations est supérieur ou égal au nombre d'inconnues.

**Théorème.**

Soient  $E$  un espace vectoriel de dimension finie et  $(u_1, \dots, u_p)$  une famille de  $p$  vecteurs de  $E$ .  
Si  $(u_1, \dots, u_p)$  est libre alors  $p \leq \dim(E)$ .

**Démonstration :**

Soit  $(u_1, \dots, u_p)$  une famille libre de vecteurs d'un espace vectoriel  $E$  de dimension  $n$ .

(On note  $\mathcal{B}$  une base de  $E$ )

on a donc : (définition de la liberté)

$$\forall (x_1, \dots, x_p) \in \mathbb{K}^p, \quad \sum_{k=1}^p x_k u_k = 0_E \iff (x_1, \dots, x_p) = (0, \dots, 0)$$

ce qui donne dans la base  $\mathcal{B}$  : en notant  $M = \text{Mat}_{\mathcal{B}}(u_1, \dots, u_p)$

$$\forall (x_1, \dots, x_p) \in \mathbb{K}^p, \quad M \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \iff (x_1, \dots, x_p) = (0, \dots, 0)$$

donc le système  $MX = 0$  a une unique solution et ainsi il a autant ou plus d'équations que d'inconnues.

(on utilise ici le rappel précédent)

autrement dit  $p \leq n$ .

**Remarque :** En dimension  $n$ , une famille libre contient au plus  $n$  vecteurs.

**Corollaire**

Si pour tout  $n \in \mathbb{N}^*$ , s'il existe une famille libre de  $n$  vecteurs de  $E$  alors  $E$  n'est pas de dimension finie.

**Exemples :**

Pour tout  $n \in \mathbb{N}^*$ , la famille  $(x \mapsto x^k)_{1 \leq k \leq n}$  est une famille libre de  $n$  vecteurs de  $\mathbb{R}^{\mathbb{R}}$  donc

$\mathbb{R}^{\mathbb{R}}$  n'est pas de dimension finie

Pour tout  $n \in \mathbb{N}^*$ , la famille  $(X^k)_{1 \leq k \leq n}$  est une famille libre de  $n$  vecteurs de  $\mathbb{R}[X]$  donc

$\mathbb{R}[X]$  n'est pas de dimension finie

**3.7.2 Dans une famille libre chaque vecteur donne une nouvelle direction.****Théorème**

Soient  $E$  un espace vectoriel et  $(u_1, \dots, u_n, v)$  des vecteurs de  $E$ .

Si  $(u_1, \dots, u_n)$  est libre, alors les deux propriétés suivantes sont équivalentes :

$$v \notin \text{Vect}(u_1, \dots, u_n) \iff (u_1, \dots, u_n, v) \text{ est libre.}$$

**Démonstration :**

Supposons que  $(u_1, \dots, u_n)$  soit une famille libre.

- Supposons que  $v \notin \text{Vect}(u_1, \dots, u_n)$  et montrons que  $(u_1, \dots, u_n, v)$  est libre.

Soient  $\lambda_1, \dots, \lambda_n, \mu \in \mathbb{K}$  tels que  $\lambda_1 u_1 + \dots + \lambda_n u_n + \mu v = 0_E$  (\*)

Si  $\mu \neq 0$ , alors on aurait  $v = -\frac{\lambda_1}{\mu} u_1 - \dots - \frac{\lambda_n}{\mu} u_n \in \text{Vect}(u_1, \dots, u_n)$ , ce qui est impossible donc  $\mu = 0$ .

L'égalité (\*) devient  $\lambda_1 u_1 + \dots + \lambda_n u_n = 0$ , mais  $(u_1, \dots, u_n)$  est libre, donc  $\lambda_1 = \dots = \lambda_n = 0$ .

ainsi,  $\lambda_1 = \dots = \lambda_n = \mu = 0$ , ce qui achève la démonstration de  $(u_1, \dots, u_n, v)$  est libre.

• Supposons que  $v \in \text{Vect}(u_1, \dots, u_n)$  et montrons que  $(u_1, \dots, u_n, v)$  est liée.

$v \in \text{Vect}(u_1, \dots, u_n)$  donc il existe  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  tels que  $v = \lambda_1 u_1 + \dots + \lambda_n u_n$ .

On obtient alors  $\lambda_1 u_1 + \dots + \lambda_n u_n - v = 0_E$ , qui est une relation linéaire avec des coefficients non tous nuls, ce qui montre que  $(u_1, \dots, u_n, v)$  est liée.

On a bien montré :

$$v \notin \text{Vect}(u_1, \dots, u_n) \iff (u_1, \dots, u_n, v) \text{ est libre.}$$

**Remarque :** Dans une famille libre chaque vecteur donne une nouvelle direction.

**Corollaire**

Soient  $E$  un espace vectoriel,  $n \in \mathbb{N}^*$  et  $(u_1, \dots, u_n)$  une famille de  $n$  vecteurs de  $E$ .

$u_1 \neq 0_E$  et  $\forall k \in \llbracket 1, n-1 \rrbracket$ ,  $u_{k+1} \notin \text{Vect} \langle u_1, \dots, u_k \rangle$  si, et seulement si,  $(u_1, \dots, u_n)$  est libre

**En effet :**  $\boxed{\Rightarrow}$  par récurrence sur  $k$  avec le théorème et  $\boxed{\Leftarrow}$  avec le théorème est le fait qu'une sous-famille d'une famille libre est libre.

**Remarque :**

On peut avec ce corollaire démontrer le théorème sur les familles de polynômes non nuls de degrés distincts.

### 3.7.3 Compléter une famille libre ou extraire d'une famille génératrice.

**Théorème.**

Soit  $E$  un espace vectoriel de dimension finie.

- ❶ De toute famille génératrice de  $E$  on peut extraire une base de  $E$ .
- ❷ Toute famille libre de  $E$  peut être complétée en une base de  $E$ .

**Démonstrations :** (*algorithmiques*)

On note  $\mathcal{F}$  la famille considérée.

❶ On part d'une famille génératrice  $\mathcal{F}$ .

Tant que  $\mathcal{F}$  n'est pas libre, on retire un vecteur de  $\mathcal{F}$  qui s'exprime comme combinaison linéaire des autres.

À chaque étape,  $\mathcal{F}$  reste génératrice.

Le processus s'arrête en un nombre fini d'étapes, et la famille obtenue est une base de  $E$ .

❷ On part d'une famille libre  $\mathcal{F}$ .

Tant que  $\mathcal{F}$  n'est pas génératrice, on ajoute à  $\mathcal{F}$  un vecteur de  $E$  qui ne s'exprime pas comme combinaison linéaire des autres.

À chaque étape,  $\mathcal{F}$  reste libre.

Le processus s'arrête en un nombre fini d'étapes, et la famille obtenue est une base de  $E$ .

**Remarques :**

- Une famille génératrice d'un sous-espace vectoriel de dimension  $n$  a au moins  $n$  vecteurs.
- Une famille libre d'un sous-espace vectoriel de dimension  $n$  a au plus  $n$  vecteurs.

## 3.8 Famille libre, famille génératrice et dimension.

$n$  désigne ici un entier naturel non nul.

### 3.8.1 Famille génératrice en dimension $n$ .

**Théorème :**

Soit  $E$  un espace vectoriel de dimension finie et différent de  $\{0_E\}$ .

Si  $E$  est dimension  $n$ , toute famille génératrice de  $E$  formée de  $n$  vecteurs est une base de  $E$ .

**En effet :** si  $\mathcal{F}$  n'était pas libre on pourrait retirer un vecteur tout en restant génératrice, ce qui donnerait une famille génératrice de cardinal  $< n$ , impossible donc  $\mathcal{F}$  est libre et c'est une base.

**Rédaction type :** On a montré que  $\begin{cases} \mathcal{F} \text{ est génératrice de } E \\ \mathcal{F} \text{ contient } n \text{ vecteurs} \\ \dim(E) = n \end{cases} \quad \text{donc } \mathcal{F} \text{ est une base de } E.$

### 3.8.2 Famille libre en dimension $n$ .

**Théorème :**

Soit  $E$  un espace vectoriel de dimension finie et différent de  $\{0_E\}$ .  
 Si  $E$  est dimension  $n$ , toute famille libre de  $E$  formée de  $n$  vecteurs est une base de  $E$ .

**En effet :** si  $\mathcal{F}$  n'était pas génératrice on pourrait ajouter un vecteur sans perdre la liberté, ce qui donnerait une famille libre de cardinal  $> n$ , impossible donc  $\mathcal{F}$  est génératrice et c'est une base.

**Rédaction type :** On a montré que  $\begin{cases} \mathcal{F} \text{ est une famille libre} \\ \mathcal{F} \text{ contient } n \text{ vecteurs de } E \\ \dim(E) = n \end{cases} \quad \text{donc } \mathcal{F} \text{ est une base de } E.$

### 3.8.3 Dimension et sous-espace vectoriel.

**Proposition.**

Soit  $E$  un espace vectoriel et  $F$  une partie de  $E$ ,  
 Si  $F$  est un sous-espace vectoriel de  $E$  et  $E$  est de dimension finie alors  $F$  est de dimension finie.

**Démonstration :** (*admis*)

**Théorème :**

Soit  $E$  un espace vectoriel de dimension finie,  
 et soient  $F$  et  $G$  deux sous-espaces vectoriels de  $E$ ,  
 ① Si  $F \subset G$  alors  $\dim(F) \leq \dim(G)$   
 ②  $[ F \subset G \text{ et } \dim(F) = \dim(G) ] \iff F = G$

**Démonstration.**

- ①  $F$  est de dimension finie d'après la proposition précédente donc  $F$  admet une base  $(e_1, \dots, e_m)$  d'après 3.7.1. et comme  $F \subset G$  on a  $(e_1, \dots, e_m)$  une famille libre de  $G$ , ce qui entraîne  $m \leq \dim(G)$  d'après 3.7.1.
- ② Seule  $\Rightarrow$  nécessite une démonstration.  
 On suppose que :  $F \subset G$  et  $\dim(F) = \dim(G)$  et on note  $(e_1, \dots, e_m)$  une base de  $F$   
 d'une part :  $F \subset G$  on a  $(e_1, \dots, e_m)$  une famille libre de  $G$ , d'autre part :  $\dim(F) = \dim(G) = m$   
 donc  $(e_1, \dots, e_m)$  est une base de  $G$  d'après 3.8.2. , mais c'est aussi une base de  $F$  donc on a bien  $F = G$ .

## 3.9 Systèmes linéaires homogènes et dimension.

Soit  $M$  une matrice de  $\mathcal{M}_{p,n}(\mathbb{K})$ , on s'intéresse ici à la résolution du système linéaire  $MX = 0$ .  
*p équations et n inconnues.*

**Rappels :**

- L'algorithme du pivot de Gauss permet de réduire le système  $MX = 0$  en un système échelonné  $TX = 0$ .
- Le nombre de lignes non nuls de  $TX = 0$  est égal au nombre d'inconnues principales noté  $r$ .  
*(C'est le rang du système)*
- Le nombre d'inconnues secondaires vaut  $n - r$ .

**Théorème.**

L'ensemble des solutions du système  $MX = 0$  est un sous-espace vectoriel de dimension  $n - r$ .

**Démonstration :**

Notons  $S$  l'ensemble des solutions du système  $MX = 0$ .

- $0_{n,1} \in S$  car  $M0_{n,1} = 0_{n,1}$ .
- Soient  $(X, Y) \in S^2$  et  $(\alpha, \beta) \in \mathbb{K}^2$ . Alors  $MX = 0$  et  $MY = 0$ , donc  $M(\alpha X + \beta Y) = \alpha MX + \beta MY = 0$ . ainsi  $\alpha X + \beta Y \in S$ , on peut alors en déduire que  $S$  est un sous-espace vectoriel de  $\mathcal{M}_{n,1}(\mathbb{K})$ .

Par opérations élémentaires sur les lignes, le système  $MX = 0$  est équivalent à un système échelonné  $TX = 0$ . Ainsi, toute solution s'écrit comme combinaison linéaire de  $n - r$  solutions particulières obtenues en donnant successivement à l'une des inconnues secondaires la valeur 1 et aux autres la valeur 0.

L'ensemble  $S$  est donc engendré par  $n - r$  vecteurs, et ces vecteurs forment une famille libre.

Donc  $S$  est un sous-espace vectoriel de dimension  $n - r$ .

*Conseil général : N'utilisez ce théorème uniquement si vous êtes sûr de vous et que cela est nécessaire.*

-----

Illustration sur un exemple : *Extrait de la feuille \_calcul\_4*

Soit  $u = (x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$ ,

$$\begin{aligned}
 u \in F &\iff \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0 \\ 2x_1 + 4x_2 + x_3 + 5x_5 = 0 \\ x_3 + x_4 + x_5 = 0 \end{cases} \\
 &\iff \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0 \\ -5x_3 - 8x_4 - 5x_5 = 0 \\ x_3 + x_4 + x_5 = 0 \end{cases} \\
 &\iff \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0 \\ -5x_3 - 8x_4 - 5x_5 = 0 \\ -3x_4 = 0 \end{cases} \quad (3 \text{ inc. principales et } 2 \text{ inc. secondaires}) \\
 &\iff \begin{cases} x_1 = -2x_2 - 2x_5 \\ x_3 = -x_5 \\ x_4 = 0 \end{cases}
 \end{aligned}$$

$$\begin{aligned}
 \text{donc } F &= \{(-2x_2 - 2x_5, x_2, -x_5, 0, x_5) \mid (x_2, x_5) \in \mathbb{R}^2\} \\
 &= \{x_2(-2, 1, 0, 0, 0) + x_5(-2, 0, -1, 0, 1) \mid (x_2, x_5) \in \mathbb{R}^2\}
 \end{aligned}$$

Donc  $F = \text{Vect} \langle (-2, 1, 0, 0, 0), (2, 0, -1, 0, 1) \rangle$  et  $((-2, 1, 0, 0, 0), (-2, 0, -1, 0, 1))$  est libre.

$F$  est un sous-espace vectoriel et  $((-2, 1, 0, 0, 0), (-2, 0, -1, 0, 1))$  est une base de  $F$ .

*Remarque : cet espace est de dimension 2, mais il est infini (l'ensemble des solutions est infini)*

**Corollaire.**

- ❶ S'il y a strictement plus d'inconnues que d'équations ( $n > p$ ) alors le système  $MX = 0$  a une infinité de solutions.
- ❷ Si  $(0, \dots, 0)$  est l'unique solution de  $MX = 0$ , alors il y a autant ou plus d'équations que d'inconnues ( $n \leq p$ ).

**En effet :** Pour ❶ : Si  $n > p$  alors il y a au moins une inconnue secondaire donc une infinité de solution. et ❷ est la contraposée de ❶ .

### 3.10 Rang.

#### 3.10.1 Rang d'une famille de vecteurs.

Soient  $E$  un espace vectoriel et  $n$  un entier naturel non nul.

**Définition.**

Pour  $(u_1, \dots, u_n)$  une famille de  $n$  vecteurs de  $E$ .

$$\text{rg}(u_1, u_2, \dots, u_n) = \dim(\text{Vect}(u_1, u_2, \dots, u_n))$$

**Remarque :** On ne change pas le rang en faisant des opérations élémentaires sur les vecteurs.

**Théorèmes :**

Soit  $\mathcal{F}$  une famille de vecteurs de  $E$

- ❶  $\mathcal{F}$  est libre si, et seulement si,  $\text{rg}(\mathcal{F})$  est égal au nombre de vecteurs de  $\mathcal{F}$ .
- ❷  $\mathcal{F}$  est génératrice de  $E$  si, et seulement si,  $\text{rg}(\mathcal{F})$  est égal à  $\dim(E)$ .
- ❸  $\mathcal{F}$  est une base de  $E$  si, et seulement si,  $\begin{cases} \text{rg}(\mathcal{F}) \text{ est égal au nombre de vecteurs de } \mathcal{F} \\ \text{et est égal à la dimension de } E \end{cases}$

**Démonstration.**

❶ Si  $\mathcal{F}$  est libre, alors c'est une base de  $\text{Vect}(\mathcal{F})$  donc  $\text{rg}(\mathcal{F})$  est égal au nombre de vecteurs de  $\mathcal{F}$ .

Réciproquement si  $\text{rg}(\mathcal{F})$  est égal au nombre  $m$  de vecteurs de  $\mathcal{F}$  alors  $\mathcal{F}$  est une famille génératrice de  $m$  vecteurs de  $\text{Vect}(\mathcal{F})$  qui est de dimension  $m$ , donc c'est une base de  $\text{Vect}(\mathcal{F})$  d'où  $\mathcal{F}$  est libre.

❷ Si  $\mathcal{F}$  est génératrice de  $E$ , alors  $E = \text{Vect}(\mathcal{F})$  donc  $\dim(E) = \text{rg}(\mathcal{F})$ .

Réciproquement : Si  $\text{rg}(\mathcal{F}) = \dim(E)$  alors  $\text{Vect}(\mathcal{F}) \subset E$  et  $\dim(\text{Vect}(\mathcal{F})) = \dim(E)$  donc (d'après 3.8.3)  $E = \text{Vect}(\mathcal{F})$  autrement dit :  $\mathcal{F}$  est génératrice de  $E$ .

❸ C'est une conséquence directe des ❶ et ❷.

### 3.10.2 Rang d'une matrice.

**Définition. (Définition du rang d'une matrice)**

Soit  $M$  une matrice de  $\mathcal{M}_{n,m}(\mathbb{K})$ ,

On appelle rang de la matrice  $M$ , la dimension de l'espace engendré par les colonnes de  $M$ .

Notation : on note  $\text{rg}(M)$  le rang de la matrice  $M$ .

**Théorèmes :**

- ❶ Pour toute matrice  $M$  de  $\mathcal{M}_{n,m}(\mathbb{K})$ ,  $\text{rg}(M^\top) = \text{rg}(M)$ .
- ❷ Pour toute matrice  $M$  de  $\mathcal{M}_n(\mathbb{K})$ ,  $\text{rg}(M) = n$  si, et seulement si,  $M$  est inversible.

**Démonstration. (admis)**

**Remarques :**

- Le rang d'une matrice est (aussi) la dimension de l'espace engendré par ses lignes.
- On ne modifie pas le rang d'une matrice en faisant des opérations élémentaires sur ses colonnes et ses lignes.
- Si  $M \in \mathcal{M}_{n,p}(\mathbb{K})$  alors  $\text{rg}(M) \leq n$  et  $\text{rg}(M) \leq p$  (ou encore  $\text{rg}(M) \leq \min(n, p)$ )
- On obtient ce rang en appliquant l'algorithme de Gauss sur les lignes ou sur les colonnes de la matrice.

### 3.10.3 Matrices et familles de vecteurs.

Soit  $E$  un espace vectoriel de dimension finie.

**Théorème.**

Soient  $(u_1, u_2, \dots, u_n)$  une famille de vecteurs de  $E$  et  $\mathcal{B}$  une base de  $E$ .

$$\text{rg}(u_1, u_2, \dots, u_n) = \text{rg}(\text{Mat}_{\mathcal{B}}(u_1, u_2, \dots, u_n))$$

**Démonstration :** Ici on utilise ❶ et ❷ de la proposition de 3.6.2.

Notons  $r = \text{rg}(u_1, \dots, u_n)$ .

3.7.3 montre qu'il existe donc une sous-famille de  $r$  vecteurs, par exemple  $(u_{i_1}, \dots, u_{i_r})$ , qui est libre et telle que  $\text{Vect}(u_1, \dots, u_n) = \text{Vect}(u_{i_1}, \dots, u_{i_r})$ .

D'après ❷,  $(\text{Coord}_{\mathcal{B}}(u_{i_1}), \dots, \text{Coord}_{\mathcal{B}}(u_{i_r}))$  est libre.

Et d'après ❶,  $\text{Vect}(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n)) = \text{Vect}(\text{Coord}_{\mathcal{B}}(u_{i_1}), \dots, \text{Coord}_{\mathcal{B}}(u_{i_r}))$ .

Donc  $(\text{Coord}_{\mathcal{B}}(u_{i_1}), \dots, \text{Coord}_{\mathcal{B}}(u_{i_r}))$  est une base de  $\text{Vect}(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$ .

On a bien :  $\text{rg}(u_1, \dots, u_n) = \text{rg}(\text{Mat}_{\mathcal{B}}(u_1, \dots, u_n))$ .

**Théorèmes.**

Soient  $(u_1, u_2, \dots, u_n)$  une famille de vecteurs de  $E$  et  $\mathcal{B}$  une base de  $E$ .

On note  $M = \text{Mat}_{\mathcal{B}}(u_1, u_2, \dots, u_n)$

❶  $(u_1, u_2, \dots, u_n)$  est libre si, et seulement si,  $\text{rg}(M) = n$

❷  $(u_1, u_2, \dots, u_n)$  est génératrice si, et seulement si,  $\text{rg}(M) = \dim(E)$

❸  $(u_1, u_2, \dots, u_n)$  est une base de  $E$  si, et seulement si,  $M$  est inversible.

**Remarque :** Ici  $n$  est le nombre de vecteurs de la famille et non la dimension de  $E$ .

**Démonstration :** Conséquence des théorèmes de 3.10.1.

### 3.10.4 Rang d'un système.

**Définition :**

Le rang d'un système est le nombre d'inconnues principales après réduction par l'algorithme du pivot de Gauss.

**Proposition :**

Soient  $M \in \mathcal{M}_{n,m}(\mathbb{K})$ ,  $B \in \mathcal{M}_{n,1}(\mathbb{K})$  et  $\Sigma$  le système  $MX = B$

$$\text{rg}(M) = \text{rg}(\Sigma)$$

# Probabilités

## Plan du chapitre

---

4.1	Ensemble des événements et probabilités. . . . .	38
4.1.1	Ensemble des événements. . . . .	38
4.1.2	Notion de tribu. . . . .	38
4.1.3	Système complet d'événements . . . . .	39
4.1.4	Probabilité. . . . .	39
4.1.5	Négligeable, presque sûr, quasi-complet. . . . .	41
4.1.6	Formule des probabilités totales (Version 1) . . . . .	41
4.2	Conditionnement et indépendance. . . . .	42
4.2.1	Définition. . . . .	42
4.2.2	Formule des probabilités composées . . . . .	43
4.2.3	Formule des probabilités totales (Version 2) . . . . .	43
4.2.4	Formule de Bayes . . . . .	43
4.2.5	Indépendance . . . . .	44
4.3	Variables aléatoires réelles. . . . .	45
4.3.1	Définition. . . . .	45
4.3.2	Fonction de répartition. . . . .	45
4.3.3	Indépendance de variables aléatoires. . . . .	45
4.3.4	Propriétés de l'indépendance mutuelle. . . . .	46
4.4	Espérance, variance et moments d'une variable aléatoire réelle. . . . .	46
4.4.1	Espérance. . . . .	46
4.4.2	Variance. . . . .	48
4.4.3	Formule de Kœnig-Huygens. . . . .	48
4.4.4	Variance de $aX + b$ : . . . . .	48
4.4.5	Ecart-type. . . . .	49
4.4.6	Moments d'ordres supérieurs. . . . .	49
4.4.7	Avec des variables aléatoires indépendantes. . . . .	49

---

## 4.1 Ensemble des événements et probabilités.

On cherche à modéliser une expérience aléatoire, on nomme  $\Omega$  l'ensemble des résultats (l'univers).

### 4.1.1 Ensemble des événements.

### 4.1.2 Notion de tribu.

Soit  $\mathcal{T}$  une partie de  $\mathcal{P}(\Omega)$ ,  
 dire que  $\mathcal{T}$  est une **tribu** sur  $\Omega$  signifie que :

- ❶  $\Omega \in \mathcal{T}$  (Contient l'événement certain)
- ❷  $\forall A \in \mathcal{T}, \bar{A} \in \mathcal{T}$ . ( $\mathcal{T}$  est stable par passage au complémentaire).
- ❸  $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}, \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{T}$ . ( $\mathcal{T}$  est stable par union dénombrable).

- Remarques :**
- Cette définition ne fera pas l'objet de questions directes, mais elle doit être connue.
  - La donnée de  $(\Omega, \mathcal{F})$  est appelée : espace probabilisable.
  - $\mathcal{F}$  est l'ensemble des événements.
  - Un ensemble  $E$  est dénombrable s'il existe une bijection de  $\mathbb{N}$  dans  $E$ .

autrement dit : s'il existe une suite  $(x_n)$  telle que 
$$\begin{cases} E = \{ x_n \mid n \in \mathbb{N} \} \\ \text{et } \forall (i, j) \in \mathbb{N}^2, i \neq j \implies x_i \neq x_j \end{cases}$$

Avec cette définition les ensembles dénombrables sont infinis. Vous trouverez d'autres définitions.

**Propositions.**

Soit  $\mathcal{F}$  une partie de  $\mathcal{P}(\Omega)$ ,  
 si  $\mathcal{F}$  est une **tribu** alors :

- ❶  $\emptyset \in \mathcal{F}$  (Contient l'événement impossible).
- ❷  $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}, \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{F}$ . ( $\mathcal{F}$  est stable par intersection dénombrable).
- ❸  $\mathcal{F}$  est stable par réunion finie et par intersection finie.

**Démonstrations.** (Faites au tableau)

**Propriétés des intersections et des réunions dénombrables.**

Soient  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements,  $B$  un événement et  $N$  un entier naturel non nul,

**Complémentaire :**  $\overline{\bigcap_{n=0}^{+\infty} A_n} = \bigcup_{n=0}^{+\infty} \overline{A_n}$        $\overline{\bigcup_{n=0}^{+\infty} A_n} = \bigcap_{n=0}^{+\infty} \overline{A_n}$

**Distributivité,**  $B \cup \left( \bigcap_{n=0}^{+\infty} A_n \right) = \bigcap_{n=0}^{+\infty} (B \cup A_n)$        $B \cap \left( \bigcup_{n=0}^{+\infty} A_n \right) = \bigcup_{n=0}^{+\infty} (B \cap A_n)$ .

**Inclusion,**  $\bigcap_{n=0}^{+\infty} A_n \subset \bigcap_{n=0}^N A_n$        $\bigcup_{n=0}^N A_n \subset \bigcup_{n=0}^{+\infty} A_n$

**Démonstrations.** (Voir cours outils sur logique et ensembles)

**4.1.3 Système complet d'événements**

**Définition.**

Soit  $(A_n)_{n \in I}$  une suite finie ou dénombrable d'événements ( $I = \underbrace{[1; n]}_{\text{finie}}$  ou  $I = \underbrace{\mathbb{N}}_{\text{dénombrable}}$ ),

Dire que  $(A_n)_{n \in I}$  est un système complet d'événements signifie que :

- ❶  $\forall (i, j) \in I^2, i \neq j \implies A_i \cap A_j = \emptyset$       et      ❷  $\bigcup_{n \in I} A_n = \Omega$

**Remarque :** ❶ se dit :  $(A_n)$  est une suite d'événements deux à deux disjoints.  
 ou encore :  $(A_n)$  est une suite d'événements deux à deux incompatibles.

**4.1.4 Probabilité.**

**Définition.**

Soient  $(\Omega, \mathcal{F})$  un espace probabilisable et  $P$  une application de  $\mathcal{F}$  dans  $\mathbb{R}$ .

Dire que  $P$  est une **probabilité** sur  $(\Omega, \mathcal{F})$  signifie que :

- ❶  $\forall A \in \mathcal{F}, \quad P(A) \in [0, 1]$ .
- ❷  $P(\Omega) = 1$
- ❸ Quel que soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements. (*Axiome de  $\sigma$ -additivité*)  
si les  $A_n$  sont deux à deux incompatibles,

$$\text{alors la série } \sum_{n \geq 0} P(A_n) \text{ converge et : } P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$$

**Remarque :**    ❸ entraîne : Si  $A \cap B = \emptyset$  alors  $P(A \cup B) = P(A) + P(B)$ .

On retrouve donc les mêmes résultats que dans le cours de première année lorsque  $\Omega$  est fini et  $\mathcal{F} = \mathcal{P}(\Omega)$  :

- $n$  étant fixé dans  $\mathbb{N}$ , quel que soit  $(A_k)_{1 \leq k \leq n}$  une suite de  $n$  événements,  
si les  $A_k$  sont deux à deux incompatibles,    alors     $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$
- $\forall A \in \mathcal{F}, \quad P(\bar{A}) = 1 - P(A)$     En effet :  $A \cup \bar{A} = \Omega$  et  $A \cap \bar{A} = \emptyset$  donc  $P(A) + P(\bar{A}) = P(\Omega) = 1$
- $\forall (A, B) \in \mathcal{F}^2, \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$ .  
*En effet :*  $\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (\bar{A} \cap B)) = P(A) + P(\bar{A} \cap B) = P(A) + P(B) - P(A \cap B)$
- Pour tout  $(A, B) \in \mathcal{F}^2$ ,    si  $A \subset B$  alors  $P(A \setminus B) = P(B) - P(A)$ .
- Pour tout  $(A, B) \in \mathcal{F}^2$ ,    si  $A \subset B$  alors  $P(A) \leq P(B)$ .  
*En effet :*  $P(B) = P(A \cap B) + P(\bar{A} \cap B) = P(A) + P(\bar{A} \cap B) \geq 0$
- Quel que soit  $(A_k)_{1 \leq k \leq n}$  une suite de  $n$  événements,  $P\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n P(A_k)$

**Proposition.** (*complément, à savoir redémontrer*)

Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements,

- ❶ si  $\lim_{N \rightarrow +\infty} P\left(\bigcup_{n=0}^N A_n\right) = 1$     alors :     $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1$
- ❷ si  $\lim_{N \rightarrow +\infty} P\left(\bigcap_{n=0}^N A_n\right) = 0$     alors :     $P\left(\bigcap_{n=0}^{+\infty} A_n\right) = 0$

**Démonstration :** (*faite au tableau*)

- ❷ Pour tout  $n \in \mathbb{N}$ ,  $\bigcup_{k=0}^n A_k \subset \bigcup_{k=0}^{+\infty} A_k$     donc     $P\left(\bigcup_{k=0}^n A_k\right) \leq P\left(\bigcup_{k=0}^{+\infty} A_k\right)$
- on a  $\forall n \in \mathbb{N}, \quad P\left(\bigcup_{k=0}^n A_k\right) \leq P\left(\bigcup_{k=0}^{+\infty} A_k\right) \leq 1$     et     $\lim_{n \rightarrow +\infty} \left(P\left(\bigcup_{k=0}^n A_k\right)\right) = 1$

$$\text{donc (passage à la limite) } \boxed{P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1}$$

- ❶ Pour tout  $n \in \mathbb{N}$ ,  $\bigcap_{k=0}^{+\infty} A_k \subset \bigcap_{k=0}^n A_k$     donc     $P\left(\bigcap_{k=0}^{+\infty} A_k\right) \leq P\left(\bigcap_{k=0}^n A_k\right)$
- on a  $\forall n \in \mathbb{N}, \quad 0 \leq P\left(\bigcap_{k=0}^{+\infty} A_k\right) \leq P\left(\bigcap_{k=0}^n A_k\right)$     et     $\lim_{n \rightarrow +\infty} \left(P\left(\bigcap_{k=0}^n A_k\right)\right) = 0$

$$\text{donc (passage à la limite) } \boxed{P\left(\bigcap_{n=0}^{+\infty} A_n\right) = 0}$$

**Proposition.**

Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements,

- si  $(A_n)_{n \in \mathbb{N}}$  est un système complet d'événements alors :  $\sum_{n=0}^{+\infty} P(A_n) = 1$

**Remarques en lien avec le cours sur les séries :**

- Si  $(A_n)$  est une suite d'événements deux à deux incompatibles, la définition d'une probabilité de première année suffit pour montrer que la série  $\sum_{n \geq 0} P(A_n)$  converge.
- Les séries convergentes sont ici toutes absolument convergentes, l'ordre de sommation n'a pas d'importance, on pourra sans ambiguïté, pour  $I \subset \mathbb{N}$  utiliser la notation  $\sum_{i \in I} P(A_i)$  pour les sommes.

**4.1.5 Négligeable, presque sûr, quasi-complet.**

**Définitions.**

Soit  $A$  un événement,  
 dire que  $A$  est **négligeable** signifie que  $P(A) = 0$ .  
 dire que  $A$  est **presque sûr** signifie que  $P(A) = 1$ .  
 Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements,  
 dire que  $(A_n)_{n \in \mathbb{N}}$  est un système **quasi-complet** d'événements signifie que :

**❶**  $\forall (i, j) \in \mathbb{N}^2, \quad i \neq j \implies A_i \cap A_j = \emptyset$   
**❷**  $\sum_{n=0}^{+\infty} P(A_n) = 1$  (La série converge et sa somme vaut 1)    ou     $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1$

**Remarque :** Si un système est complet alors il est quasi-complet.    *La réciproque n'est pas toujours vraie.*

**Propriétés.**

Soit  $A$  un événement,

**❶** Si  $A$  est quasi-impossible alors :  
 pour tout événement  $B$ ,     $P(A \cap B) = 0$     et     $P(A \cup B) = P(B)$   
  
**❷** Si  $A$  est quasi-certain alors :  
 pour tout événement  $B$ ,     $P(A \cap B) = P(B)$     et     $P(A \cup B) = 1$

**Démonstrations :**

- ❶** On suppose  $P(A) = 0$ ,  
 On a  $A \cap B \subset A$  donc  $0 \leq P(A \cap B) \leq P(A)$  donc  $P(A \cap B) = 0$   
 et  $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0 + P(B) - 0$  donc  $P(A \cup B) = P(B)$
- ❷** On suppose  $P(A) = 1$ ,  
 On a  $A \subset A \cup B$  donc  $P(A) \leq P(A \cup B) \leq 1$  donc  $P(A \cup B) = 1$   
 et  $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 1 + P(B) - 1$  donc  $P(A \cap B) = P(B)$

**4.1.6 Formule des probabilités totales (Version 1)**

**❶** Si  $(A_k)_{1 \leq k \leq n}$  est un système complet d'événements alors pour tout événement  $B$  on a :

$$P(B) = \sum_{k=1}^n P(A_k \cap B)$$

**❷** Si  $(A_n)_{n \in \mathbb{N}}$  est un système quasi-complet d'événements alors pour tout événement  $B$  on a :

$$\left( \text{la série } \sum_{n \geq 0} P(A_n \cap B) \text{ est convergente et} \right) \quad P(B) = \sum_{n=0}^{+\infty} P(A_n \cap B)$$

**Remarques :**

- Le théorème **❷** est aussi vrai avec un système complet.

- On sait que lorsque une suite d'événements deux à deux incompatibles alors la série  $\sum_{n \geq 0} P(A_n)$  converge.

Ce qui permet dans le théorème précédent d'oublier de préciser :  $\sum_{n \geq 0} P(A_n \cap B)$  est convergent.

**Démonstrations :**

- ❶  $(A_k)_{1 \leq k \leq n}$  est un système complet d'événements et  $B$  est un événement :

$$\begin{aligned}
 P(B) &= P(B \cap \Omega) \\
 &= P\left(B \cap \left(\bigcup_{k=1}^n A_k\right)\right) && \text{car } (A_k)_{1 \leq k \leq n} \text{ est un système complet} \\
 &= P\left(\bigcup_{k=1}^n (A_k \cap B)\right) \\
 &= \sum_{k=1}^n P(A_k \cap B) && \text{car les } (A_k \cap B) \text{ sont 2 à 2 incompatibles } (\sigma\text{-additivité}).
 \end{aligned}$$

- ❷  $(A_n)_{n \in \mathbb{N}}$  est un système quasi-complet d'événements et  $B$  est un événement :

$$\begin{aligned}
 P(B) &= P\left(B \cap \left(\bigcup_{n=0}^{+\infty} A_n\right)\right) && \text{car } P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1 \text{ (voir la proposition précédente).} \\
 &= P\left(\bigcup_{n=0}^{+\infty} (A_n \cap B)\right) \\
 &= \sum_{n=0}^{+\infty} P(A_n \cap B) && \text{car les } (A_n \cap B) \text{ sont 2 à 2 incompatibles. } (\sigma\text{-additivité})
 \end{aligned}$$

## 4.2 Conditionnement et indépendance.

On se place dans un espace probabilisé  $(\Omega, \mathcal{F}, P)$

### 4.2.1 Définition.

**Définition.**

Soient  $A$  et  $B$  deux événements tels que  $P(A) \neq 0$ ,  
 On appelle probabilité de  $B$  sachant  $A$ , notée  $P_A(B)$ , le réel :  $P_A(B) = \frac{P(A \cap B)}{P(A)}$

**Théorème.**

Quel que soit  $A \in \mathcal{F}$  tel que  $P(A) \neq 0$ , l'application  $P_A$  est une probabilité sur  $(\Omega, \mathcal{F})$ .

**Démonstration :** (non faite)

Il s'agit ici de démontrer que :

- ❶ Pour tout événement  $B$ ,  $0 \leq P_A(B) \leq 1$ ,
- ❷  $P_A(\Omega) = 1$

- ❸ lorsque  $(B_n)$  est une suite d'événements deux à deux incompatibles alors  $P_A\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} P_A(B_n)$

**Remarques :**

- $P_A$  est appelée "probabilité conditionnelle sachant  $A$ ".
- Calculer une probabilité conditionnelle revient à se placer dans la situation où l'événement  $A$  est réalisé.

### 4.2.2 Formule des probabilités composées

**Théorème :** (*Formule des probabilités composées*)

Soient  $n$  un entier supérieur à 2 et  $(A_k)_{1 \leq k \leq n}$  une liste d'événements vérifiant  $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$ ,

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

**Démonstration :**

On utilise les conventions :  $\bigcap_{i=1}^0 A_i = \Omega$  et  $\mathbb{P}_\Omega(A_1) = \mathbb{P}(A_1)$ .

$$\begin{aligned} P(A_1) \times P_{A_1}(A_2) \times \dots \times P_{A_1 \cap \dots \cap A_{n-1}}(A_n) &= \prod_{k=1}^n \mathbb{P}_{\bigcap_{i=1}^{k-1} A_i}(A_k) \\ &= \prod_{k=1}^n \frac{\mathbb{P}\left(\bigcap_{i=1}^k A_i\right)}{\mathbb{P}\left(\bigcap_{i=1}^{k-1} A_i\right)} \\ &= \mathbb{P}\left(\bigcap_{k=1}^n A_k\right) \quad (\text{Produit télescopique}) \end{aligned}$$

### 4.2.3 Formule des probabilités totales (Version 2)

**Théorème :**

❶ Si  $(A_k)_{1 \leq k \leq n}$  est un système complet d'événements *non négligeables pour P* (i.e : pour tout  $k$ ,  $P(A_k) \neq 0$ ) alors pour tout événement  $B$  on a :  $P(B) = \sum_{k=1}^n P(A_k) \cdot P_{A_k}(B)$

❷ Si  $(A_n)_{n \in \mathbb{N}}$  est un système quasi-complet d'événements *non négligeables pour P* (i.e : pour tout  $n$ ,  $P(A_n) \neq 0$ ) alors pour tout événement  $B$  on a :

$$\left( \text{la série } \sum_{n \geq 0} P(A_n) \cdot P_{A_n}(B) \text{ est convergente et} \right) \quad P(B) = \sum_{n=0}^{+\infty} P(A_n) \cdot P_{A_n}(B)$$

**En effet :** (*Immédiat à partir de la version 1 et la définition d'une probabilité conditionnelle*)

**Remarques :**

- Le théorème ❷ est vrai avec un système complet ou un système quasi-complet d'événements.
- On pourra ne pas vérifier  $P(A_n) \neq 0$  et quand  $P(A_n) = 0$ , on pose  $P(A_n) \cdot P_{A_n}(B) = 0$ .
- Ici aussi on peut oublier de préciser que la série converge.

### 4.2.4 Formule de Bayes

**Proposition :**

Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements,  
Si  $(A_n)_{n \in \mathbb{N}}$  est un système quasi-complet d'événements ( tel que : pour tout  $n$ ,  $P(A_n) \neq 0$ )  
alors pour tout événement  $B$  (tel que  $P(B) \neq 0$ ) et pour tout  $k \in \mathbb{N}$  on a :  $P_B(A_k) = \frac{P(A_k) \cdot P_{A_k}(B)}{\sum_{n=0}^{+\infty} P(A_n) \cdot P_{A_n}(B)}$

**Remarques :**

- En pratique il est souvent plus simple de redémontrer cette formule.
- Ce théorème est vrai pour un système complet et pour un système quasi-complet d'événements.
- Ce théorème est vrai avec un système complet d'événements fini. (*Théorème vu en première année*).

### 4.2.5 Indépendance

#### Définition :

Soient  $A$  et  $B$  deux événements,

Dire que  $A$  et  $B$  sont **indépendants** signifie que :  $P(A \cap B) = P(A)P(B)$

#### Propositions :

Soient  $A$  et  $B$  deux événements,

❶ Si  $P(B) \neq 0$  alors  $A$  et  $B$  sont indépendants si, et seulement si,  $P_B(A) = P(A)$ .

❷ Si  $A$  et  $B$  sont indépendants alors  $A$  et  $\bar{B}$  sont indépendants,  $\bar{A}$  et  $B$  sont indépendants et  $\bar{A}$  et  $\bar{B}$  sont indépendants.

#### Démonstrations :

❶ Immédiat avec la définition d'une probabilité conditionnelle.

❷ Pour montrer ces trois équivalences, il suffit de montrer l'implication suivante :

Quels que soient  $M$  et  $N$  deux événements de  $\Omega$ ,

Si  $M$  et  $N$  sont indépendants alors  $M$  et  $\bar{N}$  sont indépendants.

Prenons deux événements  $M$  et  $N$  deux événements indépendants de  $\Omega$ ,

On sait que  $\mathbb{P}(M) = \mathbb{P}(M \cap N) + \mathbb{P}(M \cap \bar{N})$

et comme  $M$  et  $N$  sont indépendants il vient :  $\mathbb{P}(M) = \mathbb{P}(M) \times \mathbb{P}(N) + \mathbb{P}(M \cap \bar{N})$ , puis :

$$\begin{aligned} \mathbb{P}(M \cap \bar{N}) &= \mathbb{P}(M) - \mathbb{P}(M) \times \mathbb{P}(N) \\ &= \mathbb{P}(M)(1 - \mathbb{P}(N)) \\ &= \mathbb{P}(M)\mathbb{P}(\bar{N}) \end{aligned}$$

**Remarque :** Dire que  $(A_n)$  est une suite d'événements **2 à 2 indépendants** signifie que

$$\forall (i, j) \in I^2, \quad i \neq j \implies P(A_i \cap A_j) = P(A_i)P(A_j)$$

#### Définition :

Soit  $(A_n)_{n \in I}$  une suite finie ou dénombrable d'événements.

Dire que les  $A_n$  sont **mutuellement indépendants** signifie que :

$$\text{pour toute partie finie } J \text{ de } I, \quad P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

#### Autrement dit :

Pour tout entier  $k \geq 2$ , et  $k$  indices  $i_1 < \dots < i_k$  on a :  $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times \dots \times P(A_{i_k})$ .

**Remarque :** si les  $A_n$  sont mutuellement indépendants alors les  $A_n$  sont 2 à 2 indépendants, *la réciproque n'est pas toujours vraie.*

#### Proposition :

Soient  $(A_n)_{n \in I}$  une suite finie ou dénombrable d'événements.

et  $(B_n)_{n \in I}$  une suite d'événements vérifiant :  $\forall n \in I, B_n = A_n$  ou  $B_n = \bar{A}_n$

❶ Si les  $A_n$  sont 2 à 2 indépendants alors les  $B_n$  sont 2 à 2 indépendants.

❷ Si les  $A_n$  sont mutuellement indépendants alors les  $B_n$  sont mutuellement indépendants.

**Autrement dit :** Si dans une suite d'événements deux à deux (resp. mutuellement) indépendants on remplace un ou plusieurs événements par leur contraire, on obtient toujours une suite d'événements deux à deux (resp. mutuellement) indépendants.

### 4.3 Variables aléatoires réelles.

On considère ici un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .

#### 4.3.1 Définition.

##### Définition

Une variable aléatoire réelle est une application de  $\Omega$  dans  $\mathbb{R}$  vérifiant :  
 Pour tout  $a \in \mathbb{R}$ ,  $[X \leq a]$  est un événement.

##### Remarques :

- Comme pour les tribus vous ne devriez pas être interrogés sur cette définition.
- On note  $[X \leq a]$  l'événement  $\{\omega \in \Omega \mid X(\omega) \leq a\}$
- L'ensemble des valeurs prises par  $X$  est  $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$
- Pour tout  $(a, b) \in \mathbb{R}^2$ ,  $[X \geq a]$ ,  $[a \leq X \leq b]$  sont des événements.

**Proposition :** Pour tout  $I$  intervalle de  $\mathbb{R}$ ,  $[X \in I]$  est un événement.

#### 4.3.2 Fonction de répartition.

##### Définition

Soit  $X$  une variable aléatoire de  $(\Omega, \mathcal{F}, P)$ ,  
 On appelle fonction de répartition de  $X$  la fonction  $F : \mathbb{R} \rightarrow [0, 1]$   
 $x \mapsto \mathbb{P}([X \leq x])$

##### Remarques :

- Quand plusieurs variables aléatoires sont présentes on note  $F_X : \mathbb{R} \rightarrow [0, 1]$   
 $x \mapsto \mathbb{P}([X \leq x])$
- C'est à rapprocher de la notion de fréquences cumulées en statistiques.

##### Proposition :

Pour tout  $(a, b) \in \mathbb{R}^2$  tel que  $a \leq b$  on a :

$$\mathbb{P}(X \in ]a, b]) = F(b) - F(a)$$

**En effet :**  $(X \leq b)$  est l'union disjointe de  $(X \leq a)$  et  $(a < X \leq b)$  donc  $\mathbb{P}(X \leq b) = \mathbb{P}(X \leq a) + \mathbb{P}(a < X \leq b)$

##### Proposition :

Soit  $X$  une variable aléatoire de fonction de répartition  $F$  alors :

- ❶  $F$  est une fonction croissante sur  $\mathbb{R}$ .
- ❷  $F$  est partout continue à droite.
- ❸  $\lim_{x \rightarrow -\infty} F(x) = 0$  et  $\lim_{x \rightarrow +\infty} F(x) = 1$

##### Démonstration :

- ❶ Si  $x_1 < x_2$  alors  $(X \leq x_1) \subset (X \leq x_2)$  donc  $\mathbb{P}(X \leq x_1) \leq \mathbb{P}(X \leq x_2)$
- ❷ et ❸ admis.

#### 4.3.3 Indépendance de variables aléatoires.

##### Indépendance de deux variables aléatoires.

##### Définition (définition du programme)

Soient  $X$  et  $Y$  deux variables aléatoires sur le même espace probabilisé  $(\Omega, \mathcal{F}, P)$ ,  
 Dire que  $X$  et  $Y$  sont indépendantes signifie que :  
 quel que soit le couple  $(I, J)$  d'intervalles de  $\mathbb{R}$ ,  $P((X \in I) \cap (Y \in J)) = P(X \in I) \times P(Y \in J)$

##### Théorème (caractérisation)

Soient  $X$  et  $Y$  deux variables aléatoires de  $(\Omega, \mathcal{F}, P)$ ,  
 $X$  et  $Y$  sont indépendantes si, et seulement si,  $\forall (x, y) \in \mathbb{R}^2$ ,  $P((X \leq x) \cap (Y \leq y)) = P(X \leq x) \times P(Y \leq y)$

### Indépendance de $n$ variables aléatoires.

#### Définition

Soient  $n$  un entier supérieur ou égal à 2 et  $(X_k)_{1 \leq k \leq n}$  une liste de variables aléatoires de  $(\Omega, \mathcal{F}, P)$ , Dire que les  $X_k$  sont (*mutuellement*) indépendantes signifie que :

$$\text{quelle que soit la liste } (I_k)_{1 \leq k \leq n} \text{ d'intervalles de } \mathbb{R}, \quad P\left(\bigcap_{k=1}^n (X_k \in I_k)\right) = \prod_{k=1}^n P(X_k \in I_k)$$

#### Théorème (*caractérisation*)

Soient  $n$  un entier supérieur ou égal à 2 et  $(X_k)_{1 \leq k \leq n}$  une liste de variables aléatoires sur le même espace probabilisé  $(\Omega, \mathcal{F}, P)$ ,

Les  $X_k$  sont (*mutuellement*) indépendantes si, et seulement si, :

$$\text{quelle que soit la liste } (x_k)_{1 \leq k \leq n} \text{ de réels,} \quad P\left(\bigcap_{k=1}^n (X_k \leq x_k)\right) = \prod_{k=1}^n P(X_k \leq x_k)$$

#### Remarque :

On parle aussi de "liste mutuellement indépendante" (*pour "liste de variables aléatoires mutuellement indépendantes"*).

### Indépendance d'une suite de variables aléatoires.

#### Définition

Soit  $(X_n)_{n \in \mathbb{N}}$  une liste de variables aléatoires de  $(\Omega, \mathcal{F}, P)$ ,

Dire que les variables de  $(X_n)_{n \in \mathbb{N}}$  sont (*mutuellement*) indépendantes signifie que :

toute liste finie extraite de la suite est (*mutuellement*) indépendante.

#### 4.3.4 Propriétés de l'indépendance mutuelle.

#### Théorème :

Soit  $(X_k)_{1 \leq k \leq n}$  une liste de variables aléatoires de  $(\Omega, \mathcal{F}, P)$ ,

❶ Si  $X_1, \dots, X_n$  sont (*mutuellement*) indépendantes alors

toute sous-famille de  $(X_1, \dots, X_n)$  l'est aussi.

❷ (*Lemme des coalitions*)

Soient  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  et  $g : \mathbb{R}^{n-p} \rightarrow \mathbb{R}$  deux fonctions,

si  $X_1, \dots, X_p, \dots, X_n$  sont (*mutuellement*) indépendantes alors

$f(X_1, \dots, X_p)$  et  $g(X_{p+1}, \dots, X_n)$  sont indépendantes.

❸ Soient  $f_1, \dots, f_n$  des fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$ ,

si  $X_1, \dots, X_n$  sont (*mutuellement*) indépendantes alors

$f_1(X_1), \dots, f_n(X_n)$  sont (*mutuellement*) indépendantes.

## 4.4 Espérance, variance et moments d'une variable aléatoire réelle.

On considère ici un espace probabilisé  $(\Omega, \mathcal{F}, P)$ .

#### 4.4.1 Espérance.

*Un introduction un peu théorique.*

On admet l'existence d'une fonction  $E$  définie sur une partie de l'ensemble des variables aléatoires sur  $(\Omega, \mathcal{F}, P)$  et à valeurs dans  $\mathbb{R}$  qui possèdent les propriétés suivantes :

❶  $E(1) = 1$

❷ Si  $X$  est à valeurs positives ou nulles et  $X$  admet une espérance alors  $E(X) \geq 0$ .

❸ Si  $a$  et  $b$  sont deux réels et  $X$  et  $Y$  deux variables aléatoires possédant une espérance alors

$$aX + bY \text{ admet une espérance et } E(aX + bY) = aE(X) + bE(Y)$$

#### Remarques :

- On ne vous posera pas de question sur cette introduction.
- Retenir que certaines variables aléatoires n'ont pas d'espérance.
- Une variable est dite **centrée** lorsque son espérance est nulle.
- Si  $X$  est une grandeur physique alors  $E(X)$  et  $X$  ont la même dimension.

### Définitions

#### Situation 1 : Variable discrète finie.

Toutes les variables aléatoires finies admettent une espérance :

Pour  $X$  une variable aléatoire avec  $X(\Omega) = \{x_k \mid k \in \llbracket 1; n \rrbracket\}$ ,  $E(X) = \sum_{k=1}^n x_k \mathbb{P}(X = x_k)$

#### Situation 2 : Variable discrète dénombrable.

Soit  $X$  un variable aléatoire avec  $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ ,

$X$  admet une espérance si, et seulement si,  $\sum_{n \in \mathbb{N}} x_n \mathbb{P}(X = x_n)$  est absolument convergente.

et alors  $E(X) = \sum_{n=0}^{+\infty} x_n \mathbb{P}(X = x_n)$

#### Situation 3 : Variable à densité.

Soit  $X$  un variable aléatoire réelle de densité  $f$ ,

$X$  admet une espérance si, et seulement si,  $\int_{-\infty}^{+\infty} x f(x) dx$  est absolument convergente.

et alors  $E(X) = \int_{-\infty}^{+\infty} x f(x) dx$

### Propriétés.

#### Proposition (Généralisation de la linéarité)

Soient  $n \in \mathbb{N}^*$ ,  $(X_k)_{1 \leq k \leq n}$  une liste de variables aléatoires sur un même espace probabilisé et  $(a_k)_{1 \leq k \leq n}$  une liste de réels.

Si les  $X_k$  possèdent toutes une espérance alors  $\sum_{k=1}^n a_k X_k$  admet une espérance et

$$E\left(\sum_{k=1}^n a_k X_k\right) = \sum_{k=1}^n a_k E(X_k)$$

**Démonstration :** *Raisonnement par récurrence sur  $n$ .*

#### Proposition (Croissance de l'espérance)

$$\text{Si } X \leq Y \quad \text{alors} \quad E(X) \leq E(Y)$$

**En effet :**  $X \leq Y$  entraîne  $Y - X \geq 0$  donc  $E(Y - X) \geq 0$  et avec la linéarité il vient  $E(Y) - E(X) \geq 0$ .

**Théorèmes de transfert.**

**Théorème.** (*Variable discrète*)

Soient  $X$  une variable aléatoire réelle discrète et  $f$  une fonction définie sur  $X(\Omega)$ ,

**Dans le cas fini.** on note :  $n$  le cardinal de  $X(\Omega)$  et  $X(\Omega) = \{x_k \mid k \in \llbracket 1; n \rrbracket\}$ ,

l'espérance de  $f(X)$  vérifie :  $E(f(X)) = \sum_{k=1}^n f(x_k) P(X = x_k)$ .

**Dans le cas dénombrable.** on note :  $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ , ( avec les  $(x_n)$  2 à 2 distincts )

$f(X)$  admet une espérance si, et seulement si,  $\sum_{n \in \mathbb{N}} f(x_n) P(X = x_n)$  est absolument convergente.

et alors l'espérance de  $f(X)$  vérifie :  $E(f(X)) = \sum_{n=0}^{+\infty} f(x_n) P(X = x_n)$ .

**Théorème** (*Variable à densité*)

Soient  $X$  une variable aléatoire réelle de densité  $f$ , telle que  $X$  est à valeurs dans  $]a, b[$   
( $-\infty \leq a < b \leq +\infty$ )

et  $\varphi : ]a, b[ \rightarrow \mathbb{R}$  une fonction continue sauf éventuellement en un nombre fini de points.

$\varphi(X)$  admet une espérance si, et seulement si,  $\int_a^b \varphi(t) f(t) dt$  est absolument convergente.

et alors :  $E(\varphi(X)) = \int_a^b \varphi(t) f(t) dt$

**4.4.2 Variance.**

**Définition.**

**Définition**

Soit  $X$  une variable aléatoire admettant une espérance  $E(X)$ .  
 Lorsque la variable  $(X - E(X))^2$  admet une espérance on dit que  $X$  admet une variance et on définit sa variance par

$$V(X) = E((X - E(X))^2)$$

**Remarques :**

- Si une variable aléatoire  $X$  admet une variance alors nécessairement elle admet une espérance.
- Quelle que soit la variable aléatoire  $X$  admettant une variance, on a :  $V(X) \geq 0$ .
- Si  $V(X) = 0$  alors  $\mathbb{P}([X = m]) = 1$  (où  $m = E(X)$ ), *On dit que  $X$  est une variable quasi-certaine.*
- Dire qu'une variable aléatoire est **réduite** signifie que sa variance est égale à 1.

**4.4.3 Formule de Koenig-Huygens.**

**Théorème :**

Soit  $X$  est une variable aléatoire réelle,  
 Si  $X$  et  $X^2$  admettent une espérance alors  $X$  admet une variance et

$$V(X) = E(X^2) - (E(X))^2$$

**4.4.4 Variance de  $aX + b$  :**

**Théorème :**

Pour toute variable aléatoire  $X$  admettant une variance et  $a$  et  $b$  deux réels, on a :

$$aX + b \text{ admet une variance et } V(aX + b) = a^2 V(X)$$

#### 4.4.5 Ecart-type.

**Définition.**

Quand  $X$  admet une variance, l'écart-type d'une variable aléatoire  $X$  est le réel :  $\sigma_X = \sqrt{V(X)}$

**Remarque :** Si  $X$  est une grandeur physique alors  $\sigma_X$  et  $X$  ont la même dimension.

**Définition :**

Pour  $X$  une variable aléatoire admettant une variance  $V(X) \neq 0$ ,  
 $X^* = \frac{X - \mathbb{E}(X)}{\sigma_X}$  est appelée **variable aléatoire centrée réduite** associée à  $X$ .

**Remarque :**  $X^*$  est sans dimension.

#### 4.4.6 Moments d'ordres supérieurs.

**Définition.**

Pour  $X$  une variable aléatoire et  $n$  un entier naturel,  
 Quand  $X^n$  admet une espérance on appelle  $E(X^n)$  le moment de  $X$  d'ordre  $n$ .

#### 4.4.7 Avec des variables aléatoires indépendantes.

**Théorème :**

Soient  $X$  et  $Y$  deux variables aléatoires sur  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

① Si  $X$  et  $Y$  sont indépendantes et admettent des espérances alors  
 $XY$  admet une espérance et  $E(XY) = E(X)E(Y)$ .

② Si  $X$  et  $Y$  sont indépendantes et admettent des variances alors  
 $X + Y$  admet une variance et  $V(X + Y) = V(X) + V(Y)$

**Démonstration.** (*admis*)

**Généralisation :**

Soit  $(X_k)_{1 < k \leq n}$  une suite de variables aléatoires sur  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

① Si les  $X_k$  sont indépendantes et admettent des espérances alors

$$\prod_{k=1}^n X_k \text{ admet une espérance et } E\left(\prod_{k=1}^n X_k\right) = \prod_{k=1}^n E(X_k)$$

② Si les  $X_k$  sont indépendantes et admettent des variances alors

$$\sum_{k=1}^n X_k \text{ admet une variance et } V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k)$$

# Variables aléatoires discrètes

## Plan du chapitre

5.1	Variables aléatoires discrètes. . . . .	<b>50</b>
5.1.1	Définition . . . . .	50
5.1.2	Système complet d'événements associé à $X$ . . . . .	51
5.1.3	Loi de probabilité . . . . .	51
5.1.4	Espérance et variance d'une variable aléatoire réelle discrète. . . . .	51
5.2	Indépendance. . . . .	<b>55</b>
5.2.1	Caractérisations. . . . .	55
5.2.2	Indépendance d'une suite de variables aléatoires. . . . .	55
5.2.3	Propriétés de l'indépendance mutuelle . . . . .	55
5.2.4	Théorèmes. . . . .	55
5.2.5	Stabilité des lois de Poisson. . . . .	56
5.2.6	Complément. . . . .	57
5.3	Lois finies usuelles. . . . .	<b>57</b>
5.3.1	Loi certaine. . . . .	57
5.3.2	Loi de Bernoulli. . . . .	57
5.3.3	Loi uniforme. . . . .	58
5.3.4	Loi binomiale. . . . .	59
5.4	Lois discrètes infinies usuelles. . . . .	<b>61</b>
5.4.1	Lois géométriques . . . . .	61
5.4.2	Loi de Poisson . . . . .	63

## 5.1 Variables aléatoires discrètes.

On considère un espace probabilisé  $(\Omega; \mathcal{F}, \mathbb{P})$  quelconque.

### 5.1.1 Définition

**Définition.**

Soit  $X$  une variable aléatoire sur  $(\Omega; \mathcal{F}, \mathbb{P})$   
 Dire que  $X$  est une variable aléatoire discrète signifie que  $X(\Omega)$  est un ensemble de réels fini ou dénombrable.

**Remarques :**

- Lorsque  $X$  est une variable aléatoire discrète, on note  $X(\Omega) = \{ x_i \mid i \in I \}$  avec  $I = \overbrace{[1; n]}^{\text{fini}}$  ou  $I = \overbrace{\mathbb{N}}^{\text{dénombrable}}$   
 ( et  $\forall (i, j) \in I^2, \quad i \neq j \implies x_i \neq x_j$  )
- *Extrait du programme :* Une variable aléatoire est dite discrète si l'ensemble  $X(\Omega)$  de ses valeurs est inclus dans un sous-ensemble  $\mathcal{N}$  de  $\mathbb{R}$  indexé par une partie de  $\mathbb{N}$ .
- *Comment bien définir une variable aléatoire ?*

**En pratique 1 :** On nous donne  $X(\Omega) = \{ x_i \mid i \in I \}$  et  $P(X = x_i) = p_i$

Vérifier si  $X$  est bien définie revient à :

- ❶ Vérifier que les  $x_i$  sont 2 à 2 distincts.
- ❷ que  $\forall i \in I, p_i \geq 0$
- ❸ et que  $\sum_{i \in I} p_i$  vaut 1.

Exemples : Feuille Cours\_5

**Proposition.**

Si  $(x_n)_{n \in \mathbb{N}}$  est une suite de réels deux à deux distincts et  $(p_n)_{n \in \mathbb{N}}$  une suite de réels positifs tels que  $\sum p_n$  converge et a pour somme 1, alors il existe une variable aléatoire réelle discrète  $X$  sur  $(\Omega; \mathcal{F}, \mathbb{P})$  vérifiant  $P(X = x_n) = p_n$  pour tout entier naturel  $n$ .

**En pratique 2 :** (On décrit une expérience aléatoire)

On décrit une expérience aléatoire avec  $\Omega$  quelconque. On note  $E$  la partie de  $\Omega$  où  $X$  n'est pas définie.

$X$  est une variable aléatoire bien définie lorsque  $P(E) = 0$ .

Dans ce cas là on confond  $\Omega$  et  $\Omega \setminus E$

Exemples : Feuille Cours\_5

- *Extrait du programme :* On tolère qu'une variable aléatoire issue d'une expérience aléatoire puisse ne pas être définie sur un événement de probabilité nulle.

**5.1.2 Système complet d'événements associé à  $X$**

**Théorème.**

Soit  $X$  une variable aléatoire sur  $(\Omega; \mathcal{F}, \mathbb{P})$   
 Si  $X$  est une variable aléatoire discrète avec  $X(\Omega) = \{ x_i \mid i \in I \}$  ( avec  $i \neq j \implies x_i \neq x_j$  )  
 alors  $([X = x_i])_{i \in I}$  est un système *quasi-complet* d'événements.

**5.1.3 Loi de probabilité**

**Définition.**

Soit  $X$  est une variable aléatoire discrète.  
 L'application  $X(\Omega) \longrightarrow \mathbb{R}$  est appelée **loi de probabilité** de  $X$ .  
 $x \longmapsto \mathbb{P}(X = x)$

**Définition. Loi conditionnelle.**

Soit  $X$  est une variable aléatoire discrète et  $A$  un événement de probabilité non nulle.  
 L'application  $X(\Omega) \longrightarrow \mathbb{R}$  est appelée **loi de probabilité de  $X$  sachant  $A$** .  
 $x \longmapsto \mathbb{P}_A(X = x)$

**5.1.4 Espérance et variance d'une variable aléatoire réelle discrète.**

**Espérance**

**Définition** (Espérance mathématique d'une variable aléatoire discrète.)

- Lorsque  $X(\Omega)$  est fini, en notant  $n = \text{card}(X(\Omega))$  et  $X(\Omega) = \{x_k \mid k \in \llbracket 1; n \rrbracket\}$   
 on appelle espérance mathématique de  $X$  le réel :  $E(X) = \sum_{k=1}^n x_k P(X = x_k)$
- Lorsque  $X(\Omega)$  est dénombrable, en notant  $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$  ( avec les  $(x_n)$  2 à 2 distincts ),  
 $X$  admet une espérance si, et seulement si, la série  $\sum_{n \geq 0} x_n P(X = x_n)$  est absolument convergente.  
 alors l'espérance mathématique de  $X$  est le réel :  $E(X) = \sum_{n=0}^{+\infty} x_n P(X = x_n)$

**En pratique.**

- *Cas fini.*

" $X(\Omega)$  est fini donc  $X$  admet une espérance et  $E(X) = \sum_{k=1}^n x_k P(X = x_k)$ "

- *Cas dénombrable.*

❶ " $X$  admet une espérance si et seulement si,  $\sum_{n \geq 0} x_n P(X = x_n)$  est absolument convergente."

On montre que :  $\sum_{n \geq 0} |x_n| P(X = x_n)$  est convergente et on dit que  $X$  admet une espérance.

❷ On calcule l'espérance  $E(X) = \sum_{n=0}^{+\infty} x_n P(X = x_n)$

**Exemples :** Feuille Cours\_5

**Remarques :**

- La convergence absolue est nécessaire pour que la somme  $\sum_{n=0}^{+\infty} x_n P(X = x_n)$  ne dépende pas du choix des  $(x_n)$  pour décrire  $X(\Omega)$ . (*Voir le théorème 1.5.2 sur les séries*).
- Retenir que certaines variables aléatoires n'ont pas d'espérance. (*Exemple : feuille Cours\_5*)
- Une variable aléatoire est dite **centrée** lorsque son espérance est nulle.

**Théorème :** (*linéarité*)

Soient  $X$  et  $Y$  deux variables aléatoires d'un même espace probabilisé,  
Si  $X$  et  $Y$  admettent chacune une espérance alors  
pour tout réel  $a$  et  $b$ ,  $aX + bY$  admet une espérance et  $E(aX + bY) = aE(X) + bE(Y)$ .

**Proposition :** (*Croissance*)

Soient  $X$  et  $Y$  deux variables aléatoires d'un même espace probabilisé,  
❶ Si  $X \geq 0$  et  $X$  admet une espérance alors  $E(X) \geq 0$ .  
❷ Si  $X \leq Y$  et si  $X$  et  $Y$  admettent chacune une espérance alors  $E(X) \leq E(Y)$

**Généralisation**

Soient  $n \in \mathbb{N}^*$ ,  $(X_k)_{1 \leq k \leq n}$  une liste de variables aléatoires sur un même espace probabilisé  
et  $(a_k)_{1 \leq k \leq n}$  une liste de réels.

Si les  $X_k$  possèdent toutes une espérance alors  $E\left(\sum_{k=1}^n a_k X_k\right) = \sum_{k=1}^n a_k E(X_k)$

**Théorème de transfert.**

**Théorème.** (*Théorème de transfert*)

Soient  $X$  une variable aléatoire réelle discrète et  $f$  une fonction définie sur  $X(\Omega)$ ,

**Dans le cas fini.** on note :  $n$  le cardinal de  $X(\Omega)$  et  $X(\Omega) = \{x_k \mid k \in \llbracket 1; n \rrbracket\}$ ,

$$\text{l'espérance de } f(X) \text{ vérifie : } E(f(X)) = \sum_{k=1}^n f(x_k) P(X = x_k).$$

**Dans le cas dénombrable.** on note :  $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ , ( avec les  $(x_n)$  2 à 2 distincts )

$f(X)$  admet une espérance si, et seulement si,  $\sum_{n \in \mathbb{N}} f(x_n) P(X = x_n)$  est absolument convergente.

$$\text{et alors l'espérance de } f(X) \text{ vérifie : } E(f(X)) = \sum_{n=0}^{+\infty} f(x_n) P(X = x_n).$$

### Rédaction.

- *Cas fini.*

$X(\Omega)$  est fini et  $f$  est définie sur  $X(\Omega)$  donc (théorème de transfert)  $f(X)$  admet une espérance

$$\text{et } E(f(X)) = \sum_{k=1}^n f(x_k) P(X = x_k)$$

- *Cas dénombrable.*

❶ ( $f$  est définie sur  $X(\Omega)$ )

❷  $f(X)$  admet une espérance si et seulement si,  $\sum_{n \geq 0} f(x_n) P(X = x_n)$  est absolument convergente.

On montre que :  $\sum_{n \geq 0} |f(x_n)| P(X = x_n)$  est convergente et on dit que  $f(X)$  admet une espérance.

❸ On calcule l'espérance  $E(f(X)) = \sum_{n=0}^{+\infty} f(x_n) P(X = x_n)$

**Exemples :** feuille Cours\_5

**Remarque :**  $X$  peut admettre une espérance alors que  $f(X)$  non.

### Variance d'une variable aléatoire discrète.

#### Définition

Soit  $X$  une variable aléatoire discrète.

