# Classes préparatoires aux grandes écoles

# Filière économique

# Voie générale
# ECG

## Annexe II
## Programmes de mathématiques approfondies - informatique

MINISTÈRE
DE L'ENSEIGNEMENT SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION

# Classes préparatoires aux grandes écoles

# Programme de mathématiques approfondies – informatique de la classe d'ECG 2nde année

# Table des matières

# 1 Objectifs généraux de la formation

Les mathématiques jouent un rôle important en sciences économiques et en gestion, dans les domaines notamment de la finance ou de la gestion d'entreprise, de la finance de marché, des sciences 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 la formation dans les classes préparatoires économiques et commerciales n'est pas de former des professionnels des mathématiques, mais des personnes capables d'utiliser des outils mathématiques ou d'en comprendre l'usage dans diverses situations de leur parcours académique et professionnel.

Les programmes définissent les objectifs de l'enseignement de ces classes et décrivent les connaissances et les capacités exigibles des étudiants. Ils précisent 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.

Une fonction fondamentale de l'enseignement des mathématiques dans ces classes est de structurer la pensée des étudiants et de les former à la rigueur et à la logique en insistant sur les divers types de raisonnement (par équivalence, implication, l'absurde, analyse-synthèse...).

# 2 Compétences développées

L'enseignement de mathématiques en classes préparatoires économiques et commerciales vise en particulier à développer chez les étudiants les compétences 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) 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. Utiliser avec discernement l'outil informatique.

 

- **Communiquer par écrit et oralement :** comprendre les énoncés mathématiques, savoir rédiger une solution rigoureuse, présenter une production mathématique.

## 3 Architecture des programmes

Le programme de mathématiques de deuxième année de la filière EC de mathématiques approfondies se situe dans le prolongement de celui de première année et permet d'en consolider les acquis. Son objectif est de fournir aux étudiants le bagage nécessaire pour suivre les enseignements spécialisés de mathématiques, économie et gestion dispensés en Grande École ou en troisième année de Licence à l'université.

Il s'organise autour de quatre points forts :

- En algèbre linéaire et bilinéaire, on introduit la réduction des endomorphismes et des matrices carrées ainsi que les structures euclidiennes. Ces notions d'algèbre linéaire trouveront des applications en analyse lors de l'optimisation des fonctions de plusieurs variables, mais aussi en probabilités et en analyse de données (statistiques descriptives bivariées).
- En analyse, on complète l'étude des intégrales généralisées débutée en première année de classe préparatoire et on introduit les fonctions de plusieurs variables définies sur $\mathbf{R}^n$ ainsi que la notion de gradient. Au quatrième semestre, on poursuit cette étude dans le but de résoudre des problèmes d'optimisation avec ou sans contraintes, cruciaux en économie et en finance.
- En probabilités, l'étude des variables aléatoires discrètes, initiée au lycée et poursuivie en première année, se prolonge au troisième semestre par l'étude des couples et des suites de variables aléatoires discrètes ; au quatrième semestre, les notions sur les variables aléatoires à densité sont complétées. L'ensemble des notions sera présenté en lien avec la simulation informatique des phénomènes aléatoires. Un des objectifs est de permettre, en fin de formation, une approche plus rigoureuse (et une compréhension plus aboutie) des méthodes de l'estimation ponctuelle ou par intervalles de confiance.
- L'informatique est enseignée tout au long de l'année en lien direct avec le programme de mathématiques. Cette pratique régulière permettra aux étudiants de construire ou de reconnaître des algorithmes relevant par exemple de la simulation de lois de probabilité, de la recherche d'extrema en analyse ou de différentes techniques d'estimation.

Au fur et à mesure de la progression, on aura à cœur de mettre en valeur l'interaction entre les différentes parties du programme.

Le programme de mathématiques est organisé en deux semestres. Ce découpage en deux semestres d'enseignement doit être respecté. En revanche, au sein de chaque semestre, aucun ordre particulier n'est imposé et chaque professeur conduit en toute liberté l'organisation de son enseignement, bien que la présentation par blocs soit fortement déconseillée.

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 ou des exemples d'activités ou d'applications.

Les développements formels ou trop théoriques doivent être évités. Ils ne correspondent pas au cœur de formation de ces classes préparatoires.

Les résultats mentionnés dans le programme seront admis ou démontrés selon les choix didactiques faits par le professeur. Pour certains résultats, marqués comme «admis», la présentation d'une démonstration en classe est déconseillée.

Les séances de travaux dirigés permettent de privilégier la prise en main, puis la mise en œuvre par les étudiants, des techniques usuelles et bien délimitées, inscrites dans le corps du programme. Cette maîtrise s'acquiert notamment par l'étude de problèmes que les étudiants doivent *in fine* être capables de résoudre par eux-mêmes. Le symbole $\boxed{\blacktriangleright}$ indique les parties du programme pouvant être traitées en liaison avec l'informatique.

Le langage Python comporte de nombreuses fonctionnalités permettant d'illustrer simplement certaines notions mathématiques. Ainsi, on utilisera avec pertinence l'outil informatique en cours de mathématiques pour visualiser et illustrer les notions étudiées.
Les travaux pratiques de mathématiques avec Python sont organisés autour de quatre thèmes faisant intervenir divers points du programme de mathématiques. L'objectif est d'apprendre aux étudiants à utiliser Python de manière judicieuse et autonome ainsi que de leur permettre d'illustrer ou de modéliser des situations concrètes en mobilisant leurs connaissances mathématiques. Les savoir-faire et compétences que les étudiants doivent acquérir lors de ces séances de travaux pratiques sont spécifiés dans la liste des exigibles et rappelés en préambule de chaque thème. Les nouvelles notions mathématiques introduites dans certains thèmes ne font pas partie des exigibles du programme. L'enseignement de ces travaux pratiques se déroulera sur les créneaux horaires dédiés à l'informatique.

# ENSEIGNEMENT DE MATHÉMATIQUES DU TROISIÈME SEMESTRE

## I - Algèbre linéaire et bilinéaire

### 1 - Compléments d'algèbre linéaire

#### a) Somme directe de sous-espaces vectoriels

| | |
|---|---|
| Dimension d'une somme directe de $k$ espaces vectoriels. | Base adaptée à une somme directe. |
| Concaténation de bases de sous espaces vectoriels. | Caractérisation de sommes directes par concaténation des bases. |

#### b) Changement de base

| | |
|---|---|
| Matrice d'un endomorphisme dans une base. | Rappels. |
| Matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$. | Notation $P_{\mathcal{B},\mathcal{B}'}$. |
| $P^{-1}_{\mathcal{B}',\mathcal{B}} = P_{\mathcal{B},\mathcal{B}'}$. | |
| Formules de changement de base. | $X_{\mathcal{B}} = P_{\mathcal{B},\mathcal{B}'} X_{\mathcal{B}'}$.<br>$\mathrm{Mat}_{\mathcal{B}'}(f) = P^{-1}_{\mathcal{B},\mathcal{B}'} \,\mathrm{Mat}_{\mathcal{B}}(f)\, P_{\mathcal{B},\mathcal{B}'}$. |
| Matrices semblables. | Deux matrices $A$ et $B$ carrées sont semblables s'il existe une matrice inversible $P$ telle que $B = P^{-1}AP$.<br>$A$ et $B$ sont semblables si et seulement si elles représentent les matrices d'un même endomorphisme dans des bases différentes. |

#### c) Trace

*La trace d'une matrice carrée est introduite uniquement comme outil simple et efficace en vue de la recherche de valeurs propres. Tout développement théorique est exclu. Aucun autre résultat concernant la trace n'est au programme.*

| | |
|---|---|
| Trace d'une matrice carrée.<br>Linéarité de la trace et $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. | Notation $\mathrm{Tr}(A)$. |
| Invariance de la trace par changement de base. | $\mathrm{Tr}(A) = \mathrm{Tr}(P^{-1}AP)$. |

## 2 - Éléments propres des endomorphismes et des matrices carrées, réduction

*Les espaces vectoriels considérés dans ce chapitre sont définis sur* $\mathbf{R}$*. Dans toute cette partie,* $f$ *désignera un endomorphisme d'un espace vectoriel* $E$ *de dimension finie, et* $A$ *une matrice carrée.*

### a) Vecteurs propres et espaces propres

Valeurs propres, vecteurs propres, sous-espaces propres d'un endomorphisme de $E$ et d'une matrice carrée.

