

Formulaire de Mathématiques

1	Analyse	2
1.1	Suites réelles	2
1.2	Séries réelles	4
1.3	Fonctions usuelles	5
1.4	Plan d'étude d'une fonction	6
1.5	Fonctions continues sur un intervalle	9
1.6	Étude locale de fonctions	10
1.7	Calcul intégral	11
1.8	Équations différentielles	15
1.9	Fonctions de deux variables	17
2	Algèbre	18
2.1	Sommes et produits	18
2.2	Polynômes	19
2.3	Calcul matriciel	20
2.4	Applications	23
2.5	Espaces vectoriels	24
2.6	Applications linéaires	25
3	Probabilités	28
3.1	Espaces probabilisés	28
3.2	Variables aléatoires	30
3.3	Variables aléatoires discrètes	32
3.4	Variables aléatoires à densité	35
3.5	Chaînes de Markov	37
3.6	Convergence, approximation et estimation	39

1 Analyse

1.1 Suites réelles

Suites usuelles

- Une suite $(u_n)_{n \geq n_0}$ est **arithmétique** lorsqu'il existe $r \in \mathbb{R}$ tel que :

$$\forall n \geq n_0, \quad u_{n+1} = u_n + r.$$

Forme explicite : Pour tout $n \geq n_0$, $u_n = u_{n_0} + (n - n_0)r$.

- Une suite $(u_n)_{n \geq n_0}$ est **géométrique** lorsqu'il existe $q \in \mathbb{R}$ tel que :

$$\forall n \geq n_0, \quad u_{n+1} = qu_n.$$

Forme explicite : Pour tout $n \geq n_0$, $u_n = q^{n-n_0}u_{n_0}$.

- Une suite $(u_n)_{n \in \mathbb{N}}$ est **arithmético-géométrique** s'il existe $a, b \in \mathbb{R}$ avec $a \neq 1$ tels que :

$$\forall n \in \mathbb{N}, \quad u_{n+1} = au_n + b$$

Forme explicite : Pour obtenir la forme explicite de (u_n) ,

- on cherche l'unique solution x_0 de l'équation au point fixe : $x = ax + b$.
 - on pose pour tout $n \in \mathbb{N}$, $v_n = u_n - x_0$ et on démontre que $(v_n)_{n \in \mathbb{N}}$ est géométrique de raison a .
 - on donne la forme explicite de $(v_n)_{n \in \mathbb{N}}$ puis celle de $(u_n)_{n \in \mathbb{N}}$.
- Une suite (u_n) est **récurrente linéaire d'ordre 2** s'il existe $a, b \in \mathbb{R}$ tels que :

$$\forall n \in \mathbb{N}, \quad u_{n+2} = au_{n+1} + bu_n.$$

Forme explicite : On résout l'équation caractéristique $x^2 = ax + b$.

- S'il y a deux solutions distinctes x_1 et x_2 , alors il existe deux réels λ et μ tels que :

$$\forall n \in \mathbb{N}, \quad u_n = \lambda x_1^n + \mu x_2^n.$$

- S'il y a une unique solution x_0 , alors il existe deux réels λ et μ tels que :

$$\forall n \in \mathbb{N}, \quad u_n = (\lambda + \mu n)x_0^n$$

Pour déterminer λ et μ , on utilise deux termes de la suite (habituellement u_0 et u_1).

Suites convergentes, divergentes

- La limite d'une suite donnée sous forme explicite s'obtient par opérations sur les limites sauf dans les cas suivants où on ne peut pas conclure directement :

$$\infty - \infty, \quad 0 \times \infty, \quad \frac{\infty}{\infty}, \quad \frac{0}{0}.$$

Si on obtient une de ces **formes indéterminées**, il faut transformer l'expression de la suite pour lever l'indétermination ou/et utiliser les équivalents ou/et les développements limités.

- Limite des sous-suites d'une suite convergente : Si la suite (u_n) admet une limite $\ell \in \mathbb{R} \cup \{\pm\infty\}$ alors pour toute fonction strictement croissante φ de \mathbb{N} dans \mathbb{N} , la suite $(u_{\varphi(n)})$ tend vers la même limite ℓ .
- Si (u_{2n}) et (u_{2n+1}) convergent **vers la même limite** ℓ , alors la suite (u_n) converge vers ℓ .
- Passage à la limite dans les inégalités **larges** :

Si (u_n) et (v_n) sont deux suites **convergentes** et m et M sont deux nombres réels, alors :

- Si $u_n \leq M$ à partir d'un certain rang n_0 , alors $\lim_{n \rightarrow +\infty} u_n \leq M$.
- Si $u_n \geq m$ à partir d'un certain rang n_0 , alors $\lim_{n \rightarrow +\infty} u_n \geq m$.

- Si $u_n \leq v_n$ à partir d'un certain rang n_0 , alors $\lim_{n \rightarrow +\infty} u_n \leq \lim_{n \rightarrow +\infty} v_n$.

- Technique de majoration ou de minoration :

Soient (u_n) et (v_n) deux suites réelles telles que $u_n \leq v_n$ à partir d'un certain rang n_0 .

- Si $\lim_{n \rightarrow +\infty} u_n = +\infty$, alors (v_n) est divergente et $\lim_{n \rightarrow +\infty} v_n = +\infty$.

- Si $\lim_{n \rightarrow +\infty} v_n = -\infty$, alors (u_n) est divergente et $\lim_{n \rightarrow +\infty} u_n = -\infty$.

- Théorème d'encadrement :

Soient (u_n) , (v_n) et (w_n) trois suites telles que :

$$u_n \leq v_n \leq w_n \quad \text{à partir d'un certain rang } n_0.$$

Si (u_n) et (w_n) convergent vers une même limite ℓ , alors (v_n) converge aussi vers ℓ .

- Théorème des suites monotones :

Toute suite **croissante** et **majorée** (par une **constante** M) est **convergente**. Toute suite **croissante** et **non majorée** diverge vers $+\infty$.

Toute suite est **décroissante** et **minorée** (par une **constante** m) est **convergente**. Toute suite **décroissante** et **non minorée** diverge vers $-\infty$.

- Théorème des suites adjacentes :

Deux suites (u_n) et (v_n) sont **adjacentes** si elles vérifient les trois hypothèses suivantes :

$$(1) \ (u_n) \text{ est croissante,} \quad (2) \ (v_n) \text{ est décroissante,} \quad (3) \ \lim_{n \rightarrow +\infty} (v_n - u_n) = 0.$$

Dans ce cas, (u_n) et (v_n) **convergent vers une même limite** $\ell \in \mathbb{R}$ et on a :

$$\forall n \in \mathbb{N}, \quad u_n \leq \ell \leq v_n.$$

- Cas des suites récurrentes $u_{n+1} = f(u_n)$:

Si f est **continue** sur un intervalle stable J et si (u_n) converge vers $\ell \in J$ alors $f(\ell) = \ell$, et ℓ est un point fixe de f .

Relations de comparaison

- Sous réserve que v_n ne s'annule pas à partir d'un certain rang,

$$u_n = o(v_n) \Leftrightarrow \lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0 \quad \text{et} \quad u_n \sim v_n \Leftrightarrow \lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 1.$$

- Soient (u_n) et (v_n) deux suites. On a :

$$u_n \sim v_n \Leftrightarrow u_n = v_n + o(v_n)$$

On peut donc négliger les termes négligeables dans une somme.

- Croissances comparées : Si $\alpha, \beta, q \in \mathbb{R}$ avec $\alpha, \beta > 0$ et $q > 1$, alors :

$$\ln(n)^\beta = o(n^\alpha), \quad n^\alpha = o(q^n), \quad q^n = o(n!), \quad n! = o(n^n).$$

- Équivalents usuels : Soit (u_n) une suite telle que $\lim_{n \rightarrow +\infty} u_n = 0$. Alors :

$$e^{u_n} - 1 \sim u_n, \quad \ln(1 + u_n) \sim u_n, \quad (1 + u_n)^\alpha - 1 \sim \alpha u_n \quad \text{pour } \alpha \neq 0.$$

- Opérations sur les équivalents :

On peut faire le produit, le quotient, la puissance ou la valeur absolue des équivalents. Attention, ce sont les seules opérations autorisées sur les équivalents !

- Si (u_n) converge vers un réel ℓ **non nul**, alors $u_n \sim \ell$.
- Si deux suites (u_n) et (v_n) sont **équivalentes**, alors leurs termes u_n et v_n sont de **même signe** à partir d'un certain rang.
- Si deux suites (u_n) et (v_n) sont **équivalentes**, alors elles sont de **même nature**, c'est-à-dire :
 - Si $u_n \sim v_n$ et si (v_n) converge vers un réel ℓ , alors (u_n) converge également vers ℓ .
 - Si $u_n \sim v_n$ et si (v_n) diverge vers $\pm\infty$, alors (u_n) diverge également vers $\pm\infty$.

1.2 Séries réelles

Généralités sur les séries réelles

- Une série $\sum_{n \geq 0} u_n$ converge si et seulement si la suite des sommes partielles $\left(\sum_{k=0}^n u_k \right)_{n \in \mathbb{N}}$ converge.

Dans ce cas, on définit la somme de la série par : $\sum_{n=0}^{+\infty} u_n = \lim_{n \rightarrow +\infty} \sum_{k=0}^n u_k$.

- Si on est capable d'expliciter la somme partielle S_n en fonction de n (c'est en particulier le cas si elle est de type "télescopique"), alors on détermine $\lim_{n \rightarrow +\infty} S_n$:

- Si on obtient une limite finie S , alors la série converge et sa somme est égale à S .
- Sinon, la série diverge.

- Opérations sur les séries convergentes :

- Pour tout $\lambda \neq 0$, $\sum u_n$ **converge** si et seulement si $\sum \lambda u_n$ **converge** et on a :

$$\sum_{n=0}^{+\infty} \lambda u_n = \lambda \sum_{n=0}^{+\infty} u_n.$$

- Si $\sum u_n$ et $\sum v_n$ sont **convergentes**, alors la série $\sum (u_n + v_n)$ **converge** et on a :

$$\sum_{n=0}^{+\infty} (u_n + v_n) = \sum_{n=0}^{+\infty} u_n + \sum_{n=0}^{+\infty} v_n.$$

Attention, on ne peut pas scinder la somme d'une série convergente en deux sommes sans avoir vérifié au préalable que ces deux séries convergent.

- Condition nécessaire de convergence :

Pour que la série $\sum u_n$ converge, **il faut que** $\lim_{n \rightarrow +\infty} u_n = 0$. Mais **ce n'est pas suffisant !**

Séries usuelles

SÉRIES USUELLES		CAS DE CONVERGENCE	SOMME DANS CE CAS
SÉRIES GÉOMÉTRIQUES	$\sum_{n \geq p} q^n$	$ q < 1$	$\sum_{n=p}^{+\infty} q^n = \frac{q^p}{1-q}$
SÉRIES GÉOMÉTRIQUES DÉRIVÉES D'ORDRE 1	$\sum_{n \geq 1} nq^{n-1}$	$ q < 1$	$\sum_{n=1}^{+\infty} nq^{n-1} = \frac{1}{(1-q)^2}$
SÉRIES GÉOMÉTRIQUES DÉRIVÉES D'ORDRE 2	$\sum_{n \geq 2} n(n-1)q^{n-2}$	$ q < 1$	$\sum_{n=2}^{+\infty} n(n-1)q^{n-2} = \frac{2}{(1-q)^3}$
SÉRIES EXPONENTIELLES	$\sum_{n \geq 0} \frac{x^n}{n!}$	quelque soit $x \in \mathbb{R}$	$\sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x$
SÉRIES DE RIEMANN	$\sum_{n \geq 1} \frac{1}{n^\alpha}$	$\alpha > 1$	

Critères de comparaison des séries à termes positifs

COMPARAISON PAR	HYPOTHÈSES	CONCLUSION
<i>Inégalité</i>	<ul style="list-style-type: none"> • $u_k \leq v_k$ à partir d'un certain rang • $u_k \geq 0$ à partir d'un certain rang • $\sum_{k \geq 0} v_k$ converge 	$\sum_{k \geq 0} u_k$ converge
	<ul style="list-style-type: none"> • $u_k \leq v_k$ à partir d'un certain rang • $u_k \geq 0$ à partir d'un certain rang • $\sum_{k \geq 0} u_k$ diverge 	$\sum_{k \geq 0} v_k$ diverge
<i>Négligeabilité</i>	<ul style="list-style-type: none"> • $u_k = o_{+\infty}(v_k)$ • $v_k \geq 0$ à partir d'un certain rang • $\sum_{k \geq 0} v_k$ converge 	$\sum_{k \geq 0} u_k$ converge
	<ul style="list-style-type: none"> • $u_k = o_{+\infty}(v_k)$ • $v_k \geq 0$ à partir d'un certain rang • $\sum_{k \geq 0} u_k$ diverge 	$\sum_{k \geq 0} v_k$ diverge
<i>Équivalent</i>	<ul style="list-style-type: none"> • $u_k \sim_{+\infty} v_k$ • $v_k \geq 0$ (ou $u_k \geq 0$) à partir d'un certain rang • $\sum_{k \geq 0} v_k$ converge (respectivement diverge) 	$\sum_{k \geq 0} u_k$ converge (resp. diverge)

Séries absolument convergentes

- Une série $\sum u_n$ est **absolument convergente** si la série $\sum |u_n|$ converge.
- Dans ce cas, on a :

$$\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n| \quad (\text{Inégalité triangulaire}).$$

- La **convergence absolue** implique la **convergence** (attention, la réciproque est fausse).

1.3 Fonctions usuelles

Fonction partie entière

La fonction **partie entière** est la fonction **définie sur** \mathbb{R} qui, à tout réel x , associe l'unique entier relatif, noté $[x]$, qui vérifie :

$$[x] \leq x < [x] + 1.$$

En d'autres termes, $[x]$ est le plus grand entier relatif qui est inférieur ou égal à x .

Fonction valeur absolue

- La fonction **valeur absolue** est la fonction **définie sur** \mathbb{R} qui, à tout réel x , associe la valeur absolue $|x|$, c'est-à-dire :

$$\forall x \in \mathbb{R}, |x| = \begin{cases} x & \text{si } x \geq 0, \\ -x & \text{si } x < 0. \end{cases}$$

$|x|$ représente la distance de x à 0 et $|y - x|$ la distance de x à y .

