

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 voie scientifique 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, 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 (études de chaînes de Markov) 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, abordées dès la première année, 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.
- Les travaux pratiques de mathématiques avec Scilab sont organisés autour de six thèmes faisant intervenir divers points du programme de mathématiques. L'objectif est d'apprendre aux étudiants à utiliser Scilab 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.

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 logiciel Scilab comporte de nombreuses fonctionnalités permettant d'illustrer simplement certaines notions mathématiques. Ainsi, on utilisera dès que possible l'outil informatique en cours de mathématiques pour visualiser et illustrer les notions étudiées.

ENSEIGNEMENT DE MATHÉMATIQUES DU TROISIÈME SEMESTRE

I - Algèbre linéaire et bilinéaire

Dans tout ce chapitre \mathbf{K} désignera \mathbf{R} ou \mathbf{C} .

1 - Compléments d'algèbre linéaire

a) Changement de base

Matrice d'un endomorphisme dans une base.

Rappels.

Matrice de passage de \mathcal{B} vers \mathcal{B}' .

Notation $P_{\mathcal{B},\mathcal{B}'}$.

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

Formules de changement de base.

$$X_{\mathcal{B}} = P_{\mathcal{B},\mathcal{B}'} X_{\mathcal{B}'}$$

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

Matrices semblables.

Deux matrices A et B carrées sont semblables s'il existe une matrice inversible P telle que $B = P^{-1}AP$.

A et B peuvent être interprétées comme les matrices d'un même endomorphisme dans des bases différentes.

Définition d'un sous-espace stable par un endomorphisme.

Seule la définition est exigible des étudiants.

b) 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.

Notation $\text{Tr}(A)$.

Linéarité de la trace.

Invariance de la trace par changement de base.

$$\text{Tr}(A) = \text{Tr}(P^{-1}AP)$$

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

Les espaces vectoriels considérés dans ce chapitre sont définis sur \mathbf{K} . 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.

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

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

Valeurs propres des matrices triangulaires.

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

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.

Si Q est un polynôme annulateur de f (respectivement A) et λ une valeur propre de f (respectivement A), alors λ 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.

Exemples des homothéties, des projecteurs et des symétries.

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.

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

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

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 .

Caractérisation des endomorphismes diagonalisables à l'aide des dimensions des sous-espaces propres.

$\text{Mat}_{\mathcal{B}}(f)$ est alors une matrice diagonale.

f est diagonalisable si et seulement si

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

f est diagonalisable si et seulement si E est somme directe des sous-espaces propres de f .

Matrices diagonalisables, diagonalisation d'une matrice carrée.

Application au calcul des puissances d'un endomorphisme ou 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.

Inégalité de Cauchy-Schwarz.

Vecteurs orthogonaux, sous-espaces orthogonaux.

Familles orthogonales, familles orthonormales ou orthonormées.

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.

Existence de bases orthonormées.

Coordonnées et norme d'un vecteur dans une base orthonormée.

Expression matricielle du produit scalaire et de la norme euclidienne en base orthonormée.

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.

Interprétation matricielle des résultats précédents.

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 .

Un produit scalaire est une forme bilinéaire symétrique, définie positive.

Produit scalaire canonique sur \mathbf{R}^n ; exemples de produits scalaires.

Cas de l'égalité.

On ne considèrera que des familles finies.

Toute famille orthogonale ne contenant pas le vecteur nul est libre.

Un espace euclidien est un espace vectoriel de dimension finie sur \mathbf{R} , muni d'un produit scalaire.

On pourra introduire la méthode de l'orthonormalisation de Schmidt.

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

$$\langle x, y \rangle = {}^tXY; \|x\|^2 = {}^tXX.$$

Changement de bases orthonormées.

La matrice de passage est orthogonale : $P^{-1} = {}^tP$.

Aucune autre connaissance sur les matrices orthogonales n'est au programme.

Notation F^\perp .

Supplémentaire orthogonal d'un sous-espace vectoriel.

Complétion d'une famille orthonormée en une base orthonormée.

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 du logiciel Scilab.

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 .

Cas des fonctions affines de n variables.

Équation du graphe d'une fonction définie sur \mathbf{R}^n .

On se limitera à des exemples simples.

Lignes de niveau pour les fonctions de deux variables.

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 : $\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathbf{R}^n$,

$$\|x - x_0\| \leq \alpha \implies |f(x) - f(x_0)| \leq \varepsilon.$$

f est continue sur \mathbf{R}^n si et seulement si f est continue en tout point de \mathbf{R}^n .

Aucune difficulté ne sera soulevée sur ce sujet.

On mettra en avant l'analogie avec la notion de continuité des fonctions d'une variable vue en première année.

Les fonctions polynomiales de n variables sont continues sur \mathbf{R}^n . Résultat admis.

Somme, produit, quotient.

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. Résultats admis.

Opérations sur les fonctions continues.

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.

Gradient en un point x .

Fonctions de classe C^1 sur \mathbf{R}^n .

Opérations sur les fonctions de classe C^1 .

Pour une fonction de classe C^1 : existence et unicité d'un développement limité d'ordre 1 en un point.

Notation $\partial_i(f)$.

Notation $\nabla(f)(x)$.

$\nabla(f)(x)$ est l'élément de \mathbf{R}^n égal à $(\partial_1(f)(x), \dots, \partial_n(f)(x))$.

