A.P.M.L

ASSOCIATION DES PROFESSEURS DE MATHÉMATIQUES EN CLASSES PRÉPARATOIRES LITTÉRAIRES

site : association-apml.fr

NOTATIONS UTILISÉES ET PRÉCISIONS CONCERNANT LA DÉCLINAISON DU PROGRAMME DU CONCOURS DE L'ENS DANS LES CLASSES PRÉPARATOIRES BL

Le présent document vise à préciser les notations utilisées en cours de mathématiques dans les classes préparatoires d'hypokhâgne et khâgne BL, ainsi que la manière dont est décliné le programme du concours de l'ENS, notamment en ce qui concerne les exemples donnés dans le cours ou les situations classiques abordées en exercice. Le but est de permettre aux jurys d'avoir une idée précise de la manière dont le programme de mathématiques au concours de l'ENS est enseigné dans les différentes classes préparatoires BL, et d'éviter l'utilisation dans les sujets de notations avec lesquelles les candidats ne sont pas familiarisés. Ceci est d'autant plus important que les étudiants de BL sont nombreux à présenter à la fois le concours de la banque d'épreuves BLSES et celui de la banque BCE.

Afin de faciliter la lecture, les indications apparaissent en italique et en couleur bleu dans le corps du texte du paragraphe auquel elles se rapportent.

Classe préparatoire B/L
Programme de mathématiques

Table des matières

Objectifs généraux de la formation

Les mathématiques jouent un rôle important dans la société et ont une importance grandissante dans les sciences humaines et sociales. Les probabilités et la statistique interviennent dans tous les secteurs de l'économie et dans une grande variété de contextes (actuariat, biologie, épidémiologie, finance quantitative, prévision économique...) où la modélisation de phénomènes aléatoires à partir de bases de données est indispensable.

L'objectif de ce programme est de permettre de manière équilibrée

- une formation par les mathématiques en tant que telles ;
- l'acquisition d'outils utiles notamment aux sciences sociales et à l'économie (probabilités et statistiques, introduction aux fonctions de deux variables par exemple).

L'objectif de la formation dans les classes préparatoires B/L n'est pas de former des professionnels des mathématiques. L'enseignement de mathématiques concourt à structurer la pensée des étudiants, à développer leurs capacités d'imagination et d'abstraction, et à les former à la rigueur et à la logique en insistant sur les divers types de raisonnement (par équivalence, implication, l'absurde, analyse-synthèse...). Il permet aux étudiants d'utiliser des outils mathématiques ou d'en comprendre l'usage dans diverses situations de leur parcours académique et professionnel. L'état de l'art en sciences sociales et économie a été un guide important pour donner aux étudiants de B/L les bases dont ils auront besoin pour aller plus loin.

Le programme définit les objectifs de l'enseignement de ces classes et décrit les connaissances et les capacités exigibles des étudiants. Il précise également certains points de terminologie et certaines notations. Les limites du programme sont clairement précisées. Elles doivent être respectées aussi bien dans le cadre de l'enseignement en classe que dans l'évaluation.

1 - Compétences développées

L'enseignement de mathématiques en classes préparatoires B/L permet de développer chez les étudiants les compétences générales suivantes :

- **Rechercher et mettre en œuvre des stratégies adéquates :** savoir analyser un problème, émettre des conjectures notamment à partir d'exemples, choisir des concepts et des outils mathématiques pertinents.
- **Modéliser :** savoir conceptualiser des situations concrètes (phénomènes aléatoires ou déterministes notamment issus de problèmes de sciences sociales ou économiques) et les traduire en langage mathématique, élaborer des algorithmes.
- **Interpréter :** être en mesure d'interpréter des résultats mathématiques dans des situations concrètes, avoir un regard critique sur ces résultats.
- **Raisonner et argumenter :** savoir conduire une démonstration, confirmer ou infirmer des conjectures.
- **Maîtriser le formalisme et les techniques mathématiques :** savoir employer les symboles mathématiques à bon escient, être capable de mener des calculs de manière pertinente et efficace.
- **Communiquer par écrit et oralement :** comprendre les énoncés mathématiques, savoir rédiger une solution rigoureuse, présenter une production mathématique.

2 - Architecture des programmes

Par rapport au programme précédent, le programme d'algèbre linéaire donne une place plus importante aux aspects matriciels et introduit des bases de géométrie euclidienne qui pourront être illustrées par d'autres parties du programme. Il est important de mettre en valeur l'interaction entre les différentes parties du programme. À titre d'exemple, l'algèbre linéaire trouvera ainsi son application dans les problèmes d'optimisation, l'analyse et les probabilités dans les problèmes d'estimation.

Le programme a été rédigé sur deux années. Au sein de chaque année, aucun ordre particulier n'est imposé et chaque professeur conduit en toute liberté l'organisation de son enseignement. Le programme tient compte de l'évolution des programmes de Terminale tout en maintenant une exigence intellectuelle élevée adaptée aux étudiants de la filière B/L et à la place que prennent aujourd'hui les techniques quantitatives en sciences humaines et sociales.

Le programme se présente de la manière suivante : dans la colonne de gauche figurent les contenus exigibles des étudiants ; la colonne de droite comporte des précisions sur ces contenus et des exemples d'activités ou d'applications.

Généralités

Concernant cette partie, le vocabulaire doit être connu et un savoir-faire est attendu. Aucune difficulté théorique ne sera soulevée. Certaines des notions peuvent être introduites en situation sans faire l'objet de chapitres spécifiques.

1 - Logique

Connecteurs : et, ou, non, implication, réciproque, contraposée.	Négation d'une phrase mathématique utilisant connecteurs et quantificateurs.
Quantificateurs $\exists, \forall$.	
Raisonnement par l'absurde.	Introduit sur des exemples.
Notion de condition nécessaire et de condition suffisante.	

2 - Vocabulaire ensembliste

Appartenance, inclusion, notations $\in, \subset$. Ensemble $\mathcal{P}(E)$ des parties de E.	
Complémentaire, notation $\bar{A}$. Union, intersection, notations $\cup, \cap$. Distributivité, lois de Morgan.	Lien entre les opérations ensemblistes et les connecteurs logiques.
Définition du produit cartésien d'ensembles.	Exemples : $\mathbf{R}^2, \mathbf{R}^n$.
Applications, composition, restriction. Applications injectives, surjectives, bijectives.	Introduire ces notions en situation. Le vocabulaire doit être connu.

3 - Les nombres entiers

Notations $\mathbf{N}$ et $\mathbf{Z}$.	
Raisonnement par récurrence.	Il sera introduit sur des exemples *Sont vues en cours et utilisées en exercices, la **récurrence simple**, la **récurrence double**, et la **récurrence généralisée** (Si la propriété $\mathcal{P}(n)$ est vraie au rang initial et si, pour tout n, le fait qu'elle soit vraie pour tout entier $k \leqslant n$ implique qu'elle est vraie au rang $n+1$, alors elle est vraie pour tout n).* *Ces propriétés sont présentées sous le nom de « principe de récurrence »et ne sont pas démontrées.*
Notations $\sum, \prod$. Définition de $n!$.	
Formules $$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}; \quad \sum_{k=0}^{n} x^k = \frac{x^{n+1}-1}{x-1}.$$	Savoir les retrouver. *Exemples de formules démontrées par récurrence : $\sum\limits_{k=1}^{n} k^2$, $\sum\limits_{k=1}^{n} k^3$.*

4 - La droite réelle

Propriétés élémentaires des opérations de $(\mathbf{R}, +, \times)$.	*On revoit à cette occasion les identités remarquables classiques : factorisation de $x^n - y^n$, en particulier pour $n = 2$ et $n = 3$.* *Le développement de $(a + b)^n$ par la formule du binôme de Newton peut être démontré dans ce cadre. On signale dans le chapitre sur* $\mathbf{C}$ *que la démonstration reste valable avec a et b complexes.* *De même, en algèbre linéaire, on remarque, sans la refaire, que la démonstration reste valable avec deux matrices qui commutent pour la multiplication et avec deux endomorphismes qui commutent pour la loi de composition.*
Manipulation d'inégalités. Intervalles.	*On donne la définition basique des différents intervalles (fermé, ouvert, semi-ouvert ; infinis). La notion de convexité, qui n'est pas au programme, n'est pas évoquée. Un intervalle fermé borné est appelé « segment ».*
Valeur absolue, inégalité triangulaire.	La construction de $\mathbf{R}$ n'est pas au programme. *Sont rappelées dans le cours, souvent sans démonstration, la majoration de $\lvert a+b\rvert$ et de $\lvert a-b\rvert$ par la somme des valeurs absolues, et on généralise la propriété à n termes par récurrence.* *On démontre également que $\lvert\lvert a\rvert - \lvert b\rvert\rvert$ est majoré par $\lvert a + b\rvert$ et $\lvert a - b\rvert$.* *Dans certaines classes, afin de gagner du temps, ces propriétés sont démontrées uniquement dans le paragraphe sur* $\mathbf{C}$.
Majorant, minorant, maximum, minimum, borne supérieure, borne inférieure.	L'existence d'une borne supérieure (inférieure) pour toute partie non vide majorée (minorée) de $\mathbf{R}$ est admise. *On fait remarquer aux étudiants sur des exemples simples que la borne supérieure et la borne inférieure n'appartiennent pas forcément à l'ensemble. On définit le plus grand élément et le plus petit élément, notés Max et Min*

5 - Le plan complexe

Partie réelle, partie imaginaire, conjugué d'un nombre complexe.	On donnera l'interprétation géométrique de ces notions.
Opérations, propriétés élémentaires de $(\mathbf{C}, +, \times)$, calcul du quotient en coordonnées cartésiennes.	
Module, argument, notation exponentielle, calcul du produit et du quotient en coordonnées polaires.	Lien avec la notion informelle de coordonnées polaires.
Formules d'Euler et de Moivre.	Lien avec les formules trigonométriques. *Les formules trigonométriques que l'on demande aux étudiants de connaître sont les formules les plus utiles :* $\cos^2(x) + \sin^2(x) = 1$ *;* $\cos(a + b)$*,* $\sin(a + b)$*,* $\tan(a + b)$ *;* *(en particulier,* $\cos(2x)$ *et* $\sin(2x)$*)* *ainsi que les formules de symétrie que les étudiants retrouvent par lecture sur le cercle trigonométrique :* $\cos(\pi + x)$*,* $\cos(-x)$*, etc.* *Les autres formules de trigonométrie (exemples :* $\cos(3x)$*,* $\cos(p) + \cos(q)$*,* $\cos(a)\cos(b)$*, ...) font parfois l'objet d'exercices mais les étudiants n'ont pas à les connaître.* *La linéarisation des polynômes trigonométriques (que l'on utilise en calcul intégral) est vue en exercices.* *Pour illustrer l'utilisation de la forme trigonométrique d'un nombre complexe, on résout en exercice l'équation* $z^n = a$*, en particulier pour* $a = 1$*.*
Résolution des équations du second degré à coefficients réels.	Les formules de résolution sont exigibles, ainsi que leur démonstration.