- Pour tout réel x ,

$$|x| \geq 0, \quad |-x| = |x|, \quad \sqrt{x^2} = |x|.$$

- Si $r \geq 0$,

$$|x| = r \Leftrightarrow x = \pm r, \quad |x| \leq r \Leftrightarrow -r \leq x \leq r, \quad |x| \geq r \Leftrightarrow x \leq -r \text{ ou } x \geq r.$$

- Pour tous réels x, y ,

$$\begin{aligned} & - |x| \times |y| = |x \times y| \text{ et, si } y \neq 0, \frac{|x|}{|y|} = \left| \frac{x}{y} \right|. \\ & - |x + y| \leq |x| + |y| \text{ (inégalité triangulaire).} \end{aligned}$$

Fonctions logarithme et exponentielle

- Pour tous $x, y > 0$ et pour tout $\alpha \in \mathbb{R}$,

$$\ln(x \times y) = \ln(x) + \ln(y), \quad \ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y), \quad \ln(x^\alpha) = \alpha \ln(x), \quad \ln(1) = 0, \quad e^{\ln(x)} = x.$$

- Pour tous $x, y \in \mathbb{R}$ et pour tout $\alpha \in \mathbb{R}$,

$$e^{x+y} = e^x \times e^y, \quad e^{x-y} = \frac{e^x}{e^y}, \quad e^{\alpha x} = (e^x)^\alpha, \quad e^0 = 1, \quad \ln(e^x) = x.$$

Fonctions puissances

- Soit $\alpha \in \mathbb{R}$. On appelle **fonction puissance** d'exposant α la fonction f définie sur \mathbb{R}_+^* par :

$$f(x) = x^\alpha = \exp(\alpha \ln(x)).$$

- Soient α et β des réels quelconques. Alors :

$$\begin{aligned} & - \text{Pour tout } x > 0, x^\alpha \times x^\beta = x^{\alpha+\beta}; \frac{x^\alpha}{x^\beta} = x^{\alpha-\beta}; (x^\alpha)^\beta = x^{\alpha \times \beta}. \\ & - \text{Pour tout } x, y > 0, x^\alpha \times y^\alpha = (x \times y)^\alpha; \frac{x^\alpha}{y^\alpha} = \left(\frac{x}{y}\right)^\alpha. \end{aligned}$$

1.4 Plan d'étude d'une fonction

Pour les fonctions puissances

Dans le cas des fonctions puissances du type $x \mapsto (u(x))^\alpha$ (avec α non entier) ou $x \mapsto u(x)^{v(x)}$, on commencera **toujours** l'étude en les écrivant sous forme exponentielle :

$$(u(x))^\alpha = e^{\alpha \ln(u(x))} \quad \text{ou} \quad u(x)^{v(x)} = e^{v(x) \ln(u(x))}.$$

Domaine de définition

On détermine le **domaine de définition** \mathcal{D}_f de f à partir de l'expression de $f(x)$. S'il y a :

- Un dénominateur : il faut supprimer de \mathcal{D}_f les racines de ce dénominateur.
- $\sqrt{u(x)}$: il faut résoudre l'inéquation $u(x) \geq 0$ et restreindre \mathcal{D}_f à l'ensemble des solutions.
- $\ln(u(x))$: il faut résoudre l'inéquation $u(x) > 0$ et restreindre \mathcal{D}_f à l'ensemble des solutions.

Parité

Pour étudier la parité de f (uniquement si l'énoncé le demande), on procède en deux étapes :

- On vérifie que \mathcal{D}_f est symétrique par rapport à 0 (il suffit de le constater).
- On exprime $f(-x)$ à l'aide de $f(x)$:
 - Si pour tout $x \in \mathcal{D}_f$, $f(-x) = f(x)$, alors f est **paire** et sa courbe représentative \mathcal{C}_f est symétrique par rapport à l'axe des ordonnées (Oy).
 - Si pour tout $x \in \mathcal{D}_f$, $f(-x) = -f(x)$, alors f est **impaire** et sa courbe représentative \mathcal{C}_f est symétrique par rapport à l'origine du repère O .

Limites aux bornes du domaine de définition

La limite d'une fonction s'obtient par opérations sur les limites sauf dans les cas suivants où on ne peut pas conclure directement :

$$\infty - \infty, \quad 0 \times \infty, \quad \frac{\infty}{\infty}, \quad \frac{0}{0}.$$

Si on obtient une de ces **formes indéterminées**, il faut transformer l'expression de la fonction pour lever l'indétermination ou/et utiliser les équivalents ou/et les développements limités.

Asymptotes

Après avoir déterminé les limites aux bornes du domaine de définition, on précise s'il y a des asymptotes horizontales ou verticales à la courbe représentative \mathcal{C}_f de f :

- Si $\lim_{x \rightarrow x_0} f(x) = \pm\infty$ (où $x_0 \in \mathbb{R}$), alors \mathcal{C}_f admet une **asymptote verticale** d'équation $x = x_0$.
- Si $\lim_{x \rightarrow \pm\infty} f(x) = \ell$, alors \mathcal{C}_f admet une **asymptote horizontale** d'équation $y = \ell$ en $\pm\infty$.

Uniquement si l'énoncé le demande, on étudiera si \mathcal{C}_f admet une **asymptote oblique** Δ d'équation $y = ax + b$ en $+\infty$ (même chose en $-\infty$). Pour cela, on étudie l'écart algébrique $\varepsilon(x) = f(x) - (ax + b)$ entre \mathcal{C}_f et Δ :

- Si $\lim_{x \rightarrow +\infty} \varepsilon(x) = 0$, alors Δ est asymptote de \mathcal{C}_f en $+\infty$.
- Le signe de $\varepsilon(x)$ donne la position relative de \mathcal{C}_f et Δ .

Étude de la continuité

On pensera à toujours préciser le **domaine de continuité** : c'est le même que le domaine de définition sauf s'il y a une **partie entière**.

Si l'énoncé le demande, il faudra étudier la continuité en un point $x_0 \in \mathcal{D}_f$ en dehors du domaine de continuité. Pour cela, on calcule la limite $\lim_{x \rightarrow x_0} f(x)$:

- Si $\lim_{x \rightarrow x_0} f(x) = f(x_0)$, alors f est continue en x_0 .
- Sinon, f n'est pas continue en x_0 .

On peut éventuellement distinguer continuité à gauche et à droite si les limites sont différentes en x_0^- et x_0^+ .

Si l'énoncé le suggère, on pourra aussi étudier le prolongement par continuité de la fonction f en un point $x_0 \notin \mathcal{D}_f$ qui est une borne du domaine de définition. Pour cela, on calcule $\lim_{x \rightarrow x_0} f(x)$:

- Si $\lim_{x \rightarrow x_0} f(x) = \ell$, alors f est prolongeable par continuité en x_0 en posant $f(x_0) = \ell$.
- Sinon, f n'est pas prolongeable par continuité en x_0 .

On peut éventuellement prolonger f par continuité à gauche ou à droite si les limites sont différentes en x_0^- et x_0^+ .

Étude de la dérivabilité

Avant de dériver la fonction, on pensera systématiquement à préciser le **domaine de dérivabilité** : c'est le même que le domaine de continuité sauf s'il y a une **valeur absolue** ou une **racine carrée**. Dans ce cas, il faut supprimer les valeurs qui annulent la valeur absolue ou la racine carrée.

Si l'énoncé le demande, il faudra étudier la dérivabilité en un point x_0 en dehors du domaine de dérivabilité où f est continue. Pour cela, on calcule la limite du **taux d'accroissement** en x_0 :

- Si $\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \ell$, alors f est dérivable en x_0 et $f'(x_0) = \ell$.
- Sinon, f n'est pas dérivable en x_0 .

On peut éventuellement distinguer dérivabilité à gauche et à droite si les limites sont différentes en x_0^- et x_0^+ .

Calcul de la dérivée

Pour tout x dans le domaine de dérivabilité de f , on peut alors calculer $f'(x)$. Pour cela, rappelons les différentes formules à utiliser :

Fonction $f(x) = \dots$	Fonction dérivée $f'(x) = \dots$
constante	0
$e^{u(x)}$	$u'(x)e^{u(x)}$
$\ln(u(x))$	$\frac{u'(x)}{u(x)}$
$(u(x))^\alpha, \alpha \in \mathbb{R}$	$\alpha u'(x)(u(x))^{\alpha-1}$
$\sqrt{u(x)}$	$\frac{u'(x)}{2\sqrt{u(x)}}$
$u(x) + v(x)$	$u'(x) + v'(x)$
$\lambda u(x), \lambda \in \mathbb{R}$	$\lambda u'(x)$
$u(x)v(x)$	$u'(x)v(x) + u(x)v'(x)$
$\frac{u(x)}{v(x)}$	$\frac{u'(x)v(x) - u(x)v'(x)}{(v(x))^2}$
$u(v(x))$	$v'(x)u'(v(x))$

Inégalité des accroissements finis

- Première version :

Soit f une fonction **continue** et **dérivable** sur un intervalle I telle que pour tout $x \in I, m \leq f'(x) \leq M$.
Alors :

$$\forall a, b \in I, m(b - a) \leq f(b) - f(a) \leq M(b - a).$$

- Deuxième version :

Soit f une fonction **continue** et **dérivable** sur un intervalle I telle que pour tout $x \in I, |f'(x)| \leq M$.
Alors :

$$\forall a, b \in I, |f(b) - f(a)| \leq M |b - a|$$

Nous utiliserons généralement la deuxième version de l'inégalité des accroissements finis pour l'étude des suites récurrentes du type $u_{n+1} = f(u_n)$.

Tableau de variation

Dans un même tableau :

- On détermine le signe de $f'(x)$ pour tout x dans le domaine de dérivabilité de f .
- On en déduit les variations de f sur son domaine de définition.
- On complète le tableau de variation avec les limites aux bornes du domaine de définition (et on vérifie que c'est bien cohérent avec les variations de f).

Convexité

Pour étudier la convexité de la fonction f , il faut calculer sa dérivée seconde f'' (en justifiant avant que f' est continue et dérivable !) puis étudier son signe :

- Si $f''(x) \geq 0$, f est **convexe**. Dans ce cas :
 - \mathcal{C}_f est toujours en dessous de ses cordes (définition à utiliser si f n'est pas deux fois dérivable) :

$$\forall a, b \in I, \forall \lambda \in [0, 1], \quad f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b).$$

- \mathcal{C}_f est toujours au dessus de ses tangentes.

- Si $f'' \leq 0$, f est **concave**. Dans ce cas :
 - \mathcal{C}_f est toujours au dessus de ses cordes (définition à utiliser si f n'est pas deux fois dérivable) :

$$\forall a, b \in I, \forall \lambda \in [0, 1], \quad f(\lambda a + (1 - \lambda)b) \geq \lambda f(a) + (1 - \lambda)f(b).$$

- \mathcal{C}_f est toujours en dessous de ses tangentes.

- Si f'' s'annule et change de signe en a , \mathcal{C}_f admet un **point d'inflexion** $A = (a, f(a))$ (qui caractérise un changement de convexité). Dans ce cas, \mathcal{C}_f traverse sa tangente T_A en A d'équation :

$$T_A : y = f'(a)(x - a) + f(a).$$

Tracé de la courbe représentative

On termine l'étude en traçant la courbe représentative \mathcal{C}_f de la fonction (si l'énoncé le demande) :

- On commence par placer les différents points remarquables de la courbe (maximum, minimum, points d'intersections avec les axes, points d'inflexions...).
- On trace les tangentes en ces points.
- On trace les éventuelles asymptotes horizontales, verticales, obliques.
- On trace enfin \mathcal{C}_f (en faisant attention à respecter la parité, la convexité...).

1.5 Fonctions continues sur un intervalle

- Théorème des valeurs intermédiaires :

Soit f une fonction **continue** sur un intervalle $[a, b]$ et qui **change de signe** sur cet intervalle (c'est-à-dire que $f(a).f(b) \leq 0$).

Alors il existe $c \in [a, b]$ tel que $f(c) = 0$, c'est-à-dire que f **s'annule au moins une fois sur** $[a, b]$.

- Image continue d'un intervalle :

Soit f une fonction **continue** sur un intervalle $[a, b]$.

Pour tout réel k compris entre $f(a)$ et $f(b)$, il existe (au moins) un réel $c \in [a, b]$ tel que $f(c) = k$.

Ainsi, **l'image d'un intervalle par une fonction continue est un intervalle**.

- Image continue d'un segment :

Si f est continue sur $[a, b]$, on a :

$$f([a, b]) = \left[\min_{x \in [a, b]} f(x), \max_{x \in [a, b]} f(x) \right].$$

Ainsi, **l'image d'un segment par une fonction continue est un segment**.

- Théorème de la bijection :

Soit f une fonction **continue** et **strictement monotone** sur un intervalle I de \mathbb{R} .

Alors $J = f(I)$ est un intervalle et f réalise une **bijection** de I dans J .

- Définition de la bijection réciproque :

Si f est une bijection d'un intervalle I dans $J = f(I)$, alors la **bijection réciproque** de f est définie par :

$$f^{-1} : \begin{cases} J & \rightarrow I \\ y & \mapsto f^{-1}(y) = \text{l'unique antécédant de } y \text{ par } f \end{cases}$$

Ainsi, on a pour tout $x \in I$ et $y \in J$:

$$y = f(x) \Leftrightarrow f^{-1}(y) = x.$$

- Propriétés de la bijection réciproque :

– Pour tout $x \in I$ et $y \in J$,

$$f^{-1} \circ f(x) = x \quad \text{et} \quad f \circ f^{-1}(y) = y.$$

En particulier, f^{-1} réalise une **bijection** de J sur I .

– Dans un repère orthonormé, les courbes représentatives \mathcal{C}_f et $\mathcal{C}_{f^{-1}}$ sont **symétriques par rapport à la première bissectrice du plan d'équation $y = x$** .

– f^{-1} est elle-même **continue** sur J , **strictement monotone** et de **même sens de variation** que f .

– Si $a, b \in \mathbb{R} \cup \{\pm\infty\}$, alors :

$$\lim_{x \rightarrow a} f(x) = b \Leftrightarrow \lim_{x \rightarrow b} f^{-1}(x) = a.$$

– Si f est **dérivable** sur I et si f' **ne s'annule pas** sur I , alors f^{-1} est **dérivable** sur J et

$$\forall y \in J, (f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

1.6 Étude locale de fonctions

Comparaison de fonctions au voisinage d'un point

- Sous réserve que g ne s'annule pas au voisinage de x_0 ,

$$f(x) \underset{x \rightarrow x_0}{=} o(g(x)) \Leftrightarrow \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0 \quad \text{et} \quad f(x) \underset{x \rightarrow x_0}{\sim} g(x) \Leftrightarrow \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 1.$$

- Soient f et g deux fonctions. On a :

$$f(x) \underset{x \rightarrow x_0}{\sim} g(x) \quad \Leftrightarrow \quad f(x) \underset{x \rightarrow x_0}{=} g(x) + o(g(x)).$$

On peut donc négliger les termes négligeables dans une somme.

- Croissances comparées : Si $\alpha, \beta \in \mathbb{R}_+^*$, alors :

$$(\ln(x))^\beta \underset{x \rightarrow +\infty}{=} o(x^\alpha), \quad x^\alpha \underset{x \rightarrow +\infty}{=} o(e^{\alpha x}), \quad |\ln(x)|^\beta \underset{x \rightarrow 0}{=} o\left(\frac{1}{x^\alpha}\right), \quad e^{\beta x} \underset{x \rightarrow -\infty}{=} o\left(\frac{1}{x^\alpha}\right).$$

- Équivalents usuels au voisinage de 0 :

$$\bullet e^x - 1 \underset{x \rightarrow 0}{\sim} x, \quad \bullet \ln(1+x) \underset{x \rightarrow 0}{\sim} x, \quad \bullet (1+x)^\alpha - 1 \underset{x \rightarrow 0}{\sim} \alpha x \quad \text{pour } \alpha \neq 0.$$

- Opérations sur les équivalents :

On peut faire le produit, le quotient, la puissance ou la valeur absolue des équivalents. Attention, ce sont les seuls opérations autorisées sur les équivalents !