Les fonctions polynomiales de n variables sont des fonctions de classe C^1 sur \mathbf{R}^n . Résultat admis.

Somme, produit, quotient.

La composition d'une fonction de classe C^1 sur \mathbf{R}^n à valeurs dans un intervalle I de \mathbf{R} par une fonction de classe C^1 sur I à valeurs dans \mathbf{R} est de classe C^1 .

Résultats admis.

$f(x+h) = f(x) + \langle \nabla(f)(x), h \rangle + \|h\|\varepsilon(h)$ où $\varepsilon(0) = 0$ et ε continue en 0. Résultat admis.

b) Dérivée directionnelle

Droite \mathcal{D} passant par x , de vecteur directeur u .

Si f est de classe C^1 , dérivée de la fonction g définie sur \mathbf{R} par :

$$g(t) = f(x + th).$$

Dérivée directionnelle de f au point x dans la direction h .

Paramétrisation : $t \mapsto x + tu$, $t \in \mathbf{R}$.

$$g'(t) = \langle \nabla(f)(x + th), h \rangle.$$

$$g'(0) = \langle \nabla(f)(x), h \rangle.$$

On en déduira une interprétation géométrique du gradient dans le cas où h est un vecteur de norme 1.

c) 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. Toutes les dérivées directionnelles en ces points sont nulles.

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.

Pour l'étude du cas discret, on pourra utiliser les notions et les énoncés classiques suivants sur les familles sommables absolument convergentes. Tout cours théorique sur les familles sommables est fortement déconseillé et on se limitera à une approche heuristique.

On admet que les manipulations ensemblistes classiques (produits finis, réunions dénombrables) d'ensembles dénombrables fournissent encore des ensembles dénombrables. On remarquera en particulier que l'ensemble $\mathbf{N} \times \mathbf{N}$ est dénombrable. 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.

Soit I un ensemble dénombrable infini, indexé par \mathbf{N} sous la forme $I = \{\varphi(n); n \in \mathbf{N}\}$ où φ est une bijection de \mathbf{N} dans I . Si la série $\sum u_{\varphi(n)}$ converge absolument, alors sa somme est indépendante de l'indexation φ , et pourra également être notée $\sum_{i \in I} u_i$. 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. Ainsi :

- Si $I = \bigsqcup_{j \in J} I_j$ (union disjointe) avec J un ensemble dénombrable et I_j des ensembles dénombrables pour tout j , alors : $\sum_{i \in I} u_i = \sum_{j \in J} \sum_{k \in I_j} u_k$.

- Si I et J sont des ensembles dénombrables, alors : $\left(\sum_{i \in I} u_i \right) \times \left(\sum_{j \in J} v_j \right) = \sum_{(i,j) \in I \times J} u_i v_j$.

On admettra que les théorèmes ou les techniques classiques concernant les séries s'étendent dans ce cadre.

1 - Compléments sur les variables aléatoires réelles

a) Généralités sur les variables aléatoires réelles

σ -algèbre \mathcal{B} des boréliens.

Aucun développement théorique sur la tribu des boréliens n'est au programme.

On admettra que pour tout borélien B et pour toute variable aléatoire réelle X définie sur (Ω, \mathcal{A}) , $[X \in B]$ appartient à \mathcal{A} .

σ -algèbre associée à une variable aléatoire X .

Notation \mathcal{A}_X . C'est la plus petite tribu contenant les événements $[X \leq x]$ pour tout réel x . Elle représente l'information fournie par X .

Une somme, un produit de variables aléatoires sont des variables aléatoires.

Résultat admis.

b) Espérance et conditionnement pour les variables aléatoires discrètes

Existence d'une espérance par domination.

Si X et Y sont deux variables aléatoires discrètes vérifiant $0 \leq |X| \leq Y$ presque sûrement, et si Y admet une espérance, alors X admet également une espérance. Dans ce cas, $|E(X)| \leq E(Y)$. Résultat admis.

Croissance de l'espérance pour les variables aléatoires discrètes.

Résultat admis.

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 définie sur (Ω, \mathcal{A}, P) , soit (A_n) un système complet d'événements et J l'ensemble des entiers n tels que $P(A_n) \neq 0$. Alors X admet une espérance pour P si et seulement si la série :

$$\sum_{(x,n) \in X(\Omega) \times J} x P_{A_n}([X = x]) P(A_n)$$

converge absolument. Dans ce cas, pour tout n dans J , l'espérance $E(X|A_n)$ est définie et

$$E(X) = \sum_{n \in J} E(X|A_n) P(A_n).$$

c) Compléments d'analyse

Reste d'une intégrale convergente.

L'intégration par parties sera pratiquée pour des intégrales sur un segment, on effectuera ensuite un passage à la limite.

Pratique de l'intégration par parties pour les intégrales sur un intervalle quelconque.

Si f est continue sur $]a, b[$, si φ est une bijection de $] \alpha, \beta [$ sur $] a, b [$, croissante et de classe C^1 , alors les intégrales $\int_a^b f(u) du$ et

$\int_\alpha^\beta f(\varphi(t)) \varphi'(t) dt$ sont de même nature et en cas de convergence sont égales.

Changement de variables.

Énoncé analogue dans le cas où φ est décroissante.