Dire que  $X$  admet une variance signifie que  $X$  et  $(X - E(X))^2$  admettent une espérance.

on définit alors sa variance par :  $V(X) = E((X - E(X))^2)$

#### Remarques :

- Retenir que certaines variables aléatoires n'ont pas de variance.
- Dire qu'une variable aléatoire est **réduite** signifie que sa variance est égale à 1.

### Formule de Kœnig-Huygens.

**Lemme.** (*Exercice*)

Soit  $X$  une variable aléatoire discrète,

Si  $X^2$  admet une espérance alors  $X$  admet une espérance.

**Démonstration.** (*Voir feuille cours\_5\_2*)

**Théorème :** (*Formule de Kœnig-Huygens*).

Soit  $X$  est une variable aléatoire discrète,

$X$  admet une variance si, et seulement si,  $X^2$  admet une espérance

et alors

$$V(X) = E(X^2) - (E(X))^2$$

**Démonstration.** *Faite au tableau.*

**Rédaction.**

**Cas fini.**

" $X(\Omega)$  est fini donc  $X$  admet une variance et  $E(X) = \dots, E(X^2) = \dots$ , et ainsi  $V(X) = E(X^2) - E(X)^2$ "

**Cas dénombrable.**

❶ " $X^2$  admet une espérance si, et seulement si,  $\sum_{n \geq 0} x_n^2 P(X = x_n)$  est absolument convergente."

On montre que :  $\sum_{n \geq 0} |x_n^2| P(X = x_n)$  est convergente et on dit que  $X$  admet une variance.

❷ On calcule les deux espérances  $E(X) = \sum_{n=0}^{+\infty} x_n P(X = x_n)$  et  $E(X^2) = \sum_{n=0}^{+\infty} x_n^2 P(X = x_n)$

❸ On applique le théorème de Koenig-Huygens :  $X$  admet une variance et  $V(X) = E(X^2) - E(X)^2$

**Exemples :** *Feuille Cours\_5\_2*

### Propriétés de la variance.

**Proposition :**

Quelle que soit la variable aléatoire  $X$  admettant une variance,

❶  $V(X) \geq 0$ .

❷ Si  $V(X) = 0$  alors  $\mathbb{P}([X = m]) = 1$  (où  $m = E(X)$ ),

*On dit que  $X$  est une variable quasi-certaine.*

❸ Pour toute variable  $X$  admettant une variance et  $a$  et  $b$  deux réels, on a :

$$aX + b \text{ admet une variance et } V(aX + b) = a^2 V(X)$$

**Démonstrations.**

### Moments d'ordre supérieurs.

**Définition :**

Pour  $X$  une variable aléatoire et  $n$  un entier naturel,

Quand  $X^n$  admet une espérance on appelle  $E(X^n)$  le moment de  $X$  d'ordre  $n$ .

### Ecart-type.

**Définition :**

Quand  $X$  admet une variance, l'écart-type d'une variable aléatoire  $X$  est le réel :  $\sqrt{V(X)}$

On note souvent :  $\sigma = \sqrt{V(X)}$  ou  $\sigma_X = \sqrt{V(X)}$ ,

d'où la notation  $V(X) = \sigma^2$

**Définition :**

Pour  $X$  une variable aléatoire admettant une variance  $V(X) \neq 0$ ,

$X^* = \frac{X - \mathbb{E}(X)}{\sigma_X}$  est appelée **variable aléatoire centrée réduite** associée à  $X$ .

**Démonstration.** À montrer  $E(X^*) = 0$  et  $V(X^*) = 1$ .

## 5.2 Indépendance.

### 5.2.1 Caractérisations.

Revoir la définition générale de l'indépendance.

**Caractérisation. Indépendance de deux variables aléatoires discrètes.**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes,  
 Dire que  $X$  et  $Y$  sont indépendantes signifie que :  
 quel que soit le couple  $(x, y) \in \mathbb{R}^2$ ,  $\mathbb{P}((X = x) \cap (Y = y)) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y)$

**Caractérisation. Indépendance de  $n$  variables aléatoires discrètes.**

Soient  $n$  un entier supérieur ou égal à 2 et  $(X_k)_{1 \leq k \leq n}$  une liste (*finie*) de variables aléatoires discrètes,  
 Dire que les  $X_k$  sont (*mutuellement*) indépendantes signifie que :  
 quel que soit la liste de réels  $(x_k)_{1 \leq k \leq n}$ ,  $\mathbb{P}\left(\bigcap_{k=1}^n (X_k = x_k)\right) = \prod_{k=1}^n \mathbb{P}(X_k = x_k)$

### 5.2.2 Indépendance d'une suite de variables aléatoires.

(Déjà vu dans le chapitre Probabilité).

#### Définition

Soit  $(X_n)_{n \in \mathbb{N}}$  une liste de variables aléatoires de  $(\Omega, \mathcal{F}, P)$ ,  
 Dire que les variables de  $(X_n)_{n \in \mathbb{N}}$  sont (*mutuellement*) indépendantes signifie que :  
 toute liste finie extraite de la suite est (*mutuellement*) indépendante.

### 5.2.3 Propriétés de l'indépendance mutuelle

(Déjà vu dans le chapitre Probabilité).

#### Théorème :

Soit  $(X_k)_{1 \leq k \leq n}$  une liste de variables aléatoires de  $(\Omega, \mathcal{F}, P)$ ,

- ❶ Si  $X_1, \dots, X_n$  sont (*mutuellement*) indépendantes **alors**  
 toute sous-famille de  $(X_1, \dots, X_n)$  l'est aussi.
- ❷ (**Lemme des coalitions**)  
 Soient  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  et  $g : \mathbb{R}^{n-p} \rightarrow \mathbb{R}$  deux fonctions,  
 si  $X_1, \dots, X_p, \dots, X_n$  sont (*mutuellement*) indépendantes **alors**  
 $f(X_1, \dots, X_p)$  et  $g(X_{p+1}, \dots, X_n)$  sont indépendantes.
- ❸ Soient  $f_1, \dots, f_n$  des fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$ ,  
 si  $X_1, \dots, X_n$  sont (*mutuellement*) indépendantes **alors**  
 $f_1(X_1), \dots, f_n(X_n)$  sont (*mutuellement*) indépendantes.

### 5.2.4 Théorèmes.

#### Théorème :

Soient  $X$  et  $Y$  deux variables aléatoires discrètes sur  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

- ❶ Si  $X$  et  $Y$  sont indépendantes et admettent des espérances alors  
 $XY$  admet une espérance et  $E(XY) = E(X)E(Y)$ .
- ❷ Si  $X$  et  $Y$  sont indépendantes et admettent des variances alors  
 $X + Y$  admet une variance et  $V(X + Y) = V(X) + V(Y)$

On admet ❶ car il nous manque un théorème hors-programme, mais en admettant ❶ on peut démontrer ❷ .  
 Démonstration de ❷ :

**Généralisation :**

Soit  $(X_k)_{1 < k \leq n}$  une suite de variables aléatoires discrètes sur  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

- ❶ Si les  $X_k$  sont indépendantes et admettent des espérances alors
 
$$\prod_{k=1}^n X_k \text{ admet une espérance et } E\left(\prod_{k=1}^n X_k\right) = \prod_{k=1}^n E(X_k)$$
- ❷ Si les  $X_k$  sont indépendantes et admettent des variances alors
 
$$\sum_{k=1}^n X_k \text{ admet une variance et } V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k)$$

**Démonstration :**

**5.2.5 Stabilité des lois de Poisson.**

**Proposition.**

Soient  $\lambda_1, \lambda_2$  deux réels non nuls et  $X_1$  et  $X_2$  deux variables aléatoires (sur le même espace probabilisé),

Si  $\begin{cases} X_1 \hookrightarrow \mathcal{P}(\lambda_1), X_2 \hookrightarrow \mathcal{P}(\lambda_2) \\ \text{et } X_1, X_2 \text{ sont indépendantes} \end{cases}$  alors  $X_1 + X_2 \hookrightarrow \mathcal{P}(\lambda_1 + \lambda_2)$

**Démonstration :**

On note  $Z = X_1 + X_2$ , on sait que :  $X_1(\Omega) = \mathbb{N}$  et  $X_2(\Omega) = \mathbb{N}$  donc  $Z(\Omega) \subset \mathbb{N}$

en appliquant la formule des proba. totales avec le système complet :  $([X = k])_{k \in \mathbb{N}}$ , on obtient pour  $n \in \mathbb{N}$  :

$$\begin{aligned} P([X_1 + X_2 = n]) &= \sum_{k=0}^{+\infty} P([X_1 = k] \cap [X_1 + X_2 = n]) \\ &= \sum_{k=0}^{+\infty} P([X_1 = k] \cap [X_2 = n - k]) \\ &= \sum_{k=0}^n P([X_1 = k] \cap [X_2 = n - k]) \quad \text{car pour } k > n \quad [X_2 = n - k] = \emptyset \\ &= \sum_{k=0}^n P([X_1 = k]) \times P([X_2 = n - k]) \quad \text{car } X_1 \text{ et } X_2 \text{ sont indépendantes} \\ &= \sum_{k=0}^n \frac{\lambda_1^k}{k!} e^{-\lambda_1} \times \frac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2} \\ &= \frac{e^{-\lambda_1 - \lambda_2}}{n!} \sum_{k=0}^n \binom{n}{k} \lambda_1^k \lambda_2^{n-k} \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n \quad \text{Formule du binôme.} \end{aligned}$$

Remarque : Après ce calcul on peut affirmer que  $Z(\Omega) = \mathbb{N}$ .

$$X_1 + X_2 \hookrightarrow \mathcal{P}(\lambda_1 + \lambda_2)$$

**Généralisation .**

Soient  $(X_k)_{1 < k \leq N}$  une liste de variables aléatoires,  $(\lambda_k)_{1 < k \leq N}$  une liste de réels non nuls, (sur le même espace probabilisé)

on note :  $X = \sum_{k=1}^N X_k$  et  $\lambda = \sum_{k=1}^N \lambda_k$ .

Si les  $X_k$  sont mutuellement indépendantes et si  $\forall k \in \mathbb{N}, X_k \hookrightarrow \mathcal{P}(\lambda_k)$  alors  $X \hookrightarrow \mathcal{P}(\lambda)$

**Démonstration :**

Soit  $(X_n)$  une suite de variable aléatoire suivant tous une loi de Poisson, on note pour chaque  $n$ ,  $\lambda_n$  le paramètre de la loi de  $X_n$ .

Montrons par récurrence sur  $n$  que pour tout  $n \in \mathbb{N}^*$ ,  $Z_n = \underbrace{\sum_{k=1}^n X_k}_{\mathcal{P}(n)}$  suit la loi de Poisson de paramètre  $\sum_{k=1}^n \lambda_k$

- Pour  $n = 1$ ,

$X_1$  suit la loi de Poisson de paramètre  $\lambda_1$  et  $\sum_{k=1}^1 \lambda_k = \lambda_1$  donc  $\mathcal{P}(1)$  est vérifiée

- Soit  $n \in \mathbb{N}^*$  tel que  $\mathcal{P}(n)$  est vraie,

$$Z_{n+1} = \sum_{k=1}^{n+1} X_k = \sum_{k=1}^n X_k + X_{n+1}$$

or (*lemme de coalition*) les  $X_n$  sont *mutuellement* indépendantes donc  $\sum_{k=1}^n X_k$  et  $X_{n+1}$  sont indépendantes

et *hypothèse de récurrence*  $\sum_{k=1}^n X_k \hookrightarrow \mathcal{P}\left(\sum_{k=1}^n \lambda_k\right)$  et on a :  $X_{n+1} \hookrightarrow \mathcal{P}(\lambda_{n+1})$

on peut en déduire avec le résultat de la question 1) que :  $\underbrace{\sum_{k=1}^{n+1} X_k}_{\mathcal{P}(n+1)} \hookrightarrow \mathcal{P}\left(\sum_{k=1}^n \lambda_k + \lambda_{n+1}\right)$

En conclusion : pour tout  $n \in \mathbb{N}^*$ ,  $\sum_{k=1}^n X_k$  suit la loi de Poisson de paramètre  $\sum_{k=1}^n \lambda_k$

**5.2.6 Complément.**

**Théorème**

Soit  $(A_k)_{1 \leq k \leq n}$  une famille d'événements,  
 Les événements  $A_k$  sont mutuellement indépendants  
 si, et seulement si, les variables aléatoires  $\mathbb{1}_{A_k}$  sont indépendantes.

**5.3 Lois finies usuelles.**

**5.3.1 Loi certaine.**

Soit  $a$  un nombre réel.

Dire que  $X$  est la variable certaine égale à  $a$  signifie que  $X$  est la fonction constante  $X : \Omega \longrightarrow \mathbb{R}$   
 $\omega \longmapsto a$

On a :  $X(\Omega) = \{a\}$  et  $\mathbb{P}([X = a]) = 1, \mathbb{E}(X) = a \quad V(X) = 0.$

**Remarque :** Dire que  $X$  est "quasi-certaine égale à  $a$ " signifie que  $\mathbb{P}([X = a]) = 1$

**5.3.2 Loi de Bernoulli.**

**Définition :**

Soit  $p$  un réel de  $]0, 1[$ ,  
 dire que  $X$  suit une loi de Bernoulli de paramètre  $p$  signifie que :  $\begin{cases} X(\Omega) = \{0, 1\} \\ \mathbb{P}([X = 0]) = 1 - p \\ \mathbb{P}([X = 1]) = p \end{cases}$

On note :  $X \hookrightarrow \mathcal{B}(p)$  et on a :  $\mathbb{E}(X) = p$  et  $V(X) = p(1 - p)$

**En effet :**  $X(\Omega) = \{0, 1\}$  est fini donc  $X$  admet une espérance et une variance et

$$E(X) = 0 \times (1 - p) + 1 \times p \text{ donc } \boxed{E(X) = p} \quad \text{et} \quad E(X^2) = 0 \times (1 - p) + 1^2 \times p = p$$

donc (avec la formule Kœnig-Huygens)  $\boxed{V(X) = p - p^2 = p(1 - p)}$

**Remarque :** Si  $A$  est un événement de  $\Omega$  alors  $\mathbb{1}_A$  est une variable aléatoire et  $\mathbb{1}_A \hookrightarrow \mathcal{B}(\mathbb{P}(A))$

**Simulation avec une fonction Python :**

```
def bernoulli(p):
    if rd.random() <= p :
        return 1
    return 0
```

### 5.3.3 Loi uniforme.

**Définition :**

Soient  $a$  et  $b$  deux entiers relatifs tels que  $a \leq b$ ,  
 Dire que  $X$  suit une loi uniforme sur  $\llbracket a; b \rrbracket$  signifie que :  $\begin{cases} X(\Omega) = \llbracket a; b \rrbracket \\ \forall i \in \llbracket a; b \rrbracket, \mathbb{P}([X = i]) = \frac{1}{b - a + 1} \end{cases}$

On note :  $X \hookrightarrow \mathcal{U}(\llbracket a; b \rrbracket)$ , (*Attention : Ne pas confondre avec  $X \hookrightarrow \mathcal{U}([a; b])$  du cours sur les VAR à densité*)

**Proposition :**

si  $X \hookrightarrow \mathcal{U}(\llbracket a; b \rrbracket)$  alors  $E(X) = \frac{a + b}{2}$  et  $V(X) = \frac{(b - a)(b - a + 2)}{12}$   
 (la variance n'est pas à retenir)

**Démonstration.** (*A savoir refaire*)

❶ Commençons par montrer que pour  $X \hookrightarrow \mathcal{U}(\llbracket 0; n \rrbracket)$ ,  $E(X) = \frac{n}{2}$  et  $V(X) = \frac{n(n + 2)}{12}$ .

$$\begin{aligned} E(X) &= \sum_{k=0}^n \frac{k}{n + 1} & E(X^2) &= \sum_{k=0}^n \frac{k^2}{n + 1} & V(X) &= E(X^2) - E(X)^2 \\ &= \frac{1}{n + 1} \frac{n(n + 1)}{2} & &= \frac{1}{n + 1} \frac{n(n + 1)(2n + 1)}{6} & &= \frac{n(2n + 1)}{6} - \frac{n^2}{4} \\ &= \frac{n}{2} & &= \frac{n(2n + 1)}{6} & &= \frac{n}{2} \left( \frac{2n + 1}{3} - \frac{n}{2} \right) \\ & & & & &= \frac{n(n + 2)}{12} \end{aligned}$$

❷ Prenons maintenant une variable aléatoire  $X$  suivant la loi uniforme sur  $\llbracket a; b \rrbracket$ .

En notant  $Y = X - a$ , on a :  $Y \hookrightarrow \llbracket 0; b - a \rrbracket$

(En effet  $Y(\Omega) = \llbracket 0; b - a \rrbracket$  et toutes les valeurs sont équiprobables)

on a  $X = Y + a$  donc (linéarité)  $E(X) = E(Y) + a$  or  $E(Y) = \frac{b - a}{2}$  (en utilisant ❶) donc  $\boxed{E(X) = \frac{a + b}{2}}$

on a  $X = Y + a$  donc (comme  $V(\alpha X + \beta) = \alpha^2 V(X)$ )  $V(X) = V(Y)$

or (en utilisant ❶)  $V(Y) = \frac{(b - a)^2 + 2(b - a)}{12}$  donc  $\boxed{V(X) = \frac{(b - a)^2 + 2(b - a)}{12}}$

**Remarque :** pour  $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$ ,  $E(X) = \frac{n + 1}{2}$  et  $V(X) = \frac{n^2 - 1}{12}$ .

**Simulation avec une fonction Python (sans randint) :**

```
def uniforme(a, b):
    return floor( a+ (b-a+1)*rd.random() )           # floor importé de math
```

**Fonction de répartition**

Si  $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$  alors 
$$\begin{cases} \forall x < 1, & F_X(x) = 0 \\ \forall x \in [1, n], & F_X(x) = \frac{\lfloor x \rfloor}{n} \\ \forall x > n, & F_X(x) = 1 \end{cases}$$

**5.3.4 Loi binomiale.****Définition :**

Soit  $n \in \mathbb{N}^*$  et  $p$  un réel de  $]0, 1[$ ,  $X$  une variable aléatoire,  
Dire que  $X$  suit une loi binomiale de paramètres  $(n, p)$  signifie que :

$$\begin{cases} X(\Omega) = \llbracket 0; n \rrbracket \\ \forall k \in \llbracket 0; n \rrbracket, & \mathbb{P}([X = k]) = \binom{n}{k} p^k (1-p)^{n-k} \end{cases}$$

On note :  $X \hookrightarrow \mathcal{B}(n, p)$  et on a  $\mathbb{E}(X) = np$  et  $V(X) = np(1-p)$

**Démonstration.**

$X(\Omega)$  est fini donc  $X$  admet une espérance et  $E(X) = \sum_{k=0}^n kP(X = k)$  donc

$$\begin{aligned} E(X) &= \sum_{k=1}^n k \binom{n}{k} p^k q^{n-k} && \text{(On a enlevé le terme pour } k=0 \text{ car il est nul)} \\ &= \sum_{k=1}^n n \binom{n-1}{k-1} p^k q^{n-k} && \text{(formule d'absorption, du chef, du capitaine, ...)} \\ &= n \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} q^{n-k-1} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k q^{n-1-k} \\ &= np(p+q)^{n-1} \end{aligned}$$

$$\boxed{E(X) = np}$$

$X(\Omega)$  est fini donc  $X^2$  admet une espérance (théorème de transfert).

$$\begin{aligned} E(X^2) &= \sum_{k=0}^n k^2 P(X = k) \\ &= \sum_{k=2}^n k(k-1)P(X = k) + \sum_{k=0}^n kP(X = k) && \text{(les termes pour } k=0, 1 \text{ sont nuls)} \\ &= \sum_{k=2}^n k(k-1) \binom{n}{k} p^k q^{n-k} + np && \text{(on reconnaît l'espérance)} \\ &= \sum_{k=2}^n n(n-1) \binom{n-2}{k-2} p^k q^{n-k} + np && \text{(double formule du chef)} \\ &= n(n-1) \sum_{k=0}^{n-2} \binom{n-2}{k} p^{k+2} q^{n-k-2} + np && \text{(réindexation)} \end{aligned}$$

$$\begin{aligned}
&= n(n-1)p^2 \sum_{k=0}^{n-2} \binom{n-2}{k} p^k q^{n-2-k} + np \\
&= n(n-1)p^2(p+q)^{n-2} + np \\
&= n(n-1)p^2 + np.
\end{aligned}$$

donc (Formule de Koenig-Huygens)  $X$  admet une variance et

$$\begin{aligned}
V(X) &= n(n-1)p^2 + np - n^2p^2 \\
&= -np^2 + np
\end{aligned}$$

$$\boxed{V(X) = np(1-p)}$$

### Situation type

Si une expérience est constituée de  $n$  épreuves de Bernoulli identiques et indépendantes et si  $X$  désigne le nombre de succès alors  $X$  suit la loi binomiale de paramètres  $(n, p)$  où  $p$  est la probabilité du succès.

**Remarque :** On se retrouve dans cette situation lors d'un **tirage successif avec remise**.

### Simulation avec une fonction Python.

```
def binomiale(n,p):
    x = 0
    for k in range(n):
        x += bernoulli(p)
    return x
```

### Somme de $n$ variables de Bernoulli identiques et indépendantes.

**Théorème :**

La somme de  $n$  variables de Bernoulli de paramètre  $p$  (*mutuellement*) indépendantes est une variable aléatoire suivant la loi binomiale de paramètres  $(n, p)$

Soient  $n$  un entier naturel non nuls et  $p$  un réel de l'intervalle  $]0, 1[$

$$\text{Si } \left\{ \begin{array}{l} \textcircled{1} X_1, \dots, X_n \text{ sont indépendantes} \\ \textcircled{2} \forall i \in \llbracket 1; n \rrbracket, X_i \hookrightarrow \mathcal{B}(p) \\ \textcircled{3} X = X_1 + X_2 + \dots + X_n \end{array} \right. , \quad \text{alors } X \hookrightarrow \mathcal{B}(n, p)$$

### Stabilité de la loi binomiale. (Complément)

Soient  $X_1$  et  $X_2$  deux variables aléatoires,  $n_1, n_2$  deux entiers naturels non nuls et  $p$  un réel de  $]0, 1[$

Si  $X_1$  et  $X_2$  sont indépendantes et si  $X_1 \hookrightarrow \mathcal{B}(n_1, p)$  et  $X_2 \hookrightarrow \mathcal{B}(n_2, p)$

alors

$$X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$$

**Généralisation .**

Soient  $(X_k)_{1 < k \leq N}$  une liste de variables aléatoires,  $(n_k)_{1 < k \leq N}$  une liste d'entiers naturels non nuls et

$p$  un réel de  $]0, 1[$ , on note :  $X = \sum_{k=1}^N X_k$  et  $n = \sum_{k=1}^N n_k$ .

Si les  $X_k$  sont mutuellement indépendantes et si  $\forall k \in \mathbb{N}, X_k \hookrightarrow \mathcal{B}(n_k, p)$  alors  $X \hookrightarrow \mathcal{B}(n, p)$

## 5.4 Lois discrètes infinies usuelles.

### 5.4.1 Lois géométriques

#### Définition

Soient  $X$  une variable aléatoire réelle et  $p$  un réel de  $]0, 1[$ ,  
Dire que  $X$  suit une loi géométrique de paramètre  $p$  signifie :

$$X(\Omega) = \mathbb{N}^* \quad \forall n \in \mathbb{N}^*, \quad P(X = n) = (1 - p)^{n-1} p$$

#### Remarques :

- $X(\Omega) = \mathbb{N}^*$  et  $\sum_{k=1}^{+\infty} (1-p)^{k-1} p = \frac{p}{1-(1-p)} = 1$  donc  $X$  est une variable aléatoire discrète bien définie.
- On note  $X \leftrightarrow \mathcal{G}(p)$ .
- En notant  $q = 1 - p$ , on a :  $\forall n \in \mathbb{N}^*, \quad P(X = n) = q^{n-1} p$

#### Situation type.

##### (Rédaction)

"Cette expérience est la succession d'un nombre indéfini d'épreuves de Bernoulli identiques et indépendantes, le succès est ..... la probabilité du succès vaut  $p$ ,

$X$  est le rang du premier succès donc  $X \leftrightarrow \mathcal{G}(p)$ ."

#### Fonction de répartition.

##### Proposition

Soit  $F$  la fonction de répartition d'une variable aléatoire  $X$  suivant une loi géométrique de paramètre  $p$ ,

- ❶ Pour tout  $n < 1$ ,  $F(n) = 0$
- ❷ Pour tout  $n \in \mathbb{N}^*$ ,  $F$  est constante sur  $[n; n + 1[$  et  $F(n) = 1 - q^n$

##### Démonstration

Soit  $n \in \mathbb{N}^*$ ,

$$\begin{aligned} P([X > n]) &= 1 - P([X \leq n]) \\ &= 1 - \sum_{k=1}^n P([X = k]) \\ &= 1 - \sum_{k=1}^n pq^{k-1} \\ &= 1 - p \frac{1 - q^n}{1 - q} \\ &= q^n \end{aligned}$$

##### Interprétation :

$[X > n]$  : "avoir des échecs au cours des  $n$  premières épreuves", donc  $P(X > n) = q^n$

##### Théorème

$$X \leftrightarrow \mathcal{G}(p) \quad \text{si, et seulement si,} \quad \forall n \in \mathbb{N}^*, \quad P(X \leq n) = 1 - q^n \quad \text{et} \quad P(X > n) = q^n$$

**Démonstration.** La loi d'une variable aléatoire est entièrement définie par la donnée de sa fonction de répartition.

**Espérance et variance.**

**Proposition.**

Soit  $X$  un variable aléatoire réelle,  
 si  $X$  suit une loi géométrique de paramètre  $p$  alors  $X$  admet une espérance et une variance et

$$E(X) = \frac{1}{p} \quad \text{et} \quad V(X) = \frac{q}{p^2}$$

**Démonstration.**

•  $X$  admet une espérance si, et seulement si,  $\sum_{n \geq 1} x_n P(X = x_n)$  est absolument convergente.

or  $\sum_{n \geq 1} |x_n| P(X = x_n) = \sum_{n \geq 1} n(1-p)^{n-1} p$  qui est une série géométrique dérivée qui converge

donc  $X$  admet une espérance et

$$\begin{aligned} E(X) &= \sum_{n=1}^{+\infty} n(1-p)^{n-1} p \\ &= p \sum_{n=1}^{+\infty} n(1-p)^{n-1} \\ &= p \times \frac{1}{(1-(1-p))^2} \end{aligned}$$

$$E(X) = \frac{1}{p}$$

**Remarque :**

En lançant indéfiniment un dé, le nombre moyen de lancers pour obtenir le premier 1 est égal à 6.

*c'est l'inverse de  $\frac{1}{6}$*

•  $X^2$  admet une espérance si, et seulement si,  $\sum_{n \geq 1} x_n^2 P(X = x_n)$  est absolument convergente.

si, et seulement si,  $\sum_{n \geq 1} x_n^2 P(X = x_n)$  est convergente. ( car  $x_n^2 \geq 0$  )

(Simplifions les sommes partielles, puis faisons tendre  $n$  vers  $+\infty$ )

$$\begin{aligned} \sum_{k=1}^n x_k^2 P(X = x_k) &= \sum_{k=1}^n k^2 p q^{k-1} \\ &= p q \sum_{k=1}^n k(k-1) q^{k-2} + p \sum_{k=1}^n k q^{k-1} \\ &\xrightarrow{n \rightarrow +\infty} p q \times \frac{2}{p^3} + \frac{p}{p^2} \end{aligned}$$

donc  $E(X^2)$  existe et  $E(X^2) = \frac{2q+p}{p^2}$

on en déduit (formule Kœnig-Huygens) que  $X$  admet une variance et

$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 \\ &= \frac{2q+p}{p^2} - \frac{1}{p^2} \\ &= \frac{2q+p-1}{p^2} \\ &= \frac{2q-q}{p^2} \end{aligned}$$

$$V(X) = \frac{q}{p^2}$$

**Loi sans mémoire.****Proposition**

Soient  $X$  une variable aléatoire réelle et  $p$  un réel de  $]0, 1[$ ,  
si  $X$  suit une loi géométrique de paramètre  $p$  alors

$$\forall (k, n) \in \mathbb{N} \times \mathbb{N}^*, \quad P([X > n+k] | [X > k]) = P([X > n])$$

**Démonstration.**

Par définition (de la probabilité conditionnelle) : 
$$P([X > n+k] | [X > k]) = \frac{P([X > n+k] \cap [X > k])}{P([X > k])}$$

or  $n \geq 0$  donc  $[X > n+k] \subset [X > k]$  et ainsi  $[X > n+k] \cap [X > k] = [X > n+k]$  et

$$P([X > n+k] | [X > k]) = \frac{P([X > n+k])}{P([X > k])}$$

or  $\underbrace{P([X > k]) = q^k}$  et  $P([X > n+k]) = q^{n+k}$ , donc  $P([X > n+k] | [X > k]) = q^n$ ,  
vu avec la fonct. de répartition

en conclusion on a bien :

$$\forall (k, n) \in \mathbb{N} \times \mathbb{N}^*, \quad P([X > n+k] | [X > k]) = P([X > n])$$

**Remarques :** (*interprétation*)

- En se plaçant dans la situation type :  $[X > k]$  : " Pas de succès pendant les  $k$  premiers lancers " on observe les  $n$  prochains lancers le rang du premier succès suit la loi  $\mathcal{G}(p)$
- Si on n'observe l'expérience de 1 à  $k$  et qu'il n'y a pas eu de succès alors la loi du temps d'attente du premier succès est la même qu'au début.
- Encore une autre interprétation :  
On fixe  $n \in \mathbb{N}^*$ , et on note  $Y = X - n$ ,  
Sachant  $(X > n)$  ("il n'y a eu que des échecs du rang 1 au rang  $n$ "),  
 $Y$  est alors le temps d'attente du premier succès à partir du rang  $n+1$ ,  
La proposition a montré que : la loi conditionnelle sachant  $(X > n)$  de  $Y$  est la loi  $\mathcal{G}(p)$
- On retrouvera cette propriété avec la loi exponentielle.

**Simulation numérique.****Simulation avec une fonction Python :**

```
def geometrique(p):
    x = 1
    while rd.random() > p :
        x += 1
    return x
```

**5.4.2 Loi de Poisson****Définition****Définition.**

Soient  $X$  une variable aléatoire réelle et  $\lambda$  un réel strictement positif,  
Dire que  $X$  suit une loi de Poisson de paramètre  $\lambda$  signifie que :

$$X(\Omega) = \mathbb{N} \quad \text{et} \quad \forall n \in \mathbb{N}, \quad P(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

$X(\Omega) = \mathbb{N}$ , pour tout  $n \in \mathbb{N}$ ,  $\frac{\lambda^n}{n!} e^{-\lambda} \geq 0$  et

$$\begin{aligned} \sum_{k=0}^n \frac{\lambda^k}{k!} e^{-\lambda} &= e^{-\lambda} \sum_{k=0}^n \frac{\lambda^k}{k!} \\ &\xrightarrow{n \rightarrow +\infty} e^{-\lambda} \times e^\lambda \quad \text{série exponentielle} \\ &\xrightarrow{n \rightarrow +\infty} 1 \end{aligned}$$

On a bien défini une variable aléatoire discrète.

**Remarques :**

- on note :  $X \hookrightarrow \mathcal{P}(\lambda)$
- Siméon Denis Poisson (1781-1840)
- Ne pas oublier de préciser que  $\lambda \in \mathbb{R}_+^*$  au début de cette définition.
- Savoir vérifier rapidement que c'est bien une loi de probabilité .

Espérance et variance.

**Proposition.**

Soient  $X$  une variable aléatoire réelle et  $\lambda$  un réel strictement positif,  
 Si  $X$  suit une loi de Poisson de paramètre  $\lambda$  alors, admet une espérance et une variance et  
 $E(X) = \lambda$  et  $V(X) = \lambda$

**Démonstration :**

- $X$  admet une espérance si, et seulement si,  $\sum_{n \geq 0} x_n P(X = x_n)$  est absolument convergente.  
 si, et seulement si,  $\sum_{n \geq 0} x_n P(X = x_n)$  est convergente. (car  $x_n \geq 0$ )

Utilisons les sommes partielles.

Pour  $n \in \mathbb{N}$ ,

$$\begin{aligned} \sum_{k=0}^n k P(X = k) &= \sum_{k=0}^n k \frac{\lambda^k}{k!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{k=1}^n \frac{\lambda^k}{(k-1)!} \\ &= \lambda e^{-\lambda} \sum_{k=1}^n \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda e^{-\lambda} \sum_{k=0}^{n-1} \frac{\lambda^k}{k!} \quad \text{on réindice} \\ \text{(on fait tendre } n \text{ vers } +\infty) &\xrightarrow{n \rightarrow +\infty} \lambda e^{-\lambda} \times e^\lambda \quad \text{somme partielle de la série exponentielle} \\ &\xrightarrow{n \rightarrow +\infty} \lambda \end{aligned}$$

Donc  $X$  admet une espérance et

$E(X) = \lambda$

- $X^2$  admet une espérance si, et seulement si,  $\sum_{n \geq 0} x_n^2 P(X = x_n)$  est absolument convergente.  
 si, et seulement si,  $\sum_{n \geq 0} x_n^2 P(X = x_n)$  est convergente. (car  $x_n^2 \geq 0$ )

Avec les sommes partielles.

$$\sum_{k=0}^n x_k^2 P(X = x_k) = \sum_{k=0}^n k^2 \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\begin{aligned}
&= \sum_{k=0}^n (k(k-1) + k) \frac{\lambda^k}{k!} e^{-\lambda} \\
&= e^{-\lambda} \lambda^2 \sum_{k=2}^n \frac{\lambda^{k-2}}{(k-2)!} + e^{-\lambda} \lambda \sum_{k=1}^n \frac{\lambda^{k-1}}{(k-1)!} \\
&= e^{-\lambda} \lambda^2 \sum_{k=0}^{n-2} \frac{\lambda^k}{k!} + e^{-\lambda} \lambda \sum_{k=0}^{n-1} \frac{\lambda^k}{k!} \\
&\xrightarrow{n \rightarrow +\infty} e^{-\lambda} \lambda^2 e^{\lambda} + e^{-\lambda} \lambda e^{\lambda} \quad (\text{Série exponentielle}) \\
&\xrightarrow{n \rightarrow +\infty} \lambda^2 + \lambda
\end{aligned}$$

donc  $X^2$  admet une espérance et :  $E(X^2) = \lambda^2 + \lambda$

ainsi (*formule Kœnig-Huygens*)  $X$  admet une variance et

$$\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \lambda^2 + \lambda - \lambda^2
\end{aligned}$$

donc  $V(X) = \lambda$

### Approximation de lois binomiales par des lois de Poisson.

**Proposition.**

Soient  $X$  une variable aléatoire réelle et  $\lambda$  un réel strictement positif,  
 Pour tout  $k \in \mathbb{N}$ ,

$$\binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \xrightarrow{n \rightarrow +\infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

En posant  $p_n = \frac{\lambda}{n}$  on obtient :

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} \xrightarrow{n \rightarrow +\infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

**Démonstration :**

On fixe  $k \in \mathbb{N}$  et on prend un  $n \geq k$ ,

$$\begin{aligned}
\binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} &\sim \frac{n^k}{k!} \times \frac{\lambda^k}{n^k} \times \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{\lambda}{n}\right)^{-k} \\
&\sim \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \\
&\xrightarrow{n \rightarrow +\infty} \frac{\lambda^k}{k!} e^{-\lambda}
\end{aligned}$$

Dans le calcul précédent on a utilisé :  $\left(1 - \frac{\lambda}{n}\right)^n \xrightarrow{n \rightarrow +\infty} e^{-\lambda}$  et  $\binom{n}{k} \underset{n \rightarrow +\infty}{\sim} \frac{n^k}{k!}$

### **Approximation**

Quand  $n \geq 30$  et  $p \leq 0,1$  on peut approcher la loi binomiale  $\mathcal{B}(n, p)$  par la loi de Poisson  $\mathcal{P}(np)$

### Simulation numérique.

### **Approximation**

Quand  $n \geq 30$  et  $\frac{\lambda}{n} \leq 0,1$  on peut approcher par la loi de Poisson  $\mathcal{P}(\lambda)$  par la loi binomiale  $\mathcal{B}\left(n, \frac{\lambda}{n}\right)$

Remarque :  $n \geq 30$  et  $\frac{\lambda}{n} \leq 0,1$  équivaut à  $n \geq \max(30, 10\lambda)$

**Simulation avec une fonction Python :**

```
def loi_poisson(lbd):          # attention lambda est un mot réservé du langage Python
    n = int(max(10*lbd, 30))  # on choisit n pour avoir n >= 30 et p <= 0.1
    p = lbd/n
    return binomiale(n, p)
```

Il existe aussi `rd.poisson(lbd)` du module `random`.

# Applications linéaires.

## Plan du chapitre

6.1	Applications linéaires. . . . .	<b>67</b>
6.1.1	Exemples. . . . .	68
6.1.2	Définitions et vocabulaire. . . . .	68
6.1.3	Propriétés. . . . .	69
6.2	Opérations et applications linéaires. . . . .	<b>70</b>
6.2.1	Combinaison linéaire. . . . .	70
6.2.2	Composition. . . . .	71
6.2.3	Puissance d'un endomorphisme . . . . .	71
6.2.4	Réciproque d'un isomorphisme. . . . .	72
6.3	Noyau et Image. . . . .	<b>72</b>
6.3.1	Définitions. . . . .	72
6.3.2	Propriétés. . . . .	72
6.3.3	Image et surjectivité. . . . .	73
6.3.4	Noyau et injectivité. . . . .	74
6.4	Image d'une base. . . . .	<b>74</b>
6.5	Représentation matricielle. . . . .	<b>77</b>
6.5.1	Introduction. . . . .	77
6.5.2	Matrice d'une application linéaire. . . . .	77
6.5.3	Matrices et opérations. . . . .	79
6.6	Isomorphe à $\mathbb{K}^n$ . . . . .	<b>80</b>
6.7	Rang d'une application linéaire. . . . .	<b>81</b>
6.7.1	Définition. . . . .	81
6.7.2	Lien avec les autres notions de rang. . . . .	81
6.7.3	Théorème du rang . . . . .	81
6.7.4	Caractérisation des isomorphismes . . . . .	82
6.8	Une matrice vue comme une application linéaire. . . . .	<b>82</b>
6.8.1	Application linéaire canoniquement associée à une matrice. . . . .	82
6.8.2	Noyau, image. . . . .	82
6.8.3	Rang d'une matrice. . . . .	83
6.8.4	Inverse à gauche, inverse à droite. . . . .	84
6.9	Changement de base. . . . .	<b>84</b>
6.9.1	Matrice de passage. . . . .	84
6.9.2	Changement de bases, action sur les coordonnées d'un vecteur. . . . .	84
6.9.3	Changement de bases, action sur la matrice d'un endomorphisme. . . . .	85
6.10	Matrices semblables . . . . .	<b>86</b>

## 6.1 Applications linéaires.

Dans ce cours  $E$  et  $F$  désignent deux espaces vectoriels quelconques sur  $\mathbb{K}$ . ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ )

### 6.1.1 Exemples.

- Les applications linéaires de  $\mathbb{R}$  dans  $\mathbb{R}$  sont les fonctions  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$  vérifiant :

$$\begin{aligned} \text{il existe } a \in \mathbb{R} \text{ tel que : } f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto ax \end{aligned}$$

Remarque : Ce sont les seules fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$  qui possèdent la propriété :

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad \forall (\alpha, \beta) \in \mathbb{R}^2, \quad f(\alpha x_1 + \beta x_2) = \alpha f(x_1) + \beta f(x_2)$$

- "Linéarité de l'intégrale".

Soient  $a$  et  $b$  deux réels tels que  $a < b$ , on considère l'application :  $C^0([a, b]) \longrightarrow \mathbb{R}$

$$f \longmapsto \int_a^b f(t) dt$$

En notant  $\Psi$  cette application on a :

$$\forall (f_1, f_2) \in C^0([a, b])^2, \quad \forall (\alpha, \beta) \in \mathbb{R}^2, \quad \Psi(\alpha f_1 + \beta f_2) = \alpha \Psi(f_1) + \beta \Psi(f_2)$$

- "Linéarité de l'espérance".

Soit  $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$  un espace probabilisé, on note  $\mathcal{V}$  l'ensemble des variables aléatoires sur  $\Omega$ , on considère l'application définie dans le cours de probabilité :

$$\begin{aligned} \mathbb{E} : \mathcal{V} &\longrightarrow \mathbb{R} \\ X &\longmapsto \mathbb{E}(X) \end{aligned}$$

Cette application  $\mathbb{E}$  vérifie :  $\forall (X_1, X_2) \in \mathcal{V}^2, \quad \forall (\alpha, \beta) \in \mathbb{R}^2, \quad \mathbb{E}(\alpha X_1 + \beta X_2) = \alpha \mathbb{E}(X_1) + \beta \mathbb{E}(X_2)$

- "Le passage aux coordonnées dans une base est linéaire".

Soit  $E$  un espace vectoriel de dimension  $n$  et  $B$  une base de  $E$ ,

$$\begin{aligned} E &\longrightarrow \mathcal{M}_{n,1}(\mathbb{R}) \\ u &\longmapsto \text{Coord}_B(u) \end{aligned}$$

Cette application  $\text{Coord}_B$  vérifie :

$$\forall (u_1, u_2) \in E^2, \quad \forall (\alpha, \beta) \in \mathbb{R}^2, \quad \text{Coord}(\alpha u_1 + \beta u_2) = \alpha \text{Coord}_B(u_1) + \beta \text{Coord}_B(u_2)$$

- "Le passage à la dérivée est linéaire".

On considère l'application :  $C^\infty(\mathbb{R}) \longrightarrow C^\infty(\mathbb{R})$

$$f \longmapsto f'$$

En notant  $d$  cette application on a :

$$\forall (f_1, f_2) \in C^\infty(\mathbb{R})^2, \quad \forall (\alpha, \beta) \in \mathbb{R}^2, \quad d(\alpha f_1 + \beta f_2) = \alpha d(f_1) + \beta d(f_2)$$

- " $X \longmapsto AX$ ".

Soient  $p$  et  $n$  deux entiers naturels non nuls et  $A$  une matrice de  $\mathcal{M}_{n,p}(\mathbb{R})$

On considère l'application :

$$\begin{aligned} \mathcal{M}_{p,1}(\mathbb{R}) &\longrightarrow \mathcal{M}_{n,1}(\mathbb{R}) \\ X &\longmapsto AX \end{aligned}$$

En notant  $\Psi$  cette application on a :

$$\forall (X_1, X_2) \in \mathcal{M}_{n,1}(\mathbb{R})^2, \quad \forall (\alpha, \beta) \in \mathbb{R}^2, \quad \Psi(\alpha X_1 + \beta X_2) = \alpha \Psi(X_1) + \beta \Psi(X_2)$$

### 6.1.2 Définitions et vocabulaire.

#### Définition :

Soit  $f$  une application de  $E$  dans  $F$ .

Dire que  $f$  est une application linéaire signifie que :

$$\forall (\alpha, \beta) \in \mathbb{K}^2, \quad \forall (u, v) \in E^2, \quad f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$$

**Vocabulaire et notations :**

- On note :  $\mathcal{L}(E, F)$  l'ensemble des applications linéaires de  $E$  dans  $F$ .
- (Lorsque  $F = \mathbb{K}$ ). On appelle **forme linéaire** sur  $E$  les applications linéaires de  $E$  dans  $\mathbb{K}$ .
- (Lorsque  $F = E$ ). On appelle **endomorphisme** de  $E$  les applications linéaires de  $E$  dans  $E$ .
- On note  $\mathcal{L}(E)$  l'ensemble des endomorphismes de  $E$ .
- Les applications linéaires bijectives sont appelées **isomorphismes**.
- Les endomorphismes bijectifs sont appelés **automorphismes**.
- On note  $GL(E)$  l'ensemble des automorphismes de  $E$ . (*Groupe Linéaire*)

**Des cas particuliers importants.**

L'application nulle de  $E$  dans  $F$  est linéaire.

L'application identité de  $E$  est un endomorphisme de  $E$ . (*c'est même un isomorphisme*)

L'application  $\text{Coord}_B$  définie dans le cours sur les espaces vectoriels est un isomorphisme.

*Autres exemples dans la feuille Cours\_6.*

**6.1.3 Propriétés.****Propositions :**

❶ Si  $f$  est linéaire de  $E$  dans  $F$ , on a alors :  $f(0_E) = 0_F$

❷ Si  $f$  est linéaire de  $E$  dans  $F$ ,

$$\text{pour tout } (u_1, \dots, u_n) \in E^n \text{ et } (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \quad f\left(\sum_{i=1}^n \lambda_i u_i\right) = \sum_{i=1}^n \lambda_i f(u_i)$$

**Démonstration :**

❶ Si  $f \in \mathcal{L}(E, F)$ , on a  $\forall (\alpha, \beta) \in \mathbb{K}^2, \forall (u, v) \in E^2, f(\alpha.u + \beta.v) = \alpha.f(u) + \beta.f(v)$   
donc en particulier  $f(0.0_E + 0.0_E) = 0.f(0_E) + 0.f(0_E)$  et comme  $0.0_E = 0_E$  et  $0.f(0_E) = 0_F$  il vient  $f(0_E) = 0_F$

❷ On suppose que  $f$  est linéaire de  $E$  dans  $F$ ,  
pour chaque entier  $n$  on note :  $\mathcal{P}_n$  la proposition :

$$\text{pour tout } (u_1, \dots, u_n) \in E^n \text{ et } (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \quad f\left(\sum_{i=1}^n \lambda_i u_i\right) = \sum_{i=1}^n \lambda_i f(u_i)$$

montrons par récurrence sur  $n$  que pour tout entier  $n \geq 2$ ,  $\mathcal{P}_n$  est vraie.

- Pour  $n = 2$ ,

La proposition est vraie, c'est exactement la définition d'une application linéaire.

- Soit  $n \geq 2$  tel que  $\mathcal{P}_n$  est vraie,  
Soient  $(u_1, \dots, u_{n+1}) \in E^{n+1}$  et  $(\lambda_1, \dots, \lambda_{n+1}) \in \mathbb{K}^{n+1}$ ,

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} \lambda_i u_i\right) &= f\left(\sum_{i=1}^n \lambda_i u_i + \lambda_{n+1} u_{n+1}\right) \\ &= f\left(\sum_{i=1}^n \lambda_i u_i\right) + \lambda_{n+1} f(u_{n+1}) && \text{car } f \in \mathcal{L}(E, F) \\ &= \sum_{i=1}^n \lambda_i f(u_i) + \lambda_{n+1} f(u_{n+1}) && \text{en utilisant } \mathcal{P}_n \\ f\left(\sum_{i=1}^{n+1} \lambda_i u_i\right) &= \sum_{i=1}^{n+1} \lambda_i f(u_i) && \text{on a bien } \mathcal{P}_{n+1} \end{aligned}$$

On a bien montré que pour tout entier  $n \geq 2$  :

$$\text{pour tout } (u_1, \dots, u_n) \in E^n \text{ et } (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \quad f\left(\sum_{i=1}^n \lambda_i u_i\right) = \sum_{i=1}^n \lambda_i f(u_i)$$

**Remarque :**

Pour démontrer qu'une application de  $E$  dans  $F$  n'est pas linéaire, on montre au choix :

- ❶  $f(0_E) \neq 0_F$
- ❷ En prenant  $\alpha = \dots\dots\dots$  dans  $\mathbb{K}$  et  $u = \dots\dots\dots$  dans  $E$ , on a  $f(\alpha u) \neq \alpha f(u)$
- ❸ En prenant  $u = \dots\dots\dots$  dans  $E$  et  $v = \dots\dots\dots$  dans  $E$ , on a  $f(u + v) \neq f(u) + f(v)$

Exemples dans la feuille Cours\_6.

## 6.2 Opérations et applications linéaires.

### 6.2.1 Combinaison linéaire.

Des précisions sur l'espace vectoriel  $F^E$  où  $E$  et  $F$  désignent deux espaces vectoriels.

**Définition :**

Pour  $f$  une application  $E$  dans  $F$  et  $\alpha$  un scalaire on définit les applications :

$$\begin{array}{ll} \alpha f : E \longrightarrow F & f + g : E \longrightarrow F \\ u \longmapsto \alpha f(u) & u \longmapsto f(u) + g(u) \end{array}$$

**Remarques :**

- On admet que  $F^E$  muni de ces deux lois est un espace vectoriel.
- Le vecteur nul de  $F^E$  est la fonction nulle, plus précisément c'est l'application  $E \longrightarrow F$   
 $u \longmapsto 0_F$

**Théorème :**

Soient  $f$  et  $g$  deux applications de  $E$  dans  $F$ ,

Si  $f$  et  $g$  sont dans  $\mathcal{L}(E, F)$  et  $\alpha$  et  $\beta$  deux scalaires alors  $\alpha f + \beta g \in \mathcal{L}(E, F)$ .

**Démonstration :**

Soient  $(\alpha, \beta) \in \mathbb{K}^2$  et  $(f, g) \in \mathcal{L}(E, F)^2$ , on note  $h = \alpha f + \beta g$ ,

Soient  $(\lambda, \mu) \in \mathbb{K}^2$  et  $(u, v) \in E^2$ ,

$$\begin{aligned} h(\lambda u + \mu v) &= \alpha f(\lambda u + \mu v) + \beta g(\lambda u + \mu v) && \text{(Définitions des opérations ci-dessus)} \\ &= \alpha \lambda f(u) + \alpha \mu f(v) + \beta \lambda g(u) + \beta \mu g(v) && \text{(f et g sont linéaires)} \\ &= \lambda(\alpha f(u) + \beta g(u)) + \mu(\alpha f(v) + \beta g(v)) \\ &= \lambda h(u) + \mu h(v) \end{aligned}$$

$$\text{donc } h = \alpha f + \beta g \in \mathcal{L}(E, F)$$

On a bien démontré que :  $\boxed{\text{si } f \in \mathcal{L}(E, F) \text{ et } g \in \mathcal{L}(E, F) \text{ alors } \alpha f + \beta g \in \mathcal{L}(E, F)}$

**Remarques :**

- Le vecteur nul étant l'application nulle de  $E$  dans  $F$ ,  $\mathcal{L}(E, F)$  est un sous-espace vectoriel de  $F^E$ .
- Plus généralement : Toute combinaison linéaire d'applications linéaires est une application linéaire.

$$\text{Si pour tout } k \in \llbracket 1; n \rrbracket, \lambda_k \in \mathbb{K} \text{ et } f_k \in \mathcal{L}(E, F) \text{ alors } \sum_{k=1}^n \lambda_k f_k \in \mathcal{L}(E, F)$$

### 6.2.2 Composition.

**Théorème :**

Soient  $E, F$  et  $G$  trois espaces vectoriels et  $f : E \rightarrow F$  et  $g : F \rightarrow G$ .  
**Si**  $f \in \mathcal{L}(E, F)$  et  $g \in \mathcal{L}(F, G)$  **alors**  $g \circ f \in \mathcal{L}(E, G)$

**Démonstration :**

On suppose  $f \in \mathcal{L}(E, F)$  et  $g \in \mathcal{L}(F, G)$ ,  
 Soient  $(\alpha, \beta) \in \mathbb{K}^2$  et  $(u, v) \in E^2$ ,

$$\begin{aligned} g \circ f(\alpha u + \beta v) &= g(\alpha f(u) + \beta f(v)) && \text{car } f \in \mathcal{L}(E, F) \\ &= \alpha g(f(u)) + \beta g(f(v)) && \text{car } g \in \mathcal{L}(F, G) \\ &= \alpha g \circ f(u) + \beta g \circ f(v) \end{aligned}$$

donc  $g \circ f \in \mathcal{L}(E, G)$

**Plus généralement :** Toute composée d'applications linéaires est une application linéaire.

**Propriétés :**

- ❶ Si  $g \in \mathcal{L}(F, G)$ ,  $(\alpha_1, \alpha_2) \in \mathbb{K}^2$ ,  $f_1 \in \mathcal{L}(E, F)$   $f_2 \in \mathcal{L}(E, F)$   
**alors**  $g \circ (\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 g \circ f_1 + \alpha_2 g \circ f_2$
- ❷ Si  $g \in \mathcal{L}(E, F)$ ,  $(\alpha_1, \alpha_2) \in \mathbb{K}^2$ ,  $f_1 \in \mathcal{L}(F, G)$   $f_2 \in \mathcal{L}(F, G)$   
**alors**  $(\alpha_1 f_1 + \alpha_2 f_2) \circ g = \alpha_1 f_1 \circ g + \alpha_2 f_2 \circ g$

**En effet :** (égalité de deux applications par égalité en tout  $u \in E$ )

- ❶  $g \circ (\alpha_1 f_1 + \alpha_2 f_2)(u) = g(\alpha_1 f_1(u) + \alpha_2 f_2(u)) \underset{g \in \mathcal{L}(F, G)}{=} \alpha_1 g(f_1(u)) + \alpha_2 g(f_2(u)) = (\alpha_1 g \circ f_1 + \alpha_2 g \circ f_2)(u)$
- ❷  $(\alpha_1 f_1 + \alpha_2 f_2) \circ g(u) = \alpha_1 f_1(g(u)) + \alpha_2 f_2(g(u)) = \alpha_1 f_1 \circ g(u) + \alpha_2 f_2 \circ g(u) = (\alpha_1 f_1 \circ g + \alpha_2 f_2 \circ g)(u)$

**Remarque :** ici, on utilise uniquement la linéarité de  $g$  dans ❶. Dans ❷, aucune linéarité supplémentaire n'intervient : on utilise seulement la définition de la combinaison linéaire d'applications.

### 6.2.3 Puissance d'un endomorphisme

**Définition :**

Soit  $f \in \mathcal{L}(E)$ , on définit pour  $n \in \mathbb{N}$  la notation  $f^n$  par la relation de récurrence :

$$f^0 = Id_E \quad \forall n \in \mathbb{N}, \quad f^{n+1} = f^n \circ f$$

Plus simplement :  $f^n = \underbrace{f \circ \dots \circ f}_n$   
 n applications

**Théorème :**

Soit  $E$  un espace vectoriel,  $f : E \rightarrow E$  et  $n \in \mathbb{N}$ . **Si**  $f \in \mathcal{L}(E)$  **alors**  $f^n \in \mathcal{L}(E)$

**Propriétés :**

Soit  $f \in \mathcal{L}(E)$ ,  
 pour tout  $(n_1, n_2) \in \mathbb{N}^2$ ,  $f^{n_1+n_2} = f^{n_1} \circ f^{n_2}$  et  $(f^{n_1})^{n_2} = f^{n_1 n_2}$

**Conséquences :**

Soit  $f \in \mathcal{L}(E)$ ,  
 pour tout  $(n_1, n_2) \in \mathbb{N}^2$ ,  $f^{n_1} \circ f^{n_2} = f^{n_2} \circ f^{n_1}$  et  $(f^{n_1})^{n_2} = (f^{n_2})^{n_1}$

**Remarque :** en général, la composition des applications n'est pas commutative. Ici, la commutativité vient du fait que toutes les applications sont des puissances d'une même application  $f$ .

### 6.2.4 Réciproque d'un isomorphisme.

**Théorème :**

La réciproque d'une bijection linéaire est linéaire.

**Démonstration :**

On suppose connaître  $f$  une application linéaire bijective de  $E$  dans  $F$

Soient  $(v_1, v_2) \in F^2$  et  $(\alpha_1, \alpha_2) \in \mathbb{K}^2$ ,

comme  $f$  est bijective on peut définir  $u_1 = f^{-1}(v_1)$  et  $u_2 = f^{-1}(v_2)$ ,

ce qui nous donne :  $v_1 = f(u_1)$  et  $v_2 = f(u_2)$ .

$$\begin{aligned} f^{-1}(\alpha_1 v_1 + \alpha_2 v_2) &= f^{-1}(\alpha_1 f(u_1) + \alpha_2 f(u_2)) \\ &= f^{-1}(f(\alpha_1 u_1 + \alpha_2 u_2)) && \text{car } f \text{ est linéaire} \\ &= \alpha_1 u_1 + \alpha_2 u_2 && \text{car } f^{-1} \circ f = Id_E \\ &= \alpha_1 f^{-1}(v_1) + \alpha_2 f^{-1}(v_2) \end{aligned}$$

*Ce qui achève la démonstration*

**Autrement dit sachant que la réciproque d'une bijection est une bijection :**

Si  $f$  est un isomorphisme de  $E$  dans  $F$ , alors  $f^{-1}$  est un isomorphisme de  $F$  dans  $E$ .

**Remarque :** Si  $f$  est un automorphisme de  $E$  alors pour tout  $n \in \mathbb{N}$ , on note  $f^{-n} = (f^{-1})^n$

## 6.3 Noyau et Image.

### 6.3.1 Définitions.

**Définition :**

Etant donné une application linéaire  $f : E \rightarrow F$ , on appelle :

- **noyau de  $f$** , le sous-ensemble de  $E$  suivant :  $\ker(f) = \{u \in E \mid f(u) = 0_F\}$
- **image de  $f$** , le sous-ensemble de  $F$  suivant :  $\text{Im}(f) = \{f(u) \mid u \in E\}$

**Remarques :**

- $\ker(f) \subset E$  et  $\text{Im}(f) \subset F$ .
- $\ker(f)$  est l'ensemble des antécédents de  $0_F$  par  $f$ .  $\text{Im}(f)$  est l'image directe de  $E$  par  $f$ .
- $\text{Im}(f) = \{v \in F \mid \exists u \in E : f(u) = v\}$
- Si  $f$  est l'application nulle alors  $\text{Im}(f) = \{0_F\}$  et  $\ker(f) = E$ .
- Si  $f$  est l'application identité de  $E$  alors  $\text{Im}(f) = E$  et  $\ker(f) = \{0_E\}$ .

### 6.3.2 Propriétés.

**Théorème :**

Si  $f$  est une application linéaire de  $E$  dans  $F$  alors :

- ❶  $\ker(f)$  est un sous-espace vectoriel de  $E$
- ❷  $\text{Im}(f)$  est un sous-espace vectoriel de  $F$ .

**Démonstrations :**

- ❶ •  $\ker(f) \subset E$ ,
- $f(0_E) = 0_F$  donc  $0_E \in \ker(f)$
- Soient  $(\alpha, \beta) \in \mathbb{K}^2$  et  $(u, v) \in \ker(f)^2$ ,

$$\begin{aligned} f(\alpha u + \beta v) &= \alpha f(u) + \beta f(v) && \text{car } f \in \mathcal{L}(E, F) \\ &= \alpha 0_F + \beta 0_F && \text{car } u, v \in \ker(f) \\ &= 0_F \end{aligned}$$

donc  $\alpha u + \beta v \in \ker(f)$

En conclusion :  $\ker(f)$  est un sous-espace vectoriel de  $E$

② •  $\text{Im}(f) \subset F$ ,

•  $0_F = f(0_E)$  donc  $0_F \in \text{Im}(f)$

• Soient  $(\alpha, \beta) \in \mathbb{K}^2$  et  $(v_1, v_2) \in \text{Im}(f)^2$ ,

on note  $v_1 = f(u_1)$  et  $v_2 = f(u_2)$  avec  $u_1, u_2 \in E$ ,

$$\begin{aligned} \alpha v_1 + \beta v_2 &= \alpha f(u_1) + \beta f(u_2) \\ &= f(\alpha u_1 + \beta u_2) \quad \text{car } f \in \mathcal{L}(E, F) \\ &\in \text{Im}(f) \end{aligned}$$

En conclusion :  $\text{Im}(f)$  est un sous-espace vectoriel de  $F$

### 6.3.3 Image et surjectivité.

**Proposition :** Pour  $f \in \mathcal{L}(E, F)$ ,  $f$  est surjective si, et seulement si,  $\text{Im}(f) = F$ .

**En effet :**

$\Rightarrow$  Si  $f$  est surjective, alors  $\forall y \in F, \exists x \in E$  tel que  $y = f(x) \in \text{Im}(f)$  et ainsi  $F \subset \text{Im}(f)$ .

Or  $\text{Im}(f) \subset F$ , donc  $\text{Im}(f) = F$ .

$\Leftarrow$  Si  $\text{Im}(f) = F$ , alors  $\forall y \in F, \exists x \in E$  tel que  $f(x) = y$ . Ainsi  $f$  est surjective.

**Remarque :** Sachant que  $\text{Im}(f) = f(E)$  (*image directe*) cette propriété est vraie pour toute application de  $E$  dans  $F$ .

**Théorème :**

Soit  $f \in \mathcal{L}(E, F)$ , Si  $(e_1, \dots, e_n)$  est une base de  $E$  alors,  $\text{Im}(f) = \text{Vect}(f(e_1), \dots, f(e_n))$ .

**Démonstration 1 :** Double inclusion

$\supseteq$  Sachant que pour tout  $i \in \llbracket 1, n \rrbracket$ ,  $f(e_i) \in \text{Im}(f)$  et que  $\text{Im}(f)$  est un sous-espace vectoriel de  $F$ , on a bien  $\text{Vect}(f(e_1), \dots, f(e_n)) \subset \text{Im}(f)$

$\subseteq$  Soit  $y \in \text{Im}(f)$ , on note  $y = f(x)$  avec  $x = \sum_{i=1}^n x_i e_i$ ,

comme  $f$  est linéaire, il vient  $y = \sum_{i=1}^n x_i f(e_i)$ , et ainsi  $y \in \text{Vect}(f(e_1), \dots, f(e_n))$

on a bien  $\text{Im}(f) \subset \text{Vect}(f(e_1), \dots, f(e_n))$

En conclusion :

$$\text{Im}(f) = \text{Vect}(f(e_1), \dots, f(e_n))$$

**Démonstration 2 :** Une autre démonstration plus efficace mais moins élémentaire.

$$\begin{aligned} \text{Im}(f) &= \{f(u) \mid u \in E\} \\ &= \left\{ f\left(\sum_{i=1}^n x_i e_i\right) \mid (x_1, \dots, x_n) \in \mathbb{K}^n \right\} \\ &= \left\{ \sum_{i=1}^n x_i f(e_i) \mid (x_1, \dots, x_n) \in \mathbb{K}^n \right\} \quad (\text{car } f \text{ est linéaire}) \\ &= \text{Vect}(f(e_1), \dots, f(e_n)) \end{aligned}$$

### 6.3.4 Noyau et injectivité.

**Théorème :**

Soit  $f$  une application linéaire de  $E$  dans  $F$ ,  
 $f$  est injective si, et seulement si,  $\ker(f) = \{0_E\}$

**Démonstration**

$\Rightarrow$  On suppose que  $f$  est injective.

Pour  $x \in E$ ,

$$\begin{aligned} x \in \ker(f) &\iff f(x) = 0_F \\ &\iff f(x) = f(0_E) \\ &\iff x = 0_E \quad (\text{car } f \text{ est injective}) \end{aligned}$$

donc  $\ker(f) = \{0_E\}$  ■.

*Remarque : au tableau, nous n'avons pas fait un raisonnement par équivalence.*

$\Leftarrow$  On suppose que  $\ker(f) = \{0_E\}$ .

Soit  $(x_1, x_2) \in E^2$  tel que  $f(x_1) = f(x_2)$ ,

on a alors :  $f(x_1) - f(x_2) = 0_F$ , et comme  $f$  est linéaire il vient  $f(x_1 - x_2) = 0_F$

d'où  $(x_1 - x_2) \in \ker(f)$  et en utilisant l'hypothèse  $\ker(f) = \{0_E\}$  il vient  $x_1 - x_2 = 0_E$

ou encore  $x_1 = x_2$ .

donc  $f$  est injective ■.

## 6.4 Image d'une base.

Soient  $E$  et  $F$  deux  $\mathbb{K}$  espaces vectoriels de **dimension finie**.

**Théorème :**

Soit  $\mathcal{B} = (e_1, e_2, \dots, e_n)$  une base de  $E$ ,  
 quel que soit  $(v_1, v_2, \dots, v_n)$  une famille de vecteurs de  $F$ , il existe une et une seule application linéaire  $f$  de  $E$  dans  $F$  vérifiant :

$$\forall i \in \llbracket 1; n \rrbracket, \quad f(e_i) = v_i$$

**Autrement dit :** Une application linéaire est entièrement définie par l'image d'une base.

*Extrait du programme :* "Détermination d'une application linéaire par l'image d'une base."

**Démonstration :** (ce qui a été fait au tableau)

**Idée :** Comme  $f$  est linéaire alors si  $u = \sum_{i=1}^n x_i e_i$  alors  $f(u) = \sum_{i=1}^n x_i f(e_i)$

*La seule connaissance des coordonnées de  $u$  et les  $f(e_i)$  permet de calculer  $f(u)$ .*

**Existence :** L'application suivante convient : (ie : elle est linéaire et vérifie  $\forall i f(e_i) = v_i$ )

$$\begin{aligned} f : E &\longrightarrow F \\ u &\longmapsto \sum_{i=1}^n x_i v_i \quad \text{où} \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \text{Coord}_{\mathcal{B}}(u) \end{aligned}$$

*Au tableau j'ai défini  $f : E \longrightarrow F$  c'est une autre approche mais c'est peut-être plus simple avec les  $v_i$ .*

$$u \longmapsto \sum_{i=1}^n x_i f(e_i)$$

**Unicité :** Supposons connaître deux applications  $f_1$  et  $f_2$  qui conviennent.

*(ie : elles sont linéaires et vérifient  $\forall i f(e_i) = v_i$ )*

Soit  $u \in E$ , on note  $u = \sum_{i=1}^n x_i e_i$ ,

$$\begin{aligned}
 f_1(u) &= f_1\left(\sum_{i=1}^n x_i e_i\right) \\
 &= \sum_{i=1}^n x_i f_1(e_i) && \text{car } f_1 \in \mathcal{L}(E, F) \\
 &= \sum_{i=1}^n x_i f_2(e_i) && \text{car } f_1(e_i) = f_2(e_i) = v_i \\
 &= f_2\left(\sum_{i=1}^n x_i e_i\right) && \text{car } f_2 \in \mathcal{L}(E, F) \\
 &= f_2(u)
 \end{aligned}$$

Donc  $\forall u \in E, f_1(u) = f_2(u)$  donc  $f_1 = f_2$  ■

**Théorème :**

Soit  $f \in \mathcal{L}(E, F)$  et  $\mathcal{B}$  une base de  $E$ .

$f$  est bijective si, et seulement si, l'image par  $f$  de la base  $\mathcal{B}$  est une base de  $F$ .

**Démonstration :**

On note  $\mathcal{B} = (e_1, \dots, e_n)$  et on désigne alors par  $f(\mathcal{B})$  la famille  $(f(e_1), \dots, f(e_n))$

Montrons séparément :

❶  $f$  injective  $\iff f(\mathcal{B})$  est libre et ❷  $f$  surjective  $\iff f(\mathcal{B})$  est génératrice de  $F$ .

Démo de ❶ (je fais comme au tableau mais on peut aussi faire une double implication)

$$\begin{aligned}
 f \text{ est injective} &\iff \ker(f) = \{0_E\} \\
 &\iff \forall u \in E, f(u) = 0_F \iff u = 0_E \\
 &\iff \forall (x_1, \dots, x_n) \in \mathbb{K}^n, f\left(\sum_{k=1}^n x_k e_k\right) = 0_F \iff \sum_{k=1}^n x_k e_k = 0_E \\
 &\iff \forall (x_1, \dots, x_n) \in \mathbb{K}^n, \underbrace{\sum_{k=1}^n x_k f(e_k)}_{\text{car } f \text{ est linéaire}} = 0_F \iff \underbrace{\forall i \in [1, n], x_i = 0}_{\text{car } (e_1, \dots, e_n) \text{ est libre}}
 \end{aligned}$$

On retrouve la définition de la liberté de  $f(\mathcal{B})$ .

$f$  est injective si, et seulement si,  $f(\mathcal{B})$  est libre

Démo de ❷ (Ici je suis plus efficace qu'au tableau en utilisant le résultat important donnant  $\text{Im}(f) = \text{Vect}(f(\mathcal{B}))$ )

$$\begin{aligned}
 f \text{ est surjective} &\iff \text{Im}(f) = F \\
 &\iff \text{Vect}(f(\mathcal{B})) = F
 \end{aligned}$$

$f$  est surjective si, et seulement si,  $f(\mathcal{B})$  est une famille génératrice de  $F$

En conclusion :

$f$  est bijective si, et seulement si,  $f(\mathcal{B})$  est une base de  $F$

**Conséquence :**

Soit  $f \in \mathcal{L}(E, F)$  avec  $E$  un espace de dimension finie.