- Si $f(x)$ converge vers une limite finie ℓ **non nulle** lorsque x tend vers x_0 , alors $f(x) \underset{x \rightarrow x_0}{\sim} \ell$.

- Si $f(x) \underset{x \rightarrow x_0}{\sim} g(x)$ et si $\lim_{x \rightarrow x_0} g(x) = \ell$ ($\ell \in \mathbb{R} \cup \{\pm\infty\}$), alors $\lim_{x \rightarrow x_0} f(x) = \ell$.

Développement limité au voisinage d'un point

Soit $f : I \rightarrow \mathbb{R}$ et $x_0 \in I$.

- Une fonction f admet un **développement limité d'ordre 2 en** x_0 s'il existe trois réels a_0, a_1 et a_2 tels que :

$$f(x) \underset{x \rightarrow x_0}{=} \underbrace{a_0 + a_1(x - x_0) + a_2(x - x_0)^2}_{\text{partie régulière}} + \underbrace{o((x - x_0)^2)}_{\text{reste}}.$$

S'il existe, ce développement limité est unique.

- Formule de Taylor-Young à l'ordre 2 :

Si f est de classe \mathcal{C}^2 sur I , alors f admet un développement limité d'ordre 2 en x_0 donné par la formule :

$$f(x) \underset{x \rightarrow x_0}{=} f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + o((x - x_0)^2).$$

- Développements limités usuels au voisinage de 0 :

$$- e^x \underset{x \rightarrow 0}{=} 1 + x + \frac{x^2}{2} + o(x^2)$$

$$- \ln(1 + x) \underset{x \rightarrow 0}{=} x - \frac{x^2}{2} + o(x^2)$$

$$- \forall \alpha \in \mathbb{R}, (1 + x)^\alpha \underset{x \rightarrow 0}{=} 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2}x^2 + o(x^2)$$

- Une fonction est équivalente au premier terme non nul de son développement limité.

1.7 Calcul intégral

Primitives

- Si f est une fonction **continue** sur un intervalle I , alors f admet une **primitive** sur I .
- Soit F_1 une primitive de f sur I . Alors :

$$F_2 \text{ est une primitive de } f \text{ sur } I \iff \text{Il existe } k \in \mathbb{R} \text{ tel que : } \forall x \in I, F_2(x) = F_1(x) + k.$$

- Si f est continue sur un intervalle I et si $c \in I$, alors il existe une unique primitive de f sur I s'annulant en c .
- Primitives et opérations :

Soient f et g deux fonctions admettant des primitives F et G sur un intervalle I et $\lambda \in \mathbb{R}$.

- $F + G$ est une primitive de $f + g$ sur I .
- λF est une primitive de λf sur I .

Soient f une fonction admettant une primitive F sur un intervalle J et u une fonction continue et dérivable sur un intervalle I à valeurs dans J . Alors $F(u)$ est une primitive de $u' \times f(u)$ sur I .

- Primitives usuelles :

Fonction $f(x) = \dots$	Primitives $F(x) = \dots$
$u'(x)e^{u(x)}$	$e^{u(x)} + k, k \in \mathbb{R}$
$u'(x)(u(x))^{-1} = \frac{u'(x)}{u(x)}$	$\ln u(x) + k, k \in \mathbb{R}$
$u'(x)(u(x))^\alpha \ (\alpha \in \mathbb{R} \setminus \{-1\})$	$\frac{(u(x))^{\alpha+1}}{\alpha+1} + k, k \in \mathbb{R}$
$\frac{u'(x)}{\sqrt{u(x)}} \text{ (cas particulier où } \alpha = -\frac{1}{2})$	$2\sqrt{u(x)} + k, k \in \mathbb{R}$

Techniques de calcul d'intégrales sur un segment

- Calcul direct : Soit f une fonction **continue** sur un intervalle I , F une primitive de f sur I et $a, b \in I$. Alors :

$$\int_a^b f(t)dt = [F(t)]_a^b = F(b) - F(a).$$

- Intégration par parties : Soient u et v deux fonctions de **classe** \mathcal{C}^1 sur un segment $[a, b]$. Alors :

$$\int_a^b u'(t)v(t)dt = [u(t)v(t)]_a^b - \int_a^b u(t)v'(t)dx.$$

- Changement de variable : Soit f une fonction continue sur un intervalle I et φ une fonction de **classe** \mathcal{C}^1 sur un segment $[a, b]$, telle que $\varphi([a, b]) \subset I$. Alors :

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(u)du.$$

Propriétés de l'intégrale sur un segment

Soient f et g deux fonctions continues sur un intervalle I et $a, b \in I$.

- **Linéarité de l'intégrale** : Pour tous $\lambda, \mu \in \mathbb{R}$,

$$\int_a^b (\lambda f(t) + \mu g(t))dt = \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt.$$

- **Relation de Chasles** : Pour tout $c \in I$,

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

- **Positivité de l'intégrale** : Supposons que, pour tout réel $t \in I$, $f(t) \geq 0$.

– Bornes croissantes : Si $a \leq b$, alors $\int_a^b f(t)dt \geq 0$.

– Bornes décroissantes : Si $a \geq b$, alors $\int_a^b f(t)dt \leq 0$.

- **Intégration des inégalités** : Supposons que, pour tout réel $t \in I$, $f(t) \leq g(t)$.

– Bornes croissantes : Si $a \leq b$, alors $\int_a^b f(t)dt \leq \int_a^b g(t)dt$.

– Bornes décroissantes : Si $a \geq b$, alors $\int_a^b f(t)dt \geq \int_a^b g(t)dt$.

Techniques de calcul des intégrales généralisées

Soit f est une fonction continue sur $[a, +\infty[$, avec $a \in \mathbb{R}$. Pour calculer l'intégrale $\int_a^{+\infty} f(t)dt$ **généralisée** en $+\infty$, on utilise l'une des méthodes suivantes :

- En revenant à la définition : L'intégrale généralisée $\int_a^{+\infty} f(t)dt$ **converge** si $\int_a^x f(t)dt$ admet une limite finie lorsque x tend vers $+\infty$ et on a alors :

$$\int_a^{+\infty} f(t)dt = \lim_{x \rightarrow +\infty} \int_a^x f(t)dt.$$

- Par intégration par parties ou changement de variable :
 - On commence par faire cette intégration par parties ou ce changement de variable sur un segment $[a, x]$ avec $x \in [a, +\infty[$;
 - On fait ensuite tendre x vers $+\infty$ en veillant à bien justifier que toutes les limites considérées existent.

On n'effectuera JAMAIS d'intégration par parties ou de changement de variables directement sur une intégrale généralisée.

Propriétés des intégrales généralisées

Soient $a \in \mathbb{R}$ et f, g deux fonctions définies et continues sur $[a, +\infty[$.

- **Linéarité de l'intégrale** : Si les intégrales $\int_a^{+\infty} f(t)dt$ et $\int_a^{+\infty} g(t)dt$ convergent, alors :

– Pour tous réels λ, μ , l'intégrale $\int_a^{+\infty} (\lambda f(t) + \mu g(t))dt$ converge.

– De plus, on a :

$$\int_a^{+\infty} (\lambda f(t) + \mu g(t))dt = \lambda \int_a^{+\infty} f(t)dt + \mu \int_a^{+\infty} g(t)dt.$$

Attention, on ne peut pas scinder une intégrale généralisée en deux intégrales sans avoir vérifié au préalable que ces deux intégrales convergent.

- **Relation de Chasles** : Si l'intégrale $\int_a^{+\infty} f(t)dt$ converge, alors :

– Pour tout $b \in [a, +\infty[$, $\int_a^b f(t)dt$ et $\int_b^{+\infty} f(t)dt$ convergent.

– De plus, on a :

$$\int_a^{+\infty} f(t)dt = \int_a^b f(t)dt + \int_b^{+\infty} f(t)dt.$$

- **Positivité de l'intégrale** : Si l'intégrale $\int_a^{+\infty} f(t)dt$ converge et que $f(t) \geq 0$ pour tout $t \in [a, +\infty[$, alors :

$$\int_a^{+\infty} f(t)dt \geq 0.$$

- **Intégration des inégalités** : Si $\int_a^{+\infty} f(t)dt$ et $\int_a^{+\infty} g(t)dt$ convergent et que $f(t) \leq g(t)$ pour tout $t \in [a, +\infty[$, alors :

$$\int_a^{+\infty} f(t)dt \leq \int_a^{+\infty} g(t)dt.$$

Intégrales généralisées et parité

Soit f une fonction continue sur \mathbb{R} .

- **Fonctions paires** :

– Si f est une fonction paire, alors $\int_{-\infty}^0 f(t)dt$ et $\int_0^{+\infty} f(t)dt$ sont de même nature.

– En cas de convergence, on a :

$$\int_{-\infty}^0 f(t)dt = \int_0^{+\infty} f(t)dt \quad \text{et} \quad \int_{-\infty}^{+\infty} f(t)dt = 2 \int_0^{+\infty} f(t)dt.$$

- **Fonctions impaires** :

– Si f est une fonction impaire, alors $\int_{-\infty}^0 f(t)dt$ et $\int_0^{+\infty} f(t)dt$ sont de même nature.

– En cas de convergence, on a :

$$\int_{-\infty}^0 f(t)dt = - \int_0^{+\infty} f(t)dt \quad \text{et} \quad \int_{-\infty}^{+\infty} f(t)dt = 0.$$

Critères de convergence des intégrales généralisées

- Intégrales de Riemann : Soit $\alpha \in \mathbb{R}$.
L'intégrale $\int_1^{+\infty} \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha > 1$.

- Critères de comparaison des intégrales de fonctions positives :

COMPARAISON PAR	HYPOTHÈSES	CONCLUSION
<i>Inégalité</i>	<ul style="list-style-type: none"> • Au voisinage de $+\infty$, $0 \leq f(t) \leq g(t)$ • $\int_a^{+\infty} g(t)dt$ converge 	$\int_a^{+\infty} f(t)dt$ converge
	<ul style="list-style-type: none"> • Au voisinage de $+\infty$, $0 \leq f(t) \leq g(t)$ • $\int_a^{+\infty} f(t)dt$ diverge 	$\int_a^{+\infty} g(t)dt$ diverge
<i>Négligeabilité</i>	<ul style="list-style-type: none"> • $f(t) = o_{+\infty}(g(t))$ • Au voisinage de $+\infty$, $g(t) \geq 0$ et $f(t) \geq 0$ • $\int_a^{+\infty} g(t)dt$ converge 	$\int_a^{+\infty} f(t)dt$ converge
	<ul style="list-style-type: none"> • $f(t) = o_{+\infty}(g(t))$ • Au voisinage de $+\infty$, $g(t) \geq 0$ et $f(t) \geq 0$ • $\int_a^{+\infty} f(t)dt$ diverge 	$\int_a^{+\infty} g(t)dt$ diverge
<i>Équivalent</i>	<ul style="list-style-type: none"> • $f(t) \sim_{+\infty} g(t)$ • Au voisinage de $+\infty$, $g(t) \geq 0$ (ou $f(t) \geq 0$) • $\int_a^{+\infty} g(t)dt$ converge (respectivement diverge) 	$\int_a^{+\infty} f(t)dt$ converge (resp. diverge)

- Intégrales absolument convergentes :

- L'intégrale $\int_a^{+\infty} f(t)dt$ généralisée en $+\infty$ est **absolument convergente** si $\int_a^{+\infty} |f(t)| dt$ converge.
- Dans ce cas, on a :

$$\left| \int_a^{+\infty} f(t)dt \right| \leq \int_a^{+\infty} |f(t)| dt \quad (\text{Inégalité triangulaire}).$$

- La **convergence absolue** implique la **convergence** (attention, la réciproque est fausse).

Toutes les propriétés précédentes, données pour des intégrales du type $\int_a^{+\infty}$, sont également valables pour des intégrales du type $\int_{-\infty}^a$ et $\int_{-\infty}^{+\infty}$.

1.8 Équations différentielles

Équations différentielles linéaires

- Principe de superposition :

Soient

$$(E_1) : a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = b_1$$

et

$$(E_2) : a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = b_2$$

deux équations différentielles linéaires à coefficients constants qui diffèrent seulement par leur second membre b_1 et b_2 .

Si y_1 est une solution de (E_1) et y_2 est une solution de (E_2) , alors, pour tous réels λ et μ , $\lambda y_1 + \mu y_2$ est une solution de l'équation différentielle

$$(E) : a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = \lambda b_1 + \mu b_2.$$

- Le principe de superposition dit en substance que :

$$\text{résoudre l'équation différentielle } (E) \Leftrightarrow \begin{cases} \text{résoudre l'équation différentielle homogène } (E_0) \\ \text{et} \\ \text{trouver une solution particulière de } (E) \end{cases}$$

- Structure algébrique de l'ensemble des solutions de l'équation homogène :

On considère une équation différentielle **linéaire homogène** à coefficients constants d'ordre n :

$$(E_0) : a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0.$$

L'ensemble des solutions de (E_0) est **stable par combinaisons linéaires**. Autrement dit, c'est un sous-espace vectoriel de l'espace vectoriel des fonctions de \mathbb{R} dans \mathbb{R} .

- Existence et unicité d'une solution à un problème de Cauchy :

Pour tout $t_0 \in I$ et pour tout $(y_0, y_1, \dots, y_{n-1}) \in \mathbb{R}^n$, il **existe une unique solution** au problème de Cauchy

$$\begin{cases} a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = b \\ y(t_0) = y_0, y'(t_0) = y_1, \dots, y^{(n-1)}(t_0) = y_{n-1} \end{cases}$$

- On appelle **trajectoire** d'une équation différentielle (E) sur I , tout sous-ensemble de \mathbb{R}^2

$$\{(t, y(t)) \mid t \in I\},$$

où y est une solution de l'équation différentielle (E) .

On dit que c'est une **trajectoire d'équilibre** si y est constante et que c'est une **trajectoire convergente** si $y(t)$ admet une limite finie quand t tend vers $+\infty$.

Si une trajectoire est convergente, alors elle converge nécessairement vers une trajectoire d'équilibre.

- Pour résoudre une équation différentielle $(E) : y' + ay = b$:

– On résout l'équation homogène $(E_0) : y' + ay = 0$ associée à (E) .

Les solutions sont de la forme $y_0(t) = \lambda e^{-at}$, avec λ une constante réelle.

– On cherche une équation particulière y_p de (E) (en se laissant guider par l'exercice).

– On donne l'ensemble des solutions de (E) . Elles sont de la forme $y(t) = y_p(t) + y_0(t)$.