Les changements de variables non affines devront être indiqués aux candidats.

d) Compléments sur les variables aléatoires à densité

La notion générale d'espérance ou des moments d'ordre supérieur d'une variable aléatoire réelle quelconque est hors d'atteinte dans le cadre de ce programme. Néanmoins, on admettra que le théorème de transfert permet de calculer l'espérance de $g(X)$ dans le cas où X est une variable aléatoire à densité.

Exemples simples de calculs de fonctions de répartition et de densités de fonctions d'une variable aléatoire à densité.

Théorème de transfert.

Croissance de l'espérance pour les 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).

Moment d'ordre r ($r \in \mathbf{N}^*$).

e) Compléments sur les lois usuelles

Lois γ . Espérance et variance d'une variable aléatoire suivant une loi γ .

Transformées affines de variables aléatoires suivant des lois normales.

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

Rappels de première année pour des densités de variables aléatoires de la forme $aX + b$ ($a \neq 0$). En complément de la première année, les étudiants devront savoir retrouver les lois de X^2 et $\varphi(X)$ avec φ de classe C^1 strictement monotone sur $X(\Omega)$.

Si X est une variable aléatoire ayant une densité f_X nulle en dehors de l'intervalle $]a, b[$, avec $-\infty \leq a < b \leq +\infty$, 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 $\int_a^b g(t)f_X(t)dt$ si et seulement si cette intégrale converge absolument.

On pourra le démontrer 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.

Résultat admis.

Exemples de variables aléatoires n'admettant pas d'espérance ou de variance.

Notation $m_r(X) = E(X^r)$.

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

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

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

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.

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

2 - Couples de variables aléatoires

a) Cas général ; indépendance

Loi d'un couple de variables aléatoires réelles.

Si deux couples (X_1, Y_1) et (X_2, Y_2) ont même loi et si g est une fonction continue sur \mathbf{R}^2 à valeurs dans \mathbf{R} , alors les variables aléatoires $g(X_1, Y_1)$ et $g(X_2, Y_2)$ ont la même loi.

Indépendance de deux variables aléatoires.

Caractérisations de l'indépendance de deux variables aléatoires.

Espérance conditionnelle dans le cas de l'indépendance.

b) Couples de variables aléatoires réelles discrètes

Caractérisation de la loi d'un couple (X, Y) de variables aléatoires discrètes.

Caractérisation de l'indépendance de deux variables aléatoires discrètes.

Loi de la somme de deux variables aléatoires discrètes et indépendantes, produit de convolution discret.

Stabilité des lois binomiales et de Poisson.

La loi d'un couple (X, Y) de variables aléatoires réelles est donnée par la fonction $F_{(X,Y)}$ définie sur \mathbf{R}^2 par :

$$F_{(X,Y)}(x, y) = P([X \leq x] \cap [Y \leq y]).$$

Résultat admis.

Deux variables aléatoires X et Y sont indépendantes si et seulement si pour tous réels x et y :

$$P([X \leq x] \cap [Y \leq y]) = P([X \leq x]) P([Y \leq y]).$$

- Deux variables aléatoires X et Y sont indépendantes si et seulement si $P([X \in I] \cap [Y \in J]) = P([X \in I])P([Y \in J])$ pour tous intervalles I et J de \mathbf{R} .
- X et Y sont deux variables aléatoires indépendantes si et seulement si tout événement A de \mathcal{A}_X est indépendant de tout événement B de \mathcal{A}_Y .

Résultats admis.

Soit X une variable aléatoire discrète. Si Y est indépendante de X et si $A \in \mathcal{A}_Y$ est de probabilité non nulle, alors $E(X) = E(X|A)$.

La loi d'un couple (X, Y) de variables aléatoires discrètes est caractérisée par la donnée des valeurs $P([X = x] \cap [Y = y])$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.

Deux variables aléatoires X et Y discrètes sont indépendantes si et seulement si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,

$$P([X = x] \cap [Y = y]) = P([X = x])P([Y = y]).$$

$$P([X + Y = z]) = \sum_{x \in X(\Omega)} P([X = x])P([Y = z - x]).$$

- Si X_1 et X_2 sont deux variables aléatoires indépendantes suivant respectivement des lois $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$, alors $X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$.
- Si X_1 et X_2 sont deux variables aléatoires indépendantes suivant respectivement des lois $\mathcal{P}(\lambda_1)$ et $\mathcal{P}(\lambda_2)$, alors $X_1 + X_2 \hookrightarrow \mathcal{P}(\lambda_1 + \lambda_2)$.

Loi d'une variable aléatoire $Z = g(X, Y)$ où g est une fonction définie sur l'ensemble des valeurs prises par le couple (X, Y) .

Espérance de $Z = g(X, Y)$ et théorème de transfert.

Linéarité.

Espérance d'un produit de variables aléatoires discrètes indépendantes.

Covariance de deux variables aléatoires discrètes admettant un moment d'ordre 2. Propriétés.

Formule de Huygens.

Coefficient de corrélation linéaire.

Propriétés.

Variance de la somme de deux variables aléatoires discrètes.

c) Couples de variables aléatoires réelles à densité

En cas d'utilisation du produit de convolution, la preuve de sa légitimité n'est pas exigible des candidats.

Linéarité, positivité et croissance de l'espérance.

Densité de la somme $Z = X + Y$ de deux variables aléatoires à densité indépendantes, produit de convolution.