Si  $f$  est bijective alors  $F$  est de dimension finie et  $\dim(E) = \dim(F)$ .

**En effet :** On vient de montrer que si  $f$  est bijective alors  $f(\mathcal{B})$  est une base de  $F$ .  
Or le nombre de vecteurs de  $f(\mathcal{B})$  est égal à  $\dim(E)$  donc  $\dim(E) = \dim(F)$ .

**Théorème :** (*Quand  $E$  et  $F$  ont la même dimension*)

Soit  $f \in \mathcal{L}(E, F)$  avec  $\dim(E) = \dim(F)$ .

- ❶  $f$  est injective si, et seulement si,  $f$  est bijective.
- ❷  $f$  est surjective si, et seulement si,  $f$  est bijective.

**Démonstration :** (*Il suffit de montrer :  $f$  est injective  $\iff f$  est surjective*)

On suppose que  $\dim(E) = \dim(F) = n$ , on sait donc que  $f(\mathcal{B})$  possède  $n$  vecteurs.

On utilise deux fois la propriété fondamentale (noté  $(*)$ ) :

"Dans un espace vectoriel de dimension  $n$ , toute famille libre (ou génératrice) de  $n$  vecteurs est une base".

$$\begin{aligned}
 f \text{ est injective} &\iff f(\mathcal{B}) \text{ est une famille libre} && \text{(vue dans la démonstration du th. précédent)} \\
 & \quad (*) \updownarrow \\
 &\iff f(\mathcal{B}) \text{ est une base de } F \\
 & \quad (*) \updownarrow \\
 &\iff f(\mathcal{B}) \text{ est une famille génératrice de } F \\
 &\iff f \text{ est surjective} && \text{(vue dans la démonstration du th. précédent)}
 \end{aligned}$$

En conclusion : Lorsque  $\dim(E) = \dim(F)$  on a

$$f \text{ est bijective} \iff f \text{ est injective} \quad \text{et} \quad f \text{ est surjective} \iff f \text{ est bijective}$$

**En pratique :**

- quand  $\dim(E) = \dim(F)$ , il suffit de montrer que  $f$  est injective pour montrer qu'elle est bijective.
- quand  $\dim(E) = \dim(F)$ , il suffit de montrer que  $f$  est surjective pour montrer qu'elle est bijective.

**Théorème :** *En particulier et le plus utile en pratique.*

Soit  $E$  un espace de dimension finie et  $f \in \mathcal{L}(E)$  (un endomorphisme de  $E$ ).

- ❶  $f$  est injective si, et seulement si,  $f$  est bijective.
- ❷  $f$  est surjective si, et seulement si,  $f$  est bijective.

**En effet :** C'est juste un cas particulier du théorème précédent dans le cas  $F = E$ .

**Remarque :** Attention ce théorème est faux en dimension infinie comme le montre les 2 exemples suivants.

- L'application linéaire  $f : \mathbb{R}[X] \longrightarrow \mathbb{R}[X]$  est injective, mais non surjective.  
 $P \longmapsto XP$

**En effet :**

- Soit  $P \in \mathbb{R}[X]$ ,

$$\begin{aligned}
 P \in \ker(f) &\iff f(P) = 0_{\mathbb{R}[X]} \\
 &\iff XP = 0_{\mathbb{R}[X]}
 \end{aligned}$$

$$\iff P = 0_{\mathbb{R}[X]} \quad \text{intégrité car } X \neq 0_{\mathbb{R}[X]}$$

donc  $\ker(f) = \{0_{\mathbb{R}[X]}\}$  et ainsi  $f$  est injective

- 1 n'a pas d'antécédent par  $f$ , en effet si  $XP = 1$  on aurait  $0 = 1$  donc  $f$  n'est pas surjective

- L'application linéaire  $g : \mathbb{R}[X] \longrightarrow \mathbb{R}[X]$  est surjective, mais non injective.  
 $P \longmapsto P'$

**En effet :**

- Pour  $Q = \sum_{k=0}^n a_k X^k$  un polynôme quelconque de  $\mathbb{R}[X]$ ,

en prenant  $P = \sum_{k=0}^n \frac{a_k}{k+1} X^{k+1}$  on a  $g(P) = Q$  donc tout élément de  $\mathbb{R}[X]$  admet un antécédent par  $g$ ,

$g$  est surjective

- $g(1) = 0_{\mathbb{R}[X]}$  donc  $\text{Ker}(g) \neq \{0_{\mathbb{R}[X]}\}$  et ainsi  $g$  n'est pas injective

*Cette dernière remarque est difficile pour un élève de BCPST car on fait essentiellement des exercices en dimension finie. A vous de montrer que vous êtes capable de retenir ces parties du cours, je vous assure c'est possible vous pouvez le faire.*

## 6.5 Représentation matricielle.

Soient  $E$  et  $F$  deux  $\mathbb{K}$  espaces vectoriels de dimension finie.

On notera si nécessaire :  $p = \dim(E)$ ,  $\mathcal{B} = (e_1, \dots, e_p)$  une base de  $E$ ,  
 $n = \dim(F)$ ,  $\mathcal{B}' = (e'_1, \dots, e'_n)$  une base de  $F$ .

### 6.5.1 Introduction.

Soit  $f$  une application linéaire de  $E$  dans  $F$ .

Pour  $u = \sum_{k=1}^p x_k e_k \in E$ , comme  $f$  est linéaire on obtient :  $f(u) = \sum_{k=1}^p x_k f(e_k)$  et en appliquant  $\text{Coord}_{\mathcal{B}'}$  qui est

aussi linéaire il vient :  $\text{Coord}_{\mathcal{B}'}(f(u)) = \sum_{k=1}^p x_k \text{Coord}_{\mathcal{B}'}(f(e_k))$

Ce qui donne (*Raisonnement déjà vu* :  $AX = \sum_{i=1}^p x_i C_i$ ) :

$$\text{Coord}_{\mathcal{B}'}(f(u)) = \begin{pmatrix} \vdots & & \vdots & & \vdots \\ \text{Coord}_{\mathcal{B}'}(f(e_1)) & \cdots & \text{Coord}_{\mathcal{B}'}(f(e_j)) & \cdots & \text{Coord}_{\mathcal{B}'}(f(e_p)) \\ \vdots & & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

### 6.5.2 Matrice d'une application linéaire.

#### Définition

Soient  $\mathcal{B} = (e_1, \dots, e_p)$  une base de  $E$ ,  $\mathcal{B}'$  une base de  $F$  et  $f \in \mathcal{L}(E, F)$ ,  
on appelle **matrice de  $f$  dans les bases  $\mathcal{B}$  et  $\mathcal{B}'$**  la matrice (notée :  $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$ )

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) = \text{Mat}_{\mathcal{B}'}(f(e_1), \dots, f(e_p)) \quad \text{ou encore} \quad \text{Mat}_{\mathcal{B}'}(f(\mathcal{B}))$$

#### Remarques :

- La matrice  $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$  dépend du choix des bases  $\mathcal{B}$  et  $\mathcal{B}'$ .
- On dit que la matrice  $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$  représente  $f$  dans les bases  $\mathcal{B}$  et  $\mathcal{B}'$ .
- Lorsque  $f$  est un endomorphisme de  $E$  et que  $\mathcal{B} = \mathcal{B}'$  on note :  $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) = \text{Mat}_{\mathcal{B}}(f)$ .
- La matrice de l'application nulle est la matrice nulle.

Illustration permettant de retenir la définition.

$$\text{Mat}_{\mathcal{B},\mathcal{B}'}(f) = \begin{pmatrix} \vdots & \vdots & \vdots \\ \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix} \left. \begin{array}{l} \rightarrow e'_1 \\ \vdots \\ \rightarrow e'_i \\ \vdots \\ \rightarrow e'_n \end{array} \right\} \mathcal{B}'$$

$f(\mathcal{B})$   
 $\overbrace{f(e_1) \quad f(e_j) \quad f(e_p)}$   
 $\downarrow \quad \downarrow \quad \downarrow$

**Théorème :**

Pour  $\mathcal{B}$  une base de  $E$  et  $\mathcal{B}'$  une base de  $F$  (deux bases fixées)  
 l'application  $f \mapsto \text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  est une bijection de  $\mathcal{L}(E, F)$  dans  $\mathcal{M}_{n,p}(\mathbb{R})$

**En effet :** Une application linéaire est entièrement définie par l'image d'une base.

**Remarques :**

- ❶ Pour montrer que  $f = g$ , il suffit de montrer que pour des bases  $\mathcal{B}$  et  $\mathcal{B}'$  quelconques,

$$\text{Mat}_{\mathcal{B},\mathcal{B}'}(f) = \text{Mat}_{\mathcal{B},\mathcal{B}'}(g)$$

- ❷ On peut définir une application linéaire en donnant sa matrice dans deux bases quelconques.

**Théorème :** (Théorème très utilisé dans ce cours, plus rarement en exercice) On le notera "**Théorème (\*)**"

Soient  $\mathcal{B}$  une base de  $E$ , et  $\mathcal{B}'$  une base de  $F$  et  $f$  une application linéaire de  $E$  dans  $F$ ,  
 $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  est l'unique matrice  $A \in \mathcal{M}_{n,p}(\mathbb{K})$  vérifiant :

$$\forall u \in E, \quad \text{Coord}_{\mathcal{B}'}(f(u)) = A \text{Coord}_{\mathcal{B}}(u)$$

**En effet :** (Fait au tableau)

- On a vu dans l'introduction que  $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  convient
- Si  $A_1$  et  $A_2$  conviennent alors  $\forall u \in E, A_1 \text{Coord}_{\mathcal{B}}(u) = A_2 \text{Coord}_{\mathcal{B}}(u)$  donc  $\forall X \in \mathcal{M}_{p,1}(\mathbb{K}), A_1 X = A_2 X$   
 En prenant  $X_1, \dots, X_p$  la base canonique de  $\mathcal{M}_{p,1}$  on montre que  $A_1$  et  $A_2$  ont les mêmes colonnes donc  $A_1 = A_2$ .  
 (d'où l'unicité.)

Dans la suite de ce chapitre on utilisera **Théorème (\*)** sous la forme :

Si on trouve une matrice  $A$  vérifiant  $\forall u \in E, \text{Coord}_{\mathcal{B}'}(f(u)) = A \text{Coord}_{\mathcal{B}}(u)$

alors on pourra affirmer que  $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f) = A$

**Relation importante :** Pour  $f \in \mathcal{L}(E, F)$ ,  $\mathcal{B}$  une base de  $E$  et  $\mathcal{B}'$  une base de  $F$

$$\forall u \in E, \quad \text{Coord}_{\mathcal{B}'}(f(u)) = \text{Mat}_{\mathcal{B},\mathcal{B}'}(f) \text{Coord}_{\mathcal{B}}(u)$$

Avec son corollaire plus couramment utilisé : Pour  $f \in \mathcal{L}(E)$ ,  $\mathcal{B}$  une base de  $E$

$$\forall u \in E, \quad \text{Coord}_{\mathcal{B}}(f(u)) = \text{Mat}_{\mathcal{B}}(f) \text{Coord}_{\mathcal{B}}(u)$$

## 6.5.3 Matrices et opérations.

Matrice d'une combinaison de deux applications linéaires.**Théorème**

Soient  $\mathcal{B}$  une base de  $E$  et  $\mathcal{B}'$  une base de  $F$ ,  
 Pour tout  $(f, g) \in \mathcal{L}(E, F)^2$  et pour tout  $(\alpha, \beta) \in \mathbb{K}^2$  on a :

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(\alpha f + \beta g) = \alpha \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) + \beta \text{Mat}_{\mathcal{B}, \mathcal{B}'}(g)$$

**En effet :**  $\forall u \in E$ ,  $\text{Coord}_{\mathcal{B}'}((\alpha f + \beta g)(u)) = \text{Coord}_{\mathcal{B}'}(\alpha f(u) + \beta g(u))$   
 $= \alpha \text{Coord}_{\mathcal{B}'}(f(u)) + \beta \text{Coord}_{\mathcal{B}'}(g(u))$   
 $= \alpha \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) \text{Coord}_{\mathcal{B}}(u) + \beta \text{Mat}_{\mathcal{B}, \mathcal{B}'}(g) \text{Coord}_{\mathcal{B}}(u)$   
 $= \underbrace{(\alpha \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) + \beta \text{Mat}_{\mathcal{B}, \mathcal{B}'}(g))}_{\text{Mat}_{\mathcal{B}, \mathcal{B}'}(\alpha f + \beta g)} \text{Coord}_{\mathcal{B}}(u)$

En utilisant le **Théorème (\*)** il vient :  $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(\alpha f + \beta g) = \alpha \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) + \beta \text{Mat}_{\mathcal{B}, \mathcal{B}'}(g)$

**Conséquence :**

Soient  $\mathcal{B}$  une base de  $E$  et  $\mathcal{B}'$  une base de  $F$ ,  
 l'application  $f \mapsto \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$  est un isomorphisme de  $\mathcal{L}(E, F)$  dans  $\mathcal{M}_{n,p}(\mathbb{K})$

**En effet :** On vient de démontrer qu'elle est linéaire et un peu plus haut on a vu qu'elle est bijective.

Matrice de la composée de deux applications linéaires.**Théorème**

Soient  $\mathcal{B}_1$  une base de  $E$ ,  $\mathcal{B}_2$  une base de  $F$  et  $\mathcal{B}_3$  une base de  $G$ ,  
 Pour tout  $f \in \mathcal{L}(E, F)$  et  $g \in \mathcal{L}(F, G)$  on a :  $\text{Mat}_{\mathcal{B}_1, \mathcal{B}_3}(g \circ f) = \text{Mat}_{\mathcal{B}_2, \mathcal{B}_3}(g) \text{Mat}_{\mathcal{B}_1, \mathcal{B}_2}(f)$

**Démonstration :** pour tout  $u \in E$ ,

$$\begin{aligned} \text{Coord}_{\mathcal{B}_3}((g \circ f)(u)) &= \text{Coord}_{\mathcal{B}_3}(g(f(u))) \\ &= \text{Mat}_{\mathcal{B}_2, \mathcal{B}_3}(g) \text{Coord}_{\mathcal{B}_2}(f(u)) \\ &= \text{Mat}_{\mathcal{B}_2, \mathcal{B}_3}(g) \text{Mat}_{\mathcal{B}_1, \mathcal{B}_2}(f) \text{Coord}_{\mathcal{B}_1}(u). \end{aligned}$$

D'après le théorème (\*),

$$\text{Mat}_{\mathcal{B}_1, \mathcal{B}_3}(g \circ f) = \text{Mat}_{\mathcal{B}_2, \mathcal{B}_3}(g) \text{Mat}_{\mathcal{B}_1, \mathcal{B}_2}(f).$$

**Corollaire** (trois endomorphismes et quatre bases)

Soient  $f, g, h \in \mathcal{L}(E)$  et  $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$  et  $\mathcal{B}_4$  quatre bases de  $E$ .

$$\text{Mat}_{\mathcal{B}_1, \mathcal{B}_4}(h \circ g \circ f) = \text{Mat}_{\mathcal{B}_3, \mathcal{B}_4}(h) \text{Mat}_{\mathcal{B}_2, \mathcal{B}_3}(g) \text{Mat}_{\mathcal{B}_1, \mathcal{B}_2}(f)$$

**En effet :** (à faire)

**Corollaire** (Puissance d'un endomorphisme avec une seule base).

Pour  $f$  un endomorphisme de  $E$ ,  $m$  un entier naturel et  $\mathcal{B}$  une base de  $E$ ,

$$\text{Mat}_{\mathcal{B}}(f^m) = (\text{Mat}_{\mathcal{B}}(f))^m$$

**En effet :** (à faire)

**Matrice de la réciproque d'une application linéaire bijective.****Théorème**

Soient  $\mathcal{B}$  une base de  $E$ ,  $\mathcal{B}'$  une base de  $F$  et  $f$  une application linéaire de  $E$  dans  $F$ .

$f$  est bijective si, et seulement si,  $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  est inversible

et alors :

$$\text{Mat}_{\mathcal{B}',\mathcal{B}}(f^{-1}) = (\text{Mat}_{\mathcal{B},\mathcal{B}'}(f))^{-1}$$

**Démonstration :**

- Supposons que  $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  est inversible.

Prenons un vecteur  $v$  de  $F$  et cherchons ses antécédents par  $f$  :

Pour  $u \in E$ ,

$$\begin{aligned} f(u) = v &\iff \text{Mat}_{\mathcal{B},\mathcal{B}'}(f)\text{Coord}_{\mathcal{B}}(u) = \text{Coord}_{\mathcal{B}'}(v) \\ &\iff \text{Coord}_{\mathcal{B}}(u) = \text{Mat}_{\mathcal{B},\mathcal{B}'}(f)^{-1}\text{Coord}_{\mathcal{B}'}(v) \\ &\iff u = \text{Coord}_{\mathcal{B}}^{-1}(\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)^{-1}\text{Coord}_{\mathcal{B}'}(v)) \end{aligned}$$

Tout élément de  $F$  admet un unique antécédent par  $f$ , donc  $f$  est une bijection ■

- Supposons que  $f$  est une bijection,
  - En utilisant le théorème 1, on sait que  $f(\mathcal{B})$  est une base de  $F$  et ainsi  $\dim(E) = \dim(F)$  donc  $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f) \in \mathcal{M}_n(\mathbb{K})$  (matrice carrée)
  - On a aussi vu que  $f^{-1}$  est linéaire. (Voir ci-dessous)
  - En utilisant le résultat sur la matrice d'une composée :

$$\begin{aligned} \text{Mat}_{\mathcal{B},\mathcal{B}'}(f)\text{Mat}_{\mathcal{B}',\mathcal{B}}(f^{-1}) &= \text{Mat}_{\mathcal{B}'}(f \circ f^{-1}) \\ &= \text{Mat}_{\mathcal{B}'}(Id_F) \\ &= I_n \end{aligned}$$

$$\begin{aligned} \text{Mat}_{\mathcal{B}',\mathcal{B}}(f^{-1})\text{Mat}_{\mathcal{B},\mathcal{B}'}(f) &= \text{Mat}_{\mathcal{B}}(f^{-1} \circ f) \\ &= \text{Mat}_{\mathcal{B}}(Id_E) \\ &= I_n \end{aligned}$$

on peut en conclure que  $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  est inversible et que  $(\text{Mat}_{\mathcal{B},\mathcal{B}'}(f))^{-1} = \text{Mat}_{\mathcal{B}',\mathcal{B}}(f^{-1})$  ■

Ce qui achève la démonstration de :

$f$  est bijective si, et seulement si,  $\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  est inversible

**6.6 Isomorphe à  $\mathbb{K}^n$ .****Définition.**

Soient  $E$  et  $F$  deux espaces vectoriels,

Dire que  $E$  et  $F$  sont isomorphes signifie qu'il existe un isomorphisme de  $E$  dans  $F$ .

**Théorème.**

Pour  $n$  un entier naturel non nul,

tout espace vectoriel de dimension  $n$  est isomorphe à  $\mathbb{K}^n$ .

**Conséquence.**

Soient  $E$  et  $F$  deux espaces vectoriels de dimension finie,

$E$  et  $F$  sont isomorphes si, et seulement si,  $\dim(E) = \dim(F)$ .

**Théorème :**

Soient  $E$  un espace de dimension  $n$  et  $\mathcal{B}$  une famille de vecteurs de  $E$ .  
 Si  $\mathcal{B}$  est une base de  $E$  alors  
 l'application  $E \rightarrow \mathcal{M}_{n,1}(\mathbb{K})$  est un isomorphisme.  
 $v \mapsto \text{Coord}_{\mathcal{B}}(v)$

**Démonstration :** (Vue dans le cours sur les espaces vectoriels)

**Théorème :**

Soient  $E$  un espace de dimension finie,  $\mathcal{B}$  une base de  $E$  et  $u_1, \dots, u_n$  des vecteurs de  $E$ .

- ❶ Pour tout  $v \in E$ ,  $v \in \text{Vect}(u_1, \dots, u_n) \iff \text{Coord}_{\mathcal{B}}(v) \in \text{Vect}(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$
- ❷  $(u_1, \dots, u_n)$  est libre si, et seulement si,  $(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$  est libre.
- ❸  $\text{rg}(u_1, \dots, u_n) = \text{rg}(\text{Coord}_{\mathcal{B}}(u_1), \dots, \text{Coord}_{\mathcal{B}}(u_n))$ .

**Démonstration.** (Vue dans le cours sur les espaces vectoriels)

## 6.7 Rang d'une application linéaire.

Soient  $E$  et  $F$  deux  $\mathbb{K}$  espaces vectoriels de dimension finie. On note  $p = \dim(E)$  et  $n = \dim(F)$ .

### 6.7.1 Définition.

**Définition :**

Soit  $f$  une application linéaire de  $E$  dans  $F$ ,  
 on appelle rang de l'application linéaire  $f$  l'entier naturel noté  $\text{rg}(f)$  et défini par :

$$\text{rg}(f) = \dim(\text{Im}(f))$$

### 6.7.2 Lien avec les autres notions de rang.

On note  $\mathcal{B} = (e_1, e_2, \dots, e_p)$  une base de  $E$  et  $\mathcal{B}'$  une base de  $F$ .

$$\begin{aligned}
 \text{rg}(f) &= \dim(\text{Im}(f)) \\
 &= \dim(\text{Vect}(f(e_1), \dots, f(e_p))) \\
 &= \text{rg}(f(e_1), \dots, f(e_p)) \\
 &= \text{rg}(\text{Coord}_{\mathcal{B}'}(f(e_1)), \dots, \text{Coord}_{\mathcal{B}'}(f(e_p))) \\
 &= \text{rg}(\text{Mat}_{\mathcal{B}'}(f(e_1), \dots, f(e_p))) \\
 &= \text{rg}(\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)).
 \end{aligned}$$

### 6.7.3 Théorème du rang

**Théorème :**

Soient  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension finie et  $F$  un  $\mathbb{K}$ -espace vectoriel.

$$\forall f \in \mathcal{L}(E, F), \quad \text{rg}(f) + \dim(\ker(f)) = \dim(E)$$

**Démonstration :** Ex 4 de la feuille Cours\_6\_2.

### 6.7.4 Caractérisation des isomorphismes

#### Théorème :

Soient  $E$  et  $F$  deux  $\mathbb{K}$ -espaces vectoriels de dimension  $n$  et  $f \in \mathcal{L}(E, F)$   
 $f$  est bijective si, et seulement si,  $\text{rg}(f) = n$

**En effet :** Le théorème du rang donne  $\text{rg}(f) = n \iff \ker(f) = \{0_E\}$  et ainsi  $\text{rg}(f) = n \iff f$  injective.  
 Avec le paragraphe 6.4 il vient  $\text{rg}(f) = n$  si, et seulement si,  $f$  est bijective.

#### Conséquences :

Soient  $E$  et  $F$  deux  $\mathbb{K}$ -espaces vectoriels de dimension finie et  $f \in \mathcal{L}(E, F)$

- $f$  injective  $\iff \text{rg}(f) = \dim(E)$
- $f$  surjective  $\iff \text{rg}(f) = \dim(F)$

## 6.8 Une matrice vue comme une application linéaire.

Ici  $n$  et  $p$  désignent deux entiers naturels non nuls.

### 6.8.1 Application linéaire canoniquement associée à une matrice.

#### Définition.

Pour toute matrice  $M$  de  $\mathcal{M}_{n,p}(\mathbb{K})$  on appelle application canoniquement associée à  $M$ , l'application :

$$\begin{array}{ccc} \mathcal{M}_{p,1}(\mathbb{K}) & \longrightarrow & \mathcal{M}_{n,1}(\mathbb{K}) \\ X & \longmapsto & MX \end{array}$$

#### Propriété

Pour toute matrice  $M$  de  $\mathcal{M}_{n,p}(\mathbb{K})$  l'application :  $\mathcal{M}_{p,1}(\mathbb{K}) \longrightarrow \mathcal{M}_{n,1}(\mathbb{K})$  est linéaire.  
 $X \longmapsto MX$

#### Démonstration :

On note  $f$  cette application.

Soient  $(\alpha, \beta) \in \mathbb{K}^2$  et  $(X_1, X_2) \in \mathcal{M}_{p,1}(\mathbb{K})^2$ ,

$$\begin{aligned} f(\alpha X_1 + \beta X_2) &= M(\alpha X_1 + \beta X_2) \\ &= \alpha M X_1 + \beta M X_2 && \text{(cours sur les matrices)} \\ &= \alpha f(X_1) + \beta f(X_2) \end{aligned}$$

$f$  est une application linéaire

#### Remarque :

En notant :  $\mathcal{B}$  la base canonique de  $\mathcal{M}_{p,1}(\mathbb{K})$ ,  $\mathcal{B}'$  la base canonique de  $\mathcal{M}_{n,1}(\mathbb{K})$ .

La matrice de  $f$  dans les bases  $\mathcal{B}$  et  $\mathcal{B}'$  est égale à  $M$ .

$$(\mathcal{M}_{\mathcal{B}, \mathcal{B}'}(f) = M)$$

### 6.8.2 Noyau, image.

#### Définition :

Pour toute matrice  $M$  de  $\mathcal{M}_{n,p}(\mathbb{K})$ , on définit le noyau et l'image de  $M$  par :

$$\ker(M) = \left\{ X \in \mathcal{M}_{p,1}(\mathbb{K}) \mid MX = 0 \right\} \quad \text{Im}(M) = \left\{ MX \mid X \in \mathcal{M}_{p,1}(\mathbb{K}) \right\}$$

**Remarques :**

- En notant  $f : \mathcal{M}_{p,1}(\mathbb{K}) \longrightarrow \mathcal{M}_{n,1}(\mathbb{K})$  on a :  $\ker(M) = \ker(f)$  et  $\text{Im}(M) = \text{Im}(f)$   

$$X \longmapsto MX$$
- $\ker(M) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K}) \mid M \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}$  et  $\text{Im}(M) = \left\{ M \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \mid (x_1, \dots, x_p) \in \mathbb{K}^p \right\}$
- $\ker(M)$  et  $\text{Im}(M)$  sont des espaces de matrices colonnes.  $\ker(M) \subset \mathcal{M}_{p,1}(\mathbb{K})$  et  $\text{Im}(M) \subset \mathcal{M}_{n,1}(\mathbb{K})$
- Pour tout  $Y \in \mathcal{M}_{n,1}(\mathbb{K})$ ,  $Y \in \text{Im}(M) \iff \exists X \in \mathcal{M}_{p,1}(\mathbb{K}) : Y = MX$
- Le plus souvent pour déterminer une base de  $\ker(M)$  on résout le système homogène  $MX = 0$ .  
*(Méthode du pivot sur les lignes)*
- Le plus souvent pour déterminer une base  $\text{Im}(M)$  on utilise  $\text{Im}(M) = \text{Vect}(\underbrace{C_1, \dots, C_p}_{\text{Le colonnes de } M})$  et ensuite :
  - ❶ On utilise les relations entre les colonnes données par l'étude du noyau pour supprimer des colonnes.
  - ou
  - ❷ On utilise la méthode du pivot sur les colonnes de  $M$ .
  - ou
  - ❸ On connaît la dimension  $m$  avec le théorème du rang et on trouve une famille libre de  $m$  vecteurs de  $\text{Im}(M)$ .

**Proposition :**

*Lien entre noyau et image d'une matrice et d'une application linéaire représentée par cette matrice dans des bases.*

Soient  $f \in \mathcal{L}(E, F)$ ,  $\mathcal{B}$  une base de  $E$  et  $\mathcal{B}'$  une base de  $F$  :

$$\forall u \in E, \quad u \in \ker(f) \iff \text{Coord}_{\mathcal{B}}(u) \in \ker(\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f))$$

$$\forall v \in F, \quad v \in \text{Im}(f) \iff \text{Coord}_{\mathcal{B}'}(v) \in \text{Im}(\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f))$$

**Exemples :**

1. Soit  $\Delta$  l'application linéaire de  $\mathbb{R}_2[X]$  dans  $\mathbb{R}_2[X]$  qui à  $P(X)$  associe  $P(X + 1) - P(X - 1)$ .  
*(on ne démontrera pas que  $\Delta$  est bien linéaire).*  
 Déterminer une base du noyau et une base de l'image de  $\Delta$ .
2. Soit  $f$  l'application linéaire de  $\mathbb{R}_3[X]$  dans  $\mathbb{R}^3$  qui à  $P(X)$  associe  $(P(0), P(1), P(2))$ .  
*(on ne démontrera pas que  $f$  est bien linéaire).*  
 Déterminer une base du noyau et une base de l'image de  $f$ .

**6.8.3 Rang d'une matrice.**

**Définition :** Le rang de  $M$  est la dimension de l'espace engendré par les colonnes de  $M$

C'est le rang de l'application linéaire :  $\mathcal{M}_{p,1}(\mathbb{K}) \longrightarrow \mathcal{M}_{n,1}(\mathbb{K})$   

$$X \longmapsto MX$$

Dans l'algorithme du pivot de Gauss pour la résolution du système homogène  $MX = 0$  :  
*(système  $n$  équations,  $p$  inconnues)*

- $\text{rg}(M)$  est le nombre d'inconnues principales, (ou encore le nombre de pivots)
- $\dim(\ker(M))$  est le nombre d'inconnues secondaires,

**Théorème :**

Pour toute matrice  $M$  de  $\mathcal{M}_{n,p}(\mathbb{K})$  :  $\dim(\text{Im}(M)) + \dim(\ker(M)) = p$

**En effet :**

**Attention :**  $p$  est le nombre de colonnes de  $M$ .

**Corollaire :**

Pour toute matrice  $M$  de  $\mathcal{M}_n(\mathbb{K})$  :  $\dim(\text{Im}(M)) + \dim(\ker(M)) = n$

### 6.8.4 Inverse à gauche, inverse à droite.

**Théorème.**

Pour  $A \in \mathcal{M}_n(\mathbb{K})$  et  $B \in \mathcal{M}_n(\mathbb{K})$  deux matrices carrées,  $AB = I_n \iff BA = I_n$

**Démonstration :** (Il suffit de montrer une deux implications).

On suppose que  $AB = I_n$  et on note :  $f : \mathcal{M}_{n,1}(\mathbb{K}) \rightarrow \mathcal{M}_{n,1}(\mathbb{K})$ .

$$X \mapsto BX$$

$\text{Ker}(B) \subset \text{ker}(AB)$  cela entraîne ici que  $\text{ker}(B) = \{0_{n \times 1}\}$  donc  $\text{rg}(B) = n$  et ainsi  $B$  est inversible. et alors la relation  $AB = I_n$  entraîne  $ABB^{-1} = B^{-1}$  donc  $A = B^{-1}$  et ainsi  $BA = I_n$

## 6.9 Changement de base.

### 6.9.1 Matrice de passage.

**Définition et notation**

Soient  $\mathcal{B}$  et  $\mathcal{B}'$  deux bases de  $E$ ,  
la matrice  $\text{Mat}_{\mathcal{B}}(\mathcal{B}')$  est appelée matrice de passage de  $\mathcal{B}$  à  $\mathcal{B}'$ , on la note  $P_{\mathcal{B},\mathcal{B}'}$

**Remarques :**

- ❶ Quand on donne une nouvelle base  $\mathcal{B}'$ , on donne la matrice  $\text{Mat}_{\mathcal{B}}(\mathcal{B}')$ .  
*La matrice des vecteurs de la nouvelle base dans l'ancienne.*
- ❷ Pour les démonstrations on pourra utiliser la remarque  $P_{\mathcal{B},\mathcal{B}'} = \text{Mat}_{\mathcal{B}',\mathcal{B}}(Id_E)$

**Propriétés.**

Quelles que soient  $\mathcal{B}$  et  $\mathcal{B}'$  des bases de  $E$ ,

- ❶  $P_{\mathcal{B},\mathcal{B}} = I_n$
- ❷  $P_{\mathcal{B},\mathcal{B}'}$  est inversible et  $P_{\mathcal{B},\mathcal{B}'}^{-1} = P_{\mathcal{B}',\mathcal{B}}$

**En effet :** ❶ Pour tout  $i \in \llbracket 1, n \rrbracket$ ,  $\text{Coord}_{\mathcal{B}}(e_i) = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{\text{ième}} \text{ donc } \boxed{P_{\mathcal{B},\mathcal{B}} = I_n}$

❷  $P_{\mathcal{B}',\mathcal{B}}P_{\mathcal{B},\mathcal{B}'} = \text{Mat}_{\mathcal{B},\mathcal{B}'}(Id_E)\text{Mat}_{\mathcal{B}',\mathcal{B}}(Id_E) = \text{Mat}_{\mathcal{B}',\mathcal{B}'}(Id_E) = I_n$  donc

$P_{\mathcal{B},\mathcal{B}'}$  est inversible et  $P_{\mathcal{B},\mathcal{B}'}^{-1} = P_{\mathcal{B}',\mathcal{B}}$

**Remarque :** Toute matrice inversible est la matrice de passage entre deux bases bien choisies.

**Propriétés.** (Complément)

Quelles que soient  $\mathcal{B}$ ,  $\mathcal{B}'$  et  $\mathcal{B}''$  trois bases de  $E$ ,  $P_{\mathcal{B},\mathcal{B}'}P_{\mathcal{B}',\mathcal{B}''} = P_{\mathcal{B},\mathcal{B}''}$

**Démonstration.** Idée :  $\text{Mat}_{\mathcal{B},\mathcal{B}''}(Id_E) = \text{Mat}_{\mathcal{B}',\mathcal{B}''}(Id_E) \text{Mat}_{\mathcal{B},\mathcal{B}'}(Id_E)$

### 6.9.2 Changement de bases, action sur les coordonnées d'un vecteur.

**Théorème**

Quelles que soient  $\mathcal{B}$  et  $\mathcal{B}'$  deux bases de  $E$ ,  $\forall u \in E$ ,  $\text{Coord}_{\mathcal{B}}(u) = P_{\mathcal{B},\mathcal{B}'} \text{Coord}_{\mathcal{B}'}(u)$

**Démonstration :**

Soit  $u = \sum_{k=1}^n x'_k e'_k$  un vecteur quelconque exprimé sur la base  $\mathcal{B}'$ .

donc en appliquant l'application linéaire  $\text{Coord}_{\mathcal{B}}(\cdot)$  on obtient  $\text{Coord}_{\mathcal{B}}(u) = \sum_{k=1}^n x'_k \text{Coord}_{\mathcal{B}}(e'_k)$

or (remarque du paragraphe 3.6.2)  $\sum_{k=1}^n x'_k \text{Coord}_{\mathcal{B}}(e'_k) = \text{Mat}_{\mathcal{B}}(\mathcal{B}') \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$

donc

$$\text{Coord}_{\mathcal{B}}(u) = P_{\mathcal{B},\mathcal{B}'} \text{Coord}_{\mathcal{B}'}(u)$$

**Remarques :**

- en notant :  $X = \text{Coord}_{\mathcal{B}}(u)$ ,  $X' = \text{Coord}_{\mathcal{B}'}(u)$  et  $P = P_{\mathcal{B},\mathcal{B}'}$

$$X = PX' \quad X' = P^{-1}X$$

- Certains trouvent le nom "matrice de passage de  $\mathcal{B}$  à  $\mathcal{B}'$  contre-intuitif.

en effet : pour passer de  $X$  à  $X'$  on applique  $X \mapsto P_{\mathcal{B}',\mathcal{B}}X$

**6.9.3 Changement de bases, action sur la matrice d'un endomorphisme.****Théorème**

Soient  $\mathcal{B}$  et  $\mathcal{B}'$  deux bases de  $E$ .

$$\forall f \in \mathcal{L}(E) \quad \text{Mat}_{\mathcal{B}'}(f) = P_{\mathcal{B}',\mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B},\mathcal{B}'}$$

**En effet :** Pour  $u \in E$ ,

$$\begin{aligned} \text{Coord}_{\mathcal{B}'}(f(u)) &= P_{\mathcal{B}',\mathcal{B}} \text{Coord}_{\mathcal{B}}(f(u)) \\ &= P_{\mathcal{B}',\mathcal{B}} \text{Mat}_{\mathcal{B}}(f) \text{Coord}_{\mathcal{B}}(u) \end{aligned}$$

$$\forall u \in E, \quad \text{Coord}_{\mathcal{B}'}(f(u)) = \underbrace{P_{\mathcal{B}',\mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B},\mathcal{B}'}}_{\text{Mat}_{\mathcal{B}'}(f)} \text{Coord}_{\mathcal{B}'}(u)$$

On peut en déduire que :  $\text{Mat}_{\mathcal{B}'}(f) = P_{\mathcal{B}',\mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B},\mathcal{B}'}$

Ici il faut se rappeler que :

$\text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$  est l'unique matrice  $A \in \mathcal{M}_{n,p}(\mathbb{R})$  vérifiant :  $\forall u \in E, \quad \text{Coord}_{\mathcal{B}'}(f(u)) = A \text{Coord}_{\mathcal{B}}(u)$

- Il y a plusieurs démonstrations de cette formule, mais à la fin il faut trouver un moyen de la retenir .

$$\text{Mat}_{\mathcal{B}'}(f) = P_{\mathcal{B}',\mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B},\mathcal{B}'}$$

- Si on note  $M = \text{Mat}_{\mathcal{B}}(f)$ ,  $M' = \text{Mat}_{\mathcal{B}'}(f)$  et  $P = P_{\mathcal{B},\mathcal{B}'}$  alors :

$$M' = P^{-1}MP$$

et

$$M = PM'P^{-1}$$

**Remarque :** (Plus compréhensible après le chapitre sur la **diagonalisation**).

On passe souvent d'une base  $\mathcal{B}$  à une base  $\mathcal{B}'$  dans laquelle la matrice  $M' = \Delta$  est plus simple

(diagonale ou triangulaire).

on a alors avec la formule du changement de base en notant  $P = P_{\mathcal{B},\mathcal{B}'}$  :

$$\Delta = P^{-1}MP$$

on en déduit la relation :

$$M = P\Delta P^{-1}$$

## 6.10 Matrices semblables

Dans ce paragraphe toutes les matrices sont carrées.

**Définition :**

Soient  $M$  et  $N$  deux matrices de  $\mathcal{M}_n(\mathbb{K})$ ,

dire que  $M$  est semblable à  $N$  signifie qu'il existe une matrice  $P$  inversible de  $\mathcal{M}_n(\mathbb{K})$  telle que :

$$M = P^{-1}NP$$

**Exemple et contre exemple.**

1.  $M = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$  et  $N = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$  sont semblables.
2.  $M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  et  $N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  ne sont pas semblables.

**Attention :** une telle matrice  $P$  n'est pas unique.

**En effet :**

**Proposition**

1.  $M$  est semblable à  $M$ . (*réflexif*)
2. Si  $M$  est semblable à  $N$  alors  $N$  est semblable à  $M$ . (*symétrique*)
3. Si  $M_1$  est semblable à  $M_2$  et  $M_2$  est semblable à  $M_3$  alors  $M_1$  est semblable à  $M_3$ .  
(*transitive*)

**Démonstration.**

**Caractérisation.**

Soient  $M$  et  $N$  deux matrices de  $\mathcal{M}_n(\mathbb{K})$ ,

$M$  et  $N$  sont semblables si, et seulement si,

elles sont les matrices d'un même endomorphisme dans deux bases.

**Démonstration.**

**Autrement dit :**

Deux matrices  $M$  et  $N$  de  $\mathcal{M}_n(\mathbb{K})$  sont semblables si, et seulement si, :

pour un espace  $E$  de dimension  $n$ , un endomorphisme  $f$  de  $E$  et deux bases  $\mathcal{B}$ ,  $\mathcal{B}'$  de  $E$ , on a :

$$M = \text{Mat}_{\mathcal{B}'}(f) \quad \text{et} \quad N = \text{Mat}_{\mathcal{B}}(f)$$

$M$  et  $N$  représentent le même endomorphisme dans deux bases.

**Théorème :**

Si deux matrices  $M$  et  $N$  sont semblables avec pour  $P$  inversible la relation  $N = P^{-1}MP$  alors

$$\forall n \in \mathbb{N}, \quad N^n = P^{-1}M^nP$$

**Sauf indication contraire vous devez refaire la démonstration dans une copie.**

**Démonstration 1.**

Montrons par récurrence sur  $n \in \mathbb{N}$  que pour tout  $n \in \mathbb{N}$ ,  $\underbrace{M^n = P^{-1}N^nP}_{\text{notée } \mathcal{R}(n)}$

- Pour  $n = 0$ ,  
d'une part  $M^0 = I_n$  et d'autre part  $PN^0P^{-1} = P^{-1}I_nP = P^{-1}P = I_n$

on a bien  $\mathcal{R}(0)$  vraie

- Soit  $n \in \mathbb{N}$  tel que  $\mathcal{R}(n)$  est vraie,

$$\begin{aligned} M^{n+1} &= M^n \times M \\ &= P^{-1}N^n P \times M && \text{(d'après l'hypothèse de récurrence)} \\ &= P^{-1}N^n P \times P^{-1}NP && \text{(d'après la donnée de l'énoncé)} \\ &= P^{-1}N^n I_n NP \\ &= P^{-1}N^{n+1}P && \text{ce qui donne } \mathcal{R}(n+1) \end{aligned}$$

En conclusion :

$$\boxed{\text{pour tout } n \in \mathbb{N}, \quad M^n = P^{-1}N^n P}$$

**Démonstration 2.** Pour les champions du changement de base.

*Je ne vous conseillerais pas encore cette démonstration, mais c'est bien de savoir qu'elle existe.*

On note  $\mathcal{B}, \mathcal{B}'$  deux bases de  $\mathcal{M}_{n,1}(\mathbb{K})$  et  $f$  l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbb{K})$  tels que :

$$P = P_{\mathcal{B},\mathcal{B}'} \quad \text{et} \quad N = \text{Mat}_{\mathcal{B}}(f)$$

comme  $M = P^{-1}NP$ , il vient :  $M = \text{Mat}_{\mathcal{B}'}(f)$ , (Effet du changement de base sur la matrice d'un endomorphisme)  
 Pour  $n \in \mathbb{N}$ ,

$$N = \text{Mat}_{\mathcal{B}}(f) \text{ donc } N^n = \text{Mat}_{\mathcal{B}}(f^n) \quad \text{et} \quad M = \text{Mat}_{\mathcal{B}'}(f) \quad \text{donc} \quad M^n = \text{Mat}_{\mathcal{B}'}(f^n)$$

or on sait (Effet du changement de base sur la matrice d'un endomorphisme) que :

$$\text{Mat}_{\mathcal{B}'}(f^n) = P_{\mathcal{B},\mathcal{B}'}^{-1} \text{Mat}_{\mathcal{B}}(f^n) P_{\mathcal{B},\mathcal{B}'}$$

on obtient bien :

$$\boxed{\text{pour tout } n \in \mathbb{N}, \quad M^n = P^{-1}N^n P}$$

**Remarques :**

- si  $M$  et  $N$  sont semblables alors pour tout  $n \in \mathbb{N}$ ,  $M^n$  et  $N^n$  le sont aussi.
- On utilise les matrices semblables pour calculer les puissances de matrice.

**Proposition**

$$\boxed{\text{Soit } \lambda \in \mathbb{K}, \quad M \text{ est semblable à } \lambda I_n \text{ si, et seulement si, } M = \lambda I_n.}$$

**En effet :**

$$\begin{aligned} M \text{ est semblable à } \lambda I_n &\iff \exists P \in \mathcal{M}_n(\mathbb{K}) \text{ (invertible)} : M = P^{-1}(\lambda I_n)P \\ &\iff \exists P \in \mathcal{M}_n(\mathbb{K}) \text{ (invertible)} : M = \lambda P^{-1}I_n P \\ &\iff \exists P \in \mathcal{M}_n(\mathbb{K}) \text{ (invertible)} : M = \lambda P^{-1}P \\ &\iff M = \lambda I_n \end{aligned}$$

**Théorème :** (complément)

$$\boxed{\begin{aligned} &\text{Soient } A \text{ et } B \text{ deux matrices de } \mathcal{M}_n(\mathbb{K}), \\ &\text{Si } A \text{ et } B \text{ sont semblables alors } \text{rg}(A) = \text{rg}(B). \end{aligned}}$$

**Démonstration.** voir feuille Cours\_6\_5 Ex 10

**Remarque :** La réciproque est fautive, voir feuille Cours\_6\_5 Ex 9

**Théorème :** (complément)

$$\boxed{\begin{aligned} &\text{Soient } A \text{ et } B \text{ deux matrices de } \mathcal{M}_n(\mathbb{K}), \\ &\text{Si } A \text{ et } B \text{ sont semblables alors } \text{tr}(A) = \text{tr}(B). \end{aligned}}$$

**Démonstration.** voir feuille Cours\_6\_5 Ex 6

# Intégrales généralisées

## Plan du chapitre

7.1	Introduction	88
7.2	Définitions. vocabulaire. notation.	89
7.2.1	Intégrales sur un intervalle semi-ouvert.	89
7.2.2	Intégrales sur un intervalle ouvert.	90
7.2.3	Cas particulier d'une fonction prolongeable par continuité.	91
7.2.4	Généralisation.	92
7.3	Théorèmes de convergence.	93
7.3.1	Théorème de convergence par comparaison des intégrandes positifs	93
7.3.2	Théorème de convergence par équivalence des intégrandes positifs.	94
7.4	Propriétés	95
7.4.1	Relation de Chasles.	95
7.4.2	Linéarité.	96
7.4.3	Positivité	97
7.4.4	Croissance de l'intégrale.	97
7.4.5	Stricte positivité.	97
7.4.6	Parité.	98
7.5	Absolute convergence.	99
7.6	Calculs.	100
7.6.1	Intégrations par parties.	100
7.6.2	Changement de variable.	100
7.7	Intégrales impropres usuelles.	102
7.7.1	Loi uniforme	102
7.7.2	Loi exponentielle	102
7.7.3	Loi normale	102
7.7.4	Compléments : Intégrales de Riemann.	102

## 7.1 Introduction

Ce que nous avons déjà vu :

- Interprétation géométrique de l'intégrale d'une fonction continue sur un intervalle.
- Lien avec les sommes de Riemann.
- Théorème fondamental de l'analyse et sa conséquence principale :  $\int_a^b f(t) dt = F(b) - F(a)$

Les objectifs de ce cours :

- Définir sous conditions :  $\int_a^b f(t) dt$  pour  $f$  continue sur  $]a, b[$ .
- Définir sous conditions :  $\int_{-\infty}^{+\infty} f(t) dt$  pour  $f$  continue sur  $\mathbb{R}$  sauf en un nombre fini de points.

- Etudier sous conditions :  $x \mapsto \int_{-\infty}^x f(t) dt$  pour  $f$  continue sur  $\mathbb{R}$  sauf en un nombre fini de points.

Comme pour les séries, il y faudra établir la convergence (l'existence) avant de faire le calcul.

On va retrouver les propositions du cours sur l'intégrale d'une fonction continue sur un segment :

Relation de Chasles, linéarité, positivité, croissance de l'intégrale, stricte positivité, parité.

La représentation reste l'aire de domaines du plan, mais ici ils ne sont pas bornés (*le plus souvent*).

Pour l'étude de la nature (*convergence, existence*) nous verrons deux théorèmes comme pour les séries.

La configuration principale de ces objectifs étant l'étude de deux fonctions  $f$  et  $F$  définies sur  $\mathbb{R}$  et vérifiant :

$$\forall x \in \mathbb{R}, \quad F(x) = \int_{-\infty}^x f(t) dt$$

où :  $f$  continue sur  $\mathbb{R}$  sauf éventuellement sur un ensemble fini.

$F$  continue sur  $\mathbb{R}$  tout entier et de classe  $C^1$  sur  $\mathbb{R}$  sauf éventuellement sur un ensemble fini.

En pratique pour l'étude de la nature et pour les calculs on va devoir s'adapter aux différentes situations :

- ❶  $\int_a^b f(t) dt$  pour  $f$  continue sur  $[a, b[$ . avec  $b \in \mathbb{R}$  ou  $b = +\infty$ .
- ❷  $\int_a^b f(t) dt$  pour  $f$  continue sur  $]a, b]$ . avec  $a \in \mathbb{R}$  ou  $a = -\infty$ .
- ❸  $\int_a^b f(t) dt$  pour  $f$  continue sur  $]a, b[$ . avec ( $a \in \mathbb{R}$  ou  $a = -\infty$ ) et ( $b = +\infty$  ou  $b \in \mathbb{R}$ ).
- ❹  $\int_a^b f(t) dt$  pour  $f$  continue sauf en un nombre fini de points sur  $]a, b[$ .  
avec ( $a \in \mathbb{R}$  ou  $a = -\infty$ ) et ( $b = +\infty$  ou  $b \in \mathbb{R}$ ).

**Remarque :** Dans ce cours les propositions sont présentées pour  $\int_a^b f(t) dt$  lorsque  $a < b$ ,

mais  $\int_b^a f(t) dt$  et  $\int_a^b f(t) dt$  sont de même nature (*CV ou DV*) et quand elles représentent un réel on a :

$$\int_b^a f(t) dt = - \int_a^b f(t) dt$$

## 7.2 Définitions. vocabulaire. notation.

Toutes les intégrales définies dans ce cours sont appelées *intégrales généralisées* (ou *impropres*).

### 7.2.1 Intégrales sur un intervalle semi-ouvert.

Les fonctions continues sur  $[a, b[$

Ici  $a$  est un réel et  $b$  désigne un réel  $> a$  ou  $+\infty$ .

On dit que l'intégrale est *impropre* en  $b$ .

**Définition.**

Soit  $f$  une fonction continue sur l'intervalle  $[a, b[$ ,

- Si  $\lim_{x \rightarrow b^-} \int_a^x f(t) dt = \ell \in \mathbb{R}$  alors on dit que  $\underbrace{\int_a^b f(t) dt}_{\text{on dit aussi que l'intégrale converge}}$  existe et vaut  $\lim_{x \rightarrow b^-} \int_a^x f(t) dt$ .
- Quand  $\int_a^x f(t) dt$  n'a pas de limite réelle quand  $x$  tend vers  $b^-$ , on dit que  $\int_a^b f(t) dt$  est divergente.

**Caractérisation.** (Avec une primitive de  $f$  sur  $[a, b[$ )

Soit  $f$  une fonction continue sur l'intervalle  $[a, b[$ , on note  $F$  une primitive de  $f$  sur  $[a, b[$ ,

- Si  $\lim_{x \rightarrow b^-} F(x) = \ell \in \mathbb{R}$  alors  $\underbrace{\int_a^b f(t) dt}_{\text{on dit aussi que l'intégrale converge}}$  existe et vaut  $\lim_{x \rightarrow b^-} F(x) - F(a)$ .

- Quand  $F(x)$  n'a pas de limite réelle quand  $x$  tend vers  $b^-$ , on dit que  $\int_a^b f(t) dt$  est divergente.

**En effet :** pour  $x \in [a, b[$ ,  $f$  est continue sur le segment  $[a, x]$  donc  $\int_a^x f(t) dt = F(x) - F(a)$

**Remarques :**

- Cette caractérisation ne dépend pas du choix de la primitive choisie.
- Ici on détermine la nature en trouvant une primitive de l'intégrande.

Dans la suite nous verrons des raisonnements permettant d'étudier la nature sans faire le calcul d'une primitive.

**Exemples :** voir la feuille\_Cours\_8.

**Proposition.** (On peut remplacer la borne non impropre)

Soit  $f$  une fonction continue sur l'intervalle  $[a, b[$ .

Pour  $c$  un élément de  $[a, b[$ , les intégrales  $\int_a^b f(t) dt$  et  $\int_c^b f(t) dt$  ont la même nature.

**En effet :**  $\int_a^b f(t) dt$  converge si, et seulement si,  $\lim_{x \rightarrow b^-} F(x) = \ell \in \mathbb{R}$  si, et seulement si,  $\int_c^b f(t) dt$ .

**Les fonctions continues sur  $]a, b]$**

Ici  $a$  est un réel ou  $-\infty$  et  $b$  un réel vérifiant  $a < b$ .

On dit que l'intégrale est impropre en  $a$ .

**Définition.**

Soit  $f$  une fonction continue sur l'intervalle  $]a, b]$ ,

- Si  $\lim_{x \rightarrow a^+} \int_x^b f(t) dt = \ell \in \mathbb{R}$  alors on dit que  $\underbrace{\int_a^b f(t) dt}_{\text{on dit aussi que l'intégrale converge}}$  existe et vaut  $\lim_{x \rightarrow a^+} \int_x^b f(t) dt$ .

- Quand  $\int_x^b f(t) dt$  n'a pas de limite réelle quand  $x$  tend vers  $a^+$ , on dit que  $\int_a^b f(t) dt$  est divergente.

On peut adapter la caractérisation et la proposition du paragraphe précédent

### 7.2.2 Intégrales sur un intervalle ouvert.

Ici  $a$  est un réel ou  $-\infty$  et  $b$  un réel  $> a$  ou  $+\infty$ .

**Définition.**

Soit  $f$  une fonction continue sur l'intervalle  $]a, b[$  et  $c$  un élément quelconque de  $]a, b[$ ,

$\int_a^b f(t) dt$  converge signifie que :  $\int_a^c f(t) dt$  et  $\int_c^b f(t) dt$  convergent.  
et alors elle vaut  $\int_a^c f(t) dt + \int_c^b f(t) dt$ .

**Caractérisation.** (Avec une primitive de  $f$  sur  $]a, b[$ ).

Soit  $f$  une fonction continue sur l'intervalle  $]a, b[$ , on note  $F$  une primitive de  $f$  sur  $]a, b[$ ,

• Si  $\begin{cases} \lim_{x \rightarrow a^+} F(x) = \ell_1 \in \mathbb{R} \\ \lim_{x \rightarrow b^-} F(x) = \ell_2 \in \mathbb{R} \end{cases}$  alors  $\underbrace{\int_a^b f(t) dt}_{\text{on dit aussi que l'intégrale converge}}$  existe et vaut  $\lim_{x \rightarrow b^-} F(x) - \lim_{x \rightarrow a^+} F(x)$ .

• Quand  $F$  n'a pas de limite réelle en  $a^+$  ou en  $b^-$ , on dit que  $\int_a^b f(t) dt$  est divergente.

**En effet :**  $\int_a^c f(t) dt$  et  $\int_c^b f(t) dt$  convergent si, et seulement si,  $\lim_{x \rightarrow a^+} F(x) = \ell_1 \in \mathbb{R}$  et  $\lim_{x \rightarrow b^-} F(x) = \ell_2 \in \mathbb{R}$

**Attention :** Dans un exercice : ne pas commencer par  $\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$ ,  
c'est à la fin lorsqu'on a montré la convergence des deux intégrales.

Voir la rédaction dans la *feuille\_cours\_8*

**Propositions :**

- Cette définition est cohérente avec la définition des intégrales d'une fonction continue sur un segment.

Soit  $f$  une fonction continue sur le segment  $[a, b]$ , on note  $F$  une primitive de  $f$  sur  $[a, b]$ ,

L'intégrale  $\int_a^b f(t) dt$  existe et on a aussi :  $\int_a^b f(t) dt = \lim_{x \rightarrow b^-} F(x) - \lim_{x \rightarrow a^+} F(x)$ .

**En effet :** Ici  $F$  est continue sur le segment  $[a, b]$ . (*elle est même dérivable*)

- *Cas particulier de la fonction nulle.*

Si  $f$  est la fonction nulle sur  $]a, b[$  alors  $\int_a^b f(t) dt$  est convergente et  $\int_a^b f(t) dt = 0$

**En effet :**  $f$  est la fonction nulle sur  $]a, b[$  alors il existe  $k \in \mathbb{R}$  tel que  $\forall x \in ]a, b[, F(x) = k$

*On utilise cette proposition dans la plupart des exemples avec les fonctions indicatrices.*

- (*Peu importe les valeurs de  $f$  aux bornes de  $[a, b]$* )

Soient  $f$  et  $g$  deux fonctions continues sur  $]a, b[$

Si  $\forall t \in ]a, b[, f(t) = g(t)$

alors  $\int_a^b f(t) dt$  et  $\int_a^b g(t) dt$  sont de même nature et égales en cas de convergence.

**En effet :** Les primitives de  $f$  et  $g$  sur l'ouvert  $]a, b[$  sont les mêmes.

### 7.2.3 Cas particulier d'une fonction prolongeable par continuité.

**Proposition.**

Soient  $a$  et  $b$  deux réels,  $f$  une fonction continue sur l'intervalle  $]a, b[$ ,

Si  $f(x)$  admet une limite réelle quand  $x$  tend vers  $a^+$  et quand  $x$  tend vers  $b^-$ ,

alors l'intégrale  $\int_a^b f(t) dt$  est convergente.

**En effet :** En notant  $g$  le prolongement par continuité de  $f$  à  $[a, b]$  on a  $\int_a^b f(t) dt$   $\int_a^b g(t) dt$  sont de même nature et égales en cas de convergence.

**Exemples :**

$$\int_0^1 \frac{\sin(t)}{t} dt \quad \int_0^1 t \ln(t) dt \quad \int_1^2 \frac{\ln(t)}{t-1} dt$$

**Remarque :**

Certains disent dans cette situation que l'intégrale est "faussetment impropre" en  $a$  ou/et en  $b$ .  
(terme non utilisé dans le programme)

### 7.2.4 Généralisation.

**Définition.**

Quand  $a$  désigne un réel ou  $-\infty$ ,  $b$  un réel ou  $+\infty$ ,  $n$  un entier naturel non nul et  $(a_0, \dots, a_n)$  tels que :

$$a = a_0 < a_1 < \dots < a_{n-1} < a_n = b$$

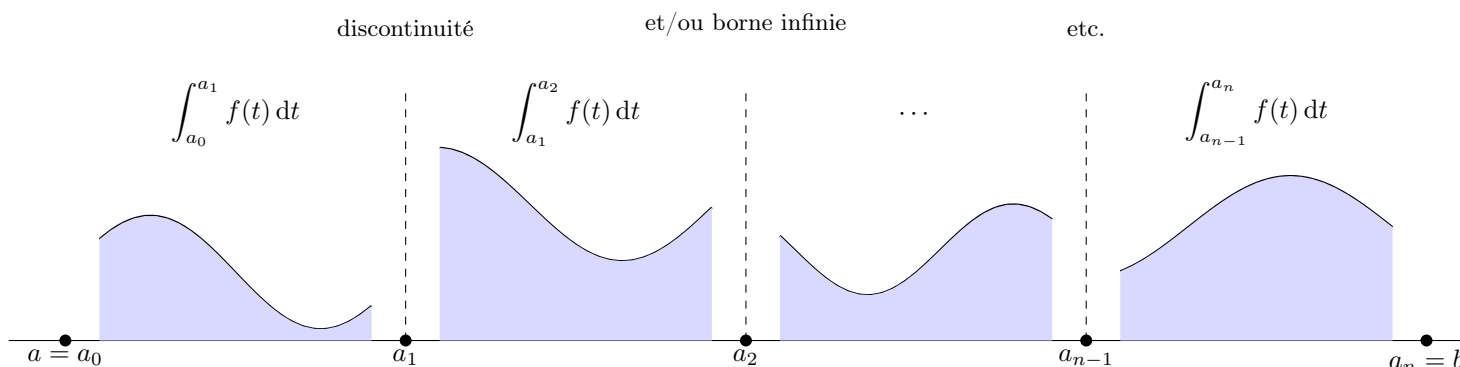
Pour  $f$  une fonction continue sur  $]a, b[$  sauf (éventuellement) aux réels  $a_1, \dots, a_{n-1}$

Dire que l'intégrale  $\int_a^b f(t) dt$  est convergente signifie que :

pour chaque  $k \in \llbracket 1, n \rrbracket$ , l'intégrale  $\int_{a_{k-1}}^{a_k} f(t) dt$  est convergente.

et en cas de convergence de toutes ces intégrales, on définit :  $\int_a^b f(t) dt = \sum_{k=1}^n \int_{a_{k-1}}^{a_k} f(t) dt$ .

*Illustration graphique.*



On découpe donc l'intégrale impropre sur chaque intervalle  $]a_{k-1}, a_k[$ , puis on dit que  $\int_a^b f(t) dt$  converge si, et seulement si, toutes les intégrales obtenues convergent.

**Notation.** (Extrait du programme)

Les notations  $\int_I f$ ,  $\int_I f(t) dt$ ,  $\int_a^b f$ ,  $\int_a^b f(t) dt$  pourront, selon le contexte, désigner l'intégrale généralisée ou sa valeur.

**Méthode pratique générale.** voir exemples dans la feuille\_Cours\_8

❶ Déterminer les intervalles  $]a_{k-1}, a_k[$   $_{1 \leq k \leq n}$ .

❷ Etudier pour chaque  $k$  entre 0 et  $n - 1$ , la nature de l'intégrale  $\int_{a_{k-1}}^{a_k} f(t) dt$ .

③ En cas de convergence de toutes ces intégrales impropres, on conclut :

$$\int_a^b f(t) dt \text{ est convergente et } \int_a^b f(t) dt = \sum_{k=1}^n \int_{a_{k-1}}^{a_k} f(t) dt$$

**Remarques :**

- Si au moins une des intégrales  $\int_{a_{k-1}}^{a_k} f(t) dt$  diverge alors  $\int_a^b f(t) dt$  diverge.
- On ne peut écrire  $\int_a^b f(t) dt = \sum_{k=1}^n \int_{a_{k-1}}^{a_k} f(t) dt$  qu'à la fin du raisonnement en cas de convergence.

**Proposition.** (*Intégrales sur un plus petit intervalle*)

Soit  $f$  une fonction continue sur  $]a, b[$  sauf (*éventuellement*) en nombre fini de réels de  $]a, b[$ ,  
 Si l'intégrale  $\int_a^b f(t) dt$  est convergente alors  
 pour tous réels  $c$  et  $d$  vérifiant  $a \leq c < d \leq b$ , l'intégrale  $\int_c^d f(t) dt$  est convergente.

**Proposition.**

Soient  $\alpha$  un réel et  $f$  une fonction continue sur  $]a, b[$  sauf (*éventuellement*) en nombre fini de réels de  $]a, b[$ ,  
 Si  $\alpha \neq 0$  alors les intégrales  $\int_a^b f(t) dt$  et  $\int_a^b \alpha f(t) dt$  sont de même nature.

## 7.3 Théorèmes de convergence.

### 7.3.1 Théorème de convergence par comparaison des intégrandes positifs

**Théorème.**

Soient  $f$  et  $g$  deux fonctions continues sur  $[a, b[$ ,  
 Si  $\forall t \in [a, b[, 0 \leq f(t) \leq g(t)$   
 et si  $\int_a^b g(t) dt$  est convergente alors  $\int_a^b f(t) dt$  est convergente.

**Démonstration.**

$$F : x \mapsto \int_a^x f(t) dt \quad \text{et} \quad G : x \mapsto \int_a^x g(t) dt \text{ sont croissantes (car les intégrandes sont positifs)}$$

L'encadrement et la "croissance de l'intégrale" donne  $\forall x \in [a, b[, 0 \leq F(x) \leq G(x)$

$$\text{donc } \forall x \in [a, b[, 0 \leq F(x) \leq \underbrace{\int_a^b g(t) dt}_{\in \mathbb{R}}$$

Il suffit alors d'appliquer le théorème de limite monotone à  $F$  pour conclure.

**Remarque :** En prenant la contraposée, on obtient :

**Corollaire.**

Si  $\forall t \in [a, b[, 0 \leq f(t) \leq g(t)$  et si  $\int_a^b f(t) dt$  est divergente, alors  $\int_a^b g(t) dt$  est divergente.

**Corollaires.**

- ❶ Soient  $f$  et  $g$  deux fonctions continues sur  $]a, b]$ ,  
 Si  $\forall t \in ]a, b]$ ,  $0 \leq f(t) \leq g(t)$  et si  $\int_a^b g(t) dt$  est convergente alors  $\int_a^b f(t) dt$  est convergente
- ❷ Soient  $f$  et  $g$  deux fonctions continues sur  $]a, b[$  (sauf en un nombre fini de réels),  
 Si  $\forall t \in ]a, b[$ ,  $0 \leq f(t) \leq g(t)$  (sauf en un nombre fini de réels).  
 et si  $\int_a^b g(t) dt$  est convergente alors  $\int_a^b f(t) dt$  est convergente.

**Remarques.**

- Comme pour les séries : On ne passe pas à l'intégrale!!!
- On compare les intégrandes positifs pour en déduire la nature d'une intégrale.
- Ce théorème permet uniquement de trouver la nature de  $\int_a^b f(t) dt$  ou de  $\int_a^b g(t) dt$
- Ce théorème ne permet pas de calculer la valeur de  $\int_a^b f(t) dt$  quand il y a convergence.
- Attention ici la conclusion du théorème n'est pas : "elles sont de même nature".

**7.3.2 Théorème de convergence par équivalence des intégrandes positifs.****Théorème.**

- Soient  $f$  et  $g$  deux fonctions continues sur  $[a, b[$ ,  
 Si  $\forall t \in [a, b[$ ,  $g(t) \geq 0$  et si  $f(t) \underset{t \rightarrow b^-}{\sim} g(t)$ , alors
- ❶ Si  $\int_a^b g(t) dt$  converge alors  $\int_a^b f(t) dt$  converge.
- ❷ Si  $\int_a^b g(t) dt$  diverge alors  $\int_a^b f(t) dt$  diverge.

**Démonstration.** (Rapidement)

On se place dans le cas où  $g$  ne s'annule pas sur  $[a, b[$ ,

- $f(t) \underset{t \rightarrow b^-}{\sim} g(t)$  donc il existe  $c \in [a, b[$ ,  $\forall t \in [c, b[$ ,  $\frac{1}{2} \leq \frac{f(t)}{g(t)} \leq \frac{3}{2}$
- et  $g(t) \geq 0$  donc  $\forall t \in [c, b[$ ,  $0 \leq \frac{1}{2}g(t) \leq f(t) \leq \frac{3}{2}g(t)$

Il suffit alors d'appliquer le théorème de convergence par comparaison pour conclure.

**Corollaire.**

- Soient  $f$  et  $g$  deux fonctions continues sur  $]a, b]$ ,  
 Si  $\forall t \in ]a, b]$ ,  $g(t) \geq 0$  et si  $f(t) \underset{t \rightarrow a^+}{\sim} g(t)$ , alors
- ❶ Si  $\int_a^b g(t) dt$  converge alors  $\int_a^b f(t) dt$  converge.
- ❷ Si  $\int_a^b g(t) dt$  diverge alors  $\int_a^b f(t) dt$  diverge.

**Remarques.**

- Ce théorème permet uniquement de trouver la nature de  $\int_a^b f(t) dt$ .

- Ce théorème ne permet pas de calculer la valeur de  $\int_a^b f(t) dt$ , ni même son signe.
- On peut résumer ❶ et ❷ en :

$$\int_a^b g(t) dt \text{ converge si, et seulement si, } \int_a^b f(t) dt \text{ converge.}$$

- On peut résumer ❶ et ❷ en :

$$\int_a^b g(t) dt \quad \text{et} \quad \int_a^b f(t) dt \quad \text{sont de même nature.}$$

## 7.4 Propriétés

$a$  est un réel ou  $-\infty$  et  $b$  un réel  $> a$  ou  $+\infty$ ,

et  $f, g, f_1, \dots, f_n$  sont des fonctions continues sur  $]a, b[$  [sauf éventuellement en un nombre fini de réels.