Première année

I - Suites et séries de nombres réels

	Selon les notations habituelles, la suite de terme général u_n est notée (u_n). La notation u_n (sans parenthèses) désigne le terme d'indice n de la suite.
Limite (finie ou infinie) d'une suite. Unicité de la limite. Toute suite convergente est bornée. Opérations algébriques sur les suites convergentes. Passage à la limite dans des inégalités.	On donnera la définition sans en faire un usage systématique. *Les suites extraites sont utilisées en exercices et se limitent aux suites (u_{2n}) et (u_{2n+1}). On les utilise notamment dans les exercices faisant appel au théorème des suites adjacentes.*
Étude de convergence par encadrement.	
Comparaison entre les suites $n!, a^n, n^b, (\ln n)^c$.	
Suites monotones, limites des suites monotones.	Toute suite croissante admet une limite (finie ou infinie) qui est, si elle est majorée, la borne supérieure de ses termes (démonstration non exigible). *On démontre en cours le théorème des suites adjacentes.*
Suites définies par récurrence, suites géométriques, suites arithmétiques, suites arithmético-géométriques.	*Les suites à récurrence linéaire d'ordre 2 sont vues en cours, soit dans le paragraphe sur les espaces vectoriels, soit dans le cours des suites en admettant leur formule explicite. On les utilise en exercices par exemple pour le calcul de la puissance n-ième d'une matrice et pour décrire certains processus en probabilités.* Aucune théorie générale n'est exigible sur les suites définies par récurrence. *Ainsi, l'étude des suites du type $u_{n+1} = f(u_n)$ est vue uniquement en exercices. La méthode d'étude, utile notamment dans certains exercices de probabilités, est connue des étudiants.*
Séries à termes positifs. Somme (finie ou infinie).	On pourra montrer l'existence d'une valeur limite (finie ou infinie) par la croissance des sommes partielles. Paradoxe de Zénon. *Pour la série de terme général u_n, on note généralement S_n la somme partielle d'ordre n. On montre en cours que toute série dont les sommes partielles sont majorées par une même constante est convergente.*
$\sum_k \sum_l a_{k,l} = \sum_l \sum_k a_{k,l}$	Résultat admis (les termes sont positifs).

Séries géométriques.

La règle de d'Alembert, permettant d'étudier simplement les séries géométriques, est vue en cours. Dans certaines classes, elle est admise donc la démonstration n'est pas exigible. On applique ce théorème notamment à la convergence de la série exponentielle.

Convergence des séries de Riemann $\sum n^{-a}$.

Résultat admis, la démonstration par comparaison avec l'intégrale pourra être traitée en exercice le moment venu.

Le critère de Riemann peut être énoncé en cours et démontré en utilisant la négligeabilité. Auquel cas, il est formulé de la façon suivante :

S'il existe $\alpha > 1$ tel que $n^\alpha u_n$ tend vers $\ell \in \mathbf{R}$, alors $\sum u_n$ converge.

S'il existe $\alpha \leqslant 1$ tel que $n^\alpha u_n$ tend vers $\ell \in \mathbf{R}^*$ ou vers l'infini, alors $\sum u_n$ diverge.

Sommation d'inégalités.

Si $0 \leqslant u_n \leqslant v_n$ pour tout n, $\sum_n u_n \leqslant \sum_n v_n$. En particulier, si $\sum_n v_n$ est finie, alors $\sum_n u_n$ aussi.

Conformément à l'esprit du programme sur ce point, les méthodes vues en cours et en exercices visent à permettre d'étudier le plus simplement possible la convergence d'une série à termes positifs, en limitant la technicité des calculs. Ainsi, on montre en cours que deux séries à termes positifs dont les termes généraux sont équivalents sont de même nature. De même, on démontre que si $u_n = o(v_n)$, alors $\sum v_n$ converge $\Rightarrow \sum u_n$ converge.

$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$; pour $|x| < 1$, $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$.

Admis.

À part ces exemples, on ne fera aucune théorie sur les séries à termes non positifs.

Le théorème des séries alternées n'est pas au programme. Aussi les suites du type $\sum \frac{(-1)^n}{n}$ sont étudiées en exercice en montrant que (u_{2n}) et (u_{2n+1}) sont adjacentes.

Afin de toujours pouvoir se ramener à des séries à termes positifs, on définit la convergence absolue et on admet sans démonstration que toute série absolument convergente est convergente.

Conformément au programme, aucune théorie n'est faite sur les séries à termes de signe quelconque.

II - **Algèbre linéaire**

L'accent sera mis dans la présentation de l'algèbre linéaire sur les sous-espaces de $\mathbf{R}^n$, avec de nombreux exemples en dimension 2 ou 3 visant à développer l'intuition géométrique. Représenter un sous-espace vectoriel comme noyau d'une matrice revient à en donner un système d'équations. Représenter un sous-espace vectoriel comme image d'une matrice revient à en donner une description paramétrique. On montrera notamment comment passer d'un point de vue à l'autre.

Selon les classes, la présentation de l'algèbre linéaire est faite soit en se limitant à $\mathbf{R}^n$, soit en introduisant la notion générale d'espace vectoriel. Cette liberté pédagogique de présentation de l'algèbre linéaire a été reconnue aux professeurs à la demande de l'APML par courrier électronique de la DGESIP en date du 13 juillet 2016.

1 - L'espace $\mathbf{R}^n$

Définition de l'espace $\mathbf{R}^n$ des n-uplets de réels, interprétation géométrique comme vecteurs.

Les opérations $+ : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$ et $\cdot : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n$.

On définira les opérations par leur expression algébrique en en donnant l'interprétation géométrique.

Les propriétés :
commutativité $x + y = y + x$;
associativités $(x + y) + z = x + (y + z)$;
$(ab) \cdot x = a \cdot (b \cdot x)$;
opposé $x + (-1) \cdot x = 0$;
distributivité $(a+b)\cdot(x+y) = a\cdot x+a\cdot y+b\cdot x+b\cdot y$.

Les propriétés pourront être démontrées à partir de la définition des opérations.

Notion de combinaison linéaire, définition d'un sous-espace vectoriel de $\mathbf{R}^n$.

Classification des sous-espaces vectoriels de $\mathbf{R}^2$.

Si le sous-espace n'est pas $\{0\}$, il contient un vecteur non nul e. Si tous les éléments du sous-espace sont proportionnels à e, c'est une droite vectorielle, sinon, c'est $\mathbf{R}^2$.

Représentation d'une droite vectorielle ou affine par équation ou par paramétrisation.

2 - Matrices et systèmes linéaires

Matrices à coefficients réels. Application associée de $\mathbf{R}^m$ dans $\mathbf{R}^n$.

$(x_i) \mapsto (\sum_j A_{ij} x_j)$

Définition d'une application linéaire de $\mathbf{R}^m$ dans $\mathbf{R}^n$.

L'application associée à une matrice est linéaire.

Toute application linéaire est associée à une matrice.

Somme de matrices, produit par un réel, propriétés.

Les opérations sur $\mathbf{M}_{n,m}(\mathbf{R})$ vérifient les mêmes propriétés que celles de $\mathbf{R}^{mn}$.

Produit de matrices et composition.

Systèmes linéaires, écriture sous la forme $A(x) = y$.

En vue d'utiliser les mêmes notations que celles en pratique dans les classes de ECS et ECE, classes débouchant sur le concours de la BCE que présentent nos étudiants, les vecteurs sont écrits systématiquement avec des lettres minuscules, les matrices et les vecteurs colonnes avec des lettres majuscules.
Ainsi, si X est la matrice colonne représentant le vecteur x dans une base $\mathcal{B}$, si Y est la matrice colonne représentant y dans une base $\mathcal{B}'$, si M est la matrice de l'application linéaire f dans les bases $\mathcal{B}$ et $\mathcal{B}'$, on a l'équivalence :
$y = f(x) \iff Y = MX$.
On note avec les lettres minuscules le terme générique $a_{i,j}$ de la matrice A, ainsi que les coordonnées $x_1, x_2, \ldots, x_n$ du vecteur x. On note X la matrice colonne dont les composantes sont les coordonnées du vecteur x ; lorsqu'on écrit cette matrice en ligne, c'est sa transposée ${}^tX$.

Noyau d'une application linéaire. Une application linéaire est injective si et seulement si son noyau est réduit à $\{0\}$.

Algorithme de Gauss pour réduire un système linéaire (une matrice) à une forme échelonnée par opérations sur les lignes.

On donnera des exemples de résolution de systèmes linéaires (homogènes ou non) en utilisant l'algorithme de Gauss. Des exemples où il existe une solution unique, où il n'existe pas de solution, et où il existe plusieurs solutions seront traités. Dans ce dernier cas, on donnera une représentation paramétrique de l'ensemble des solutions.

Le système $A(x) = 0$ a des solutions non triviales si A a moins de lignes que de colonnes.

Ce résultat sera utilisé en théorie de la dimension.

Détermination d'un système d'équations pour l'image d'une matrice (écriture de l'image d'une matrice comme noyau d'une autre matrice).

Matrice transposée, produit des transposées.

$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.
La transposée d'une matrice est notée ^{t}A et non A^{T}. Ainsi :
$^{t}(AB) = {}^{t}B\,{}^{t}A$

3 - Matrices carrées inversibles

Dans cette section les matrices sont carrées (et réelles).