– Si on a une condition initiale $y(t_0) = y_0$, on évalue la solution y obtenue pour déterminer la valeur de λ .

- Pour résoudre une équation différentielle $(E) : y'' + ay' + by = c$:

– On résout l'équation homogène $(E_0) : y'' + ay' + by = 0$ associée à (E) . Pour cela, on détermine les racines de l'équation caractéristique $r^2 + ar + b = 0$.

* Si elle admet deux racines r_1 et r_2 , les solutions sont de la forme $y_0(t) = \lambda e^{r_1 t} + \mu e^{r_2 t}$, avec λ et μ deux constantes réelles.

- * Si elle admet une unique racine r_0 , les solutions sont de la forme $y_0(t) = (\lambda t + \mu)e^{r_0 t}$, avec λ et μ deux constantes réelles.
- On cherche une équation particulière y_p de (E) (en se laissant guider par l'exercice).
- On donne enfin l'ensemble des solutions de (E) . Elles sont de la forme $y(t) = y_p(t) + y_0(t)$.
- Si on a des conditions initiales $y(t_0) = y_0$ et $y'(t_0) = y_1$, on évalue la solution y obtenue pour déterminer les valeurs de λ et μ .

Systèmes différentiels linéaires

- Existence et unicité d'une solution à un problème de Cauchy :

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Pour tout $t_0 \in \mathbb{R}$ et pour tout $X^0 \in \mathcal{M}_{n,1}(\mathbb{R})$, il **existe une unique solution** au problème de Cauchy :

$$\begin{cases} X' = AX \\ X(t_0) = X^0 \end{cases}$$

- Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice **diagonalisable**. On note :

- $\alpha_1, \dots, \alpha_n$ les valeurs propres de A (non nécessairement distinctes, chaque valeur propre apparaît autant de fois que la dimension du sous-espace propre associé) ;
- (U_1, \dots, U_n) une base de $\mathcal{M}_{n,1}(\mathbb{R})$ constituée de vecteurs propres de A telle que, pour tout $i \in \llbracket 1, n \rrbracket$, U_i est un vecteur propre associé à la valeur propre α_i .

Alors l'ensemble des solutions du systèmes différentiel linéaire $X' = AX$ est :

$$\begin{aligned} S &= \{t \mapsto \lambda_1 e^{\alpha_1 t} U_1 + \dots + \lambda_n e^{\alpha_n t} U_n \mid (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n\} \\ &= \left\{ t \mapsto \sum_{i=1}^n \lambda_i e^{\alpha_i t} U_i \mid (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n \right\} \\ &= \text{Vect} (t \mapsto e^{\alpha_1 t} U_1, \dots, t \mapsto e^{\alpha_n t} U_n). \end{aligned}$$

- Lien entre équation différentielle d'ordre 2 et système différentiel :

Soient a et b deux réels, avec $b \neq 0$. On considère l'équation différentielle linéaire **homogène** d'ordre 2 suivante :

$$y'' + ay' + by = 0.$$

En posant $X = \begin{pmatrix} y \\ y' \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}$, on a l'équivalence :

$$y'' + ay' + by = 0 \Leftrightarrow X' = AX.$$

- Soit $A \in \mathcal{M}_n(\mathbb{R})$. On appelle **trajectoire** du système différentiel linéaire $X' = AX$ tout sous-ensemble de \mathbb{R}^n de la forme

$$\{(x_1(t), x_2(t), \dots, x_n(t)) \mid t \in \mathbb{R}\}$$

où (x_1, \dots, x_n) est une solution du système différentiel $X' = AX$.

On dit qu'une trajectoire est un **état d'équilibre** si x_1, x_2, \dots, x_n sont constantes.

On dit qu'une trajectoire est **convergente** si x_1, x_2, \dots, x_n admettent toutes une limite finie en $+\infty$. Sinon on dit que la trajectoire est **divergente**.

- Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice **diagonalisable**. On considère le système différentiel linéaire $X' = AX$.

- Si toutes les valeurs propres de A sont négatives ou nulles, alors toutes les trajectoires du système convergent vers un point d'équilibre et on dit que ces points d'équilibres sont **stables**.
- Si A possède au moins une valeur propre strictement positive, alors il existe des trajectoires divergentes.

1.9 Fonctions de deux variables

Sauf mention contraire, on considère f une fonction de classe \mathcal{C}^2 sur un **ouvert** Ω de \mathbb{R}^2 .

- On appelle **point critique** de f tout point $(x_0, y_0) \in \Omega$ où le **gradient** de f s'annule :

$$\nabla(f)(x_0, y_0) = 0 \Leftrightarrow \begin{pmatrix} \partial_1(f)(x_0, y_0) \\ \partial_2(f)(x_0, y_0) \end{pmatrix} = 0 \Leftrightarrow \begin{cases} \partial_1(f)(x_0, y_0) = 0 \\ \partial_2(f)(x_0, y_0) = 0 \end{cases}$$

- Condition nécessaire d'extremum local :

Si f admet un extremum local en (x_0, y_0) , **alors** (x_0, y_0) est un point critique de f .

- Théorème de Schwarz :

Pour tout $(x, y) \in \Omega$, on a :

$$\partial_{2,1}^2(f)(x, y) = \partial_{1,2}^2(f)(x, y).$$

On rappellera toujours les hypothèses de ce théorème avant de l'utiliser : f est \mathcal{C}^2 sur l'ouvert Ω .

- Condition suffisante d'extremum local :

Soit $(x_0, y_0) \in \Omega$ un point critique de f et $\nabla^2(f)(x_0, y_0)$ la **matrice hessienne** de f en ce point définie par :

$$\nabla^2(f)(x_0, y_0) = \begin{pmatrix} \partial_{1,1}^2(f)(x_0, y_0) & \partial_{1,2}^2(f)(x_0, y_0) \\ \partial_{2,1}^2(f)(x_0, y_0) & \partial_{2,2}^2(f)(x_0, y_0) \end{pmatrix}.$$

Alors :

- Si les valeurs propres de $\nabla^2(f)(x_0, y_0)$ sont **strictement positives**, alors le point critique est un **minimum local** ;
 - Si les valeurs propres de $\nabla^2(f)(x_0, y_0)$ sont **strictement négatives**, alors le point critique est un **maximum local** ;
 - Si les valeurs propres de $\nabla^2(f)(x_0, y_0)$ sont **non nulles** et de **signes opposés**, alors le point critique **n'est pas un extremum local** ;
 - Si une des valeurs propre de $\nabla^2(f)(x_0, y_0)$ est nulle, alors on ne peut rien conclure par l'étude de la hessienne.
- Soit f une fonction **continue** sur une partie Ω **fermée** et **bornée** de \mathbb{R}^2 . Alors f admet un maximum global et un minimum global sur Ω , de sorte qu'il existe (x_0, y_0) et $(x_1, y_1) \in \Omega$ tels que :

$$\forall (x, y) \in \Omega, \quad f(x_0, y_0) \leq f(x, y) \leq f(x_1, y_1).$$

2 Algèbre

2.1 Sommes et produits

Opérations sur les sommes

- Soient $n \in \mathbb{N}$, $x_0, x_1, \dots, x_{n-1}, x_n$ et $y_0, y_1, \dots, y_{n-1}, y_n$ des réels. Alors :

$$\sum_{k=0}^n (x_k + y_k) = \sum_{k=0}^n x_k + \sum_{k=0}^n y_k.$$

- Soient $n \in \mathbb{N}$, $x_0, x_1, \dots, x_{n-1}, x_n$ des réels et $\lambda \in \mathbb{R}$. Alors :

$$\sum_{k=0}^n \lambda x_k = \lambda \sum_{k=0}^n x_k \quad (\text{factorisation par } \lambda \text{ qui ne dépend pas de l'indice}).$$

- Soient $n, m \in \mathbb{N}$ avec $0 \leq m < n$ et $x_0, x_1, \dots, x_{n-1}, x_n$ des réels. Alors :

$$\sum_{k=0}^n x_k = \sum_{k=0}^m x_k + \sum_{k=m+1}^n x_k \quad (\text{relation de Chasles}).$$

Sommes usuelles

Soit $n, p \in \mathbb{N}$, avec $p \leq n$.

- Somme d'une constante a (indépendante de l'indice de sommation k) :

$$\sum_{k=p}^n a = (n - p + 1) \times a.$$

- Sommes des premières puissances des n premiers entiers naturels :

$$\sum_{k=0}^n k = \frac{n(n+1)}{2}, \quad \sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6}, \quad \sum_{k=0}^n k^3 = \frac{n^2(n+1)^2}{4}.$$

- Somme géométrique (avec $q \in \mathbb{R}^*$) :

$$\sum_{k=p}^n q^k = \begin{cases} n - p + 1 & \text{si } q = 1, \\ q^p \times \frac{1 - q^{n-p+1}}{1 - q} & \text{si } q \neq 1. \end{cases}$$

Opérations sur les produits

- Soient $n \in \mathbb{N}$, $x_0, x_1, \dots, x_{n-1}, x_n$ et $y_0, y_1, \dots, y_{n-1}, y_n$ des réels. Alors :

$$\prod_{k=0}^n (x_k \times y_k) = \left(\prod_{k=0}^n x_k \right) \times \left(\prod_{k=0}^n y_k \right).$$

- Soient $n \in \mathbb{N}$, $x_0, x_1, \dots, x_{n-1}, x_n$ des réels et $y_0, y_1, \dots, y_{n-1}, y_n$ des réels non nuls. Alors :

$$\prod_{k=0}^n \frac{x_k}{y_k} = \frac{\prod_{k=0}^n x_k}{\prod_{k=0}^n y_k}.$$

- Soient $n, m \in \mathbb{N}$ avec $0 \leq m < n$ et $x_0, x_1, \dots, x_{n-1}, x_n$ des réels. Alors :

$$\prod_{k=0}^n x_k = \left(\prod_{k=0}^m x_k \right) \times \left(\prod_{k=m+1}^n x_k \right) \quad (\text{relation de Chasles}).$$

Coefficients binomiaux

Soit $n \in \mathbb{N}$.

- Factorielle n : $n! = \prod_{k=1}^n k$. Par convention, $0! = 1$.

- $\binom{n}{k}$ est le nombre de parties de k éléments dans un ensemble de n éléments.

En particulier, $\binom{n}{0} = 1$, $\binom{n}{1} = n$, $\binom{n}{n-1} = n$, $\binom{n}{n} = 1$ et $\binom{n}{k} = 0$ si $k > n$.

- Formule explicite des coefficients binomiaux : pour tout $k \in \llbracket 0, n \rrbracket$,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

- Symétrie des coefficients binomiaux : pour tout $k \in \llbracket 0, n \rrbracket$,

$$\binom{n}{k} = \binom{n}{n-k}.$$

- Formule du triangle de Pascal : pour tout $k \in \llbracket 1, n-1 \rrbracket$,

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

- Formule du binôme de Newton : pour tout $a, b \in \mathbb{R}$,

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Sommes doubles

Soient $n, m \in \mathbb{N}^*$.

- Interversion de sommes à indices indépendants :

$$\sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} x_{i,j} = \sum_{i=1}^n \left(\sum_{j=1}^m x_{i,j} \right) = \sum_{j=1}^m \left(\sum_{i=1}^n x_{i,j} \right).$$

- Interversion de sommes à indices dépendants :

$$\sum_{1 \leq i \leq j \leq n} x_{i,j} = \sum_{i=1}^n \left(\sum_{j=i}^n x_{i,j} \right) = \sum_{j=1}^n \left(\sum_{i=1}^j x_{i,j} \right)$$

$$\sum_{1 \leq i < j \leq n} x_{i,j} = \sum_{i=1}^{n-1} \left(\sum_{j=i+1}^n x_{i,j} \right) = \sum_{j=2}^n \left(\sum_{i=1}^{j-1} x_{i,j} \right)$$

2.2 Polynômes

Degré d'un polynôme

Soient $P = \sum_{k=0}^n a_k X^k$ et $Q = \sum_{k=0}^m b_k X^k$ deux polynômes et $\lambda \in \mathbb{R}^*$. Alors :

- $\deg(P) = \max \{i \in \llbracket 0, n \rrbracket \mid a_i \neq 0\}$ (si cela existe).
- $\deg(0) = -\infty$ (par convention).
- $\deg(P+Q) \leq \max(\deg(P), \deg(Q))$.
- $\deg(\lambda.P) = \deg(P)$.
- $\deg(PQ) = \deg(P) + \deg(Q)$.
- $\deg(P \circ Q) = \deg(P) \times \deg(Q)$.

Division euclidienne

Soient A, B deux polynômes tels que $B \neq 0$.

- Division euclidienne : Il existe un unique couple (Q, R) de polynômes tel que :

$$\begin{cases} A = BQ + R \\ \deg(R) < \deg(B) \end{cases}$$

Q et R sont appelés le **quotient** et le **reste** de la **division euclidienne** de A par B .

- Si $R = 0$, alors on dit que B divise A et on note $B|A$.

Racines d'un polynôme

- Soit $P \in \mathbb{R}_n[X]$. Alors :

$$a \text{ est une racine de } P \Leftrightarrow P(a) = 0 \Leftrightarrow (X - a)|P \Leftrightarrow \exists Q \in \mathbb{R}_{n-1}[X], P = (X - a) \times Q.$$

On pourra alors déterminer Q par identification ou par division euclidienne.

- Un polynôme non nul de $\mathbb{R}_n[X]$ admet au plus n racines distinctes.
En particulier, si un polynôme de $\mathbb{R}_n[X]$ admet plus de $n + 1$ racines distinctes, alors il est nul.

2.3 Calcul matriciel

Opérations sur les matrices

- Produit matriciel : Soient $A \in \mathcal{M}_{n,p}(\mathbb{R})$ et $B \in \mathcal{M}_{p,q}(\mathbb{R})$. On définit la matrice $C = A \times B$ de $\mathcal{M}_{n,q}(\mathbb{R})$ par :

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, q \rrbracket, [C]_{i,j} = \sum_{k=1}^p [A]_{i,k} [B]_{k,j}$$

Rappelons que le produit matriciel **n'est pas commutatif**.