### 7.4.1 Relation de Chasles.

**Théorème :**

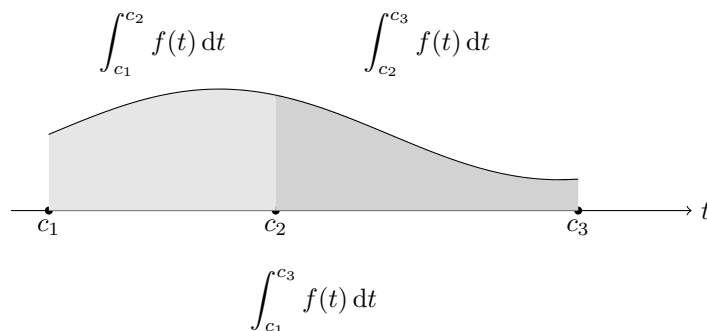
Si  $\int_a^b f(t) dt$  est convergente,

alors, pour tout  $(c_1, c_2, c_3) \in [a, b]^3$ ,  $\int_{c_1}^{c_2} f(t) dt$ ,  $\int_{c_2}^{c_3} f(t) dt$  et  $\int_{c_1}^{c_3} f(t) dt$  convergent

et  $\int_{c_1}^{c_3} f(t) dt = \int_{c_1}^{c_2} f(t) dt + \int_{c_2}^{c_3} f(t) dt$

**Démonstration :** (*admis*) (On ajoute  $c_1, c_2, c_3$  à la subdivision ...)

*Illustration graphique de la relation de Chasles.*



L'aire totale entre  $c_1$  et  $c_3$  est la somme des aires entre  $c_1$  et  $c_2$ , puis entre  $c_2$  et  $c_3$ .

**Généralisation :** pour  $n \in \mathbb{N}^*$

Si  $\int_a^b f(t) dt$  est convergente,

alors, quel que soit  $(c_0, \dots, c_n) \in [a, b]^{n+1}$ , on a :  $\forall k \in [0; n-1]$ ,  $\int_{c_k}^{c_{k+1}} f(t) dt$  convergent

et  $\int_{c_0}^{c_n} f(t) dt = \sum_{k=0}^{n-1} \int_{c_k}^{c_{k+1}} f(t) dt$

*Remarque :* Les  $c_k$  ne sont pas nécessairement dans l'ordre croissant est peuvent être égaux à  $a$  ou  $b$ .

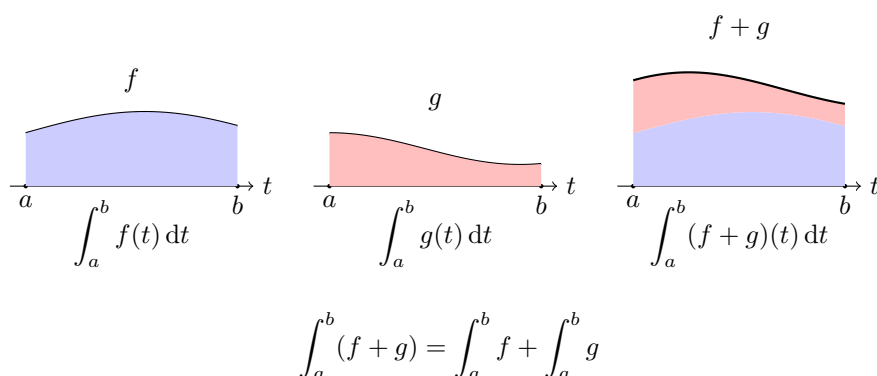
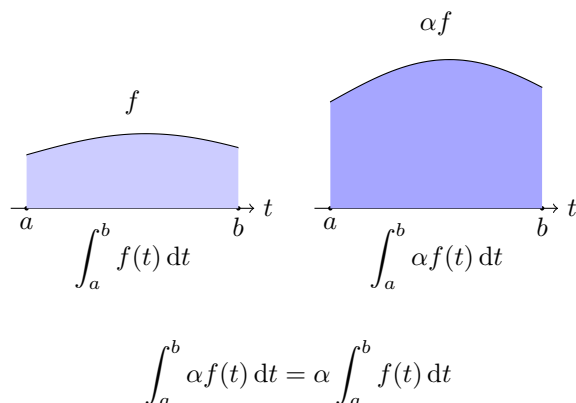
## 7.4.2 Linéarité.

Théorème :

Si  $\int_a^b f(t) dt$  et  $\int_a^b g(t) dt$  sont convergentes,  
 alors pour tout  $(\alpha, \beta) \in \mathbb{R}^2$ ,  $\int_a^b (\alpha f(t) + \beta g(t)) dt$  converge et

$$\int_a^b (\alpha f(t) + \beta g(t)) dt = \alpha \int_a^b f(t) dt + \beta \int_a^b g(t) dt$$
**Démonstration :** (admis) (Travailler sur des subdivisions de  $]a, b[$  adaptées à  $f$  et  $g$ )

Illustration graphique de la linéarité de l'intégrale.

 $f + g$  $\alpha f$ .**Généralisation :**

Si pour tout  $k \in \llbracket 1; n \rrbracket$ , l'intégrale  $\int_a^b f_k(t) dt$  est convergente,  
 alors pour tout  $(\alpha_k)_{1 \leq k \leq n} \in \mathbb{R}^n$ , l'intégrale  $\int_a^b \left( \sum_{k=1}^n \alpha_k f_k(t) \right) dt$  converge et

$$\int_a^b \left( \sum_{k=1}^n \alpha_k f_k(t) \right) dt = \sum_{k=1}^n \alpha_k \left( \int_a^b f_k(t) dt \right)$$
**Démonstration :** (admis) (Raisonnement par récurrence sur  $n$ )

### 7.4.3 Positivité

On rappelle qu'on est sous l'hypothèse  $a < b$ .

**Théorème :**

$$\text{Lorsque } \int_a^b f(t) dt \text{ est convergente,}$$

$$\text{Si } \forall t \in ]a, b[, f(t) \geq 0 \quad \text{alors} \quad \int_a^b f(t) dt \geq 0$$

**Démonstration :** (*admis*)

### 7.4.4 Croissance de l'intégrale.

**Théorème :**

$$\text{Lorsque } \int_a^b f(t) dt \quad \text{et} \quad \int_a^b g(t) dt \quad \text{sont convergentes,}$$

$$\text{Si } \forall t \in ]a, b[, f(t) \leq g(t) \quad \text{alors} \quad \int_a^b f(t) dt \leq \int_a^b g(t) dt$$

**En effet :** Appliquer le théorème de 7.4.3 à la fonction  $g - f$ .

### 7.4.5 Stricte positivité.

$a$  désigne un réel ou  $-\infty$ ,  $b$  un réel ou  $+\infty$  et  $a < b$ .

**Théorème :**

$$\text{Lorsque } \int_a^b f(t) dt \text{ est convergente.}$$

$$\text{Si } \begin{cases} f \text{ est continue sur } ]a, b[, \\ \forall t \in ]a, b[, f(t) \geq 0, \\ \int_a^b f(t) dt = 0 \end{cases} \quad \text{alors} \quad \forall t \in ]a, b[, f(t) = 0$$

**Démonstration.**

$f$  est continue sur  $]a, b[$ , il existe  $F$  un primitive de  $f$  sur  $]a, b[$ .

d'une part  $f \geq 0$  donc  $F$  est croissante,

d'autre part  $\int_a^b f(t) dt$  est convergente donc  $\lim_{b^-} F$  et  $\lim_{a^+} F$  existent, sont réelles et  $\int_a^b f(t) dt = \lim_{b^-} F - \lim_{a^+} F$

et comme  $\int_a^b f(t) dt = 0$  il vient  $\lim_{b^-} F = \lim_{a^+} F$

$F$  est croissante et  $\lim_{b^-} F = \lim_{a^+} F$  entraîne que  $F$  est constante sur  $]a, b[$  et ainsi  $\forall t \in ]a, b[, f(t) = 0$ .

**Remarques :**

- On peut remplacer dans ce théorème l'intervalle  $]a, b[$  par tout intervalle  $I : ]a, b], [a, b[$  ou  $[a, b]$

**En effet :**

- Attention à l'argument " $f$  est continue sur  $I$ ".

On peut trouver une fonction positive sur  $]0, 1[$ , telle que  $\int_0^1 f(t) dt = 0$  et qui n'est pas nulle sur  $]0, 1[$ .

(qui n'est pas la fonction nulle)

**Corollaire 1 :**

Lorsque  $\int_a^b f(t) dt$  est convergente,

Si  $\begin{cases} f \text{ est continue sur } ]a, b[, \\ \forall t \in ]a, b[, f(t) \geq 0, \\ f \text{ n'est pas la fonction nulle sur } ]a, b[ \end{cases}$  alors  $\int_a^b f(t) dt > 0$

**En effet :** C'est une forme contraposée du théorème précédent.

**Exemple d'application :** L'intégrale  $\int_0^1 t \cos^2\left(\frac{1}{t}\right) dt$  converge et  $\int_0^1 t \cos^2\left(\frac{1}{t}\right) dt > 0$

**Corollaire 2 :**

Pour  $f$  une fonction continue sur  $]a, b[$  sauf peut-être en un nombre fini de points.

Lorsque  $\int_a^b f(t) dt$  est convergente,

Si  $\forall t \in ]a, b[, f(t) > 0$  alors  $\int_a^b f(t) dt > 0$

**En effet :**

*En raisonnant sur une subdivision bien choisie, l'hypothèse ici entraîne les hypothèses du corollaire précédent.*

**Exemples d'utilisation :**

- Pour tout  $n \in \mathbb{N}$ ,  $\int_0^{\frac{\pi}{2}} \cos^n(t) dt > 0$
- $\int_0^1 (\ln(t))^3 dt < 0$

**7.4.6 Parité.**

*Ici  $a$  est un réel positif ou  $+\infty$  et  $f$  est continue sur  $] - a, a[$  sauf en un nombre fini de réels.*

**Proposition :**

Lorsque  $f$  est paire,

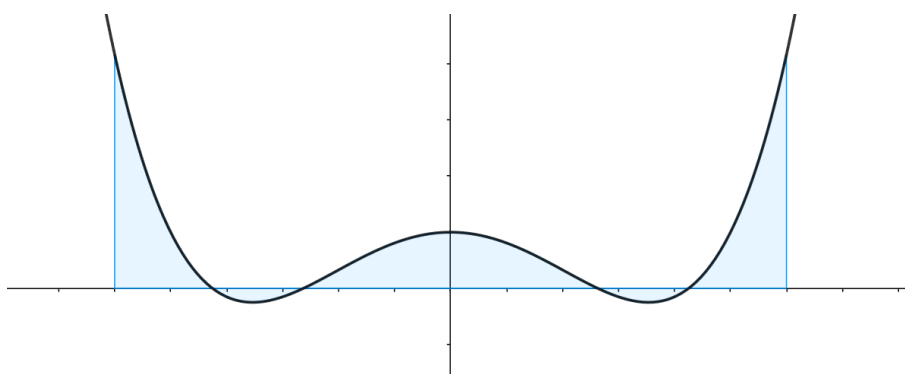
$\int_{-a}^a f(t) dt$  est convergente **si, et seulement si,**  $\int_0^a f(t) dt$  est convergente

et alors  $\int_{-a}^a f(t) dt = 2 \int_0^a f(t) dt$

**En effet :** Sachant que  $f$  est paire, si  $F$  est une primitive de  $f$  sur  $] - a, a[$ , alors  $x \mapsto -F(-x)$  est aussi une primitive de  $f$ , donc il existe  $k \in \mathbb{R}$  tel que  $\forall x \in ] - a, a[, F(x) + F(-x) = k. (= 2F(0))$

Ainsi,  $\lim_{x \rightarrow a^-} F(x) \in \mathbb{R} \iff \lim_{x \rightarrow (-a)^+} F(x) \in \mathbb{R}$ . ce qui donne l'équivalence de convergence.

Enfin, en notant  $\ell = \lim_{x \rightarrow a^-} F(x)$ , il vient :  $\int_{-a}^a f(t) dt = \ell - (k - \ell) = 2\ell - k = 2(\ell - F(0)) = 2 \int_0^a f(t) dt$



**Proposition :**

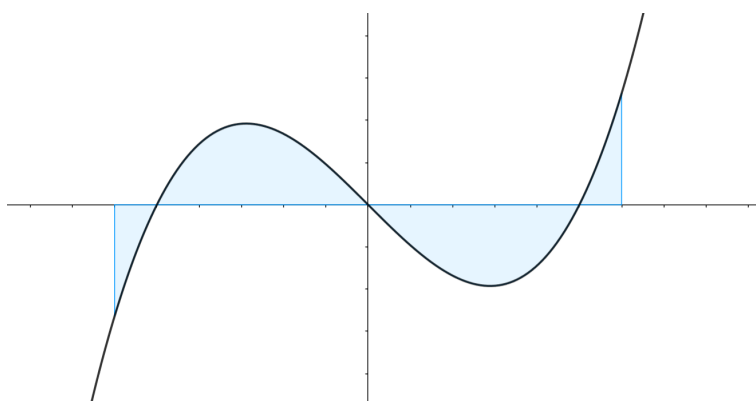
Lorsque  $f$  est impaire,  
 $\int_{-a}^a f(t) dt$  est convergente **si, et seulement si**,  $\int_0^a f(t) dt$  est convergente  
 et alors  $\int_{-a}^a f(t) dt = 0$

**En effet :**

Sachant que  $f$  est impaire, si  $F$  est une primitive de  $f$  sur  $] -a, a[$ , alors  $x \mapsto F(-x)$  est aussi une primitive de  $f$ , donc il existe  $k \in \mathbb{R}$  tel que  $\forall x \in ] -a, a[$ ,  $F(x) - F(-x) = k$ . ( $= 0$ )

Ainsi,  $\lim_{x \rightarrow a^-} F(x) \in \mathbb{R} \iff \lim_{x \rightarrow (-a)^+} F(x) \in \mathbb{R}$ . ce qui donne l'équivalence de convergence.

Enfin, en notant  $\ell = \lim_{x \rightarrow a^-} F(x)$ , il vient :  $\int_{-a}^a f(t) dt = \ell - (k + \ell) = k = 0$

**7.5 Absolue convergence.**

$a$  désigne un réel ou  $-\infty$  et  $b$  un réel  $> a$  ou  $+\infty$ ,

**Définition.**

Soit  $f$  une fonction continue sur  $]a, b[$  sauf éventuellement en un nombre fini de réels.

Dire que  $\int_a^b f(t) dt$  est absolument convergente signifie que  $\int_a^b |f(t)| dt$  est convergente

**Théorème :**

Soit  $f$  une fonction continue sur  $]a, b[$  sauf éventuellement en un nombre fini de réels.

Si  $\int_a^b f(t) dt$  est absolument convergente

alors  $\int_a^b f(t) dt$  est convergente et  $\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$

**En effet :**

- Pour la convergence on applique le théorème de convergence par comparaison avec l'encadrement

$$\forall t \in ]a, b[, 0 \leq f(t) + |f(t)| \leq 2|f(t)|$$

On montre la convergence de  $\int (f + |f|)$  puis avec la linéarité celle de  $\int f$

- Ensuite on applique la croissance de l'intégrale à l'encadrement :  $\forall t \in ]a, b[, -|f(t)| \leq f(t) \leq |f(t)|$ .

**Proposition**

Soient  $f$  et  $g$  des fonctions continues sur  $]a, b[$  sauf éventuellement en un nombre fini de réels, et  $\alpha, \beta$  deux réels,

Si  $\int_a^b f(t) dt$  et  $\int_a^b g(t) dt$  sont absolument convergentes

alors  $\int_a^b (\alpha f(t) + \beta g(t)) dt$  est absolument convergente.

**En effet :** Cela découle directement de  $\forall t \in ]a, b[, 0 \leq |\alpha f(t) + \beta g(t)| \leq |\alpha| |f(t)| + |\beta| |g(t)|$  avec le théorème de convergence par comparaison.

*Autrement dit :* L'ensemble des fonctions continues sur  $]a, b[$ , sauf éventuellement en un nombre fini de réels, dont l'intégrale est absolument convergente sur  $]a, b[$  est stable par combinaison linéaire. C'est un sous espace vectoriel de l'espace des fonctions

**7.6 Calculs.**

Dans ce paragraphe,  $a$  et  $b$  vérifient  $-\infty \leq a < b \leq +\infty$ .

**7.6.1 Intégrations par parties.**

**Théorème.** (*Intégration par parties*)

Soient  $u$  et  $v$  deux fonctions de classe  $C^1$  sur  $]a, b[$ ,

Si  $\lim_{x \rightarrow a^+} u(x)v(x) \in \mathbb{R}$ ,  $\lim_{x \rightarrow b^-} u(x)v(x) \in \mathbb{R}$  et  $\int_a^b u'(t)v(t) dt$  converge (**Trois conditions**)

alors  $\int_a^b u(t)v'(t) dt$  est convergente et

$$\int_a^b u(t)v'(t) dt = \lim_{x \rightarrow b^-} u(x)v(x) - \lim_{x \rightarrow a^+} u(x)v(x) - \int_a^b u'(t)v(t) dt$$

**Démonstration.** (*admis*)

**En pratique** on n'utilise pas souvent ce théorème, il est plus simple de chercher une primitive.

**Corollaire.**

Soient  $u$  et  $v$  deux fonctions de classe  $C^1$  sur  $]a, b[$ ,

Si  $\lim_{x \rightarrow a^+} u(x)v(x) \in \mathbb{R}$ , et  $\lim_{x \rightarrow b^-} u(x)v(x) \in \mathbb{R}$ ,

alors  $\int_a^b u(t)v'(t) dt$  et  $\int_a^b u'(t)v(t) dt$  sont de même nature.

**7.6.2 Changement de variable.**

**Théorème.** (*Changement de variable  $x = \varphi(t)$* )

Soient  $\varphi$  et  $f$  deux fonctions à valeurs dans  $\mathbb{R}$ .

Si  $\varphi$  est de classe  $C^1$  et strictement monotone sur  $]a, b[$ , on note  $\alpha = \lim_{t \rightarrow a^+} \varphi(t)$  et  $\beta = \lim_{t \rightarrow b^-} \varphi(t)$

et si  $f$  est continue sur  $]\alpha, \beta[$ ,

alors  $\int_\alpha^\beta f(x) dx$  converge si, et seulement si,  $\int_a^b f(\varphi(t)) \varphi'(t) dt$  converge

et en cas de convergence :  $\int_\alpha^\beta f(x) dx = \int_a^b f(\varphi(t)) \varphi'(t) dt$

**Démonstration.** (*admis*)

**Corollaire.** (*f* strictement croissante)

Soient  $\varphi$  et  $f$  deux fonctions à valeurs dans  $\mathbb{R}$  et  $\alpha, \beta$  des réels ou  $-\infty$  ou  $+\infty$ .

Si  $\varphi : ]a, b[ \rightarrow ]\alpha, \beta[$  est de classe  $C^1$ , strictement croissante et bijective, et si  $f$  est continue sur  $] \alpha, \beta [$ ,

alors  $\int_{\alpha}^{\beta} f(x) dx$  converge si, et seulement si,  $\int_a^b f(\varphi(t)) \varphi'(t) dt$  converge

et en cas de convergence :

$$\int_{\alpha}^{\beta} f(x) dx = \int_a^b f(\varphi(t)) \varphi'(t) dt$$

**En effet :**  $\alpha = \lim_{t \rightarrow a^+} \varphi(t)$  et  $\beta = \lim_{t \rightarrow b^-} \varphi(t)$

**Corollaire.** (*f* strictement décroissante)

Soient  $\varphi$  et  $f$  deux fonctions à valeurs dans  $\mathbb{R}$  et  $\alpha, \beta$  des réels ou  $-\infty$  ou  $+\infty$ .

Si  $\varphi : ]a, b[ \rightarrow ]\beta, \alpha[$  est de classe  $C^1$ , strictement décroissante et bijective, et si  $f$  est continue sur  $] \beta, \alpha [$ ,

alors  $\int_{\alpha}^{\beta} f(x) dx$  converge si, et seulement si,  $\int_a^b f(\varphi(t)) \varphi'(t) dt$  converge

et en cas de convergence :

$$\int_{\alpha}^{\beta} f(x) dx = \int_a^b f(\varphi(t)) \varphi'(t) dt$$

**En effet :**  $\alpha = \lim_{t \rightarrow a^+} \varphi(t)$  et  $\beta = \lim_{t \rightarrow b^-} \varphi(t)$

**En pratique :**

- ❶ On identifie la fonction  $\varphi$  et on vérifie qu'elle possède les propriétés requises.
- ❷ on cherche  $a, b, \alpha$  et  $\beta$  tels que  $t$  va de  $a$  à  $b$  et  $x$  de  $\alpha$  à  $\beta$ .
- ❸ on remplace  $dx$  par  $\varphi'(t)dt$  ou  $\varphi'(t)dt$  par  $dx$ .
- ❹ on remplace  $x$  par  $\varphi(t)$  ou  $\varphi(t)$  par  $x$ .

$$\underbrace{\int_{\alpha}^{\beta}}_{\text{❷}} \underbrace{f(x)}_{\text{❹}} \underbrace{dx}_{\text{❸}} = \underbrace{\int_a^b}_{\text{❷}} \underbrace{f(\varphi(t))}_{\text{❹}} \underbrace{\varphi'(t) dt}_{\text{❸}}$$

**Corollaire.** *Changement de variable affine*  $x = mt + p$ .

Soient  $f$  une fonction à valeurs dans  $\mathbb{R}$  et  $m, p$  deux réels tels que  $m \neq 0$

On note :  $\alpha = \lim_{t \rightarrow a^+} (mt + p)$  et  $\beta = \lim_{t \rightarrow b^-} (mt + p)$

Si  $f$  est continue sur  $] \alpha, \beta [$ ,

alors  $\int_a^b f(mt + p) m dt$  converge si, et seulement si,  $\int_{\alpha}^{\beta} f(x) dx$  converge

et en cas de convergence :

$$\int_a^b f(mt + p) dt = \int_{\alpha}^{\beta} f(x) \frac{1}{m} dx$$

**Remarque :**

Ici les conditions d'utilisation du théorème sont plus simples, on peut alléger la rédaction.

## 7.7 Intégrales impropres usuelles.

### 7.7.1 Loi uniforme

$$\int_{-\infty}^{+\infty} \mathbb{1}_{[a,b]}(t) dt = b - a$$

### 7.7.2 Loi exponentielle

Pour  $\lambda \in \mathbb{R}_+^*$

$$\int_{-\infty}^{+\infty} \mathbb{1}_{\mathbb{R}_+}(t) e^{-\lambda t} dt = \frac{1}{\lambda}$$

### 7.7.3 Loi normale

Centrée réduite. (*admis*)

$$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}$$

Soient  $\mu \in \mathbb{R}$  et  $\sigma \in \mathbb{R}_+^*$ ,

$$\int_{-\infty}^{+\infty} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \sigma\sqrt{2\pi}$$

*A savoir redémontrer.*

### 7.7.4 Compléments : Intégrales de Riemann.

*Hors programme, voir feuille\_Exo\_15*

	$\alpha \leq 0$	$0 < \alpha < 1$	$\alpha = 1$	$1 < \alpha$
$\int_0^1 \frac{1}{t^\alpha} dt$	CV	CV	DV	DV
$\int_0^1 \frac{1}{t^\alpha} dt =$	$\frac{1}{1-\alpha}$	$\frac{1}{1-\alpha}$		
$\int_1^{+\infty} \frac{1}{t^\alpha} dt$	DV	DV	DV	CV
$\int_1^{+\infty} \frac{1}{t^\alpha} dt =$				$\frac{1}{\alpha-1}$

Des exemples de convergentes.

$$\int_0^1 \frac{1}{\sqrt{t}} dt = 2 \quad \int_1^{+\infty} \frac{1}{t^2} dt = 1 \quad \int_1^{+\infty} \frac{1}{t^3} dt = \frac{1}{2} \quad \int_1^{+\infty} \frac{1}{t\sqrt{t}} dt = 2$$

Des exemples de divergentes.

$$\int_0^1 \frac{1}{t^2} dt \text{ diverge} \quad \int_0^1 \frac{1}{t} dt \text{ diverge} \quad \int_1^{+\infty} \frac{1}{t} dt \text{ diverge} \quad \int_1^{+\infty} \frac{1}{\sqrt{t}} dt \text{ diverge}$$

# Diagonalisation

## Plan du chapitre

8.1	Introduction . . . . .	103
8.2	Eléments propres. . . . .	104
8.2.1	Eléments propres d'un endomorphisme. . . . .	104
8.2.2	Eléments propres d'une matrice carrée. . . . .	105
8.2.3	Lien entre matrice et endomorphisme. . . . .	107
8.3	Propriétés des familles de vecteurs propres. . . . .	108
8.3.1	Vecteurs propres associés à des valeurs propres distinctes. . . . .	108
8.3.2	Juxtaposition des bases des sous-espaces propres. . . . .	108
8.3.3	En dimension finie. . . . .	110
8.4	Diagonalisation. . . . .	110
8.4.1	Définition. . . . .	110
8.4.2	Condition suffisante. . . . .	112
8.4.3	Condition nécessaire et suffisante. . . . .	112
8.4.4	En pratique pour les matrices carrées. . . . .	113
8.4.5	Application au calcul de puissance de matrices carrées. . . . .	114
8.4.6	Application aux systèmes différentielles linéaires. . . . .	114

## 8.1 Introduction

Dans ce paragraphe on travaille dans  $E$  un espace vectoriel sur  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ .

Lors d'une mesure d'absorbance, on met en évidence des fréquences caractéristiques propre à une molécule. Ici nous allons mettre en évidence des *valeurs propres* à des endomorphismes. (*Etude spectrale*)

Sachant qu'une matrice carrée peut être vue comme un endomorphisme, toutes les définitions sur les endomorphismes seront encore valables pour les matrices.

On étudie souvent des phénomènes modélisés par des relations  $u_{n+1} = f(u_n)$  ou  $X_{n+1} = MX_n$ .

On s'intéresse ici au "direction fixe" : autrement dit les vecteurs  $u$  vérifiant  $u \neq 0_E$  et  $f(u)$  et  $u$  sont colinéaires.

Ces directions privilégiées sont dirigées par les vecteurs appelés *vecteurs propres* de  $f$ .

### Exemples :

- Avec  $f : (x, y) \mapsto (2x + y, x + 2y)$  avec  $u = (1, 1)$  et  $v = (1, -1)$

$$f(u) = 3u \quad f(v) = v$$

- Avec  $M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  avec  $X_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$  ou  $X_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$$MX_1 = 2X_1 \quad MX_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (= 0X_2)$$

## 8.2 Eléments propres.

### 8.2.1 Eléments propres d'un endomorphisme.

#### Valeurs propres. Spectre.

**Définition** (*Valeur propre d'un endomorphisme*)

Soient  $E$  un espace vectoriel,  $f$  un endomorphisme de  $E$  et  $\lambda \in \mathbb{K}$   
 Dire que  $\lambda$  est une **valeur propre** de  $f$  signifie que :  
 il existe un vecteur  $u$  **non nul** de  $E$  tel que  $f(u) = \lambda u$ .

**En pratique :** (*caractérisations*)

Montrer qu' "un scalaire  $\lambda$  est une valeur propre de  $f$ " revient à savoir s'  
 "il existe un vecteur  $u$  non nul vérifiant  $f(u) = \lambda u$ "  
 c'est équivalent à "pour un certain vecteur  $u$  non nul on a :  $f(u) - \lambda u = 0_E$ "  
 c'est équivalent à "il existe un vecteur  $u$  non nul vérifiant  $(f - \lambda Id_E)(u) = 0_E$ "  
 c'est équivalent à "il existe un vecteur  $u$  non nul appartenant à  $\ker(f - \lambda Id_E)$ "  
 c'est équivalent à " $\ker(f - \lambda Id_E) \neq \{0_E\}$ "  
 c'est équivalent à "l'endomorphisme  $f - \lambda Id_E$  n'est pas injectif".

Si de plus  $E$  est de dimension finie ( $\dim(E) = n$ ) :

c'est équivalent à "l'endomorphisme  $f - \lambda Id_E$  n'est pas bijectif".  
 c'est équivalent à " $\text{rg}(f - \lambda Id_E) < n$ ".

**Remarques autour de la valeur propre 0 :**

- 0 est une valeur propre de  $f$  si, et seulement si,  $f$  n'est pas injective.
- 0 est une valeur propre de  $f$  si, et seulement si,  $\ker(f) \neq \{0_E\}$

**En effet :** Pour  $\lambda = 0$ ,  $(f - \lambda Id_E) = f$  donc  $\ker(f - \lambda Id_E) \neq \{0_E\} \iff \ker(f) \neq \{0_E\}$

**Définition** (*Spectre.*)

Soient  $E$  un espace vectoriel et  $f$  un endomorphisme de  $E$ ,  
 On appelle **spectre** de  $f$  l'ensemble des valeurs propres de  $f$ . (on note :  $Sp(f)$ ).

#### Vecteurs propres. Sous-espaces propres.

**Définition** (*Vecteur propre d'un endomorphisme*)

Soient  $E$  un espace vectoriel,  $f$  un endomorphisme de  $E$  et  $u \in E$   
 Dire que  $u$  est un **vecteur propre** de  $f$  signifie que :  
 ❶  $u$  est non nul et ❷ il existe un  $\lambda \in \mathbb{K}$  tel que  $f(u) = \lambda u$ .

**Définition** (*Vecteur propre et valeur propre associés*)

Soient  $E$  un espace vectoriel,  $f$  un endomorphisme de  $E$ ,  $u \in E$  et  $\lambda \in \mathbb{K}$   
 Dire que  $u$  est un **vecteur propre de  $f$  associé** à la valeur propre  $\lambda$  signifie que :  
 ❶  $u$  est non nul et que ❷  $f(u) = \lambda u$ .

**Définition** (*Sous-espace propre associé à une valeur propre.*)

Soient  $E$  un espace vectoriel,  $f$  un endomorphisme de  $E$  et  $\lambda \in \mathbb{K}$   
 Lorsque  $\lambda$  est une valeur propre de  $f$   
 l'ensemble des vecteurs  $u$  vérifiant  $f(u) = \lambda u$  est appelé **sous-espace propre** de  $f$  associé à  $\lambda$ .  
 On le note :  $E_\lambda(f)$

$$E_\lambda(f) = \{ u \in E \mid f(u) = \lambda u \}$$

**Remarques :**

- $E_\lambda(f) = \ker(f - \lambda Id_E)$  En effet :  $f(u) = \lambda u \iff f(u) - \lambda u = 0_E \iff (f - \lambda Id_E)(u) = 0_E$
- $0_E$  appartient à  $E_\lambda(f)$
- $E_\lambda(f)$  est l'ensemble des vecteurs propres de  $f$  complété par  $0_E$ .
- $E_\lambda(f)$  est un sous-espace vectoriel de  $E$ . **En effet :** c'est le noyau d'une application linéaire.

**Théorème** *En dimension finie*

Soient  $E$  un espace vectoriel de dimension  $n$  et  $f$  un endomorphisme de  $E$ .

Pour tout  $\lambda \in \text{Sp}(f)$ ,    ❶  $\dim(E_\lambda(f)) = n - \text{rg}(f - \lambda Id_E)$

                                  ❷  $\dim(E_\lambda(f)) \geq 1$

**En effet :** *Théorème du rang appliqué à  $f - \lambda Id_E$ ,*

**8.2.2 Eléments propres d'une matrice carrée.**

*Dans ce paragraphe comme dans le cours sur les applications linéaires, on adapte toutes les définitions en confondant la matrice carrée  $M$  et l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbb{K})$  qui à  $X$  associe  $MX$ .*

Ici  $n$  est un entier naturel non nul.

**Valeurs propres. Spectre.**

**Définition** (*Valeur propre d'une matrice carrée*)

Soient  $M$  une matrice de  $\mathcal{M}_n(\mathbb{K})$  et  $\lambda \in \mathbb{K}$

Dire que  $\lambda$  est une **valeur propre** de  $M$  signifie que :

il existe une matrice colonne  $X$  **non nulle** de  $\mathcal{M}_{n,1}(\mathbb{K})$  telle que  $MX = \lambda X$ .

**En pratique :** (*caractérisation*).

Montrer qu' "un scalaire  $\lambda$  est une valeur propre de  $M$ " revient à savoir s'

"il existe une colonne  $X$  non nulle vérifiant  $MX = \lambda X$ ".

c'est équivalent à "le système  $(M - \lambda I_n)X = 0$  admet une solution non nulle"

c'est équivalent à " la matrice  $(M - \lambda I_n)$  n'est pas inversible ".

c'est équivalent à "  $\text{rg}(M - \lambda I_n) < n$  ".

c'est équivalent à "  $\ker(M - \lambda I_n) \neq \{0_{n,1}\}$  "

**Remarques :** (*cas particulier de la valeur propre 0*)

- 0 est une valeur propre de  $M$  si, et seulement si,  $\text{rg}(M) \neq n$ .
- 0 est une valeur propre de  $M$  si, et seulement si,  $M$  n'est pas inversible.

**Définition** (*Spectre.*)

Soit  $M \in \mathcal{M}_n(\mathbb{K})$ ,

On appelle **spectre** de  $M$  l'ensemble des valeurs propres de  $M$ . (on note :  $\text{Sp}(M)$ ).

**Vecteurs propres. Sous-espaces propres.**

**Définition** (*Vecteur propre d'une matrice*)

Soient  $M \in \mathcal{M}_n(\mathbb{K})$  et  $X \in \mathcal{M}_{n,1}(\mathbb{K})$

Dire que  $X$  est une **vecteur propre** de  $M$  signifie que :

- ❶  $X$  est non nulle et
- ❷ il existe un  $\lambda \in \mathbb{K}$  tel que  $MX = \lambda X$ .

**Définition** (*Vecteur propre et valeur propre associée*)

Soient  $M \in \mathcal{M}_n(\mathbb{K})$ ,  $X \in \mathcal{M}_{n,1}(\mathbb{K})$  et  $\lambda \in \mathbb{K}$ .

Dire que  $X$  est un **vecteur propre de  $M$  associé** à la valeur propre  $\lambda$  signifie que :

- ❶  $X$  est non nulle    et    ❷  $MX = \lambda X$ .

**Définition** (*Sous-espace propre associée à une valeur propre.*)

Soient  $M \in \mathcal{M}_n(\mathbb{K})$  et  $\lambda \in \mathbb{K}$ .

Si  $\lambda$  est une valeur propre de  $M$ , alors

l'ensemble des  $X \in \mathcal{M}_{n,1}(\mathbb{K})$  telles que  $MX = \lambda X$  est appelé **sous-espace propre** de  $M$  associé à  $\lambda$ .  
On le note :  $E_\lambda(M)$

$$E_\lambda(M) = \{ X \in \mathcal{M}_{n,1}(\mathbb{K}) \mid MX = \lambda X \}$$

**Remarques :**

- $E_\lambda(M) = \ker(M - \lambda I_n)$
- C'est l'ensemble des vecteurs propres associés à la valeur propre  $\lambda$  complété par la colonne nulle.
- C'est un sous-espace vectoriel de l'ensemble des matrices colonnes :  $\mathcal{M}_{n,1}(\mathbb{K})$   
(*Démonstration : c'est le noyau d'un endomorphisme de  $\mathcal{M}_{n,1}(\mathbb{K})$* )

**Proposition.** (*Dimension d'un sous-espace propre.*)

Soit  $M \in \mathcal{M}_n(\mathbb{K})$ ,

- pour tout  $\lambda \in Sp(M)$ ,
- ❶  $\dim(E_\lambda(M)) = n - \text{rg}(M - \lambda I_n)$
  - ❷  $\dim(E_\lambda(M)) \geq 1$

**Démonstration :** *Théorème du rang appliqué à la matrice  $M - \lambda I_n$ .*

**Cas des matrices triangulaires.**

**Théorème.**

Soit  $M \in \mathcal{M}_n(\mathbb{K})$ ,

si  $M$  est triangulaire (*supérieure ou inférieure*)

alors les valeurs propres de  $M$  sont les coefficients diagonaux de  $M$ .

**En effet.** On applique à la matrice triangulaire  $M - \lambda I_n$  le théorème fondamental :

"Une matrice triangulaire est inversible si, et seulement si, elle n'a pas de zéros sur sa diagonale".

**Autrement dit :** un scalaire  $\lambda$  est une valeur propre de  $M = (m_{i,i})$  si, et seulement si,  $\lambda \in \{ m_{i,i} \mid i \in \llbracket 1; n \rrbracket \}$

**Cas des matrices diagonales.**

**Proposition.**

Soit  $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ , on note  $\Delta = \text{Diag}(\alpha_1, \dots, \alpha_n)$

- ❶  $Sp(\Delta) = \{ \alpha_i \mid i \in \llbracket 1; n \rrbracket \}$
- ❷ pour tout  $\lambda \in Sp(\Delta)$ ,  $\dim(E_\lambda(\Delta)) = \text{card}(\{ i \in \llbracket 1; n \rrbracket \mid \alpha_i = \lambda \})$

**Démonstration.**

**Autrement dit ❷ :**

la dimension de l'espace propre de  $\Delta$  associé à  $\lambda$  est égale au nombre d'apparitions de  $\lambda$  sur la diagonale.

**Attention :** La proposition ❷ est fausse pour une matrice triangulaire.

**Résultats classiques, mais hors programme.***(A savoir redémontrer) Voir la feuille\_cours\_7*Transposée (**Ex 7**). Somme des lignes (resp. des colonnes) constantes (**Ex 11**). Polynôme annulateur (**Ex 9, 10**)**8.2.3 Lien entre matrice et endomorphisme.**Ici  $E$  est un  $\mathbb{K}$ -espace vectoriel de dimension  $n \in \mathbb{N}^*$ .**Théorème** (*Valeur propre.*)

Soient  $f \in \mathcal{L}(E)$ ,  $\lambda \in \mathbb{K}$  et  $\mathcal{B}$  une base quelconque de  $E$ ,  
 $\lambda$  est une valeur propre de  $f$  si, et seulement si,  $\lambda$  est une valeur propre de  $\text{Mat}_{\mathcal{B}}(f)$ .

**En effet :** En notant :  $M = \text{Mat}_{\mathcal{B}}(f)$ ,

$$\lambda \in \text{Sp}(f) \iff \text{rg}(f - \lambda \text{Id}_E) < n \iff \text{rg}(M - \lambda I_n) < n \iff \lambda \in \text{Sp}(M)$$

**Autrement dit :**

Soient  $f \in \mathcal{L}(E)$  et  $\lambda \in \mathbb{K}$ , pour toute base  $\mathcal{B}$  de  $E$ ,  $\text{Sp}(f) = \text{Sp}(\text{Mat}_{\mathcal{B}}(f))$ .

**Corollaires :**

- ❶ Si  $\mathcal{B}$  et  $\mathcal{B}'$  sont deux bases de  $E$ , alors  $\text{Sp}(\text{Mat}_{\mathcal{B}}(f)) = \text{Sp}(\text{Mat}_{\mathcal{B}'}(f))$ .
- ❷ Pour  $A$  et  $B$  deux matrices carrées, si  $A$  et  $B$  sont semblables alors  $\text{Sp}(A) = \text{Sp}(B)$

**En effet :**

- ❶  $\text{Sp}(\text{Mat}_{\mathcal{B}}(f)) = \text{Sp}(f) = \text{Sp}(\text{Mat}_{\mathcal{B}'}(f))$
- ❷ Si  $A$  et  $B$  sont semblables alors elles sont les matrices d'un même endomorphisme dans deux bases donc en utilisant ❶ il vient  $\text{Sp}(A) = \text{Sp}(B)$

**Théorème** (*Vecteurs propres .*)

Soient  $f \in \mathcal{L}(E)$ ,  $u \in E$ ,  $\lambda \in \mathbb{K}$  et  $\mathcal{B}$  une base de  $E$ .  
 $u$  est un vecteur propre de  $f$  associé à la valeur propre  $\lambda$   
si, et seulement si,  $\text{Coord}_{\mathcal{B}}(u)$  est un vecteur propre de  $\text{Mat}_{\mathcal{B}}(f)$  associé à la valeur propre  $\lambda$

**En effet :** Avec les propriétés de l'application  $\text{Coord}_{\mathcal{B}}(\cdot)$  on a :

$$f(u) = \lambda u \iff \text{Mat}_{\mathcal{B}}(f)\text{Coord}_{\mathcal{B}}(u) = \lambda \text{Coord}_{\mathcal{B}}(u)$$

**Propositions** (*Sous-espaces propres.*)

Soient  $f \in \mathcal{L}(E)$ ,  $u_1, \dots, u_m$  des vecteurs de  $E$ ,  $\lambda \in \text{Sp}(f)$  et  $\mathcal{B}$  une base de  $E$ ,  
on note  $M = \text{Mat}_{\mathcal{B}}(f)$  et pour tout  $i \in \llbracket 1; m \rrbracket$ ,  $X_i = \text{Coord}_{\mathcal{B}}(u_i)$ ,

- ❶  $(X_1, \dots, X_m)$  est une base de  $E_{\lambda}(M)$  si, et seulement si,  $(u_1, \dots, u_m)$  est une base de  $E_{\lambda}(f)$
- ❷  $\dim(E_{\lambda}(M)) = \dim(E_{\lambda}(f))$

**En effet :** On utilise les propriétés de l'application  $\text{Coord}_{\mathcal{B}}(\cdot)$  (3.6.2)**Corollaires :**

- ❶ Si  $\mathcal{B}$  et  $\mathcal{B}'$  sont deux bases de  $E$ , alors  $\forall \lambda \in \text{Sp}(f)$ ,  $\dim(E_{\lambda}(\text{Mat}_{\mathcal{B}}(f))) = \dim(E_{\lambda}(\text{Mat}_{\mathcal{B}'}(f)))$ .
- ❷ Pour  $A$  et  $B$  deux matrices carrées,  
Si  $A$  et  $B$  sont semblables alors  $\forall \lambda \in \text{Sp}(A)$ ,  $\dim(E_{\lambda}(A)) = \dim(E_{\lambda}(B))$

**En effet :**  $\dim(E_{\lambda}(\text{Mat}_{\mathcal{B}}(f))) = \dim(\ker(f - \lambda \text{Id}_E)) = \dim(E_{\lambda}(\text{Mat}_{\mathcal{B}'}(f)))$

### 8.3 Propriétés des familles de vecteurs propres.

Soient  $E$  un espace vectoriel,  $f$  un endomorphisme de  $E$  et  $m$  un entier naturel non nul. et soient  $n$  un entier non nul et  $M$  une matrice de  $\mathcal{M}_n(\mathbb{K})$

#### 8.3.1 Vecteurs propres associés à des valeurs propres distinctes.

**Théorème.**

Soient  $u_1, \dots, u_m$  des vecteurs de  $E$  et  $\lambda_1, \dots, \lambda_m$  des scalaires.  
 Si  $\lambda_1, \dots, \lambda_m$  sont  $m$  valeurs propres distinctes de  $f$  et  
 si  $u_1, \dots, u_m$  sont des vecteurs propres de  $f$  associés respectivement à  $\lambda_1, \dots, \lambda_m$ ,  
 alors  $(u_1, \dots, u_m)$  est une famille libre.

**Démonstration :**

On fixe  $m$  dans  $\mathbb{N}^*$ . et on considère  $u_1, \dots, u_m$  des vecteurs propres associés respectivement à  $m$  valeurs propres distinctes  $\lambda_1, \dots, \lambda_m$ .

Montrons par une récurrence finie que  $(u_1, \dots, u_m)$  est une famille libre.

- Pour  $k = 1$ ,  
 comme  $u_1$  est un vecteur propre  $u_1 \neq 0_E$  donc  $(u_1)$  est une famille libre.

- Soit  $k \in \llbracket 1, m - 1 \rrbracket$  tel que  $(u_1, \dots, u_k)$  est une famille libre,

*(Montrons que  $(u_1, \dots, u_{k+1})$  est une famille libre)*

Soient  $(\alpha_1, \dots, \alpha_{k+1}) \in \mathbb{K}^{k+1}$  tel que  $\sum_{j=1}^{k+1} \alpha_j u_j = 0_E$ , **(1)**

comme  $f$  est linéaire on en déduit  $\sum_{j=1}^{k+1} \alpha_j f(u_j) = 0_E$ , puis  $\sum_{j=1}^{k+1} \alpha_j \lambda_j u_j = 0_E$  **(2)**

en faisant **(2)** -  $\lambda_{k+1}$ **(1)** on obtient :  $\sum_{j=1}^k \alpha_j (\lambda_j - \lambda_{k+1}) u_j = 0_E$ ,

et comme on a supposé que  $(u_1, \dots, u_k)$  est libre on en déduit  $\forall j \in \llbracket 1, k \rrbracket, \alpha_j (\lambda_j - \lambda_{k+1}) = 0$

sachant de plus que  $\forall j \in \llbracket 1, k \rrbracket, \lambda_j - \lambda_{k+1} \neq 0$  on a  $\forall j \in \llbracket 1, k \rrbracket, \alpha_j = 0$ .

en reprenant **(1)** on en déduit  $\alpha_{k+1} u_{k+1} = 0_E$  qui entraîne  $\alpha_{k+1} = 0$  car  $u_{k+1} \neq 0_E$ .

on a bien montré que :  $\sum_{j=1}^{k+1} \alpha_j u_j = 0_E \implies \forall j \in \llbracket 1, k \rrbracket, \alpha_j = 0$

*(La famille  $(u_1, \dots, u_{k+1})$  est libre)*

- En conclusion *(de ce raisonnement par récurrence)* :

$(u_1, \dots, u_m)$  est une famille libre

**Corollaire.**

Soient  $X_1, \dots, X_m$  des matrices de  $\mathcal{M}_{n,1}(\mathbb{K})$  et  $\lambda_1, \dots, \lambda_m$  des scalaires.  
 Si  $\lambda_1, \dots, \lambda_m$  sont  $m$  valeurs propres distinctes de la matrice  $M$  et  
 si  $X_1, \dots, X_m$  sont des vecteurs propres de  $M$  associés respectivement à  $\lambda_1, \dots, \lambda_m$ ,  
 alors  $(X_1, \dots, X_m)$  est une famille libre.

**En effet :** Il suffit d'appliquer le théorème précédent à l'endomorphisme  $X \mapsto MX$ .

#### 8.3.2 Juxtaposition des bases des sous-espaces propres.

**Théorème.**

Soient  $\mathcal{B}_1, \dots, \mathcal{B}_m$  des familles de vecteurs de  $E$  et  $\lambda_1, \dots, \lambda_m$  des scalaires.  
 Si  $\lambda_1, \dots, \lambda_m$  sont  $m$  valeurs propres distinctes de  $f$  et  
 si  $\mathcal{B}_1, \dots, \mathcal{B}_m$  sont respectivement des bases des sous espaces propres  $E_{\lambda_1}(f), \dots, E_{\lambda_m}(f)$ ,  
 alors  $\underbrace{(\mathcal{B}_1, \dots, \mathcal{B}_m)}_{\text{juxtaposition des bases}}$  est une famille libre.

**Démonstration :**

1. Montrons le lemme suivant :

$$\text{Si } \forall i \in \llbracket 1; m \rrbracket, u_i \in E_{\lambda_i}(f) \quad \text{et} \quad \sum_{i=1}^m u_i = 0_E \quad \text{alors} \quad \forall i \in \llbracket 1; m \rrbracket, u_i = 0_E$$

Soient  $m \in \mathbb{N}^*$ ,  $\lambda_1, \dots, \lambda_m$   $m$  valeurs propres distinctes de  $f$  et  $u_1, \dots, u_m$  des vecteurs vérifiant :

$$\forall i \in \llbracket 1; m \rrbracket, u_i \in E_{\lambda_i}(f) \quad \text{et} \quad \sum_{i=1}^m u_i = 0_E$$

Montrons par l'absurde que :  $\forall i \in \llbracket 1; m \rrbracket, u_i = 0_E$ ,

On suppose qu'il existe au moins un  $i$  pour lequel  $u_i \neq 0_E$ ,

on note alors  $u_{i_1}, u_{i_2}, \dots, u_{i_r}$  les vecteurs non nuls de la famille  $(u_1, \dots, u_m)$

on a alors  $\sum_{k=1}^r u_{i_k} = 0_E$  et  $u_{i_1}, u_{i_2}, \dots, u_{i_r}$  sont des vecteurs propres associés à  $r$  valeurs propres distinctes.

ce qui est impossible car le théorème démontré en 8.3.1 montre que la famille  $(u_{i_1}, u_{i_2}, \dots, u_{i_r})$  est une famille libre.

En conclusion :

$$\text{Si } \forall i \in \llbracket 1; m \rrbracket, u_i \in E_{\lambda_i}(f) \quad \text{et} \quad \sum_{i=1}^m u_i = 0_E \quad \text{alors} \quad \forall i \in \llbracket 1; m \rrbracket, u_i = 0_E$$

2. Montrons maintenant le théorème :

On suppose que :  $\lambda_1, \dots, \lambda_m$  sont  $m$  valeurs propres distinctes de  $f$

et que  $\mathcal{B}_1, \dots, \mathcal{B}_m$  sont respectivement des bases des sous espaces propres  $E_{\lambda_1}(f), \dots, E_{\lambda_m}(f)$ .

On utilise les notations suivantes.

Pour  $i \in \llbracket 1, m \rrbracket$ ,  $\mathcal{B}_i = (u_{i,j})_{1 \leq j \leq n_i}$  une base de  $E_{\lambda_i}(f)$

et  $I = \{(i, j) \in \mathbb{N}^2 \mid 1 \leq i \leq m \text{ et } 1 \leq j \leq n_i\}$ ,

Montrons que  $\mathcal{B} = (u_{i,j})_{(i,j) \in I}$  est une famille libre.

Soit  $(\alpha_{i,j}) \in \mathbb{K}^I$  tel  $\sum_{(i,j) \in I} \alpha_{i,j} u_{i,j} = 0_E$ ,

$$\text{on a alors : } \sum_{i=1}^m \underbrace{\left( \sum_{j=1}^{n_i} \alpha_{i,j} u_{i,j} \right)}_{\in E_{\lambda_i}(f)} = 0_E \quad \text{ce qui entraîne avec le lemme de la question 1) que :}$$

$$\forall i \in \llbracket 1; m \rrbracket, \sum_{j=1}^{n_i} \alpha_{i,j} u_{i,j} = 0_E$$

mais on sait que pour chaque  $i$ , la famille  $(u_{i,j})_{1 \leq j \leq n_i}$  est libre donc on peut en déduire que :

$$\forall i \in \llbracket 1; m \rrbracket, \forall j \in \llbracket 1; n_i \rrbracket, \alpha_{i,j} = 0$$

En conclusion :

$$\text{La juxtaposition } (\mathcal{B}_1, \dots, \mathcal{B}_m) \text{ est libre.}$$

**Corollaire.**

Soient  $\mathcal{B}_1, \dots, \mathcal{B}_m$  des familles de vecteurs de  $\mathcal{M}_{n,1}(\mathbb{K})$  et  $\lambda_1, \dots, \lambda_m$  des scalaires.  
 Si  $\lambda_1, \dots, \lambda_m$  sont  $m$  valeurs propres distinctes de  $M$  et  
 si  $\mathcal{B}_1, \dots, \mathcal{B}_m$  sont respectivement des bases des sous espaces propres  $E_{\lambda_1}(M), \dots, E_{\lambda_m}(M)$ ,  
 alors  $\underbrace{(\mathcal{B}_1, \dots, \mathcal{B}_m)}_{\text{juxtaposition des bases}}$  est une famille libre.

**En effet :** Il suffit d'appliquer le théorème précédent à l'endomorphisme  $X \mapsto MX$ .

### 8.3.3 En dimension finie.

#### Nombre de valeurs propres distinctes.

##### **Théorème.**

Soient  $n$  un entier naturel non nul,  $E$  un espace vectoriel,  $f$  un endomorphisme de  $E$ .  
Si  $E$  est de dimension  $n$  alors  $f$  a au plus  $n$  valeurs propres distinctes.

**En effet :** C'est application immédiate du théorème :

En dimension  $n$  une famille libre contient au plus  $n$  vecteurs.

*Remarque :* Si on a trouvé  $n$  valeurs propres distinctes, il est alors inutile de chercher, on les a toutes.

##### **Corollaire.**

Soient  $n$  un entier naturel non nul, les matrices de  $\mathcal{M}_n(\mathbb{K})$  ont au plus  $n$  valeurs propres distinctes.

**En effet :** Il suffit d'appliquer le théorème précédent à l'endomorphisme  $X \mapsto MX$ .

#### Somme des dimensions des sous-espaces propres.

##### **Théorème.**

Soient  $E$  un espace vectoriel de dimension finie,  $f$  un endomorphisme de  $E$  et  $m$  un entier naturel non nul,  $\lambda_1, \dots, \lambda_m$  des scalaires.  
Si  $\lambda_1, \dots, \lambda_m$  sont  $m$  valeurs propres distinctes de  $f$  alors  $\sum_{k=1}^m \dim(E_{\lambda_k}(f)) \leq \dim(E)$

**En effet :** C'est une conséquence directe du théorème 8.3.2.

*Remarque :* Si on a trouvé  $m$  valeurs propres  $\lambda_1, \dots, \lambda_m$  distinctes et que  $\sum_{k=1}^m \dim(E_{\lambda_k}(f)) = \dim(E)$ , il est alors inutile de chercher, on les a toutes.

##### **Corollaire.**

Soit  $M \in \mathcal{M}_n(\mathbb{K})$  et  $m$  un entier naturel non nul,  $\lambda_1, \dots, \lambda_m$  des scalaires.  
Si  $\lambda_1, \dots, \lambda_m$  sont  $m$  valeurs propres distinctes de  $M$  alors  $\sum_{k=1}^m \dim(E_{\lambda_k}(M)) \leq n$

**En effet :** Il suffit d'appliquer le théorème précédent à l'endomorphisme  $X \mapsto MX$ .

*Remarque :* Si on a trouvé  $m$  valeurs propres  $\lambda_1, \dots, \lambda_m$  distinctes et que  $\sum_{k=1}^m \dim(E_{\lambda_k}(f)) = \dim(E)$ , il est alors inutile de chercher, on les a toutes.

**Rappel utile ici :** on peut calculer  $\dim(E_{\lambda}(M))$  avec la relation :  $\dim(E_{\lambda}(M)) = n - \text{rg}(M - \lambda I_n)$

## 8.4 Diagonalisation.

Soit  $E$  un espace vectoriel de dimension finie.

### 8.4.1 Définition.

#### Pour un endomorphisme.

##### **Définition**

Soit  $f$  un endomorphisme de  $E$ ,  
dire que  $f$  est diagonalisable signifie qu'il existe une base  $\mathcal{B}$  de  $E$  telle que  $\text{Mat}_{\mathcal{B}}(f)$  est diagonale.

**Autrement dit :**

$f$  est diagonalisable si, et seulement si,  $E$  possède une base constituée de vecteurs propres de  $f$ .

**Vocabulaire.** "Diagonaliser  $f$ " signifie :

- trouver une base de  $E$  pour laquelle  $Mat_{\mathcal{B}}(f)$  est diagonale.
- ou
- trouver une base de  $E$  formée de vecteurs propres de  $f$ .

**Remarque :** Certains endomorphismes ne sont pas diagonalisables.

**Exemple :**  $f : (x, y) \mapsto (x + y, y)$  n'est pas diagonalisable.

**Pour une matrice carrée.****Définition.**

Soit  $M \in \mathcal{M}_n(\mathbb{K})$ , ,  
dire que  $M$  est diagonalisable signifie que  $M$  est semblable à une matrice diagonale.

**Remarque :** Certaines matrices ne sont pas diagonalisables.

**Vocabulaire.** "Diagonaliser  $M$ " signifie trouver deux matrices :  $\left\{ \begin{array}{l} P \text{ inversible} \\ \Delta \text{ diagonale} \end{array} \right.$ , telles que  $M = P\Delta P^{-1}$

Lien avec "Diagonaliser  $f : X \mapsto MX$ " :

$P$  est la matrice d'une base ( $P$  inversible), formée de vecteurs propre de  $M$  ( $M = P\Delta P^{-1}$ ).

**Remarques :**

- Les matrices diagonales sont diagonalisables.  
(En particulier la matrice nulle et la matrice identité de  $\mathcal{M}_n(\mathbb{K})$  sont diagonalisables).
- Pour  $A$  et  $B$  sont semblables alors ( $A$  est diagonalisable si, et seulement si,  $B$  est diagonalisable).
- Vu en exercice : Si  $M = P\Delta P^{-1}$  alors la trace<sup>1</sup> de  $M$  est égale à celle de  $\Delta$

Autrement dit : Si  $M$  est diagonalisable alors  $\text{tr}(M) = \sum_{\lambda \in \text{sp}(M)} \lambda \dim(E_{\lambda}(M))$

A la fin d'un exercice, vérifier que la trace de  $M$  est égale à celle de la matrice diagonale obtenue.

**Théorème.**

Soient  $f$  un endomorphisme de  $E$  et  $\mathcal{B}$  une base de  $E$   
 $f$  est diagonalisable si, et seulement si,  $Mat_{\mathcal{B}}(f)$  est diagonalisable.

**En effet :**

$\Rightarrow$  Si  $f$  est diagonalisable, il existe une base  $\mathcal{B}'$  de  $E$  formée de vecteurs propres de  $f$ ,  
alors  $Mat_{\mathcal{B}'}(f) = \Delta$  est diagonale.

En notant  $P$  la matrice de passage de  $\mathcal{B}$  vers  $\mathcal{B}'$ , on a :  $Mat_{\mathcal{B}}(f) = P\Delta P^{-1}$ , donc  $Mat_{\mathcal{B}}(f)$  est diagonalisable.

$\Leftarrow$  Si  $Mat_{\mathcal{B}}(f)$  est diagonalisable, il existe une matrice inversible  $P$  et une matrice diagonale  $\Delta$  telles que  $Mat_{\mathcal{B}}(f) = P\Delta P^{-1}$ .

Notons  $\mathcal{B}'$  la base dont les vecteurs ont pour coordonnées dans  $\mathcal{B}$  les colonnes de  $P$ .  
Alors  $P$  est la matrice de passage de  $\mathcal{B}$  vers  $\mathcal{B}'$ , et  $Mat_{\mathcal{B}'}(f) = \Delta$ , qui est diagonale.  
Ainsi  $\mathcal{B}'$  est une base de vecteurs propres de  $f$ , donc  $f$  est diagonalisable.

1. On rappelle que la notion de trace n'est pas au programme.

### 8.4.2 Condition suffisante.

**Théorème.** (Condition suffisante pour qu'un endomorphisme soit diagonalisable).

Soit  $f \in \mathcal{L}(E)$  ( où  $E$  un espace vectoriel de dimension  $n$  ),  
 Si  $f$  possède  $n$  valeurs propres distinctes alors  $f$  est diagonalisable.  
 et tous les sous-espaces propres sont de dimension 1.

**En effet :** • Les  $n$  vecteurs propres associés aux  $n$  valeurs propres distinctes forment une famille libre de  $n$  vecteurs de  $E$ , de dimension  $n$  donc ces  $n$  vecteurs forment une base de vecteurs propres de  $f$ .  $f$  est diagonalisable.

• Comme  $\forall \lambda \in Sp(f)$ ,  $\dim(E_\lambda(f)) \geq 1$  et  $\sum_{\lambda \in Sp(f)} \dim(E_\lambda(f)) \leq n$  il vient que  $\forall \lambda \in Sp(f)$ ,  $\dim(E_\lambda(f)) = 1$ .

**Corollaire.** (Condition suffisante pour qu'une matrice carrée soit diagonalisable)

Soit  $M \in \mathcal{M}_n(\mathbb{K})$ , Si  $M$  possède  $n$  valeurs propres distinctes alors  $M$  est diagonalisable.  
 et tous les sous-espaces propres sont de dimension 1.

**En effet :** Il suffit d'appliquer le théorème précédent à l'endomorphisme  $X \mapsto MX$ .

**Attention :**

Si  $M$  (ou  $f$ ) n'a pas  $n$  valeurs propres distinctes alors on ne peut rien conclure avec ce théorème.

### 8.4.3 Condition nécessaire et suffisante.

**Théorème.** (Condition nécessaire et suffisante pour qu'un endomorphisme soit diagonalisable).

Soit  $f \in \mathcal{L}(E)$  ( où  $E$  un espace vectoriel de dimension  $n$  ),  
 $f$  est diagonalisable si, et seulement si,  
 la somme des dimensions des sous-espaces propres de  $f$  est égale à  $n$ .

**Démonstration.** (admise)

**Autrement dit :** (Encore une occasion de comprendre la notion d'équivalence)

- Si  $\sum_{\lambda \in Sp(f)} \dim(E_\lambda(f)) = n$  alors  $f$  est diagonalisable
- Si  $\sum_{\lambda \in Sp(f)} \dim(E_\lambda(f)) \neq n$  alors  $f$  n'est pas diagonalisable

**Remarque :**

si  $f$  n'a pas de valeur propre alors  $\sum_{\lambda \in Sp(f)} \dim(E_\lambda(f)) = 0$  et ainsi  $f$  n'est pas diagonalisable.

**Corollaire.** (Condition nécessaire et suffisante pour qu'une matrice carrée soit diagonalisable).

Soit  $M \in \mathcal{M}_n(\mathbb{R})$ ,  
 $M$  est diagonalisable si, et seulement si,  
 la somme des dimensions des sous-espaces propres de  $M$  est égale à  $n$ .

**En effet :** Il suffit d'appliquer le théorème précédent à l'endomorphisme  $X \mapsto MX$ .

**Autrement dit :**

- Si  $\sum_{\lambda \in Sp(M)} \dim(E_\lambda(M)) = n$  alors  $M$  est diagonalisable
- Si  $\sum_{\lambda \in Sp(M)} \dim(E_\lambda(M)) \neq n$  alors  $M$  n'est pas diagonalisable

**8.4.4 En pratique pour les matrices carrées.**

$M \in \mathcal{M}_n(\mathbb{K})$ .

On a montré que  $M$  est diagonalisable avec l'un des théorèmes précédents et on a trouvé une base de chaque sous-espace propre et en les juxtaposant on a une base de vecteurs propres,

autrement dit on a :

$(U_1, U_2, \dots, U_n)$  une base de  $\mathcal{M}_{n,1}(\mathbb{K})$  et  $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$  tels que :

$$MU_1 = \lambda_1 U_1, \quad MU_2 = \lambda_2 U_2, \quad \dots, \quad MU_n = \lambda_n U_n \quad (*)$$

On note  $P$  la matrice de  $(U_1, U_2, \dots, U_n)$  dans la base canonique de  $\mathcal{M}_{n,1}(\mathbb{K})$  et  $\Delta = \text{Diag}(\lambda_1, \dots, \lambda_n)$ ,

d'une part  $(U_1, U_2, \dots, U_n)$  est une base donc  $P$  est inversible

et d'autre part on a :

$$M = P\Delta P^{-1}$$

**❶ Première explication.**

Les relations (\*) peuvent s'écrire :

$$M(U_1 | U_2 | \dots | U_n) = (\lambda_1 U_1 | \lambda_2 U_2 | \dots | \lambda_n U_n)$$

ou encore

$$MP = P\Delta$$

et comme  $P$  est inversible, on obtient

$$\boxed{M = P\Delta P^{-1}}$$

**❷ Deuxième explication.**

On note  $f : X \mapsto MX$ ,  $\mathcal{C}$  la base canonique de  $\mathcal{M}_{n,1}(\mathbb{K})$  et  $\mathcal{B}$  la base  $(U_1, U_2, \dots, U_n)$

$P$  est la matrice de changement de base de  $\mathcal{C}$  à  $\mathcal{B}$ ,  $M = \text{Mat}_{\mathcal{C}}(f)$  et  $\Delta = \text{Mat}_{\mathcal{B}}(f)$ ,

la formule de changement de variable donne :  $\Delta = P^{-1}MP$  ou encore :

$$\boxed{M = P\Delta P^{-1}}$$

**Théorème :**

Si  $P$  est la matrice d'une base de vecteurs propres de  $M \in \mathcal{M}_n(\mathbb{R})$  alors  $M = P\Delta P^{-1}$

où  $\Delta = \text{Diag}(\lambda_1, \dots, \lambda_n)$  avec  $\lambda_j$  la valeur propre associée à la colonne  $C_j$  de  $P$ .

**Rédaction** (*A adapter à chaque situation*)

**Exemple :** Diagonaliser la matrice  $M = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$

- On a trouvé le spectre en cherchant les  $\lambda$  pour lesquels  $\text{rg}(M - \lambda I_3) < 3$   $\boxed{\text{Sp}(M) = \{1, 2\}}$
- On a trouvé des bases des sous-espaces propres en résolvant les systèmes :  $(M - \lambda I_3)X = 0$  :

$$\left( \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right) \text{ est une base de } E_2(M) \quad \text{et} \quad \left( \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right) \text{ est une base de } E_1(M).$$

On remarque que la somme des dimensions des sous-espaces propres de  $M$  est égale à 3 donc  $\boxed{M \text{ est diagonalisable}}$

en juxtaposant les bases on obtient une base de vecteurs propres de  $M$ , plus précisément :

En prenant :  $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & -1 \end{pmatrix}$  et  $\Delta = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  on a  $P$  inversible et :

$$\boxed{M = P\Delta P^{-1}}$$

### 8.4.5 Application au calcul de puissance de matrices carrées.

Lorsque  $M$  est diagonalisable il est simple de calculer les puissances de  $M$

$$\text{Si } M = P \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \lambda_{n-1} & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix} P^{-1} \quad \text{alors pour tout } k \in \mathbb{N}, \quad M^k = P \begin{pmatrix} \lambda_1^k & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \lambda_{n-1}^k & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n^k \end{pmatrix} P^{-1}$$

**Sauf indication contraire vous devez refaire la démonstration dans une copie.**

Revoir le paragraphe 6.10 sur les matrices semblables

### 8.4.6 Application aux systèmes différentielles linéaires.

On s'intéresse ici aux systèmes du type :

$$X'(t) = MX(t) \quad \text{où } M \in \mathcal{M}_n(\mathbb{R})$$

En posant  $Y(t) = P^{-1}X(t)$

$$\begin{aligned} X'(t) = MX(t) &\iff X'(t) = P\Delta P^{-1}X(t) \\ &\iff P^{-1}X'(t) = \Delta P^{-1}X(t) \\ &\iff Y'(t) = \Delta Y(t) \end{aligned}$$

Ce dernier système est simple à résoudre : chaque ligne est une équation du type :  $y_i'(t) = \lambda_i y_i(t)$

**Exemple :** Voir l'Ex 8 de la feuille\_Exo\_16.

# Variables aléatoires réelles à densité.

## Plan du chapitre

9.1	Densités de probabilité et fonction de répartition. . . . .	<b>115</b>
9.1.1	Densités de probabilité. . . . .	115
9.1.2	Propriétés de la fonction de répartition. . . . .	116
9.1.3	Caractérisation des variables à densité. . . . .	117
9.1.4	Loi de $X$ . . . . .	117
9.2	Méthodes pratiques. . . . .	<b>117</b>
9.2.1	De la fonction de répartition à une densité. . . . .	118
9.2.2	D'une densité à la fonction de répartition. . . . .	118
9.3	Calculs de probabilité. . . . .	<b>119</b>
9.3.1	Des événements négligeables. . . . .	119
9.3.2	Avec la fonction de répartition. . . . .	120
9.3.3	Avec une densité. . . . .	120
9.3.4	Valeurs prises par $X$ . . . . .	120
9.3.5	Indépendance . . . . .	121
9.4	Exemples de recherche de la loi de $u(X)$ . . . . .	<b>121</b>
9.5	Espérance. Propriétés. . . . .	<b>122</b>
9.5.1	Théorème de transfert. . . . .	123
9.5.2	Variance d'une variable aléatoire. . . . .	123
9.6	Lois à densité usuelles. . . . .	<b>124</b>
9.6.1	Lois uniformes à densité. . . . .	124
9.6.2	Lois exponentielles. . . . .	126
9.6.3	Lois normales. . . . .	129
9.7	Somme de variables aléatoires à densité indépendantes. . . . .	<b>139</b>
9.7.1	Produit de convolution. . . . .	139
9.7.2	Somme de variables aléatoires normales indépendantes. . . . .	139

## 9.1 Densités de probabilité et fonction de répartition.

### 9.1.1 Densités de probabilité.

#### Définition

Soit  $f$  une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$ .

Dire que  $f$  est une densité de probabilité signifie que :

- ❶  $f$  est positive ou nulle sur  $\mathbb{R}$ ,
- ❷  $f$  est continue sur  $\mathbb{R}$  sauf en un nombre fini de points
- ❸  $\int_{-\infty}^{+\infty} f(t) dt$  converge et vaut 1.