Matrice diagonale, triangulaire.

Résolution des systèmes linéaires associés.

Trace, trace d'un produit.

On démontre en temps voulu, en application de cette propriété, que deux matrice semblables ont même trace, qui est la trace de tout endomorphisme représenté par ces deux matrices dans deux bases différentes.
Ainsi la trace est une propriété de l'endomorphisme, indépendante de la base choisie pour représenter cet endomorphisme par une matrice.

Matrice inversible, équivalence entre :

– le système $A(x) = y$ a une unique solution pour tout y ;
– il existe une matrice carrée B telle que $AB = I = BA$.

La matrice B est alors unique, c'est l'inverse de A, notée A^{-1}.
La solution du système $A(x) = y$ est $x = A^{-1}(y)$.

Calcul de l'inverse par pivot de Gauss.
Inverse d'un produit.
Les opérations sur les lignes correspondent à des multiplications à gauche par des matrices inversibles, matrices des opérations élémentaires.

*Afin de limiter les calculs laborieux, sources d'erreurs, **les étudiants sont incités en classe à ne pas utiliser sans discernement la méthode du pivot de Gauss**, dans les exercices, pour résoudre un système ou inverser une matrice, cela en particulier lorsque des opérations simples (substitutions des inconnues ou combinaisons judicieuses d'équations) simplifient les calculs.*
En particulier, les étudiants peuvent choisir d'inverser une matrice en résolvant le système $AX = Y$ soit par la méthode du pivot de Gauss, soit par la méthode de leur choix.
On note $A \sim B$ pour signifier que A et B sont équivalentes.

Déterminant des matrices 2×2.

Une matrice 2×2 est inversible si et seulement si son déterminant est non nul.
La théorie générale des déterminants est hors programme.
*En application de cette instruction, les étudiants **ne connaissent pas** la notion de déterminant pour $n > 2$.*

4 - Sous-espaces vectoriels de $\mathbf{R}^n$

Le noyau et l'image d'une matrice sont des sous-espaces vectoriels.

Le noyau est l'ensemble des solutions du système $A(x) = 0$, l'image est l'ensemble des seconds membres y tels que le système $A(x) = y$ a une solution. L'image est aussi le sous-espace vectoriel engendré par les colonnes.

Sous-espace vectoriel engendré par une famille de vecteurs, famille génératrice.
Famille libre de vecteurs de $\mathbf{R}^n$, base d'un sous-espace vectoriel, coordonnées dans une base. Base canonique de $\mathbf{R}^n$.

Dans un sous-espace vectoriel de $\mathbf{R}^n$, une famille libre a un nombre d'éléments inférieur ou égal à celui d'une famille génératrice. Toutes les bases ont le même nombre d'éléments.
Toute famille libre de $\mathbf{R}^n$ a au plus n éléments.

Pour la démonstration, on pourra utiliser qu'un système linéaire homogène ayant moins d'équations que d'inconnues a des solutions non triviales.

Existence d'une base pour tout sous-espace vectoriel de $\mathbf{R}^n$, théorème de la base incomplète.

Trouver une base du noyau d'une matrice est une compétence exigible.

Dimension d'un sous-espace vectoriel de $\mathbf{R}^n$.

Si $E \subset F \subset \mathbf{R}^n$ sont des sous-espaces vectoriels, alors la dimension de E est inférieure ou égale à celle de F.

Dans un sous-espace vectoriel de dimension d, une famille libre à d éléments est une base, une famille génératrice à d éléments est une base.

5 - Applications linéaires entre sous-espaces vectoriels de $\mathbf{R}^n$.

Noyau et image d'une application linéaire, rang.

Représentation par une matrice dans des bases.

Changement de bases, formule $A' = Q^{-1}AP$.

Toute application linéaire de rang r peut être représentée dans des bases appropriées par la matrice

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Cet énoncé pourra être démontré en utilisant le théorème de la base incomplète ou en utilisant les opérations élémentaires. Il conduit au théorème du rang.

Théorème du rang.

La formule de Grassmann : $\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G)$, *très pratique dans les exercices, est démontrée en application du théorème du rang, ou par des considérations de bases, ou encore en considérant un supplémentaire de $F \cap G$ dans F. Les étudiants ont l'habitude de l'utiliser régulièrement en exercice.*

Isomorphismes.
Une application linéaire entre sous-espaces vectoriels de même dimension est un isomorphisme si elle est injective, si elle est surjective.

Deux sous-espaces vectoriels isomorphes ont même dimension.

En particulier $\mathbf{R}^n$ et $\mathbf{R}^m$ ne sont pas isomorphes pour $n \neq m$.

6 - Rang d'une matrice

Le rang en lignes est égal au rang en colonnes.

On pourra se ramener à la matrice équivalente de la forme $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Le rang est le nombre de lignes non nulles dans les formes échelonnées.
Toute matrice carrée inversible à droite (à gauche) est inversible.

7 - Espaces vectoriels

Un espace vectoriel de dimension n est un ensemble E muni d'une opération interne $+$, d'une opération externe $\cdot$, et d'une bijection f de E dans $\mathbf{R}^n$ qui préserve les combinaisons linéaires.

On ne considérera donc que des espaces vectoriels de dimension finie. Cette définition a pour but de permettre de discuter d'espaces autres que $\mathbf{R}^n$, mais aucune difficulté abstraite ne sera soulevée.

Applications linéaires entre espaces vectoriels, endomorphismes, isomorphismes.
Exemples : $\mathbf{R}^n, \mathbf{R}_n[x], \mathbf{M}_{n,m}(\mathbf{R}), \mathcal{L}(E, F)$.

La notion d'indéterminée ne sera pas introduite.

Sous-espaces vectoriels et restrictions d'applications linéaires.
Bases d'espaces vectoriels, la dimension est aussi le cardinal des bases.

La bijection structurelle f permet de se ramener à la théorie de $\mathbf{R}^n$.
Les étudiants peuvent démontrer qu'un sous-ensemble d'un espace vectoriel est un sous-espace vectoriel, de plusieurs manières :

- *soit en se ramenant à des équations cartésiennes qui font apparaître une famille génératrice,*
- *soit en utilisant le théorème de caractérisation des sous-espaces vectoriels,*
- *soit en montrant que l'ensemble est l'image ou l'image réciproque d'un sous-espace vectoriel par une application linéaire*

Selon les classes, l'espace des fonctions polynômes et sa base canonique pourront être donnés à titre d'exemple pour illustrer la généralité du concept d'espace vectoriel, mais conformément au programme, les exercices ont pour cadre la dimension finie.

III - **Fonctions d'une variable réelle**

Un des objectifs principaux est de savoir étudier une fonction, tracer et interpréter son graphe.

Domaine de définition, tableau de variations, graphe d'une fonction.

On notera indifféremment une fonction de la variable x, quand le contexte est sans ambiguïté, par $x \mapsto f(x)$ ou simplement $f(x)$.
On insiste en classe sur la différence entre la fonction f et le réel $f(x)$. On écrit : « la fonction f telle que $f(x) = \dots$ » ou bien « $f : x \mapsto f(x)$ » mais on se garde d'employer l'expression « la fonction $f(x) = \dots$ », l'expérience montrant que cela est source de confusion pour les étudiants.

Graphe des fonctions x^2, x^3, $\sqrt{x}$, $\sqrt[3]{x}$, $\ln x$, e^x, $\sin x$, $\cos x$, $\tan x$, $|x|$.

Ces graphes serviront d'illustrations aux concepts introduits dans cette section.
Le graphe de la fonction f telle que $f(x) = 1/x$ se rajoute à la liste des fonctions usuelles.
Les fonctions sinus, cosinus et tangente hyperboliques sont hors programme et ne sont pas connues des étudiants.
La fonction partie entière ne figurant pas explicitement au programme, elle n'est pas vue en cours. Tout énoncé l'utilisant devra donc en rappeler la définition.

Fonctions périodiques, paires, impaires, majorées, minorées, bornées.
Fonctions monotones, strictement monotones.

***La convexité est hors programme** : aucune notion relative aux parties ou aux fonctions convexes (ou concaves) n'est abordée en classe.*

1 - Limites et continuité

Limite (finie ou infinie) d'une fonction en un point, limite à gauche, limite à droite, limites en $\pm\infty$.

Les définitions seront interprétées graphiquement et illustrées par des exemples.
On remarquera, sans s'y attarder formellement, que la notion de limite en un point s'étend au cas d'une fonction non définie en ce point.

Opérations algébriques sur les limites.
Limites et composition (de deux fonctions ou d'une fonction et d'une suite).
Unicité de la limite, passage à la limite dans des inégalités.

Étude de convergence par encadrement.

$\lim_0 x^a|\ln x|^b$, $\lim_\infty x^a|\ln x|^b$, $\lim_{\pm\infty}|x|^a e^x$.

La démonstration de l'existence et de la valeur de ces limites n'est pas exigible.

L'énoncé des théorèmes de croissances comparées données aux étudiants dans le cours est le suivant : Pour tout $\alpha > 0$ et $\beta > 0$,

$$\frac{e^{\alpha x}}{x^\beta} \xrightarrow[x\to+\infty]{} +\infty, \qquad x^\beta e^{\alpha x} \xrightarrow[x\to-\infty]{} 0,$$

$$\frac{(\ln x)^\alpha}{x^\beta} \xrightarrow[x\to+\infty]{} 0, \qquad x^\beta|\ln x|^\alpha \xrightarrow[x\to0^+]{} 0$$

Continuité, opérations sur les fonctions continues.

La notion de continuité uniforme est totalement hors programme. Les étudiants ne connaissent pas cette notion. Même chose pour la notion de fonctions lipchitziennes.

Les fonctions classiques sont continues sur leur domaine de définition : ln, exp, sin, cos, x^a, $|x|$.

La continuité est vraie par définition pour exp, elle est admise pour $\sin, \cos$. Elle sera démontrée pour ln en tant que fonction réciproque de exp.

On remarque en cours, sans forcément le démontrer, que les fonctions classiques sont continues sur leur domaine de définition.