Stabilité de la loi γ pour la somme.

Stabilité de la loi normale pour la somme.

Espérance d'un produit de variables aléatoires à densité indépendantes.

On remarquera que $\mathcal{A}_Z \subset \mathcal{A}_{(X, Y)}$.

On se limitera à des cas simples tels que $\min(X, Y)$, $\max(X, Y)$, $X + Y$.

Sous réserve de convergence absolue :

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

Résultat admis.

Résultat admis.

Si X et Y sont deux variables aléatoires discrètes indépendantes admettant une espérance, alors XY admet également une espérance et $E(XY) = E(X)E(Y)$. On pourra admettre ce résultat.

Notation $\text{Cov}(X, Y)$.

Bilinéarité, symétrie, positivité de la covariance.

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Si X et Y sont indépendantes et admettent un moment d'ordre 2, leur covariance est nulle. La réciproque est fautive.

Notation $\rho(X, Y)$.

$|\rho(X, Y)| \leq 1$. Interprétation dans le cas où $\rho(X, Y) = \pm 1$.

Résultat admis.

Si la fonction h définie par la relation

$h(x) = \int_{-\infty}^{+\infty} f_X(t) f_Y(x - t) dt$ est définie et continue sauf peut-être en un nombre fini de points, c'est une densité de Z .

C'est le cas si f_X (ou f_Y) est bornée.

Si X_1 et X_2 sont deux variables aléatoires indépendantes suivant respectivement des lois $\gamma(\nu_1)$ et $\gamma(\nu_2)$, alors $X_1 + X_2 \leftrightarrow \gamma(\nu_1 + \nu_2)$.

Si X et Y sont deux variables aléatoires à densité indépendantes admettant une espérance, alors XY admet également une espérance et $E(XY) = E(X)E(Y)$. Résultat admis.

Variance de la somme de deux variables aléatoires à densité indépendantes.

Résultat admis.

3 - 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 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.

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

Loi marginale.

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

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

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, \dots, X_n) et (Y_1, Y_2, \dots, 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 $g(X_1, X_2, \dots, X_n)$ et $g(Y_1, Y_2, \dots, Y_n)$ ont même loi.

Espérance d'une somme de variables aléatoires.

Aucune difficulté ne sera soulevée.

Résultat admis.

Croissance de l'espérance.

Si X et Y admettent une espérance, $X + Y$ admet une espérance et $E(X + Y) = E(X) + E(Y)$. Généralisation à n variables aléatoires.

Résultats admis.

Si $X \leq Y$ presque sûrement et si X et Y admettent une espérance, alors $E(X) \leq E(Y)$.

Résultat admis.

Existence d'une espérance par domination.

Si X et Y sont deux variables aléatoires vérifiant $0 \leq |X| \leq Y$ presque sûrement, et si Y admet une espérance, alors X admet également une espérance. Dans ce cas, $|E(X)| \leq E(Y)$.

Résultat admis.

Indépendance mutuelle de n variables aléatoires réelles.

X_1, \dots, X_n sont mutuellement indépendantes si et seulement si :

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

pour tous réels x_1, \dots, x_n .

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

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

Lemme des coalitions.

Espérance du produit de variables aléatoires indépendantes.

Variance d'une somme de variables aléatoires indépendantes.

Somme de variables aléatoires de Bernoulli indépendantes de même paramètre.

Sommes de variables aléatoires indépendantes suivant des lois de Poisson, des lois binomiales. Loi de la somme de n variables aléatoires indépendantes de loi $\mathcal{E}(1)$.

Indépendance mutuelle d'une suite infinie de variables aléatoires réelles discrètes.

- X_1, \dots, X_n sont mutuellement indépendantes si et seulement si :

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

pour tous intervalles I_1, \dots, I_n de \mathbf{R} .

- X_1, \dots, X_n sont mutuellement indépendantes si et seulement si toute famille d'événements (A_1, \dots, A_n) , avec A_k élément de \mathcal{A}_{X_k} , est une famille d'événements mutuellement indépendants.

Résultat admis.

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

Résultat admis.

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

Résultat admis.

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.

Si X et Y sont indépendantes et admettent une variance, $X + Y$ admet une variance et $V(X + Y) = V(X) + V(Y)$.

Généralisation à n variables aléatoires mutuellement indépendantes.

Résultats admis.

La somme de n variables aléatoires de Bernoulli indépendantes et de même espérance p suit la loi binomiale $\mathcal{B}(n, p)$.

Pour étudier la somme de n variables aléatoires indépendantes de loi $\mathcal{E}(\lambda)$, on se ramènera après multiplication par λ à une somme de n variables aléatoires indépendantes de loi $\mathcal{E}(1)$.

ENSEIGNEMENT DE MATHÉMATIQUES DU QUATRIÈME SEMESTRE

I - Compléments d'algèbre bilinéaire

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

Endomorphismes symétriques.

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.

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 :

$$\langle f(x), y \rangle = \langle x, f(y) \rangle.$$

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

2 - Projection orthogonale

Projection orthogonale sur un sous-espace vectoriel F .

Si (u_1, \dots, 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.

Notation p_F .

Si \mathcal{B} est une base orthonormée de E et si U_1, \dots, U_k sont les vecteurs colonnes associés aux vecteurs u_1, \dots, u_k dans la base \mathcal{B} , alors :

$$\text{Mat}_{\mathcal{B}}(p_F) = \sum_{i=1}^k U_i {}^t U_i.$$

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

Application au problème des moindres carrés : 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 $X \in \mathcal{M}_{p,1}(\mathbf{R})$.