Remarques :

- si  $f$  est une densité de probabilité alors quels que soient  $a$  et  $b$  vérifiant  $-\infty \leq a, b \leq +\infty$ ,  $\int_a^b f(t) dt$  converge.
- si  $f$  est une densité de probabilité alors on peut définir sur  $\mathbb{R}$ , la fonction  $x \mapsto \int_{-\infty}^x f(t) dt$

**Définition**

On considère  $X$  une variable aléatoire réelle d'un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .  
 Dire que  $X$  est une variable aléatoire réelle à densité signifie qu'il existe une densité de probabilité  $f$  telle que :

$$\forall x \in \mathbb{R}, \quad \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t) dt$$

**Remarques.**

- ❶ Une telle fonction n'est pas unique et est appelée densité de  $X$ .
- ❷ En notant  $F_X$  la fonction de répartition de  $X$ , on a :  $\forall x \in \mathbb{R}, \quad F_X(x) = \int_{-\infty}^x f(t) dt$
- ❸ Dire que deux variables aléatoires suivent la même loi signifie qu'elles ont la même fonction de répartition.

**Proposition.**

Soient  $f$  et  $g$  deux fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$  à valeurs positives.  
 Si  $f$  est une densité de  $X$  et si  $g$  coïncide avec  $f$  sauf un nombre fini de points alors  $g$  est une densité de  $X$ .

**En effet :** Les fonctions coïncident sauf sur un ensemble fini donc les deux intégrales convergent et valent 1.

**En pratique :** Quand on choisit une densité on peut choisir arbitrairement l'image d'un nombre fini de points.

**Théorème.**

Soit  $f$  une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$ .  
 Si  $f$  est une densité de probabilité alors  
 il existe un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et une variable aléatoire sur cet espace dont  $f$  est une densité.

**Démonstration (Admis)**

**9.1.2 Propriétés de la fonction de répartition.**

**Propriétés des fonctions de répartition des variables aléatoires à densité.**

Si  $F$  est la fonction de répartition d'une variable à densité alors :

- ❶  $F$  est continue sur  $\mathbb{R}$ .
- ❷  $F$  est de classe  $C^1$  sur  $\mathbb{R}$  sauf éventuellement en un nombre fini de points.
- ❸  $F$  est croissante sur  $\mathbb{R}$ .
- ❹  $\lim_{x \rightarrow -\infty} F(x) = 0$  et  $\lim_{x \rightarrow +\infty} F(x) = 1$
- ❺ En tout  $x$  où  $f$  est continue,  $F$  est dérivable,  $F'(x) = f(x)$

**Démonstration : (Rapidement)**

• La continuité à droite, ❸ et ❹ sont des propriétés communes à toutes les fonctions de répartition (*pas seulement celles des variables aléatoires à densité*)

• On subdivise  $]a, b[$  avec  $a_0 < \dots < a_n$  tels que  $f$  continue sur les ouverts  $]a_i, a_{i+1}[$ .

◊ Comme  $\forall x \in ]a_i, a_{i+1}[$ ,  $F(x) = \int_{-\infty}^x f(t) dt$  on a  $F$  de classe  $C^1$  et  $F' = f$  sur  $]a_i, a_{i+1}[$

◊  $\lim_{x \rightarrow a_{i+1}^-} \int_{-\infty}^x f(t) dt = \int_{-\infty}^{a_{i+1}} f(t) dt = F(a_{i+1})$

**Proposition :**

Les variables aléatoires discrètes ne sont pas des variables aléatoires à densité

**En effet :** Le fonction de répartition des variables aléatoires discrètes n'est pas continue.

**Proposition :**

Si  $X$  est de densité  $f$  et que  $f$  est continue sur  $\mathbb{R}$  alors  $F_X$  est  $C^1$  sur  $\mathbb{R}$  et  $F' = f$

**En effet :** Ici  $f$  est continue sur  $\mathbb{R}$ ,  $F : x \mapsto \int_{-\infty}^x f(t) dt$  est une primitive de  $f$  sur  $\mathbb{R}$

**9.1.3 Caractérisation des variables à densité.**

**Théorème.**

Soit  $X$  une variable aléatoire, on note  $F_X : \mathbb{R} \rightarrow \mathbb{R}$   
 $x \mapsto \mathbb{P}(X \leq x)$   
 $X$  est à densité **si, et seulement si**,  
 ❶  $F_X$  est continue sur  $\mathbb{R}$  (en entier) et  
 ❷  $F_X$  est de classe  $C^1$  sur  $\mathbb{R}$  sauf en un nombre fini de points.

**Démonstration :**

**En pratique.**

Lorsqu'une variable aléatoire est définie par sa fonction de répartition  $F_X$  :

Si ❶  $F_X$  est continue sur  $\mathbb{R}$  et ❷ de classe  $C^1$  sauf en un nombre fini de points.  
 alors on peut affirmer que  $X$  est une variable à densité.

Si ❶  $F$  discontinue en (au moins) un  $x_0 \in \mathbb{R}$  ou si ❷  $F$  n'est pas de classe  $C^1$  en un nombre infini de points  
 alors on peut affirmer que  $X$  n'est pas une variable à densité.

**Remarques :**

- Dans ce théorème on fait l'hypothèse que  $F$  est une fonction de répartition donc on sait déjà que :

$$F \text{ est croissante, que } \lim_{-\infty} F = 0 \text{ et que } \lim_{+\infty} F = 1$$

- Les variables discrètes ne sont pas à densité, mais il en existe d'autres, des variables ni discrètes, ni à densité.

**9.1.4 Loi de  $X$ .**

**Dans ce chapitre, donner la loi d'une variable aléatoire  $X$ , c'est justifier que  $X$  admet une densité et en donner une.**

**En pratique :** (Le plus souvent).

- ❶ On détermine  $F_X$  la fonction de répartition de  $X$ .
- ❷ On vérifie que  $F_X$  est continue sur  $\mathbb{R}$  en entier, qu'elle est  $C^1$  sauf en un nombre fini de points.
- ❸ On peut alors affirmer que  $X$  est une variable à densité.
- ❹ On détermine alors une densité de  $X$  en dérivant  $F_X$  sur les intervalles où elle est dérivable.

**9.2 Méthodes pratiques.**

Dans les deux propositions suivantes on note :

Soient  $n \in \mathbb{N}$  et  $(a_1, a_2, a_3, \dots, a_n) \in \mathbb{R}^n$  des réels vérifiant :  $a_1 < a_2 < a_3 < \dots < a_n$ .  
un nombre fini de points de  $\mathbb{R}$ .

on note :  $D = \{a_1, a_2, a_3, \dots, a_n\}$

on remarque que  $\mathbb{R} \setminus D$  ( $\mathbb{R}$  privé d'un nombre fini de points.) peut s'écrire :

$$\mathbb{R} \setminus D = ] - \infty, a_1 [ \cup ] a_1, a_2 [ \cup \dots \cup ] a_{n-1}, a_n [ \cup ] a_n, +\infty [$$

(C'est une réunion d'intervalles)

### 9.2.1 De la fonction de répartition à une densité.

#### Théorème

Soient  $F$  une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$  et  $X$  une variable aléatoire à densité.  
 si  $F$  est la fonction de répartition de  $X$  et si  $F$  est de classe  $C^1$  sur  $\mathbb{R} \setminus D$   
 alors la fonction  $f$  définie par 
$$\begin{cases} \forall x \in \mathbb{R} \setminus D, & f(x) = F'(x) \\ \forall x \in D, & f(x) = 0 \end{cases}$$
 est une densité de  $X$ .

#### Remarques :

- On définit ici une des densités de  $X$ .
- Lorsque  $F$  est constante sur un intervalle alors  $f$  est nulle sur cet intervalle.
- Bien vérifier (*au brouillon*) à la fin de vos calculs que la fonction obtenue est bien une densité de probabilité.  
 La fonction  $f$  doit être positive, continue sauf sur  $D$  et son intégrale sur  $\mathbb{R}$  vaut 1.

#### En pratique :

$$\begin{cases} \forall x \in ]-\infty, a_1[, & f(x) = F'(x) \\ \forall x \in ]a_1, a_2[, & f(x) = F'(x) \\ \vdots & \\ \forall x \in ]a_n, +\infty[, & f(x) = F'(x) \end{cases} \quad \text{et} \quad \forall i \in \llbracket 1, n \rrbracket, \quad f(a_i) = 0 \quad (\text{choix arbitraire dans } \mathbb{R}^+)$$

#### Exemples : Feuille\_Calculs\_8

### 9.2.2 D'une densité à la fonction de répartition.

**Option 1.** On utilise la définition : 
$$\forall x \in \mathbb{R}, \quad F(x) = \int_{-\infty}^x f(t) dt$$

En notant  $f_0$  la restriction de  $f$  sur  $]-\infty, a_1[$ , ...,  $f_i$  sur  $]a_i, a_{i+1}[$ , ...,  $f_n$  sur  $]a_n, +\infty[$  on a :

$$\begin{cases} \forall x \in ]-\infty, a_1[, & F(x) = \int_{-\infty}^x f_0(t) dt \\ \forall x \in ]a_1, a_2[, & F(x) = F(a_1) + \int_{a_1}^x f_1(t) dt \\ \vdots & \\ \forall x \in [a_n, +\infty[, & F(x) = F(a_n) + \int_{a_n}^x f_n(t) dt \end{cases}$$

*Remarque :* On retrouve ci-dessus une présentation qui permet de retrouver les propriétés de régularité de  $F$  :

$F$  de classe  $C^1$  partout sauf (éventuellement) sur  $D$ ,  $F$  est continue sur  $\mathbb{R}$ .

**Option 2.** On utilise la proposition suivante.

Soient  $f$  une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$  et  $X$  une variable aléatoire à densité.  
 si  $f$  est une densité de  $X$  et si  $f$  est continue sur  $\mathbb{R} \setminus D$  alors  
 la fonction de répartition de  $X$  est une primitive de  $f$  sur tous les intervalles :  

$$]-\infty, a_1[, ]a_1, a_2[, \dots, ]a_{n-1}, a_n[ \text{ et } ]a_n, +\infty[$$

Sur chaque intervalle on choisit la primitive de  $f$  qui vérifie :

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \forall i \in \llbracket 1, n \rrbracket, \quad \lim_{x \rightarrow a_i^-} F(x) = \lim_{x \rightarrow a_i^+} F(x) = F(a_i) \quad \text{et} \quad \lim_{x \rightarrow +\infty} F(x) = 1$$

**Remarques :**

- Lorsque  $f$  est nulle sur  $]a_i, a_{i+1}[$  alors  $F$  est constante sur  $[a_i, a_{i+1}]$   
(encore vraie sur  $] -\infty, a_1[$  et  $]a_n, +\infty[$ )
- Bien vérifier (au brouillon) à la fin de vos calculs que la fonction obtenue est bien une fonction de répartition.

Elle doit être croissante, continue sur tout  $\mathbb{R}$ ,  $C^1$  sur  $\mathbb{R} \setminus D$ ,  $\lim_{x \rightarrow -\infty} F(x) = 0$  et  $\lim_{x \rightarrow +\infty} F(x) = 1$ .

**Exemples :** *Feuille\_Calculs\_7***En résumé.**

- On passe de la fonction de répartition à une densité en dérivant sur chaque intervalle.
- On passe d'une densité à la fonction de répartition en déterminant une primitive sur chaque intervalle.  
(La primitive qui permet à  $F$  d'être continue)

**Attention.**

- Dans le cours des fonctions dérivables sur un intervalle :
  - Quand on dérive une fonction  $F$  sur  $I$ , sa dérivée  $f$  est unique.
  - Quand on cherche une primitive de  $f$  sur  $I$ , il y a une infinité de primitives.
- Dans ce cours, c'est le contraire :
  - Pour une densité de probabilité  $f$ , il y a une unique fonction de répartition  $F$ .
  - Pour une fonction de répartition  $F$ , il y a une infinité de densité de probabilité.

### 9.3 Calculs de probabilité.

#### 9.3.1 Des événements négligeables.

**Proposition.**

Soit  $X$  une variable aléatoire réelle d'un espace probablisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .

On suppose  $X$  est une variable à densité, on note  $f$  une de ses densités.

- ❶ Pour tout  $a \in \mathbb{R}$  on a :  $\mathbb{P}(X = a) = 0$ .
- ❷ Pour tout ensemble fini  $D$ ,  $\mathbb{P}(X \in D) = 0$ .
- ❸ Pour tout ensemble dénombrable  $D$ ,  $\mathbb{P}(X \in D) = 0$ .

**Démonstration.**

- ❶ Pour  $\varepsilon > 0$ , on a :  $(X = a) \subset (a - \varepsilon < X \leq a)$  donc  $\forall \varepsilon > 0$ ,  $0 \leq \mathbb{P}(X = a) \leq F(a) - F(a - \varepsilon)$   
or  $X$  est une variable aléatoire à densité donc  $F$  est continue en  $a$  et ainsi  $\lim_{\varepsilon \rightarrow 0} F(a - \varepsilon) = F(a)$   
cela entraîne que  $\boxed{\forall a \in \mathbb{R}, \mathbb{P}(X = a) = 0}$

- ❷ • On note  $D = \{d_i \mid i \in \llbracket 1; n \rrbracket\}$  (avec  $i \neq j \Rightarrow d_i \neq d_j$ )

$$\begin{aligned}
 P(X \in D) &= P\left(\bigcup_{i=1}^n (X = d_i)\right) \\
 &= \sum_{i=1}^n P(X = d_i) \quad \text{car les } (X = d_i) \text{ sont deux à deux incompatibles} \\
 &= \sum_{i=1}^n 0 \quad \text{d'après } \text{❶}
 \end{aligned}$$

$\boxed{\text{Si } D \text{ un ensemble fini et } X \text{ est à densité alors } \mathbb{P}(X \in D) = 0}$

- On note  $D = \{d_n \mid n \in \mathbb{N}\}$  (avec  $i \neq j \Rightarrow d_i \neq d_j$ )

$$\begin{aligned}
 P(X \in D) &= P\left(\bigcup_{n=0}^{+\infty} (X = d_n)\right) \\
 &= \sum_{n=0}^{+\infty} P(X = d_n) \quad \text{car les } (X = d_n) \text{ sont deux à deux incompatibles} \\
 &= \sum_{n=0}^{+\infty} 0 \quad \text{d'après 1)}
 \end{aligned}$$

Si  $D$  un ensemble dénombrable et  $X$  est à densité alors  $P(X \in D) = 0$

### 9.3.2 Avec la fonction de répartition.

**Proposition :**

On considère  $X$  une variable aléatoire réelle d'un espace probablisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .  
 On suppose  $X$  est une variable à densité, on note  $F$  sa fonction de répartition.

- ❶ Pour tout  $(a, b) \in \mathbb{R}^2$  tel que  $a \leq b$  on a :  $\mathbb{P}(X \in [a, b]) = F(b) - F(a)$ .  
*on peut remplacer  $[a, b]$  par  $]a, b]$ ,  $[a, b[$  ou  $]a, b[$ .*
- ❷ Pour tout  $a \in \mathbb{R}$ ,  $\mathbb{P}(X \in ]-\infty; a]) = F(a)$ .  
*on peut remplacer  $] - \infty; a]$  par  $] - \infty; a[$ .*
- ❸ Pour tout  $a \in \mathbb{R}$ ,  $\mathbb{P}(X \in [a; +\infty[) = 1 - F(a)$ .  
*on peut remplacer  $[a; +\infty[$  par  $]a; +\infty[$ .*

**En effet :** On sait que :  $P(X \in ]a, b]) = F(b) - F(a)$ ,  $P(X \leq a) = F(a)$  et  $P(X > a) = 1 - F(a)$ .  
 Et ici on a  $P(X = a) = 0$  d'où les résultats énoncés.

### 9.3.3 Avec une densité.

**Proposition.**

Soit  $X$  une variable aléatoire réelle d'un espace probablisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .  
 On suppose  $X$  est une variable à densité, on note  $f$  une de ses densités.

- ❶ Pour tout  $(a, b) \in \mathbb{R}^2$  tel que  $a \leq b$  on a :  $\mathbb{P}(X \in [a, b]) = \int_a^b f(t) dt$ .
- ❷ Pour tout  $(a, b)$  tel que  $-\infty \leq a \leq b \leq +\infty$  on a :  $\mathbb{P}(X \in ]a, b]) = \int_a^b f(t) dt$ .

**En effet :**  $F(b) - F(a) = \int_{-\infty}^b f(t) dt - \int_{-\infty}^a f(t) dt = \int_a^b f(t) dt$  (Relation de Chasles)

**Remarque :** Dans ce contexte on pourra utiliser la notation :  $\mathbb{P}(X \in I) = \int_I f(t) dt$ .

### 9.3.4 Valeurs prises par X.

**Définition.** (Valeurs de  $X$ )

Soient  $X$  une variable aléatoire réelle d'un espace probablisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et  $I$  un intervalle de  $\mathbb{R}$ ,  
 dire que  $X$  est (quasiment) à valeurs dans  $I$  signifie que  $\mathbb{P}(X \in I) = 1$

**Remarques :**

- On peut étendre cette définition en remplaçant  $I$  par une réunion dénombrable d'intervalles.
- L'événement  $(X \in I)$  est presque sûr.
- Lorsqu'on observe la réalisation de cette expérience, on est presque sûr que  $X$  prend une valeur dans  $I$ .
- En pratique on pourra supprimer "(quasiment)"

**Proposition.** (Valeurs prises par  $X$  et densité de  $X$ )

Soit  $X$  une variable aléatoire à densité,  
 $X$  est (quasiment) à valeurs dans  $I$   
 si, et seulement si, il existe une densité de  $X$  nulle en dehors de  $I$ .

**Démonstration :** (admis)

**Proposition.** (Valeurs prises de  $X$  et fonction de répartition  $X$ )

Soit  $X$  une variable aléatoire à densité et  $a, b$  deux réels,  
 $X$  est (quasiment) à valeurs dans  $[a, b]$   
 si, et seulement si,  $\forall x \leq a, F(x) = 0$  et  $\forall x \geq b, F(x) = 1$

**Démonstration :** (admis)

**Remarque :** Certains s'autorisent  $X(\Omega) = I$ , mais cette notation peut amener des questions inutiles.

### 9.3.5 Indépendance

**Définition.** Indépendance de deux variables aléatoires à densité.

Soient  $X$  et  $Y$  deux variables aléatoires à densité sur le même espace probabilisé  $(\Omega, \mathcal{F}, P)$ ,  
 $X$  et  $Y$  sont indépendantes si, et seulement si :  
 quel que soit le couple  $(x, y)$  de réels,  $P((X \leq x) \cap (Y \leq y)) = P(X \leq x) \times P(Y \leq y)$

Quel que soit le couple  $(x, y)$  de réels,  $P((X \leq x) \cap (Y \leq y)) = F_X(x) \times F_Y(y)$

Quel que soit le couple  $(x, y)$  de réels,  $P((X \leq x) \cap (Y \leq y)) = \int_{-\infty}^x f_X(t) dt \times \int_{-\infty}^y f_Y(t) dt$

**Définition** Indépendance de  $n$  variables aléatoires à densité.

Soient  $n$  un entier supérieur ou égal à 2 et  $(X_k)_{1 \leq k \leq n}$  une liste de variables aléatoires à densité sur le même espace probabilisé  $(\Omega, \mathcal{F}, P)$ .  
 Les  $X_k$  sont (mutuellement) indépendantes si, et seulement si, :  
 quel que soit la liste  $(x_k)_{1 \leq k \leq n}$  de réels,  $P\left(\bigcap_{k=1}^n (X_k \leq x_k)\right) = \prod_{k=1}^n P(X_k \leq x_k)$

Revoir le cours donnant les propriétés communes à toutes les variables aléatoires . (lemme de coalition ...)

## 9.4 Exemples de recherche de la loi de $u(X)$ .

**Notation de  $u(X)$ .**

Soient  $X$  une variable à densité,  $I$  un intervalle et  $u$  une fonction usuelle définie sur  $I$ .  
 Si  $X$  est (quasiment) à valeurs dans  $I$ , alors on peut définir la variable aléatoire  $\omega \mapsto u(X(\omega))$  .  
 on note :  $u(X)$  cette variable aléatoire.

**Remarques :**

- Dans le cours sur les applications on aurait noté  $u \circ X$  cette variable aléatoire.
- $u(X)$  est bien une variable aléatoire si  $u$  est simple, mais il arrive que  $u(X)$  n'en soit pas une, mais cela dépasse le cadre de ce cours.

**Exemples.** Feuille Cours\_9\_3.

**Proposition.**

Soient  $X$  une variable à densité,  $I$  un intervalle et  $u$  une fonction définie sur  $I$ .  
 Si  $X$  est à valeurs dans  $I$ , alors  $u(X)$  est à valeurs dans  $u(I)$

**En effet :**  $(X \in I) \subset (u(X) \in u(I))$  donc  $1 = P(X \in I) \leq P(u(X) \in u(I)) \leq 1$

**Remarque :**  $u(I)$  est l'image directe de  $I$  par  $u$ , on l'obtient en dressant le tableau de variations de  $u$  sur  $I$ .

**Exemples.** Feuille Cours\_9\_3.

**En pratique :**

On connaît la densité de  $X$  et on répond à la question "Donner la loi de  $Y = u(X)$ ."

- 1 On détermine les valeurs prises par  $X$ . (On note  $I$ )
- 2 On détermine la fonction de répartition  $F_X$  de  $X$ .
- 3 On détermine  $J$  l'ensemble des valeurs prises par  $Y$ . (Tableau de variations de  $u$  sur  $I$ )

En posant  $u : x \mapsto x^2$

$x$	-1	0	1
$u$	1	↘ 0	↗ 1

- 4 Pour  $y \in J$ ,  
on exprime  $F_Y(y) = P(Y \leq y)$  avec  $F_X$   
puis  $F_Y(y)$  en fonction de  $y$ .
- 5 On donne les expressions définissant  $F_Y$  sur  $\mathbb{R}$  en entier.
- 6 On justifie que  $Y$  est une variable densité avec sa fonction de répartition.  
 $F_Y$  est de classe  $C^1$  sur  $\mathbb{R}$  sauf (éventuellement) en un nombre fini de réels et  $F_Y$  est continue sur  $\mathbb{R}$  en entier.
- 7 On déduit de  $F_Y$  une densité de  $Y$  en dérivant sur les intervalles ouverts où elle est dérivable.

## 9.5 Espérance. Propriétés.

**Définition.**

On considère  $X$  une variable aléatoire réelle d'un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .  
 On suppose  $X$  est une variable à densité, on note  $f$  une de ses densités.  
 Dire que  $X$  admet une espérance signifie que l'intégrale  $\int_{-\infty}^{+\infty} tf(t) dt$  est absolument convergente.  
 et lorsque cette condition est vérifiée on définit l'espérance de  $X$  par :  $E(X) = \int_{-\infty}^{+\infty} tf(t) dt$ .

**Remarques :**

- Pour  $f$  une densité de probabilité on a l'équivalence : (mais en pratique je ne m'en sers pas).  
 $\int_{-\infty}^{+\infty} tf(t) dt$  est absolument convergente si, et seulement si,  $\int_{-\infty}^{+\infty} tf(t) dt$  est convergente.
- Certaines variables à densité n'admettent pas d'espérance.
- Quand  $E(X) = 0$ , on dit que la variable est centrée.

**Exemples.** Feuille Cours\_9\_4

### 9.5.1 Théorème de transfert.

Soient  $X$  une variable aléatoire réelle de densité  $f$ , telle que  $X$  est à valeurs dans  $]a, b[$   
 $(-\infty \leq a < b \leq +\infty)$   
 et  $\varphi : ]a, b[ \rightarrow \mathbb{R}$  une fonction continue sauf éventuellement en un nombre fini de points.  
 $\varphi(X)$  admet une espérance si, et seulement si,  $\int_a^b \varphi(t)f(t)dt$  est absolument convergente.  
 et alors :  $E(\varphi(X)) = \int_a^b \varphi(t)f(t) dt$

#### Remarques :

- Ici on a bien besoin de la condition d'absolue convergence. (voir la remarque précédente sur l'espérance)
- On obtient l'espérance de  $\varphi(X)$  avec la densité de  $X$ .
- Toujours la même remarque sur la notion d'équivalence :

Si  $\int_a^b |\varphi(t)| \cdot f(t) dt$  ne converge pas alors  $\varphi(X)$  n'admet pas d'espérance.

Si  $\int_a^b |\varphi(t)| \cdot f(t) dt$  converge alors  $\varphi(X)$  admet une d'espérance.

Exemples. Feuille Cours\_9\_4

### 9.5.2 Variance d'une variable aléatoire.

#### Définition

##### Définition Rappel

Soit  $X$  une variable aléatoire admettant une espérance  $E(X)$ .  
 Lorsque la variable  $(X - E(X))^2$  admet une espérance on dit que  $X$  admet une variance et on définit sa variance par

$$V(X) = E((X - E(X))^2)$$

#### Remarques :

- Retenir que certaines variables aléatoires n'ont pas d'espérance.
- Dire qu'une variable aléatoire est **réduite** signifie que sa variance est égale à 1.
- En notant  $m = E(X)$ , cela donne pour  $X$  une variable aléatoire à densité : (théorème de transfert)

$$V(X) = \int_{-\infty}^{+\infty} (t - m)^2 f(t) dt \quad (\text{en cas de convergence})$$

- Si  $X$  est une variable à densité alors  $V(X) \neq 0$ .

Exemples. Feuille Cours\_9\_4

#### Formule de Kœnig-Huygens.

**Théorème :** (Formule de Kœnig-Huygens).

Soit  $X$  est une variable aléatoire réelle,  
 $X$  admet une variance si, et seulement si,  $X$  et  $X^2$  admettent une espérance.  
 et alors :  $V(X) = E(X^2) - (E(X))^2$  (en cas de convergence)

#### Remarque :

En notant  $m = E(X)$ , cela donne pour  $X$  une variable aléatoire à densité : (théorème de transfert)

$$V(X) = \int_{-\infty}^{+\infty} t^2 f(t) dt - m^2 \quad (\text{en cas de convergence})$$

Exemples. Feuille Cours\_9\_4

## 9.6 Lois à densité usuelles.

On considère  $X$  une variable aléatoire d'un espace probabilisé  $(\Omega; \mathcal{F}, \mathbb{P})$  quelconque.

### 9.6.1 Lois uniformes à densité.

**Définition :**

Soient  $a, b$  deux réels vérifiant  $a < b$

Dire que  $X$  suit une loi uniforme sur  $[a, b]$  signifie que la fonction  $\frac{1}{b-a} \mathbf{1}_{[a,b]}$  est une densité de  $X$ .

On note :  $X \hookrightarrow \mathcal{U}([a, b])$

La fonction  $t \mapsto \frac{1}{b-a} \mathbf{1}_{[a,b]}(t)$  est une densité de  $X$ .

**Proposition :** *Caractérisation avec la fonction de répartition.*

Soient  $X$  une variable aléatoire et  $a, b$  deux réels ( $a < b$ ) on note  $F$  la fonction de répartition de  $X$ ,

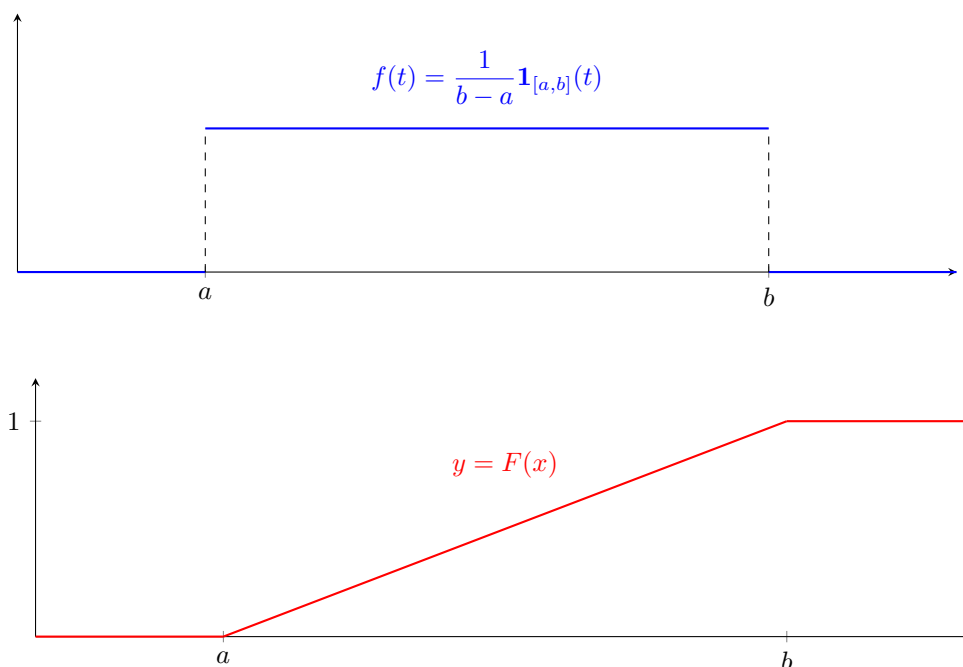
$X \hookrightarrow \mathcal{U}([a, b])$  si, et seulement si,  $F$  est la fonction définie par

$$\begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } b < x \end{cases} .$$

**En effet :** Il suffit de calculer  $\int_{-\infty}^x f(t) dt$  pour  $x$  dans les différents intervalles  $] -\infty; a[$ ,  $[a, b]$  et  $[b; +\infty[$ .

**Remarque :** on peut reconnaître une loi uniforme avec sa fonction de répartition.

*Représentation graphique.*



**Proposition.** (*Stabilité des lois uniformes par transformation affine  $t \mapsto mt + p$* )

Soient  $X$  une variable aléatoire,  $a, b, m$  et  $p$  quatre réels vérifiant  $a < b$  et  $m \neq 0$

- lorsque  $m > 0$ , Si  $X \hookrightarrow \mathcal{U}([a, b])$  alors  $mX + p \hookrightarrow \mathcal{U}([ma + p, mb + p])$ .
- lorsque  $m < 0$ , Si  $X \hookrightarrow \mathcal{U}([a, b])$  alors  $mX + p \hookrightarrow \mathcal{U}([mb + p, ma + p])$ .

**Démonstration.** (Faites au tableau)

**A retenir :** Si  $X$  suit une loi uniforme sur un segment alors  $mX + p$  aussi.

**Remarques.**

- Les réciproques sont vraies.

- $[a, b] \rightarrow [ma + p, mb + p]$  est une bijection et sa réciproque est  $[ma + p, mb + p] \rightarrow [a, b]$   
 $x \mapsto mx + p$   $x \mapsto \frac{x - p}{m}$

Cela autorise les raisonnements suivants :

①  $X \hookrightarrow \mathcal{U}([a, b])$  donc  $X - a \hookrightarrow \mathcal{U}([0, b - a])$  donc  $\frac{X - a}{b - a} \hookrightarrow \mathcal{U}([0, 1])$

②  $X \hookrightarrow \mathcal{U}([0, 1])$  donc  $(b - a)X \hookrightarrow \mathcal{U}([0, b - a])$  donc  $a + (b - a)X \hookrightarrow \mathcal{U}([a, b])$

**Exemples :**

Si on sait que  $X \hookrightarrow \mathcal{U}([0, 2])$  alors on peut en déduire que  $2X - 1 \hookrightarrow \mathcal{U}([-1, 3])$

Si on sait que  $2X + 3 \hookrightarrow \mathcal{U}([-1, 2])$  alors on peut en déduire que  $X \hookrightarrow \mathcal{U}\left(\left[-2, -\frac{1}{2}\right]\right)$

Si on sait que  $X \hookrightarrow \mathcal{U}([-1, 1])$  alors on peut en déduire que  $2 - X \hookrightarrow \mathcal{U}([1, 3])$

Si on sait que  $2 - 3X \hookrightarrow \mathcal{U}([0, 1])$  alors on peut en déduire que  $X \hookrightarrow \mathcal{U}\left(\left[\frac{1}{3}, \frac{2}{3}\right]\right)$

(Simulation des lois uniformes)

En Python, le module `random` peut être importé via `import random as rd`,  
`rd.random()` - - - - - Simule une réalisation d'une variable  $X \hookrightarrow \mathcal{U}([0, 1])$ .

Pour simuler une réalisation d'une variable  $X \hookrightarrow \mathcal{U}([a, b])$ , il suffit d'utiliser l'instruction

`a + (b-a) * rd.random()`

Pour simuler une réalisation d'une variable  $X \hookrightarrow \mathcal{U}\left(\left[m - \frac{h}{2}, m + \frac{h}{2}\right]\right)$ , on utilise :

`m + h/2*( -1 + 2*rd.random() )`

**Remarque :** En physique vous utilisez le module `numpy.random` et la fonction `uniform(a, b, N)`

**Proposition.**

$X \hookrightarrow \mathcal{U}([a, b])$  ( $a < b$ ), alors elle admet une espérance et une variance et :

$$E(X) = \frac{a + b}{2} \qquad V(X) = \frac{(b - a)^2}{12} \qquad \sigma(X) = \frac{b - a}{\sqrt{12}}$$

**Démonstrations.**

- (Existence de l'espérance)

$f$  est nulle en dehors de  $[a, b]$  donc  $\int_{-\infty}^{+\infty} |t|f(t) dt = \int_a^b \frac{|t|}{b - a} dt$

Cette dernière intégrale est celle d'une fonction continue sur un segment donc elle converge,

$X$  admet une espérance

- (Calcul de l'espérance)

$$E(X) = \int_{-\infty}^{+\infty} tf(t) dt$$

$$\begin{aligned}
&= \int_a^b \frac{t}{b-a} dt \\
&= \frac{1}{b-a} \left[ \frac{t^2}{2} \right]_a^b \\
&= \frac{b^2 - a^2}{2(b-a)}
\end{aligned}$$

$$E(X) = \frac{a+b}{2}$$

- (Existence du moment d'ordre 2)

$X^2$  admet une espérance si, et seulement si,  $\int_{-\infty}^{+\infty} t^2 f(t) dt$  est absolument convergente.

$$\int_{-\infty}^{+\infty} t^2 f(t) dt = \int_a^b \frac{t^2}{b-a} dt$$

Cette dernière intégrale est celle d'une fonction continue sur un segment donc elle converge,

- (Calcul de la variance)

Théorème de transfert

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{+\infty} t^2 f(t) dt \\
&= \int_a^b \frac{t^2}{b-a} dt \\
&= \frac{1}{b-a} \left[ \frac{t^3}{3} \right]_a^b \\
&= \frac{b^3 - a^3}{3(b-a)} \\
&= \frac{a^2 + ab + b^2}{3} \quad (\text{Formule de Bernoulli})
\end{aligned}$$

En conclusion :

$X$  possède une variance et (formule de Kœnig-Huygens)

$$\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \frac{a^2 + ab + b^2}{3} - \left( \frac{a+b}{2} \right)^2 \\
&= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} \\
&= \frac{a^2 - 2ab + b^2}{12}
\end{aligned}$$

$$V(X) = \frac{(a-b)^2}{12} \quad \text{et} \quad \sigma_X = \frac{b-a}{2\sqrt{3}}$$

### 9.6.2 Lois exponentielles.

**Définition :**

Soit  $\lambda$  un réel strictement positif, ( $\lambda \in \mathbb{R}_+^*$ )

Dire que  $X$  suit une loi exponentielle de paramètre  $\lambda$  signifie que :

la fonction  $t \mapsto \lambda e^{-\lambda t} \mathbb{1}_{[0; +\infty[}(t)$  est une densité de  $X$ .

On note :  $X \hookrightarrow \mathcal{E}(\lambda)$

**Proposition :**

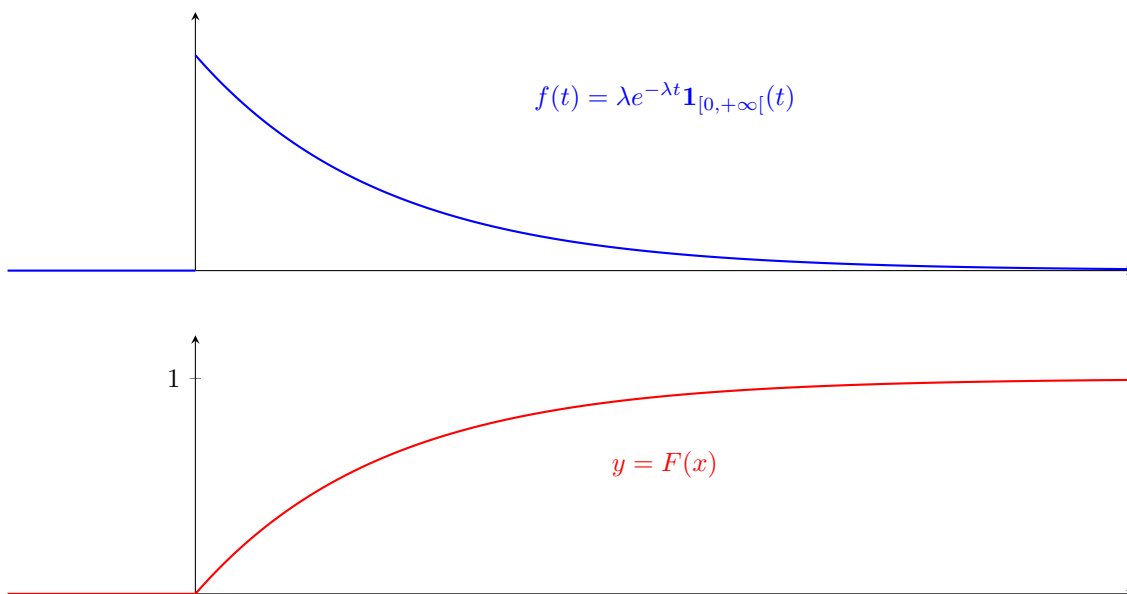
Pour tout  $\lambda \in \mathbb{R}_+^*$ ,

$X \leftrightarrow \mathcal{E}(\lambda)$  si, et seulement si, sa fonction de répartition est  $x \mapsto (1 - e^{-\lambda x}) \mathbb{1}_{[0; +\infty[}(x)$ .

**En effet :** Il suffit de calculer  $\int_{-\infty}^x f(t) dt$  pour  $x$  dans les intervalles  $] -\infty; 0[$  et  $[0; +\infty[$ .

**Remarque :** on peut reconnaître une loi exponentielle avec sa fonction de répartition.

Représentation graphique.



**Proposition.**

Pour tout  $\lambda \in \mathbb{R}_+^*$ , si  $X \leftrightarrow \mathcal{E}(\lambda)$  alors elle admet une espérance et une variance, et :

$$E(X) = \frac{1}{\lambda} \quad V(X) = \frac{1}{\lambda^2} \quad \sigma(X) = \frac{1}{\lambda}$$

**Démonstrations.**

• (Existence de l'espérance)

$f$  est nulle en dehors de  $[0; +\infty[$  donc  $\int_{-\infty}^{+\infty} |t|f(t) dt = \int_0^{+\infty} \lambda t e^{-\lambda t} dt$

Pour  $x > 0$ ,

$$\begin{aligned} \int_0^x \lambda t e^{-\lambda t} dt &= \left[ -t e^{-\lambda t} \right]_0^x - \int_0^x (-e^{-\lambda t}) dt && \text{Intégration par parties} \\ &= -x e^{-\lambda x} + \int_0^x e^{-\lambda t} dt \\ &= -x e^{-\lambda x} + \left[ -\frac{1}{\lambda} e^{-\lambda t} \right]_0^x \\ &= -x e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x} + \frac{1}{\lambda} \\ &\xrightarrow{x \rightarrow +\infty} \frac{1}{\lambda} && \text{Croissance comparée} \end{aligned}$$

ainsi  $\int_{-\infty}^{+\infty} t f(t) dt$  est absolument convergente donc

$X$  admet une espérance

• (Calcul de l'espérance)

$$E(X) = \int_{-\infty}^{+\infty} t f(t) dt$$

$$\begin{aligned}
&= \int_0^{+\infty} \lambda t e^{-\lambda t} dt \\
&= \frac{1}{\lambda}
\end{aligned}$$

$$E(X) = \frac{1}{\lambda}.$$

- (Existence du moment d'ordre 2)

$X^2$  admet une espérance si, et seulement si,  $\int_{-\infty}^{+\infty} t^2 f(t) dt$  est absolument convergente.

$$\int_{-\infty}^{+\infty} |t^2| f(t) dt = \int_0^{+\infty} \lambda t^2 e^{-\lambda t} dt$$

Pour  $x > 0$ ,

$$\begin{aligned}
\int_0^x \lambda t^2 e^{-\lambda t} dt &= \left[ t^2 (-e^{-\lambda t}) \right]_0^x - \int_0^x 2t (-e^{-\lambda t}) dt \\
&= -x^2 e^{-\lambda x} + \frac{2}{\lambda} \int_0^x t \lambda e^{-\lambda t} dt \\
&\xrightarrow{x \rightarrow +\infty} 0 + \frac{2}{\lambda} \times \underbrace{\frac{1}{\lambda}}_{E(X)}
\end{aligned}$$

$X^2$  admet une espérance

$$E(X^2) = \int_{-\infty}^{+\infty} t^2 f(t) dt = \int_0^{+\infty} \lambda t^2 e^{-\lambda t} dt = \frac{2}{\lambda^2} \quad \boxed{E(X^2) = \frac{2}{\lambda^2}}$$

En conclusion :

$X$  possède une variance et (formule de Kœnig-Huygens)

$$\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\
&= \frac{1}{\lambda^2}
\end{aligned}$$

$$\boxed{V(X) = \frac{1}{\lambda^2} \quad \text{et} \quad \sigma_X = \frac{1}{\lambda}}$$

(Simulation des lois exponentielles)

En Python, on importe `random` et `math` via `import random as rd` et `from math import *`  
 Pour simuler une réalisation d'une variable  $X \hookrightarrow \mathcal{E}(a)$ , ( $a > 0$ ) il suffit d'utiliser l'instruction  
`-1/a*log(1 - rd.random())` ou `-1/a*log(rd.random())`

**Démonstration :**

Soit  $U$  une variable aléatoire suivant la loi uniforme sur  $]0, 1[$ , on note  $X = -\frac{1}{a} \ln(1 - U)$

$U$  prend ses valeurs dans  $]0, 1[$  (car  $f$  est nulle en dehors de  $]0, 1[$ ) donc

$1 - U$  est à valeurs dans  $]0, 1[$  et  $\ln(1 - U)$  est à valeurs dans  $] -\infty; 0[$  et enfin comme  $a > 0$  :

$$X = -\frac{1}{a} \ln(1 - U) \text{ prend ses valeurs dans } ]0, +\infty[$$

Pour  $x \in ]0, +\infty[$

$$\begin{aligned}
 F_X(x) &= P(X \leq x) \\
 &= P\left(-\frac{1}{a} \ln(1-U) \leq x\right) \\
 &= P(\ln(1-U) \geq -ax) \quad \text{car } a > 0 \\
 &= P(1-U \geq \exp(-ax)) \quad \text{car } t \mapsto e^t \text{ est strictement croissante sur } \mathbb{R} \\
 &= P(U \leq 1 - \exp(-ax)) \\
 &= F_U(1 - \exp(-ax)) \\
 &= 1 - \exp(-ax) \quad \text{car } 1 - \exp(-ax) \in ]0, 1[ \text{ et } U \hookrightarrow \mathcal{U}(]0, 1[)
 \end{aligned}$$

$$F_X : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - \exp(-ax) & \text{si } 0 < x \end{cases}$$

**Proposition.** (*Absence de mémoire*)

$$\text{S'il existe } \lambda \in \mathbb{R}_+^* \text{ tel que } X \hookrightarrow \mathcal{E}(\lambda) \quad \text{alors} \quad \forall (t_1, t_2) \in \mathbb{R}_+^*, P_{X > t_1}(X > t_1 + t_2) = P(X > t_2)$$

**Démonstration.**

Soit  $(t_1, t_2) \in \mathbb{R}_+^*$ ,

$$\begin{aligned}
 P_{X > t_1}(X > t_1 + t_2) &= \frac{P((X > t_1) \cap (X > t_1 + t_2))}{P(X > t_1)} \\
 &= \frac{P(X > t_1 + t_2)}{P(X > t_1)} \quad (\text{car } (X > t_1 + t_2) \subset (X > t_1)) \\
 &= \frac{e^{-\lambda(t_1+t_2)}}{e^{-\lambda t_1}} \\
 &= e^{-\lambda t_2} = P(X > t_2)
 \end{aligned}$$

$$P_{X > t_1}(X > t_1 + t_2) = P(X > t_2)$$

La loi exponentielle est sans mémoire :

Si  $X$  est un temps d'attente et si à  $t_1$  on attend toujours alors le temps d'attente suit la même loi qu'au début.

**Remarques :**

❶ Attention à l'autre notation des probabilités conditionnelles :

$$\forall (t_1, t_2) \in \mathbb{R}_+^*, P(X > t_1 + t_2 | X > t_1) = P(X > t_2)$$

❷ Les lois exponentielles permettent de modéliser la durée de vie d'un phénomène sans mémoire, exemples classiques : composant électronique, désintégration d'un atome radioactif.

❸ Si on sait qu'à l'instant  $t_1$  le composant fonctionne toujours, sa durée de vie suit toujours la même loi qu'au début de sa vie.

En particulier :

Si on sait qu'à l'instant  $t_1$  le composant fonctionne toujours, son espérance de vie est toujours la même.

❹ On pourrait montrer que cette implication est une équivalence.

### 9.6.3 Lois normales.

#### Loi normale centrée réduite.

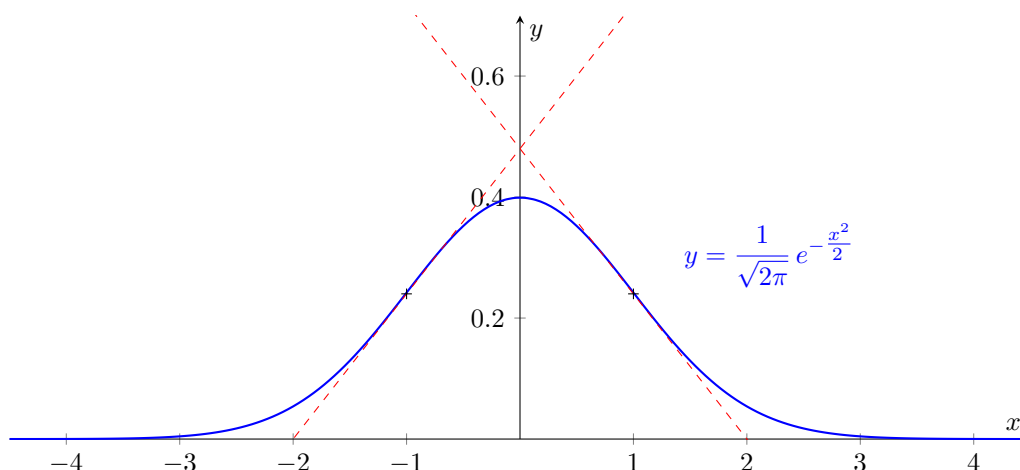
**Définition :**

Dire que  $X$  suit la loi normale centrée réduite signifie que :

$$\text{la fonction } t \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \text{ est une densité de } X.$$

On note :  $X \hookrightarrow \mathcal{N}(0, 1)$

**Remarque :** on a admis que  $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1$  pour les autres propriétés de cette densité Voir feuille Cours\_10



**Proposition.**

Si  $X \hookrightarrow \mathcal{N}(0, 1)$ , alors elle admet une espérance et une variance et :

$$E(X) = 0 \quad V(X) = 1 \quad \sigma(X) = 1$$

**Démonstrations**

• (Existence de l'espérance)

$$\int_{-\infty}^{+\infty} |t|f(t) dt = \int_{-\infty}^{+\infty} \frac{|t|}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

L'intégrande est pair, étudions :  $\int_0^{+\infty} \frac{|t|}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} t e^{-\frac{t^2}{2}} dt$

Pour  $x > 0$ ,

$$\begin{aligned} \int_0^x t e^{-\frac{t^2}{2}} dt &= \left[ -e^{-\frac{t^2}{2}} \right]_0^x \\ &= 1 - e^{-\frac{x^2}{2}} \\ &\xrightarrow{x \rightarrow +\infty} 1 \end{aligned}$$

ainsi  $\int_{-\infty}^{+\infty} t f(t) dt$  est absolument convergente donc

X admet une espérance

• (Calcul de l'espérance)

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} t f(t) dt \\ &= \int_{-\infty}^{+\infty} \frac{t}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\ &= 0 \quad \text{car l'intégrande est impair} \end{aligned}$$

$E(X) = 0.$

$X^2$  admet une espérance si, et seulement si,  $\int_{-\infty}^{+\infty} t^2 f(t) dt$  est absolument convergente.

si, et seulement si,  $\int_{-\infty}^{+\infty} \frac{t^2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$  est convergente.

si, et seulement si,  $\int_0^{+\infty} \frac{t^2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$  est convergente.

(l'intégrande est pair)

Pour  $x > 0$ ,

$$\begin{aligned} \int_0^x t^2 e^{-\frac{t^2}{2}} dt &= \left[ t(-e^{-\frac{t^2}{2}}) \right]_0^x - \int_0^x (-e^{-\frac{t^2}{2}}) dt \\ &= -xe^{-\frac{x^2}{2}} + \int_0^x e^{-\frac{t^2}{2}} dt \\ &\xrightarrow{x \rightarrow +\infty} 0 + \frac{1}{2} \times \sqrt{2\pi} \end{aligned}$$

$X^2$  admet une espérance

$$E(X^2) = \int_{-\infty}^{+\infty} t^2 f(t) dt = 2 \times \frac{1}{\sqrt{2\pi}} \times \int_0^{+\infty} t^2 e^{-\frac{t^2}{2}} dt = 1$$

$E(X^2) = 1$

En conclusion :

$X$  possède une variance et (formule de Kœnig-Huygens)

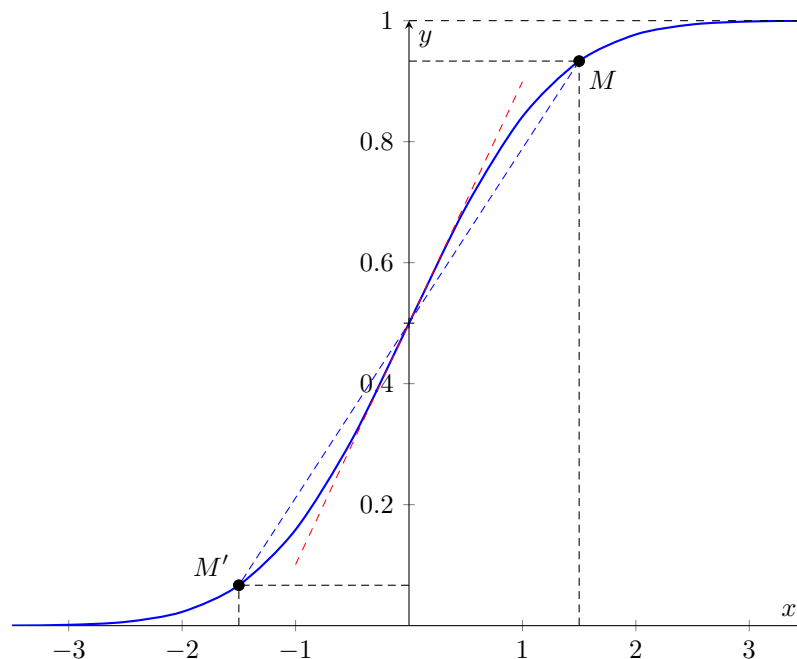
$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 \\ &= 1 - 0 \end{aligned}$$

$V(X) = 1$  et  $\sigma_X = 1$

### Fonction de répartition de la loi normale centrée réduite

On note dans ce chapitre  $\Phi$  la fonction de répartition de la loi centrée réduite.

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



**Remarque :** On ne sait pas exprimer  $\Phi$  à l'aide des fonctions usuelles.

Pour les calculs numériques on utilise une table de valeurs ou un outil informatique.

Voir la feuille Cours\_10\_3.

**Proposition.**

- ❶  $\Phi$  est dérivable sur  $\mathbb{R}$  et  $\forall x \in \mathbb{R}, \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- ❷  $\Phi$  réalise une bijection continue et strictement croissante de  $\mathbb{R}$  dans  $]0, 1[$ .
- ❸ Pour tout  $x \in \mathbb{R}, \Phi(-x) = 1 - \Phi(x)$ . En particulier,  $\Phi(0) = \frac{1}{2}$ .
- ❹ Pour tout  $a \in \mathbb{R}_+, P(|X| \leq a) = 2\Phi(a) - 1$

**Démonstration**

❶ Comme  $f : t \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$  est continue sur  $\mathbb{R}$  (en entier),  $\Phi$  est une primitive sur  $\mathbb{R}$  de  $f$ .

Ainsi  $\Phi$  est dérivable sur  $\mathbb{R}$  et  $\forall x \in \mathbb{R}, \Phi'(x) = f(x)$

❷ on a donc  $\forall x \in \mathbb{R}, \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} > 0$  et ainsi  $\Phi$  est strictement croissante

Donc  $\Phi$  est strictement croissante et continue sur l'intervalle  $\mathbb{R}, \lim_{-\infty} \Phi = 0$  et  $\lim_{+\infty} \Phi = +\infty$   
donc (Théorème de la bijection) :  $\Phi$  réalise une bijection de  $] -\infty; +\infty[$  sur  $]0, 1[$

❸  $\Phi(0) = \int_{-\infty}^0 f(t)dt$ ; or  $f$  est paire donc  $\int_{-\infty}^{+\infty} f(t)dt = 2 \int_{-\infty}^0 f(t)dt$  et comme  $\int_{-\infty}^{+\infty} f(t)dt = 1$  on obtient

$$\Phi(0) = \frac{1}{2}$$

On note  $g : x \mapsto \Phi(-x) + \Phi(x)$ ,

Cette fonction  $g$  est dérivable et  $\forall x \in \mathbb{R}, g'(x) = -f(-x) + f(x) = 0$  car  $f$  est paire.

La fonction  $g$  est donc une fonction constante sur  $\mathbb{R}$  et comme  $g(0) = 2\Phi(0)$  il vient que  $g$  est la fonction constante égale à 1.

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

❹ Soient  $X \hookrightarrow \mathcal{N}(0, 1)$  et  $a \in \mathbb{R}_+$ ,

$$\begin{aligned} P(|X| \leq a) &= P(-a \leq X \leq a) \\ &= \Phi(a) - \Phi(-a) \\ &= \Phi(a) - (1 - \Phi(a)) \\ &= 2\Phi(a) - 1 \end{aligned}$$

$$P(|X| \leq a) = 2\Phi(a) - 1$$

**Remarque :** La courbe représentative de  $\Phi$  est symétrique par rapport à  $A(0, \frac{1}{2})$ .

**En effet :** la relation précédente s'écrit :  $\forall x \in \mathbb{R}, \frac{\Phi(0-x) + \Phi(0+x)}{2} = \frac{1}{2}$

**Proposition.**

Pour toute variable aléatoire réelle  $X$ ,  $X \hookrightarrow \mathcal{N}(0, 1)$  si, et seulement si,  $-X \hookrightarrow \mathcal{N}(0, 1)$

**Démonstration.**

On suppose que  $X \hookrightarrow \mathcal{N}(0, 1)$  et on note a  $Y = -X$ ,

Pour  $x \in \mathbb{R}$ ,

$$\begin{aligned} F_Y(x) &= P(-X \leq x) \\ &= P(X \geq -x) \\ &= 1 - P(X \leq -x) \\ &= 1 - \Phi(-x) \\ &= \Phi(x) \quad \text{D'après la proposition précédente} \end{aligned}$$

$\Phi$  est la fonction de répartition de  $Y$  donc  $-X \hookrightarrow \mathcal{N}(0, 1)$

**Définition**

On appelle **fonction des quantiles** de la loi centrée réduite, la réciproque de la bijection  $\mathbb{R} \rightarrow ]0, 1[$   
 $x \mapsto \Phi(x)$

On note dans ce cours  $\Phi^{-1} : \alpha \mapsto u_\alpha$ .  $\Phi^{-1}(\alpha) = u_\alpha$

**Remarques :**

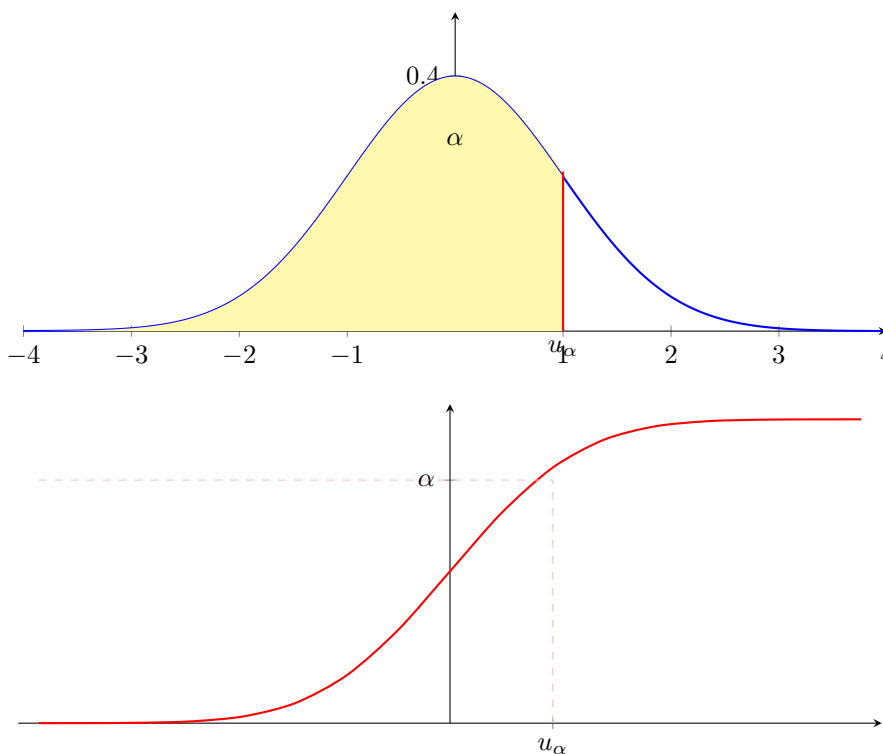
- On peut aussi noter cette fonction comme dans le cours sur les applications  $\Phi^{-1}$ .
- Pour  $\alpha \in ]0, 1[$ ,  $u_\alpha$  est le **quantile d'ordre  $\alpha$  de la loi  $\mathcal{N}(0, 1)$** .
- Pour  $\alpha \in ]0, 1[$ ,  $u_\alpha$  est l'unique réel  $x$  vérifiant  $\Phi(x) = \alpha$ .

**Propositions.**

- ❶  $\alpha \mapsto u_\alpha$  est une bijection continue et strictement croissante de  $]0, 1[$  dans  $\mathbb{R}$ .
- ❷ Pour tout  $\alpha \in ]0, 1[$ ,  $\Phi(u_\alpha) = \alpha$
- ❸ Pour tout  $\alpha \in ]0, 1[$ ,  $u_\alpha = -u_{1-\alpha}$
- ❹ Pour tout  $\alpha \in ]0, 1[$ ,  $\frac{1}{\sqrt{2\pi}} \int_{-u_{1-\frac{\alpha}{2}}}^{u_{1-\frac{\alpha}{2}}} e^{-\frac{t^2}{2}} dt = 1 - \alpha$

**Démonstrations.**

❶ et ❷ viennent des propriétés générales des fonctions réciproques.

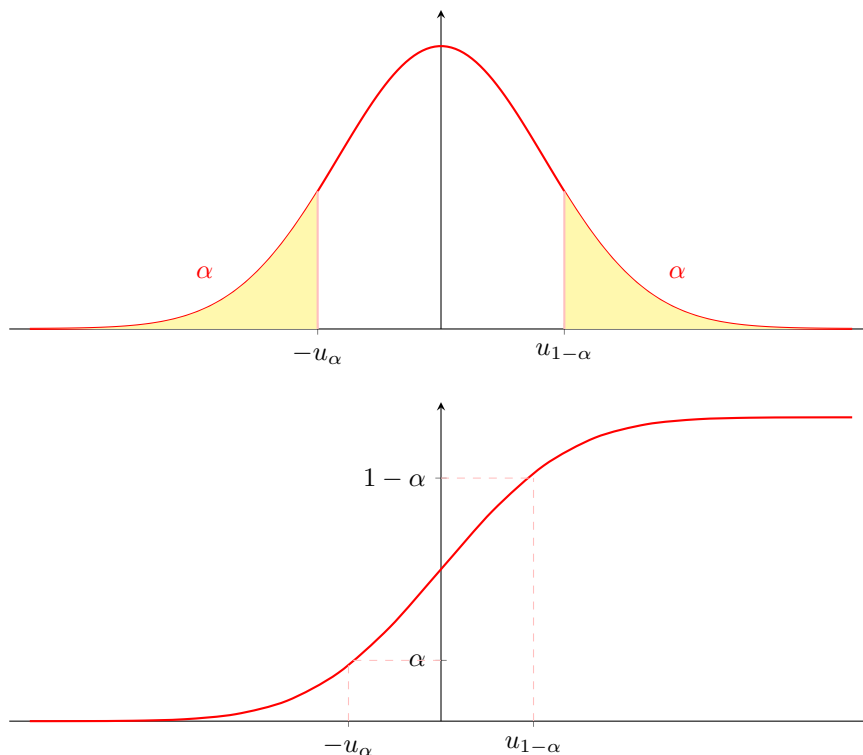


❸ Pour  $\alpha \in ]0, 1[$ ,

$$\begin{aligned}
 \Phi(-u_\alpha) &= 1 - \Phi(u_\alpha) && \text{d'après Ex4. 4)} \\
 &= 1 - \alpha \\
 &= \Phi(u_{1-\alpha})
 \end{aligned}$$

or  $\Phi$  est strictement croissante sur  $\mathbb{R}$  donc

$$-u_\alpha = u_{1-\alpha}$$

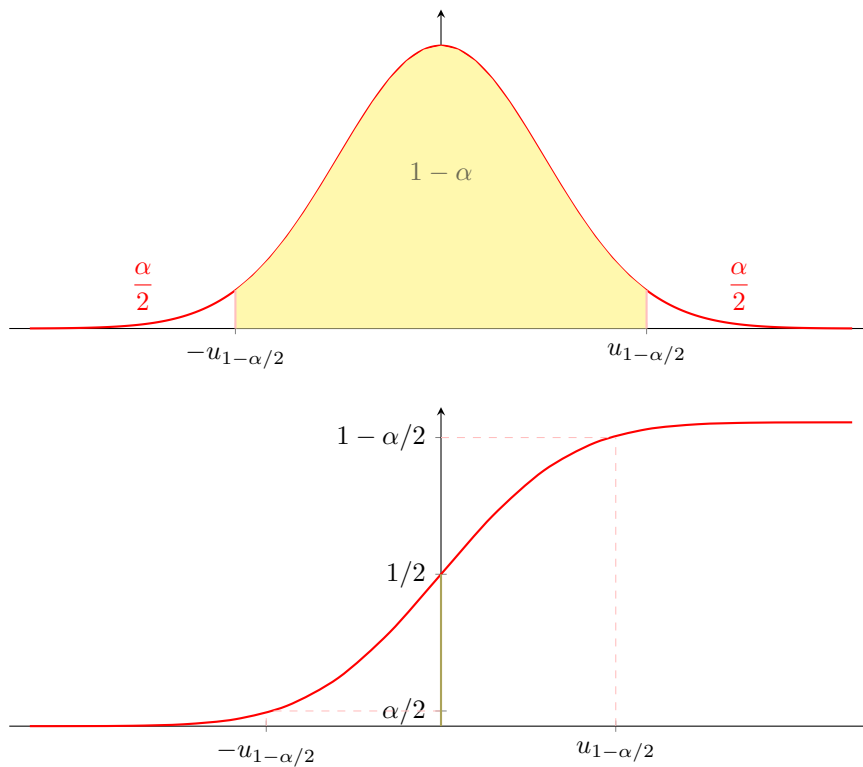


④

$$P(-u_{1-\frac{\alpha}{2}} \leq X \leq u_{1-\frac{\alpha}{2}}) = P(|X| \leq u_{1-\frac{\alpha}{2}}) = 2\Phi(u_{1-\frac{\alpha}{2}}) - 1 = 2\left(1 - \frac{\alpha}{2}\right) - 1 = 1 - \alpha$$

On en déduit directement :

$$\frac{1}{\sqrt{2\pi}} \int_{-u_{1-\frac{\alpha}{2}}}^{u_{1-\frac{\alpha}{2}}} e^{-\frac{t^2}{2}} dt = 1 - \alpha$$



**Proposition.**

Soit  $X$  une variable aléatoire,

Si  $X \hookrightarrow \mathcal{U}([0, 1])$  alors  $\Phi^{-1}(X) \hookrightarrow \mathcal{N}(0, 1)$

**Démonstration.**

$$\begin{aligned}
 F_Y(x) &= P(Y \leq x) \\
 &= P(\Phi^{-1}(X) \leq x) \\
 &= P(X \leq \Phi(x)) \quad \text{car } \Phi \text{ est une bijection strict. croissante} \\
 &= \Phi(x) \quad \text{car } X \hookrightarrow \mathcal{U}(]0, 1[) \text{ et } \Phi(x) \in ]0, 1[.
 \end{aligned}$$

$\Phi$  est la fonction de répartition de  $Y$  donc

$$\boxed{\text{Si } X \hookrightarrow \mathcal{U}(]0, 1[) \text{ alors } \Phi^{-1}(X) \hookrightarrow \mathcal{N}(0, 1)}$$

**Remarque :** cette propriété est plus générale, ce n'est pas spécifique à la loi normale, elle se nomme méthode d'inversion, elle nous avait déjà permis de simuler la loi exponentielle. Voir la feuille info\_21.

(Simulation de la loi normale centrée normale)

```

Option 1. (Dans le formulaire donné à l'oral de l'agro-véto.) En Python, le module random peut être
importé via import random as rd,
rd.gauss(0, 1) - - - - - Simule une réalisation d'une variable  $X \hookrightarrow \mathcal{N}(0, 1)$ .

Option 2. (En utilisant la proposition précédente .)
En Python, on peut importer  $\Phi^{-1}$  via from scipy.stats import norm

norm.ppf # percent point function

norm.ppf(rd.random()) - - - - - Simule une réalisation d'une variable  $X \hookrightarrow \mathcal{N}(0, 1)$ .
    
```

**Remarque :** En physique vous utilisez le module numpy.random et la fonction normal(0, 1, N)

Le programme suivant illustre l'utilisation du module scipy.stats.

```

from scipy.stats import norm # norm : normal continuous random variable

f = norm.pdf # probability density function
F = norm.cdf # cumulative distribution function
u = norm.ppf # percent point function

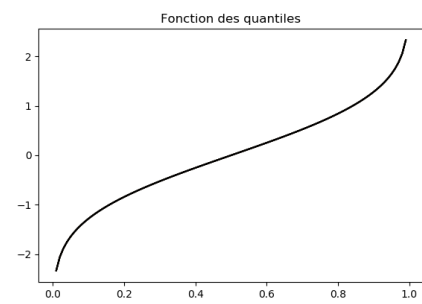
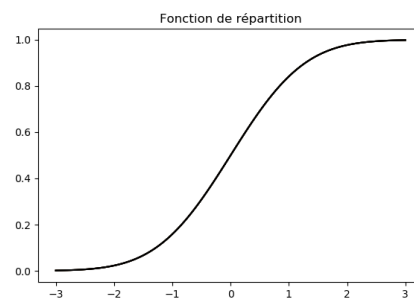
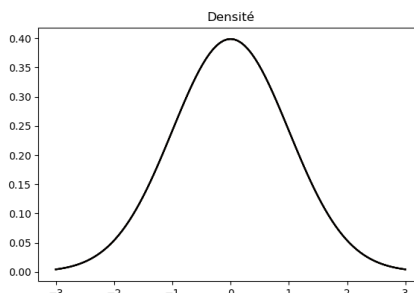
N= 100
a = -3
b = 3

x = [a+k*(b-a)/N for k in range(N+1)]
plt.subplot(221)
y = [ f(t) for t in x ]
plt.plot(x, y, 'k')
plt.title('Densité')

plt.subplot(223)
y = [ F(t) for t in x ]
plt.plot(x, y, 'k')
plt.title('Fonction de répartition')

x = [ 0 + k *1/N for k in range(1,N)]
plt.subplot(224)
y = [ u(t) for t in x ]
plt.plot(x, y, 'k')
plt.title('Fonction des quantiles')

plt.show()
    
```

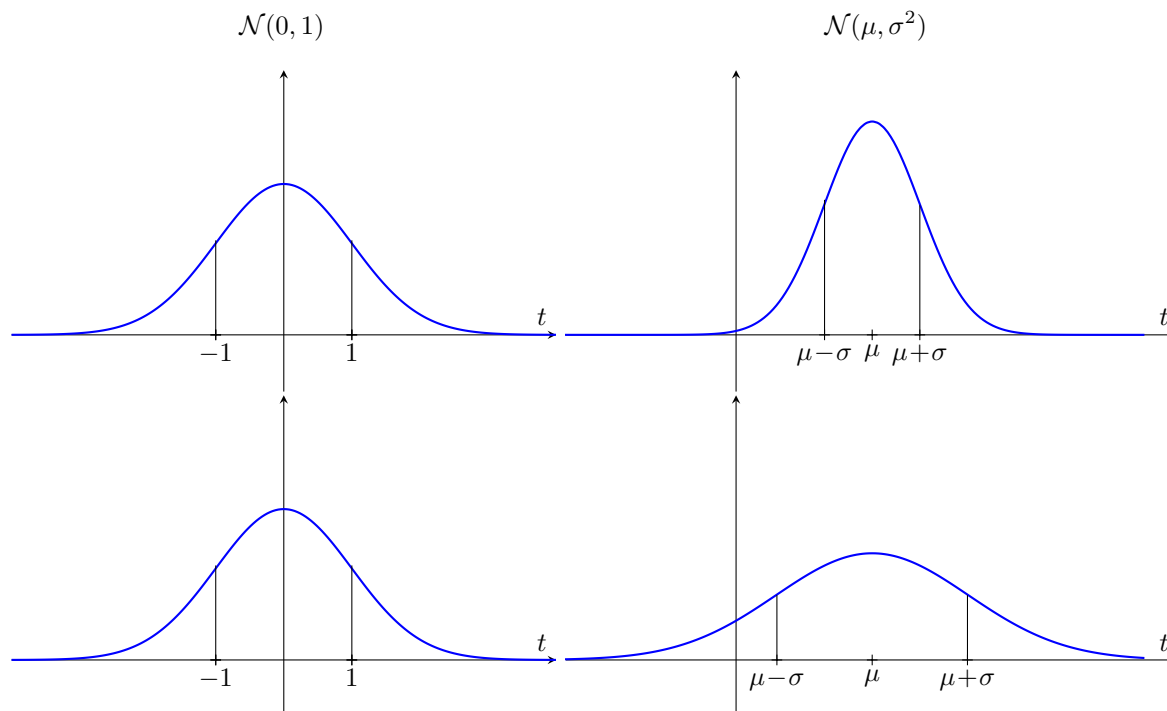


**Lois normales quelconques.**

**Définition.**

Soient  $X$  une variable aléatoire,  $\sigma$  un réel strictement positif et  $\mu$  un réel quelconque.  
 Dire que  $X$  suit la loi normale d'espérance  $\mu$  et d'écart-type  $\sigma$  signifie que  
 $X$  est une variable aléatoire à densité et la fonction  $t \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$  est une densité de  $X$ .