Valeurs propres des matrices triangulaires.

Spectre d'un endomorphisme et d'une matrice carrée.

Notations $\mathrm{Sp}(f)$ et $\mathrm{Sp}(A)$.

Si $Q$ est un polynôme, obtention d'éléments propres de $Q(f)$ à partir d'éléments propres de $f$.

Si $f(x) = \lambda x$ alors $Q(f)(x) = Q(\lambda)x$.
Si $AX = \lambda X$ alors $Q(A)X = Q(\lambda)X$.

### b) Recherche d'éléments propres

Polynômes annulateurs d'un endomorphisme, d'une matrice carrée.

On pourra donner les exemples des homothéties, des projecteurs et des symétries.

Si $Q$ est un polynôme annulateur de $f$ (respectivement $A$) et $\lambda$ une valeur propre de $f$ (respectivement $A$), alors $\lambda$ est racine de $Q$.

Tout endomorphisme d'un espace de dimension finie admet au moins un polynôme annulateur non nul.
Toute matrice carrée admet au moins un polynôme annulateur non nul.

Aucune autre connaissance sur les polynômes annulateurs ne figure au programme.

### c) Propriétés générales

Un endomorphisme d'un espace de dimension finie admet un nombre fini de valeurs propres et ses sous-espaces propres sont en somme directe.

$$\sum_{\lambda \in \mathrm{Sp}(f)} \dim \ker(f - \lambda \mathrm{Id}_E) \leqslant \dim(E).$$

Une concaténation de familles libres de sous-espaces propres associés à des valeurs propres distinctes forme une famille libre de $E$.

En particulier, une famille de vecteurs propres associés à des valeurs propres distinctes est une famille libre.
Un endomorphisme d'un espace vectoriel de dimension $n$ a au plus $n$ valeurs propres.

### d) Réduction des endomorphismes et des matrices carrées

| | |
|---|---|
| $f$ est diagonalisable si et seulement s'il existe une base $\mathcal{B}$ de $E$ composée de vecteurs propres de $f$. | $\mathrm{Mat}_{\mathcal{B}}(f)$ est alors une matrice diagonale. |
| Caractérisation des endomorphismes diagonalisables à l'aide des dimensions des sous-espaces propres. | $f$ est diagonalisable si et seulement si $\displaystyle\sum_{\lambda\in \mathrm{Sp}(f)} \dim\ker(f-\lambda \mathrm{Id}_E)=\dim(E)$. |
| $f$ est diagonalisable si et seulement si $E$ est somme directe des sous-espaces propres de $f$. | Si $\dim(E)=n$, tout endomorphisme de $E$ admettant $n$ valeurs propres distinctes est diagonalisable et les sous-espaces propres sont tous de dimension 1. |
| Matrices diagonalisables, diagonalisation d'une matrice carrée. | Interprétation matricielle des résultats précédents.<br>$A$ est diagonalisable si et seulement s'il existe une matrice $P$ inversible telle que $P^{-1}AP$ est une matrice diagonale. Les colonnes de $P$ forment une base de $\mathcal{M}_{n,1}(\mathbf{R})$ constituée de vecteurs propres de $A$. |
| Application au calcul des puissances d'une matrice carrée. | |

### 3 - Algèbre bilinéaire

*L'objectif de ce chapitre est d'introduire les notions fondamentales de l'algèbre bilinéaire dans le cadre euclidien, utilisées en particulier lors de l'étude des fonctions de $n$ variables. L'étude des endomorphismes symétriques sera faite au quatrième semestre.*

*Les espaces vectoriels considérés dans ce chapitre sont des $\mathbf{R}$-espaces vectoriels. On identifiera $\mathbf{R}$ et $\mathcal{M}_1(\mathbf{R})$.*

#### a) Produit scalaire

| | |
|---|---|
| Produit scalaire, norme associée. | Un produit scalaire est une forme bilinéaire symétrique, définie positive.<br>Produit scalaire canonique sur $\mathbf{R}^n$ ; exemples de produits scalaires. |
| Inégalité de Cauchy-Schwarz. | Cas de l'égalité. |
| Vecteurs orthogonaux, sous-espaces orthogonaux. | |
| Familles orthogonales, familles orthonormales ou orthonormées. | On ne considèrera que des familles finies.<br>Toute famille orthogonale ne contenant pas le vecteur nul est libre. |
| Théorème de Pythagore. | |

#### b) Espaces euclidiens

*Dans ce paragraphe $x$, $y$ désignent des vecteurs d'un espace vectoriel et $X$, $Y$ sont les colonnes coordonnées correspondantes dans une base.*

| | |
|---|---|
| Espace euclidien. | Un espace euclidien est un espace vectoriel de dimension finie sur $\mathbf{R}$, muni d'un produit scalaire. |
| Existence de bases orthonormées. | On pourra introduire la méthode de l'orthonormalisation de Schmidt sur des exemples en petite dimension, mais cette méthode n'est pas exigible. |
| Coordonnées et norme d'un vecteur dans une base orthonormée. | $x = \sum_i \langle x, e_i \rangle e_i \ , \ \|x\|^2 = \sum_i \langle x, e_i \rangle^2.$ |
| Expression matricielle du produit scalaire et de la norme euclidienne en base orthonormée. | $\langle x, y \rangle = {}^tXY \ ; \ \|x\|^2 = {}^tXX.$ |
| Changement de bases orthonormées. | La matrice de passage d'une base orthonormée à une base orthonormée est orthogonale : $P^{-1} = {}^tP$. Aucune autre connaissance sur les matrices orthogonales n'est au programme. |
| Supplémentaire orthogonal d'un sous-espace vectoriel. Complétion d'une famille orthonormée en une base orthonormée. | Notation $F^\perp$. |

## II - **Fonctions réelles définies sur $\mathbf{R}^n$**

### 1 - Introduction aux fonctions définies sur $\mathbf{R}^n$

*Au troisième semestre, l'objectif est de confronter les étudiants à la notion de fonction réelle de $n$ variables, aux principales définitions tout en évitant les problèmes de nature topologique. C'est pourquoi le domaine de définition des fonctions sera systématiquement $\mathbf{R}^n$, muni de sa structure euclidienne canonique. L'étude de la continuité d'une fonction en un point pathologique est hors programme, ainsi que l'étude des recollements de formules lorsque $f$ est définie sur $\mathbf{R}^n$ par plusieurs formules.*

*Dès que possible, les notions introduites seront illustrées à l'aide de la libraire matplotlib.pyplot de Python.*

| | |
|---|---|
| Fonctions définies sur $\mathbf{R}^n$ à valeurs dans $\mathbf{R}$. | On donnera de nombreux exemples de fonctions de 2, 3 ou $n$ variables réelles. Les fonctions polynomiales de $n$ variables donnent des exemples simples de fonctions définies sur $\mathbf{R}^n$. |
| Équation du graphe d'une fonction définie sur $\mathbf{R}^n$. | Cas des fonctions affines de $n$ variables. |
| Lignes de niveau pour les fonctions de deux variables. | On se limitera à des exemples simples. |

| | |
|---|---|
| Continuité d'une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$. | Une fonction $f$, définie sur $\mathbf{R}^n$, est continue au point $x_0$ de $\mathbf{R}^n$ si : $\forall \varepsilon > 0$, $\exists \alpha > 0$, $\forall x \in \mathbf{R}^n$, $\|x - x_0\| \leqslant \alpha \Longrightarrow \|f(x) - f(x_0)\| \leqslant \varepsilon$.<br>$f$ est continue sur $\mathbf{R}^n$ si et seulement si $f$ est continue en tout point de $\mathbf{R}^n$.<br>Aucune difficulté ne sera soulevée sur ce sujet.<br>On mettra en avant l'analogie avec la notion de continuité des fonctions d'une variable vue en première année.<br>Les fonctions polynomiales de $n$ variables sont continues sur $\mathbf{R}^n$. Résultat admis. |
| Opérations sur les fonctions continues. | Somme, produit, quotient.<br>La composition d'une fonction continue sur $\mathbf{R}^n$ à valeurs dans un intervalle $I$ de $\mathbf{R}$ par une fonction continue de $I$ à valeurs dans $\mathbf{R}$ est continue.<br>Résultats admis. |