La fonction logarithme népérien est définie comme la primitive sur $\mathbf{R}_+^*$ s'annulant en 1 de la fonction $1/x$. Elle est donc continue. On en déduit la continuité de l'exponentielle, présentée comme l'application réciproque du logarithme népérien. Les équations différentielles n'étant pas au programme, on ne présente pas l'exponentielle comme solution de l'équation différentielle $y' = y$.

Exemples de prolongement par continuité.

L'image d'un intervalle fermé borné par une fonction continue est un intervalle fermé borné.

Résultat admis.

Théorème des valeurs intermédiaires.

Toute fonction continue sur un intervalle fermé borné atteint son maximum et son minimum.

Toute fonction continue et strictement monotone sur un intervalle admet une fonction réciproque sur l'intervalle image, qui est continue et strictement monotone.

La fonction arctan sera introduite comme exemple.

Graphe de la fonction réciproque par symétrie.

2 - Dérivées

Dérivée, dérivée à gauche, dérivée à droite, interprétation graphique.	La dérivabilité implique la continuité. *Les étudiants peuvent utiliser au choix la notation f' ou $\dfrac{df}{dx}$, mais la première notation est la plus utilisée en classe.* *On s'attache en classe à utiliser des notations rigoureuses sur les dérivées, en évitant toute notation du type $\left(\ln(x^2+1)\right)'$...* *Le nombre dérivée à droite est noté $f'_d(a)$, plutôt que $f'(a^+)$ qui conduit aux étudiants à penser que a^+ est un réel. Même chose pour le nombre dérivée à gauche $f'_g(a)$ au lieu de $f'(a^-)$.*
Équation de la droite tangente au graphe en un point. Fonction dérivée. Cas des fonctions classiques : ln, exp, sin, cos, x^a. Méthodes de calcul (linéarité, produit, quotient, composition).	
Dérivée de la fonction réciproque.	Application au calcul de la dérivée de arctan.
Développement limité d'ordre 1.	
Dérivée et extremums.	Une fonction dérivable sur $[a,b]$ atteint son minimum en un point x_0. Si $x_0 \in \,]a,b[$ alors $f'(x_0)=0$, si $x_0=a$ alors $f'(a^+)\geqslant 0$, si $x_0=b$ alors $f'(b^-)\leqslant 0$. Démonstration à partir du développement limité. Exemples. On fera remarquer qu'un point critique (c'est-à-dire un point où la fonction dérivée s'annule) n'est pas forcément un extremum.
Théorème de Rolle, égalité des accroissements finis.	Démonstration ce qui précède.
Inégalité des accroissements finis.	On donnera les versions de l'inégalité pour $m \leqslant f' \leqslant M$ et pour $\lvert f'\rvert \leqslant k$. Exemples d'application à la convergence de suites récurrentes.
Utilisation de la dérivée pour l'étude des variations.	Une fonction est constante sur un intervalle si et seulement si sa dérivée est identiquement nulle. Une fonction est croissante sur un intervalle si et seulement si sa dérivée est positive ou nulle.
Une fonction dont la dérivée est strictement positive est strictement croissante.	On donnera la démonstration. On remarquera, sans formalisation, que le résultat reste vrai si la dérivée s'annule en un nombre fini de points.
Dérivées d'ordre supérieur, fonctions $\mathcal{C}^k$, $\mathcal{C}^\infty$.	On insistera plus sur les conclusions et l'utilisation des résultats que sur leurs hypothèses de régularité. *La formule de Leibniz ne figure pas au programme, et n'est donc pas exigible des étudiants*

3 - Exemple d'étude de fonction : régression linéaire

On considère des données se présentant comme des couples de variables $(x_i, y_i), 1 \leqslant i \leqslant n$, où x_i est vue comme une variable explicative de y_i. En notant $\bar{x}$ et $\bar{y}$ les moyennes, on cherche un coefficient a tel que $\bar{y} + a(x - \bar{x})$ soit une bonne approximation de y. On peut pour cela minimiser la somme des écarts quadratiques

$$\sum_i (y_i - \bar{y} - a(x_i - \bar{x}))^2.$$

En étudiant la fonction de a ci-dessus, on montre que la valeur optimale de a est

$$a = \frac{\sum_i ((x_i - \bar{x})(y_i - \bar{y}))}{\sum_i (x_i - \bar{x})^2}.$$

C'est le britannique Francis Galton qui a introduit la méthode au 19^{e} siècle pour étudier la taille des enfants y_i en fonction de la taille des parents x_i, obtenant un coefficient $a \approx 2/3$. Le fait que $a > 0$ signifie que les enfants sont (en moyenne) plus grands que la moyenne lorsque les parents le sont. Le fait que $a < 1$ signifie un retour vers la moyenne, « regression towards the mean » en anglais, d'où le nom de la méthode. Les explications génétiques que Galton donna de ce phénomène sont aujourd'hui considérées comme incorrectes.
On pourra interpréter cette valeur comme

$$a = \frac{\mathrm{Cov}(x, y)}{\mathrm{Var}(x)},$$

en lien avec le cours de probabilités.

Sur ce sujet, on se limite strictement au programme : on donne uniquement cet exemple en cours, en dérivant la fonction et en dressant son tableau de variations. L'interprétation en termes de covariance est évoquée en seconde année lors de la définition de la covariance. Aucune autre notion relative à la régression linéaire n'est abordée en cours ou en exercice.

4 - Intégration

Définition informelle de l'intégrale $\int_a^b f$ comme aire algébrique.

On ne soulèvera pas de difficulté sur la régularité de f, on limitera la discussion aux fonctions continues.
Selon les classes, l'intégrale peut être définie de façon informelle comme aire algébrique ou à l'aide des primitives. Dans le premier cas, on montre que l'intégrale fonction de sa borne supérieure est une primitive de f ; dans le second cas, on admet que toute fonction continue sur un segment y possède une primitive F, on définit $\int_a^b f$ par $F(b) - F(a)$ puis on démontre que cette quantité représente l'aire sous la courbe.
Quelle que soit la présentation de l'intégrale, le fait qu'une somme de Riemann permette d'approcher l'aire sous la courbe pour les fonctions continues sur un segment est admis et justifié de façon informelle dans certaines classes, et démontré pour une fonction continue et croissante dans d'autres. La démonstration n'est donc pas exigible des étudiants.

Propriétés de l'intégrale : linéarité, relation de Chasles, monotonie. Inégalité de la moyenne.	Les propriétés ne seront pas démontrées, mais interprétées en termes géométriques. *On démontre en cours les résultats incontournables suivants :* — *Inégalité triangulaire :* *si* $a \leqslant b$, $\left\| \int_a^b f(t)dt \right\| \leqslant \int_a^b \|f(t)\|dt$ — *Pour une fonction continue et positive, il y a équivalence entre* $\int_a^b f = 0$ *et* f *est identiquement nulle sur le segment* $[a, b]$.
La fonction $\int^x f$ est une primitive de f. Les primitives de f diffèrent d'une constante additive.	Démonstration à partir des propriétés de l'intégrale.
Relation $\int_a^b f = F(b) - F(a)$.	Vraie pour toute primitive F de f.
Primitives de fonctions usuelles : e^x, x^a, $\ln x$, $\sin x$, $\cos x$.	*Se rajoute à la liste :* $1/(x + a)$. *Les méthodes d'intégration des fractions rationnelles du type* $\dfrac{\alpha x + \beta}{ax^2 + bx + c}$ *sont vues en exercices. Aucune notion générale sur la décomposition d'une fraction rationnelle en éléments simples n'est exigibles des étudiants.*
Calcul d'intégrales : intégration par parties, changements de variable.	Il n'est pas attendu des étudiants qu'ils sachent trouver eux-mêmes les bons changements de variable, sauf dans quelques cas simples comme les changements affines.

IV - **Probabilités**

L'esprit du programme de probabilités est de familiariser les étudiants au concept de variables aléatoires et de leur indépendance, dont la partie statistique, en deuxième année, peut être vue comme un aboutissement. Les variables aléatoires finies ou discrètes sont plus à envisager comme un contexte dans lequel certaines des propriétés importantes peuvent être démontrées de manière simple que comme une source d'exercices de dénombrement.

1 - Événements aléatoires

Univers Ω, ensemble $\mathcal{E}$ des événements.

$\mathcal{E}$ est un ensemble de parties de Ω. On ne soulèvera pas de difficultés sur l'ensemble $\mathcal{E}$.
On précise que l'ensemble des événements est un ensemble de parties de Ω, contenant Ω lui-même, stable par passage au complémentaire, et par intersection finie ou dénombrable.

Événement A et B, événement A ou B, événement contraire, événements incompatibles, famille complète d'événements.

Lien avec les opérations ensemblistes. On mentionnera que la réunion d'une suite d"événements est un événement.
On note $\overline{A}$ le complémentaire de A.

Probabilité $P : \mathcal{E} \to [0,1]$, axiomes :
$P(\Omega) = 1, P(\emptyset) = 0$;
$P(A) \leqslant P(B)$ si $A \subset B$;
$P(A \cup B) = P(A) + P(B)$ si $A \cap B = \emptyset$;
$P(\cup_i A_i) = \sum_i P(A_i)$ si A_i est une suite d'événements deux à deux disjoints.

Dans la mesure du possible, on note $\mathbf{P}$ *(double barre) la probabilité afin d'éviter des confusions avec la probabilité p réel de $[0,1]$.*
La formule du crible de Poincaré n'est pas au programme donc n'est pas exigible des étudiants. Les énoncés nécessitant cette formule devront donc la redonner. Le cas particulier de deux événements est évidemment traité en cours.
Les théorèmes de la limite monotone (limite de la probabilité d'une suite croissante ou décroissante d'événements) sont vus en cours, mais selon les classes, leur démonstration peut ou non être admise

Probabilité conditionnelle sachant un événement B de probabilité non nulle :

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

On remarque que $P_B : \mathcal{E} \to [0,1]$ est une probabilité.

Formule des probabilités composées :

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}).$$

Formule des probabilités totales :
$P(A) = \sum_i P(A|B_i)P(B_i)$.

(B_i) est une famille complète d'événements de probabilité strictement positive.

Formules de Bayes :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

On donnera des applications concrètes de ces formules.

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_i P(B|A_i)P(A_i)}$$

(A_i) est un système complet d'événements, tous les événements sont de probabilité non nulle. On donnera des applications concrètes.