La formule donnant la valeur de X réalisant le minimum n'est pas exigible.

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.

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 $D = P^{-1}AP = {}^tPAP$.

Si X_1, \dots, X_n sont les colonnes de P , alors $(X_i)_{1 \leq i \leq n}$ est une base orthonormée de $\mathcal{M}_{n,1}(\mathbf{R})$, formée de vecteurs propres de A associés aux valeurs propres $\lambda_1, \dots, \lambda_n$. On a :

$$A = \sum_{i=1}^n \lambda_i X_i {}^tX_i.$$

II - Fonctions réelles de n variables ; recherche d'extrema

L'objectif est d'arriver à une bonne maîtrise des problèmes d'extrema 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 $\{x \in \mathbf{R}^n / \varphi(x) \leq a\}$ où φ 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 - Extension de la notion de fonction réelle de n variables

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 aux fonctions définies sur un ouvert de \mathbf{R}^n .

Aucune difficulté théorique ne sera soulevée.

Pour les fonctions C^1 définies sur un ouvert de \mathbf{R}^n : extension des notions de dérivées partielles, gradient, dérivée directionnelle, développement limité d'ordre 1.

2 - Fonctions de classe C^2

Dérivées partielles d'ordre 2.

Notation $\partial_{i,j}^2(f)$.

$\partial_{i,j}^2(f) = \partial_i(\partial_j(f))$.

Fonctions de classe C^2 sur un ouvert \mathcal{O} de \mathbf{R}^n .

Les fonctions polynomiales de n variables sont de classe C^2 sur \mathbf{R}^n . Résultat admis.

Opérations sur les fonctions de classe C^2 .

Théorème de Schwarz.

Matrice hessienne en un point x .

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

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

Dérivée seconde directionnelle de f au point x dans la direction h .

Somme, produit, quotient.

La composition d'une fonction de classe C^2 sur \mathcal{O} à valeurs dans un intervalle I de \mathbf{R} par une fonction de classe C^2 sur I à valeurs dans \mathbf{R} est de classe C^2 .

Résultats admis.

Si f est de classe C^2 sur un ouvert \mathcal{O} de \mathbf{R}^n , alors pour tout point x de \mathcal{O} et pour tout couple (i, j) d'entiers compris entre 1 et n :

$$\partial_{i,j}^2(f)(x) = \partial_{j,i}^2(f)(x).$$

Théorème admis.

Notation $\nabla^2(f)(x)$.

Si f est de classe C^2 sur un ouvert \mathcal{O} , alors la matrice hessienne est symétrique en tout point de \mathcal{O} .

$$q(h) = {}^tH A H.$$

$$f(x+h) = f(x) + \langle \nabla(f)(x), h \rangle + \frac{1}{2} q_x(h) + \|h\|^2 \varepsilon(h)$$

où $\varepsilon(0) = 0$, ε continue en $\bar{0}$ et q_x est la forme quadratique associée à la matrice hessienne $\nabla^2(f)(x)$.

Résultat admis.

La dérivée seconde directionnelle de f au point x dans la direction h est $q_x(h)$.

Si $g(t) = f(x+th)$, alors $g''(t) = q_{x+th}(h)$ 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.

Application à l'encadrement d'une forme quadratique sur \mathbf{R}^n .

Si q est une forme quadratique associée à une matrice symétrique A , alors q admet un maximum global et un minimum global sur le fermé borné $\{x \in \mathbf{R}^n / \|x\| = 1\}$.

Il existe alors deux réels α et β tels que pour tout élément h de \mathbf{R}^n :

$$\alpha \|h\|^2 \leq q(h) \leq \beta \|h\|^2.$$

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. Toutes les dérivées directionnelles en ces points sont nulles.

d) Exemples de recherches d'extrema sous une contrainte quelconque

Dans tout ce paragraphe, \mathcal{O} désigne un ouvert de \mathbf{R}^n et φ désigne une fonction de classe C^1 sur \mathcal{O} . On note alors $\mathcal{C} = \{x \in \mathcal{O} / \varphi(x) = c\}$ l'ensemble des points vérifiant la contrainte $\varphi(x) = c$. On se placera dans le cadre où $\nabla(\varphi)(x)$ est non nul pour tout élément x de \mathcal{C} ; on dira alors que la contrainte est non critique. L'étude d'une fonction le long de sa frontière est l'une des applications de la maximisation sous contrainte.

Sur des exemples, on sensibilisera les étudiants à une interprétation graphique de l'optimisation sous contrainte.

Définition d'un extremum local ou global sous la contrainte $\varphi(x) = c$ d'une fonction f définie sur \mathcal{O} .

Condition nécessaire du premier ordre pour un extremum sous la contrainte non critique \mathcal{C} .

Si f est une fonction de classe C^1 sur \mathcal{O} , pour que f atteigne un extremum local en x_0 sous la contrainte non critique \mathcal{C} , il faut qu'il existe un réel λ tel que :

$$\begin{cases} \varphi(x_0) = c \\ \nabla(f)(x_0) = \lambda \nabla(\varphi)(x_0) \end{cases}$$

Résultat admis.