### 2 - Calcul différentiel

*L'introduction des notions différentielles concernant les fonctions numériques de plusieurs variables réelles se fait en se limitant aux fonctions définies sur $\mathbf{R}^n$. La détermination de la classe d'une fonction n'est pas au programme.*
*La recherche d'extremum est abordée ici, jusqu'à la condition nécessaire du premier ordre.*

*Les fonctions sont désormais supposées définies et continues sur $\mathbf{R}^n$.*

#### a) Dérivées partielles, gradient

| | |
|---|---|
| Fonctions partielles en un point. | |
| Dérivées partielles d'ordre 1. | Notation $\partial_i f$. |
| Gradient en un point $x$. | Notation $\nabla f(x)$.<br>$\nabla f(x)$ est l'élément de $\mathbf{R}^n$ égal à $(\partial_1 f(x), \ldots, \partial_n f(x))$. |
| Dérivées partielles d'ordre 2. | Notation $\partial^2_{i,j} f$. |
| Fonctions de classe $C^1$ et $C^2$ sur $\mathbf{R}^n$. | Les fonctions polynomiales de $n$ variables sont des fonctions de classe $C^2$ sur $\mathbf{R}^n$. Résultat admis. |
| Opérations sur les fonctions de classe $C^1$ et $C^2$. | Somme, produit, quotient.<br>La composition d'une fonction de classe $C^1$ [resp. $C^2$] sur $\mathbf{R}^n$ à valeurs dans un intervalle $I$ de $\mathbf{R}$ par une fonction de classe $C^1$ [resp. $C^2$] sur $I$ à valeurs dans $\mathbf{R}$ est de classe $C^1$ [resp. $C^2$].<br>Résultats admis. |
| Pour une fonction de classe $C^1$ : existence et unicité d'un développement limité d'ordre 1 en un point. | $f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \|h\|\varepsilon(h)$ où $\varepsilon(0) = 0$ et $\varepsilon$ continue en 0. Résultat admis. |
| Si $f$ est de classe $C^1$, dérivée de la fonction $g$ définie sur $\mathbf{R}$ par : $g(t) = f(x + th)$. | $g'(t) = \langle \nabla f(x + th), h \rangle$ et donc $g'(0) = \nabla f(x_0)$.<br>Interprétation géométrique du gradient. |

#### b) Recherche d'extremum : condition d'ordre 1

| | |
|---|---|
| Définition d'un extremum local, d'un extremum global. | |
| Condition nécessaire du premier ordre. Point critique. | Si une fonction $f$ de classe $C^1$ sur $\mathbf{R}^n$ admet un extremum local en un point $x$, alors $\nabla f(x) = 0$. Les points où le gradient s'annule sont appelés points critiques. |

## III - Compléments de probabilités ; couples et $n$-uplets de variables aléatoires réelles

*L'objectif est double :*

- *d'une part, consolider les acquis de première année concernant les variables aléatoires discrètes, et enrichir le champ des problèmes étudiés, avec, en particulier, l'étude simultanée de variables aléatoires (vecteurs aléatoires de $\mathbf{R}^n$) ;*
- *d'autre part, effectuer une étude élémentaire des lois continues usuelles discrètes ou à densité.*

*On fera des liens entre certaines lois dans le cadre des approximations et des convergences, ainsi que les liens entre statistique et probabilités dans le cadre de l'estimation.*

*La théorie des familles sommables n'est pas au programme. Aucune difficulté concernant la dénombrabilité ne sera soulevée (on pourra si besoin admettre que $\mathbf{N}^k$ est dénombrable.) On admettra le théorème suivant :*
*Soit $I$ un ensemble dénombrable infini, indexé par $\mathbf{N}$ sous la forme $I = \{\varphi(n); n \in \mathbf{N}\}$ où $\varphi$ est une bijection de $\mathbf{N}$ dans $I$. Si la série $\sum u_{\varphi(n)}$ converge absolument, alors sa somme est indépendante de l'indexation $\varphi$, et pourra également être notée $\sum\limits_{i \in I} u_i$. L'étude de cette convergence n'est pas un objectif du programme. On dira alors que la série est absolument convergente (ou converge absolument). Toutes les opérations (somme, produit, regroupement par paquets, etc.) sont alors licites. Aucune difficulté ne sera soulevée sur ces notions, qui ne sont pas exigibles des étudiants, et tout exercice ou problème y faisant référence devra impérativement les rappeler.*

### 1 - Compléments sur les variables aléatoires réelles

#### a) Généralités sur les variables aléatoires réelles

| | |
|---|---|
| On rappellera la signification de la notation $(\Omega, \mathcal{A}, P)$. | |
| Définition d'une variable aléatoire. | Une variable aléatoire sur $(\Omega, \mathcal{A}, P)$ est une application $X : \Omega \to \mathbf{R}$ telle que, pour tout $x$ dans $\mathbf{R}$, $[X \leq x]$ est un événement. Le fait de vérifier qu'une fonction est une variable aléatoire n'est pas un des objectifs du programme. |
| Fonction de répartition d'une variable aléatoire. | |

| | |
|---|---|
| Loi d'une variable aléatoire | C'est la donnée des probabilités $P(X \in I)$ où $I$ est intervalle.<br>La loi est caractérisée par la fonction de répartition. |
| Une combinaison linéaire, un produit de variables aléatoires sont des variables aléatoires.<br>Le maximum et le minimum de variables aléatoires sont des variables aléatoires. | Résultat admis. |

**b) Espérance et conditionnement pour les variables aléatoires discrètes**

| | |
|---|---|
| Espérance conditionnelle. | Si $A$ est un événement de probabilité non nulle, $E(X/A)$ est l'espérance de $X$, si elle existe, pour la probabilité conditionnelle $P_A$. |
| Formule de l'espérance totale. | Soit $X$ une variable aléatoire discrète, soit $(A_n)$ un système complet d'événements tels que, pour tout $n$ dans $\mathbf{N}$, $P(A_n) \neq 0$. Alors $X$ admet une espérance pour $P$ si et seulement si :<br>— pour tout $n \in \mathbf{N}$ l'espérance conditionnelle $E(X/A_n)$ existe ;<br>— la série $\sum_{n \in \mathbf{N}} E(\|X\|/A_n)P(A_n)$ converge.<br>Dans ce cas, $E(X) = \sum_{n \in \mathbf{N}} E(X/A_n)P(A_n)$. |

**2 - Introduction aux variables aléatoires à densité**

**a) Densités et fonction de répartition d'une variable aléatoire**

| | |
|---|---|
| Définition d'une densité de probabilité sur $\mathbf{R}$. | Une fonction $f : \mathbf{R} \to \mathbf{R}$ est une densité de probabilité lorsqu'elle est continue sauf en nombre fini de points, positive et vérifie :<br>$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$ |
| Définition d'une variable aléatoire à densité. | On dit qu'une variable aléatoire $X$ est à densité si sa fonction de répartition $F_X$ est continue sur $\mathbf{R}$ et de classe $C^1$ sur $\mathbf{R}$ éventuellement privé d'un ensemble fini de points.<br>Toute fonction égale à $F'_X$ sauf éventuellement en un nombre finie de point est une densité de probabilité et on dit que c'est une densité de $X$.<br>Pour tout $x$ de $\mathbf{R}$, $F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$. |
| Caractérisation de la loi d'une variable aléatoire à densité par la donnée d'une densité $f_X$. | |
| Toute densité de probabilité sur $\mathbf{R}$ est la densité d'une variable aléatoire. | Résultat admis. |

| | |
|---|---|
| Transformation affine d'une variable à densité. | Les étudiants devront savoir calculer la fonction de répartition et la densité de $aX + b$ ($a \neq 0$). |
| Exemples simples de calculs de fonctions de répartition et de densités de fonctions d'une variable aléatoire à densité. | Les étudiants devront savoir retrouver les lois de $X^2$ et $\varphi(X)$ avec $\varphi$ de classe $C^1$ strictement monotone sur $X(\Omega)$. |