- Transposition : Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$. La **matrice transposée** ${}^tA \in \mathcal{M}_{p,n}(\mathbb{R})$ de A est définie par :

$$\forall (i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, n \rrbracket, [{}^tA]_{i,j} = [A]_{j,i}.$$

Rappelons que, sous réserve que les calculs suivants aient un sens,

$${}^t({}^tA) = A, \quad {}^t(\lambda \cdot A + \mu \cdot B) = \lambda \cdot {}^tA + \mu \cdot {}^tB, \quad {}^t(A \times B) = {}^tB \times {}^tA.$$

Puissances d'une matrice carrée

- Pour calculer la puissances d'une matrice carrée M , on utilisera l'une des méthodes suivantes :
 - **Si elle est diagonale** : C'est la diagonale des puissances.
 - **Par récurrence** : En calculant les premières puissances de M , on obtient une formule générale qui semble être valable pour tout entier naturel. On démontre alors cette formule par récurrence.
 - **Avec la formule du binôme de Newton** : On écrit $M = A + B$ où A et B sont deux matrices qui **commutent** (c'est-à-dire telles que $AB = BA$) et dont les puissances sont plus simples à calculer. On utilise alors la formule :

$$M^p = (A + B)^p = \sum_{k=0}^p \binom{p}{k} A^k B^{p-k} = \sum_{k=0}^p \binom{p}{k} B^k A^{p-k}.$$

- **En se ramenant aux puissances d'une matrice diagonale** : On écrit $M = PDP^{-1}$ avec D une matrice diagonale, on démontre par récurrence que, pour tout entier naturel k , $M^k = PD^kP^{-1}$ et on obtient finalement M^p en faisant le produit matriciel.
- **A l'aide d'un polynôme annulateur** $P \in \mathbb{R}_d[X]$ de M (en faisant une division euclidienne de X^k par P ou par récurrence en écrivant M^k comme combinaison linéaire de I, M, \dots, M^{d-1}).

Matrices carrées inversibles

- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est inversible s'il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $A \times B = B \times A = I_n$.
- Caractérisations des matrices inversibles : Soit $A \in \mathcal{M}_n(\mathbb{R})$. Les assertions suivantes sont équivalentes :
 - A est une matrice inversible ;
 - Il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $B \times A = I_n$ (et alors $A^{-1} = B$) ;
 - Il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $A \times B = I_n$ (et alors $A^{-1} = B$) ;
 - Les colonnes (ou les lignes) de A sont linéairement indépendantes ;
 - $\forall B \in \mathcal{M}_{n,1}(\mathbb{R})$, le système $AX = B$ admet une unique solution (qui est $X = A^{-1}B$).

- Propriétés de l'inverse : Étant données deux matrices inversibles A et B ,

$$(A^{-1})^{-1} = A, \quad ({}^t A)^{-1} = {}^t (A^{-1}), \quad (A \times B)^{-1} = B^{-1} \times A^{-1}$$

- Pour montrer qu'une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est inversible, on utilisera l'une des méthodes suivantes :
 - **Si elle est diagonale** : $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ est inversible si et seulement si $\lambda_i \neq 0$ pour tout $1 \leq i \leq n$, et on a alors $A^{-1} = \text{diag}(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n})$.
 - **Si elle est triangulaire** : Une matrice triangulaire est inversible si et seulement si tous ses coefficients diagonaux sont non nuls.
 - **Si elle est carrée de taille 2** : Une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible si et seulement si son déterminant $\det(A) = ad - bc$ est non nul. Dans ce cas,

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

- **Si on nous suggère l'inverse** : Il suffit de faire le produit des deux matrices.
- **Si on a trouvé un polynôme annulateur P de A** :
 - * Si $P(0) \neq 0$, alors A est inversible. Pour le prouver, on met l'équation $P(A) = 0$ sous la forme $A \times B = I_n$ (et $A^{-1} = B$ est un polynôme en A).
 - * Si $P(0) = 0$, alors A n'est en général pas inversible. Pour le prouver, on raisonne par l'absurde : on met l'équation $P(A) = 0$ sous la forme $A \times B = 0$, puis on multiplie par A^{-1} à gauche pour aboutir à une contradiction.
- **Si on veut montrer que A n'est pas inversible** : On vérifie si l'une des colonnes (ou des lignes) de A est combinaison linéaire des autres colonnes (ou lignes).
- **Dans le cas général** : On applique la méthode du pivot de Gauss.

Noyau, image et rang d'une matrice

- Le **noyau** d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{R})$, noté $\text{Ker}(A)$, est l'ensemble :

$$\text{Ker}(A) = \{X \in \mathcal{M}_{p,1}(\mathbb{R}) \mid A \cdot X = 0_{n,1}\}.$$

C'est un sous-espace vectoriel de $\mathcal{M}_{p,1}$.

- L'**image** d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{R})$, noté $\text{Im}(A)$, est l'ensemble :

$$\text{Im}(A) = \{Y \in \mathcal{M}_{n,1}(\mathbb{R}) \mid \exists X \in \mathcal{M}_{p,1}(\mathbb{R}), Y = A \cdot X\}.$$

C'est un sous-espace vectoriel de $\mathcal{M}_{n,1}$ et, si on note C_i la i -ième colonne de A ,

$$\text{Im}(A) = \text{Vect}(C_1, \dots, C_p).$$

- Le **rang** $\text{rg}(A)$ d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{R})$ est le rang de la famille de ses vecteurs colonnes C_1, \dots, C_p . Autrement dit :

$$\text{rg}(A) = \dim(\text{Im}(A)).$$

- Le rang d'une matrice est aussi le rang de ses vecteurs lignes :

$$\text{rg}({}^t A) = \text{rg}(A).$$

- Pour calculer le rang d'une matrice A , on échelonne la matrice par la méthode du pivot de Gauss. Le rang de A est alors le nombre de pivots non nuls de la matrice échelonnée obtenue.
- Théorème du rang : Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$. Alors :

$$p = \dim(\text{Ker}(A)) + \text{rg}(A).$$

- CNS d'inversibilité : Si $A \in \mathcal{M}_n(\mathbb{R})$, alors :

$$A \text{ inversible} \Leftrightarrow \text{Ker}(A) = \{0_{n,1}\} \Leftrightarrow \text{Im}(A) = \mathcal{M}_{n,1}(\mathbb{R}) \Leftrightarrow \text{rg}(A) = n.$$

Éléments propres d'une matrice carrée

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

- $\lambda \in \mathbb{R}$ est une **valeur propre** de A lorsque :

$$\exists X \in \mathcal{M}_{n,1}(\mathbb{R}) \setminus \{0_{n,1}\}, AX = \lambda X.$$

- $X \in \mathcal{M}_{n,1}(\mathbb{R})$ est un **vecteur propre** de A associé à la valeur propre λ si :

$$X \text{ est non nul et } AX = \lambda X.$$

- Le **spectre** $\text{Sp}(A)$ de A est l'ensemble des valeurs propres de A .
- Si λ est une valeur propre de A , le sous-espace vectoriel

$$E_\lambda(A) = \{X \in \mathcal{M}_{n,1}(\mathbb{R}) \mid AX = \lambda X\} = \text{Ker}(A - \lambda I_n).$$

est le **sous-espace propre** de A associé à la valeur propre λ .

- Caractérisation des valeurs propres de A :

$$\begin{aligned} \lambda \in \text{Sp}(A) &\Leftrightarrow E_\lambda(A) \neq \{0_{n,1}\} \\ &\Leftrightarrow \dim(\text{Ker}(A - \lambda I_n)) \geq 1 \\ &\Leftrightarrow \text{rg}(A - \lambda I_n) < n \\ &\Leftrightarrow (A - \lambda I_n) \text{ n'est pas inversible.} \end{aligned}$$

De plus, on a : $\dim(E_\lambda(A)) = n - \text{rg}(A - \lambda I_n)$.

- Cas particulier de la valeur propre 0 :

$$0 \in \text{Sp}(A) \Leftrightarrow \dim(\text{Ker}(A)) \geq 1 \Leftrightarrow \text{rg}(A) < n \Leftrightarrow A \text{ n'est pas inversible.}$$

De plus, on a : $\dim(E_0(A)) = n - \text{rg}(A)$.

- Concaténer de familles libres de sous-espaces propres :

Soient $\lambda_1, \lambda_2, \dots, \lambda_p$ des valeurs propres **deux à deux distinctes** d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$.

Si \mathcal{L}_1 est une famille libre de $E_{\lambda_1}(A)$, \mathcal{L}_2 est une famille libre de $E_{\lambda_2}(A)$, ... , \mathcal{L}_p est une famille libre de $E_{\lambda_p}(A)$, alors la famille \mathcal{L} obtenue en concaténant (c'est-à-dire en juxtaposant) les familles $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_p$ est libre.

- $A \in \mathcal{M}_n(\mathbb{R})$ possède au plus n valeurs propres distinctes.
- Si A est une matrice **triangulaire**, alors ses valeurs propres sont exactement ses coefficients diagonaux.
- Si $P \in \mathbb{R}[X]$ un polynôme annulateur de A , alors les valeurs propres de A sont **parmi** les racines de P .
- Si $A \in \mathcal{M}_n(\mathbb{R})$, alors :

$$\lambda \in \text{Sp}(A) \Leftrightarrow A - \lambda I_2 \text{ non inversible} \Leftrightarrow \det(A - \lambda I_2) = 0.$$

Réduction des matrices carrées

- $A \in \mathcal{M}_n(\mathbb{R})$ est **diagonalisable** si elle est semblable à une matrice diagonale, c'est-à-dire s'il existe une matrice diagonale $D \in \mathcal{M}_n(\mathbb{R})$ et une matrice inversible $P \in \mathcal{M}_n(\mathbb{R})$ telles que :

$$A = PDP^{-1}.$$

- Condition nécessaire et suffisante de diagonalisabilité :

Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est diagonalisable si et seulement si il existe une base de $\mathcal{M}_{n,1}(\mathbb{R})$ formée de vecteurs propres de A .

- Pour montrer qu'une matrice A est diagonalisable, on procèdera de la façon suivante :

– Si la matrice A est symétrique (c'est-à-dire si ${}^tA = A$), alors A est diagonalisable.

– Sinon, il faut déterminer les valeurs propres de A :

* Si A ne possède aucune valeur propre, alors A n'est pas diagonalisable.

* Si A possède une seule valeur propre, alors on raisonne par l'absurde pour montrer que A n'est pas diagonalisable (sauf si $A = \lambda I_n$).

* Si A possède p valeurs propres $\lambda_1, \dots, \lambda_p$ (avec $1 < p \leq n$), on détermine une base \mathcal{B}_i de chacun de ses sous-espaces propres $E_{\lambda_i}(A)$ pour $i \in \llbracket 1, p \rrbracket$.

La matrice A est diagonalisable si et seulement si la concaténation des bases de ses sous-espaces

propres $\mathcal{B} = \bigcup_{i=1}^n \mathcal{B}_i$ est une base de $\mathcal{M}_{n,1}(\mathbb{R})$.

- Supposons que $A \in \mathcal{M}_n(\mathbb{R})$ est diagonalisable. Pour expliciter les matrices D et $P \in \mathcal{M}_n(\mathbb{R})$ telles que D est diagonale, P est inversible et $A = PDP^{-1}$:

– Pour la matrice D :

C'est la matrice diagonale avec sur sa diagonale les valeurs propres λ_i de A apparaissant $\dim(E_{\lambda_i}(A))$ fois.

– Pour la matrice P :

Comme A est diagonalisable, la famille $\mathcal{B} = \bigcup_{i=1}^n \mathcal{B}_i$, obtenue par concaténation des bases des $E_{\lambda_i}(A)$,

est une base de $\mathcal{M}_{n,1}(\mathbb{R})$. P est alors la matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ à la base \mathcal{B} .

Attention à l'ordre : il faut placer les vecteurs de la base \mathcal{B} dans P dans le même ordre que les valeurs propres auxquelles ils sont associés dans D .

2.4 Applications

Soit f une application d'un ensemble E dans un ensemble F .

- Applications injectives :

f est **injective** \Leftrightarrow tout élément de F admet **au plus** un antécédent dans E par f
 $\Leftrightarrow (\forall x_1, x_2 \in E, f(x_1) = f(x_2) \Rightarrow x_1 = x_2).$

- Applications surjectives :

f est **surjective** \Leftrightarrow tout élément de F admet **au moins** un antécédent dans E par f
 $\Leftrightarrow (\forall y \in F, \exists x \in E, y = f(x)).$

- Applications bijectives :

f est **bijective** \Leftrightarrow tout élément de F admet **un unique** antécédent dans E par f
 $\Leftrightarrow (\forall y \in F, \exists! x \in E, y = f(x))$
 $\Leftrightarrow f$ est **injective** et **surjective**
 \Leftrightarrow il existe une application $g : F \rightarrow E$ telle que : $g \circ f = Id_E$ et $f \circ g = Id_F$.

Dans ce cas, g est unique et c'est la **bijection réciproque** de f .

- Composée d'applications bijectives :

Si f est une **bijection** de E dans F et si g est une **bijection** de F dans G , alors $g \circ f$ est une **bijection** de E dans G et on a :

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

2.5 Espaces vectoriels

- Caractérisation des sous-espaces vectoriels :

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

- $F \subset E$
- F est une partie non vide de E .
- $\forall u, v \in F, \forall \lambda \in \mathbb{R}, \lambda \cdot u + v \in F$.

- Sous-espaces vectoriels engendrés :

Si (u_1, \dots, u_p) est une famille de p vecteurs de E , on note $\text{Vect}(u_1, \dots, u_p)$ l'ensemble de toutes les combinaisons linéaires de la famille (u_1, \dots, u_p) :

$$\text{Vect}(u_1, \dots, u_p) = \{\lambda_1 \cdot u_1 + \dots + \lambda_p \cdot u_p \mid (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p\}.$$

C'est le **sous-espace vectoriel de E engendré** par la famille (u_1, \dots, u_p) et c'est le plus petit sous-espace vectoriel contenant les vecteurs u_1, \dots, u_p .

- Familles de vecteurs :

- Une famille (u_1, \dots, u_p) de p vecteurs de E est **génératrice** de E si

$$E = \text{Vect}(u_1, \dots, u_p).$$

- Une famille (u_1, \dots, u_p) de p vecteurs de E est **libre** si :

$$\forall (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p, \lambda_1 \cdot u_1 + \dots + \lambda_p \cdot u_p = 0_E \Rightarrow \lambda_1 = \dots = \lambda_p = 0.$$

- Une famille (u_1, \dots, u_p) est une **base** de E si et seulement si tout vecteur de E peut s'écrire de manière unique comme combinaison linéaire de (u_1, \dots, u_p) , c'est-à-dire :

$$\forall v \in E, \exists! (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n, v = \lambda_1 \cdot u_1 + \dots + \lambda_p \cdot u_p.$$

Toutes les bases de E ont le même cardinal qui est la **dimension** de E .

- Espaces vectoriels de références :

- L'espace vectoriel \mathbb{R}^n des n -uplets de réels.

- * Élément neutre : $(0, \dots, 0)$.
- * Base canonique :

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

- * Dimension : $\dim(\mathbb{R}^n) = n$.

- L'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{R})$ des matrices de taille $n \times p$.

- * Élément neutre : la matrice nulle.
- * Base canonique :

$$\forall i, j \in \llbracket 1, n \rrbracket, \quad E_{i,j} = \begin{matrix} & & & j & & \\ & & & \downarrow & & \\ \left(\begin{array}{cccccc} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{array} \right) \end{matrix}$$

- * Dimension : $\dim(\mathcal{M}_{n,p}(\mathbb{R})) = n \times p$.

- L'espace vectoriel $\mathbb{R}_n[X]$ des polynômes de degré $\leq n$.

- * Élément neutre : le polynôme nul.
- * Base canonique : $(1, X, X^2, \dots, X^n)$
- * Dimension : $\dim(\mathbb{R}_n[X]) = n + 1$.