On ne soulèvera pas de difficulté théorique et on se limitera à des exemples simples.

Si q est une forme quadratique associée à une matrice symétrique A , alors q admet un maximum global (respectivement un minimum global) sous la contrainte $\|x\| = 1$, en un point correspondant à un vecteur propre de la matrice A associé à la plus grande valeur propre (respectivement la plus petite).

Application à l'encadrement d'une forme quadratique : cas d'égalité.

e) Condition d'ordre 2

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

Point selle (ou col).

Exemples de recherche d'extrema globaux.

Si x_0 est un point critique de f :

- si $\text{Sp}(\nabla^2 f(x_0)) \subset \mathbf{R}_+^*$, alors f admet un minimum local en x_0 ,
- si $\text{Sp}(\nabla^2 f(x_0)) \subset \mathbf{R}_-^*$, alors f admet un maximum local en x_0 ,
- si $\text{Sp}(\nabla^2 f(x_0))$ 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 forme quadratique q_{x_0} .

On pourra faire une étude directe du signe sur \mathcal{O} de $f(x) - f(x_0)$.

Dans les situations qui s'y prêtent, on pourra étudier le cas où pour tout x de \mathcal{O} , q_x est positive (ou négative), par exemple en appliquant la formule de Taylor avec reste intégral à la fonction g définie par $g(t) = f(x_0 + th)$.

f) Recherche d'extrema sous contrainte d'égalités linéaires

Dans tout ce paragraphe \mathcal{C} désigne l'ensemble des solutions d'un système linéaire $\begin{cases} g_1(x) = b_1 \\ \vdots \\ g_p(x) = b_p \end{cases}$ 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 $\text{Vect}(\nabla(g_1)(x_0), \dots, \nabla(g_p)(x_0))$. Pour tout h de \mathcal{H} , la dérivée de f en x_0 dans la direction h est nulle.

On remarquera que :

$$\mathcal{H}^\perp = \text{Vect}(\nabla(g_1)(x_0), \dots, \nabla(g_p)(x_0)).$$

Point critique pour l'optimisation sous contrainte.

Exemples de recherche d'extrema globaux sous contrainte d'égalités linéaires.

On pourra faire une étude directe du signe sur \mathcal{O} de $f(x) - f(x_0)$. On pourra, dans les situations qui s'y prêtent, étudier le cas où pour tout x de $\mathcal{C} \cap \mathcal{O}$ et tout h de \mathcal{H} on a $q_x(h) \geq 0$ (respectivement $q_x(h) \leq 0$).

III - Probabilités : convergences, estimation

1 - Convergences et approximations

a) Convergence en probabilité

On pourra démontrer ces inégalités dans le cas d'une variable aléatoire discrète ou à densité.

Inégalité de Markov.

Si X est une variable aléatoire positive admettant une espérance, alors pour tout réel a strictement positif :

$$P([X \geq a]) \leq \frac{E(X)}{a}.$$

On pourra appliquer cette inégalité à $Y = |X|^r$, $r \in \mathbf{N}^*$.

Inégalité de Bienaymé-Tchebychev.

Si X est une variable aléatoire admettant un moment d'ordre 2, alors :

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

Convergence en probabilité.

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

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P(|X_n - X| \geq \varepsilon) = 0.$$

Notation $X_n \xrightarrow{P} X$.

Loi faible des grands nombres pour une suite de variables aléatoires indépendantes admettant une même espérance et une même variance.

Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes ayant même espérance m et même variance et soit $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$.

Alors :

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P(|\bar{X}_n - m| \geq \varepsilon) = 0.$$

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.

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 :

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

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

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

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

Convergence en loi d'une suite de variables aléatoires suivant la loi binomiale $\mathcal{B}(n, \lambda/n)$ vers une variable aléatoire suivant la loi de Poisson de paramètre λ .

Théorème de Slutsky.

Si $(X_n)_{n \in \mathbf{N}^*}$ converge en loi vers X et si $(Y_n)_{n \in \mathbf{N}^*}$ converge en probabilité vers une constante c , alors $(X_n + Y_n)_{n \in \mathbf{N}^*}$ converge en loi vers $X + c$ et $(X_n Y_n)_{n \in \mathbf{N}^*}$ converge en loi vers cX .

Composition par une fonction continue.

Résultats admis.

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 σ^2 non nulle, si on note : $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$, alors la suite de variables aléatoires centrées réduites $\bar{X}_n^* = \sqrt{n} \left(\frac{\bar{X}_n - m}{\sigma} \right)$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.

D'où, on a pour tout (a, b) tel que $-\infty \leq a \leq b \leq +\infty$:

$$\lim_{n \rightarrow +\infty} P([a \leq \bar{X}_n^* \leq b]) = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Exemples d'approximations de la loi binomiale et de la loi de Poisson par la loi normale.

Résultats admis.

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 θ décrivant un sous-ensemble Θ de \mathbf{R} (éventuellement de \mathbf{R}^2). Le paramètre θ 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 θ ou de $g(\theta)$ (fonction à valeurs réelles du paramètre θ), à partir d'un échantillon de données x_1, \dots, 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, \dots, X_n définies sur un même espace probabilisable (Ω, \mathcal{A}) muni d'une famille de probabilités $(P_\theta)_{\theta \in \Theta}$. Les X_1, \dots, X_n seront supposées P_θ -indépendantes et de même loi que X pour tout θ .