On note :  $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$  (la loi normale d'espérance  $\mu$  et de variance  $\sigma^2$ ).



**Cloche de Gauss**

**Remarques :**

- les points d'inflexion sur la courbe de la densité de  $\mathcal{N}(\mu, \sigma^2)$  sont aux points de coordonnées  $\mu - \sigma$  et  $\mu + \sigma$ .
- On dit aussi que  $X$  suit la loi gaussienne  $\mathcal{N}(\mu, \sigma^2)$  ou encore que  $X$  est une gaussienne.
- Attention certains auteurs notent :  $X \hookrightarrow \mathcal{N}(\mu, \sigma)$  (la loi normale d'espérance  $\mu$  et d'écart-type  $\sigma$ ).

*C'est le cas de fonctions Python*

**Proposition.**

Soient  $X$  une variable aléatoire,  $\sigma$  un réel strictement positif et  $\mu$  un réel quelconque.  
 $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$  si, et seulement si, sa fonction de répartition est  $F_X : x \mapsto \Phi\left(\frac{x - \mu}{\sigma}\right)$

**En effet :** En notant  $G : x \mapsto \Phi\left(\frac{x - \mu}{\sigma}\right)$  on a  $G$  et  $F_X$  de même dérivée, donc diffèrent d'une constante. et comme elles ont la même limite en  $+\infty$ , elles sont égales.

**Théorème.**

Soient  $X$  une variable aléatoire,  $\sigma$  un réel strictement positif et  $\mu$  un réel quelconque.  
 en notant :  $X^* = \frac{X - \mu}{\sigma}$  on a :  $X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \iff X^* \hookrightarrow \mathcal{N}(0, 1)$

**Démonstration.**

- $\Rightarrow$  On suppose que  $X^* = \frac{X - \mu}{\sigma}$  suit la loi normale centrée réduite,

$$\begin{aligned} \text{pour } x \in \mathbb{R}, \quad F_X(x) &= P(X \leq x) \\ &= P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) \\ &= P\left(X^* \leq \frac{x - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{x - \mu}{\sigma}\right) \end{aligned}$$

donc  $F_X$  est de classe  $C^1$  sur  $\mathbb{R}$  ( $X$  est bien à densité) et  $\forall x \in \mathbb{R}, \quad F'_X(x) = \frac{1}{\sigma} \Phi'\left(\frac{x - \mu}{\sigma}\right),$

$$\text{or } \forall t \in \mathbb{R}, \quad \Phi'(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

donc la fonction  $x \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$  est une densité de  $X$  et ainsi  $X \hookrightarrow \mathcal{N}(\mu, \sigma^2).$

- $\Leftarrow$  On suppose que la fonction  $t \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$  est une densité de  $X,$

Pour  $x \in \mathbb{R},$

$$\begin{aligned} F_{X^*}(x) &= P(X^* \leq x) \\ &= P\left(\frac{X - \mu}{\sigma} \leq x\right) \\ &= P(X \leq \sigma x + \mu) \\ &= \int_{-\infty}^{\sigma x + \mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \\ &= \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{v^2}{2}} \sigma dv \quad (v = \frac{t - \mu}{\sigma}) \\ &= \Phi(x) \end{aligned}$$

donc  $\Phi$  est la fonction de répartition de  $X^*$  et ainsi on a bien  $X^* \hookrightarrow \mathcal{N}(0, 1).$

**Proposition.**

Soient  $X$  une variable aléatoire,  $\sigma$  un réel strictement positif et  $\mu$  un réel quelconque.

$$\text{si } X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \quad \text{alors} \quad E(X) = \mu \quad V(X) = \sigma^2 \quad \sigma(X) = \sigma$$

**Démonstration.**

- $X = \sigma X^* + \mu$  donc (linéarité)  $E(X) = \sigma E(X^*) + \mu$  et sachant que  $E(X^*) = 0$  il vient :

$$\boxed{E(X) = \mu}$$

.

- $X = \sigma X^* + \mu$  donc ( $V(aX + b) = a^2 V(X)$ )  $V(X) = \sigma^2 V(X^*)$  et sachant que  $V(X^*) = 1$  il vient :

$$\boxed{V(X) = \sigma^2}$$

.

**Théorème.** (Stabilité de l'ensemble des lois normales par transformation affine).

Soient  $X$  une variable aléatoire,  $\sigma$  un réel strictement positif et  $\mu$  un réel quelconque.  
 Pour tout  $(a, b) \in \mathbb{R}^2$  avec  $(a \neq 0)$   

$$X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \quad \text{si, et seulement si,} \quad aX + b \hookrightarrow \mathcal{N}(a\mu + b, a^2\sigma^2)$$

**Démonstration :** On pose  $Y = aX + b$ .

$E(Y) = aE(X) + b$  et  $V(Y) = a^2V(X)$  donc  $E(Y) = a\mu + b$  et  $V(Y) = a^2\sigma^2$

$Y^* = \frac{Y - E(Y)}{\sqrt{V(Y)}}$  or  $E(Y) = aE(X) + b$  et  $V(Y) = a^2V(X)$  donc  $Y^* = \frac{a}{|a|} X^*$

• si  $a > 0$ ,

$$\begin{aligned} Y \hookrightarrow \mathcal{N}(a\mu + \sigma, a^2\sigma^2) &\iff Y^* \hookrightarrow \mathcal{N}(0, 1) \\ &\iff X^* \hookrightarrow \mathcal{N}(0, 1) && \text{en effet ici : } \frac{a}{|a|} = 1 \\ &\iff X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \end{aligned}$$

• si  $a < 0$ ,

$$\begin{aligned} Y \hookrightarrow \mathcal{N}(a\mu + \sigma, a^2\sigma^2) &\iff Y^* \hookrightarrow \mathcal{N}(0, 1) \\ &\iff -X^* \hookrightarrow \mathcal{N}(0, 1) && \text{en effet ici : } \frac{a}{|a|} = -1 \\ &\iff X^* \hookrightarrow \mathcal{N}(0, 1) && \text{vu en cours} \\ &\iff X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \end{aligned}$$

**Remarques :**

- Ce qu'il faut retenir : **Si  $X$  suit une loi normale alors  $aX + b$  suit une loi normale**  
 pour les paramètres il suffit d'utiliser :  $E(aX + b) = aE(X) + b$  et  $V(aX + b) = a^2V(X)$
- Si  $X \hookrightarrow \mathcal{N}(0, 1)$  alors  $\mu + \sigma X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$

(Simulation de la loi normale)

**Option 1.**

En Python, le module `random` peut être importé via `import random as rd`

`m + sqrt(v)*rd.gauss(0, 1)` - - - - - Simule une réalisation d'une variable  $X \hookrightarrow \mathcal{N}(m, v)$ .

**Option 2.**

En Python, on peut importer  $\Phi^{-1}$  via `from scipy.stats import norm`

`norm.ppf` # percent point function

`m+sqrt(v)*norm.ppf(rd.random())` - - - - - Simule une réalisation d'une variable  $X \hookrightarrow \mathcal{N}(m, v)$ .

**Remarques :**

- `rd.gauss(mu, sigma)` renvoie :  
 un flottant aléatoire suivant une distribution gaussienne d'espérance `mu` et de variance `sigma`<sup>2</sup>
- En physique certains utilisent le module `numpy.random` et la fonction `normal(a, b, N)`  
 Attention :  $a$  est l'espérance et  $b$  est l'écart-type.

## 9.7 Somme de variables aléatoires à densité indépendantes.

### 9.7.1 Produit de convolution.

**Définition.**

Soient  $f$  et  $g$  deux fonctions continues sur  $\mathbb{R}$  sauf peut-être en un nombre fini de points.  
On appelle produit de convolution de  $f$  et de  $g$  la fonction :

$$h : x \mapsto \int_{-\infty}^{+\infty} f(t)g(x-t) dt$$

**Théorème.**

Soient  $X$  et  $Y$  deux variables aléatoires de densité respective  $f$  et  $g$ ,

Si  $X$  et  $Y$  sont indépendantes alors :

- $X + Y$  est une variable à densité.
- $x \mapsto \int_{-\infty}^{+\infty} f(t)g(x-t) dt$  est une densité de  $X + Y$ .

**Remarques :**

- Les sujets devront vous rappeler ce théorème.
- $X + Y = Y + X$  donc

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

- Rappel : Pour deux variables discrètes indépendantes  $X$  et  $Y$ , en notant  $Z = X + Y$  :

$$\text{Pour tout } z \in Z(\Omega), \quad P(Z = z) = \sum_{x \in X(\Omega)} P(X = x)P(Y = z - x)$$

*formule des probabilités totales et indépendance*

**Exemples :** (Voir la feuille Cours\_10\_2)

*Revoir la somme de deux lois exponentielles de même paramètre ou de deux uniformes sur  $[0, 1]$ .*

*Attention à l'usage des fonctions indicatrices dans ce contexte.*

### 9.7.2 Somme de variables aléatoires normales indépendantes.

**Théorème.** (Stabilité de l'ensemble des lois normales par combinaison linéaire)

Soient  $X_1$  et  $X_2$  deux variables aléatoires,  $\sigma_1, \sigma_2$  deux réels strictement positifs et  $\mu_1, \mu_2$  deux réels quelconques.

Si  $X_1 \hookrightarrow \mathcal{N}(\mu_1, \sigma_1^2)$  et  $X_2 \hookrightarrow \mathcal{N}(\mu_2, \sigma_2^2)$  et si  $X_1$  et  $X_2$  sont indépendantes

❶ alors  $X_1 + X_2 \hookrightarrow \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Si de plus  $\alpha_1, \alpha_2$  sont deux réels (tels que  $(\alpha_1, \alpha_2) \neq (0, 0)$ )

❷ alors  $\alpha_1 X_1 + \alpha_2 X_2 \hookrightarrow \mathcal{N}(\alpha_1 \mu_1 + \alpha_2 \mu_2, \alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2)$

**Démonstration :** Voir l'Ex\_7 feuille Cours\_10\_2, non fait en classe.

*Vous pouvez aussi regarder le sujet ENS 2025.*

**Généralisation.**

Soient  $n \in \mathbb{N}^*$ ,  $(X_k)_{1 \leq k \leq n}$  des variables aléatoires,  $(\sigma_k)_{1 \leq k \leq n}$  des réels strictement positifs,  $(\mu_k)_{1 \leq k \leq n}$  et  $(\alpha_k)_{1 \leq k \leq n}$  des réels quelconques tels que les  $\alpha_k$  sont tous nuls.

Si pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $X_k \hookrightarrow \mathcal{N}(\mu_k, \sigma_k^2)$  et si les  $X_k$  sont (*mutuellement*) indépendantes

alors

$$\sum_{k=1}^n \alpha_k X_k \hookrightarrow \mathcal{N} \left( \sum_{k=1}^n \alpha_k \mu_k, \sum_{k=1}^n \alpha_k^2 \sigma_k^2 \right)$$

**Démonstration.** (*Récurrence + lemme de coalition*)

**Corollaire.**

Toute combinaison linéaire (*non nulle*) de gaussiennes mutuellement indépendantes est une gaussienne.

**En pratique :** *Il suffit de connaître ce corollaire.*

Avec pour  $k \in \llbracket 1, n \rrbracket$ ,  $X_k \hookrightarrow \mathcal{N}(\mu_k, \sigma_k^2)$  et  $Y = \sum_{k=1}^n \alpha_k X_k$

❶ Les  $X_k$  sont des gaussiennes mutuellement indépendantes donc  $Y$  suit une loi normale.

❷ Par linéarité de l'espérance on obtient  $E(Y) = \sum_{k=1}^n \alpha_k E(X_k)$ , on note  $\mu$  ce réel.

❸ Les variables  $X_k$  sont mutuellement indépendantes donc  $V(Y) = \sum_{k=1}^n V(\alpha_k X_k)$

et comme  $V(aX + b) = a^2 V(X)$  il vient :  $V(Y) = \sum_{k=1}^n \alpha_k^2 V(X_k)$ , on note  $\sigma^2$  ce réel.

On peut alors conclure :  $\boxed{Y \hookrightarrow \mathcal{N}(\mu; \sigma^2)}$  où  $\mu = \sum_{k=1}^n \alpha_k \mu_k$  et  $\sigma^2 = \sum_{k=1}^n \alpha_k^2 \sigma_k^2$

**Exemples :** *Exercices 1 et 2 de la Feuille Exo\_19.*

# Statistique.

## Plan du chapitre

10.1	Statistique univariée. . . . .	<b>141</b>
10.1.1	Les résultats d'une expérience. . . . .	141
10.1.2	Des nombres caractéristiques. . . . .	142
10.2	Statistique bivariée. . . . .	<b>144</b>
10.2.1	Nuage de points. . . . .	144
10.2.2	Covariance et coefficient de corrélation linéaire. . . . .	144
10.2.3	Droite de régression linéaire . . . . .	145
10.3	Loi faible des grands nombres. . . . .	<b>145</b>
10.3.1	Inégalité de Markov. . . . .	145
10.3.2	Inégalité de Bienaymé-Tchebychev. . . . .	146
10.3.3	Loi faible des grands nombres. . . . .	147
10.4	Echantillonnage. Estimation. . . . .	<b>149</b>
10.4.1	Echantillon . . . . .	149
10.4.2	Base de la statistique inférentielle. . . . .	149
10.4.3	Estimateur. . . . .	150
10.4.4	Moyenne empirique. . . . .	150
10.4.5	Ecart-type empirique. . . . .	151
10.5	Convergence en loi . . . . .	<b>152</b>
10.6	Théorème central limite . . . . .	<b>153</b>
10.6.1	Première forme. . . . .	153
10.6.2	Seconde forme . . . . .	155
10.7	Intervalle de confiance de la moyenne. . . . .	<b>156</b>
10.8	Test de conformité à la moyenne. . . . .	<b>158</b>
10.8.1	Pour une espérance. . . . .	158
10.8.2	Sur une proportion. . . . .	159

## 10.1 Statistique univariée.

### 10.1.1 Les résultats d'une expérience.

Les résultats d'une expérience peuvent être qualitatives ou quantitatives.

Nous nous intéresserons ici aux résultats numériques. (souvent le résultat d'une mesure).

On note :  $(x_1, x_2, \dots, x_n)$  la liste des données ( $n$ -uplets). Chaque  $x_i$  est un nombre réel.

On commence souvent par rassembler les résultats identiques pour présenter les données sous la forme :

valeurs	$\tilde{x}_1$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_{p-1}$	$\tilde{x}_p$
effectifs	$e_1$	$e_2$	$\cdots$	$e_{p-1}$	$e_p$

Les valeurs distinctes sont classées dans l'ordre croissant.  $\tilde{x}_1 < \tilde{x}_2 < \dots < \tilde{x}_{p-1} < \tilde{x}_p$

L'effectif  $e_i$  est le nombre de résultats égaux à  $\tilde{x}_i$ ,  $n = \sum_{i=1}^p e_i$  est l'effectif total.

Souvent on préfère souvent utiliser les fréquences plutôt que les effectifs.

valeurs	$\tilde{x}_1$	$\tilde{x}_2$	$\dots$	$\tilde{x}_{p-1}$	$\tilde{x}_p$
fréquences	$f_1$	$f_2$	$\dots$	$f_{p-1}$	$f_p$

La fréquence  $f_i = \frac{e_i}{n}$  est la proportion de données dont la valeur est égale  $\tilde{x}_i$ .      Remarque :  $\sum_{i=1}^p f_i = 1$

On présente aussi les résultats avec les effectifs cumulés.

L'effectif cumulé en un réel  $x$  est égal au nombre de données dont la valeur est inférieure ou égale à  $x$ .

valeurs	$\tilde{x}_1$	$\tilde{x}_2$	$\dots$	$\tilde{x}_{p-1}$	$\tilde{x}_p$
effectifs cumulés	$c_1$	$c_2$	$\dots$	$c_{p-1}$	$c_p$

pour  $i$  compris entre 1 et  $p$ ,  $c_i = \sum_{k=1}^i e_k$

On utilise aussi les fréquences cumulées :

valeurs	$\tilde{x}_1$	$\tilde{x}_2$	$\dots$	$\tilde{x}_{p-1}$	$\tilde{x}_p$
fréquences cumulées	$F_1$	$F_2$	$\dots$	$F_{p-1}$	$F_p$

pour  $i$  compris entre 1 et  $p$ ,  $F_i = \sum_{k=1}^i f_k$

c'est la proportion de données dont la valeur est inférieure ou égale à  $\tilde{x}_i$  et  $F_p = 1$ .

### 10.1.2 Des nombres caractéristiques.

#### Caractéristiques de position.

**Moyenne :**

$$\bar{x} = \frac{1}{n} \left( \sum_{k=1}^n x_k \right) = \frac{1}{\sum_{k=1}^p e_k} \left( \sum_{k=1}^p e_k \tilde{x}_k \right) = \sum_{k=1}^p f_k \tilde{x}_k$$

**Médiane :** (*Définition*)

On suppose ici que les données brutes sont triées :  $x_1 \leq x_2 \leq \dots \leq x_{n-1} \leq x_n$ .  
 si  $n$  est impair alors on choisit pour médiane le réel  $x_{\frac{n+1}{2}}$   
 si  $n$  est pair alors on choisit pour médiane le réel  $\frac{1}{2} (x_{\frac{n}{2}} + x_{\frac{n}{2}+1})$

Si on note  $\text{med}(x)$  la médiane on a :

$$\text{card}\{ i \mid x_i < \text{med}(x) \} = \text{card}\{ i \mid x_i > \text{med}(x) \}$$

**Interprétation :** La médiane est un réel qui partage la série en deux parties de même effectifs. (*Il n'y a pas unicité de ce paramètre*)

Si on change la numérotation comme avec Python :  $x_0 \leq x_1 \leq \dots \leq x_{n-2} \leq x_{n-1}$

si  $n$  est impair alors on choisit pour médiane le réel  $x_{\frac{n-1}{2}}$   
 si  $n$  est pair alors on choisit pour médiane le réel  $\frac{1}{2}(x_{\frac{n}{2}-1} + x_{\frac{n}{2}})$

**Premier et troisième quartiles.** (notées  $Q_1$  et  $Q_3$ )

Le **premier quartile** est l'unique valeur  $x_i$  où  $i$  est l'unique  $i$  de  $\llbracket 1, p \rrbracket$  vérifiant :

$$\sum_{k=1}^{i-1} e_k < \frac{n}{4} \quad \sum_{k=1}^i e_k \geq \frac{n}{4}$$

Le **troisième quartile** est l'unique valeur  $x_i$  où  $i$  est l'unique  $i$  de  $\llbracket 1, p \rrbracket$  vérifiant :

$$\sum_{k=1}^{i-1} e_k < \frac{3n}{4} \quad \sum_{k=1}^i e_k \geq \frac{3n}{4}$$

**Interprétation :** Les quartiles partagent en 4 parts *approximativement* égales.

**Premier et neuvième déciles** (notées  $D_1$  et  $D_9$ ) :

Le **premier décile** est l'unique valeur  $x_i$  où  $i$  est l'unique  $i$  de  $\llbracket 1, p \rrbracket$  vérifiant :

$$\sum_{k=1}^{i-1} e_k < \frac{n}{10} \quad \sum_{k=1}^i e_k \geq \frac{n}{10}$$

Le **neuvième décile** est l'unique valeur  $x_i$  où  $i$  est l'unique  $i$  de  $\llbracket 1, p \rrbracket$  vérifiant :

$$\sum_{k=1}^{i-1} e_k < \frac{9n}{10} \quad \sum_{k=1}^i e_k \geq \frac{9n}{10}$$

**Interprétation :** Les déciles partagent en 10 parts *approximativement* égales.

**Remarques :**

- L'ensemble de ses valeurs se présente souvent sous la forme d'un **boxplot** ou encore **une boîte à moustaches**.
- On rencontre une autre définition des fractiles : les quartiles sont alors les quatre intervalles séparant la population en quatre (Voir sujet 0 de Biologie).

**Modes ou valeurs modales :**

L'ensemble des valeurs des valeurs modales :

$$\{ \tilde{x}_i \mid 1 \leq i \leq p \text{ et } e_i = \max_{1 \leq k \leq p} (e_k) \}$$

**Interprétation :**

- C'est un  $\tilde{x}_i$  correspondant à la plus grande valeur des  $e_i$ . (Il n'y a pas unicité de ce paramètre).
- C'est une valeur qui apparaît le plus souvent.

**Caractéristiques de dispersion.**

L'**écart-type** est donné par chacune des formules :

$$\sigma_x = \sqrt{\frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2} = \sqrt{\sum_{k=1}^p f_k (\tilde{x}_k - \bar{x})^2} = \sqrt{\frac{1}{\sum_{k=1}^p e_k} \left( \sum_{k=1}^p e_k (\tilde{x}_k - \bar{x})^2 \right)}$$

**Remarques :**

- On a toujours  $\sigma_x \geq 0$ .
- Si les  $x_i$  ont une unité alors  $\sigma_x$  a la même unité.
- $\sigma_x = 0$  si, et seulement si,  $(x_1, \dots, x_n) \in \text{vect} < (1, \dots, 1) >$

On appelle **variance** (empirique) le carré de l'écart-type :

$$\sigma_x^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 = \sum_{k=1}^p f_k (\tilde{x}_k - \bar{x})^2 = \frac{1}{\sum_{k=1}^p e_k} \left( \sum_{k=1}^p e_k (\tilde{x}_k - \bar{x})^2 \right)$$

**Remarque :**

Dans la suite du cours nous allons définir une variable aléatoire que l'on appellera aussi variance empirique.

## 10.2 Statistique bivariée.

### 10.2.1 Nuage de points.

Au cours d'une expérience on recueille des valeurs de deux caractères :

$$x_1, x_2, \dots, x_n \quad \text{et} \quad y_1, y_2, \dots, y_n$$

Lorsqu'on étudie la relation éventuelle entre les deux caractères on commence par représenter tous les points de coordonnées  $(x_i, y_i)$ . On obtient alors un nuage de points.

On place le **point moyen** du nuage : le point de coordonnées  $(\bar{x}, \bar{y})$ .

On peut organiser ces données dans le tableau suivant :

valeurs du premier caractère	$\tilde{x}_1$	$\tilde{x}_2$	$\dots$	$\tilde{x}_{p-1}$	$\tilde{x}_p$
valeurs du deuxième caractère	$\tilde{y}_1$	$\tilde{y}_2$	$\dots$	$\tilde{y}_{p-1}$	$\tilde{y}_p$
effectifs	$e_1$	$e_2$	$\dots$	$e_{p-1}$	$e_p$

**Remarque :**  $n = \sum_{k=1}^p e_k$

### 10.2.2 Covariance et coefficient de corrélation linéaire.

**Définition :**

On appelle **covariance** des séries  $x$  et  $y$  le réel :

$$\text{cov}(x, y) = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y}) \quad \text{ou encore} \quad \text{cov}(x, y) = \frac{1}{n} \sum_{k=1}^p e_k (\tilde{x}_k - \bar{x})(\tilde{y}_k - \bar{y})$$

**Proposition :**

$$\text{cov}(x, y) = \frac{1}{n} \left( \sum_{k=1}^n x_k y_k \right) - \bar{x} \bar{y} = \frac{1}{n} \left( \sum_{k=1}^p e_k \tilde{x}_k \tilde{y}_k \right) - \bar{x} \bar{y}$$

**Démonstration :**

$$\begin{aligned} \text{cov}(x, y) &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y}) \\ &= \frac{1}{n} \sum_{k=1}^n (x_k y_k - x_k \bar{y} - \bar{x} y_k + \bar{x} \bar{y}) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n} \left( \sum_{k=1}^n x_k y_k \right) - \frac{1}{n} \bar{y} \left( \sum_{k=1}^n x_k \right) - \frac{1}{n} \bar{x} \left( \sum_{k=1}^n y_k \right) + \bar{x} \bar{y} \\
&= \frac{1}{n} \left( \sum_{k=1}^n x_k y_k \right) - \bar{x} \bar{y} - \bar{x} \bar{y} + \bar{x} \bar{y}
\end{aligned}$$

**Définition**

On appelle **coefficient de corrélation** (*linéaire*) le nombre :

$$r_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$

**Remarques :**

- Ce coefficient est sans unité.
- Savoir calculer cette valeur avec votre calculatrice, un tableur et en faisant une fonction Python.

**Propositions :**

$$\textcircled{1} \quad -1 \leq r_{xy} \leq 1$$

$$\textcircled{2} \quad r_{xy} = 1 \text{ ou } -1 \quad \text{si, et seulement si,} \quad \text{les points } M_1, M_2, \dots, M_n \text{ sont alignés.}$$

**10.2.3 Droite de régression linéaire****Définition :**

La **droite de régression linéaire de  $y$  en  $x$**  est la droite d'équation  $y = ax + b$  où :

$$a = \frac{\text{cov}(x, y)}{\sigma_x^2} \quad \text{et} \quad b = \bar{y} - a\bar{x}$$

**Démonstration :** (*Voir Annexe C*)

**Savoir faire :** calculer  $a$  et  $b$  avec une calculatrice, un tableur **et** en faisant une fonction Python.

En donnant cette droite on dit que réalise un ajustement affine selon *la méthode des moindres carrés*.

- (*Il existe une deuxième droite de régression linéaire :*)

On définit la droite de **régression linéaire de  $x$  en  $y$**  qui est la droite d'équation  $x = \alpha y + \beta$  où :

$$\alpha = \frac{\text{cov}(x, y)}{\sigma_y^2} \quad \text{et} \quad \beta = \bar{x} - \alpha \bar{y}$$

- Ces deux droites de régression passent par le point milieu.

**10.3 Loi faible des grands nombres.****10.3.1 Inégalité de Markov.****Théorème :**

Soient  $X$  une variable aléatoire réelle et  $a$  un nombre réel,

$$\text{Si} \quad \begin{cases} X \text{ admet une espérance} \\ X \text{ est à valeurs positives} \\ a > 0 \end{cases} \quad \text{alors} \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

**Démonstration :** On pose  $Y = a \mathbb{1}_{X \geq a}$

•  $\mathbb{1}_{X \geq a}$  suit la loi de Bernoulli de paramètre  $\mathbb{P}(X \geq a)$  et ainsi son espérance est  $\mathbb{E}(\mathbb{1}_{X \geq a}) = \mathbb{P}(X \geq a)$ .

• On a :  $Y \leq X$

En effet pour  $\omega \in \Omega$ ,

si  $X(\omega) \geq a$  alors  $Y(\omega) = a$  donc on a bien  $Y(\omega) \leq X(\omega)$

si  $X(\omega) < a$  alors  $Y(\omega) = 0$  et comme  $X$  est à valeurs positives ou nulles, ici aussi on a bien  $Y(\omega) \leq X(\omega)$

•  $Y \leq X$  et ces deux variables admettent une espérance donc  $\mathbb{E}(Y) \leq \mathbb{E}(X)$

or (linéarité)  $\mathbb{E}(Y) = a \mathbb{E}(\mathbb{1}_{X \geq a})$  ou encore  $\mathbb{E}(Y) = a \mathbb{P}(X \geq a)$

et comme  $a > 0$  on peut en conclure que :  $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$  ■

**Interprétations :**

• Quand  $X$  est à valeurs positives, il est peu probable que  $X$  prenne des valeurs grandes devant à  $\mathbb{E}(X)$ .

• Quand  $X$  est à valeurs positives, la probabilité que  $X$  dépasse largement son espérance est faible.

• Quand  $X$  est à valeurs positives, pour tout  $k > 0$ ,  $\mathbb{P}(X \geq k \mathbb{E}(X)) \leq \frac{1}{k}$

En particulier :  $\mathbb{P}(X \geq 2 \mathbb{E}(X)) \leq \frac{1}{2}$ ,  $\mathbb{P}(X \geq 3 \mathbb{E}(X)) \leq \frac{1}{3}$ ,  $\mathbb{P}(X \geq 10 \mathbb{E}(X)) \leq \frac{1}{10}$

**Exemples :**

• Si  $X$  est à valeurs positives et  $\mathbb{E}(X) = 2$  alors  $\mathbb{P}(X \geq 10) \leq \frac{1}{5}$

• Si  $X$  est un temps d'attente de moyenne 10 minutes alors  $\mathbb{P}(X \geq 25) \leq \frac{10}{25} = \frac{2}{5}$

En répétant cette expérience un grand nombre de fois,

la proportion d'attentes dépassant 25 minutes sera proche de  $\mathbb{P}(X \geq 25)$  et donc inférieure à  $\frac{2}{5}$ .

**Remarques :**

Ce théorème donne des informations sur  $X$  sans connaître la loi de  $X$ , seul  $\mathbb{E}(X)$  intervient.

Attention à bien rappeler les conditions d'application de ce théorème<sup>1</sup> :

$a > 0$ ,  $X$  à valeurs dans  $\mathbb{R}_+$  et  $X$  admet une espérance.

### 10.3.2 Inégalité de Bienaymé-Tchebychev.

**Théorème :**

Soit  $X$  une variable aléatoire réelle,

Si  $X$  admet une variance alors  $\forall t > 0$ ,  $\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{\mathbb{V}(X)}{t^2}$ .

**Démonstration :**

$X$  admet une variance donc  $Y = (X - \mathbb{E}(X))^2$  admet une espérance et  $Y = (X - \mathbb{E}(X))^2$  est à valeurs positives ou nulles. La variable  $Y = (X - \mathbb{E}(X))^2$  vérifie les conditions d'utilisation de l'inégalité de Markov.

On en déduit avec  $a = t^2 > 0$  :

$$\mathbb{P}\left((X - \mathbb{E}(X))^2 \geq t^2\right) \leq \frac{\mathbb{E}(X - \mathbb{E}(X))^2}{t^2}$$

or  $((X - \mathbb{E}(X))^2 \geq t^2) = (|X - \mathbb{E}(X)| \geq t)$  et  $\mathbb{E}(X - \mathbb{E}(X))^2 = \mathbb{V}(X)$  donc

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq t\right) \leq \frac{\mathbb{V}(X)}{t^2}$$

1. Comme pour tous les théorèmes du cours

**Interprétations :**

- La variable aléatoire  $|X - \mathbb{E}(X)|$  est l'écart de  $X$  à l'espérance  $\mathbb{E}(X)$ .  
On majore dans ce théorème la probabilité que  $X$  s'écarte de  $\mathbb{E}(X)$  de plus de  $t$ .
- Plus  $t$  est grand et plus la probabilité est faible.
- $V(X)$  est un critère de dispersion : pour  $t$  fixé, plus  $V(X)$  est faible et plus le majorant fourni ici est faible.

**Exemples :**

- Si  $\mathbb{E}(X) = 10$  et  $\sigma(X) = 2$  alors  $\mathbb{P}(|X - 10| \geq 6) \leq \frac{4}{36} = \frac{1}{9}$ .
- Si  $X$  est une mesure expérimentale de moyenne  $m$  et d'écart-type  $\sigma$ , alors la probabilité que la mesure s'écarte de plus de  $3\sigma$  de la moyenne est inférieure à  $\frac{1}{9}$ .

**Remarques :**

Ce théorème donne des informations sur  $X$  sans connaître la loi de  $X$ , seules  $\mathbb{E}(X)$  et  $V(X)$  interviennent.

Bien rappeler les conditions d'application de ce théorème<sup>2</sup> :  $t > 0$  et  $X$  admet une variance.

**Corollaires.**

Soit  $X$  une variable aléatoire réelle,

- ❶ Si  $X$  admet une variance  $\sigma^2$  alors  $\forall t > 0, \mathbb{P}\left(|X - \mathbb{E}(X)| < t\right) \geq 1 - \frac{\sigma^2}{t^2}$ .
- ❷ Si  $X$  admet une variance  $\sigma^2$  alors  $\forall k > 0, \mathbb{P}\left(|X - \mathbb{E}(X)| \geq k\sigma\right) \leq \frac{1}{k^2}$ .

**En effet :****Lien avec les intervalles de confiance.**

$$\mathbb{P}\left(|X - m| < t\right) = \mathbb{P}\left(m \in ]X - t; X + t[ \right) \quad \left( \leq \mathbb{P}\left(m \in [X - t; X + t] \right) \right)$$

Une des applications importantes de ce théorème est la démonstration de **la loi faible des grands nombres**.

**10.3.3 Loi faible des grands nombres.****Théorème :**

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles, on définit :  $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$

Si les  $X_n$  sont mutuellement indépendantes, admettant la même espérance  $m$  et la même variance alors :

$$\text{pour tout } \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} \mathbb{P}(|M_n - m| \geq \varepsilon) = 0$$

**Démonstration :**

- La linéarité de  $E(\cdot)$  donne  $E(M_n) = m$  et l'indépendance des  $X_k$  donne  $V(M_n) = \frac{\sigma^2}{n}$

En appliquant à  $M_n$  l'inégalité de Bienaymé-Tchebychev on obtient :

$$\forall \varepsilon > 0, \quad P(|M_n - m| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

Pour  $\varepsilon$  un réel strictement positif fixé, on a :

$$\forall n \in \mathbb{N}^*, \quad 0 \leq P(|M_n - m| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

2. Comme pour tous les théorèmes du cours

or on sait que  $\lim_{n \rightarrow +\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$ , donc (*théorème des gendarmes*)

$$\text{Quel que soit } \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P(|M_n - m| \geq \varepsilon) = 0$$

**Interprétation :**  $(|M_n - m| \geq \varepsilon)$  est l'événement : " $M_n$  s'éloigne de  $m$  de plus de  $\varepsilon$ "

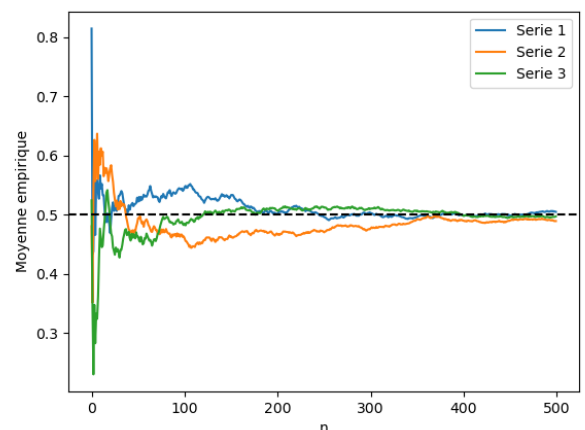
Quand  $n$  est suffisamment grand,  $(|M_n - m| \geq \varepsilon)$  est négligeable, ceci même pour un petit  $\varepsilon > 0$ .

**Remarques :**

- Si  $(X_1, \dots, X_n)$  est un échantillon d'une loi mère  $X$  alors,  
la moyenne empirique  $M_n$  converge en probabilité vers  $E(X)$ . (*Notion hors programme*)
- Quand on observe une expérience aléatoire et qu'on obtient  $n$  valeurs numériques :  $x_1, \dots, x_n$ ,  
pour  $n$  suffisamment grand,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  est une approximation de l'espérance  $m$ .
- C'est ce qui justifie qu'en informatique en réalisant un grand nombre de simulations on obtient une valeur approchée d'une espérance (*Une valeur moyenne*). C'est le principe de la Méthode de Monte-Carlo.

```
def M_(n):
    M = [rd.random()]
    for k in range(1, n):
        M.append( (M[-1]*k+ rd.random())/(k+1) )
    return M

for i in range(1, 4):
    E = M_(500)
    plt.plot(E, label = "Serie "+ str(i))
plt.xlabel("n")
plt.ylabel("Moyenne empirique")
plt.legend()
plt.axhline(0.5, color='black', linestyle='--')
plt.show()
```



**Remarques :**

Ce théorème donne des informations sur la suite  $(M_n)$  sans connaître la loi de  $X$ .

Bien rappeler les conditions d'applications de ce théorème<sup>3</sup> :  $X$  admet *une espérance et une variance*.

**Corollaire.**

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires suivant toutes la loi de Bernoulli de paramètre  $p \in ]0, 1[$ ,

$$\text{on définit : } F_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Si les  $X_n$  sont mutuellement indépendantes, alors

$$\text{pour tout } \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} \mathbb{P}(|F_n - p| \geq \varepsilon) = 0$$

**En effet :**

$(X_n)$  est une suite de variables aléatoires mutuellement indépendantes avec  $E(X_k) = p$  et  $V(X_k) = p(1 - p)$ .

On peut alors appliquer le théorème avec :  $E(F_n) = p$

$$\text{Quel que soit } \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P(|F_n - p| \geq \varepsilon) = 0$$

**Interprétation :** En considérant que l'événement  $(X_k = 1)$  représente un succès,  $F_n$  est la proportion de succès.

$(|F_n - p| \geq \varepsilon)$  est l'événement : " $F_n$  s'éloigne de  $p$  de plus de  $\varepsilon$ "

Quand  $n$  est suffisamment grand,  $(|F_n - p| \geq \varepsilon)$  est négligeable, ceci même pour un petit  $\varepsilon > 0$ .

3. Comme pour tous les théorèmes du cours

**Remarques :**

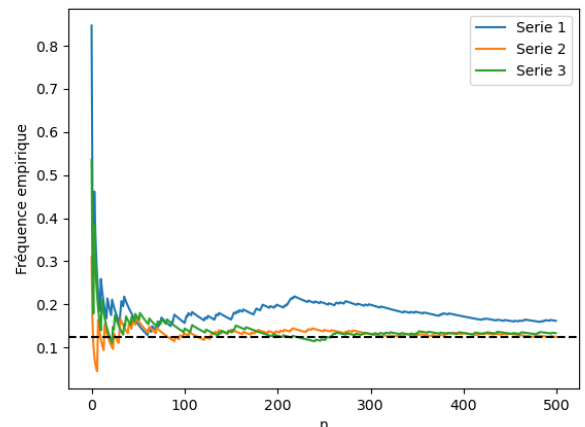
- Si  $(X_1, \dots, X_n)$  est un échantillon de la loi de Bernoulli  $\mathcal{B}(p)$ , alors la fréquence empirique  $F_n$  converge en probabilité vers  $p$ . (Notion hors programme)
- Si  $A$  est un événement d'une expérience que l'on répète  $n$  fois et  $X_k = \mathbb{1}_A$  alors  $F_n$  est le nombre de réalisations de  $A$  divisé par le nombre total d'expériences (la fréquence de réalisation de  $A$ )
- Pour  $n$  suffisamment grand, la réalisation de  $F_n$  est une approximation de  $p$ .
- C'est ce qui justifie qu'en informatique on réalise un grand nombre de simulations pour obtenir une valeur approchée d'une probabilité. (Méthode de Monte-Carlo)

```

p = 1/8
def F_(n):
    f = [ bernoulli(p)]
    for k in range(1, n):
        f.append( (f[-1]*k+ bernoulli(p))/(k+1) )
    return f

for i in range(1, 4):
    E = F_(500)
    plt.plot(E, label = "Serie "+ str(i))
plt.xlabel("n")
plt.ylabel("Fréquence empirique")
plt.legend()
plt.axhline(p, color='black', linestyle='--')
plt.show()

```



## 10.4 Echantillonnage. Estimation.

### 10.4.1 Echantillon

#### Définition. (complément)

Soient  $n \in \mathbb{N}^*$  et  $X$  une variable aléatoire, On appelle  $n$ -échantillon de  $X$  un  $n$ -uplet  $(X_1, \dots, X_n)$  de variables aléatoires mutuellement indépendantes et de même loi que  $X$ .

#### Remarques :

- la loi de  $X$  est appelée la loi mère de l'échantillon.
- On dit que  $X_1, \dots, X_n$  sont  $n$  variables aléatoires indépendantes et identiquement distribuées. (i.i.d)
- On dit que le  $n$ -uplet  $(x_1, \dots, x_n) \in \mathbb{R}^n$  est une réalisation de l'échantillon de  $(X_1, \dots, X_n)$  lorsqu'il existe  $\omega \in \Omega$  tel que :

$$(x_1, \dots, x_n) = (X_1(\omega), \dots, X_n(\omega))$$

### 10.4.2 Base de la statistique inférentielle.

En observant un phénomène aléatoire on obtient une série statistique  $(x_1, \dots, x_n)$  de  $n$  mesures.

On admet l'existence d'un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et d'une loi mère.

La statistique inférentielle a pour but de déterminer des informations sur cette loi mère à partir d'une réalisation  $(x_1, \dots, x_n)$ .

### 10.4.3 Estimateur.

#### Définition. Complément

##### Définition :

Soit  $X$  une variable aléatoire dont la loi dépend d'un paramètre inconnu  $\theta$ .

Soit  $(X_1, \dots, X_n)$  un échantillon de cette loi.

On appelle **estimateur** du paramètre  $\theta$  toute variable aléatoire  $T_n$  qui est une fonction de l'échantillon  $(X_1, \dots, X_n)$  :

$$T_n = g(X_1, \dots, X_n).$$

**Remarque :** Un estimateur est une variable aléatoire, tandis qu'une estimation est la valeur obtenue lorsqu'on remplace l'échantillon aléatoire par des données observées.

#### Biais. Complément

##### Définition.

Soit  $T_n$  un estimateur d'un paramètre  $\theta$ .

On appelle **bias de l'estimateur**  $T_n$ , le réel  $\mathbb{E}(T_n) - \theta$

##### Remarques :

- Un estimateur  $T_n$  de  $\theta$  est dit sans biais lorsque :  $E(T_n) = \theta$
- Un estimateur  $T_n$  de  $\theta$  est dit asymptotiquement sans biais lorsque :  $\lim_{n \rightarrow +\infty} E(T_n - \theta) = 0$

### 10.4.4 Moyenne empirique.

#### Définition.

Soient  $n \in \mathbb{N}^*$  et  $(X_1, \dots, X_n)$  une liste de  $n$  variables aléatoires.

On appelle **moyenne empirique**, notée  $M_n$ , la variable aléatoire :  $M_n = \frac{1}{n} \sum_{k=1}^n X_k$

#### Remarques :

- Parfois cette variable est notée :  $\bar{X}_n$  (*Cette notation est source d'erreurs*)
- Quand on observe une expérience on obtient  $n$  valeurs :  $x_1, \dots, x_n$ ,  $M_n$  prend alors la valeur  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

#### Théorème.

Soient  $n \in \mathbb{N}^*$  et  $(X_1, \dots, X_n)$ , une liste de  $n$  variables aléatoires indépendantes de même loi, admettant une espérance  $\mu$  et une variance  $\sigma^2$  non nulle.

❶  $M_n$  admet une espérance et  $E(M_n) = \mu$ ,

❷  $M_n$  admet une variance et  $V(M_n) = \frac{\sigma^2}{n}$ ,

❸ La variable aléatoire centrée réduite associée à  $M_n$ , notée  $M_n^*$ , est égale à  $\frac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}}$

**Démonstrations.** (*Faites de nombreuses fois en classe, et à savoir refaire*)

$$M_n = \frac{1}{n} \sum_{k=1}^n X_k \text{ donc (linéarité) } E(M_n) = \frac{1}{n} \sum_{k=1}^n E(X_k) \quad \boxed{E(M_n) = \mu}$$

$$\text{et (indépendance) } V(M_n) = \frac{1}{n^2} \sum_{k=1}^n V(X_k) \quad \boxed{E(V_n) = \frac{\sigma^2}{n}}$$

$$\text{et comme } M_n^* \text{, est égale à } \frac{M_n - E(M_n)}{\sqrt{V(M_n)}} \text{ il vient : } \boxed{M_n^* = \frac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}}}$$

**Remarques :**

- ❶  $M_n$  est un estimateur sans biais de  $\mu$ .
- ❷ La variance de  $M_n - \mu$  tend vers 0 quand  $n$  tend vers  $+\infty$ .
- ❸ On a vu que  $\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} P(|M_n - \mu| \geq \varepsilon) = 0$  (*Loi Faible des grands nombres*)

**Interprétations :**

- Quand  $n$  est grand, la dispersion de  $M_n$  est petite.
- Quand  $n$  est grand,  $M_n$  fournit une approximation de  $\mu$ . (*Faire le lien avec les approximations faites avec Python*)

**10.4.5 Ecart-type empirique.****Définition.**

Soient  $n \in \mathbb{N}^*$  et  $(X_1, \dots, X_n)$ , une liste de  $n$  variables aléatoires.

On appelle variance empirique la variable aléatoire :  $S_n^2 = \frac{1}{n} \sum_{k=1}^n (X_k - M_n)^2$

On appelle **écart-type empirique** la variable aléatoire :  $S_n = \sqrt{\frac{1}{n} \sum_{k=1}^n (X_k - M_n)^2}$

**Proposition.** (*Autre expression de la variance empirique*).

Soient  $n \in \mathbb{N}^*$  et  $(X_1, \dots, X_n)$ , une liste de  $n$  variables aléatoires.

$$\frac{1}{n} \sum_{k=1}^n (X_k - M_n)^2 = \frac{1}{n} \left( \sum_{k=1}^n X_k^2 \right) - M_n^2$$

**Démonstration.** (*Faite en classe et à savoir refaire*)

$$\begin{aligned} S_n^2 &= \frac{1}{n} \sum_{k=1}^n (X_k - M_n)^2 \\ &= \frac{1}{n} \sum_{k=1}^n (X_k^2 - 2M_n X_k + M_n^2) \\ &= \frac{1}{n} \sum_{k=1}^n X_k^2 - \frac{2M_n}{n} \sum_{k=1}^n X_k + \frac{1}{n} \sum_{k=1}^n M_n^2 \\ &= \frac{1}{n} \sum_{k=1}^n X_k^2 - 2M_n^2 + M_n^2 \end{aligned}$$

$$S_n^2 = \frac{1}{n} \left( \sum_{k=1}^n X_k^2 \right) - M_n^2$$

**Proposition.** (*complément*)

Soient  $n \in \mathbb{N}^*$  et  $(X_1, \dots, X_n)$ , une liste de  $n$  variables aléatoires indépendantes de même loi, admettant une espérance  $\mu$  et une variance  $\sigma^2$  non nulle.

La variance empirique admet une espérance égale à  $\frac{n-1}{n} \sigma^2$ . (ou encore  $\sigma^2 - \frac{\sigma^2}{n}$ )

**Démonstration.** (*Faite en classe et à savoir refaire*)

On vient de montrer que  $S_n^2 = \frac{1}{n} \left( \sum_{k=1}^n X_k^2 \right) - M_n^2$  on en déduit (avec la linéarité de  $E(\cdot)$ ) que :

$$\begin{aligned} E(S_n^2) &= \frac{1}{n} \sum_{k=1}^n E(X_k^2) - E(M_n^2) \\ &= \frac{1}{n} \sum_{k=1}^n (V(X_k) + E(X_k)^2) - (V(M_n) + E(M_n)^2) \quad \text{En effet : } \underbrace{E(X^2) = V(X) + E(X)^2}_{\text{Koenig - Huygens}} \\ &= \frac{1}{n} \sum_{k=1}^n (\sigma^2 + \mu^2) - \left( \frac{\sigma^2}{n} + \mu^2 \right) \\ &= \sigma^2 - \frac{\sigma^2}{n} \end{aligned}$$

$$S_n^2 = \frac{(n-1)\sigma^2}{n}$$

**Remarques :**

- $E(S_n^2 - \sigma^2) = -\frac{\sigma^2}{n}$  ( $S_n^2$  est un estimateur asymptotiquement sans biais de  $\sigma^2$ )
- En posant  $S_n'^2 = \frac{n}{n-1} S_n^2$ , on a  $E(S_n'^2 - \sigma^2) = 0$  ( $S_n'^2$  est un estimateur sans biais de  $\sigma^2$ )

**Définition.**

On appelle **écart-type empirique corrigé** la variable aléatoire :  $S_n' = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (X_k - M_n)^2}$

## 10.5 Convergence en loi

**Définition**

Soient  $X$  une variable aléatoire réelle et  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles.

On note  $F$  la fonction de répartition de  $X$  et  $D$  l'ensemble des réels où  $F$  est discontinue.

Dire que  $(X_n)$  converge en loi vers  $X$  signifie que :  $\forall x \in \mathbb{R} \setminus D, \lim_{n \rightarrow +\infty} \mathbb{P}(X_n \leq x) = F(x)$

**Remarques :**

- 1 On note :  $X_n \xrightarrow{\mathcal{L}} X$
- 2 On doit pour chaque  $x \in \mathbb{R} \setminus D$  calculer la limite de la suite  $((\mathbb{P}(X_n \leq x))_{n \in \mathbb{N}}$ .
- 3 On admet que l'ensemble  $D$  est au plus dénombrable.
- 4 La convergence en loi permet d'approcher la loi de  $X_n$  par celle de  $X$  lorsque  $n$  est assez grand.

**Exemples :** Voir la feuille\_info\_22.

**Caractérisation** (quand  $X$  est à densité)

Soient  $X$  une variable aléatoire réelle **à densité** et  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles.

$(X_n)$  converge en loi vers  $X$  si, et seulement si,  $\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x)$

**En effet :** Ici  $X$  est à densité donc sa fonction de répartition est continue sur  $\mathbb{R}$  en entier.

(Autrement dit  $D$  est l'ensemble vide)

**Caractérisation** (quand  $X$  et  $X_n$  sont discrètes à valeurs dans  $\mathbb{N}$ )

Soient  $X$  une variable aléatoire réelle à valeurs dans  $\mathbb{N}$  et  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles à valeurs dans  $\mathbb{N}$ .

$(X_n)$  converge en loi vers  $X$  si, et seulement si,  $\forall k \in \mathbb{N}, \lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$

**Démonstration.** (admis)

**Exemple :** Si  $X_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$  alors  $X_n \xrightarrow{\mathcal{L}} X$  avec  $X \hookrightarrow \mathcal{P}(\lambda)$ .

**Remarque :** Dans le théorème de la loi faible des grands nombres nous avons vu une autre convergence.

**Définition** (complément)

Soient  $X$  une variable aléatoire réelle et  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles.

Dire que  $(X_n)$  converge en probabilité vers  $X$  signifie que,

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

## 10.6 Théorème central limite

### 10.6.1 Première forme.

Avec la moyenne empirique.

**Théorème.**

Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires, on définit pour tout  $n \in \mathbb{N}^*$  :  $M_n = \frac{1}{n} \sum_{k=1}^n X_k$

Si les  $X_n$  sont indépendantes de même loi, admettant une espérance  $\mu$  et une variance  $\sigma^2$  non nulle

alors  $(M_n^*)_{n \geq 1}$  converge en loi vers une variable suivant la loi normale centrée réduite.

**Remarques :**

❶ On rappelle que :  $M_n^* = \frac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}}$

❷ On note :  $M_n^* \xrightarrow{\mathcal{L}} X$  avec  $X \hookrightarrow \mathcal{N}(0, 1)$  ou encore  $M_n^* \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$

❸ Autrement dit :  $\forall x \in \mathbb{R}, \mathbb{P}(M_n^* \leq x) = \Phi(x)$

❹ Dans le cas de gaussiennes identiques et indépendantes  $M_n^*$  suit exactement la loi normale centrée réduite.

❺ (complément)

$\left[ M_n - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; M_n + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$  est un intervalle de confiance asymptotique au niveau  $1 - \alpha$  pour  $\mu$ .

**Interprétation :**

La moyenne d'un grand nombre de variables aléatoires indépendantes devient approximativement gaussienne.

**Théorème.** (Autre formulation)

Si  $(X_n)_{n \geq 1}$  est une suite de variables aléatoires indépendantes de même loi admettant une espérance  $\mu$  et une variance  $\sigma^2$  (non nulle) alors :

Pour tout  $(a, b)$  tel que  $a \leq b$ ,  $\lim_{n \rightarrow +\infty} \mathbb{P}(a \leq M_n^* \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt$

**Avec la somme.****Théorème.**

Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires, on définit pour tout  $n \in \mathbb{N}^*$  :  $Y_n = \sum_{k=1}^n X_k$

Si les  $X_n$  sont indépendantes de même loi, admettant une espérance  $\mu$  et une variance  $\sigma^2$  non nulle  
alors  $(Y_n^*)_{n \geq 1}$  converge en loi vers une variable suivant la loi normale centrée réduite.

**En effet :** En remarquant que  $Y_n = nM_n$  on montre facilement que  $Y_n^* = M_n^*$ . *Remarque :*  $Y_n^* = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$

**Interprétation :** La somme de nombreuses contributions indépendantes devient approximativement gaussienne.

**Un cas particulier**

**Théorème de Moivre-Laplace.** (corollaire du théorème précédent).

Soient  $p \in ]0, 1[$  et  $(Y_n)_{n \geq 1}$  une suite de variables aléatoires réelles,  
Si  $Y_n$  suit la loi binomiale de paramètres  $(n, p)$  alors  
 $\frac{Y_n - np}{\sqrt{np(1-p)}}$  converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.

**En effet :** On applique le théorème précédent avec  $X_k \hookrightarrow \mathcal{B}(p)$

**Remarque :**  $E(Y_n) = np$ ,  $V(Y_n) = np(1-p)$  donc  $Y_n^* = \frac{Y_n - np}{\sqrt{np(1-p)}}$ .

**Autrement dit :**

Si  $Y_n \hookrightarrow \mathcal{B}(n, p)$  alors : pour tout  $x \in \mathbb{R}$ ,  $\lim_{n \rightarrow +\infty} \mathbb{P}(Y_n^* \leq x) = \Phi(x)$   
où  $\Phi$  est la fonction de répartition de la loi  $\mathcal{N}(0; 1)$ .

**Approximations.**

Pour  $(X_1, \dots, X_n)$  une liste de  $n$  variables aléatoires i.i.d. d'espérance  $\mu$  et de variance  $\sigma^2 \neq 0$ .

- **Avec la moyenne empirique.**

Pour  $n$  assez grand, la loi de  $M_n^*$  peut être approchée par la loi normale centrée réduite.

Pour  $n$  assez grand, la loi de  $M_n$  peut être approchée par la loi  $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ .

Pour  $n$  assez grand,  $\left[ M_n \pm u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$  est un IC au niveau  $1 - \alpha$  pour  $\mu$ .

- **Avec la somme  $Y_n$ .**

Pour  $n$  assez grand, la loi de  $Y_n^*$  peut être approchée par la loi normale centrée réduite.

Pour  $n$  assez grand, la loi de  $Y_n = \sum_{k=1}^n X_k$  peut être approchée par la loi  $\mathcal{N}(n\mu, n\sigma^2)$ .

- **Avec  $Y_n \hookrightarrow \mathcal{B}(n, p)$ .**

Pour  $n$  assez grand, la loi de  $\frac{Y_n - np}{\sqrt{np(1-p)}}$  peut être approchée par la loi  $\mathcal{N}(0, 1)$ .

Pour  $n$  assez grand, la loi de  $Y_n$  peut être approchée par la loi  $\mathcal{N}(np, np(1-p))$ .

**Remarque :** "La loi de  $X$  peut être approchée par la loi de  $Y$ " se note souvent  $X \approx Y$  et signifie :

- si  $Y$  est à densité :  $\forall x \in \mathbb{R}$ ,  $P(X \leq x) \approx P(Y \leq x)$  (au sens valeur approchée dans  $\mathbb{R}$ ).
- si  $X$  et  $Y$  sont discrètes à valeurs dans  $\mathbb{N}$  :  $\forall k \in \mathbb{N}$ ,  $P(X = k) \approx P(Y = k)$   
(au sens valeur approchée dans  $\mathbb{R}$ ).

**Critères d'approximation pour les lois usuelles.**

Loi de départ	Conditions	Approximation
$\mathcal{B}(n, p)$	$n \geq 30$ et $np \leq 10$	$\mathcal{P}(np)$
$\mathcal{B}(n, p)$	$n \geq 20$ et $p \approx 0,5$	$\mathcal{N}(np, np(1-p))$
$\mathcal{B}(n, p)$	$np \geq 10$ et $n(1-p) \geq 10$	$\mathcal{N}(np, np(1-p))$
$\mathcal{P}(\lambda)$	$\lambda \geq 10$	$\mathcal{N}(\lambda, \lambda)$

Ces critères ne sont pas des théorèmes mais des règles pratiques. Dans un sujet ils devraient vous être rappelés.

**10.6.2 Seconde forme****Avec l'écart-type empirique.****Théorème.**

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires, on définit pour tout  $n \in \mathbb{N}^*$  :

$$M_n = \frac{1}{n} \sum_{k=1}^n X_k \quad \text{et} \quad S_n = \sqrt{\frac{1}{n} \sum_{k=1}^n (X_k - M_n)^2}$$

Si les  $X_n$  sont indépendantes de même loi, admettant une espérance  $\mu$  et une variance non nulle (*inconnue*)

alors  $\left(\frac{M_n - \mu}{\frac{S_n}{\sqrt{n}}}\right)_{n \geq 1}$  converge en loi vers une variable suivant la loi normale centrée réduite.

**Remarques :**

• Ici on remplace l'écart-type par l'écart-type empirique  $S_n$ .

• Autrement dit  $\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{M_n - \mu}{\frac{S_n}{\sqrt{n}}} \leq x\right) = \Phi(x)$

où  $\Phi$  est la fonction de répartition de la loi  $\mathcal{N}(0; 1)$ .

**Théorème. (Autre formulation)**

Si  $(X_n)_{n \in \mathbb{N}}$  est une suite de variables aléatoires indépendantes de même loi admettant une espérance  $\mu$  et une variance  $\sigma^2$  (*non nulle*) alors :

$$\text{Pour tout } a, b \text{ tel que } a \leq b, \quad \lim_{n \rightarrow +\infty} \mathbb{P}\left(a \leq \frac{M_n - \mu}{\frac{S_n}{\sqrt{n}}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt$$

**Avec une fréquence empirique (complément)****Corollaire.**

Si  $(X_n)_{n \in \mathbb{N}}$  est une suite de variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre  $p$  alors :

en notant  $F_n = \frac{1}{n} \sum_{k=1}^n X_k$  (*la fréquence empirique*)

$\frac{\sqrt{n}(F_n - p)}{\sqrt{F_n(1 - F_n)}}$  converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.

**En effet :**

**Remarques :**

- $\frac{\sqrt{n}(F_n - p)}{\sqrt{F_n(1 - F_n)}} = \frac{F_n - p}{\frac{\sqrt{F_n(1 - F_n)}}{\sqrt{n}}}$
- Ici  $S_n^2 = F_n(1 - F_n)$ . (Ici, il est plus simple de calculer  $S_n^2$ )

**Avec l'écart-type empirique corrigé.**

**Théorème.**

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires, on définit pour tout  $n \in \mathbb{N}^*$  :

$$M_n = \frac{1}{n} \sum_{k=1}^n X_k \quad \text{et} \quad S'_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (X_k - M_n)^2}$$

Si les  $X_n$  sont indépendantes de même loi, admettant une espérance  $\mu$  et une variance (*inconnue*)

alors  $\left(\frac{M_n - \mu}{\frac{S'_n}{\sqrt{n}}}\right)_{n \geq 1}$  converge en loi vers une variable suivant la loi normale centrée réduite.

**Remarques :**

- (*complément*)

Lorsque les  $X_k$  sont des lois normales, alors  $\frac{M_n - \mu}{\frac{S'_n}{\sqrt{n}}}$  suit exactement la loi de Student à  $n - 1$  degrés de liberté.

- Les intervalle de confiance de la moyenne sont une application de ce théorème.
- La loi de Student n'est rigoureusement valable que dans le cas normal. Son utilisation dans les autres cas repose sur une approximation.
- En pratique, on utilise la loi de Student à  $n - 1$  degrés de liberté pour des petites valeurs de  $n$ .

## 10.7 Intervalle de confiance de la moyenne.

On observe une expérience aléatoire, plus particulièrement une variable aléatoire  $X$  qui admet une espérance  $\mu$  et une variance  $\sigma^2$  inconnues.

**Proposition.**

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires, on définit pour tout  $n \in \mathbb{N}^*$  :

$$M_n = \frac{1}{n} \sum_{k=1}^n X_k \quad \text{et} \quad S_n = \sqrt{\frac{1}{n} \sum_{k=1}^n (X_k - M_n)^2}$$

Si les  $X_n$  sont indépendantes de même loi, admettant une espérance  $\mu$  et une variance (*inconnues*) alors

$$\forall \alpha \in ]0, 1[, \quad \lim_{n \rightarrow +\infty} \left( \mathbb{P} \left( \left[ |M_n - \mu| \leq u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \right] \right) \right) = 1 - \alpha$$

**Démonstration.** C'est une conséquence de la seconde forme du théorème central limite.

Pour  $(X_1, \dots, X_n)$  un échantillon d'une loi d'espérance  $\mu$  et de variance  $\sigma^2$

Pour  $n$  assez grand ,

$\left[ M_n - u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} ; M_n + u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \right]$  est un intervalle de confiance pour  $\mu$  au niveau de confiance  $1 - \alpha$ .

*On parle aussi d'intervalle de confiance asymptotique.*

**En pratique :** ( On ne connaît pas  $\mu$  et  $\sigma$  on cherche à estimer  $\mu$  avec un intervalle de confiance )

❶ On lit avec soin l'énoncé et on détermine :

$n$  : la taille de l'échantillon et  $(x_1, \dots, x_n)$  les valeurs prises par l'échantillon.

*(attention parfois les données sont sous la forme valeurs/effectifs)*

❷ On calcule  $u_{1-\frac{\alpha}{2}}$  avec une table ou avec Python ou avec la calculatrice.

❸ On calcule  $m_n$  la moyenne empirique,  $s_n^2$  la variance empirique,  $m_n - u_{1-\frac{\alpha}{2}} \frac{s_n}{\sqrt{n}}$  et  $m_n + u_{1-\frac{\alpha}{2}} \frac{s_n}{\sqrt{n}}$

❹ On conclut :

Au seuil de  $\alpha$  on estime que  $\mu$  est dans l'intervalle de confiance  $\left[ m_n - u_{1-\frac{\alpha}{2}} \frac{s_n}{\sqrt{n}} ; m_n + u_{1-\frac{\alpha}{2}} \frac{s_n}{\sqrt{n}} \right]$ .

**Remarque.** Cette notion est souvent représentée dans les barres d'erreurs en sciences expérimentales.

**Exemple :** Au risque de 5% la moyenne théorique est dans l'intervalle :  $\left[ m_n - 1,96 \frac{s_n}{\sqrt{n}} ; m_n + 1,96 \frac{s_n}{\sqrt{n}} \right]$ .

**Proposition.** (Avec des variables de Bernoulli)

Avec des variables de Bernoulli de paramètre  $p$  (inconnue), on obtient un intervalle de confiance de  $p$  :  
(au niveau de confiance  $1 - \alpha$ )

$$\forall \alpha \in ]0, 1[, \quad \mathbb{P} \left( p \in \left[ F_n - u_{1-\frac{\alpha}{2}} \frac{\sqrt{F_n(1-F_n)}}{\sqrt{n}} ; F_n + u_{1-\frac{\alpha}{2}} \frac{\sqrt{F_n(1-F_n)}}{\sqrt{n}} \right] \right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha$$

**Démonstration.** *l'idée c'est qu'ici  $S_n^2 = F_n(1 - F_n)$ .*

— Soit  $\omega \in \Omega$ ,

si  $X(\omega) = 0$  alors  $X^2(\omega) = X(\omega)$  et si  $X(\omega) = 1$  alors  $X^2(\omega) = X(\omega)$

or  $X(\Omega) = \{0, 1\}$  donc  $\forall \omega \in \Omega, X(\omega) = X^2(\omega)$ .

Si la variable aléatoire  $X$  suit une loi de Bernoulli alors  $X^2 = X$

$$\begin{aligned} S_n^2 &= \frac{1}{n} \sum_{k=1}^n X_k^2 - F_n^2 \\ &= \frac{1}{n} \sum_{k=1}^n X_k - F_n^2 \quad \text{car } X_k \text{ est une variable de Bernoulli} \\ &= F_n - F_n^2 \end{aligned}$$

$$\boxed{S_n^2 = F_n - F_n^2 = F_n(1 - F_n)}$$

**En pratique :** ( On ne connaît pas  $p$  on cherche à estimer  $p$  avec un intervalle de confiance )

❶ On lit avec soin l'énoncé et on détermine :

$n$  : la taille de l'échantillon et  $(x_1, \dots, x_n)$  les valeurs prises par l'échantillon.

❷ On calcule  $u_{1-\frac{\alpha}{2}}$  avec une table ou avec Python ou avec la calculatrice.

❸ On calcule  $f_n$  la fréquence empirique, on sait alors que la variance empirique est  $f_n(1 - f_n)$ .

④ On calcule  $f_n - u_{1-\frac{\alpha}{2}} \frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}}$  et  $f_n + u_{1-\frac{\alpha}{2}} \frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}}$

⑤ On conclut :

Au risque de  $\alpha$  on estime que  $p$  est dans l'intervalle de confiance  $\left[ f_n - u_{1-\frac{\alpha}{2}} \frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}} ; f_n + u_{1-\frac{\alpha}{2}} \frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}} \right]$ .

**Exemple.** Pour  $\alpha = 5\%$ , on a :  $u_{1-\frac{\alpha}{2}} \approx 2$  et comme  $f_n(1-f_n) \leq \frac{1}{4}$  alors :

Au seuil de 5% on estime que  $p$  est dans l'intervalle de confiance  $\left[ f_n - \frac{1}{\sqrt{n}} ; f_n + \frac{1}{\sqrt{n}} \right]$

### Intervalle de confiance de Student.

#### Théorème (Complément)

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires, on définit pour tout  $n \in \mathbb{N}^*$  :

$$M_n = \frac{1}{n} \sum_{k=1}^n X_k \quad \text{et} \quad S'_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (X_k - M_n)^2}$$

Si les  $X_n$  sont indépendantes suivant toute **la même loi normale** d'espérance  $\mu$  et de variance (*inconnues*) alors

$$\forall \alpha \in ]0, 1[, \quad \mathbb{P} \left( \mu \in \left[ M_n - t_{1-\frac{\alpha}{2}}^{(n-1)} \frac{S'_n}{\sqrt{n}} ; M_n + t_{1-\frac{\alpha}{2}}^{(n-1)} \frac{S'_n}{\sqrt{n}} \right] \right) = 1 - \alpha$$

où  $t_{1-\alpha}^{(n)}$  est le quantile d'ordre  $1 - \alpha$  de la Loi de Student avec  $n$  degrés de liberté.