- Cardinal d'une famille libre ou génératrice :

Si E est un espace vectoriel de dimension finie n , alors :

- Toute famille libre de E est de cardinal inférieur ou égal à n .
Une famille libre de E de cardinal n est une base de E .
- Toute famille génératrice de E est de cardinal supérieur ou égal à n .
Une famille génératrice de E de cardinal n est une base de E .
- Pour montrer qu'une famille $\mathcal{F} = (u_1, \dots, u_p)$ est une base de E , deux possibilités :
 - Si on connaît la dimension de E :
On vérifie que $\dim(E) = p = \text{Card}(\mathcal{F})$ et on montre que \mathcal{F} est libre.
 - Si on ne connaît pas la dimension de E :
On montre que \mathcal{F} est libre et génératrice. On aura alors $\dim(E) = p = \text{Card}(\mathcal{F})$.

- Rang d'une famille de vecteurs :

Soit E un espace vectoriel de dimension finie n et $\mathcal{F} = (u_1, \dots, u_p)$ une famille de vecteurs de E . Le **rang** de \mathcal{F} est défini par :

$$\text{rg}(\mathcal{F}) = \dim(\text{Vect}(u_1, \dots, u_p)).$$

On a alors :

- \mathcal{F} est libre si et seulement si $\text{rg}(\mathcal{F}) = p$.
- \mathcal{F} est génératrice si et seulement si $\text{rg}(\mathcal{F}) = n$.
- Si \mathcal{B} est une base de E , alors :

$$\text{rg}(\mathcal{F}) = \text{rg}(M_{\mathcal{B}}(\mathcal{F})).$$

2.6 Applications linéaires

Soit E et F deux espaces vectoriels réels.

- Une application $f : E \rightarrow F$ est dite **linéaire** si et seulement si

$$\forall u, v \in E, \forall \lambda \in \mathbb{R}, f(\lambda \cdot u + v) = \lambda \cdot f(u) + f(v).$$

- Soit $f \in \mathcal{L}(E, F)$ une application linéaire.
 - Si $E = F$, on dit que f est un **endomorphisme**.
 - Si f est bijective, on dit que f est un **isomorphisme**.
 - Si $E = F$ et si f est bijective, on dit que f est un **automorphisme**.
- Noyau d'une application linéaire $f \in \mathcal{L}(E, F)$:

$$\text{Ker}(f) = \{u \in E \mid f(u) = 0_F\}.$$

On a alors que :

- $\text{Ker}(f)$ est un sous-espace vectoriel de E .
- f est **injective** si et seulement si $\text{Ker}(f) = \{0_E\}$.
- Image d'une application linéaire $f \in \mathcal{L}(E, F)$:

$$\text{Im}(f) = \{v \in F \mid \exists u \in E, v = f(u)\} = \{f(u) \mid u \in E\}.$$

On a alors que :

- $\text{Im}(f)$ est un **sous-espace vectoriel** de F .
 - f est **surjective** si et seulement si $\text{Im}(f) = F$.
 - Si $\mathcal{B} = (e_1, \dots, e_p)$ est une base de E , alors $\text{Im}(f) = \text{Vect}(f(e_1), \dots, f(e_p))$.
 - Théorème du rang :
- Soit $f \in \mathcal{L}(E, F)$ avec E de **dimension finie**. Alors :

$$\dim(E) = \dim(\text{Ker}(f)) + \dim(\text{Im}(f)) = \dim(\text{Ker}(f)) + \text{rg}(f).$$

- Caractérisation des isomorphismes :

On suppose que E et F sont de **même dimension finie**, et soit $f \in \mathcal{L}(E, F)$. Alors

$$f \text{ est bijective} \Leftrightarrow f \text{ est injective} \Leftrightarrow f \text{ est surjective.}$$

- Matrices de passage :

Soit E un espace vectoriel de dimension finie n , $\mathcal{B} = (e_1, \dots, e_n)$ (ancienne base) et $\mathcal{B}' = (e'_1, \dots, e'_n)$ (nouvelle base) deux bases de E .

On appelle **matrice de passage** de \mathcal{B} à \mathcal{B}' et on note $P_{\mathcal{B}, \mathcal{B}'}$ la matrice de $\mathcal{M}_n(\mathbb{R})$ définie par :

$$P_{\mathcal{B}, \mathcal{B}'} = M_{\mathcal{B}}(\mathcal{B}') = \begin{matrix} & e'_1 & \dots & e'_j & \dots & e'_n \\ \begin{matrix} e_1 \rightarrow \\ e_2 \rightarrow \\ \vdots \\ e_n \rightarrow \end{matrix} & \begin{pmatrix} p_{1,1} & \dots & p_{1,j} & \dots & p_{1,n} \\ p_{2,1} & \dots & p_{2,j} & \dots & p_{2,n} \\ \vdots & & \vdots & & \vdots \\ p_{n,1} & \dots & p_{n,j} & \dots & p_{n,n} \end{pmatrix} \end{matrix} \quad \text{où} \quad \forall j \in \llbracket 1, n \rrbracket, e'_j = \sum_{i=1}^n p_{i,j} \cdot e_i.$$

On a alors les propriétés suivantes :

- Si $\mathcal{B}, \mathcal{B}'$ et \mathcal{B}'' sont trois bases de E , alors :

$$P_{\mathcal{B}, \mathcal{B}'} \times P_{\mathcal{B}', \mathcal{B}''} = P_{\mathcal{B}, \mathcal{B}''}.$$

- Si \mathcal{B} et \mathcal{B}' sont deux bases de E , alors $P_{\mathcal{B}, \mathcal{B}'}$ est inversible et on a :

$$(P_{\mathcal{B}, \mathcal{B}'})^{-1} = P_{\mathcal{B}', \mathcal{B}}.$$

- Matrice d'une application linéaire :

Soient $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $\mathcal{C} = (f_1, \dots, f_n)$ une base de F .

On appelle **matrice de $f \in \mathcal{L}(E, F)$ dans les bases \mathcal{B} et \mathcal{C}** , notée $M_{\mathcal{B}, \mathcal{C}}(f)$, la matrice $(a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{R})$ où pour tout $j \in \llbracket 1, p \rrbracket$, la j -ième colonne contient les coordonnées de $f(e_j)$ dans la base \mathcal{C} :

$$M_{\mathcal{B}, \mathcal{C}}(f) = \begin{matrix} & f(e_1) & \dots & f(e_j) & \dots & f(e_p) \\ \begin{matrix} f_1 \rightarrow \\ f_2 \rightarrow \\ \vdots \\ f_n \rightarrow \end{matrix} & \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,p} \\ a_{2,1} & \dots & a_{2,j} & \dots & a_{2,p} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j} & \dots & a_{n,p} \end{pmatrix} \end{matrix} \quad \text{où} \quad \forall j \in \llbracket 1, p \rrbracket, f(e_j) = \sum_{i=1}^n a_{i,j} \cdot f_i.$$

On a alors :

- Si $f, g \in \mathcal{L}(E, F)$ et si $\lambda, \mu \in \mathbb{R}$, alors :

$$M_{\mathcal{B}, \mathcal{C}}(\lambda \cdot f + \mu \cdot g) = \lambda \cdot M_{\mathcal{B}, \mathcal{C}}(f) + \mu \cdot M_{\mathcal{B}, \mathcal{C}}(g).$$

- Si $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$, alors :

$$M_{\mathcal{B}, \mathcal{C}}(g \circ f) = M_{\mathcal{C}, \mathcal{D}}(g) \times M_{\mathcal{B}, \mathcal{C}}(f).$$

- Si $f \in \mathcal{L}(E, F)$, alors :

$$f \text{ est un isomorphisme} \Leftrightarrow M_{\mathcal{B}, \mathcal{C}}(f) \text{ est inversible.}$$

Et dans ce cas, $(M_{\mathcal{B}, \mathcal{C}}(f))^{-1} = M_{\mathcal{C}, \mathcal{B}}(f^{-1})$.

- Rang d'une application linéaire et de sa matrice :

Soit $f \in \mathcal{L}(E, F)$, \mathcal{B} une base de E et \mathcal{C} une base de F .

Le rang de l'application linéaire f est égal au rang de sa matrice $M_{\mathcal{B}, \mathcal{C}}(f)$.

- Formules de changement de bases :

Soit E un espace vectoriel et \mathcal{B} et \mathcal{B}' des bases de E . On a pour tout $f \in \mathcal{L}(E)$:

$$M_{\mathcal{B}'}(f) = P_{\mathcal{B}', \mathcal{B}} M_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'} = (P_{\mathcal{B}, \mathcal{B}'})^{-1} M_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'}$$

- Matrices semblables :

- $A \in \mathcal{M}_n(\mathbb{R})$ et $B \in \mathcal{M}_n(\mathbb{R})$ sont **semblables** s'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{R})$ telle que :

$$A = PBP^{-1}.$$

- Deux matrices sont semblables si et seulement si elles représentent le même endomorphisme dans des bases différentes.
- Si A et B sont deux matrices semblables, alors :
 - * Elles ont le même rang : $\text{rg}(A) = \text{rg}(B)$.
En particulier, A est inversible si et seulement si B est inversible.
 - * Elles ont les mêmes valeurs propres : $\text{Sp}(A) = \text{Sp}(B)$.
 - * A est diagonalisable si et seulement si B est diagonalisable.

3 Probabilités

3.1 Espaces probabilisés

Opérations sur les événements

Soit I une partie de \mathbb{N} et $(A_i)_{i \in I}$ une suite d'événements d'un espace probabilisable (Ω, \mathcal{A}) .

- Définition de l'intersection et de l'union :

$$\omega \in \bigcap_{i \in I} A_i \Leftrightarrow \forall i \in I, \omega \in A_i \quad \text{et} \quad \omega \in \bigcup_{i \in I} A_i \Leftrightarrow \exists i \in I, \omega \in A_i.$$

- L'intersection et l'union sont **commutatives** et **associatives**.
- L'intersection et l'union sont **distributives** :

$$B \cap \left(\bigcup_{i \in I} A_i \right) = \bigcup_{i \in I} (B \cap A_i) \quad \text{et} \quad B \cup \left(\bigcap_{i \in I} A_i \right) = \bigcap_{i \in I} (B \cup A_i).$$

- Lois de Morgan :

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

Probabilités

- $P : \mathcal{A} \rightarrow [0, 1]$ est une **probabilité** sur un espace probabilisable (Ω, \mathcal{A}) si elle vérifie les conditions suivantes :

- $P(\Omega) = 1$.
- Si I est une partie (finie ou non) de \mathbb{N} et si $(A_i)_{i \in I}$ est une suite d'événements **deux à deux incompatibles**, alors on a :

$$P \left(\bigcup_{i \in I} A_i \right) = \sum_{i \in I} P(A_i).$$

- Pour tout événement $A \in \mathcal{A}$, $P(\overline{A}) = 1 - P(A)$. En particulier, $P(\emptyset) = 0$.
- Si $A, B \in \mathcal{A}$ sont deux événements tels que $A \subset B$, alors $P(A) \leq P(B)$.

Probabilités conditionnelles

- Soit $B \in \mathcal{A}$ un événement tel que $P(B) > 0$. Pour tout événement A , on appelle **probabilité de A sachant B** le réel :

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

- P_B est une probabilité sur l'espace probabilisable (Ω, \mathcal{A}) .
En particulier, toutes les règles de calculs de probabilités s'appliquent aussi pour P_B .

Probabilité d'une union

Cas d'une union finie :

- Si les $(A_i)_{1 \leq i \leq n}$ sont **deux à deux incompatibles**, on a :

$$P \left(\bigcup_{i=1}^n A_i \right) = \sum_{i=1}^n P(A_i).$$

- Dans le cas général, on utilisera les formules suivantes :
– Si $n = 2$,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Si $n = 3$,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - (P(A \cap B) + P(A \cap C) + P(B \cap C)) + P(A \cap B \cap C).$$

Cas d'une union dénombrable :

- Si les $(A_n)_{n \in \mathbb{N}}$ sont **2 à 2 incompatibles** alors la série $\sum_{n \geq 0} P(A_n)$ converge et on a :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n) \quad (\text{propriété de } \sigma\text{-additivité}).$$

- Si $(A_n)_{n \in \mathbb{N}}$ est une **suite croissante d'événements**, c'est-à-dire telle que $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$, alors on a :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n) \quad (\text{propriété de la limite monotone}).$$

- Dans le cas général, on a : $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P\left(\bigcup_{i=0}^n A_i\right)$.

Probabilité d'une intersection

Cas d'une intersection finie :

- Si les $(A_i)_{1 \leq i \leq n}$ sont **mutuellement indépendants**, on a :

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i).$$

- Dans le cas général, on a la **formule des probabilités composées** si $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$:

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

Cas d'une intersection dénombrable :

- Si $(A_n)_{n \in \mathbb{N}}$ est une **suite décroissante d'événements**, c'est-à-dire telle que $A_{n+1} \subset A_n$ pour tout $n \in \mathbb{N}$, alors on a :

$$P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n) \quad (\text{propriété de la limite monotone}).$$

- Dans le cas général, on a : $P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P\left(\bigcap_{i=0}^n A_i\right)$.

Système complet d'événements

Soit I une partie finie ou non de \mathbb{N} .

- Une famille $(A_i)_{i \in I}$ d'événements est un **système complet d'événements** si :

- ils sont **deux à deux incompatibles** : $\forall i \neq j, A_i \cap A_j = \emptyset$,
- $\bigcup_{i \in I} A_i = \Omega$.

- Si $(A_i)_{i \in I}$ est un **système complet d'événements**, alors $\sum_{i \in I} P(A_i) = 1$.

Formules des probabilités totales

Soit B un événement dont on cherche à calculer la probabilité. Lorsqu'on ressent un manque d'information pour obtenir $P(B)$ directement, on utilisera la démarche suivante :

- On **introduit un système complet d'événements** $(A_i)_{i \in I}$ bien choisi pour lever le manque d'information sur l'événement B (avec I une partie de \mathbb{N}).
- On **décompose l'événement B sur le système complet d'événements** $(A_i)_{i \in I}$:

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

- Comme c'est une **union d'événements deux à deux incompatibles**, on obtient :

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

- De plus, si pour tout $i \in I$, $P(A_i) \neq 0$, alors avec la **formule des probabilités composées**, on obtient :

$$P(B) = \sum_{i \in I} P(A_i)P_{A_i}(B).$$

3.2 Variables aléatoires

Fonction de répartition d'une variable aléatoire

- La **fonction de répartition** F_X d'une variable aléatoire X est définie par :

$$\forall x \in \mathbb{R}, F_X(x) = P(X \leq x).$$

La donnée de la **fonction de répartition** d'une variable aléatoire caractérise sa loi.

- Pour tout $x, y \in \mathbb{R}$ avec $x < y$,

$$P(X > x) = 1 - F_X(x) \quad \text{et} \quad P(x < X \leq y) = F_X(y) - F_X(x).$$

- F_X est la fonction de répartition d'une certaine variable aléatoire X si et seulement si
 - F_X est **croissante** sur \mathbb{R} .
 - $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
 - F_X est **continu à droite** en tout point de \mathbb{R} .