On appellera estimateur de $g(\theta)$ toute variable aléatoire réelle de la forme $\varphi(X_1, X_2, \dots, X_n)$ où φ est une fonction de \mathbf{R}^n dans \mathbf{R} , éventuellement dépendante de n , et indépendante de θ , 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_θ .

a) Estimation ponctuelle

Estimer ponctuellement $g(\theta)$ par $\varphi(x_1, \dots, x_n)$ où $\varphi(X_1, X_2, \dots, X_n)$ est un estimateur de $g(\theta)$ et (x_1, \dots, x_n) est une réalisation de l'échantillon (X_1, \dots, X_n) , c'est décider d'accorder à $g(\theta)$ la valeur $\varphi(x_1, \dots, x_n)$.

n -échantillon (X_1, \dots, X_n) de variables aléatoires réelles indépendantes et de même loi que X .

Définition d'un estimateur.

Exemples simples d'estimateurs.

Biais d'un estimateur.

Estimateur sans biais.

Suite $(T_n)_{n \geq 1}$ d'estimateurs.

Estimateur convergent.

Composition par une fonction continue.

Estimateur asymptotiquement sans biais.

Risque quadratique d'un estimateur.

Décomposition biais-variance du risque quadratique d'un estimateur.

Exemples de n -échantillons associés à une loi de Bernoulli $\mathcal{B}(1, p)$ avec $\theta = p$.

Un estimateur de $g(\theta)$ est une variable aléatoire de la forme $T_n = \varphi(X_1, \dots, X_n)$. La réalisation $\varphi(x_1, \dots, x_n)$ de l'estimateur T_n est l'estimation de $g(\theta)$. Cette estimation ne dépend que de l'échantillon (x_1, x_2, \dots, x_n) observé.

Exemple de la moyenne empirique $\frac{X_1 + X_2 + \dots + X_n}{n}$.

Si pour tout θ de Θ , T_n admet une espérance, on appelle biais de T_n en $g(\theta)$ le réel $b_\theta(T_n) = E_\theta(T_n) - g(\theta)$.

L'estimateur T_n de $g(\theta)$ est sans biais si pour tout θ de Θ , $E_\theta(T_n) = g(\theta)$.

Chaque T_n est de la forme $\varphi(X_1, X_2, \dots, X_n)$.

Une suite d'estimateurs $(T_n)_{n \geq 1}$ de $g(\theta)$ est convergente si pour tout θ , la suite $(T_n)_{n \geq 1}$ converge en probabilité vers $g(\theta)$.

Par abus de langage on dit aussi que l'estimateur est convergent.

Si $(T_n)_{n \geq 1}$ est une suite convergente d'estimateurs de $g(\theta)$ et si f est une fonction continue sur \mathbf{R} à valeurs réelles, alors $(f(T_n))_{n \geq 1}$ est une suite convergente d'estimateurs de $f(g(\theta))$.

Une suite $(T_n)_{n \geq 1}$ d'estimateurs de $g(\theta)$ est asymptotiquement sans biais si pour tout θ de Θ ,

$$\lim_{n \rightarrow +\infty} E_\theta(T_n) = g(\theta).$$

Si pour tout θ de Θ , T_n admet un moment d'ordre 2, le risque quadratique de T_n en θ , noté $r_\theta(T_n)$, est le réel :

$$r_\theta(T_n) = E_\theta((T_n - g(\theta))^2).$$

$$r_\theta(T_n) = b_\theta(T_n)^2 + V_\theta(T_n).$$

Condition suffisante de convergence.

Si pour tout θ de Θ , $\lim_{n \rightarrow +\infty} r_\theta(T_n) = 0$, alors la suite d'estimateurs $(T_n)_{n \geq 1}$ de $g(\theta)$ est convergente.

Cette convergence pourra être étudiée à l'aide de l'inégalité de Markov.

b) Estimation par intervalle de confiance, intervalle de confiance asymptotique

S'il existe des critères pour juger des qualités d'un estimateur ponctuel T_n de $g(\theta)$ (biais, risque, convergence), 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. L'utilisation dans certains cas du théorème limite central impose d'introduire la notion d'intervalle de confiance asymptotique.

Dans tout ce paragraphe $(U_n)_{n \geq 1}$ et $(V_n)_{n \geq 1}$ désigneront deux suites d'estimateurs de $g(\theta)$ tels que pour tout $\theta \in \Theta$ et pour tout $n \geq 1$, $P_\theta([U_n \leq V_n]) = 1$.

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 θ de Θ ,

$$P_\theta([U_n \leq g(\theta) \leq V_n]) \geq 1 - \alpha.$$

Sa réalisation est l'estimation de cet intervalle de confiance.

On distinguera probabilité et confiance et on éclairera ces notions à l'aide de simulations informatiques.

Intervalle de confiance asymptotique.

On appelle intervalle de confiance asymptotique de $g(\theta)$ au niveau de confiance $1 - \alpha$ une suite $([U_n, V_n])_{n \geq 1}$ vérifiant : pour tout θ de Θ , il existe une suite de réels (α_n) à valeurs dans $[0, 1]$, de limite α , telle que pour tout $n \geq 1$,

$$P_\theta([U_n \leq g(\theta) \leq V_n]) \geq 1 - \alpha_n.$$

Par abus de langage on dit aussi que $[U_n, V_n]$ est un intervalle de confiance asymptotique.