#### Remarques :

- Lorsque les  $X_k$  sont des gaussiennes  $\frac{M_n - \mu}{\frac{S'_n}{\sqrt{n}}}$  suit la loi de Student.
- Simulation des lois du khi-2 et de Student.
- Elaboration d'un test du khi-2. Voir l'annexe de la feuille Exo\_25.

## 10.8 Test de conformité à la moyenne.

### 10.8.1 Pour une espérance.

On observe une expérience aléatoire, plus particulièrement une variable aléatoire  $X$  qui admet une espérance  $\mu$  et une variance  $\sigma^2$  inconnues.

On veut tester l'hypothèse  $(H_0) : \mu = \mu_0$ . (Contre l'hypothèse alternative  $(H_1) : \mu \neq \mu_0$ )

Pour cela on prend (*la réalisation*) d'un échantillon  $(X_1, \dots, X_n)$  de la variable aléatoire  $X$ .

Sous l'hypothèse  $H_0$  on sait que : (*seconde forme du théorème central limite*)

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \left[ -u_{1-\frac{\alpha}{2}} \leq \frac{M_n - \mu_0}{\frac{S_n}{\sqrt{n}}} \leq u_{1-\frac{\alpha}{2}} \right] \right) = 1 - \alpha$$

ou encore

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \left| \frac{M_n - \mu_0}{\frac{S_n}{\sqrt{n}}} \right| > u_{1-\frac{\alpha}{2}} \right) = \alpha$$

Donc pour  $n$  assez grand,  $\frac{M_n - \mu_0}{\frac{S_n}{\sqrt{n}}}$  en dehors du segment  $[-u_{1-\frac{\alpha}{2}}; u_{1-\frac{\alpha}{2}}]$  avec une probabilité proche de  $\alpha$ .

Quand  $\frac{M_n - \mu_0}{\frac{S_n}{\sqrt{n}}}$  prend une valeur en dehors du segment  $[-u_{1-\frac{\alpha}{2}}; u_{1-\frac{\alpha}{2}}]$ , on dira qu'on rejette l'hypothèse ( $H_0$ ) au seuil de  $\alpha$ .

Le plus souvent, on prend  $\alpha = 5\%$ ,

Quand  $\frac{M_n - \mu_0}{\frac{S_n}{\sqrt{n}}}$  prend une valeur en dehors de l'intervalle  $[1,96; 1,96]$ ,  
on prend la décision de rejeter l'hypothèse ( $H_0$ ) au seuil de 5%.

### En pratique :

❶ On lit avec soin l'énoncé et on détermine :

$n$  : la taille de l'échantillon,  $(x_1, \dots, x_n)$  les valeurs de l'échantillon et  $\mu_0$  la valeur supposée de l'espérance.

❷ On calcule  $m_n = \frac{1}{n} \sum_{k=1}^n x_k$  la moyenne empirique,  $s_n^2 = \frac{1}{n} \sum_{k=1}^n x_k^2 - m_n^2$  la variance empirique et  $\frac{m_n - \mu_0}{\frac{s_n}{\sqrt{n}}}$

❸ On calcule  $u_{1-\frac{\alpha}{2}}$  avec une table ou avec Python ou avec la calculatrice.

❹ On teste l'hypothèse au seuil de  $\alpha$  :

si  $\frac{m_n - \mu_0}{\frac{s_n}{\sqrt{n}}} \notin [-u_{1-\frac{\alpha}{2}}; u_{1-\frac{\alpha}{2}}]$  alors : on rejette l'hypothèse ( $H_0$ ) au seuil de  $\alpha$ .

si  $\frac{m_n - \mu_0}{\frac{s_n}{\sqrt{n}}} \in [-u_{1-\frac{\alpha}{2}}; u_{1-\frac{\alpha}{2}}]$  alors : on ne peut rien conclure de ce test.

**Exemple :** Voir la feuille\_Cours\_14.

### 10.8.2 Sur une proportion.

Ici on note  $\mu = p$  et  $\mu_0 = p_0$ ,

et on veut tester l'hypothèse ( $H_0$ ) :  $p = p_0$ . (Contre l'hypothèse alternative ( $H_1$ ) :  $p \neq p_0$ )

### En pratique :

❶ On lit avec soin l'énoncé et on détermine :

$n$  : la taille de l'échantillon,  $(x_1, \dots, x_n)$  les valeurs de l'échantillon et  $p_0$  la valeur supposée de la proportion.

❷ On calcule  $f_n = \frac{1}{n} \sum_{k=1}^n x_k$  la fréquence empirique, on sait alors que  $s_n^2 = f_n(1 - f_n)$

❸ On calcule  $\frac{f_n - p_0}{\frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}}}$

❹ On calcule  $u_{1-\frac{\alpha}{2}}$  avec une table ou avec Python ou avec la calculatrice.

❺ On teste l'hypothèse au seuil de  $\alpha$  :

si  $\frac{f_n - p_0}{\frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}}} \notin [-u_{1-\frac{\alpha}{2}}; u_{1-\frac{\alpha}{2}}]$  alors : on rejette l'hypothèse ( $H_0$ ) au seuil de  $\alpha$ .

si  $\frac{f_n - p_0}{\frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}}} \in [-u_{1-\frac{\alpha}{2}}; u_{1-\frac{\alpha}{2}}]$  alors : on ne peut rien conclure de ce test.

**Remarques :**

- Ici on peut remplacer  $\frac{f_n - p_0}{\frac{\sqrt{f_n(1-f_n)}}{\sqrt{n}}}$  par  $\frac{f_n - p_0}{\frac{\sqrt{p_0(1-p_0)}}{\sqrt{n}}}$  (Car on est sous l'hypothèse  $p = p_0$ )

- Pour  $\alpha = 5\%$  en utilisant la majoration  $f_n(1 - f_n) \leq \frac{1}{4}$  cela devient :

si  $\frac{f_n - p_0}{\frac{1}{\sqrt{n}}} \notin [-1; 1]$  alors : on rejette l'hypothèse ( $H_0$ ) au seuil de  $\alpha$ .

si  $\frac{f_n - p_0}{\frac{1}{\sqrt{n}}} \in [-1; 1]$  alors : on ne peut rien conclure de ce test.

# Produit scalaire dans $\mathbb{R}^n$ .

## Plan du chapitre

11.1	Généralités . . . . .	<b>161</b>
11.1.1	Définitions . . . . .	161
11.1.2	Propriétés . . . . .	163
11.1.3	Inégalité de Cauchy-Schwarz. . . . .	164
11.1.4	Inégalité triangulaire. . . . .	164
11.1.5	Vecteurs orthogonaux . . . . .	165
11.1.6	Théorème de Pythagore . . . . .	165
11.2	Bases orthonormales. . . . .	<b>166</b>
11.2.1	Famille orthogonale de vecteurs non nuls . . . . .	166
11.2.2	Définition. . . . .	167
11.2.3	Coordonnées dans une base orthonormale. . . . .	168
11.2.4	Expression du produit scalaire dans une base orthonormale. . . . .	169
11.3	Matrice symétrique réelle. . . . .	<b>170</b>
11.3.1	Orthogonalité des vecteurs propres. . . . .	170
11.3.2	Théorème spectral . . . . .	170
11.4	Orthogonal d'un sous-espace vectoriel . . . . .	<b>171</b>
11.4.1	Définition. . . . .	171
11.4.2	Propriétés. . . . .	172
11.4.3	Décomposition. . . . .	173
11.5	Projection orthogonale. . . . .	<b>174</b>
11.5.1	Définition. . . . .	174
11.5.2	Caractérisation. . . . .	175
11.6	Distance. . . . .	<b>176</b>
11.6.1	Définitions. . . . .	176
11.6.2	Lien entre distance et projeté orthogonal. . . . .	177
11.6.3	Cas particulier. (complément) . . . . .	177
11.6.4	Ajustement affine par la méthode des moindres carrés. . . . .	177

## 11.1 Généralités

$n$  désigne ici un entier naturel non nul. On note  $\mathcal{B}_c$  la base canonique de  $\mathbb{R}^n$ .

### 11.1.1 Définitions

#### Produit scalaire usuel dans $\mathbb{R}^n$

**Définition.**

Le produit scalaire usuel sur  $\mathbb{R}^n$  est l'application de  $\mathbb{R}^n \times \mathbb{R}^n$  dans  $\mathbb{R}$  qui à tout couple  $((x_1, \dots, x_n), (y_1, \dots, y_n))$  associe le réel :  $\sum_{i=1}^n x_i y_i$

Pour  $u = (x_1, \dots, x_n) \in \mathbb{R}^n$  et  $v = (y_1, \dots, y_n) \in \mathbb{R}^n$  on note :  $\langle u; v \rangle$  le produit scalaire de  $u$  et  $v$ .

$$\langle u; v \rangle = \sum_{i=1}^n x_i y_i$$

**Remarque :** on parle aussi de produit scalaire canonique de  $\mathbb{R}^n$ .

**Norme euclidienne sur  $\mathbb{R}^n$**

**Définition.**

Soit  $u$  un vecteur de  $\mathbb{R}^n$ , on appelle norme (euclidienne) de  $u$  le réel noté  $\|u\|$  et défini par :

$$\|u\| = \sqrt{\langle u; u \rangle}$$

$$\|u\|^2 = \sum_{i=1}^n x_i^2 \qquad \|u\| = \sqrt{\sum_{i=1}^n x_i^2}$$

**Remarque :** nous avons vu d'autres normes en DS et en exercices.

**Ecriture matricielle.**

**Proposition.**

Soient  $u$  et  $v$  deux vecteurs de  $\mathbb{R}^n$ , on note  $X = \text{Coord}_{\mathcal{B}_c}(u)$  et  $Y = \text{Coord}_{\mathcal{B}_c}(v)$ ,

$$\langle u; v \rangle = X^T Y \qquad \|u\|^2 = X^T X$$

**Remarque :** On confond ici<sup>1</sup>  $\mathbb{R}$  et  $\mathcal{M}_1(\mathbb{R})$ .

**Proposition.** (complément)

❶ Soient  $A \in \mathcal{M}_{n,p}(\mathbb{R})$  et  $B \in \mathcal{M}_{n,m}(\mathbb{R})$ ,  
 en notant :  $C_1, \dots, C_p$  les colonnes de  $A$ ,  $C'_1, \dots, C'_m$  les colonnes de  $B$ .  
 $A^T B$  est la matrice de  $\mathcal{M}_{p,m}(\mathbb{R})$  définie par :  $(A^T B)_{i,j} = C_i^T C'_j$

❷ Soient  $(u_1, \dots, u_p)$  et  $(v_1, \dots, v_m)$  deux familles de vecteurs de  $\mathbb{R}^n$ ,  
 en notant  $A = \text{Mat}_{\mathcal{B}_c}(u_1, \dots, u_p)$ ,  $B = \text{Mat}_{\mathcal{B}_c}(v_1, \dots, v_m)$ .

$A^T B$  est la matrice de  $\mathcal{M}_{p,m}(\mathbb{R})$  définie par :  $(A^T B)_{i,j} = \langle u_i; v_j \rangle$

**En effet :**  $(A^T B)_{i,j} = \sum_{k=1}^n (A^T)_{i,k} (B)_{k,j} = \sum_{k=1}^n n_{a_{k,i}} n_{b_{k,j}} = \sum_{k=1}^n n_{(C_i)_k} n_{(C'_j)_k} = C_i^T C'_j$

**Illustration.**

On peut écrire :

$$A = \begin{pmatrix} | & & | \\ C_1 & \cdots & C_p \\ | & & | \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} | & & | \\ C'_1 & \cdots & C'_m \\ | & & | \end{pmatrix}. \qquad A^T = \begin{pmatrix} - & C_1^T & - \\ & \vdots & \\ - & C_p^T & - \end{pmatrix}.$$

$$A^T B = \begin{pmatrix} - & C_1^T & - \\ & \vdots & \\ - & C_p^T & - \end{pmatrix} \begin{pmatrix} | & & | \\ C'_1 & \cdots & C'_m \\ | & & | \end{pmatrix}.$$

Ainsi, le coefficient en position  $(i, j)$  est donné par :  $(A^T B)_{i,j} = C_i^T C'_j$ .

**Autrement dit :** Le produit matriciel  $A^T B$  est la matrice dont les coefficients sont les produits scalaires des colonnes de  $A$  par les colonnes de  $B$ .

---

1. comme dans beaucoup de sujets

### 11.1.2 Propriétés

On note  $E = \mathbb{R}^n$ ,

#### Proposition

❶ *Bilinéarité.*

$$\forall (u_1, u_2, v) \in E^3, \forall (a, b) \in \mathbb{R}^2, \quad \langle au_1 + bu_2; v \rangle = a \langle u_1; v \rangle + b \langle u_2; v \rangle$$

$$\forall (u, v_1, v_2) \in E^3, \forall (a, b) \in \mathbb{R}^2, \quad \langle u; av_1 + bv_2 \rangle = a \langle u; v_1 \rangle + b \langle u; v_2 \rangle$$

❷ *Symétrie.*

$$\forall (u, v) \in E^2, \quad \langle u; v \rangle = \langle v; u \rangle$$

❸ *Définie.*

$$\forall u \in E, \quad \langle u; u \rangle = 0 \iff u = 0_E$$

❹ *Positive.*

$$\forall u \in E, \quad \langle u; u \rangle \geq 0$$

#### Démonstration.

Pour ❶ en notant  $\text{coord}_{\mathcal{B}_c}(u) = X$ ,  $\text{coord}_{\mathcal{B}_c}(v) = Y$  et  $\text{coord}_{\mathcal{B}_c}(w) = Z$ ,

$$\begin{aligned} \langle \alpha u + \beta v; w \rangle &= (\alpha X + \beta Y)^T Z \\ &= (\alpha X^T + \beta Y^T) Z \\ &= \alpha X^T Z + \beta Y^T Z \\ &= \alpha \langle u; w \rangle + \beta \langle v; w \rangle \end{aligned}$$

#### Remarques :

- $\forall u \in E, \quad \langle u; 0_E \rangle = 0 \quad \text{et} \quad \langle 0_E; u \rangle = 0$
- $\forall a \in \mathbb{R}, \forall u \in E, \quad \|au\| = |a| \|u\|$

#### Proposition.

❶  $\forall (u, v) \in E^2, \quad \|u + v\|^2 = \|u\|^2 + 2 \langle u; v \rangle + \|v\|^2$

❷  $\forall (a, b) \in \mathbb{R}^2, \forall (u, v) \in E^2, \quad \|au + bv\|^2 = a^2 \|u\|^2 + 2ab \langle u; v \rangle + b^2 \|v\|^2$

En effet : Poit  $(u, v) \in E^2$ ,

$$\begin{aligned} \|u + v\|^2 &= \langle u + v; u + v \rangle \\ &= \langle u; u + v \rangle + \langle v; u + v \rangle \\ &= \langle u; u \rangle + \langle u; v \rangle + \langle v; u \rangle + \langle v; v \rangle \\ &= \langle u; u \rangle + 2 \langle u; v \rangle + \langle v; v \rangle \end{aligned}$$

#### Proposition. Complément

$$\forall (u_i)_{1 \leq i \leq p} \in E^p, \quad \left\| \sum_{i=1}^p u_i \right\|^2 = \sum_{i=1}^p \|u_i\|^2 + 2 \sum_{1 \leq i < j \leq p} \langle u_i; u_j \rangle$$

#### Démonstration.

$$\begin{aligned} \left\| \sum_{i=1}^p u_i \right\|^2 &= \left\langle \sum_{i=1}^p u_i; \sum_{j=1}^p u_j \right\rangle \\ &= \sum_{i=1}^p \left\langle u_i; \sum_{j=1}^p u_j \right\rangle && \text{linéarité par rapport au premier vecteur} \\ &= \sum_{i=1}^p \sum_{j=1}^p \langle u_i; u_j \rangle && \text{linéarité par rapport au deuxième vecteur} \end{aligned}$$

$$= \sum_{i=1}^p \|u_i\|^2 + 2 \sum_{1 \leq i < j \leq p} \langle u_i; u_j \rangle \quad \text{symétrie.}$$

**Remarque.** Faire le lien avec le complément donnant la variance d'une somme de variables aléatoires.

### 11.1.3 Inégalité de Cauchy-Schwarz.

**Théorème.**

$$\forall (u, v) \in E^2, \quad \langle u; v \rangle^2 \leq \|u\|^2 \|v\|^2$$

**Démonstration.**

Pour  $u = 0_E$  l'inégalité est vérifiée, il suffit de le montrer en supposant  $u \neq 0_E$ .

On définit  $P(x) = \|xu + v\|^2$  et pour  $x \in \mathbb{R}$ ,

$$\begin{aligned} P(x) &= \|xu + v\|^2 \\ &= \langle xu + v | xu + v \rangle \\ &= x \langle u | xu + v \rangle + \langle v | xu + v \rangle \\ &= x^2 \langle u | u \rangle + x \langle u | v \rangle + x \langle v | u \rangle + \langle v | v \rangle \\ &= x^2 \|u\|^2 + 2x \langle u | v \rangle + \|v\|^2 \end{aligned}$$

le discriminant de ce trinôme est égal à :  $\Delta = 4 \langle u | v \rangle^2 - 4 \|u\|^2 \|v\|^2$

de plus  $\forall x \in \mathbb{R}, P(x) = \|xu + v\|^2$  donc  $\forall x \in \mathbb{R}, P(x) \geq 0$

et ainsi le trinôme  $P$  n'a pas deux racines réelles distinctes ce qui entraîne que  $\Delta \leq 0$ , on en déduit bien :

$$\langle u | v \rangle^2 \leq \|u\|^2 \|v\|^2$$

■

**Autre formulation.**

$$|\langle u; v \rangle| \leq \|u\| \|v\| \quad \left( \sum_{i=1}^n x_i y_i \right)^2 \leq \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 \quad (X^T Y)^2 \leq X^T X Y^T Y$$

**Théorème.** (cas d'égalité)

$$\forall (u, v) \in E^2, \quad \langle u; v \rangle^2 = \|u\|^2 \|v\|^2 \iff u \text{ et } v \text{ sont colinéaires}$$

**Démonstration.** Pour  $u = 0_E$  l'équivalence est vérifiée, il suffit de la démontrer en supposant  $u \neq 0_E$ . (on reprend les notations de la démonstration précédente)

$$\begin{aligned} \langle u | v \rangle^2 = \|u\|^2 \|v\|^2 & \text{ si, et seulement si, } & \text{Le discriminant de } P \text{ est nul} \\ & \text{si, et seulement si, } & P \text{ possède une unique racine réelle} \\ & \iff & \exists ! \lambda \in \mathbb{R}, \quad P(\lambda) = 0 \\ & \iff & \exists \lambda \in \mathbb{R}, \quad P(\lambda) = 0 \quad \text{car } \forall x \in \mathbb{R}, P(x) \geq 0 \\ & \iff & \exists \lambda \in \mathbb{R}, \quad \|\lambda u + v\|^2 = 0 \\ & \iff & \exists \lambda \in \mathbb{R}, \quad v = -\lambda u \end{aligned}$$

et comme  $u \neq 0_E, \langle u | v \rangle^2 = \|u\|^2 \|v\|^2$  si, et seulement si,  $u$  et  $v$  sont colinéaires.

### 11.1.4 Inégalité triangulaire.

**Théorème.**

$$\forall (u, v) \in E^2, \quad \|u + v\| \leq \|u\| + \|v\|$$

**Démonstration.**

$$\begin{aligned} \|u + v\|^2 &= \langle u + v | u + v \rangle \\ &= \|u\|^2 + 2 \langle u | v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 && \text{d'après Cauchy-Schwarz} \\ &\leq (\|u\| + \|v\|)^2 \end{aligned}$$

or  $\|u + v\| \geq 0$  et  $\|u\| + \|v\| \geq 0$  donc ( Pour  $a \geq 0$  et  $b \geq 0$ , si  $a^2 \leq b^2$  alors  $a \leq b$  )

on a bien :  $\forall (u, v) \in E^2, \|u + v\| \leq \|u\| + \|v\|$  ■

**Théorème. (cas d'égalité)**

$$\forall (u, v) \in E^2, \|u + v\| = \|u\| + \|v\| \iff u = 0_E \text{ ou } \exists \lambda \in \mathbb{R}_+, v = \lambda u$$

**Démonstration.**

⊆ si  $u = 0_E$  c'est immédiat, si  $v = \lambda u$  avec  $\lambda \geq 0$  alors :

d'une part :  $\|u + v\| = \|u + \lambda u\| = |1 + \lambda| \|u\| = (1 + \lambda) \|u\|$  car  $(1 + \lambda) \geq 0$

d'autre part :  $\|u\| + \|v\| = \|u\| + \|\lambda u\| = \|u\| + \lambda \|u\| = (1 + \lambda) \|u\|$  car  $\lambda \geq 0$  ■

⊇ On suppose que  $\|u + v\| = \|u\| + \|v\|$ ,

on en déduit que :  $\|u + v\|^2 = \|u\|^2 + 2\|u\|\|v\| + \|v\|^2$ ,

or on sait  $\|u + v\|^2 = \|u\|^2 + 2 \langle u; v \rangle + \|v\|^2$  on en déduit que  $\langle u; v \rangle = \|u\|\|v\|$

on est dans le cas d'égalité de Cauchy-Schwartz et ainsi  $u = 0_E$  ou l'existence de  $\lambda \in \mathbb{R}$  tel que  $v = \lambda u$ ,

et alors la relation  $\langle u; v \rangle = \|u\|\|v\|$  permet d'établir  $\lambda = \frac{\|v\|}{\|u\|} \geq 0$  ■

**Théorème. (Généralisation)**

$$\forall (u_i)_{1 \leq i \leq p} \in E^p, \left\| \sum_{i=1}^p u_i \right\| \leq \sum_{i=1}^p \|u_i\|$$

**Démonstration. (Récurrence sur n)**

**11.1.5 Vecteurs orthogonaux**

**Définition**

Soient  $u$  et  $v$  deux vecteurs de  $\mathbb{R}^n$ ,  
Dire que  $u$  et  $v$  sont orthogonaux si, et seulement si,  $\langle u; v \rangle = 0$

**Remarques :**

- On pourra noter  $u \perp v$ .
- On définit de même la notion de matrices colonnes orthogonales :

"X et Y de  $\mathcal{M}_{n,1}(\mathbb{R})$  sont orthogonales" signifie que  $X^T Y = 0$

**11.1.6 Théorème de Pythagore**

**Théorème.**

Pour tout  $(u, v) \in E^2$ ,  
 $u$  et  $v$  sont orthogonaux si, et seulement si,  $\|u + v\|^2 = \|u\|^2 + \|v\|^2$

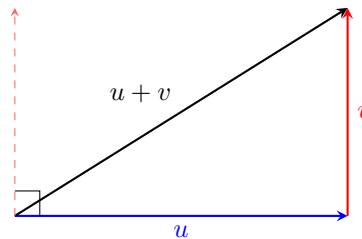
**Démonstration.**Soit  $(u, v) \in E^2$ ,

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v; u + v \rangle \\
&= \langle u; u + v \rangle + \langle v; u + v \rangle \\
&= \langle u; u \rangle + \langle u; v \rangle + \langle v; u \rangle + \langle v; v \rangle \\
&= \langle u; u \rangle + 2\langle u; v \rangle + \langle v; v \rangle
\end{aligned}$$

On en déduit que :

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 \iff \langle u; v \rangle = 0$$

Deux vecteurs  $u$  et  $v$  de  $E$  sont orthogonaux si, et seulement si,  $\|u + v\|^2 = \|u\|^2 + \|v\|^2$

*Illustration graphique.***Théorème.**

Soient  $p$  un entier naturel non nul et  $u_1, u_2, \dots, u_p$  des vecteurs de  $E$ ,

Si les  $u_i$  sont 2 à 2 orthogonaux alors,  $\left\| \sum_{i=1}^p u_i \right\|^2 = \sum_{i=1}^p \|u_i\|^2$

**En effet :** conséquence de la relation :  $\left\| \sum_{i=1}^p u_i \right\|^2 = \sum_{i=1}^p \|u_i\|^2 + 2 \sum_{1 \leq i < j \leq p} \langle u_i; u_j \rangle$

**Remarques.**

❶ Si  $ABC$  est rectangle en  $A$  alors  $AB^2 + AC^2 = BC^2$

❷ Si  $u$  et  $v$  sont orthogonaux alors  $\|u + v\|^2 = \|u\|^2 + \|v\|^2$

❸ Si les  $u_i$  sont 2 à 2 orthogonaux alors,  $\left\| \sum_{i=1}^p u_i \right\|^2 = \sum_{i=1}^p \|u_i\|^2$

- Les réciproques des implications ❶ et ❷ sont vraies, mais pas celle de ❸.

**Exercice :** Trouver un contre-exemple dans  $\mathbb{R}^2$  montrons que la réciproque de ❸ est fautive.

*Une analyse non faite au tableau a permis de trouver les vecteurs suivants :*

En prenant :  $u = (1, 0)$ ,  $v = (0, 1)$  et  $w = (1, -1)$

on a :  $\|u + v + w\|^2 = \dots = 4$ ,  $\|u\|^2 + \|v\|^2 + \|w\|^2 = \dots = 4$  et  $\langle u; w \rangle = 1 \neq 0$ .

*Ce contre-exemple montre que la réciproque de la généralisation du théorème de Pythagore est fautive pour  $p \geq 3$ .*

## 11.2 Bases orthonormales.

### 11.2.1 Famille orthogonale de vecteurs non nuls

**Définition.**

Soient  $p$  un entier naturel non nul et  $(u_1, \dots, u_p)$  une famille de  $p$  vecteurs de  $\mathbb{R}^n$ ,

Dire que  $(u_1, \dots, u_p)$  est **une famille orthogonale** signifie que :  $\forall (i, j) \in \llbracket 1; p \rrbracket^2, i \neq j \implies \langle u_i; u_j \rangle = 0$

**Remarques :**

- On dira aussi "Les  $u_i$  sont deux à deux orthogonaux".
- On définit de même une famille orthogonale de  $\mathcal{M}_{n,1}(\mathbb{R})$
- Il suffit d'avoir :  $\forall (i, j) \in \llbracket 1; p \rrbracket^2, \quad i < j \implies \langle u_i; u_j \rangle = 0$
- (*complément*) En notant  $A$  la matrice de  $p$  vecteurs  $(u_1, \dots, u_p)$  dans la base canonique de  $\mathbb{R}^n$ ,  
 $(u_1, \dots, u_p)$  est une famille orthogonale si, et seulement si  $A^\top A$  est diagonale

**Théorème.**

Soient  $p$  un entier naturel non nul et  $(u_1, \dots, u_p)$  une famille de  $p$  vecteurs de  $\mathbb{R}^n$ ,  
 Si  $(u_1, \dots, u_p)$  est une famille orthogonale de vecteurs **non nuls** alors elle est libre.

**Démonstration.**

On suppose que  $(u_1, \dots, u_p)$  est une famille orthogonale de vecteurs non nuls de  $E$ ,

Soit  $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$  tel que  $\sum_{k=1}^p \lambda_k u_k = 0_E$ ,

on en déduit pour  $j \in \llbracket 1; p \rrbracket$ ,  $\left\langle u_j; \sum_{k=1}^p \lambda_k u_k \right\rangle = 0$ , en déduit (*bilinéarité*)  $\sum_{k=1}^p \lambda_k \langle u_j; u_k \rangle = 0$

et comme la famille est orthogonale il vient :  $\lambda_j \langle u_j; u_j \rangle = 0$  ou encore  $\lambda_j \|u_j\|^2 = 0$

de plus  $u_j \neq 0_E$  donc  $\|u_j\|^2 \neq 0$  et ainsi  $\lambda_j = 0$ ,

on a bien montré que si  $\sum_{k=1}^p \lambda_k u_k = 0_E$  alors tous les  $\lambda_k$  sont nuls.

donc :

si  $(u_1, \dots, u_p)$  est une famille orthogonale de vecteurs non nuls alors  $(u_1, \dots, u_p)$  est libre

**11.2.2 Définition.**

**Définition.**

Soient  $p$  un entier naturel non nul et  $(u_1, \dots, u_p)$  une famille de  $p$  vecteurs de  $\mathbb{R}^n$ ,  
 Dire que  $(u_1, \dots, u_p)$  est **une famille orthonormale** signifie que :

$$\forall i \in \llbracket 1; p \rrbracket, \quad \|u_i\| = 1 \quad \text{et} \quad \forall (i, j) \in \llbracket 1; p \rrbracket^2, \quad i \neq j \implies \langle u_i; u_j \rangle = 0$$

**Remarques :**

- Une famille orthonormale est en particulier une famille orthogonale de vecteurs non nuls, elle est donc libre.
- Une famille orthonormale de  $p$  vecteurs d'un sous-espace  $F$  de dimension  $p$  est une base de  $F$ .  
*une base orthonormale de  $F$ .*
- On définit de même une famille orthonormale de  $\mathcal{M}_{n,1}(\mathbb{R})$
- La base canonique de  $\mathbb{R}^n$  est une base orthonormale de  $\mathbb{R}^n$ .

**Propositions.**

En notant  $P$  la matrice de  $n$  vecteurs  $(u_1, \dots, u_n)$  dans la base canonique de  $\mathbb{R}^n$ ,

$$(u_1, \dots, u_n) \text{ est une base orthonormale si, et seulement si, } P^\top P = I_n$$

Les colonnes de  $M \in \mathcal{M}_n(\mathbb{R})$  forment une famille orthonormale

$$\text{si, et seulement si, } M \text{ est inversible et } M^{-1} = M^\top$$

**En effet :**  $P^\top P = (\langle C_i, C_j \rangle)_{1 \leq i, j \leq n}$

**Propositions.**

Soient  $p \in \mathbb{N}^*$  et  $(v_1, \dots, v_p)$  une famille de  $p$  vecteurs d'une sous-espace vectoriel  $F$  de  $E$ ,

- ❶ Si  $(v_1, \dots, v_p)$  est une famille orthogonale et si les  $v_i$  sont **non nuls**

alors  $\left( \frac{v_1}{\|v_1\|}, \dots, \frac{v_p}{\|v_p\|} \right)$  est une famille orthonormale de vecteurs de  $F$ .

- ❷ Si  $(v_1, \dots, v_p)$  est une famille orthonormale et si  $\dim(F) = p$

alors  $(v_1, \dots, v_p)$  est une base orthonormale de  $F$

**En effet :**

**Théorème.**

Tout sous-espace vectoriel de  $\mathbb{R}^n$  possède une base orthonormale.

**Démonstration.** (*admis*)

**En pratique.** *On commence par trouver une base orthogonale, ensuite on norme chaque vecteur de la famille.*

**11.2.3 Coordonnées dans une base orthonormale.**

**Théorème**

Soient  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$ ,  $u$  un vecteur de  $F$  et  $\mathcal{B} = (e_1, \dots, e_p)$  une base de  $F$ .

Si  $\mathcal{B}$  est une base orthonormale de  $F$  alors  $\text{coord}_{\mathcal{B}}(u) = \begin{pmatrix} \langle u; e_1 \rangle \\ \vdots \\ \langle u; e_p \rangle \end{pmatrix}$

*Pour obtenir les coordonnées dans une base orthonormale il suffit de faire des produits scalaires.*

**Démonstration.**

on note  $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \text{Coord}_{\mathcal{B}}(u)$ , on a :  $u = \sum_{i=1}^p x_i e_i$  donc pour  $j \in \llbracket 1; p \rrbracket$ ,

$$\begin{aligned} \langle u | e_j \rangle &= \left\langle \sum_{i=1}^p x_i e_i \mid e_j \right\rangle \\ &= \sum_{i=1}^p x_i \langle e_i \mid e_j \rangle \\ &= x_j \langle e_j \mid e_j \rangle \\ &= x_j \end{aligned}$$

donc

$$\text{Coord}_{\mathcal{B}}(u) = \begin{pmatrix} \langle u; e_1 \rangle \\ \vdots \\ \langle u; e_p \rangle \end{pmatrix}$$

**Exemple :** *Voir le feuille\_cours\_11\_2*

### 11.2.4 Expression du produit scalaire dans une base orthonormale.

Extrait du programme : On souligne le fait que le produit scalaire et la norme se calculent de la même manière dans toutes les bases orthonormales.

#### Théorème

Soient  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$ ,  $u$  et  $v$  deux vecteurs de  $F$  et  $\mathcal{B}$  une base de  $F$ .

on note :  $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix} = \text{coord}_{\mathcal{B}}(u)$  et  $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} = \text{coord}_{\mathcal{B}}(v)$

Si  $\mathcal{B}$  est une base orthonormale de  $F$  alors

$$\langle u; v \rangle = \sum_{i=1}^p \alpha_i \beta_i \quad \|u\| = \sqrt{\sum_{i=1}^p \alpha_i^2}$$

**Démonstration.** Ce qu'on a fait au tableau .

On note  $\mathcal{B} = (e_1, \dots, e_p)$  et on a  $u = \sum_{i=1}^p \alpha_i e_i$  et  $v = \sum_{i=1}^p \beta_i e_i$  donc

$$\begin{aligned} \langle u; v \rangle &= \left\langle \sum_{i=1}^p \alpha_i e_i \mid \sum_{j=1}^p \beta_j e_j \right\rangle \\ &= \sum_{i=1}^p \sum_{j=1}^p \alpha_i \beta_j \langle e_i \mid e_j \rangle && \text{bilinéarité} \\ &= \sum_{i=1}^p \alpha_i \beta_i \langle e_i \mid e_i \rangle && \text{car la famille est orthogonale} \\ &= \sum_{i=1}^n \alpha_i \beta_i && \text{car } \|e_i\| = 1 \end{aligned}$$

$$\langle u; v \rangle = \sum_{i=1}^n \alpha_i \beta_i$$

Une autre rédaction .

On note  $P$  la matrice de passage de la base canonique à la base  $\mathcal{B}$ ,  $X = \text{coord}_{\mathcal{B}_c}(u)$ ,  $Y = \text{coord}_{\mathcal{B}_c}(v)$ ,  $X' = \text{coord}_{\mathcal{B}}(u)$ ,  $Y' = \text{coord}_{\mathcal{B}}(v)$ .

$$\begin{aligned} \langle u; v \rangle &= X^T Y \\ &= (PX')^T (PY') && \text{(Formule de changement de base)} \\ &= X'^T P^T P Y' \\ &= X'^T Y' && (P^T P = I_n \text{ car } \mathcal{B} \text{ est orthonormale}) \end{aligned}$$

$$\langle u; v \rangle = \sum_{i=1}^p \alpha_i \beta_i$$

On en déduit en prenant  $u = v$  :

$$\|u\|^2 = \sum_{i=1}^p \alpha_i^2$$

**Théorème (complément)**

Si  $P$  est la matrice de passage entre deux bases orthonormales alors  $P$  est inversible et  $P^{-1} = P^T$

**En effet :**  $\langle C_i \mid C_j \rangle = \langle e'_i \mid e'_j \rangle$

## 11.3 Matrice symétrique réelle.

### 11.3.1 Orthogonalité des vecteurs propres.

**Théorème.**

Soient  $n$  un entier naturel non nul,  $A \in \mathcal{M}_n(\mathbb{R})$ ,  $(X_1, X_2) \in \mathcal{M}_{n,1}(\mathbb{R})^2$  et  $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ .

Si  $\left\{ \begin{array}{l} A \text{ est symétrique} \\ X_1 \text{ est un vecteur propre de } A \text{ associée à } \lambda_1 \\ X_2 \text{ est un vecteur propre de } A \text{ associée à } \lambda_2 \\ \lambda_1 \neq \lambda_2 \end{array} \right.$  alors  $X_1$  et  $X_2$  sont orthogonales.

**Démonstration.**

Soit  $A$  une matrice symétrique de  $\mathcal{M}_n(\mathbb{R})$ ,  
on suppose connaître  $X_1, X_2$  dans  $\mathcal{M}_{n,1}(\mathbb{R})$  et  $\lambda_1, \lambda_2$  dans  $\mathbb{R}$  tels que :

$$\underbrace{X_1 \neq 0, X_2 \neq 0}_{\text{inutile ici}}, \quad AX_1 = \lambda_1 X_1 \quad \text{et} \quad AX_2 = \lambda_2 X_2 \quad \lambda_1 \neq \lambda_2$$

$$AX_1 = \lambda_1 X_1 \quad \text{donc} \quad X_1^T A^T = \lambda_1 X_1^T \quad \text{et comme } A = A^T \text{ on a : } X_1^T A = \lambda_1 X_1^T$$

en multipliant par  $X_2$  (à droite) il vient  $X_1^T A X_2 = \lambda_1 X_1^T X_2$ ,

la relation  $AX_2 = \lambda_2 X_2$  entraîne  $\lambda_2 X_1^T X_2 = \lambda_1 X_1^T X_2$  ou encore :  $(\lambda_1 - \lambda_2) X_1^T X_2 = 0$

et comme  $(\lambda_1 - \lambda_2) \neq 0$  on a bien  $X_1^T X_2 = 0$ ,

Les deux vecteurs propres  $X_1$  et  $X_2$  sont orthogonaux

**Rappel :**  $A$  symétrique signifie que  $A^\top = A$ , lire : "transposée de  $A$  égale  $A$ ".

**Corollaire.**

Soient  $n$  un entier naturel non nul,  $A \in \mathcal{M}_n(\mathbb{R})$ ,  $(X_1, X_2) \in \mathcal{M}_{n,1}(\mathbb{R})^2$  et  $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ .  
Si  $A$  est symétrique,  $X_1 \in E_{\lambda_1}(A)$ ,  $X_2 \in E_{\lambda_2}(A)$  et  $\lambda_1 \neq \lambda_2$  alors  $X_1$  et  $X_2$  sont orthogonales.

**Corollaire.**

Soient  $n$  un entier naturel non nul,  $A \in \mathcal{M}_n(\mathbb{R})$ ,

Si  $A$  est symétrique, alors en juxtaposant des bases orthogonales des sous-espaces propres de  $A$  on obtient une famille orthogonale formée de vecteurs propres de  $A$ .

### 11.3.2 Théorème spectral

**Théorème.**

Soit  $M \in \mathcal{M}_n(\mathbb{R})$ ,

Si  $M$  est symétrique alors il existe deux matrices  $P \in \mathcal{M}_n(\mathbb{R})$  et  $\Delta \in \mathcal{M}_n(\mathbb{R})$  telles que :

$$\Delta \text{ diagonale, } P \text{ inversible, } P^{-1} = P^\top \quad \text{et} \quad M = P \Delta P^\top$$

**Démonstration.** (*admis*)

**Remarques :**

- Si  $M$  est symétrique et réelle alors  $M$  possède une base orthonormale de vecteurs propres.
- " $P$  inversible,  $P^{-1} = P^\top$ " équivaut à "les colonnes de  $P$  forment une base orthonormale de  $\mathcal{M}_{n,1}(\mathbb{R})$ "
- Une matrice symétrique réelle est non seulement diagonalisable, mais on peut trouver, **parmi toutes les bases de vecteurs propres**, une base orthonormale.

- Si  $A$  est une matrice symétrique de  $\mathcal{M}_n(\mathbb{R})$  ayant  $n$  valeurs propres distinctes  
alors toutes les bases de vecteurs propres de  $A$  sont orthogonales.

**Autre formulation du théorème.**

Si  $M$  est une matrice réelle symétrique

$$\text{alors l'endomorphisme } \begin{matrix} \mathcal{M}_{n,1}(\mathbb{R}) & \longrightarrow & \mathcal{M}_{n,1}(\mathbb{R}) \\ X & \longmapsto & MX \end{matrix} \text{ est diagonalisable dans une base orthonormale de } \mathcal{M}_{n,1}(\mathbb{R}).$$

Ce qu'on simplifie par :

**Si  $M$  est une matrice réelle symétrique alors elle est diagonalisable dans une base orthonormale.**

**En pratique.** (Voir la feuille Cour\_11\_3)

Pour diagonaliser une matrice réelle symétrique  $M$  dans une base orthonormale :

- 1 "La matrice  $M$  est symétrique réelle donc  $M$  est diagonalisable".  
*(ne pas perdre du temps en utilisant la méthode générale).*
- 2 On détermine le spectre de  $M$ .
- 3 On détermine une base de chaque sous-espace propre.  
*(On vérifie (sans l'écrire) que la somme des dimensions est égale à  $n$ ).*
- 4 **1er cas :**  $M$  a  $n$  valeurs propres distinctes. *(C'est le cas le plus simple)*  
Pour chaque valeur propre on détermine un vecteur propre.  
On a alors  $(X_1, \dots, X_n)$  une base orthogonale de  $\mathcal{M}_{n,1}(\mathbb{R})$  formée de vecteurs propres de  $M$ .  
*(Cette affirmation est justifiée par le corollaire de 3.1)*

**2ème cas :**  $M$  a au moins un sous espace propre de dimension supérieure ou égale à 2.

On détermine une base orthogonale de chaque sous-espace propre. *(L'énoncé devrait, en principe, vous aider)*

En juxtaposant toutes ces bases,

on obtient  $(X_1, \dots, X_n)$  une base orthogonale de  $\mathcal{M}_{n,1}(\mathbb{R})$  formée de vecteurs propres de  $M$ .  
*(Cette affirmation est justifiée par le corollaire de 3.1)*

- 5 Il reste à normer les vecteurs :

$$\text{En prenant } (C_1, \dots, C_n) = \left( \frac{X_1}{\|X_1\|}, \dots, \frac{X_n}{\|X_n\|} \right)$$

La famille  $(C_1, \dots, C_n)$  est une base orthonormale de  $\mathcal{M}_{n,1}(\mathbb{R})$  formée de vecteurs propres de  $M$ .

**Autrement dit :**

En notant pour chaque  $i$ ,  $\lambda_i$  la valeur propre associée à  $C_i$ .

$$\text{en notant } P \text{ la matrice dont les colonnes sont } C_1, \dots, C_n \text{ et } D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

$$\underbrace{P \text{ inversible}}_{\text{c'est une base orthonormale}} \text{ et } \underbrace{P^{-1} = P^T} \text{ et } \underbrace{M = PDP^{-1}}_{\text{de vecteurs propres de } M}$$

## 11.4 Orthogonal d'un sous-espace vectoriel

On note  $E = \mathbb{R}^n$

### 11.4.1 Définition.

**Définition.**

Soit  $F$  un sous espace vectoriel de  $E$ , on définit  $F^\perp$  (l'orthogonal de  $F$ ) par :

$$F^\perp = \{ u \in E \mid \forall v \in F, \langle u; v \rangle = 0 \}$$

**Remarques :** 1  $E^\perp = \{0_E\}$       2  $\{0_E\}^\perp = E$

- 3 C'est l'ensemble des vecteurs orthogonaux à tous les vecteurs de  $F$ .

## 11.4.2 Propriétés.

## Théorèmes.

Soit  $F$  un sous espace vectoriel de  $E$ , (on note  $(e_1, \dots, e_m)$  une base de  $F$ )

- ❶  $F^\perp$  est un sous-espace vectoriel de  $E$ .
- ❷  $F \cap F^\perp = \{0_E\}$
- ❸ quel que soit  $u \in E$ ,  $u \in F^\perp \iff \forall i \in \llbracket 1, m \rrbracket, u \perp e_i$
- ❹  $\dim(F) + \dim(F^\perp) = n$
- ❺  $(F^\perp)^\perp = F$

## Démonstrations.

❶  $F^\perp$  est une partie de  $E$  et il contient le vecteur nul  $0_E$  car  $\forall v \in F, \langle v; 0_E \rangle = 0$ , montrons qu'il est stable par combinaison linéaire :

Soient  $(\alpha, \beta) \in \mathbb{R}^2$  et  $(u_1, u_2) \in F^{\perp 2}$ ,  
pour un  $v \in F$  quelconque,

$$\begin{aligned} \langle \alpha u_1 + \beta u_2; v \rangle &= \alpha \langle u_1; v \rangle + \beta \langle u_2; v \rangle && \text{(bilinearité)} \\ &= \alpha 0 + \beta 0 && \text{(car } (u_1, u_2) \in F^{\perp 2} \text{)} \\ &= 0 \end{aligned}$$

donc  $\alpha u_1 + \beta u_2 \in F^\perp$

*Ce qui achève la démonstration de :  $F^\perp$  est stable par combinaison linéaire.*

En conclusion :

$F^\perp$  est un sous-espace vectoriel de  $E$

❷

(Démonstration par double inclusion)

$\supseteq$   $F$  et  $F^\perp$  sont des sous-espaces vectoriels de  $E$  donc  $0_E \in F \cap F^\perp$ .

$\subseteq$  Soit  $u \in F \cap F^\perp$ ,

on a  $u \in F^\perp$  donc  $\forall v \in F, \langle u; v \rangle = 0$

en particulier  $u \in F$  donc on a :  $\langle u; u \rangle = 0$  ou encore  $\|u\|^2 = 0$  ce qui entraîne :  $u = 0_E$ ,

on a bien :  $F \cap F^\perp \subset \{0_E\}$ ,

En conclusion :  $F \cap F^\perp = \{0_E\}$

❸ Démonstration par double implication.

$\Rightarrow$  Si  $u \in F^\perp$  alors  $\forall v \in F, \langle u | v \rangle = 0$  en particulier cela entraîne que  $\forall i \in \llbracket 1; m \rrbracket, \langle u | e_i \rangle = 0$ .

$\Leftarrow$  Réciproquement : On suppose que pour un  $u \in E, \forall i \in \llbracket 1; m \rrbracket, \langle u | e_i \rangle = 0$ .

Prenons un  $v \in F$ , il existe  $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^m$  tel que  $v = \sum_{i=1}^p \lambda_i e_i$

$$\begin{aligned} \langle u | v \rangle &= \langle u | \sum_{i=1}^p \lambda_i e_i \rangle \\ &= \sum_{i=1}^p \lambda_i \langle u | e_i \rangle && \text{(bilinearité)} \\ &= 0 \end{aligned}$$

donc  $\forall v \in F, \langle u | v \rangle = 0$  et ainsi  $u \in F^\perp$

En conclusion :  $u \in F^\perp \iff \forall i \in \llbracket 1; m \rrbracket, \langle u | e_i \rangle = 0$

**Autrement dit** : un vecteur est dans l'orthogonal de  $F$  si, et seulement s'il est orthogonal aux vecteurs d'une base de  $F$ .

④ (Admis pour l'instant.) On le démontrera avec le théorème du rang appliqué à la  $p_F$ .

⑤ (Démonstration d'une inclusion)

Soit  $v \in F$ ,

pour tout  $u \in F^\perp$  on a  $\langle u; v \rangle = 0$  donc  $v \in (F^\perp)^\perp$ ,

$$F \subset (F^\perp)^\perp$$

En admettant ④ on a  $\dim(F) = \dim((F^\perp)^\perp)$  ce qui permet de justifier  $F = (F^\perp)^\perp$

Remarque : ④ donne  $\dim(F) + \dim(F^\perp) = n$  et  $\dim(F^\perp) + \dim((F^\perp)^\perp) = n$

**Proposition.** (Un cas particulier : vecteur normal à un hyperplan)

Soit  $v = (a_1, \dots, a_n) \in \mathbb{R}^n$  tel que  $v \neq 0_E$ ,

$$\text{Vect} \langle v \rangle = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n a_i x_i = 0 \right\}^\perp$$

$$\begin{aligned} \text{En effet : } \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n a_i x_i = 0 \right\} &= \{ u \in E \mid \langle u \mid v \rangle = 0 \} \\ &= \{ u \in E \mid \forall t \in \text{Vect} \langle v \rangle, \langle t \mid u \rangle = 0 \} \\ &= \text{Vect} \langle v \rangle^\perp \end{aligned}$$

Remarque :  $\dim \left( \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n a_i x_i = 0 \right\} \right) = n - 1$

Exemples dans  $\mathbb{R}^3$  :

$$(\text{Vect} \langle (1, 2, -4) \rangle)^\perp = \{ (x, y, z) \in \mathbb{R}^3 \mid x - 2y - 4z = 0 \} \quad \{ (x, y, z) \in \mathbb{R}^3 \mid 2x - 3z = 0 \}^\perp = \text{Vect} \langle (2, 0, -3) \rangle$$

### 11.4.3 Décomposition.

**Théorème**

Soit  $F$  un sous espace vectoriel de  $E$ ,

quel que soit  $u \in E$ , il existe un unique couple  $(u_F, u_{F^\perp}) \in F \times F^\perp$  vérifiant  $u = u_F + u_{F^\perp}$

**définition :**  $u_F$  est appelé **projeté orthogonal** de  $u$  sur  $F$ , on note le note  $p_F(u)$

**Démonstration.**

Soit  $u \in E$ ,

• (Existence) (non faite en classe)

Idée : On note  $(e_1, \dots, e_m)$  une base orthonormale de  $F$  et on pose  $v = \sum_{i=1}^m \langle u \mid e_i \rangle e_i$ ,

• d'une part on a :  $v \in F$

• d'autre part pour tout  $i \in \llbracket 1; m \rrbracket$ ,  $\langle u - v \mid e_i \rangle = \dots = 0$  donc (en utilisant ③)  $u - v \in F^\perp$

$$\langle u - v, e_i \rangle = \langle u, e_i \rangle - \sum_{j=1}^m \langle u, e_j \rangle \langle e_j, e_i \rangle = \langle u, e_i \rangle - \langle u, e_i \rangle = 0.$$

donc en posant  $(u_F, u_{F^\perp}) = (v, u - v)$  on a :  $(u_F, u_{F^\perp}) \in F \times F^\perp$  vérifiant  $u = u_F + u_{F^\perp}$

• (Unicité)

supposons qu'il existe deux couples  $(v_1, v_2) \in F \times F^\perp$  et  $(v'_1, v'_2) \in F \times F^\perp$  tels que  $\begin{cases} u = v_1 + v_2 \\ u = v'_1 + v'_2 \end{cases}$ .

On a alors  $v_1 + v_2 = v'_1 + v'_2$  ce qui entraîne  $\underbrace{v_1 - v'_1}_{\in F} = \underbrace{v'_2 - v_2}_{\in F^\perp}$  et donc  $v_1 - v'_1 \in F \cap F^\perp$

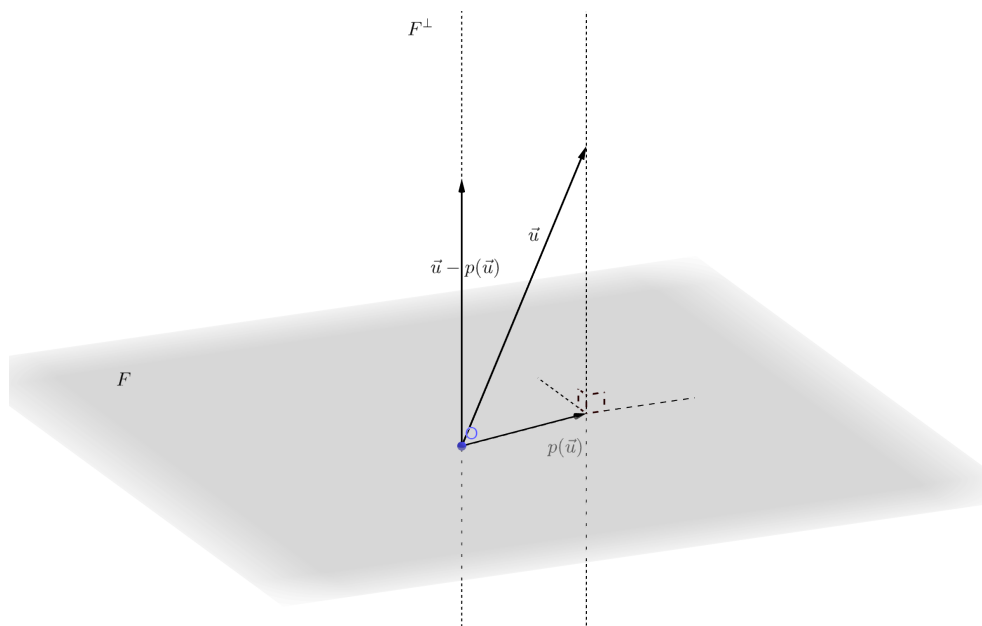
or on a vu que  $F \cap F^\perp = \{0_E\}$  donc  $v_1 = v'_1$ , puis  $v_2 = v'_2$ , le couple  $(v_1, v_2)$  est bien unique.

**Remarques :**

- ❶  $p_F(x)$  est l'unique vecteur  $v$  de  $\mathbb{R}^n$  vérifiant :  $v \in F$  et  $u - v \in F^\perp$ .
- ❷  $\forall u \in \mathbb{R}^n$ ,  $p_F(u) \in F$  et  $u - p_F(u) \in F^\perp$
- ❸ En particulier  $u - p_F(u) \perp p_F(u)$ , ce qui entraîne :  $\|u - p_F(u)\|^2 + \|p_F(u)\|^2 = \|u\|^2$ .  
(théorème de Pythagore)
- ❹  $\forall u \in F$ ,  $p_F(u) = u$  et  $\forall u \in F^\perp$ ,  $p_F(u) = 0_E$

**En effet :** tout vecteur  $u \in F$  s'écrit  $u = \underbrace{u}_{\in F} + \underbrace{0_E}_{\in F^\perp}$  donc pour tout  $u \in F$ ,  $p_F(u) = u$

tout vecteur  $u \in F^\perp$  s'écrit  $u = \underbrace{0_E}_{\in F} + \underbrace{u}_{\in F^\perp}$  donc pour tout  $u \in F^\perp$ ,  $p_F(u) = 0_E$



## 11.5 Projection orthogonale.

### 11.5.1 Définition.

**Définition.** (La projection orthogonale sur  $F$ ).

Soit  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$ ,  
 La projection orthogonale sur  $F$  est l'application de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$  qui, à tout vecteur  $u$  associe son projeté orthogonal sur  $F$ .

$$p_F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$u \longmapsto u_F \quad \text{où} \quad u = \underbrace{u_F}_{\in F} + \underbrace{u_{F^\perp}}_{\in F^\perp}$$

**Remarques.**

- Autrement dit : La projection orthogonale sur  $F$  est l'application qui, à tout vecteur, associe sa composante dans  $F$  lorsqu'il est décomposé sur  $F$  et  $F^\perp$ .
- Si  $F = \mathbb{R}^n$  alors  $p_F = Id_E$       Si  $F = \{0_E\}$  alors  $p_F$  est l'endomorphisme nul.

**Théorème.**

Soit  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$ , la projection orthogonale sur  $F$  est un endomorphisme de  $\mathbb{R}^n$ .

**Démonstration :** On note  $p$  la projection orthogonale sur  $F$ .

$p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , donc il suffit de montrer que  $p$  est linéaire

Soient  $(\alpha, \beta) \in \mathbb{R}^2$  et  $(u, v) \in E^2$ ,

en décomposant  $u$  et  $v$  :  $u = \underbrace{u_F}_{\in F} + \underbrace{u_{F^\perp}}_{\in F^\perp}$  et  $v = \underbrace{v_F}_{\in F} + \underbrace{v_{F^\perp}}_{\in F^\perp}$

on obtient :  $\alpha u + \beta v = \alpha(u_F + u_{F^\perp}) + \beta(v_F + v_{F^\perp}) = \underbrace{\alpha u_F + \beta v_F}_{\in F} + \underbrace{\alpha u_{F^\perp} + \beta v_{F^\perp}}_{\in F^\perp}$

ce qui donne  $p(\alpha u + \beta v) = \alpha u_F + \beta v_F$ , en encore :  $p(\alpha u + \beta v) = \alpha p(u) + \beta p(v)$  ■

### 11.5.2 Caractérisation.

**Théorème.**

Soit  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$ ,  
la projection orthogonale sur  $F$  est l'endomorphisme  $p$  de  $\mathbb{R}^n$  vérifiant :

- ❶  $p \circ p = p$ .
- ❷  $\ker(p)$  est égal à  $F^\perp$ .
- ❸  $\text{Im}(p)$  est égal à  $F$ .

**Démonstrations.**

- Supposons que  $p$  est la projection orthogonale sur  $F$ .

Soit  $u \in E$ ,

$$\begin{aligned} p \circ p(u) &= p(p(u)) \\ &= p(u) \quad (\text{car } p(u) \in F \text{ et } \forall v \in F, p(v) = v) \end{aligned}$$

$$p \circ p = p \quad \text{❶}$$

(Double inclusion)

On sait que  $u - p(u) \in F^\perp$  donc

- d'une part si  $p(u) = 0_E$  alors  $u \in F^\perp$  donc  $\ker(p) \subset F^\perp$ .
- d'autre part si  $u \in F^\perp$  alors  $u = \underbrace{0_E}_{\in F} + \underbrace{u}_{\in F^\perp}$  donc  $p(u) = 0_E$  d'où  $F^\perp \subset \ker(p)$ .

$$\ker(p) = F^\perp \quad \text{❷}$$

(Double inclusion)

- d'une part pour tout  $u \in F$ ,  $u = p(u)$  donc  $F \subset \text{Im}(p)$
- d'autre part pour tout  $u \in E$ ,  $p(u) \in F$  donc  $\text{Im}(p) \subset F$

$$\text{Im}(p) = F \quad \text{❸}$$

- Supposons que  $p$  vérifie ❶, ❷ et ❸.

Soit  $u \in E$ ,

d'une part,  $p(u - p(u)) = p(u) - p \circ p(u) = 0_E$  donc  $u - p(u) \in \ker(p)$  ou encore  $u - p(u) \in F^\perp$

d'autre part,  $p(u) \in \text{Im}(p)$  et comme  $F = \text{Im}(p)$ , il vient  $p(u) \in F$

donc  $p(u)$  est le projeté orthogonal de  $u$  sur  $F$ , autrement dit :  $p$  est la projection orthogonal sur  $F$ .

**Remarques :**

- Si un endomorphisme vérifie ces trois affirmations alors c'est nécessairement  $p_F$ .
- En notant  $M = \text{Mat}_{\mathcal{B}}(p_F)$  (matrice dans une base quelconque), on a :  $M^2 = M$ .
- Le théorème du rang permet enfin de justifier :  $\dim(F) + \dim(F^\perp) = n$

- $\ker(p_F)^\perp = \text{Im}(p_F)$  et  $\text{Im}(p_F)^\perp = \ker(p_F)$

### Écriture avec une base orthonormale.

#### **Théorème.**

Soient  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$  et  $(e_1, \dots, e_m)$  une base orthonormale de  $F$ .

La projection orthogonale sur  $F$  est définie par :  $\forall u \in \mathbb{R}^n, p_F(u) = \sum_{k=1}^m \langle u; e_k \rangle e_k$

**En effet :** (Voir la démonstration de l'existence de le théorème de décomposition).

#### **Remarques :**

- Pour définir la projection orthogonale de  $F$  il suffit de connaître une base orthonormale de  $F$ .
- On utilise ce théorème pour (entre autres) déterminer la matrice de  $p_F$  dans la base canonique.

**Corollaire.** (Projection sur une droite vectorielle).

Soit  $v$  un vecteur non nul de  $\mathbb{R}^n$ , on note  $F = \text{vect}(v)$  (La droite vectorielle dirigée par  $v$ )

La projection orthogonale sur  $F$  est définie par :  $\forall u \in \mathbb{R}^n, p_F(u) = \frac{\langle u; v \rangle}{\|v\|^2} v$

**En effet :** en notant :  $e_1 = \frac{v}{\|v\|}$  on a  $(e_1)$  est une base orthonormale de  $F$  donc (d'après le théorème)

$$p_F(u) = \langle u; e_1 \rangle e_1 = \left\langle u; \frac{v}{\|v\|} \right\rangle \frac{v}{\|v\|} = \frac{\langle u; v \rangle}{\|v\|^2} v$$

## 11.6 Distance.

### 11.6.1 Définitions.

**Définition.** Distance entre deux vecteurs.

Soient  $u$  et  $v$  deux vecteurs de  $\mathbb{R}^n$ ,  
on appelle distance (euclidienne) entre  $u$  et  $v$  le réel  $\|u - v\|$ .

**Propriétés :**

Soient  $u, v$  et  $w$  trois vecteurs de  $\mathbb{R}^n$ ,

- ❶  $d(u, v) \geq 0$ .    ❷  $d(u, v) = d(v, u)$ .    ❸  $d(u, v) = 0 \iff u = v$ .    ❹  $d(u, w) \leq d(u, v) + d(v, w)$

**Définition.** Distance d'un vecteur à une partie non vide de  $\mathbb{R}^n$ .

Soit  $u$  un vecteur de  $\mathbb{R}^n$  et  $S$  une partie non vide de  $\mathbb{R}^n$ ,  
on appelle distance de  $u$  à  $S$  le réel :  $\inf \{ d(u, v) \mid v \in S \}$   
(on notera  $d(u, S)$  ce réel)

**Remarque :** La borne inférieure existe car cet ensemble de réel est minorée par 0.

Revoir la définition d'une borne inférieure et le théorème de la borne supérieure dans  $\mathbb{R}$ .

Quand l'ensemble  $\{d(u, v) \mid v \in S\}$  possède un plus petit élément, (ie :  $\exists v \in S : d(u, v) = d(u, S)$ ), on note

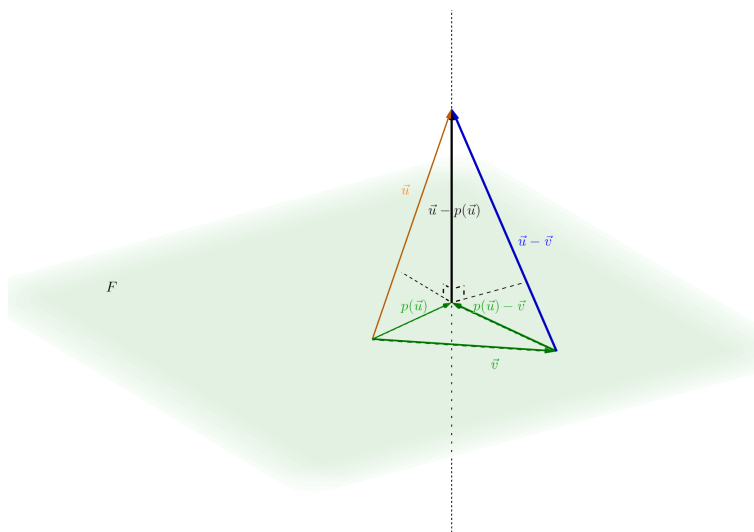
$$d(u, S) = \min(\{d(u, v) \mid v \in S\})$$

### 11.6.2 Lien entre distance et projeté orthogonal.

**Théorème.**

Si  $F$  est un sous-espace vectoriel de  $\mathbb{R}^n$  alors  $\forall u \in \mathbb{R}^n, \quad d(u, F) = \|u - p_F(u)\|$

*Autrement dit : Ce résultat signifie que la distance à  $F$  est réalisée par le projeté orthogonal.*



**Démonstration.**

- Soit  $v \in F$

$$u - v = \underbrace{(u - P_F(u))}_{\in F^\perp} + \underbrace{(P_F(u) - v)}_{\in F}$$

donc (en appliquant le théorème de Pythagore) :  $\|u - v\|^2 = \|u - P_F(u)\|^2 + \|P_F(u) - v\|^2$

- On en déduit que :  $\|u - P_F(u)\|^2 \leq \|u - v\|^2$

les normes sont positives donc on obtient :  $\forall v \in F, \quad \|u - P_F(u)\| \leq \|u - v\|$ .

- et comme  $p_F(u) \in F$  on obtient :  $\|u - p_F(u)\| = \min\{\|u - v\| \mid v \in F\}$

$$d(u, F) = \|u - p_F(u)\|$$

**Remarques :**

Pour  $u$  un vecteur de  $\mathbb{R}^n$  et  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$ ,

- $d(u, F) = \|p_{F^\perp}(u)\|$
- $p_F(u)$  est le vecteur de  $F$  réalisant le minimum sur  $F$  de la fonction  $v \mapsto d(u, v)$ .
- Pour tout  $v \in F, \quad d(u, p_F(u)) \leq d(u, v)$ , ou encore  $\|u - p_F(u)\| \leq \|u - v\|$
- Ici l'ensemble  $\{d(u, v) \mid v \in F\}$  admet un plus petit élément,  $d(u, F) = \min(\{d(u, v) \mid v \in F\})$ .

### 11.6.3 Cas particulier. (complément)

**Proposition.** (Distance d'un vecteur à un plan de  $\mathbb{R}^3$ ).

Soit  $n$  un vecteur non nul de  $\mathbb{R}^3$ , on note  $F = \{v \in \mathbb{R}^3 \mid v \perp n\}$  (Le plan de vecteur normal  $n$ )

$$\forall u \in \mathbb{R}^3, \quad d(u, F) = \frac{|\langle u; n \rangle|}{\|n\|}$$

**En effet :** En posant  $e_1 = \frac{n}{\|n\|}$ ,  $(e_1)$  est une base orthonormale de  $F^\perp$  donc  $p_{F^\perp}(u) = \langle u | e_1 \rangle e_1$

on en déduit :  $\|p_{F^\perp}(u)\| = |\langle u | e_1 \rangle| = \frac{|\langle u; n \rangle|}{\|n\|}$

### 11.6.4 Ajustement affine par la méthode des moindres carrés.

(voir Annexe C)

# Couples de variables aléatoires réelles.

## Plan du chapitre

---

12.1	Couples de variables aléatoires discrètes. . . . .	<b>178</b>
12.1.1	Introduction. . . . .	178
12.1.2	Loi conjointe. . . . .	178
12.1.3	Lois marginales. . . . .	179
12.1.4	Indépendance de variables aléatoires. . . . .	179
12.1.5	Lois conditionnelles. . . . .	180
12.1.6	Recherche de la loi de $u(X,Y)$ sur des exemples. . . . .	181
12.1.7	Espérance de $u(X,Y)$ . . . . .	182
12.2	Covariance. Variance d'une somme. . . . .	<b>182</b>
12.2.1	Définition. . . . .	182
12.2.2	Propriétés. . . . .	183
12.2.3	Variance d'une somme. . . . .	184
12.2.4	Cas particuliers des variables aléatoires indépendantes. . . . .	184
12.2.5	Coefficient de corrélation. . . . .	185

---

## 12.1 Couples de variables aléatoires discrètes.

### 12.1.1 Introduction.

Soient  $X$  et  $Y$  deux variables aléatoires discrètes sur le même espace  $(\Omega, \mathcal{T}, \mathbb{P})$ .

on note  $X(\Omega) = \{ x_i \mid i \in I \}$  avec  $\overbrace{I = \llbracket 1; n \rrbracket}^{\text{finie}}$  ou  $\overbrace{I = \mathbb{N}}^{\text{dénombrable}}$  ( et  $\forall (i, j) \in I^2, \quad i \neq j \implies x_i \neq x_j$  )  
 $Y(\Omega) = \{ y_i \mid i \in J \}$  avec  $J = \llbracket 1; n' \rrbracket$  ou  $J = \mathbb{N}$  ( et  $\forall (i, j) \in J^2, \quad i \neq j \implies y_i \neq y_j$  )

#### Définition.

L'application  $(X, Y) : \Omega \longrightarrow \mathbb{R}^2$  est appelée couple de variables aléatoires sur  $(\Omega, \mathcal{T}, \mathbb{P})$ .  
 $\omega \longmapsto (X(\omega), Y(\omega))$

**Remarque :**  $(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$   
*(autrement dit : certains couples de  $X(\omega) \times Y(\omega)$  peuvent ne pas être pris par  $(X, Y)$ )*

### 12.1.2 Loi conjointe.

#### Définition.

Soit  $(X, Y)$  un couple de variables aléatoires sur  $(\Omega, \mathcal{T}, \mathbb{P})$ .  
 On appelle loi (conjointe) du couple  $(X, Y)$  l'application  $X(\Omega) \times Y(\Omega) \longrightarrow \mathbb{R}$ .  
 $(x, y) \longmapsto \mathbb{P}([X = x] \cap [Y = y])$

#### Remarques :

- Donner la loi conjointe du couple  $(X, Y)$  c'est :

donner  $X(\Omega)$ ,  $Y(\Omega)$  et donner pour tout  $(i, j) \in I \times J$ ,  $\mathbb{P}((X = x_i) \cap (Y = y_j))$

- Lorsque le nombre de valeurs est fini on présente souvent la loi conjointe sous forme de tableau. *(Voir EX 1 dans la feuille\_cours\_13)*
- on ne donne pas toujours la loi conjointe sous la forme d'un tableau. *(Voir EX 2 dans la feuille\_cours\_13)*
- On pourra noter  $(X = x, Y = y)$  l'événement  $(X = x) \cap (Y = y)$ .