2 - Variables aléatoires discrètes

On considère ici une variable aléatoire à valeurs dans un ensemble $\mathcal{X} \subset \mathbf{R}$ de la forme $\mathcal{X} = \{x_i, i \in I\}$, où I est soit $\mathbf{N}$, soit $\mathbf{Z}$, soit un ensemble fini.

Une variable aléatoire sur $\mathcal{X}$ est une fonction $X : \Omega \to \mathcal{X}$ telle que $X^{-1}(\{x\}) = \{X = x\}$ est un événement pour tout $x \in \mathcal{X}$.

$(\{X = x\}, x \in \mathcal{X})$ est une famille complète d'événements.

*Une variable aléatoire est définie comme une **application** (et non une fonction) de Ω dans une partie dénombrable de $\mathbf{R}$. On utilise les notations probabilistes usuelles :*
L'ensemble $X^{-1}(\{x\})$ qui est l'ensemble des ω tels que $X(\omega) = x$ est noté $(X = x)$ ou $[X = x]$.

La loi de la variable aléatoire est la fonction $x \mapsto p(x) := P(X = x)$.
$\sum_{x\in\mathcal{X}} p(x) = 1$
Pour toute partie $\mathcal{X}' \subset \mathcal{X}$,

$$P(X \in \mathcal{X}') = \sum_{x\in\mathcal{X}'} p(x).$$

Fonction de répartition, quantiles.

$F_X(x) = P(X \leqslant x)$.

Variable de Bernoulli, loi de Bernoulli.

L'indicatrice de l'événement A suit une loi de Bernoulli de paramètre $P(A)$.

Loi uniforme sur un ensemble fini.

3 - Moments des variables aléatoires discrètes réelles positives

Les variables aléatoires sont positives dans cette partie.

*Au paragraphe précédent, le programme stipule que les variables aléatoires discrètes sont à valeurs dans une partie **de $\mathbf{R}$**, de la forme $\{x_i, i \in I\}$ où I est soit $\mathbf{N}$ soit $\mathbf{Z}$ soit un ensemble fini. **Il ne limite donc pas les variables aléatoires au seul cas où elles sont positives.***
Ce paragraphe 3 invite, par ailleurs, dans la colonne « commentaires », à « évoquer l'espérance des variables aléatoires finies de signe quelconque ».
Par ailleurs, les variables aléatoires continues, vues en seconde année, sont étudiées sans conditions sur le signe (exemple, la loi normale).
*Dans un souci de cohérence, on définit donc en classe les moments d'ordre k (et donc l'espérance) d'une variable discrète **de signe quelconque** par $E(X^k) = \sum_{i\in I} x_i^k \mathbf{P}(\mathbf{X} = \mathbf{k}))$, lorsque cette série est **absolument convergente**.*
De même, la covariance de X et Y est définie en cours pour deux variables aléatoires discrètes X et Y de signe quelconque.

Espérance des variables aléatoires positives.

On évoquera l'espérance des variables aléatoires finies de signe quelconque.

Moment d'ordre k (k-moment) $E(X^k) \in [0, \infty]$.

Formule de transfert.
Une variable prenant un nombre fini de valeurs a des moments finis.
Dans l'énoncé vu en cours, le théorème de transfert est réputé valable pour toute fonction f de $\mathbf{R}$ dans $\mathbf{R}$ définie sur l'ensemble $X(\Omega)$ des valeurs prises par X.

Inégalité $E(X)^2 \leqslant E(X^2)$.

Admise, mais on pourra par la suite faire le lien avec l'inégalité de Cauchy-Schwarz. Une variable de second moment fini est donc de premier moment fini.

Propriétés de l'espérance : linéarité, monotonie.

La linéarité est admise.

Variance des variables de moment d'ordre 2 fini :
$V(X) = E((X - E(X))^2)$,
$V(X) = E(X^2) - E(X)^2$, $V(aX + b) = a^2 V(X)$.

Définition de l'écart type $\sigma(X) = \sqrt{V(X)}$.
Puisque $V(|X|) \geqslant 0$ on retrouve l'inégalité $E(|X|)^2 \leqslant E(X^2)$ dans le cas des variables de moment d'ordre 2 fini.

Si $V(X) = 0$, la variable X est constante en dehors d'un événement de probabilité nulle.
Si X est une variable de Bernoulli de paramètre p, alors $E(X) = p, V(X) = p(1 - p)$.
Inégalité de Markov :

$$P(X \geqslant a) \leqslant \frac{E(X)}{a}$$

si X est une variable aléatoire à valeurs positives et $a > 0$.

L'Inégalité de Markov est, sleon les classes, démontrée soit en scindant en deux le $\sum$, soit en utilisant les indicatrices. Dans le second cas, on remarque, dans le paragraphe sur les variables continues, que cette dernière démonstration demeure valable.

Inégalité de Bienaymé-Tchebychev :

$$P(|X - E(X)| \geqslant a) \leqslant \frac{V(X)}{a^2}$$

si X est une variable aléatoire de second moment fini et $a > 0$.

L'Inégalité de Bienaymé-Tchebychev est démontrée pour toute variable aléatoire discrète possédant une variance. Elle est déduite de l'Inégalité de Markov en prenant la variable positive $(X - E(X))^2$ et a^2 au lieu de a.

Covariance de deux variables aléatoires finies :
$\operatorname{Cov}(X, Y) = E\big((X - E(X))(Y - E(Y))\big)$
$\operatorname{Cov}(X, Y) = E(XY) - E(X)E(Y)$

Les étudiants voient en cours que le calcul de la covariance nécessite la connaissance de la loi du couple (X, Y), et donc que les lois marginales ne suffisent pas pour connaître la covariance.

Coefficient de corrélation de variables aléatoires finies :

$$\operatorname{Cor}(X, Y) = \frac{\operatorname{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}.$$

Invariance d'échelle :
$\operatorname{Cor}(X, Y) = \operatorname{Cor}(aX + b, cY + d)$
pour $a > 0$, $c > 0$.
Afin d'éviter des confusions entre les notations $Cov(X, Y)$ et $Cor(X, Y)$, le coefficient de corrélation est noté $r(X, Y)$ ou $\rho(X, Y)$.
On fait remarquer aux étudiants que l'intérêt du coefficient de corrélation par rapport à la covariance est qu'il est compris entre -1 et 1. Pour cela on démontre l'inégalité de Cauchy-Schwarz, qui sera revue en seconde année dans le cadre du produit scalaire.

4 - Indépendance

Indépendance de deux événements :
$P(A \cap B) = P(A)P(B)$.

Les événements A et B (avec $P(B) \neq 0$) sont indépendants si et seulement si $P(A|B) = P(A)$.

Indépendance de variables aléatoires discrètes.

Les variables $X_1, X_2, \dots, X_n$, à valeurs dans $\mathcal{X}_1, \dots, \mathcal{X}_n$ sont dites indépendantes si, pour toutes parties $A_1 \subset \mathcal{X}_1, \dots, A_n \subset \mathcal{X}_n$,

$$P\big((X_1 \in A_1) \cap \dots \cap (X_n \in A_n)\big) = P(X_1 \in A_1) \cdots P(X_n \in A_n).$$

Deux variables de Bernoulli B_1 et B_2 sont indépendantes si et seulement si les événements $\{B_1 = 1\}$ et $\{B_2 = 1\}$ sont indépendants.
Les variables discrètes $X_1, \dots, X_n$ sont indépendantes si et seulement si

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2) \dots P(X_n = x_n)$$

pour tout $x_1 \in \mathcal{X}_1, \dots, x_n \in \mathcal{X}_n$.

La covariance de deux variables indépendantes est nulle.

Cette condition n'est pas suffisante.

Si $X_1, X_2, \dots, X_k, Y_1, \dots, Y_l$ sont des variables indépendantes, alors $f(X_1, \dots, X_k), Y_1, \dots, Y_l$ sont indépendantes pour toute fonction f.

On ne soulèvera aucune difficulté sur la notion de fonction de plusieurs variables.
Le lemme des coalitions est énoncé en cours sous la forme suivante :
Pour toutes fonctions f et g, si $X_1, X_2, \dots, X_k$, $Y_1, Y_2, \dots, Y_\ell$ sont indépendantes, alors les deux variables $f(X_1, X_2, \dots, X_k)$ et $g(Y_1, Y_2, \dots, Y_\ell)$ sont indépendantes. Ce résultat est admis.

Si les variables $X_1, \dots, X_n$ sont indépendantes et de second moment fini, alors
$E(X_1 \dots X_n) = E(X_1) \cdots E(X_n)$
et $V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n)$.

Démonstration dans le cas fini.

5 - Processus de Bernoulli

On considère dans cette section une suite $X_i, i \in \mathbf{N}^*$ de variables indépendantes suivant une loi de Bernoulli de paramètre p. L'indépendance signifie ici que, pour tout n, les variables $X_i, 1 \leqslant i \leqslant n$ sont indépendantes.

Soit T l'indice du premier 1. Alors T est une variable aléatoire à valeurs dans $\mathbf{N}^* = \{1, 2, \ldots\}$. Sa loi est la loi géométrique :

$$P(T = n) = p(1-p)^{n-1}$$

$$P(T > n) = (1-p)^n.$$

Notation $\mathcal{G}(p)$ de la loi géométrique de paramètre p.

Les lois géométriques satisfont la propriété

$$P(T > j + k \mid T > j) = P(T > k).$$

On pourra remarquer que cette propriété caractérise les lois géométriques, mais la démonstration n'est pas exigible.

Moments : $E(T) = 1/p, V(T) = (1-p)/p^2$.

On présentera le calcul de l'espérance en admettant la dérivation sous le signe somme.

On pose $S_n = X_1 + \cdots + X_n$. La loi de S_n est la loi binomiale de paramètres n et p :

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

On introduira à cette occasion les coefficients binomiaux $\binom{n}{k}$ de manière combinatoire et la notation $\mathcal{B}(n, p)$ pour la loi binomiale de paramètres n et p.

Comme cela est souligné en préambule de la partie IV, les exercices de dénombrement ne sont pas dans l'esprit du programme.