**b) Espérance des variables aléatoires à densité**

| | |
|---|---|
| Espérance d'une variable aléatoire à densité. Variables aléatoires centrées. | Une variable aléatoire $X$ de densité $f_X$ admet une espérance $E(X)$ si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x f_X(x)\mathrm{d}x$ est absolument convergente ; dans ce cas, $E(X)$ est égale à cette intégrale. Exemple de variables aléatoires n'admettant pas d'espérance. |
| Linéarité et croissance de l'espérance pour les variables aléatoires à densité. | Résultat admis. |
| Existence d'espérance par domination. | Résultat admis. |
| Théorème de transfert. | Si $X$ est une variable aléatoire ayant une densité $f_X$ nulle en dehors de l'intervalle $]a, b[$, avec $-\infty \leqslant a < b \leqslant +\infty$, et si $g$ est une fonction continue sur $]a, b[$ éventuellement privé d'un nombre fini de points, $E(g(X))$ existe et est égale à $\displaystyle\int_a^b g(t) f_X(t)\mathrm{d}t$ si et seulement si cette intégrale converge absolument. On pourra admettre ou démontrer ce résultat dans le cas où $g$ est de classe $C^1$, avec $g'$ strictement positive (ou strictement négative) et le vérifier dans des cas simples. Cette démonstration n'est pas exigible. |

**3 - Lois de variables aléatoires à densité usuelles**

| | |
|---|---|
| Loi uniforme sur un intervalle. Espérance. | Notation $X \hookrightarrow \mathcal{U}[a, b]$. $X \hookrightarrow \mathcal{U}[0, 1] \Longleftrightarrow a + (b - a)X \hookrightarrow \mathcal{U}[a, b]$. |
| Loi exponentielle. Caractérisation par l'absence de mémoire. Espérance. | Notation $X \hookrightarrow \mathcal{E}(\lambda)$. $X \hookrightarrow \mathcal{E}(1) \Longleftrightarrow \dfrac{1}{\lambda}X \hookrightarrow \mathcal{E}(\lambda) \quad (\lambda > 0)$. |
| Loi normale centrée réduite, loi normale (ou de Laplace-Gauss). Espérance. | Notation $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$. |
| Transformées affines de variables aléatoires suivant des lois normales. | Si $X$ suit une loi normale, et si $a$ et $b$ sont deux réels, avec $a \neq 0$, alors la variable aléatoire $aX + b$ suit également une loi normale. Si $\sigma > 0$, $X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \Leftrightarrow \dfrac{X - \mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$. |

Propriété de la fonction de répartition de la loi normale centrée réduite.

Pour tout réel $x$ : $\Phi(-x) = 1 - \Phi(x)$.

Exemples d'utilisation de la table de la loi normale et interprétation graphique.
On attend des étudiants qu'ils sachent représenter graphiquement les fonctions densités des lois normales et utiliser la fonction de répartition $\Phi$ de la loi normale centrée réduite. ▶

Lois $\gamma$. Espérance d'une variable aléatoire suivant une loi $\gamma$.

$X$ suit une loi $\gamma(\nu)$, avec $\nu > 0$, si $X$ admet comme densité :

$$f_X(t) = \begin{cases} 0 & \text{si } t \leqslant 0 \\ \dfrac{1}{\Gamma(\nu)} t^{\nu-1} \mathrm{e}^{-t} & \text{si } t > 0, \end{cases}$$

avec $\Gamma(\nu) = \displaystyle\int_0^{+\infty} t^{\nu-1}\mathrm{e}^{-t}\mathrm{d}t$. Pour le calcul des moments de la loi $\gamma$, on pourra établir $\Gamma(\nu+1) = \nu\Gamma(\nu)$ et $\Gamma(n+1) = n!$ pour tout entier $n$ de $\mathbf{N}$.

**4 - Variance des variables aléatoires à densité**

Variance, écart-type. Variables aléatoires centrées, centrées réduites.
Variance d'une variable aléatoire suivant une loi usuelle (uniforme sur un intervalle, exponentielle, normale).

Illustrations avec les lois usuelles.
On pourra donner un exemple de variable aléatoire n'admettant pas de variance.

On admettra que l'existence de l'espérance et de la variance d'une variable aléatoire $X$ est équivalente à l'existence de $E(X^2)$.

**5 - $n$-uplets de variables aléatoires réelles ; généralisation des propriétés de l'espérance et de la variance**

*Dans cette partie, on étend la notion de loi de couple de variables aléatoires discrètes vue en première année à un vecteur aléatoire, puis, de manière intuitive, la notion d'espérance à une somme de variables aléatoires admettant chacune une espérance. La définition de l'espérance générale ou des moments d'une variable aléatoire dans un cadre quelconque n'étant pas au programme, toute difficulté s'y ramenant est à écarter. On admettra que les propriétés opératoires usuelles de l'espérance et la variance se généralisent aux variables aléatoires quelconques.*
*L'étude, pour $n > 2$, de $n$-uplets à composantes non indépendantes n'est pas un objectif du programme.*

**a) Généralisation**

Loi d'un vecteur aléatoire à valeurs dans $\mathbf{R}^n$. Loi marginale.

La loi d'un vecteur $(X_1, \ldots, X_n)$ de variables aléatoires réelles est donné par la fonction $F_{(X_1,\ldots,X_n)}$ définie sur $\mathbf{R}^n$ par :

$$F_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) = P\left(\bigcap_{i=1}^{n} [X_i \leqslant x_i]\right).$$

Aucune difficulté ne sera soulevée sur cette notion.

Caractérisation de la loi d'un vecteur aléatoire discret à valeurs dans $\mathbf{R}^n$.

Si deux vecteurs $(X_1, X_2, \ldots, X_n)$ et $(Y_1, Y_2, \ldots, Y_n)$ ont même loi et si $g$ est une fonction continue sur $\mathbf{R}^n$ à valeurs dans $\mathbf{R}$, alors les variables aléatoires réelles $g(X_1, X_2, \ldots, X_n)$ et $g(Y_1, Y_2, \ldots, Y_n)$ ont même loi.

Aucune difficulté ne sera soulevée.
Résultat admis.

**b) Indépendance**

Indépendance mutuelle de $n$ variables aléatoires réelles.

$X_1$, ..., $X_n$ sont mutuellement indépendantes si et seulement si :

$$F_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) = \prod_{k=1}^{n} F_{X_k}(x_k)$$

pour tous réels $x_1, \ldots, x_n$.

Caractérisation de l'indépendance mutuelle de $n$ variables aléatoires réelles.

$X_1$, ..., $X_n$ sont mutuellement indépendantes si et seulement si :

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

pour tous intervalles $I_1$, ..., $I_n$ de $\mathbf{R}$.
Résultat admis.

Caractérisation de l'indépendance mutuelle de $n$ variables aléatoires réelles discrètes.

$P\left(\bigcap_{i=1}^{n} [X_i = x_i]\right) = \prod_{i=1}^{n} P([X_i = x_i])$ pour tout $(x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega)$.
Résultat admis.

Lemme des coalitions.

Si $X_1$, $X_2$ ,..., $X_n$, sont indépendantes, toute variable aléatoire fonction de $X_1$, $X_2$, ..., $X_p$ est indépendante de toute variable aléatoire fonction de $X_{p+1}$, $X_{p+2}$,..., $X_n$.
Résultat admis.

**c) Le cas particulier du couple**

On généralisera les notions de linéarité, de croissance et d'existence par domination de l'espérance à des variables aléatoires quelconques.

Résultats admis

Espérance du produit de variables aléatoires indépendantes.