**Proposition.** *(Système complet d'événements)*

Soit  $(X, Y)$  un couple de variables aléatoires finies sur  $(\Omega, \mathcal{F}, \mathbb{P})$ .

- ❶ Pour tous  $(x, y), (x', y')$  de  $X(\Omega) \times Y(\Omega)$ 

$$(x, y) \neq (x', y') \implies (X = x, Y = y) \cap (X = x', Y = y') = \emptyset$$
- ❷  $\bigcup_{(i,j) \in I \times J} (X = x_i, Y = y_j) = \Omega$
- ❸ La somme double  $\sum_{(i,j) \in I \times J} P((X = x_i, Y = y_j))$  converge et sa somme vaut 1.

### 12.1.3 Lois marginales.

**Définition.**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

- la loi de probabilité de  $X$  est appelée **première loi marginale** du couple  $(X, Y)$ ,
- la loi de probabilité de  $Y$  est appelée **deuxième loi marginale** du couple  $(X, Y)$ .

**Proposition :** *lien entre la loi conjointe et les lois marginales :*

Soient  $X$  et  $Y$  deux variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

$$\forall i \in I, \quad \mathbb{P}([X = x_i]) = \sum_{j \in J} \mathbb{P}([X = x_i] \cap [Y = y_j]) \quad \forall j \in J, \quad \mathbb{P}([Y = y_j]) = \sum_{i \in I} \mathbb{P}([X = x_i] \cap [Y = y_j])$$

**Justification :** *C'est juste l'application de la formule des probabilités totales. (version avec les intersections)*

On peut résumer les différentes propriétés dans un tableau quand les variables sont finies.

$X \backslash Y$	$y_1$	$y_2$	$\cdots$	$y_{n'}$	$\mathbb{P}(X = x_i)$
$x_1$	$p_{11}$	$p_{12}$	$\cdots$	$p_{1n'}$	$\sum p_{1\cdot}$
$x_2$	$p_{21}$	$p_{22}$	$\cdots$	$p_{2n'}$	$\sum p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$x_n$	$p_{n1}$	$p_{n2}$	$\cdots$	$p_{nn'}$	$p_{n\cdot}$
$\mathbb{P}(Y = y_j)$	$\sum p_{\cdot 1}$	$\sum p_{\cdot 2}$	$\cdots$	$\sum p_{\cdot n'}$	1

### 12.1.4 Indépendance de variables aléatoires.

**Définition :**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

Dire que  $X$  et  $Y$  sont indépendantes signifie que

$$\forall (i, j) \in I \times J, \quad \mathbb{P}([X = x_i] \cap [Y = y_j]) = \mathbb{P}([X = x_i]) \times \mathbb{P}([Y = y_j])$$

**Remarques :**

- Cela revient à dire que pour tout  $(i, j)$  les événements  $[X = x_i]$  et  $[Y = y_j]$  sont indépendants.
- Lorsque les variables sont indépendantes on peut construire la loi conjointe avec les lois marginales.
- Pour montrer que deux variables ne sont pas indépendantes, on cherche un couple  $(i, j)$  :

$$\mathbb{P}([X = x_i] \cap [Y = y_j]) \neq \mathbb{P}([X = x_i])\mathbb{P}([Y = y_j])$$

- Sur l'EX 1  $X$  et  $Y$  ne sont pas indépendantes car :
- Sur l'EX 2  $X$  et  $Y$  ne sont pas indépendantes car :

**12.1.5 Lois conditionnelles.**

**Définition :**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

- Pour tout  $x \in X(\Omega)$  tel que  $\mathbb{P}([X = x]) \neq 0$  ,  
 La loi conditionnelle de  $Y$  sachant  $[X = x]$  est l'application  $Y(\Omega) \rightarrow [0, 1]$   
 $y \mapsto \mathbb{P}_{[X=x]}([Y = y])$
- Pour tout  $y \in Y(\Omega)$  tel que  $\mathbb{P}([Y = y]) \neq 0$  ,  
 La loi conditionnelle de  $X$  sachant  $[Y = y]$  est l'application  $X(\Omega) \rightarrow [0, 1]$   
 $x \mapsto \mathbb{P}_{[Y=y]}([X = x])$

**Propositions :** (*Loi conjointe / lois conditionnelles*).

$$\forall (i, j) \in I \times J, \quad \mathbb{P}([X = x_i] \cap [Y = y_j]) = \mathbb{P}([X = x_i]) \times \mathbb{P}_{[X=x_i]}([Y = y_j])$$

$$\forall (i, j) \in I \times J, \quad \mathbb{P}([X = x_i] \cap ([Y = y_j])) = \mathbb{P}([Y = y_j]) \times \mathbb{P}_{[Y=y_j]}([X = x_i])$$

**Justification :** *Cela découle directement de la définition d'une probabilité conditionnelle.*

**Propositions :** (*Lois marginales / lois conditionnelles*).

$$\forall i \in I, \quad \mathbb{P}([X = x_i]) = \sum_{j \in J} \mathbb{P}([Y = y_j]) \times \mathbb{P}_{[Y=y_j]}([X = x_i])$$

$$\forall j \in J, \quad \mathbb{P}([Y = y_j]) = \sum_{i \in I} \mathbb{P}([X = x_i]) \times \mathbb{P}_{[X=x_i]}([Y = y_j])$$

**Justification :** *On applique la formule des probabilités totales. (Version avec les probabilités conditionnelles).*

**Exemple classique** vu dans la feuille *Exo\_8*.

Une tortue pond  $N$  œufs. On suppose que  $N$  est un variable de Poisson de paramètre  $\lambda$ .  
 Chaque œuf pondu a la probabilité  $p \in ]0, 1[$  de survivre et ceci indépendamment des autres œufs.  
 On note  $X$  le nombre de survivants.

1. Pour  $i \in \mathbb{N}$ , donner la loi conditionnelle de  $X$  sachant  $(N = i)$ .
2. En déduire pour tout  $k \in \mathbb{N}$ , la probabilité  $P(X = k)$ .

1. Sachant  $(N = i)$ , l'expérience est constituée de  $i$  épreuves de Bernoulli identiques et indépendantes, le succès étant : "l'œuf survit" a pour probabilité  $p$  et  $X$  est le nombre de succès donc  $X$  donc

$$\text{Sachant } (N = i), \quad X \hookrightarrow \mathcal{B}(i, p)$$

2. Les valeurs prises par  $X$  sont les entiers naturels,  
 pour chaque  $k \in \mathbb{N}$ , on applique la formule des probabilités totales avec le système complet  $(N = i)_{i \in \mathbb{N}}$

$$P(X = k) = \sum_{i=0}^{+\infty} P(N = i)P_{N=i}(X = k)$$

$$\begin{aligned}
 &= \sum_{i=k}^{+\infty} \frac{\lambda^i}{i!} e^{-\lambda} \times \binom{i}{k} p^k q^{i-k} \quad (\text{on supprime des termes nuls}) \\
 &= \frac{e^{-\lambda}}{k!} \lambda^k p^k \sum_{i=k}^{+\infty} \frac{(\lambda q)^{i-k}}{(i-k)!} \\
 &= \frac{e^{-\lambda}}{k!} \lambda^k p^k \sum_{i=0}^{+\infty} \frac{(\lambda q)^i}{i!} \quad (\text{Changement d'indice}) \\
 &= \frac{e^{-\lambda}}{k!} \lambda^k p^k e^{\lambda q} \quad (\text{série exponentielle}) \\
 &= \frac{(\lambda p)^k}{k!} e^{-\lambda p}
 \end{aligned}$$

$$X \hookrightarrow \mathcal{P}(\lambda p)$$

**12.1.6 Recherche de la loi de  $u(X, Y)$  sur des exemples.**

**Proposition.**

Soient  $X$  et  $Y$  deux variables aléatoires de  $(\Omega, \mathcal{F}, \mathbb{P})$  et  $u$  une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$ ,  
 Si la fonction  $u$  est définie sur  $X(\Omega) \times Y(\Omega)$  et les variables aléatoires  $X$  et  $Y$  sont discrètes  
 alors  $u(X, Y)$  est une variable aléatoire discrète de  $(\Omega, \mathcal{F}, \mathbb{P})$

**Remarques :**  $u(X, Y)$  est la variable aléatoire qui à tout  $\omega \in \Omega$  associe le réel  $u(X(\omega), Y(\omega))$ .

**Exemples :**  $X + Y$ ,  $XY$ ,  $\min(X, Y)$  et  $\max(X, Y)$ .

**En pratique.**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes et  $u$  une fonction définie sur  $X(\Omega) \times Y(\Omega)$ ,  
 on note  $Z = u(X, Y)$ ,  
 alors pour tout  $z \in Z(\Omega)$ ,  $\mathbb{P}([Z = z]) = \sum_{i \in I} \sum_{j \in J} \mathbb{P}([X = x_i] \cap [Y = y_j] \cap [Z = z])$

**Justification :** On applique juste la formule des probabilités totales.

On reprend l'**EX 1** de la feuille\_Cours\_13

$x \backslash Y$	0	1	2	3
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$
4	0	$\frac{3}{16}$	0	$\frac{3}{16}$
5	0	0	$\frac{6}{32}$	$\frac{6}{32}$

La loi de  $Z = X + Y$  :

$z$	3	4	5	6	7	8
$P(Z = z)$						

La loi de  $W = \max(X - 3, Y)$  :

$w$	0	1	2	3
$P(W = w)$				

**Un autre exemple :**

On considère un couple de variables aléatoires  $(X, Y)$  à valeurs dans  $\mathbb{N}^* \times \mathbb{N}^*$  dont la loi conjointe est donnée par

$$\forall (i, j) \in \mathbb{N}^{*2}, \quad \mathbb{P}(X = i, Y = j) = \frac{1}{2^{i+j}}.$$

Quelle est la loi de  $Z = X - Y$  ?

### 12.1.7 Espérance de $u(X, Y)$ .

Conformément au programme on ne donne le théorème de transfert que pour deux variables aléatoires discrètes finies.

**Théorème :** (Théorème de transfert)

Soient  $X$  et  $Y$  deux variables aléatoires de  $(\Omega, \mathcal{F}, \mathbb{P})$  et  $u$  une fonction de  $X(\Omega) \times Y(\Omega)$  dans  $\mathbb{R}$ ,  
 Lorsque  $X(\Omega)$  et  $Y(\Omega)$  sont finis :

la variable aléatoire  $u(X, Y)$  admet (toujours) une espérance qui vérifie :

$$E(u(X, Y)) = \sum_{i=1}^n \sum_{j=1}^m u(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j])$$

### Exemples

## 12.2 Covariance. Variance d'une somme.

### 12.2.1 Définition.

**Définition :**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,

Si  $X, Y$  admettent une variance alors  $(X, Y)$  admet une covariance donnée par :

$$\text{cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) \quad \text{ou} \quad \text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

**Démonstration.**

$$\begin{aligned} \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) &= \mathbb{E}\left(XY - \mathbb{E}(X)Y - \mathbb{E}(Y)X + \mathbb{E}(X)\mathbb{E}(Y)\right) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

**Remarques :**

- ① Quand  $\text{cov}(X, Y) = 0$  on dit qu'elles sont décorrélées.
- ② Quand  $\text{cov}(X, Y) < 0$ , en moyenne quand la valeur de  $X$  est grande, celle de  $Y$  a tendance à être petite.
- ③ Quand  $\text{cov}(X, Y) > 0$ , en moyenne quand la valeur de  $X$  est grande, celle de  $Y$  a tendance à être grande.
- ④ On peut remplacer la condition d'existence " $X, Y$  admettent une variance" par (au choix)
  - $X$  et  $Y$  admettent un moment d'ordre 2.
  - $X, Y$  et  $XY$  admettent une espérance.
- ⑤ Quand l'ensemble des valeurs est fini il n'y a pas de problème d'existence.
- ⑥ Pour  $a$  une constante :  $\text{cov}(X, a) = 0$  et  $\text{cov}(a, X) = 0$

**Exemple EX 1 :**

$X \backslash Y$	0	1	2	3	
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$	$\frac{1}{4}$
4	0	$\frac{3}{16}$	0	$\frac{3}{16}$	$\frac{3}{8}$
5	0	0	$\frac{6}{32}$	$\frac{6}{32}$	$\frac{3}{8}$
	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{2}$	1

**Exemple EX 2**

La loi conjointe de  $(X, Y)$  est :  $(X, Y)(\Omega) \subset (\mathbb{N}^*)^2$  et  $\forall (k_1, k_2) \in (\mathbb{N}^*)^2, \mathbb{P}((X = k_1) \cap (Y = k_2)) = q^{k_2-2} p^2 \mathbb{1}_{k_1 < k_2}$ ,  $X$  suit la loi géométrique de paramètre  $p$  et  $Y(\Omega) = \llbracket 2; +\infty \llbracket$  et pour tout  $k \geq 2, P(Y = k) = (k - 1)q^{k-2} p^2$   
 On pourra utiliser la remarque :

En notant  $Z$  le nombre d'épreuves entre le 1er et le 2-ième succès on a :  $X = Y + Z$  et  $Y$  et  $Z$  indépendantes.

**Proposition.**

Soit  $X$  une variable aléatoire discrète de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  
 Si  $X$  admet une variance alors  $\text{cov}(X, X) = V(X)$

**Remarque :** A mettre en parallèle avec :  $\langle u|u \rangle = \|u\|^2$  vue dans le cours : "Produit scalaire dans  $\mathbb{R}^n$ ".

**12.2.2 Propriétés.**

**Propriétés :** "C'est une forme bilinéaire positive et symétrique".

Soient  $X, Y$  et  $Z$  trois variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$  admettant une variance et  $a, b$  deux réels quelconques,

<p>❶ <math>\text{cov}(X, X) \geq 0</math></p> <p>❸ <math>\text{cov}(X, aY + bZ) = a \text{cov}(X, Y) + b \text{cov}(X, Z)</math></p>	<p>❷ <math>\text{cov}(X, Y) = \text{cov}(Y, X)</math></p> <p>❹ <math>\text{cov}(aX + bY, Z) = a \text{cov}(X, Z) + b \text{cov}(Y, Z)</math></p>
--	--

**Démonstration.** de ❸

$$\begin{aligned} \text{cov}(X, aY + bZ) &= E(X(aY + bZ)) - E(X)E(aY + bZ) \\ &= E(aXY + bXZ) - E(X)(aE(Y) + bE(Z)) \quad (\text{linéarité de l'espérance}) \\ &= aE(XY) + bE(XZ) - aE(X)E(Y) - bE(X)E(Z) \\ &= a(E(XY) - E(X)E(Y)) + b(E(XZ) - E(X)E(Z)) \\ &= a \text{cov}(X, Y) + b \text{cov}(X, Z) \end{aligned}$$

**Proposition :** (complément)

$$\text{cov} \left( \sum_{i=1}^n a_i X_i, \sum_{j=1}^{n'} b_j Y_j \right) = \sum_{i=1}^n \sum_{j=1}^{n'} a_i b_j \text{cov}(X_i, Y_j)$$

**Remarque :**

- A mettre en parallèle avec :  $\left( \sum_{i=1}^n \alpha_i \right) \left( \sum_{j=1}^m \beta_j \right)$  vue dans le cours : "Somme double".
- A mettre en parallèle avec :  $\left\langle \sum_{i=1}^p u_i ; \sum_{j=1}^{p'} v_j \right\rangle$  vue dans le cours : "Produit scalaire dans  $\mathbb{R}^n$ ".

**Théorème :** (complément)

Si  $X, Y$  admettent une variance alors  $\text{cov}(X, Y)^2 \leq V(X)V(Y)$

**Démonstration.** On se place sous l'hypothèse  $V(X) \neq 0$  et on pose  $P(\lambda) = V(\lambda X + Y)$

$$\begin{aligned} P(\lambda) &= V(\lambda X + Y) \\ &= \lambda^2 V(X) + 2\lambda \text{cov}(X, Y) + V(Y) \end{aligned}$$

et comme  $V(X) \neq 0, P$  est un polynôme de degré 2 dont le discriminant est :  $\Delta = 4\text{cov}(X, Y)^2 - 4V(X)V(Y)$ ,  $P(\lambda)$  est la variance d'un variable aléatoire donc  $\forall x \in \mathbb{R}, P(x) \geq 0$ , donc  $P$  a au plus une racine réelle et ainsi  $\Delta \leq 0$  ou encore

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

**Remarque :** A mettre en parallèle avec : l'inégalité de Cauchy-Schwarz vue dans : "Produit scalaire dans  $\mathbb{R}^n$ ".

### 12.2.3 Variance d'une somme.

#### Théorème

Soient  $X$  et  $Y$  deux variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  
 Si  $X$  et  $Y$  admettent une variance alors  $X + Y$  admet une variance et :

$$V(X + Y) = V(X) + 2 \operatorname{cov}(X, Y) + V(Y)$$

#### Démonstration.

$$\begin{aligned} V(X + Y) &= \operatorname{cov}(X + Y, X + Y) \\ &= \operatorname{cov}(X, X + Y) + \operatorname{cov}(Y, X + Y) \\ &= \operatorname{cov}(X, X) + \operatorname{cov}(X, Y) + \operatorname{cov}(Y, X) + \operatorname{cov}(Y, Y) \end{aligned}$$

#### Remarques :

- A mettre en parallèle avec :  $\|u + v\|^2 = \|u\|^2 + 2 \langle u, v \rangle + \|v\|^2$  vue dans : "Produit scalaire dans  $\mathbb{R}^n$ "
- On peut en déduire que  $\sigma(X + Y) \leq \sigma(X) + \sigma(Y)$

#### Corollaire

Pour  $a$  et  $b$  deux réels,

$$V(aX + bY) = a^2V(X) + 2ab \operatorname{cov}(X, Y) + b^2V(Y)$$

**En effet.** On applique le théorème avec  $X \leftarrow aX$  et  $Y \leftarrow bY$

#### Proposition. (Complément)

Pour  $n$  variables aléatoires discrètes  $(X_k)_{1 \leq k \leq n}$   
 Si  $(X_k)_{1 \leq k \leq n}$  admettent des variances, alors  $\sum_{i=1}^n X_i$  admet une variance et

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \operatorname{cov}(X_i, X_j)$$

Pour tout  $(a_1, \dots, a_n) \in \mathbb{R}^n$ ,

$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum_{1 \leq i < j \leq n} a_i a_j \operatorname{cov}(X_i, X_j)$$

**En effet.**

#### Remarque :

Si  $Y$  est une combinaison linéaire de variables aléatoires admettant chacune une variance alors  $Y$  admet une variance.

### 12.2.4 Cas particuliers des variables aléatoires indépendantes.

#### Théorèmes.

Soient  $X$  et  $Y$  deux variables aléatoires discrètes admettant des variances,

- ❶ Si  $X$  et  $Y$  sont indépendantes alors  $E(XY) = E(X)E(Y)$
- ❷ Si  $X$  et  $Y$  sont indépendantes alors  $\operatorname{cov}(X, Y) = 0$
- ❸ Si  $X$  et  $Y$  sont indépendantes alors  $V(X + Y) = V(X) + V(Y)$ .

#### Théorème.

Soient  $n$  un entier naturel supérieur ou égal à 2 et  $(X_k)_{1 \leq k \leq n}$  des variables aléatoires discrètes  
 Si les  $X_k$  sont (mutuellement) indépendantes et admettent chacune une variance alors

$$\sum_{i=1}^n X_i \text{ admet une variance et } V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i)$$

**Remarques :**

- Ici l'indépendance deux à deux des variables aléatoires suffit. (*complément*)
- A mettre en parallèle avec : le théorème de Pythagore vue dans : "Produit scalaire dans  $\mathbb{R}^n$ ".

**12.2.5 Coefficient de corrélation.**

**Définition :**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes de  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  
 Si  $X, Y$  admettant des variances non nulles alors  
 on appelle **coefficient corrélation linéaire** du couple  $(X, Y)$  le réel :  $\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$

**Théorème.** (*complément*)

Pour toutes variables aléatoires discrètes  $X$  et  $Y$  admettant des variances non nulles on a :

- ❶  $-1 \leq \rho_{X,Y} \leq 1$
- ❷  $\rho_{X,Y} = -1$  ou  $\rho_{X,Y} = 1 \iff \exists(a, b) \in \mathbb{R}^2 : \mathbb{P}([Y = aX + b]) = 1$

**Démonstration :**

❶ est une conséquence directe du théorème 12.2.2.

Pour ❷ : en utilisant les notations du théorème 12.2.2 :

$$\begin{aligned}
 (\text{cov}(X, Y))^2 = V(X)V(Y) &\iff \Delta = 0 \\
 &\iff P \text{ possède une racine réelle} && \text{car } \Delta \leq 0 \\
 &\iff \exists \lambda \in \mathbb{R} : V(\lambda X + Y) = 0 \\
 &\iff \exists \lambda \in \mathbb{R} : \lambda X + Y \text{ est quasi-certaine} \\
 &\iff \exists(\lambda, b) \in \mathbb{R}^2 : \mathbb{P}([\lambda X + Y = b]) = 1 \\
 &\iff \exists(a, b) \in \mathbb{R}^2 : \mathbb{P}([Y = aX + b]) = 1
 \end{aligned}$$

$$(\text{cov}(X, Y))^2 = V(X)V(Y) \iff \exists(a, b) \in \mathbb{R}^2, \mathbb{P}([Y = aX + b]) = 1$$

**Remarque :**

Si  $Y = aX + b$ , alors les valeurs de  $a$  et  $b$  sont faciles à obtenir :

$$a = \frac{\text{cov}(X, Y)}{V(X)} \quad \text{et} \quad b = E(Y) - aE(X)$$

**En effet :** Si  $Y = aX + b$ ,

- on a d'une part (*linéarité*)  $b = E(Y) - aE(X)$
- d'autre part :  $Y - aX$  et  $X$  sont indépendantes car  $Y - aX$  certaine donc  $\text{cov}(X, Y - aX) = 0$

on obtient 
$$a = \frac{\text{cov}(X, Y)}{V(X)}$$

**Remarque.**

Ces formules sont à mettre en parallèle des coefficients de la droite de régression du cours de statistique.

*(voir annexe C)*

# Annexes

## Fonctions indicatrices.

### Définition :

Soient  $E$  un ensemble et  $A$  une partie de  $E$ ,

On appelle **fonction indicatrice** de  $A$ , (notée :  $\mathbb{1}_A$ ) l'application de :  $E \longrightarrow \{0, 1\}$

$$x \longmapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

**Remarque :**  $\mathbb{1}_A$  peut prendre juste deux valeurs 0 et 1 donc  $\forall x \in E, \mathbb{1}_A(x) = 1 \iff x \in A$ .

### Une autre notation.

Soit  $E$  un ensemble et  $P(x)$  une propriété dépendant d'un élément  $x$  de  $E$ .

en notant  $A = \{x \in E \mid P(x)\}$  pour tout  $x$  de  $E$ , on note :  $\mathbb{1}_{P(x)} = \mathbb{1}_A(x)$

**Remarque :** Si  $P(x)$  vraie alors  $\mathbb{1}_{P(x)} = 1$  sinon  $\mathbb{1}_{P(x)} = 0$ .

### Cours sur les ensembles finis :

Soit  $E$  un ensemble fini,

❶ Soit  $A$  une partie de  $E$ ,  $\text{card}(A) = \sum_{x \in E} \mathbb{1}_A(x)$

❷ Soient  $A$  et  $B$  deux parties de  $E$ ,  $\text{card}(A \cap B) = \sum_{x \in A} \mathbb{1}_B(x)$

❸ Soient  $P(x)$  une propriété dépendant d'un élément  $x$  de  $E$ ,  $\sum_{x \in E} \mathbb{1}_{P(x)} = \text{card}(\{x \in E \mid P(x)\})$

### Cours sur l'indépendance des événements et des VAR :

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un système probabilisé.

❶ Si  $A$  est un événement alors  $\mathbb{1}_A$  est une variable aléatoire suivant la loi de Bernoulli  $\mathcal{B}(\mathbb{P}(A))$

❷ (Complément) Soit  $(A_k)_{1 \leq k \leq n}$  une famille d'événements,

Les événements  $A_k$  sont mutuellement indépendants

si, et seulement si, les variables aléatoires  $\mathbb{1}_{A_k}$  sont indépendantes.

### Autres relations, autres remarques :

- Produit de convolution discret.

$$\sum_{i=1}^n \mathbb{1}_{1 \leq k-i \leq n} = \sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_{i+j=k} = \text{card}(\{(i, j) \in \llbracket 1, n \rrbracket^2 \mid i+j=k\})$$

*Détail du raisonnement :*

$$\begin{aligned} \sum_{j=1}^n \mathbb{1}_{i+j=k} &= \sum_{j=1}^n \mathbb{1}_{j=k-i} \\ &= \begin{cases} 1 & \text{si } k-i \in \llbracket 1; n \rrbracket \\ 0 & \text{sinon} \end{cases} \\ &= \mathbb{1}_{1 \leq k-i \leq n} \end{aligned}$$

- Inversion d'une somme double triangulaire

$$\sum_{i=1}^n \sum_{j=1}^i a_{i,j} = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} \mathbb{1}_{j \leq i} = \sum_{j=1}^n \sum_{i=1}^n a_{i,j} \mathbb{1}_{j \leq i} = \sum_{j=1}^n \sum_{i=j}^n a_{i,j}$$

# Minimum. Maximum. Somme

## B.1 Variables aléatoires à valeurs dans $\mathbb{N}$

- Pour le max de deux VAR indépendantes à valeurs dans  $\mathbb{N}$  :

$$\begin{aligned} \forall k \in \mathbb{N}, \quad P(\max(X, Y) \leq k) &= P((X \leq k) \cap (Y \leq k)) \\ &= P(X \leq k) P(Y \leq k) \quad \text{car } X \text{ et } Y \text{ sont indépendantes} \end{aligned}$$

et

$$P(Z = k) = P(Z \leq k) - P(Z \leq k - 1)$$

- Pour le max de deux VAR indépendantes à valeurs dans  $\mathbb{N}$  :

$$\begin{aligned} \forall k \in \mathbb{N}, \quad P(\min(X, Y) > k) &= P((X > k) \cap (Y > k)) \\ &= P(X > k) P(Y > k) \quad \text{car } X \text{ et } Y \text{ sont indépendantes} \end{aligned}$$

et

$$P(Z = k) = P(Z > k - 1) - P(Z > k)$$

- Pour la somme deux VAR indépendantes à valeurs dans  $\mathbb{N}$  :

$$\begin{aligned} \forall k \in \mathbb{N}, \quad P(X + Y = k) &= \sum_{i=0}^{+\infty} P((X = i) \cap (X + Y = k)) \quad (\text{FPT avec le SCE } (X = i)_{i \in \mathbb{N}}) \\ &= \sum_{i=0}^k P((X = i) \cap (Y = k - i)) \quad (\text{On enlève des termes nuls}) \\ &= \sum_{i=0}^k P(X = i) \times P(Y = k - i) \quad \text{car } X \text{ et } Y \text{ sont indépendantes} \end{aligned}$$

$$P(X + Y = k) = \sum_{i=0}^k P(X = i) \times P(Y = k - i) \quad (\text{produit de convolution discret})$$

- Si  $X$  suit la loi  $\mathcal{G}(p)$  alors  $P(X \leq k) = 1 - q^k$  et  $P(X > k) = q^k$ . (où  $q = 1 - p$ )

- Le minimum de VAR suivant des lois géométriques suit une géométrique.

- Si  $X$  suit la loi  $\mathcal{U}(\llbracket 1, n \rrbracket)$  alors pour tout  $k \in \llbracket 0, n \rrbracket$ ,  $P(X \leq k) = \frac{k}{n}$  et  $P(X > k) = 1 - \frac{k}{n}$

- Pour l'espérance :

$$X + Y = \min(X, Y) + \max(X, Y)$$

- Si  $X$  est à valeurs dans  $\llbracket 1, n \rrbracket$ ,  $E(X) = \sum_{k=1}^n P(X \geq k)$

- Si  $X$  est à valeurs dans  $\mathbb{N}$ ,  $E(X) = \sum_{k=1}^{+\infty} P(X \geq k)$  (sous réserve de l'existence de cette somme)

**Remarques :**

- La relation  $X + Y = \min(X, Y) + \max(X, Y)$  est utile pour le calcul des espérances car elle donne une relation entre les trois espérances :  $E(X + Y) = E(\min(X, Y)) + E(\max(X, Y))$
- Les formules  $E(X) = \sum_{k=1}^n P(X \geq k)$  et  $E(X) = \sum_{k=1}^{+\infty} P(X \geq k)$  ne sont pas au programme de BCPST.
- Une idée de la démonstration de  $E(X) = \sum_{k=1}^{+\infty} P(X \geq k)$  : (Il manque des théorèmes pour justifier ces calculs)

$$\begin{aligned}
 \sum_{k=1}^{+\infty} P(X \geq k) &= \sum_{k=1}^{+\infty} \sum_{i=1}^{+\infty} P(X = i) \mathbb{1}_{i \geq k} \\
 &= \sum_{i=1}^{+\infty} \sum_{k=1}^{+\infty} P(X = i) \mathbb{1}_{i \geq k} \\
 &= \sum_{i=1}^{+\infty} \sum_{k=1}^i P(X = i) \\
 &= \sum_{i=1}^{+\infty} iP(X = i) \\
 &= \sum_{i=0}^{+\infty} iP(X = i) = E(X)
 \end{aligned}$$

- On a aussi vu dans ces exercices les égalités suivantes :

$$\sum_{i=1}^n \mathbb{1}_{1 \leq k-i \leq n} = \sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_{i+j=k} = \text{card}(\{(i, j) \in \llbracket 1, n \rrbracket^2 \mid i + j = k\})$$

**B.2 Variables aléatoires à densité**

- Pour  $Z = \max(X, Y)$  où  $X$  et  $Y$  sont deux VAR indépendantes à densité :

$$\begin{aligned}
 \forall x \in \mathbb{R}, \quad F_Z(x) &= P(\max(X, Y) \leq x) \\
 &= P((X \leq x) \cap (Y \leq x)) \\
 &= P(X \leq x) P(Y \leq x) \quad \text{car } X \text{ et } Y \text{ sont indépendantes} \\
 &= F_X(x) F_Y(x)
 \end{aligned}$$

- Pour  $Z = \min(X, Y)$  où  $X$  et  $Y$  sont deux VAR indépendantes à densité :

$$\begin{aligned}
 \forall x \in \mathbb{R}, \quad F_Z(x) &= 1 - P(\min(X, Y) > x) \\
 &= 1 - P((X > x) \cap (Y > x)) \\
 &= 1 - P(X > x) P(Y > x) \quad \text{car } X \text{ et } Y \text{ sont indépendantes} \\
 &= 1 - (1 - F_X(x))(1 - F_Y(x))
 \end{aligned}$$

- Le minimum de VAR suivant un loi exponentielle est une loi exponentielle.
- (Produit de convolution) Si  $X$  et  $Y$  sont indépendantes alors :
  - $X + Y$  est une variable à densité.
  - $x \mapsto \int_{-\infty}^{+\infty} f(t)g(x-t) dt$  est une densité de  $X + Y$ .
- La somme de gaussiennes indépendantes est une gaussienne.
- Somme de deux VAR indépendantes à valeurs dans  $\mathbb{R}_+$ , (comme par exemple la loi exponentielle)

$$\text{Pour } x \in \mathbb{R}, \quad \int_{-\infty}^{+\infty} f(t) \cdot \mathbb{1}_{\mathbb{R}_+}(t) \mathbb{1}_{\mathbb{R}_+}(x-t) dt = \int_{-\infty}^{+\infty} f(t) \cdot \mathbb{1}_{t>0} \mathbb{1}_{t<x} dt$$

$$= \int_{-\infty}^{+\infty} f(t) \cdot \mathbb{1}_{0 < t < x} dt$$

donc

Pour tout  $x \in \mathbb{R}$ , 
$$\int_{-\infty}^{+\infty} f(t) \cdot \mathbb{1}_{\mathbb{R}_+^*}(t) \mathbb{1}_{\mathbb{R}_+^*}(x-t) dt = \mathbb{1}_{\mathbb{R}_+^*}(x) \cdot \int_0^x f(t) dt$$

- Pour deux VAR indépendantes à valeurs respectives sur  $[a, b]$  et  $[c, d]$ . (comme par exemple des lois uniformes)

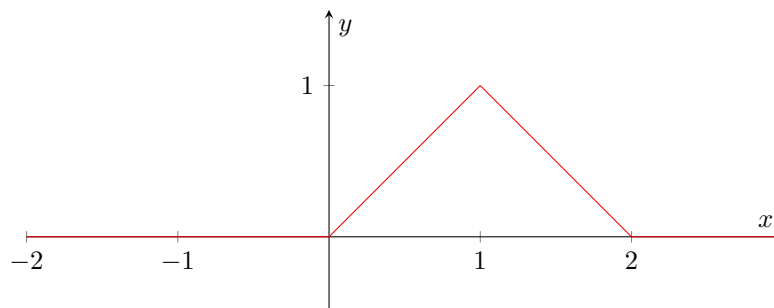
On suppose  $b - a \leq d - c$ ,

$$\begin{aligned} \varphi(x) &= \int_{-\infty}^{+\infty} f(t) \mathbb{1}_{[a,b]}(t) \cdot \mathbb{1}_{[c,d]}(x-t) dt = \int_{-\infty}^{+\infty} f(t) \mathbb{1}_{[a,b]}(t) \cdot \mathbb{1}_{[x-d, x-c]}(t) dt \\ &= \int_{-\infty}^{+\infty} f(t) \mathbb{1}_{[a,b] \cap [x-d, x-c]}(t) dt \end{aligned}$$

$x$	$-\infty$	$a + c$	$b + c$	$a + d$	$b + d$	$+\infty$
$\varphi(x)$	0	$\int_a^{x-c} f(t) dt$	$\int_a^b f(t) dt$	$\int_{x-d}^b f(t) dt$	0	0

- Un exemple avec  $X$  et  $Y$  sont indépendantes et de même densité :  $f : t \mapsto \mathbb{1}_{[0,1]}(t)$ .

$$x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 < x \leq 1 \\ 2 - x & \text{si } 1 < x \leq 2 \\ 0 & \text{si } x \geq 2 \end{cases} \text{ est une densité de } X + Y$$



- (Vu dans un DS, mais là aussi hors programme)

Pour  $X$  une variable aléatoire (quelconque) à valeurs positives admettant une espérance on a la formule suivante :

$$E(X) = \int_0^{+\infty} P(X > x) dx$$

# Régression linéaire

Cette feuille regroupe l'essentiel de ce que nous avons vu sur les sujets MMI 2017 et 2019 dans la feuille 24 d'informatique.

Cette notion apparait dans plusieurs chapitres du programme de mathématiques :

- Cours de statistiques : Coefficients de la droite de régression linéaire d'une série double de données.
- Projection orthogonale : Calcul de la distance d'un vecteur  $u$  à un plan  $\text{Vect}(v_1, v_2)$ .
- Fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$  : Calcul d'un point critique de  $(a, b) \mapsto F(a, b)$ .

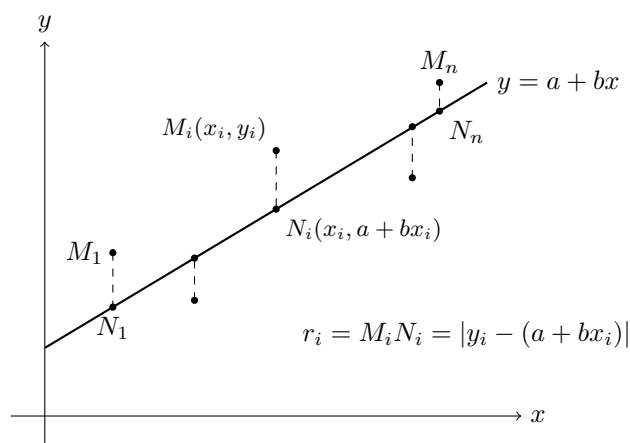
## Objectif :

Déterminer le meilleur ajustement affine au sens des moindres carrés du nuage de points :

$$((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)) \quad (\text{on suppose que les } x_i \text{ ne sont pas tous égaux})$$

Autrement dit :

on veut trouver la droite d'équation  $y = a + bx$  qui minimise la somme  $\sum_{i=1}^n r_i^2$  où les  $r_i$  sont définis sur la figure ci-dessous :



## Notations usuelles :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad \overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2 \quad \overline{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i$$

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2 \quad (\text{variance empirique de } (x_i)) \quad \text{et} \quad \text{cov}(x, y) = \overline{xy} - \bar{x} \bar{y} \quad (\text{covariance empirique de } (x_i, y_i))$$

**Remarque :** les  $x_i$  ne sont pas tous égaux donc  $\sigma_x^2 \neq 0$ .

## Résultat fondamental

Les coefficients de cette droite appelée droite de régression linéaire sont :  $b = \frac{\text{cov}(x, y)}{\sigma_x^2}$  et  $a = \bar{y} - b \bar{x}$

• **Première approche.**

On note  $u = (y_1, \dots, y_n)$ ,  $v_1 = (1, \dots, 1)$  et  $v_2 = (x_1, \dots, x_n)$

(la famille  $(v_1, v_2)$  est une famille libre car les  $x_i$  ne sont pas tous égaux)

on note  $H$  le plan  $\text{vect}(v_1, v_2)$ .

On sait que la distance de  $u$  à  $H$  est le minimum de  $\left\{ \|u - av_1 - bv_2\| \mid (a, b) \in \mathbb{R}^2 \right\}$  et ce minimum est atteint lorsque  $av_1 + bv_2$  est le projeté orthogonal de  $u$  sur  $H$ .

Or, pour tout  $(a, b) \in \mathbb{R}^2$ ,  $\|u - av_1 - bv_2\|^2 = \sum_{i=1}^n (y_i - (a + bx_i))^2 = \sum_{i=1}^n r_i^2$

Ainsi, minimiser  $\sum_{i=1}^n r_i^2$  revient à projeter  $u$  sur  $H$ . Autrement dit trouver  $(a, b)$  tel que :  $p_H(u) = av_1 + bv_2$

$$\begin{aligned} av_1 + bv_2 = p_H(u) &\iff u - (av_1 + bv_2) \in H^\perp \\ &\iff \begin{cases} u - (av_1 + bv_2) \perp v_1 \\ u - (av_1 + bv_2) \perp v_2 \end{cases} \\ &\iff \begin{cases} \langle u - (av_1 + bv_2) | v_1 \rangle = 0 \\ \langle u - (av_1 + bv_2) | v_2 \rangle = 0 \end{cases} \\ &\iff \begin{cases} \langle u | v_1 \rangle = a \langle v_1 | v_1 \rangle + b \langle v_2 | v_1 \rangle \\ \langle u | v_2 \rangle = a \langle v_1 | v_2 \rangle + b \langle v_2 | v_2 \rangle \end{cases} \\ &\iff \begin{pmatrix} \langle v_1 | v_1 \rangle & \langle v_1 | v_2 \rangle \\ \langle v_2 | v_1 \rangle & \langle v_2 | v_2 \rangle \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \langle u | v_1 \rangle \\ \langle u | v_2 \rangle \end{pmatrix} \end{aligned}$$

on remarque que

$$\langle v_1 | v_1 \rangle = n \quad \langle v_1 | v_2 \rangle = n\bar{x} \quad \langle v_2 | v_2 \rangle = n\bar{x}^2 \quad \langle u | v_1 \rangle = n\bar{y} \quad \langle u | v_2 \rangle = n\bar{x}\bar{y}$$

on obtient le système :

$$\begin{aligned} \begin{pmatrix} n & n\bar{x} \\ n\bar{x} & n\bar{x}^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} &= \begin{pmatrix} n\bar{y} \\ n\bar{x}\bar{y} \end{pmatrix} \iff \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \bar{x}^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \bar{x}\bar{y} \end{pmatrix} \\ &\iff \begin{cases} a + \bar{x}b = \bar{y} \\ \bar{x}a + \bar{x}^2b = \bar{x}\bar{y} \end{cases} \\ &\iff \begin{cases} a + \bar{x}b = \bar{y} \\ (\bar{x}^2 - (\bar{x})^2)b = \bar{x}\bar{y} - \bar{x}\bar{y} \end{cases} \quad L_2 \leftarrow L_2 - \bar{x}L_1 \end{aligned}$$

On obtient alors (car  $\sigma_x^2 \neq 0$ ) :

$$av_1 + bv_2 = p_H(u) \iff \begin{cases} a = \bar{y} - b\bar{x} \\ b = \frac{\text{cov}(x, y)}{\sigma_x^2} \end{cases}$$

• **Deuxième approche.**

On définit  $F : (a, b) \mapsto \sum_{i=1}^n (y_i - a - bx_i)^2$ ,

On remarque que  $F(a, b) = \sum_{i=1}^n r_i^2$ , on cherche donc  $a$  et  $b$  tels que  $F(a, b)$  est minimal.

La première approche justifie que  $F$  admette un minimum donc on cherche le(s) point(s) critique(s) de  $F$ .

$F : (a, b) \mapsto \sum_{i=1}^n (y_i - a - bx_i)^2$  est polynomiale en  $a$  et en  $b$  donc elle admet des dérivées partielles.

$$\begin{aligned} \frac{\partial F}{\partial a}(a, b) &= \sum_{i=1}^n -2(y_i - a - bx_i) & \frac{\partial F}{\partial b}(a, b) &= \sum_{i=1}^n -2x_i(y_i - a - bx_i) \\ &= -2 \left( \sum_{i=1}^n y_i - a \sum_{i=1}^n 1 - b \sum_{i=1}^n x_i \right) & &= -2 \left( \sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i - b \sum_{i=1}^n x_i^2 \right) \end{aligned}$$

donc

$$\begin{aligned} \begin{cases} \frac{\partial F}{\partial a}(a, b) = 0 \\ \frac{\partial F}{\partial b}(a, b) = 0 \end{cases} &\iff \begin{cases} n\bar{y} - na - bn\bar{x} = 0 \\ \sum_{i=1}^n x_i y_i - an\bar{x} - b \sum_{i=1}^n x_i^2 = 0 \end{cases} \\ &\iff \begin{cases} a + \bar{x}b = \bar{y} \\ \bar{x}a + \overline{x^2}b = \overline{xy} \end{cases} \\ &\iff \begin{cases} a = \bar{y} - b\bar{x} \\ b = \frac{\text{cov}(x, y)}{\sigma_x^2} \end{cases} \end{aligned}$$

# Annexe — Équations différentielles autonomes d'ordre 1

## Plan de l'annexe

D.1	Modèle malthusien (croissance exponentielle) . . . . .	195
D.2	Modèle logistique (croissance avec saturation) . . . . .	196
D.3	Modèle de Gompertz . . . . .	197
D.4	Modèle de Lotka–Volterra (proie–prédateur) . . . . .	197
D.5	Lien avec l'informatique (méthode d'Euler) . . . . .	198
D.5.1	Modèle logistique (extrait de la correction de la feuille info_15) . . . . .	199
D.5.2	Modèle Lotka–Volterra . . . . .	199

Nous allons ici présenter des exemples vus cette année en devoir maison (DM4) et en informatique info\_15.

### Cadre d'étude.

On s'intéresse aux équations différentielles de la forme :

$$\forall t \in I, y'(t) = F(y(t)) \quad \text{où } F \text{ est une fonction continue.}$$

### Idée fondamentale.

Le taux de variation de  $y$  ne dépend que de la valeur de  $y$ .

### Conséquence importante.

L'étude qualitative repose sur le signe de  $F$  :

- si  $F(y) > 0$  :  $y$  est croissante ;
- si  $F(y) < 0$  :  $y$  est décroissante ;
- si  $F(y) = 0$  :  $y$  est constante.

### Points d'équilibre.

Les solutions constantes sont données par :

$$F(y_0) = 0$$

On les appelle *états d'équilibre*.

### Stabilité (idée intuitive).

- Si les solutions voisines se rapprochent de  $y_0$  : équilibre stable.
- Si elles s'en éloignent : équilibre instable.

## D.1 Modèle malthusien (croissance exponentielle).

### Modélisation.

On suppose que le taux de variation de la population est proportionnel à sa taille :

$$y'(t) = k y(t) \quad \text{où } k \in \mathbb{R} \text{ est une constante.}$$

**Résolution.**

Les solutions sont les fonctions :  $y : t \mapsto y_0 e^{kt}$  où  $y_0 = y(0)$ .

**Étude suivant le signe de  $k$ .**

- Si  $k > 0$  : croissance exponentielle.
- Si  $k = 0$  : population constante.
- Si  $k < 0$  : décroissance exponentielle (extinction).

**Temps caractéristique.**

Si  $k > 0$ , le temps de doublement vérifie :

$$y(t+T) = 2y(t) \implies T = \frac{\ln 2}{k}$$

**Lecture qualitative.**

- Plus la population est grande, plus elle croît vite.
- Il n'y a aucun frein à la croissance.

**Limites du modèle.** Ne tient pas compte des ressources limitées. Non réaliste à long terme.

**Interprétation biologique.** Modèle valable sur des temps courts : croissance bactérienne initiale, population sans contrainte.

## D.2 Modèle logistique (croissance avec saturation).

Voir le DM 4

**Modélisation.**

On suppose que la croissance est freinée lorsque la population devient grande :

$$y'(t) = k y(t) \left(1 - \frac{y(t)}{K}\right)$$

où  $k > 0$  est un taux de croissance et  $K > 0$  une capacité limite.

**Interprétation.**

- si  $y \ll K$  :  $y'(t) \approx ky(t)$  (comportement malthusien) ;
- si  $y \approx K$  : la croissance ralentit ;
- si  $y > K$  : la population décroît.

**Équilibres :**  $y'(t) = 0 \iff y = 0$  ou  $y = K$

**Stabilité :** 0 est instable et  $K$  est stable.

**Résolution.**

Faite dans la feuille info\_15 et dans les sujets : MMI 2022, MCR 2023 et ENS 2025.

Pour  $y_0 > 0$ , la solution est :  $y(t) = \frac{K}{1 + Ce^{-kt}}$  où  $C = \frac{K - y_0}{y_0}$ .

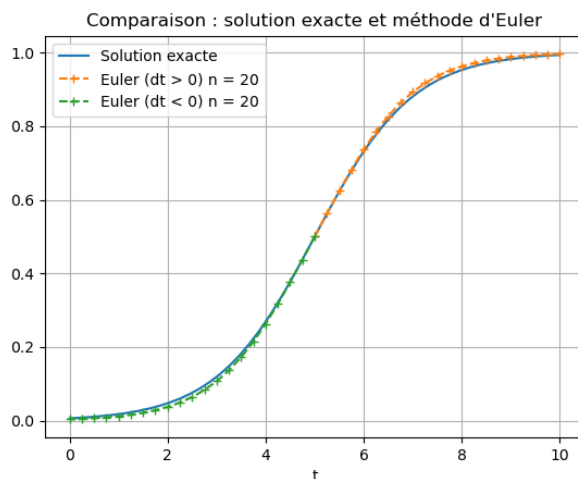


Figure extraite de la correction de la feuille info\_15

**Comportement asymptotique.**  $\lim_{t \rightarrow +\infty} y(t) = K$

**Point d'inflexion.** La croissance est maximale pour :  $y = \frac{K}{2}$  (changement de concavité de la courbe).

### D.3 Modèle de Gompertz.

Ce modèle n'a pas été vu cette année.

$$y'(t) = a y(t) \ln\left(\frac{K}{y(t)}\right)$$

**Équilibre :**  $y = K$ .

**Lecture qualitative :**

- si  $0 < y < K$  :  $\ln\left(\frac{K}{y}\right) > 0$  donc croissance ;
- si  $y > K$  :  $\ln\left(\frac{K}{y}\right) < 0$  donc décroissance ;
- $K$  est un équilibre stable.

**Solution.**

On se ramène à une équation différentielle linéaire en posant  $z(t) = \ln(y(t))$ .

$$y(t) = K \exp(-C e^{-at})$$

**Interprétation biologique :**

Modèle adapté aux phénomènes où la croissance est freinée de manière exponentielle  
(par exemple : croissance tumorale).

### D.4 Modèle de Lotka–Volterra (proie–prédateur).

**Modélisation.**

On considère deux populations :  $x(t)$  : effectif de la population de proies,  
 $y(t)$  : effectif de la population de prédateurs.

Le modèle s'écrit :

$$\begin{cases} x'(t) = ax(t) - bx(t)y(t) \\ y'(t) = -cy(t) + dx(t)y(t) \end{cases}$$

avec  $a, b, c, d > 0$ .

**Interprétation des coefficients.**

- $a$  : taux intrinsèque de reproduction des proies ;
- $b$  : taux de mortalité des proies due aux prédateurs ;
- $c$  : taux de mortalité intrinsèque des prédateurs ;
- $d$  : taux de reproduction des prédateurs liés aux proie mangées.

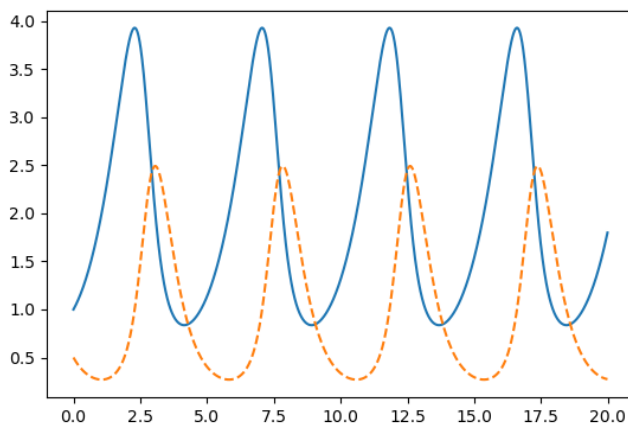


Figure extraite du sujet du DM 4

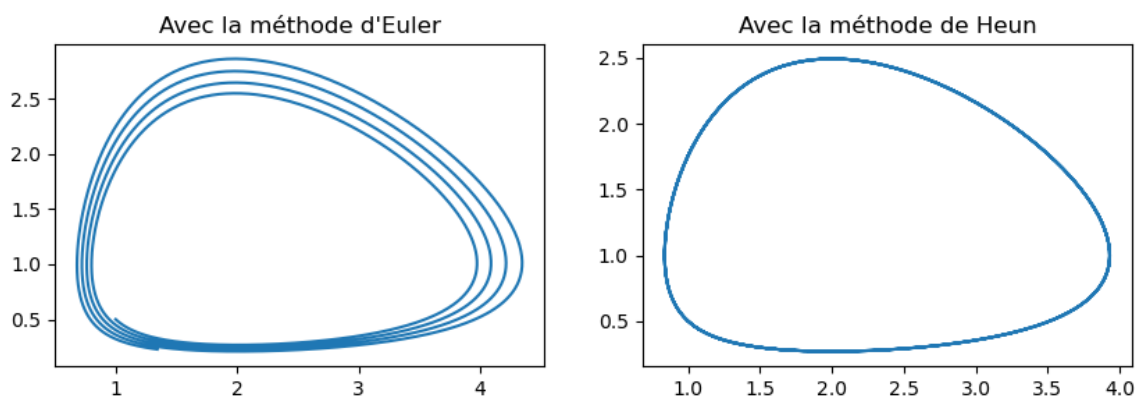


Figure extraite du sujet de la correction du DM 4

**Équilibres.**  $(x, y) = (0, 0)$  et  $\left(\frac{c}{d}, \frac{a}{b}\right)$

**Interprétation.** (Extrait de la correction du DM4)

L'effectifs des deux populations présentent des oscillations périodiques. Le nombre de lièvres augmente en premier, ce qui fait croître le nombre de Lynx. Le taux de prédation devient trop grand cela fait chuter le nombre de lièvre ce qui fait chuter à son tour le nombre de lynx.

## D.5 Lien avec l'informatique (méthode d'Euler).

Cette méthode permet d'approcher numériquement les solutions. Voir la feuille info\_13.

On cherche à tracer la courbe de l'unique solution du problème de Cauchy :

$$\forall t \in I, \quad y'(t) = F(y(t)) \quad \text{et} \quad y(t_0) = y_0$$

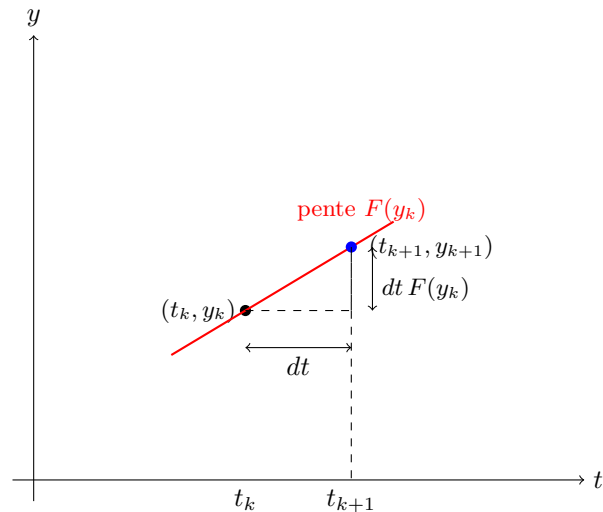
Pour cela on écrit l'approximation :

$$\frac{y(t+dt) - y(t)}{dt} \approx F(y(t)) \quad \text{ou encore} \quad y(t+dt) \approx y(t) + dt \times F(y(t))$$

La méthode pour obtenir une approximation de la courbe solution sur  $[t_0, t_f]$  consiste à :

- subdiviser  $[t_0, t_f]$  en  $n$  intervalles de même longueur  $dt = \frac{t_f - t_0}{n}$ ,
- de déterminer une approximation  $y_k$  des images  $y(t_k)$  avec la relation de récurrence suivante :

$$y_{k+1} = y_k + dt \cdot F(y_k)$$



### D.5.1 Modèle logistique (extrait de la correction de la feuille info\_15).

On applique ici la méthode d'Euler au problème différentiel suivant :

$$\text{sur } [0, 10], \quad \begin{cases} y'(t) = y(t) - y(t)^2 \\ y(5) = \frac{1}{2} \end{cases}$$

```
import matplotlib.pyplot as plt
import numpy as np
```

```
def euler(n, a, b):
    dt = (b-a)/n
    t = np.zeros(n+1)
    y = np.zeros(n+1)
    t[0], y[0] = 5, 0.5
    for k in range(n):
        t[k+1] = t[k] + dt
        y[k+1] = y[k] + dt * (y[k] - y[k]**2)    # Principe de la méthode d'Euler.
    return t, y
```

# tracé des solutions approchées

```
n = 20
t, y = euler(n,5,10)
plt.plot(t, y, '--+', label = f"Euler (dt > 0) n = {n}")
t, y = euler(n,5,0)
plt.plot(t, y, '--+', label = f"Euler (dt < 0) n = {n}")
```

```
plt.xlabel("t")
plt.ylabel("y")
```

```
plt.legend()
plt.grid(True)
plt.show()
```

### D.5.2 Modèle Lotka-Volterra

On s'intéresse au cas particulier :

$$\text{sur } [1, 9], \quad \begin{cases} x'(t) = 3x(t) - x(t)y(t) & x(1) = 1 \\ y'(t) = 2x(t)y(t) - y(t) & y(1) = 2 \end{cases}$$

```
N = 1000
ti, tf = 1, 9
dt = (tf-ti)/N
t = np.zeros(N)
x = np.zeros(N)
y = np.zeros(N)
t[0] = ti
x[0] = 1
y[0] = 2
for k in range(N-1):
    t[k+1] = t[k] + dt
    x[k+1] = x[k] + dt * (3*x[k] - x[k]*y[k] )
    y[k+1] = y[k] + dt*( 2*x[k]*y[k] - y[k])
plt.figure('Exemple 2')
plt.plot(t, x)
plt.plot(t, y)
plt.show()
```

# Statistiques transversales : test de conformité, variabilité biologique et validation d'un modèle

## Plan de l'annexe

E.1	Variabilité biologique (SVT)	201
E.2	Test de conformité à la moyenne (mathématiques)	202
E.2.1	Lien avec l'intervalle de confiance	202
E.3	Validation d'un modèle (physique)	202
E.3.1	Variante : comparaison de deux valeurs expérimentales	203
E.4	Incertitudes de type A et de type B	203
E.5	Méthode de Monte-Carlo	204
E.6	Lien avec les TIPE	204
E.7	Vocabulaire et présentation des résultats	204
E.7.1	Anglais scientifique	204
E.7.2	Le z-score comme fil conducteur	205
E.7.3	Présentation graphique des résultats	205
	Tableau de synthèse	205

*Les mêmes outils — moyenne, écart-type, intervalles — apparaissent en mathématiques, en SVT et en physique, mais avec des objectifs différents. Il est essentiel de bien identifier la question posée pour adopter la bonne démarche.*

*Ces trois approches coexistent notamment au cours des TIPE. Cette annexe a pour objectif d'en donner une vision unifiée et d'éviter les confusions fréquentes.*

## E.1 Variabilité biologique (SVT)

**Cadre.** On observe une grandeur sur un ensemble d'individus. Les valeurs sont naturellement dispersées.

La dispersion est une propriété du phénomène étudié, pas une erreur de mesure.

**Outils utilisés :** moyenne  $\bar{x}$ , écart-type  $s$ , histogramme, boîte à moustaches.

**Objectif :** *décrire* la variabilité, sans porter de jugement sur un modèle.

**Exemple.** On mesure la longueur de 200 feuilles d'une même espèce. On obtient  $\bar{x} = 8,3$  cm et  $s = 1,2$  cm. La variabilité reflète la diversité génétique et environnementale : elle n'est pas une anomalie, elle est une donnée biologique.

## E.2 Test de conformité à la moyenne (mathématiques)

**Cadre.** On dispose d'un échantillon  $X_1, \dots, X_n$  i.i.d., d'espérance inconnue  $\mu$ , d'écart-type  $s$  connu ou estimé. On souhaite tester si  $\mu$  est égale à une valeur théorique  $\mu_0$ .

**Hypothèses.**

$$H_0 : \mu = \mu_0 \quad \text{contre} \quad H_1 : \mu \neq \mu_0.$$

**Statistique de test.** On pose  $M_n = \frac{1}{n} \sum_{k=1}^n X_k$ . Par le théorème central limite, pour  $n \geq 30$ , sous  $H_0$  :

$$Z = \frac{M_n - \mu_0}{s/\sqrt{n}} \leftrightarrow \mathcal{N}(0, 1) \quad (\text{approximativement}).$$

**Règle de décision (seuil  $\alpha = 5\%$ ).**

<i>Du point de vue du statisticien</i>	<i>Du point de vue du biologiste</i>
$ Z  > 1,96 \Rightarrow$ on <b>rejette</b> $H_0$	$ Z  \leq 1,96 \Rightarrow$ résultats <b>compatibles</b> avec $\mu_0$

**Remarque de vocabulaire.** Ne jamais écrire « on accepte  $H_0$  ». La formulation correcte est : *on ne dispose pas d'une preuve suffisante pour rejeter  $H_0$* , ou encore : *les données sont compatibles avec  $\mu_0$  au seuil 5%*.

### E.2.1 Lien avec l'intervalle de confiance

Le test de conformité et l'intervalle de confiance sont deux formulations **équivalentes** :

$$\mu_0 \in \left[ M_n \pm 1,96 \frac{s}{\sqrt{n}} \right] \iff \left| \frac{M_n - \mu_0}{s/\sqrt{n}} \right| \leq 1,96.$$

Une valeur théorique est compatible avec les données si et seulement si elle appartient à l'intervalle de confiance.

**Pour aller plus loin** (hors programme BCPST). Lorsque  $n$  est petit et  $\sigma$  inconnue, on utilise la loi de Student de degré  $n - 1$  :

$$\mu_0 \in \left[ \bar{X} \pm t_{n-1, 1-\alpha/2} \frac{s}{\sqrt{n}} \right] \iff \left| \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \right| \leq t_{n-1, 1-\alpha/2}.$$

Pour  $n \geq 30$ , on a  $t_{n-1; 0,975} \approx 1,96$  et les deux approches coïncident.

## E.3 Validation d'un modèle (physique)

**Cadre.** On compare une valeur mesurée expérimentalement à une valeur  $x_{\text{th}}$  prédite par un modèle théorique. Cette valeur mesurée peut être une moyenne  $\bar{x}$  ou une valeur unique.

**Critère de compatibilité.** On calcule le **z-score** :

$$Z = \frac{\bar{x} - x_{\text{th}}}{u(x)},$$

où  $u(x)$  est l'incertitude expérimentale sur  $\bar{x}$ .

$$|Z| \leq 2 \Rightarrow \text{le modèle est compatible avec l'expérience au seuil } 95\%.$$

*Objectif : tester la cohérence d'un modèle avec les données, pas estimer un paramètre.*

**Remarque.** L'écart-type  $s$  joue un rôle différent selon le contexte :

Contexte	Incertitude	Signification
Test de conformité (maths)	$u = \frac{s}{\sqrt{n}}$	Précision sur la <i>moyenne</i>
Validation d'un modèle (physique)	$u \approx s$ (souvent)	Dispersion des <i>mesures</i>

### E.3.1 Variante : comparaison de deux valeurs expérimentales

Le  $z$ -score s'adapte naturellement à la situation où l'on confronte **deux mesures expérimentales**  $x_1 \pm u_1$  et  $x_2 \pm u_2$ , sans qu'aucune ne joue le rôle de vérité théorique. On calcule :

$$Z = \frac{|x_1 - x_2|}{\sqrt{u_1^2 + u_2^2}}.$$

Le dénominateur  $\sqrt{u_1^2 + u_2^2}$  est l'incertitude sur la différence  $x_1 - x_2$ , obtenue par composition quadratique des incertitudes indépendantes (cf. section E.4).

$$|Z| \leq 2 \Rightarrow \text{les deux mesures sont compatibles au seuil 95\%}.$$

**Exemple.** Deux groupes mesurent la célérité du son dans l'air :

$$x_1 = 340 \pm 5 \text{ m/s}, \quad x_2 = 352 \pm 8 \text{ m/s}.$$

On calcule :

$$Z = \frac{|340 - 352|}{\sqrt{5^2 + 8^2}} = \frac{12}{\sqrt{89}} \approx 1,27.$$

Puisque  $Z \approx 1,27 \leq 2$ , les deux mesures sont **compatibles** au seuil 95 % : l'écart de 12 m/s s'explique par les incertitudes expérimentales.

**Lien avec le test de conformité.** Il s'agit du même  $z$ -score : la valeur théorique  $x_{\text{th}}$  est remplacée par  $x_2$ , et l'incertitude  $u(x)$  par l'incertitude composée  $\sqrt{u_1^2 + u_2^2}$ . La logique est identique — normaliser un écart observé par l'incertitude sur cet écart.

## E.4 Incertitudes de type A et de type B

En physique expérimentale, on distingue deux sources d'incertitude.

### Incertaince de type A

Évaluée à partir de **mesures répétées** ; elle correspond à une dispersion statistique :

$$u_A(x) = \frac{s}{\sqrt{n}}.$$

*Interprétation : précision sur la valeur moyenne.*

### Incertaince de type B

Provient d'autres sources : précision de l'instrument, hypothèses du modèle, conditions expérimentales. Elle n'est **pas** obtenue par répétition des mesures.

*Exemples : résolution d'une règle, valeur approchée d'une constante, hypothèse de linéarité.*

### Incertaince totale

$$u(x) = \sqrt{u_A(x)^2 + u_B(x)^2}.$$

**Conséquence.** Lorsque les incertitudes de type B dominent, augmenter  $n$  n'améliore pas significativement la précision. La source d'erreur principale est systématique, non aléatoire.

## E.5 Méthode de Monte-Carlo

La méthode de Monte-Carlo est une technique numérique permettant d'estimer l'incertitude sur une grandeur calculée à partir de grandeurs d'entrée dispersées.

### Principe.

- On modélise chaque grandeur d'entrée par une variable aléatoire (loi choisie selon le type d'incertitude).
- On simule  $N$  réalisations indépendantes de ces grandeurs.
- On calcule la grandeur étudiée  $y$  pour chacune des  $N$  réalisations.

On obtient un échantillon  $y_1, \dots, y_N$ . On calcule :

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i, \quad s = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}.$$

$$u(y) \approx s.$$

### Point clé — pourquoi $s$ et non $s/\sqrt{N}$ ?

Les  $y_i$  ne sont pas  $N$  mesures répétées d'une même valeur : ce sont  $N$  scénarios possibles pour la grandeur  $y$ . Leur écart-type représente directement la dispersion possible de  $y$ , c'est-à-dire son incertitude. On n'estime pas une moyenne — on simule une distribution.

**Avantages.** La méthode prend en compte simultanément les incertitudes de type A et de type B, et reste applicable même lorsque la relation entre grandeurs est complexe ou non linéaire.

### Application en TIPE.

- Estimer une incertitude globale sans calcul analytique de propagation.
- Tester la robustesse d'un modèle face à la dispersion des données.
- Visualiser la distribution des résultats simulés.

## E.6 Lien avec les TIPE

Dans un TIPE, ces trois approches doivent être utilisées conjointement, mais sans être confondues.

Situation	Discipline	Objectif
Description	SVT	Présenter la variabilité : histogramme, $\bar{x}$ , $s$
Estimation	Mathématiques	Calculer un intervalle de confiance pour $\mu$
Confrontation au modèle	Physique	Tester la compatibilité entre $\bar{x}$ et $x_{th}$

Un bon TIPE distingue clairement ces trois situations.

**Remarque importante.** Le choix de l'incertitude au dénominateur dépend de la question posée :

- estimation d'une moyenne :  $\frac{s}{\sqrt{n}}$  (précision sur  $M_n$ );
- dispersion des mesures :  $s$  (variabilité individuelle).

## E.7 Vocabulaire et présentation des résultats

### E.7.1 Anglais scientifique

- Écart-type  $\sigma$  : *standard deviation* (SD).
- Écart-type de  $M_n$ , soit  $\sigma/\sqrt{n}$  : *standard error* (SE), parfois noté **ESM** (*Écart Standard à la Moyenne*) dans les rapports de concours Agro-Véto.

### E.7.2 Le z-score comme fil conducteur

Dans toutes les situations, la démarche est la même :

$$z\text{-score} = \frac{\text{écart observé}}{\text{écart-type (ou incertitude)}} \xrightarrow{\text{comparer à 1,96}} \text{décision.}$$

Ce qui change d'une discipline à l'autre : la nature de l'incertitude au dénominateur ( $s/\sqrt{n}$ ,  $s$ ,  $u(x)$  ou  $\sqrt{u_1^2 + u_2^2}$ ) et l'interprétation de la conclusion.

### E.7.3 Présentation graphique des résultats

En biologie expérimentale, les résultats sont classiquement présentés sous la forme  $\bar{x} \pm s$  (barres d'erreur = écart-type), ou parfois  $\bar{x} \pm s/\sqrt{n}$  (barres d'erreur = erreur standard). **Il est impératif de préciser laquelle des deux est utilisée.**

*Extrait type des sujets ENS :* « Les expériences présentées ont été reproduites plusieurs fois : les graphiques présentent la moyenne des résultats ainsi que l'écart-type sous forme de barres d'erreur. Les images et données brutes sont représentatives de l'ensemble des résultats obtenus. »

## Tableau de synthèse

Situation	Discipline	Formule	Dénominateur	Objectif
Variabilité	SVT	$s$	—	Décrire
Test de conformité	Maths	$\frac{M_n - \mu_0}{s/\sqrt{n}}$	$\frac{s}{\sqrt{n}}$	Décider
Validation modèle	Physique	$\frac{\bar{x} - x_{\text{th}}}{u(x)}$	$u(x)$ (type A+B)	Confronter
Deux mesures	Physique	$\frac{ x_1 - x_2 }{\sqrt{u_1^2 + u_2^2}}$	$\sqrt{u_1^2 + u_2^2}$	Comparer
Monte-Carlo	Physique/info	$u(y) \approx s_{\text{sim}}$	$s_{\text{sim}}$ (pas $s/\sqrt{N}$ )	Estimer $u(y)$