Relations :

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

Les deux notations $\binom{n}{k}$ et C_n^k peuvent être utilisées par les étudiants.

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

Triangle de Pascal.

$$\binom{n}{k} = \binom{n}{n-k}, \quad k\binom{n}{k} = n\binom{n-1}{k-1}.$$

Relation

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

Lien avec la loi binomiale de paramètres n et $p = a/(a+b)$ lorsque a et b sont positifs.

Moments :
$E(S_n) = np, V(S_n) = nV(X_1) = np(1-p)$.

La somme de deux variables binomiales indépendantes de paramètres (k, p) et (l, p) est une variable binomiale de paramètres $(k + l, p)$.

On justifiera ce résultat en interprétant cette somme comme une somme de $(k + l)$ variables de Bernoulli indépendantes.

Loi de Poisson de paramètre λ :

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Notation $\mathcal{P}(\lambda)$.

Moments d'une variable X suivant une loi de Poisson de paramètre λ :
$E(X) = \lambda, \quad V(X) = \lambda.$

On présentera le calcul de l'espérance en admettant la dérivation sous le signe somme.

Pour les lois de Poisson, le calcul de l'espérance et de la variance se fait sans avoir à admettre la dérivation sous le signe somme infinie, ce qui constituerait une source d'erreurs pour les étudiants en analyse. La démonstration se fait pas simple décalage d'indice.

On démontrera la convergence quand $n \to \infty$ de la suite de lois binomiales de paramètres n et p_n vers la loi de Poisson de paramètre λ si $np_n \to \lambda$:

$$\forall k \quad \binom{n}{k} p_n^k (1 - p_n)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}.$$

Il s'agit ici d'une convergence pour chaque k. On ne discutera pas le concept de convergence en loi en général.

Interprétation, paradigme de Poisson :
La somme S_n d'un grand nombre de variables de Bernoulli indépendantes de petit paramètre suit approximativement la loi de Poisson de paramètre $E(S_n)$.

On illustrera ce paradigme par des exemples concrets.

Conformément aux indications du programme, la convergence est démontrée à k fixé. Hormis le fait que n doit être « grand » et p « suffisamment petit », on ne précise pas en classe les conditions pratiques d'approximation (valeurs seuil de n et p), car cela ne figure pas au programme.

Exemple : dans un texte, le nombre de coquilles par page peut être modélisé par une loi de Poisson. Si, dans une collection de qualité et de pagination homogènes, le nombre moyen de coquilles par page a été estimé à $1/2$, la probabilité qu'une page donnée ne contienne pas de coquille est de 60% environ.

La loi hypergéométrique n'est pas au programme.

Deuxième année

I - Algèbre et géométrie

Comme en première année, les scalaires sont réels, et le contexte est celui de $\mathbf{R}^n$ (ou, éventuellement, d'espaces isomorphes à $\mathbf{R}^n$ comme $\mathbf{R}_k[x]$ et $\mathbf{M}_{k,l}(\mathbf{R})$) et de ses sous-espaces.

1 - Somme directe, supplémentaire

Somme de sous-espaces vectoriels.	$E+F=\{x+y, x\in E, y\in F\}$.
Somme directe de deux sous-espaces vectoriels, caractérisation par l'intersection.	La réunion d'une base de E et d'une base de F est une base de $E\oplus F$. *La définition de la somme de n sous-espaces vectoriels est donnée en cours. De même pour la somme directe. Cette notion est utilisée dans le paragraphe sur la diagonalisation.*
Supplémentaire d'un sous-espace vectoriel dans un autre, dimension des supplémentaires.	Si F est un sous-espace vectoriel, et E un sous-espace vectoriel de F, alors E admet un supplémentaire dans F.
Projection sur E parallèlement à F. Tout endomorphisme P vérifiant $P^2=P$ est une projection.	$F=\ker P, E=\ker(P-I)$ *Les deux termes « projection » ou « projecteur »peuvent être utilisés par les étudiants.* On parlera de projection **sur** E_1 et de **direction** E_2.
Symétrie par rapport à E le long de F. Tout endomorphisme S vérifiant $S^2=I$ est une symétrie.	$F=\ker(S+I), E=\ker(S-I)$ Cette définition inclut le cas $S=I$. *On parlera de symétrie **d'axe** E_1 et de **direction** E_2.* *On appelle involution tout endomorphisme f tel que $f^2=Id_E$. On démontre en cours que f est une involution si et seulement si c'est une symétrie d'axe Ker$(f+Id_E)$ et de direction Ker$(f-Id_E)$.* *Conformément aux notations utilisées en première année, les endomorphismes projecteurs et symétries sont notés avec des lettres minuscules) (p pour un projecteur et s pour une symétrie). Les lettres majuscules P et S sont utilisées pour désigner les matrices de ces endomorphismes dans une base. De même, l'application identité est notée Id_E (la matrice identité est notée I ou I_n).*

2 - Valeurs propres des endomorphismes

Représentation d'un endomorphisme dans une base par une matrice carrée, changement de base, formule $A' = P^{-1}AP$.

Valeurs propres, vecteurs propres.

Les sous-espaces propres sont définis par $E_\lambda(f) = Ker(f - \lambda Id_E)$ et $E_\lambda(A) = Ker(A - \lambda I_n)$.
Les étudiants utilisent les polynômes annulateurs dans le cadre des exercices. Si on dispose d'un polynôme annulateur de f (ou de A), on leur demande systématiquement de montrer que toute valeur propre de f (ou de A) est une racine de ce polynôme.

Des vecteurs propres dont les valeurs propres associées sont distinctes forment une famille libre.
Endomorphisme diagonalisable. Matrice diagonalisable.

Une matrice A est diagonalisable s'il existe une matrice inversible P telle que $P^{-1}AP$ soit diagonale. Un endomorphisme est diagonalisable s'il existe une base dans laquelle il est représenté par une matrice diagonale.
On donne en cours les conditions suivantes :
f est diagonalisable SSI il existe une base de E formée de vecteurs propres de f SSI E est la somme directe de ses sous-espaces propres SSI la somme des dimensions des sous-espaces propres de f est égale à $\dim(E)$.

Une application linéaire d'un sous-espace vectoriel E dans lui-même est diagonalisable si et seulement s'il existe une base de E constituée de vecteurs propres de cette application linéaire.
Si une matrice carrée $n \times n$ admet n valeurs propres distinctes, alors elle est diagonalisable.

Conformément au programme, la connaissance de la propriété « une matrice symétrique réelle est diagonalisable » n'est pas exigible des étudiants.
On démontre en classe l'équivalence suivante, que l'on demande aux étudiants de savoir justifier : « tout endomorphisme ayant une unique valeur propre est diagonalisable SSI il est de la forme λId_E. ».

Diagonalisation pratique des matrices 2×2.

3 - Produit scalaire

Seul le produit scalaire usuel de $\mathbf{R}^n$ sera étudié, aucun développement sur les espaces euclidiens plus généraux n'est au programme.

Définition du produit scalaire $\langle x, y\rangle = \sum x_i y_i$ et de la norme $\|x\| = \sqrt{\langle x, x\rangle}$.

La convergence de la suite $(\|x^{(m)}\|)$ vers 0 est équivalente à celle de chacune des suites de coordonnées $(x_i^{(m)})$.

Bilinéarité et symétrie du produit scalaire

Formule $\langle Ax, y\rangle = \langle x, A^{\mathrm{T}} y\rangle$.

On écrit le produit scalaire $\langle x, y\rangle$ lorsque x et y sont des n-uplets, c'est-à-dire des vecteurs de $\mathbf{R}^n$ représentés par leurs composantes $(x_1, x_2, \ldots, x_n)$ et $(y_1, y_2, \ldots, y_n)$ dans la base canonique.
On écrit $\langle AX, Y\rangle$ lorsque X et Y désignent des vecteurs colonnes.
On prend bien garde en classe de ne pas mélanger les deux notations, afin d'éviter aux étudiants toute confusion. Ces notations sont conformes à celles-utilisées en première année : AX désigne le produit d'une matrice (n, n) par une matrice $(n, 1)$: On ne multiplie jamais en classe une matrice A par un vecteur x

Orthogonalité entre deux vecteurs. Théorème de Pythagore : $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ si et seulement si x et y sont orthogonaux.
Familles orthogonales, orthonormées, base orthonormée.
Toute famille orthogonale de vecteurs non nuls est libre.
Toute famille orthonormée se complète en une base orthonormée.

La recherche d'une base orthogonale ou d'une base orthonormée (algorithme de Gram-Schmidt) n'est pas connue des étudiants car ce n'est pas un attendu du programme.
Aucune propriété relative aux matrices symétriques (et antisymétriques) n'est au programme.

Coordonnées dans une base orthonormée.

$x = \sum_i \langle e_i, x\rangle e_i, \quad \|x\|^2 = \sum_i \langle e_i, x\rangle^2$

Inégalité de Cauchy-Schwarz :
$\langle x, y\rangle \leqslant \|x\|\|y\|$.

Démonstration possible : Dans une base orthonormée telle que $\|x\| e_1 = x$, on a :

$$\|y\|^2 = \sum_i \langle e_i, y\rangle^2 \geqslant \langle e_1, y\rangle^2.$$

Inégalité triangulaire $\|x + y\| \leqslant \|x\| + \|y\|$.
Distance $\|x - y\|$.

Boule, sphère de centre et de rayon donnés.

Orthogonal d'une partie, d'un sous-espace vectoriel.
L'orthogonal d'un sous-espace vectoriel en est un supplémentaire.

Pour tout sous-espace vectoriel E, $E^{\perp\perp} = E$.

Une base de l'orthogonal donne un système d'équations du sous-espace.

Hyperplans.

L'orthogonal d'un vecteur non nul est un hyperplan.
Deux vecteurs orthogonaux à un même hyperplan sont colinéaires.

Projection orthogonale P_E sur un sous-espace E muni d'une base orthonormée (e_i) (c'est la projection sur E parallèlement à $E^\perp$).

$P_E(x) = \sum_i \langle e_i, x\rangle e_i$
Le projeté $P_E(x)$ de x sur E est le point de E le plus proche de x. La distance au sous-espace E est $\|x - P_E(x)\|$.