Si $X$ et $Y$ admettent une espérance et sont indépendantes, $XY$ admet une espérance et $E(XY) = E(X)E(Y)$.
Généralisation à $n$ variables aléatoires mutuellement indépendantes.
Résultats admis.

| | |
|---|---|
| Covariance de deux variables aléatoires admettant une variance. Propriétés.<br>Formule de Huygens. | Notation $\text{Cov}(X, Y)$.<br>Bilinéarité, symétrie, positivité de la covariance.<br>$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$. |
| Variance d'une somme. | $V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$. |
| Cœfficient de corrélation linéaire.<br>Propriétés. | Notation $\rho(X, Y)$.<br>$\lvert\rho(X, Y)\rvert \leqslant 1$. Interprétation dans le cas où $\rho(X, Y) = \pm 1$. |
| Si $X$ et $Y$ sont indépendantes et admettent un moment d'ordre 2, leur covariance est nulle. | La réciproque est fausse. |
| Variance d'une somme de variables aléatoires indépendantes. | Si $X$ et $Y$ sont indépendantes et admettent une variance, $X + Y$ admet une variance et $V(X + Y) = V(X) + V(Y)$.<br>Généralisation à $n$ variables aléatoires mutuellement indépendantes.<br>Résultats admis. |

**d) Sommes de variables aléatoires indépendantes**

| | |
|---|---|
| Densité de la somme $Z = X + Y$ de deux variables aléatoires à densité indépendantes, produit de convolution. | Si la fonction $h$ définie par la relation $h(x) = \displaystyle\int_{-\infty}^{+\infty} f_X(t) f_Y(x - t)\mathrm{d}t$ est définie et continue sauf peut-être en un nombre fini de points, c'est une densité de $Z$.<br>C'est le cas si $f_X$ (ou $f_Y$) est bornée. |
| Stabilité de la loi $\gamma$ pour la somme. | Si $X_1$ et $X_2$ sont deux variables aléatoires indépendantes suivant respectivement des lois $\gamma(\nu_1)$ et $\gamma(\nu_2)$, alors $X_1 + X_2 \hookrightarrow \gamma(\nu_1 + \nu_2)$. |
| Loi de la somme de $n$ variables aléatoires indépendantes de loi $\mathcal{E}(1)$. | Pour étudier la somme de $n$ variables aléatoires indépendantes de loi $\mathcal{E}(\lambda)$, on se ramènera après multiplication par $\lambda$ à une somme de $n$ variables aléatoires indépendantes de loi $\mathcal{E}(1)$. |
| Stabilité de la loi normale pour la somme. | |

# ENSEIGNEMENT DE MATHÉMATIQUES DU QUATRIÈME SEMESTRE

## I - Compléments d'algèbre bilinéaire

### 1 - Endomorphismes symétriques d'un espace euclidien, matrices symétriques

Endomorphismes symétriques.

Un endomorphisme $f$ d'un espace vectoriel euclidien $E$ est symétrique si et seulement si pour tout couple $(x, y)$ de vecteurs de $E$, on a :
$$\langle f(x)\ ,\ y\rangle \ =\ \langle x\ ,\ f(y)\rangle.$$

Un endomorphisme est symétrique si et seulement si sa matrice dans une base orthonormée est symétrique.

Si $f$ est un endomorphisme symétrique et si $F$ est un sous-espace vectoriel stable par $f$, alors $F^\perp$ est stable par $f$.

Les sous-espaces propres d'un endomorphisme symétrique $f$ d'un espace vectoriel de dimension finie sont deux à deux orthogonaux.

Si $(u_k)_{1\leqslant k\leqslant p}$ sont $p$ vecteurs propres d'un endomorphisme symétrique $f$ associés à des valeurs propres distinctes, alors la famille $(u_k)_{1\leqslant k\leqslant p}$ est une famille orthogonale.

### 2 - Projection orthogonale

Projection orthogonale sur un sous-espace vectoriel $F$.

Notation $p_F$.

Si $(u_1, \ldots, u_k)$ est une base orthonormée de $F$, alors :
$$p_F(x) = \sum_{i=1}^{k} \langle x, u_i\rangle u_i.$$

Si $p$ est un projecteur, alors $p$ est un projecteur orthogonal si et seulement si c'est un endomorphisme symétrique.

Caractérisation par minimisation de la norme.

$$v = p_F(x) \Longleftrightarrow \|x - v\| = \min_{u\in F} \|x - u\|.$$

Application au problème des moindres carrés et à la droite de régression : minimisation de $\|AX - B\|$ avec $A \in \mathcal{M}_{n,p}(\mathbf{R})$ de rang $p$, $B \in \mathcal{M}_{n,1}(\mathbf{R})$ et $X \in \mathcal{M}_{p,1}(\mathbf{R})$.
Résultats non exigibles.

### 3 - Réduction des endomorphismes et des matrices symétriques

Si $E$ est un espace vectoriel euclidien, tout endomorphisme symétrique de $E$ est diagonalisable et ses sous-espaces propres sont orthogonaux.

Résultat admis.
Si $f$ est un endomorphisme symétrique, il existe une base $\mathcal{B}$ de $E$ orthonormée composée de vecteurs propres de $f$.

| | |
|---|---|
| Toute matrice symétrique réelle est diagonalisable avec une matrice de changement de base orthogonale. | Si $A$ est symétrique réelle, il existe une matrice orthogonale $P$ et une matrice diagonale $D$ telles que $D = P^{-1}AP = {}^tPAP$. |

## II - **Fonctions réelles de $n$ variables définies sur un ouvert de $\mathbf{R}^n$ ; recherche d'extrema**

*L'objectif est de présenter la démarche de recherche d'extrema et d'en acquérir une maîtrise raisonnable à partir d'un minimum d'outils théoriques. L'espace $\mathbf{R}^n$ sera muni de la norme euclidienne usuelle.*
*La détermination de la nature topologique d'un ensemble n'est pas un objectif du programme ; elle devra toujours être précisée. Néanmoins, il est nécessaire de sensibiliser les étudiants aux notions d'ouverts et de fermés. Les étudiants ont été familiarisés avec les fonctions continues sur $\mathbf{R}^n$ au troisième semestre, aussi on s'appuiera, pour mener une initiation à la topologie de $\mathbf{R}^n$, sur les sous-ensembles de $\mathbf{R}^n$ définis par des inégalités du type $\{x \in \mathbf{R}^n / \varphi(x) < a\}$ ou $\{x \in \mathbf{R}^n / \varphi(x) \leqslant a\}$ où $\varphi$ est une fonction continue sur $\mathbf{R}^n$. On donnera également la définition d'un ensemble borné.*

*L'étude de fonctions de $n$ variables à valeurs dans $\mathbf{R}$ se limitera à des fonctions définies sur des sous-ensembles de $\mathbf{R}^n$ pouvant être définis simplement (réunion, intersection finies) à l'aide des ensembles fermés ou ouverts précédents.*

*Les résultats seront énoncés dans le cas de fonctions de $n$ variables. Pour les démonstrations, on pourra se limiter aux cas $n = 2$ ou $n = 3$.*
*Aucune des démonstrations de ce chapitre n'est exigible des étudiants.*

*Dans ce paragraphe, $h$ désigne un vecteur de $\mathbf{R}^n$ et $H$ la colonne coordonnée correspondante.*

### 1 - Fonction de $n$ variables définies sur une partie de $\mathbf{R}^n$

*Dans ce paragraphe, on étend à des fonctions définies sur un sous-ensemble de $\mathbf{R}^n$, les notions et définitions vues au troisième semestre pour des fonctions définies sur $\mathbf{R}^n$. Toute difficulté concernant la détermination de la classe d'une fonction est exclue.*

| | |
|---|---|
| Extension de la notion de continuité aux fonctions définies sur un sous-ensemble de $\mathbf{R}^n$. | Aucune difficulté théorique ne sera soulevée. |
| Extension de la notion de fonctions $C^1$ et $C^2$ aux fonctions définies sur un ouvert $\mathcal{O}$ de $\mathbf{R}^n$. | Aucune difficulté théorique ne sera soulevée.<br>Extension des notions, vues au troisième semestre, de dérivées partielles d'ordre 1 et 2, gradient, développement limité d'ordre 1, opérations sur les fonctions de classe $C^1$ ou $C^2$. |

### 2 - Compléments sur les fonctions de classe $C^2$ sur un ouvert de $\mathbf{R}^n$

| | |
|---|---|
| Matrice hessienne en un point $x$. | Notation $\nabla^2 f(x)$. |
| Théorème de Schwarz. | Si $f$ est de classe $C^2$ sur un ouvert $\mathcal{O}$, alors la matrice hessienne est symétrique en tout point de $\mathcal{O}$.<br>Résultat admis. |

Fonction quadratique définie sur $\mathbf{R}^n$ associée à une matrice symétrique réelle $A$.