Estimation par intervalle du paramètre d'une variable aléatoire de Bernoulli.

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.

Estimation par intervalle de confiance de l'espérance d'une loi admettant un moment d'ordre 2.

On pourra démontrer ce résultat dans le cas d'une loi admettant un moment d'ordre 4 : dans ce cas, la suite $(S_n)_{n \geq 1}$ des écarts-types empiriques converge en probabilité vers l'écart-type σ inconnu, ce qui permet d'utiliser le théorème de Slutsky pour remplacer σ par S_n afin d'obtenir une estimation par intervalle de confiance asymptotique de l'espérance de la loi.

TRAVAUX PRATIQUES DE MATHÉMATIQUES AVEC SCILAB

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

Les heures de travaux pratiques de mathématiques avec Scilab peuvent être organisées sous différentes formes selon les contenus à enseigner ; certaines séances, notamment celles nécessitant peu de manipulations logicielles de la part des étudiants, pourront avoir lieu en classe entière, les autres séances en groupes réduits.

Le programme d'informatique s'articule autour de six thèmes : statistiques descriptives univariées, statistiques descriptives bivariées, chaînes de Markov, 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 logiciel Scilab ne peut pas être entièrement maîtrisée par un étudiant, aussi seules les fonctions et commandes introduites en première année et celles figurant dans la sous-partie "Commandes exigibles" sont exigibles. Néanmoins, se contenter de ces seules commandes, en ignorant les nombreuses possibilités et commodités du logiciel, se révélerait rapidement contraignant et limitatif. De nouvelles commandes Scilab 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. 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 Scilab, 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 - 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 ou trois variables.

C4 : Représenter et interpréter les différentes convergences.

C5 : Utiliser à bon escient la méthode de Monte-Carlo.

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

2 - Nouvelles commandes

Toutes les commandes du programme de première année sont exigibles. Les seules nouvelles commandes exigibles des candidats sont indiquées dans ce paragraphe.

La connaissance des commandes suivantes ainsi que de leurs arguments est exigible des candidats :

`sum`, `cumsum`, `mean`, `max`, `min`, `zeros`, `ones`, `eye`, `spec`.

Les commandes suivantes devront avoir été manipulées par les étudiants mais la connaissance détaillée de leurs arguments n'est pas exigible des candidats :

`cdfnor`, `plot2d`, `fplot2d`, `plot3d`, `fplot3d`.

II - Liste des thèmes

1 - Statistiques descriptives univariées

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

Dans ce paragraphe, on analysera des données statistiques, en insistant sur les représentations graphiques. On insistera sur le rôle des différents indicateurs de position et de dispersion étudiés.

Série statistique associée à un échantillon.
Effectifs, fréquences, fréquences cumulées, diagrammes en bâtons, histogrammes.

Indicateurs de position : moyenne, médiane, mode, quantiles.

Indicateurs de dispersion : étendue, variance et écart-type empiriques, écart inter-quantile.

On pourra également utiliser les commandes : `dsearch`, `tabul`, `pie` `stdeviation`, `median`.

2 - Statistiques descriptives bivariées

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

Série statistique à deux variables, nuage de points associé.

Point moyen (\bar{x}, \bar{y}) du nuage.

Covariance et coefficient de corrélation empiriques, droites de régression.

On tracera le nuage de points et les droites de régression et on pourra effectuer des pré-transformations pour se ramener au cas linéaire.

On différenciera les variables explicatives des variables à expliquer. On pourra utiliser les commandes : `stdeviation`, `corr`.

3 - Chaînes de Markov

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

Ce thème sera l'occasion de revoir les simulations de lois discrètes étudiées en première année ainsi que d'appliquer les résultats et techniques d'algèbre linéaire étudiés au troisième semestre.

Matrice de transition.

Étude sur des exemples simples.

Comportement limite.

On pourra étudier par exemple l'indice de popularité d'une page Web (PageRank), modéliser l'évolution sociologique d'une société (passage d'individus d'une classe sociale à une autre) ou les systèmes de bonus-malus en assurances. Simulation et mise en évidence d'états stables avec la commande `grand(n, 'markov', M, x0)`.

4 - 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.

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.

À cette occasion, on pourra mettre en évidence l'orthogonalité du gradient avec les courbes de niveau d'une fonction de deux variables.

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 : minimisation du risque, maximisation du gain, etc.

5 - 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.

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 du générateur `grand`.

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.

On pourra s'intéresser au cas particulier de 12 variables aléatoires indépendantes suivant une même loi uniforme.

Simulations de variables aléatoires discrètes et à densité variées.

On pourra faire le lien entre les lois exponentielles, les lois γ et de Poisson en modélisant des temps d'attente.

6 - 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).

Ainsi, on pourra estimer par exemple les valeurs prises par la fonction de répartition de la somme ou du produit de deux variables aléatoires, ou encore, estimer le niveau réel, à rang n fini, d'intervalles de confiance asymptotiques.

Comparaison de différents estimateurs ponctuels d'un paramètre.

On pourra utiliser des données issues de situations réelles ou créer plusieurs jeux de données obtenues grâce à la commande **grand**. 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.

La comparaison pourra se faire en calculant les demi-largeurs moyennes des intervalles et leurs niveaux de confiance.