II - Étude locale des fonctions d'une variable réelle

1 - Fonctions polynomiales

Étude de polynômes, limites en $\pm\infty$.
Racines (réelles), tout polynôme de degré impair admet une racine.

Les polynômes sont à coefficients réels et sont vus comme des fonctions d'une variable réelle.
On définit en cours le degré d'une fonction polynôme (avec $\deg 0 = -\infty$), les notions de coefficient dominant, de polynôme unitaire, de polynôme irréductible et de polynôme scindé, ainsi que les notations usuelles $\mathbf{R}[x]$ et $\mathbf{R}_n[X]$, on montre que :
$\deg(PQ) = \deg(P) + \deg(Q)$
$\deg(P+Q) \leqslant \max(\deg P, \deg Q)$.

Factorisation d'un polynôme par $(x - x_0)$ si x_0 est une racine.

L'existence d'une factorisation est admise, mais savoir factoriser en pratique est exigible.
Afin de factoriser un polynôme par $x - x_0$, les étudiants peuvent soit procéder par identification, soit effectuer la division du polynôme par $x - x_0$, la formule de la division euclidienne de $P(x)$ par $Q(x)$ est donnée sans démonstration, et on montre aux étudiants comment poser la division en pratique.

Multiplicité d'une racine : la racine x_0 est de multiplicité k si $f(x) = (x - x_0)^k g(x)$, avec $g(x_0) \neq 0$.
Un polynôme change de signe en une racine si et seulement si sa multiplicité est impaire.

Si x_0 est une racine de f de multiplicité $k \geqslant 2$, alors c'est une racine de f' de multiplicité $k-1$.
Afin d'illustrer ce qui a été vu sur les nombres complexes, on fait remarquer aux étudiants sur des exemples simples qu'un polynôme à coefficients réels peut posséder des racines complexes non réelles. On signale sans démonstration, le théorème de d'Alembert-Gauss en disant que tout polynôme de degré n à coefficients réels possède dans $\mathbf{C}$ n racines, chacune comptée autant de fois que son ordre de multiplicité.

Un polynôme f admet un extremum local en x_0 si et seulement si x_0 est une racine de f' de multiplicité impaire.

Les étudiants doivent connaître le cas
$f'(x_0) = 0$ et $f''(x_0) \neq 0$.

2 - Développements limités

On se limitera autant que possible à l'ordre 2 ou 3 et on ne cherchera aucune technicité.

Développements limités, formule de Taylor Young (admise).

Quelques exemples simples de calcul de limite à l'aide de développements limités.
Les étudiants savent faire une somme, un produit, une composition et une intégration de développements limités.
La division suivant les puissances croissantes n'est pas au programme, les étudiants calculent le développement limité d'un quotient en effectuant la composition du développement du dénominateur par celui de la fonction $\frac{1}{1+x}$.

Unicité du développement limité.

Développement limité de e^x et $(1-x)^{-1}$ à tout ordre. Premiers termes de $(1+x)^a$ et $\ln(1+x)$.

Les DL des fonctions usuelles sont connus (exponentielle, sinus, cosinus, $(1+x)^\alpha$, $1/(1+x)$, $1/(1-x)$, $\ln(1+x)$, $\ln(1-x)$.
Conformément à l'indication du programme, on se limite en classe autant que possible à l'ordre 3, tout en faisant remarquer que la régularité de ces DL permet de deviner les termes suivants.

Allure locale du graphe d'une fonction admettant un développement limité du type

$$f(x) = a_0 + a_1 x + a_k x^k + x^k \varepsilon(x),$$

où $k \geqslant 2$ et $a_k \neq 0$.

La forme du graphe en un point dépend principalement du premier terme non linéaire du développement limité. Exemples avec $k = 2$ ou $k = 3$.
Pour simplifier les calculs (notamment lors des opérations sur les DL), on utilise systématiquement en classe les notations de Landau : un DL se termine par $o(x^n)$ et non par $x^n\varepsilon(x)$.
Les équivalents de toutes les fonctions usuelles sont connus des étudiants. Ils les trouvent en prenant le premier terme non nul du DL.
De façon générale, on donne la définition suivante : f et g sont équivalentes en x_0 SSI au voisinage de x_0, $f(x) = g(x)(1 + o(1))$ ce qui revient, dans le cas où g ne s'annule pas sur un voisinage de x_0 à $\lim\limits_{x\to x_0} \frac{f(x)}{g(x)} = 1$. (Cette dernière formulation simple est fréquemment utilisée par les étudiants pour prouver l'équivalence de deux fonctions).
Des formulations identiques sont données pour les suites dans le cours de première année.

Lien avec les extremums locaux et les points d'inflexion.

En un minimum local à l'intérieur du domaine de définition, le coefficient a_2 d'ordre 2 du développement limité vérifie $a_2 \geqslant 0$.
Un point où $a_1 = 0$ et $a_2 > 0$ est un minimum local.

Détermination de l'asymptote oblique d'une fonction en l'infini. Position par rapport à l'asymptote.

Étude de la fonction e^{-x^2}.

La forme du graphe doit être connue.

3 - Intégrales généralisées

Notion d'intégrale généralisée pour des fonctions positives, du type

$$\int_a^b f \in [0, \infty]$$

où a ou b peuvent être infinis et où f est continue sur $]a, b[$.

On donnera la définition comme limite tout en évoquant une notion informelle directe comme une aire (finie ou infinie).

Les deux terminologies pourront être employées : intégrale convergente (divergente) ; intégrale finie (infinie).

Les deux termes « intégrale généralisée » ou « intégrale impropre »peuvent être utilisés par les étudiants. Ainsi, on peut dire que $\int_0^1 1/x dx$ est impropre en 0.

Lorsque la fonction qui présente un problème en x_0 est prolongeable par continuité en ce point, on dit que l'intégrale est « faussement impropre en ce point »(exemple : $\int_0^{1/2} \frac{1}{\ln(x)} dx$).

L'expression « f est intégrable sur $\mathbf{R}$ », qui vient de la théorie de la mesure, n'est pas connue des étudiants. On ne dit pas que f est intégrable sur $\mathbf{R}$, mais que son intégrale sur $\mathbf{R}$ converge.

Extension des propriétés de linéarité, de monotonie, et de la relation de Chasles à ce cadre.

Admises .

On étend la définition de l'intégrale aux fonctions continues par morceaux, ce cas pouvant se rencontrer avec les densités.

On fait remarquer en cours que la convergence de l'intégrale en $+\infty$ n'impose pas à la fonction de tendre vers 0, contrairement à ce qui se passe pour les séries. Cela permet aux étudiants de comprendre la différence entre les cas discret et continu.

$\int_0^\infty e^{-t} \mathrm{d}t = 1$.

$\int_{-\infty}^\infty e^{-t^2/2} \mathrm{d}t = \sqrt{2\pi}$.

Valeur admise, on pourra toutefois démontrer la convergence par comparaison.

Conditions de convergence de $\int_0 t^{-a} \mathrm{d}t$ et $\int^\infty t^{-a} \mathrm{d}t$.

Afin de simplifier l'étude de la convergence des intégrales généralisées, on procède comme pour les séries en montrant les propriétés suivantes au point posant problème (en général 0 ou $+\infty$) :

- *Si $f \sim g$, alors leurs intégrales sont de même nature (les fonctions sont supposées positives).*
- *Si $f = o(g)$, alors la convergence de l'intégrale de g implique celle de l'intégrale de f.*
- *Critère de Riemann en $+\infty$ que l'on énonce ainsi :*
  S'il existe $\alpha > 1$ tel que $x^\alpha f(x) \underset{x \to +\infty}{\to} \ell \in \mathbf{R}$, alors l'intégrale de f en $+\infty$ converge.
  S'il existe $\alpha \leqslant 1$ tel que $x^\alpha f(x) \underset{x \to +\infty}{\to} \ell \in \mathbf{R}^* \cup \{+\infty\}$, alors l'intégrale de f en $+\infty$ diverge.

— *De même, on énonce le critère de Riemann en 0.*
Pour réaliser une intégration par parties d'une intégrale convergente généralisée en b, le professeur peut permettre aux étudiants d'intégrer directement entre a et b, à condition de penser à justifier que le crochet a une limite finie en b. Dans d'autres classes, les étudiants ont l'obligation de repasser sur un segment $[a, c]$, avec $a < c < b$ puis de faire tendre c vers b.
Afin de simplifier les calculs des moments d'une variable continue, on définit la convergence absolue d'une intégrale d'une fonction dont le signe n'est pas fixe, et on admet sans démonstration que la convergence absolue implique la convergence.

III - **Fonctions de deux variables réelles**

Le niveau de formalisme de cette partie sera minimal. On ne cherchera pas à préciser les hypothèses générales des résultats. Aucune notion précise sur la classe de régularité d'une fonction de deux variables n'est exigible. Les fonctions seront le plus souvent définies sur le plan $\mathbf{R}^2$ tout entier. Dans le cas contraire, on ne soulèvera aucune difficulté liée au bord de l'ensemble de définition.

Notation $f(x) = f(x_1, x_2)$.

1 - Exemples

Graphe, lignes de niveau. Étude d'exemples, notamment les suivants (allure du graphe et des lignes de niveau) :

Fonctions coordonnées $(x_1, x_2) \mapsto x_1$ et $(x_1, x_2) \mapsto x_2$.

Fonctions linéaires $ax_1 + bx_2$.

Le vecteur (a, b) est orthogonal aux droites de niveau.

2 - Dérivées partielles

Dérivées partielles.

Notations $\partial_1 f$, $\partial_2 f$.
Pour être en cohérence avec les notations utilisées dans le cours d'économie (où l'on précise le nom de la variable par rapport à laquelle on dérive), on utilise surtout les notations $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial^2 f}{\partial x \partial y}$, etc. plutôt que $\partial_1 f$, $\partial_{1,2} f$, etc.

Définition d'un extremum local, condition nécessaire : $\partial_1 f = 0 = \partial_2 f$.
Point critique, tout minimum local est un point critique.