$q(h) = {}^tHAH$.
On remarquera qu'il existe une base orthonormale $\mathcal{B}$ de $\mathbf{R}^n$ telle que si $h$ a pour coordonnées $h_1, \ldots, h_n$ dans $\mathcal{B}$ on a :

$$q(h) = \sum \lambda_i h_i^2,$$

où $\lambda_i, \ldots, \lambda_n$ sont les valeurs propres de $A$.

Existence et unicité d'un développement limité d'ordre 2 d'une fonction de classe $C^2$ sur un ouvert $\mathcal{O}$.

$$f(x+h) = f(x) + \langle \nabla f(x), h\rangle + \frac{1}{2}q_x(h) + \|h\|^2\varepsilon(h)$$
où $\varepsilon(0) = 0$, $\varepsilon$ continue en 0 et $q_x$ est la fonction quadratique associée à la matrice hessienne $\nabla^2 f(x)$.
Résultat admis.

Si $f$ est de classe $C^2$, dérivée seconde de la fonction $g$ définie au voisinage de 0 par :
$$g(t) = f(x+th).$$

$g''(t) = q_{x+th}(h)$ où $q_{x+th}$ est la fonction quadratique associée à la matrice hessienne $\nabla^2 f(x+th)$ et donc $g''(0) = q_x(h)$.

### 3 - Recherche d'extrema

*Dans un premier temps, on étendra rapidement les notions vues au troisième semestre à une fonction définie sur un sous-ensemble de $\mathbf{R}^n$.*

#### a) Définition

Définition d'un extremum local, d'un extremum global.

#### b) Extrema sur un ensemble fermé borné

Une fonction continue sur une partie fermée bornée admet un maximum global et un minimum global.

Résultat admis.

#### c) Condition d'ordre 1

Condition nécessaire du premier ordre.
Point critique.

Si une fonction de classe $C^1$ sur un ouvert $\mathcal{O}$ de $\mathbf{R}^n$ admet un extremum local en un point $x_0$ de $\mathcal{O}$, alors $\nabla f(x_0) = 0$.
Les points où le gradient s'annule sont appelés points critiques.

#### d) Condition d'ordre 2

Étude locale d'une fonction $f$ de classe $C^2$ sur un ouvert $\mathcal{O}$ en un point critique.

Si $x_0$ est un point critique de $f$ :
- si $\mathrm{Sp}(\nabla^2 f(x_0)) \subset \mathbf{R}_+^*$, alors $f$ admet un minimum local en $x_0$,
- si $\mathrm{Sp}(\nabla^2 f(x_0)) \subset \mathbf{R}_-^*$, alors $f$ admet un maximum local en $x_0$,
- si $\mathrm{Sp}(\nabla^2 f(x_0))$ contient deux réels non nuls de signes distincts, $f$ n'admet pas d'extremum en $x_0$.

On fera le lien avec le signe de la fonction quadratique $q_{x_0}$ associée à la hessienne de $f$ en $x_0$.

Point selle (ou col).

Une condition suffisante d'extremum global.

Si $\Omega$ est un ouvert convexe de $\mathbf{R}^n$ et si $x_0$ est un point critique de $f$ :
- si pour tout $x \in \Omega$, $\mathrm{Sp}(\nabla^2 f(x)) \subset \mathbf{R}^+$, alors $f$ admet un minimum global en $x_0$,
- si pour tout $x \in \Omega$, $\mathrm{Sp}(\nabla^2 f(x)) \subset \mathbf{R}_-$, alors $f$ admet un maximum global en $x_0$,

On introduira la notion d'ouvert convexe sans soulever aucune difficulté théorique et la vérification de cette propriété n'est pas un objectif du programme.
On admet ce résultat.

### e) Recherche d'extrema sous contrainte d'égalités linéaires

*Dans tout ce paragraphe, $\mathcal{C}$ désigne l'ensemble des solutions d'un système linéaire* $\begin{cases} g_1(x) & = b_1 \\ \vdots & \vdots \\ g_p(x) & = b_p \end{cases}$ *et $\mathcal{H}$ l'ensemble des solutions du système homogène associé.*

Condition nécessaire du premier ordre sous la contrainte $\mathcal{C}$.

Si $f$ est une fonction de classe $C^1$ sur un ouvert $\mathcal{O}$, et si la restriction de $f$ à $\mathcal{C}$ admet un extremum local en un point $x_0$, alors $\nabla f(x_0)$ est dans $\mathrm{Vect}(\nabla g_1(x_0), \ldots, \nabla g_p(x_0))$.
On remarquera que :
- $\mathcal{H}^\perp = \mathrm{Vect}(\nabla g_1(x_0), \ldots, \nabla g_p(x_0))$.
- Il existe $\lambda_1, \ldots, \lambda_p$ réels tels que :
$$\nabla f(x_0) = \sum_{i=1}^{p} \lambda_i \nabla g_i(x_0).$$

Point critique pour l'optimisation sous contrainte.
Exemples de recherche d'extrema globaux sous contrainte d'égalités linéaires dans des cas simples.

## III - Probabilités : convergences, estimation

## 1 - Convergences et approximations

### a) Convergence en probabilité

On pourra rappeler l'inégalité de Markov et l'inégalité de Bienaymé-Tchebychev vues en première année.

Convergence en probabilité.

La suite $(X_n)_{n\in\mathbf{N}^*}$ converge en probabilité vers $X$ si :

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} P([|X_n - X| \geqslant \varepsilon]) = 0.$$

Notation $X_n \xrightarrow{P} X$.

On pourra énoncer la loi faible des grands nombres en terme de convergence en probabilité.

Composition par une fonction continue.

Si $X_n \xrightarrow{P} X$ et si $f$ est une fonction continue sur $\mathbf{R}$ à valeurs réelles, alors $f(X_n) \xrightarrow{P} f(X)$. Résultat admis.

Convergence en probabilité et somme.

Si $X_n \xrightarrow{P} X$ et $Y_n \xrightarrow{P} Y$ alors $X_n+Y_n \xrightarrow{P} X+Y$.

### b) Convergence en loi

Définition de la convergence en loi d'une suite $(X_n)_{n\in\mathbf{N}^*}$ de variables aléatoires vers $X$.

Notation $X_n \xrightarrow{\mathcal{L}} X$.

La suite $(X_n)_{n\in\mathbf{N}^*}$ converge en loi vers $X$ si et seulement si en tout point de continuité $x$ de $F_X$ :

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x).$$

On illustrera cette définition à l'aide des approximations vues en première année.

Cas où les $X_n$ et $X$ prennent leurs valeurs dans $\mathbf{N}$.

La suite $(X_n)_{n\in\mathbf{N}^*}$ converge en loi vers $X$ si et seulement si :

$$\forall k \in \mathbf{N}, \quad \lim_{n\to+\infty} P([X_n = k]) = P([X = k]).$$

Composition par une fonction continue.

Si $(X_n)_{n\in\mathbf{N}^*}$ converge en loi vers $X$ et si $f$ est une fonction continue sur $\mathbf{R}$ à valeurs réelles, alors $(f(X_n))_{n\in\mathbf{N}^*}$ converge en loi vers $f(X)$. Résultat admis.

| | |
|---|---|
| Théorème limite central. | Si $(X_n)_{n\in\mathbf{N}^*}$ est une suite de variables aléatoires indépendantes et de même loi, admettant une espérance $m$ et une variance $\sigma^2$ non nulle, si on note : $\overline{X}_n = \dfrac{X_1 + \ldots + X_n}{n}$, alors la suite de variables aléatoires centrées réduites $\overline{X}_n^* = \sqrt{n}\left(\dfrac{\overline{X}_n - m}{\sigma}\right)$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.<br>D'où, on a pour tout $(a, b)$ tel que $-\infty \leqslant a \leqslant b \leqslant +\infty$ :<br>$\lim\limits_{n\to+\infty} P([a \leqslant \overline{X}_n^* \leqslant b]) = \displaystyle\int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t.$<br>Résultats admis. |
| Exemples d'approximations de la loi binomiale et de la loi de Poisson par la loi normale. | Toutes les indications devront être fournies aux candidats quant à la justification de l'utilisation des approximations. |

**2 - Estimation**

*L'objectif de cette partie est d'introduire le vocabulaire et la démarche de la statistique inférentielle en abordant, sur quelques cas simples, le problème de l'estimation, ponctuelle ou par intervalle de confiance. On se restreindra à une famille de lois de probabilités indexées par un paramètre scalaire (ou vectoriel) dont la valeur (scalaire ou vectorielle) caractérise la loi. On cherche alors à estimer la valeur du paramètre (ou une fonction simple de ce paramètre) à partir des données disponibles.*