- F_X admet une limite finie à gauche en tout point $x \in \mathbb{R}$. De plus, on a l'égalité :

$$\forall x \in \mathbb{R}, P(X = x) = P(X \leq x) - P(X < x) = F_X(x) - \lim_{t \rightarrow x^-} F_X(t).$$

- F_X est continue à gauche (et donc continue) en $x \in \mathbb{R}$ si et seulement si $P(X = x) = 0$.

Indépendance de variables aléatoires

- Deux variables aléatoires X et Y sont **indépendantes** si pour tous intervalles $I, J \subset \mathbb{R}$, les événements $(X \in I)$ et $(Y \in J)$ sont indépendants :

$$P((X \in I) \cap (Y \in J)) = P(X \in I) \times P(Y \in J).$$

- Les n variables aléatoires X_1, \dots, X_n sont **mutuellement indépendantes** si pour tous intervalles $I_1, \dots, I_n \subset \mathbb{R}$, les événements $(X_1 \in I_1), \dots, (X_n \in I_n)$ sont mutuellement indépendants :

$$P((X_1 \in I_1) \cap \dots \cap (X_n \in I_n)) = P(X_1 \in I_1) \times \dots \times P(X_n \in I_n).$$

- $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires **mutuellement indépendantes** si, pour tout $n \in \mathbb{N}$, les variables aléatoires X_1, \dots, X_n sont **mutuellement indépendantes**.
- Lemme des coalitions : Soient n variables aléatoires X_1, \dots, X_n **mutuellement indépendantes**.
 - Si f_1, \dots, f_n sont n fonctions, alors $f_1(X_1), \dots, f_n(X_n)$ sont n variables aléatoires **mutuellement indépendantes**.
 - Si Y est une variable aléatoire ne dépendant que de X_1, \dots, X_p et si Z est une variable aléatoire ne dépendant que de X_{p+1}, \dots, X_n , alors Y et Z sont **indépendantes**.

Espérance d'une variable aléatoire

- Linéarité de l'espérance : Si X et Y admettent une espérance et $\lambda, \mu \in \mathbb{R}$, alors $\lambda X + \mu Y$ admet une espérance et

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

En particulier, pour tout $a, b \in \mathbb{R}$,

$$E(aX + b) = aE(X) + b.$$

- Positivité : Si X admet une espérance et $X \geq 0$ presque sûrement, alors $E(X) \geq 0$.
- Croissance : Si X et Y admettent une espérance et $X \leq Y$ presque sûrement, alors $E(X) \leq E(Y)$.
- Si X_1, \dots, X_n sont n variables aléatoires admettant toutes une espérance, alors $X_1 + \dots + X_n$ admet une espérance, et on a :

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i).$$

- Si X_1, X_2, \dots, X_n sont n variables aléatoires **mutuellement indépendantes** admettant une espérance alors $X_1 \times \dots \times X_n$ admet une espérance, et on a :

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i).$$

Variance d'une variable aléatoire

- Définition de la variance de X :

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

En particulier, $V(X) \geq 0$ (par positivité de l'espérance).

- Formule de Koenig-Huygens :

X admet un moment d'ordre 2 si et seulement si X admet une variance et alors :

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

- Pour tout $a, b \in \mathbb{R}$,

$$V(aX + b) = a^2 V(X).$$

En particulier, la variance **n'est pas** linéaire.

- Définition de l'écart type : $\sigma(X) = \sqrt{V(X)}$.

Moments et moments centrés

Soit $r \in \mathbb{N}$.

- $m_r(X) = E(X^r)$ est, sous réserve d'existence, le moment d'ordre r de X .
- $\mu_r(X) = E((X - E(X))^r)$ est, sous réserve d'existence, le moment centré d'ordre r de X .
- Si X admet un moment d'ordre r , alors elle admet un moment de tout ordre $s \leq r$.
En particulier, si X admet une variance, alors X admet une espérance (mais la réciproque est fausse).

Variable aléatoire centrée réduite

Si X est une variable aléatoire admettant une espérance et une variance non nulle, alors

$$X^* = \frac{X - E(X)}{\sqrt{V(X)}}$$

est centrée réduite, c'est-à-dire que $E(X^*) = 0$ et $V(X^*) = 1$.

Covariance d'un couple de variables aléatoires

- Soient X, Y et Z trois variables aléatoires **admettant une variance**.

– Définition de la **covariance** de X et Y :

$$Cov(X, Y) = E((X - E(X))(Y - E(Y))).$$

– Formule de Koenig-Huygens :

$$Cov(X, Y) = E(XY) - E(X)E(Y).$$

– Symétrie : $Cov(X, Y) = Cov(Y, X)$.

– Bilinearité (c'est-à-dire linéaire par rapport à chacune des variables) :

* linéarité à gauche : $\forall \lambda, \mu \in \mathbb{R}, Cov(\lambda X + \mu Y, Z) = \lambda Cov(X, Z) + \mu Cov(Y, Z)$;

* linéarité à droite : $\forall \lambda, \mu \in \mathbb{R}, Cov(X, \lambda Y + \mu Z) = \lambda Cov(X, Y) + \mu Cov(X, Z)$.

– Positivité : $Cov(X, X) = V(X) \geq 0$.

– Si X et Y sont **indépendantes**, alors $Cov(X, Y) = 0$ (mais la réciproque est fausse).

– La variable aléatoire $X + Y$ admet une variance et on a :

$$V(X + Y) = V(X) + V(Y) + 2Cov(X, Y).$$

- Si X_1, \dots, X_n sont n variables aléatoires admettant une variance, alors $X_1 + \dots + X_n$ admet une variance, et on a :

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} Cov(X_i, X_j).$$

- Si X_1, X_2, \dots, X_n sont n variables aléatoires **mutuellement indépendantes** admettant une variance, alors $X_1 + \dots + X_n$ admet une variance, et on a :

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i).$$

Coefficient de corrélation linéaire

Soit X et Y sont deux variables aléatoires admettant chacune une variance non nulle.

- Définition du **coefficient de corrélation linéaire** de X et Y :

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}.$$

- $\rho(X, Y) \in [-1, 1]$.
- $\rho(X, Y) = 1$ si et seulement si il existe $a > 0$ et $b \in \mathbb{R}$ tels que $Y = aX + b$ presque sûrement.
- $\rho(X, Y) = -1$ si et seulement si il existe $a < 0$ et $b \in \mathbb{R}$ tels que $Y = aX + b$ presque sûrement.

3.3 Variables aléatoires discrètes

Généralité sur les variables aléatoires discrètes

- Une **variable aléatoire** X est **discrète** si son support $X(\Omega)$ est au plus dénombrable (c'est-à-dire fini ou bien en bijection avec \mathbb{N}).
- Si X est une variable aléatoire discrète telle que $X(\Omega) = \{x_i \mid i \in I\}$ avec $I \subset \mathbb{N}$, alors la **loi de probabilité** de X est l'ensemble

$$\{(x_i, p_i) \mid i \in I\}$$

où $p_i = P(X = x_i)$.

- Si X une variable aléatoire discrète de loi $\{(x_i, p_i) \mid i \in I\}$, alors
 - $(X = x_i)_{i \in I}$ est un système complet d'événements.

- En particulier, $\sum_{i \in I} p_i$ converge et vaut 1.
- Réciproquement, une famille $(p_i)_{i \in I}$ de réels est une **loi de probabilité** si :
 - pour tout $i \in I$, $p_i \geq 0$,
 - $\sum_{i \in I} p_i$ converge et vaut 1.
- A partir de la loi $\{(x_i, p_i) \mid i \in I\}$ de X , on peut déterminer sa fonction de répartition F_X :

$$\forall x \in \mathbb{R}, F_X(x) = P(X \leq x) = \sum_{x_i \leq x} p_i.$$

- A partir de la fonction de répartition F_X de X , on peut déterminer sa loi :
 - $X(\Omega)$: c'est l'ensemble (au plus dénombrable) des points x_i de discontinuité de F_X ;
 - $P(X = x_i)$ pour tout $x_i \in X(\Omega)$: c'est la hauteur du saut à l'abscisse x_i , c'est-à-dire

$$\forall i \in I, p_i = P(X = x_i) = F_X(x_i) - F_X(x_{i-1})$$

(en supposant que les x_i sont indexés de façon croissante).

- Si X une variable aléatoire discrète de loi $\{(x_i, p_i) \mid i \in I\}$, alors

$$E(X) = \sum_{i \in I} x_i p_i,$$

sous réserve de convergence absolue.

- Théorème de transfert : Si X une variable aléatoire discrète de loi $\{(x_i, p_i) \mid i \in I\}$ et si g est une application de $X(\Omega)$ dans \mathbb{R} , alors

$$E(g(X)) = \sum_{i \in I} g(x_i) p_i,$$

sous réserve de convergence absolue.

Couples de variables aléatoires discrètes

- La **loi conjointe** de (X, Y) ou **loi du couple** (X, Y) est la donnée de :
 - $X(\Omega) = \{x_i, i \in I\}$ et $Y(\Omega) = \{y_j, j \in J\}$.
 - $p_{i,j} = P((X = x_i) \cap (Y = y_j))$ pour tout $(i, j) \in I \times J$.
- Si (X, Y) est un couple de variables aléatoires discrètes, alors
 - $((X = x_i) \cap (Y = y_j))_{(i,j) \in I \times J}$ est un système complet d'événements.
 - En particulier, $\sum_{(i,j) \in I \times J} P((X = x_i) \cap (Y = y_j)) = \sum_{(i,j) \in I \times J} p_{i,j} = 1$.
- Réciproquement, une famille $(p_{i,j})_{(i,j) \in I \times J}$ de réels est une **loi conjointe de probabilité** si :
 - $\forall (i, j) \in I \times J, p_{i,j} \geq 0$.
 - $\sum_{(i,j) \in I \times J} p_{i,j} = 1$.
- Si (X, Y) est un couple de variables aléatoires discrètes tel que $P(X = x_i) \neq 0$, alors la **loi conditionnelle de Y sachant $(X = x_i)$** est la donnée :
 - des valeurs y_j prises par Y sachant $(X = x_i)$;
 - des probabilités $P_{(X=x_i)}(Y = y_j)$ pour chacune de ces valeurs y_j .
- Les **lois marginales du couple** (X, Y) sont les lois de X et de Y . Elles s'obtiennent à partir de la loi du couple (X, Y) ou d'une loi conditionnelle en appliquant la formule des probabilités totales.
- Théorème de transfert double : Si (X, Y) est un couple de variables aléatoires discrètes et g est une application de $X(\Omega) \times Y(\Omega)$ dans \mathbb{R} , alors

$$E(g(X, Y)) = \sum_{(i,j) \in I \times J} g(x_i, y_j) P((X = x_i) \cap (Y = y_j)),$$

sous réserve de la convergence absolue de la "série double".

Lois discrètes usuelles

DÉNOMINATION	NOTATION	PARAMÈTRES	SUPPORT	PROBABILITÉS	ESPÉRANCE	VARIANCE
Loi uniforme	$\mathcal{U}(\llbracket 1, n \rrbracket)$	$n \in \mathbb{N}^*$	$\llbracket 1, n \rrbracket$	$P(X = k) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
	$\mathcal{U}(\llbracket a, b \rrbracket)$	$a, b \in \mathbb{N}^*, a \leq b$	$\llbracket a, b \rrbracket$	$P(X = k) = \frac{1}{b-a+1}$	$\frac{a+b}{2}$	$\frac{(b-a)(b-a+2)}{12}$
	$\mathcal{B}(p)$	$p \in]0, 1[$	$\{0, 1\}$	$P(X = 1) = p$ $P(X = 0) = 1 - p$	p	$p(1-p)$
Loi binomiale	$\mathcal{B}(n, p)$	$n \in \mathbb{N}^*, p \in]0, 1[$	$\llbracket 0, n \rrbracket$	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Loi géométrique	$\mathcal{G}(p)$	$p \in]0, 1[$	\mathbb{N}^*	$P(X = k) = (1-p)^{k-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
	$\mathcal{P}(\lambda)$	$\lambda > 0$	\mathbb{N}	$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ
LOIS FINIES						
LOIS INFINIES						

- Stabilité par somme des lois binomiales :

Soient X_1, \dots, X_n **mutuellement indépendantes** telles que : $\forall i \in \llbracket 1, n \rrbracket, X_i \hookrightarrow \mathcal{B}(n_i, p)$. Alors :

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{B}\left(\sum_{i=1}^n n_i, p\right).$$

- Stabilité par somme des lois de Poisson :

Soient X_1, \dots, X_n **mutuellement indépendantes** telles que : $\forall i \in \llbracket 1, n \rrbracket, X_i \hookrightarrow \mathcal{P}(\lambda_i)$. Alors :

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{P}\left(\sum_{i=1}^n \lambda_i\right).$$

3.4 Variables aléatoires à densité

Généralités sur les variables aléatoires à densité

Soient X une variable aléatoire et F_X sa fonction de répartition.

- X est une **variable aléatoire à densité** si :
 - F_X est **continu** sur \mathbb{R} .
 - F_X est de **classe** \mathcal{C}^1 sauf éventuellement en un nombre fini de points.
- Si on connaît F_X et que l'on cherche une **densité de probabilité** f_X d'une variable à densité X :
 - f_X est définie et positive sur \mathbb{R} ;
 - $f_X(x) = F'_X(x)$ pour tout x appartenant à \mathbb{R} éventuellement privé d'un nombre fini de points.
- Une fonction $f_X : \mathbb{R} \rightarrow \mathbb{R}$ est une **densité** d'une variable aléatoire X si et seulement si :
 - f_X est **continu** sur \mathbb{R} sauf éventuellement en un nombre fini de points ;
 - f_X est à **valeurs positives** sur \mathbb{R} ;
 - L'intégrale $\int_{-\infty}^{+\infty} f_X(t)dt$ **converge et vaut 1**.
- Si on connaît f_X et que l'on cherche la **fonction de répartition** F_X d'une variable à densité X :

$$\forall x \in \mathbb{R}, F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t)dt.$$

En particulier, la donnée d'une densité caractérise la loi.

- Calcul de probabilités : Si X est une variable à densité,
 - Pour tout réel x , $P(X = x) = 0$.
 - Pour tous réels $x < y$,

$$P(x \leq X \leq y) = P(x < X \leq y) = P(x \leq X < y) = P(x < X < y)$$

On a ainsi la formule :

$$P(x \leq X \leq y) = F_X(y) - F_X(x) = \int_x^y f_X(t)dt.$$

- Si X une variable aléatoire à densité de densité f_X alors

$$E(X) = \int_{-\infty}^{+\infty} t f_X(t)dt,$$

sous réserve de convergence absolue.