Le point x_0 est un minimum local s'il existe $\delta > 0$ tel que $f(x) \geqslant f(x_0)$ dans le disque $B(x_0, \delta)$. On ne discutera pas d'extremum atteint au bord du domaine de définition.
Démonstration des conditions d'optimalité à l'aide des fonctions $t \mapsto f(t, x_2)$ et $t \mapsto f(x_1, t)$.

3 - Fonctions quadratiques

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

La fonction quadratique a un extremum local (strict) en 0 si et seulement si le discriminant $\Delta = b^2 - ac$ est (strictement) négatif.
Détermination (dans les cas $\Delta < 0$) du type d'extremum en fonction du signe de a et c.

4 - Retour sur la régression linéaire

Il n'est pas demandé aux étudiants de connaître les résultats de cet exemple, mais de savoir les retrouver.

Dans la présentation classique de la régression linéaire, on cherche deux coefficients a et b tels que $b + ax$ soit une aussi bonne approximation que possible de y. On minimise pour cela la fonction

$$f(a, b) = \sum_i (y_i - b - ax_i)^2,$$

et on retrouve $a = \mathrm{Cov}(x, y)/\mathrm{Var}(x)$, $b = \bar{y} - a\bar{x}$.

On trouve le point critique en calculant les dérivées partielles. On remarque que c'est la même approximation que dans la première approche.
On pourra justifier que c'est un minimum par l'interprétation géométrique suivante. Dans l'espace $\mathbf{R}^n$ (n est le nombre de données), on considère $e = (1, \ldots, 1)$, $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$. La fonction $f(a, b)$ s'interprète comme le carré de la distance entre les points y et $ax + be$. Les valeurs optimales de a et b correspondent donc à l'unique point $ax + be$ du sous-espace $E = \mathrm{Vect}(e, x)$ tel que $y - (ax + be)$ est orthogonal à E. En écrivant le système

$$\langle y - (ax + be), x \rangle = 0, \langle y - (ax + be), e \rangle = 0,$$

on retrouve bien

$$a = \frac{\langle e, e \rangle \langle x, y \rangle - \langle x, e \rangle \langle y, e \rangle}{\langle e, e \rangle \langle x, x \rangle - \langle x, e \rangle^2}, \quad b = \frac{\langle y, e \rangle - a\langle x, e \rangle}{\langle e, e \rangle}.$$

Ce qui est en vu en classe se ***limite strictement à l'exemple du programme***

5 - Étude des points critiques

Dérivées partielles d'ordre 2.

Notations $\partial^2_{1,1}f, \partial^2_{1,2}f, \partial^2_{2,1}f, \partial^2_{2,2}f$

Égalité $\partial^2_{1,2}f(x) = \partial^2_{2,1}f(x)$.

Admis.

L'égalité des dérivées partielles d'ordre 2 est présentée sous le nom de théorème de Schwarz.

Soit x un point critique de f. Si le discriminant de la fonction quadratique

$$q(y_1, y_2) = \partial^2_{1,1}f(x)y_1^2 + 2\partial^2_{1,2}f(x)y_1y_2 + \partial^2_{2,2}f(x)y_2^2$$

est strictement négatif, alors la fonction f a un extremum local en x, qui est de même nature que celui de la fonction q en 0.

Admis.

IV - **Probabilités**

1 - Variables aléatoires à densité

Soit ρ une fonction positive de $\mathbf{R}$ dans $\mathbf{R}$. La fonction $X : \Omega \to \mathbf{R}$ est une variable aléatoire de densité ρ si, pour tout intervalle $]a,b[$ de $\mathbf{R}$, l'ensemble $\{X \in]a,b[\}$ est un événement et

$$P(X \in]a,b[) = \int_a^b \rho(x)\,dx.$$

On se limitera au cas de densités ρ continues (sauf éventuellement en un nombre fini de points).
On remarquera que

$$\int_{-\infty}^{\infty} \rho = P(\Omega) = 1.$$

La densité de la variable aléatoire X est usuellement notée f (et non ρ).

Si X est une variable à densité, alors
$P(X = x) = 0$ pour tout $x \in \mathbf{R}$

Un événement de probabilité nulle n'est pas forcément impossible.

Fonction de répartition :
$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} \rho$.
Inverse (réciproque) de la fonction de répartition et quantiles.

La fonction de répartition caractérise la densité (si la densité est continue).

Exemples de calcul de fonctions de répartition et de densités de variables images $f \circ X$ avec f monotone.

Loi uniforme sur un intervalle borné.

Moments :
$\int_{\mathbf{R}} |x|^k \rho(x)dx = \int_0^{\infty} x^k(\rho(x) + \rho(-x))dx$.
Espérance des variables à densité de premier moment fini :

$$E(X) = \int_0^{\infty} x\rho(x)dx - \int_0^{\infty} x\rho(-x)dx.$$

Cette expression de l'espérance permet de ne manipuler que des intégrales de fonctions positives.
*Afin de simplifier la présentation et les calculs, on dit qu'une variable X a un moment d'ordre k si $\int x^k f(x)dx$ est **absolument convergente**. Le moment est alors simplement défini par $\int x^k f(x)dx$, sans qu'il soit besoin de se ramener à $\mathbf{R}^+$ en faisant intervenir $f(x)$ et $f(-x)$. Même chose pour l'espérance.*

Propriétés de l'espérance : linéarité, monotonie.

Admises.

Variance $V(X) = E((X - E(X))^2)$,
$V(X) = E(X^2) - E(X)^2$, $V(aX+b) = a^2 V(X)$.

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

La variable

$$X^* := \frac{X - E(X)}{\sigma(X)}$$

est centrée réduite.

Elle a pour densité

$$\rho^*(x) = \sigma(X)\rho(E(X) + \sigma(X)x).$$

Inégalité de Markov :

$$P(X \geqslant a) \leqslant \frac{E(X)}{a}$$

si X est une variable aléatoire à valeurs positives et $a > 0$.
Inégalité de Bienaymé-Tchebychev :

$$P(|X - E(X)| \geqslant a) \leqslant \frac{V(X)}{a^2}$$

si X est une variable aléatoire de second moment fini et $a > 0$.

2 - Loi normale, loi exponentielle

Une variable aléatoire gaussienne (ou normale) centrée réduite est une variable X admettant la densité

$$\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

On a $E(X) = 0$, $V(X) = 1$.
La variable $Y = \sigma X + E$ est alors une variable gaussienne (ou normale) de moyenne E et de variance σ^2, elle admet la densité

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-E)^2/2\sigma^2}.$$

On utilise la notation suivante : $\mathcal{N}(m, \sigma^2)$.

Loi exponentielle : $\rho(x) = \lambda e^{-\lambda x}$ sur $\mathbf{R}_+$
$E(X) = 1/\lambda$, $V(X) = 1/\lambda^2$.
Pour tout $x \geqslant 0$, $P(X > x) = e^{-\lambda x}$.
Les lois exponentielles satisfont la propriété

$$P(X > t + s | X > t) = P(X > s),$$

pour tous réels s et t positifs ou nuls.

Cette propriété caractérise les lois exponentielles (démonstration non exigible).
Les lois exponentielles s'interprètent comme les lois de durée de vie sans vieillissement. Ce sont des variantes continues des lois géométriques.

3 - Indépendance de variables à densité

Les variables aléatoires à densité $X_1, \ldots, X_n$ sont dites indépendantes si

$$P((X_1 \in I_1) \cap \cdots \cap (X_n \in I_n)) = P(X_1 \in I_1) \cdots P(X_n \in I_n)$$

pour tous intervalles I_k de $\mathbf{R}$.

On limitera l'utilisation de cette définition.

Si $X_1, \ldots, X_n$ sont indépendantes :
$E(X_1 \cdots X_n) = E(X_1) \cdots E(X_n)$,
$V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n)$.

Propriété admise. La loi d'une somme de variables aléatoires est hors programme.

On pourra, comme en première année, introduire la covariance et le coefficient de corrélation sans discuter les problèmes de définition des intégrales.
Cov $(X, Y) = E(X - E(X))(Y - E(Y)))$,
Cov $(X, Y) = E(XY) - E(X)E(Y)$,
Cor $(X, Y) = \dfrac{\text{Cov }(X, Y)}{\sqrt{V(X)V(Y)}}$.
La covariance de deux variables indépendantes est nulle (condition non suffisante).
Les couples de variables à densité et les produits de convolution sont ***strictement hors programme****. Les étudiants ne savent donc pas calculer une covariance dans le cas des variables à densité.*

4 - Statistiques

Soit X une variable aléatoire (discrète ou à densité). Soient X_i, $i \in \mathbf{N}^*$, des variables aléatoires indépendantes de même loi que X. On considère les variables aléatoires $\bar{X}_n = (X_1 + \cdots + X_n)/n$.
$E(\bar{X}_n) = E(X), V(\bar{X}_n) = V(X)/n.$

Loi faible des grands nombres :

$$P(|\bar{X}_n - E(X)| \geqslant \varepsilon) \to 0$$

pour tout $\varepsilon > 0$.
Intervalle de confiance :
La probabilité que l'intervalle

$$\left[\bar{X}_n - \sqrt{\frac{V}{na}}, \bar{X}_n + \sqrt{\frac{V}{na}}\right]$$

contienne $E(X)$ est supérieure ou égale à $1 - a$.

Démonstration dans le cas $V(X) < \infty$
par l'inégalité de Bienaymé-Tchebychev :

$$P(|\bar{X}_n - E(X)| \geqslant \varepsilon) \leqslant \frac{V(\bar{X}_n)}{\varepsilon^2} = \frac{V(\bar{X})}{n\varepsilon^2}.$$

On discutera la notion d'intervalle de confiance au niveau $1 - a$. On démontrera l'énoncé ci-contre à l'aide de l'inégalité de Bienaymé-Tchebychev.
En pratique :
La variance V est souvent inconnue, mais on peut la majorer :
dans le cas général d'une variable aléatoire bornée $|X| \leqslant M$, par M^2 ;
dans le cas d'une variable de Bernoulli, par $1/4$.
Application numérique :
Pour $n = 1000$, au seuil de confiance de 90%, l'incertitude est de 5% pour une variable de Bernoulli.
La notion d'intervalle de confiance vue en cours et en exercice se limite à la propriété et à l'illustration qui figurent au programme.
La notion d'estimateur n'est pas au programme.