*Dans ce contexte, on considère un phénomène aléatoire et on s'intéresse à une variable aléatoire réelle $X$ qui lui est liée, dont on suppose que la loi de probabilité n'est pas complètement spécifiée et appartient à une famille de lois dépendant d'un paramètre $\theta$ décrivant un sous-ensemble $\Theta$ de $\mathbf{R}$ (éventuellement de $\mathbf{R}^2$). Le paramètre $\theta$ est une quantité inconnue, fixée dans toute l'étude, que l'on cherche à déterminer ou pour laquelle on cherche une information partielle.*

*Le problème de l'estimation consiste alors à estimer la vraie valeur du paramètre $\theta$ ou de $g(\theta)$ (fonction à valeurs réelles du paramètre $\theta$), à partir d'un échantillon de données $x_1, \ldots, x_n$ obtenues en observant $n$ fois le phénomène. Cette fonction du paramètre représentera en général une valeur caractéristique de la loi inconnue comme son espérance, sa variance, son étendue...*

*On supposera que cet échantillon est la réalisation de $n$ variables aléatoires $X_1, \ldots, X_n$ définies sur un même espace probabilisable $(\Omega, \mathcal{A})$ muni d'une famille de probabilités $(P_\theta)_{\theta\in\Theta}$. Les $X_1, \ldots, X_n$ seront supposées $P_\theta$-indépendantes et de même loi que $X$ pour tout $\theta$.*

*On appellera estimateur de $g(\theta)$ toute variable aléatoire réelle de la forme $\varphi(X_1, X_2, \ldots, X_n)$ où $\varphi$ est une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$, éventuellement dépendante de $n$, et indépendante de $\theta$, dont la réalisation après expérience est envisagée comme estimation de $g(\theta)$.*

*Si $T_n$ est un estimateur, on notera, lorsque ces valeurs existent, $E_\theta(T_n)$ l'espérance de $T_n$ et $V_\theta(T_n)$ la variance de $T_n$, pour la probabilité $P_\theta$.*

**a) Estimation ponctuelle**

*Estimer ponctuellement $g(\theta)$ par $\varphi(x_1, \ldots, x_n)$ où $\varphi(X_1, X_2, \ldots, X_n)$ est un estimateur de $g(\theta)$ et*

*$(x_1, \ldots, x_n)$ est une réalisation de l'échantillon $(X_1, \ldots, X_n)$, c'est décider d'accorder à $g(\theta)$ la valeur $\varphi(x_1, \ldots, x_n)$.*

| | |
|---|---|
| $n$-échantillon $(X_1, \ldots, X_n)$ de variables aléatoires réelles indépendantes et de même loi que $X$. | Exemples de $n$-échantillons associés à une loi de Bernoulli $\mathcal{B}(p)$ avec $\theta = p$. |
| Définition d'un estimateur. | Un estimateur de $g(\theta)$ est une variable aléatoire de la forme $T_n = \varphi(X_1, \ldots, X_n)$. La réalisation $\varphi(x_1, \ldots, x_n)$ de l'estimateur $T_n$ est l'estimation de $g(\theta)$. Cette estimation ne dépend que de l'échantillon $(x_1, x_2, \ldots, x_n)$ observé. |
| Exemples simples d'estimateurs. | Exemple de la moyenne empirique $\dfrac{X_1 + X_2 + \cdots + X_n}{n}$. |
| Exemples simples d'estimations. | |

**b) Intervalle de confiance**

*S'il existe des critères pour juger des qualités d'un estimateur ponctuel $T_n$ de $g(\theta)$, aucune certitude ne peut jamais être apportée quant au fait que l'estimation donne la vraie valeur à estimer.*
*La démarche de l'estimation par intervalle de confiance consiste à trouver un intervalle aléatoire qui contienne $g(\theta)$ avec une probabilité minimale donnée. Dans tout ce paragraphe, $(U_n)_{n\geqslant 1}$ et $(V_n)_{n\geqslant 1}$ désigneront deux suites d'estimateurs de $g(\theta)$ telles que pour tout $\theta \in \Theta$ et pour tout $n \geqslant 1$, $P_\theta([U_n \leqslant V_n]) = 1$.*

| | |
|---|---|
| Intervalle de confiance. | Soit $\alpha \in [0,1]$. $[U_n, V_n]$ est un intervalle de confiance de $g(\theta)$ au niveau de confiance $1-\alpha$ si pour tout $\theta$ de $\Theta$, $$P_\theta([U_n \leqslant g(\theta) \leqslant V_n]) \geqslant 1-\alpha.$$ L'utilisation dans certains cas du théorème limite central impose d'introduire la notion d'intervalle de confiance asymptotique. Sa réalisation est l'estimation de cet intervalle de confiance. Les variables aléatoires $X_n$ seront supposées $P_\theta$-indépendantes et de même loi que $X$ pour tout $\theta$. On éclairera ces notions à l'aide de simulations informatiques. |
| Estimation par intervalle de confiance du paramètre d'une loi de Bernoulli. | En utilisant l'inégalité de Bienaymé-Tchebychev. |
| Estimation par intervalle de confiance de la moyenne d'une loi normale dont l'écart type est connu. | On pourra utiliser cet exemple pour introduire la variance empirique. |

**c) Estimation par intervalle de confiance asymptotique**

| | |
|---|---|
| Intervalle de confiance asymptotique. | On appelle intervalle de confiance asymptotique de $g(\theta)$ au niveau de confiance $1-\alpha$ une suite $([U_n, V_n])_{n\geqslant 1}$ vérifiant : pour tout $\theta$ de $\Theta$, il existe une suite de réels $(\alpha_n)$ à valeurs dans $[0,1]$, de limite $\alpha$, telle que pour tout $n \geqslant 1$, $P_\theta([U_n \leqslant g(\theta) \leqslant V_n]) \geqslant 1-\alpha_n$.<br>Par abus de langage on dit aussi que $[U_n, V_n]$ est un intervalle de confiance asymptotique. |
| Intervalles de confiance asymptotiques obtenus avec le théorème central limite. | Exemple du paramètre d'un loi de Bernoulli.<br>On pourra comparer, en majorant $p(1-p)$ par $\frac{1}{4}$, les intervalles de confiance obtenus par l'inégalité de Bienaymé-Tchebychev, et les intervalles de confiance asymptotiques obtenus par l'approximation normale de la loi binomiale. |

**d) Comparaison des estimateurs**

*La notion de risque quadratique n'est pas au programme.*

| | |
|---|---|
| Estimateur sans biais. | L'estimateur $T_n$ de $g(\theta)$ est sans biais si pour tout $\theta$ de $\Theta$, $E_\theta(T_n) = g(\theta)$. |
| Suite $(T_n)_{n\geqslant 1}$ d'estimateurs. | Chaque $T_n$ est de la forme $\varphi(X_1, X_2, \ldots, X_n)$. |
| Estimateur convergent. | Une suite d'estimateurs $(T_n)_{n\geqslant 1}$ de $g(\theta)$ est convergente si pour tout $\theta$, la suite $(T_n)_{n\geqslant 1}$ converge en probabilité vers $g(\theta)$.<br>Par abus de langage, on dit aussi que l'estimateur est convergent.<br>On rappellera que si $(T_n)_{n\geqslant 1}$ est une suite convergente d'estimateurs de $g(\theta)$ et si $f$ est une fonction continue sur $\mathbf{R}$ à valeurs réelles, alors $(f(T_n))_{n\geqslant 1}$ est une suite convergente d'estimateurs de $f(g(\theta))$. |
| Condition suffisante de convergence. | Une suite d'estimateurs $(T_n)_{n\geqslant 1}$ de $g(\theta)$ telle que $\lim\limits_{n\to+\infty} E(T_n) = g(\theta)$ et $\lim\limits_{n\to+\infty} V(T_n) = 0$ est convergente.<br>Cette convergence pourra être étudiée à l'aide de l'inégalité de Markov.<br>La démonstration de ce théorème donne naturellement un intervalle de confiance asymptotique de $g(\theta)$ ainsi qu'un moyen de comparer la qualité des estimateurs.<br>On illustrera en informatique ces notions $\boxed{\blacktriangleright}$. |