- Théorème de transfert : Si X est une variable aléatoire à densité de densité f_X et si $g : X(\Omega) \rightarrow \mathbb{R}$ est une fonction continue sauf éventuellement en un nombre fini de points, alors

$$E(g(X)) = \int_{-\infty}^{+\infty} g(t) f_X(t)dt,$$

sous réserve de convergence absolue.

Lois à densité usuelles

DÉNOMINATION	NOTATION	PARAMÈTRES	UNE DENSITÉ	LA FONCTION DE RÉPARTITION	ESPÉRANCE	VARIANCE
Loi uniforme	$\mathcal{U}([a, b])$	$(a, b) \in \mathbb{R}^2$ avec $a < b$	$x \mapsto \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b], \\ 0 & \text{sinon.} \end{cases}$	$x \mapsto \begin{cases} 0 & \text{si } x < a, \\ \frac{x-a}{b-a} & \text{si } x \in [a, b], \\ 1 & \text{si } b < x. \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Loi exponentielle	$\mathcal{E}(\lambda)$	$\lambda \in \mathbb{R}_+^*$	$x \mapsto \begin{cases} 0 & \text{si } x < 0, \\ \lambda e^{-\lambda x} & \text{si } x \geq 0. \end{cases}$	$x \mapsto \begin{cases} 0 & \text{si } x < 0, \\ 1 - e^{-\lambda x} & \text{si } x \geq 0. \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Loi normale centrée réduite	$\mathcal{N}(0, 1)$		$x \mapsto \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$	$x \mapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$	0	1
Loi normale	$\mathcal{N}(m, \sigma^2)$	$m \in \mathbb{R}, \sigma \in \mathbb{R}_+^*$	$x \mapsto \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$	$x \mapsto \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) dt$	m	σ^2

- Transformées affines d'une loi uniforme :

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

$$X \hookrightarrow \mathcal{U}([0, 1]) \Leftrightarrow Y = a + (b - a)X \hookrightarrow \mathcal{U}([a, b]).$$

- Propriété de la fonction de répartition Φ de la loi $\mathcal{N}(0, 1)$:

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

En particulier, $\Phi(0) = \frac{1}{2}$.

- Transformées affines d'une loi normale :

Si $X \hookrightarrow \mathcal{N}(m, \sigma^2)$, alors pour $a \in \mathbb{R}^*$ et $b \in \mathbb{R}$, on a :

$$aX + b \hookrightarrow \mathcal{N}(am + b, a^2\sigma^2).$$

En particulier, on a :

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

- Stabilité par somme des lois normales :

Soient X_1, \dots, X_n **mutuellement indépendantes** telles que : $\forall i \in \llbracket 1, n \rrbracket, X_i \hookrightarrow \mathcal{N}(m_i, \sigma_i^2)$. Alors :

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n m_i, \sum_{i=1}^n \sigma_i^2\right).$$

3.5 Chaînes de Markov

Théorie des graphes

- Un graphe est un ensemble constitué de points appelés **sommets** du graphe reliés éventuellement entre eux par des **arêtes**. On dit que :
 - Une arête reliant un sommet à lui-même est une **boucle**.
 - Un sommet est **isolé** si aucune arête ne le relie à un autre sommet.
 - Deux sommets reliés par une arête sont **adjacents**.
 - L'**ordre** d'un graphe est le nombre de sommets le constituant.
 - Le **degré** d'un sommet le nombre d'arêtes dont ce sommet est une extrémité.
- Formule d'Euler :

Soit n et p deux entiers naturels non nuls.

On considère un graphe constitué de n sommets notés A_1, \dots, A_n et de p arêtes. Alors :

$$\sum_{i=1}^n d_i = 2p$$

où, pour tout $i \in \llbracket 1, n \rrbracket$, d_i est le degré du sommet A_i .

- On appelle **chaîne** d'un graphe toute liste de sommets dans laquelle deux sommets consécutifs sont reliés par une arête.
Elle est dite **fermée** lorsque le sommet initial coïncide avec le sommet final.
Si chaque arête d'une chaîne fermée est parcourue une seule fois, on dit que c'est un **cycle**.
- Une **chaîne eulérienne** est une chaîne contenant toutes les arêtes du graphe, chacune étant parcourue une seule fois.
Un **cycle eulérien** désigne une chaîne eulérienne fermée.
Un graphe est dit **eulérien** s'il contient au moins un cycle eulérien.
- Un graphe est dit **connexe** si chaque sommet de ce graphe peut être relié par au moins une chaîne à n'importe quel autre sommet.

- CNS d'existence d'une chaîne eulérienne :
 G possède une chaîne eulérienne si et seulement si il possède zéro ou deux sommet(s) de degré impair.
- Caractérisation des graphes eulériens :
 G est un graphe eulérien si et seulement si tous ses sommets sont de degré pair.
- Matrice d'adjacence :
 Soit G un graphe dont les sommets sont les éléments de $\llbracket 1, n \rrbracket$.
 - On appelle **matrice d'adjacence** du graphe G la matrice $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ telle que, pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, $m_{i,j}$ désigne le nombre de chaînes de longueur 1 reliant les sommets i et j .
 - Soit $d \in \llbracket 1, n \rrbracket$. Pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, le coefficient situé en ligne i et colonne j de la matrice M^d est égal au nombre de chaînes de longueur d reliant le sommet i au sommet j .
 - Le graphe G est **connexe** si et seulement si la matrice $I_n + M + \dots + M^{n-1}$ a tous ses coefficients strictement positifs.
- On dit qu'un graphe est :
 - **orienté** si chaque arête est orientée, c'est-à-dire que chaque arête est dirigée d'un sommet vers un autre, à l'aide d'une flèche.
 - **pondéré** si chaque arête est pondérée par un réel positif appelé le **poids** de l'arête.

On adaptera les définitions précédentes dans ces cas (chaîne orientée, matrice d'adjacence...).

On pensera à réviser l'**algorithme de Dijkstra** du plus court chemin dans un graphe pondéré.

Chaînes de Markov

- On dit qu'une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ est une **chaîne de Markov** lorsque pour tout $n \in \mathbb{N}$, la loi de X_{n+1} ne dépend que de la loi de X_n (et non des lois des variables précédentes).
 Les **états** de la chaîne de Markov sont les valeurs possibles pour les variables aléatoires $(X_n)_{n \in \mathbb{N}}$.
- Un **graphe probabiliste** associé à une chaîne de Markov est un graphe orienté et pondéré dans lequel :
 - les sommets correspondent aux différents **états** de la chaîne de Markov ;
 - les valeurs sur les arcs, appelés **poids**, indiquent les probabilités de passage d'un état à l'autre.
- Si $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov dont les états sont $1, 2, \dots, p$, alors :
 - Le **n -ième état probabiliste** est la matrice ligne définie par :

$$U_n = (P(X_n = 1) \quad P(X_n = 2) \quad \dots \quad P(X_n = p)) \in \mathcal{M}_{1,p}(\mathbb{R}).$$

C'est un **vecteur-ligne stochastique** : ces coefficients sont tous positifs et de somme égale à 1.

- La **matrice de transition** est la matrice $A = (a_{i,j}) \in \mathcal{M}_p(\mathbb{R})$ définie par :

$$A = \begin{pmatrix} P_{(X_n=1)}(X_{n+1}=1) & P_{(X_n=1)}(X_{n+1}=2) & \dots & P_{(X_n=1)}(X_{n+1}=p) \\ P_{(X_n=2)}(X_{n+1}=1) & P_{(X_n=2)}(X_{n+1}=2) & \dots & P_{(X_n=2)}(X_{n+1}=p) \\ \vdots & \vdots & & \vdots \\ P_{(X_n=p)}(X_{n+1}=1) & P_{(X_n=p)}(X_{n+1}=2) & \dots & P_{(X_n=p)}(X_{n+1}=p) \end{pmatrix}.$$

C'est une **matrice stochastique par ligne** : ces coefficients sont tous positifs et la somme de chaque ligne égale à 1.

- Loi d'une chaîne de Markov :

$$\forall n \in \mathbb{N}, U_{n+1} = U_n \times A \quad \text{ce qui donne par récurrence} \quad \forall n \in \mathbb{N}, U_n = U_0 \times A^n.$$

- Soit A la matrice de transition d'une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$.
 - U est appelée **loi stationnaire** ou **état stable** de $(X_n)_{n \in \mathbb{N}}$ si :

$$UA = U \quad \text{et} \quad U \text{ est un vecteur-ligne stochastique.}$$

- Un **état stable** U , s'il existe, est tel que tU est un **vecteur propre** de tA associé à la valeur propre 1.
- La **loi limite** de $(X_n)_{n \in \mathbb{N}}$ définie, s'il existe, par le vecteur-ligne

$$\left(\lim_{n \rightarrow +\infty} P(X_n = 1) \quad \dots \quad \lim_{n \rightarrow +\infty} P(X_n = p) \right)$$

est un **état stable** de $(X_n)_{n \in \mathbb{N}}$.

3.6 Convergence, approximation et estimation

Convergence

- Inégalité de Markov :

Si X est une variable aléatoire **positive** admettant une espérance, alors

$$\forall a > 0, \quad P(X \geq a) \leq \frac{E(X)}{a}.$$

- Inégalité de Bienaymé-Tchebychev :

Si X une variable aléatoire admettant une variance, alors

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

- Loi faible des grands nombres :

Si $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires **mutuellement indépendantes**, admettant chacune la même espérance m et la même variance, alors

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P(|\overline{X_n} - m| \geq \varepsilon) = 0, \quad \text{où } \overline{X_n} = \frac{1}{n} \sum_{k=1}^n X_k.$$

- Convergence en loi :

- Cas général : Soient $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires de fonctions de répartition F_{X_n} et X une variable aléatoire de fonction de répartition F_X . Alors :

$$X_n \xrightarrow{\mathcal{L}} X \quad \Leftrightarrow \quad \forall x \in \mathbb{R} \text{ (où } F_X \text{ est continue), } \lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x).$$

- Cas discret : Soient $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire **toutes à valeurs dans** \mathbb{Z} . Alors :

$$X_n \xrightarrow{\mathcal{L}} X \quad \Leftrightarrow \quad \forall k \in \mathbb{Z}, \quad \lim_{n \rightarrow +\infty} P(X_n = k) = P(X = k).$$

- Convergence en loi de lois binomiales vers la loi de Poisson :

Si $X_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ pour tout $n \in \mathbb{N}^*$ et $X \hookrightarrow \mathcal{P}(\lambda)$, alors $X_n \xrightarrow{\mathcal{L}} X$.

- Théorème limite central :

Si $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires i.i.d. admettent une espérance m et une variance σ^2 non nulle et si $S_n = \sum_{k=1}^n X_k$ et $\overline{X_n} = \frac{S_n}{n}$, alors

$$S_n^* = \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \sqrt{n} \left(\frac{\overline{X_n} - m}{\sigma} \right) = \overline{X_n}^*$$

converge en loi vers une variable suivant une loi normale centrée réduite $\mathcal{N}(0, 1)$.

Autrement dit, pour tout $-\infty \leq a \leq b \leq +\infty$,

$$\lim_{n \rightarrow +\infty} P(a \leq S_n^* \leq b) = \lim_{n \rightarrow +\infty} P(a \leq \overline{X_n}^* \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{t^2}{2}\right) dt = \Phi(b) - \Phi(a).$$

Estimation

- On appelle **estimateur** (d'ordre n) de $g(\theta)$ toute variable aléatoire T_n de la forme $\varphi_n(X_1, \dots, X_n)$ indépendante de θ , où (X_1, \dots, X_n) est un n -échantillon.
- On dit que $[U_n, V_n]$ est un **intervalle de confiance** de $g(\theta)$ au **niveau de confiance** $1 - \alpha$ si pour tout θ ,

$$P(U_n \leq g(\theta) \leq V_n) \geq 1 - \alpha.$$

- Soit T_n un estimateur **sans biais** de $g(\theta)$ (c'est-à-dire tel que $E(T_n) = g(\theta)$) dont on connaît (un majorant de) la variance qui **ne fait pas intervenir** $g(\theta)$. Pour déterminer un intervalle de confiance de $g(\theta)$ à l'aide de l'inégalité de Bienaymé-Tchebychev, on procèdera comme suit :

– **Appliquer l'inégalité de Bienaymé-Tchebychev.**

On calcule $E(T_n)$ (qui doit valoir $g(\theta)$) et $V(T_n)$ et on applique l'inégalité de Bienaymé-Tchebychev à T_n :

$$\forall \varepsilon > 0, \quad \underbrace{P(|T_n - g(\theta)| \geq \varepsilon)}_{\text{Proba d'être hors de l'I. de C.}} \leq \underbrace{\frac{V(T_n)}{\varepsilon^2}}_{\text{Niveau de risque}}.$$

– **Fixer le niveau de confiance.**

On majore si nécessaire $\frac{V(T_n)}{\varepsilon^2}$ par une quantité qui **ne fait pas intervenir** $g(\theta)$. On fixe ensuite ε afin que ce majorant soit égal à α .

– **Expliciter l'intervalle de confiance.**

On résout l'inéquation $|T_n - g(\theta)| \leq \varepsilon$ afin d'isoler $g(\theta)$ et d'obtenir ainsi l'intervalle de confiance de $g(\theta)$ au niveau de confiance $1 - \alpha$.

- On appelle **intervalle de confiance asymptotique** de $g(\theta)$ au **niveau de confiance** $1 - \alpha$ toute suite $([U_n, V_n])_{n \in \mathbb{N}^*}$ vérifiant : pour tout $\theta \in \Theta$, il existe une suite de réels $(\alpha_n)_{n \in \mathbb{N}^*}$ à valeurs dans $[0, 1]$ et de limite α , telle que

$$\forall n \in \mathbb{N}^*, \quad P(U_n \leq g(\theta) \leq V_n) \geq 1 - \alpha_n.$$

- Soit T_n un estimateur **sans biais** de $g(\theta)$ (c'est-à-dire tel que $E(T_n) = g(\theta)$) dont on connaît (un majorant de) la variance qui **ne fait pas intervenir** $g(\theta)$. Pour déterminer un intervalle de confiance asymptotique d'un paramètre $g(\theta)$ à l'aide du théorème limite central, on procèdera comme suit :

– **Appliquer le théorème limite central.**

On calcule $E(T_n)$ (qui doit valoir $g(\theta)$) et $V(T_n)$ et on applique le théorème limite central.

On disposera ainsi d'une convergence en loi $T_n^* \xrightarrow{L} T$ avec $T \hookrightarrow \mathcal{N}(0, 1)$.

– **Fixer le niveau de confiance.**

Pour tout $a < b$, on a :

$$\lim_{n \rightarrow +\infty} P(a < T_n^* \leq b) = P(a < T \leq b) = \Phi(b) - \Phi(a).$$

On choisit a et b de telle sorte que $\Phi(b) - \Phi(a) = 1 - \alpha$.

– **Expliciter l'intervalle de confiance.**

On résout l'inéquation $a < T_n^* \leq b$ afin d'isoler $g(\theta)$ et d'obtenir ainsi l'intervalle de confiance asymptotique de $g(\theta)$ au niveau de confiance $1 - \alpha$.