# TRAVAUX PRATIQUES DE MATHÉMATIQUES AVEC PYTHON

*En première année, les élèves ont acquis les bases de manipulation du logiciel Python. L'objectif de l'enseignement d'informatique de seconde année est de permettre aux étudiants d'utiliser Python de manière judicieuse et autonome pour illustrer ou modéliser des situations concrètes en mobilisant leurs connaissances mathématiques.*

*Les séances de travaux pratiques doivent se faire le plus souvent possible sur ordinateur. Les étudiants, au cours de leurs études ultérieures puis de leur parcours professionnel, seront amenés à utiliser des outils informatiques divers choisis pour leurs fonctionnalités, etdés que seule une pratique régulière de ces outils informatiques peut leur permettre d'en acquérir la maîtrise. De plus, en adoptant cette démarche exploratoire permise par le dialogue interactif avec la machine, cette pratique peut s'avérer bénéfique pour les apprentissages et faciliter la compréhension de concepts plus abstraits.*

*Le programme d'informatique s'articule autour de quatre thèmes : statistiques descriptives bivariées, fonctions de plusieurs variables, simulation de lois, estimation ponctuelle ou par intervalle de confiance.*

*L'ordre dans lequel les thèmes sont abordés est libre, mais il est préférable de mener ces activités en cohérence avec la progression du cours de mathématiques.*

*Dans certains thèmes, il s'avérera nécessaire d'introduire de nouvelles notions ou approches mathématiques. Celles-ci devront être explicitées en préambule des séances d'informatique et ne pourront en aucun cas être exigibles des étudiants. Certaines seront propres à un thème particulier, d'autres (comme par exemple les méthodes de Monte-Carlo) pourront au contraire être envisagées de manière transversale. Toutes les précisions nécessaires devront toujours être données lors de leur utilisation.*

*Toute la richesse du langage Python ne peut pas être entièrement maîtrisée par un étudiant, aussi seules certaines fonctions et commandes sont exigibles. Néanmoins, se contenter de ces seules commandes, en ignorant les nombreuses possibilités et commodités du langage, se révélerait rapidement contraignant et limitatif. De nouvelles commandes Python peuvent donc être introduites, mais cela devra se faire avec parcimonie, l'objectif principal de l'activité informatique reste la mise en pratique des connaissances mathématiques. Dans les sujets, les commandes introduites devront être présentées en préambule et toutes les précisions nécessaires seront données lors de leur utilisation et leur interprétation. On favorisera à cette occasion l'autonomie et la prise d'initiatives des étudiants grâce à l'utilisation de l'aide de Python, et à l'usage d'opérations de « copier-coller »qui permettent de prendre en main rapidement des fonctions nouvelles et évitent d'avoir à connaître par cœur la syntaxe de commandes complexes.*

*L'objectif de ces travaux pratiques n'est pas l'écriture de longs programmes mais l'assimilation de savoir-faire et de compétences spécifiés dans la liste des exigibles et rappelés en préambule de chaque thème.*

*Les exemples traités dans un thème devront être tirés, autant que possible, de situations réelles (traitement de données économiques, sociologiques, historiques, démographiques, en lien avec le monde de l'entreprise ou de la finance, etc), en faisant dès que possible un rapprochement avec les autres disciplines.*

## I - **Liste des exigibles**

### 1 - Commandes

*Les commandes exigibles ont été listées dans le programme de première année. On rappellera dans les sujets toutes les syntaxes des commandes non exigibles.*

### 2 - Savoir-faire et compétences

**C1** : Produire et interpréter des résumés numériques et graphiques d'une série statistique (simple, double) ou d'une loi.

**C2** : Modéliser et simuler des phénomènes (aléatoires ou déterministes) et les traduire en langage mathématique.

**C3** : Représenter et exploiter le graphe d'une fonction d'une, deux variables.

**C4** : Représenter et interpréter différents types de convergences.

**C5** : Utiliser la méthode de Monte-Carlo sur des exemples pertinents (calcul approché d'intégrales, de probabilités).

**C6** : Porter un regard critique sur les méthodes d'estimation et de simulation.

## II - **Liste des thèmes**

### 1 - Statistiques descriptives bivariées

(Durée indicative : 3 heures. Compétences développées : **C1** et **C6**)

| | |
|---|---|
| Série statistique à deux variables, nuage de points associé.<br>Point moyen $(\bar{x}, \bar{y})$ du nuage.<br>Covariance et cœfficient de corrélation empiriques, droites de régression. | On tracera le nuage de points et les droites de régression et on pourra effectuer des prétransformations pour se ramener au cas linéaire.<br>On différenciera les variables explicatives des variables à expliquer. |

### 2 - Fonctions de plusieurs variables

(Durée indicative : 3 heures. Compétences développées : **C2** et **C3**)

| | |
|---|---|
| Graphe d'une fonction de deux variables, lignes de niveau, plan affine tangent au graphe. Dérivées partielles et dérivées directionnelles, représentation du gradient.<br>Position du graphe par rapport au plan affine tangent au graphe, lien avec les valeurs propres de la matrice hessienne, points selles. Étude d'extrema locaux et globaux. Extrema sous contrainte linéaire. | À cette occasion, on pourra mettre en évidence l'orthogonalité du gradient avec les tangentes aux lignes de niveau du graphe d'une fonction de deux variables.<br>Programmation de fonctions variées permettant de mettre en évidence les notions d'extrema locaux ou globaux, avec ou sans contrainte. On pourra prendre des exemples issus de l'économie ou de la finance. |

### 3 - Simulation de lois

(Durée indicative : 6 heures. Compétences développées : **C1**, **C2**, **C3** et **C6**)

*Dans toutes les simulations effectuées, on pourra comparer les échantillons obtenus avec les distributions théoriques, en utilisant des diagrammes en bâtons et des histogrammes. On pourra aussi tracer la fonction de répartition empirique et la comparer à la fonction de répartition théorique.*

| | |
|---|---|
| Méthode d'inversion. | Application de la méthode d'inversion pour la simulation par exemple des lois exponentielles ou de Cauchy.<br>On pourra mettre en évidence, grâce aux simulations, qu'une variable aléatoire suivant une loi de Cauchy n'admet pas d'espérance. |

| | |
|---|---|
| Méthodes de simulation d'une loi géométrique. | Utilisation d'une loi de Bernoulli et d'une boucle `while`, utilisation d'une loi exponentielle et de la fonction `floor`, utilisation de la librairie `numpy.random`. |
| Simulations informatiques d'une loi normale par utilisation du théorème limite central appliqué à différentes lois. | Comparaison entre différentes méthodes de simulation d'une loi normale.<br>Utilisation de la librairie `numpy.random`.<br>On pourra s'intéresser au cas particulier de 12 variables aléatoires indépendantes suivant une même loi uniforme. |

**4 - Estimation ponctuelle et par intervalle de confiance**

(Durée indicative : 6 heures. Compétences développées : **C2**, **C4**, **C5** et **C6**)

| | |
|---|---|
| Méthode de Monte-Carlo : principe, garanties d'approximation. | Cette méthode permet d'estimer des quantités qu'il est difficile de calculer explicitement mais qu'il est facile d'approcher par simulation (probabilités d'événements, espérances de variables aléatoires).<br>Ainsi, on pourra estimer par exemple les valeurs prises par la fonction de répartition de la somme ou du produit de deux variables aléatoires.<br>On pourra justifier par simulation la validité de l'approche par intervalle de confiance asymptotique à partir d'un certain rang. |
| Comparaison de différents estimateurs ponctuels d'un paramètre. | On pourra utiliser des données issues de situations réelles ou créer plusieurs jeux de données par simulation. Dans ce dernier cas, on pourra comparer les lois des estimateurs par exemple à l'aide d'histogrammes. |
| Comparaison des intervalles de confiance d'un paramètre obtenus par différentes méthodes. | Estimation par intervalle de confiance du paramètre d'une loi de Bernoulli et de l'espérance d'une loi normale.<br>La comparaison pourra se faire en calculant les demi-largeurs moyennes des intervalles et leurs niveaux de confiance. |