Classes préparatoires aux grandes écoles

Filière économique

Voie technologique
ECT

Programmes de mathématiques - informatique

2$^{\text{nde}}$ année

Table des matières

1 Objectifs généraux de la formation

Les mathématiques jouent un rôle important en sciences économiques et en gestion, notamment dans les domaines 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.

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

L'objectif 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.

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 technologique 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 d'économie et de gestion dispensés en Grande Ecole ou dans une formation universitaire de troisième année de Licence.

Il s'organise autour de quatre points forts :

- En algèbre linéaire, le programme se concentre sur le calcul matriciel. Le principal objectif est l'introduction de la notion de valeurs propres et de vecteurs propres et la diagonalisation des matrices carrées de taille inférieure à 3. On évitera des exemples trop calculatoires.

- En analyse, les séries et les intégrales généralisées sont étudiées en vue de leurs applications aux probabilités (variables aléatoires discrètes infinies et variables aléatoires à densité).

- En probabilités, l'étude des variables aléatoires discrètes, initiée au lycée et poursuivie en première année de classe préparatoire, 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'objectif de cette partie du programme est de permettre, en fin de formation, une approche plus rigoureuse et une compréhension plus aboutie des concepts d'estimation ponctuelle ou par intervalles de confiance que les étudiants ont rencontrés dès le lycée.

- Les travaux pratiques de mathématiques et d'informatique sont organisés autour de la poursuite de l'étude des fonctionnalités du langage SQL et, avec Python, de la simulation de lois de probabilités en continuité du programme de première année, et de thèmes de statistiques en lien avec le programme de mathématiques, avec l'objectif d'éclairer ces notions par des illustrations concrètes. 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.

Le programme de mathématiques est organisé en deux semestres de volume sensiblement équivalent. 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 dès que possible l'outil informatique en cours de mathématiques pour visualiser et illustrer les notions étudiées. Dans certaines situations, en continuité avec les programmes de lycée, l'utilisation d'un tableur peut s'avérer adaptée.

Les étudiants ont déjà une pratique algorithmique acquise au lycée. Dans leurs études futures, ils seront amenés à utiliser différents logiciels conçus pour la résolution de problématiques liées à certains contextes. Une pratique régulière d'outils informatiques les prépare utilement en ce sens. Par ailleurs, l'utilisation d'un outil informatique (programme informatique ou tableur) permet l'observation de résultats mathématiques en situation, l'exploration et la modélisation de situations non triviales plus réalistes et offre la possibilité d'expérimenter et de conjecturer.

ENSEIGNEMENT DE MATHÉMATIQUES DU TROISIÈME SEMESTRE

I - **Matrices**

Le programme exclut toute notion de structure. On ne traite que le cas des matrices réelles.

Matrices carrées d'ordre n. Ensemble $\mathcal{M}_n(\mathbf{R})$. Matrices triangulaires, matrices diagonales, matrice identité. Matrices inversibles.	
Critère d'inversibilité d'une matrice triangulaire.	Résultat admis.
Critère d'inversibilité d'une matrice carrée d'ordre 2.	$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible si et seulement si $ad - bc \neq 0$. Formule de l'inverse dans ce cas. La notation $\det(A)$ pourra être utilisée, mais elle sera limitée au cas des matrices carrées d'ordre 2. La notion de déterminant est hors-programme.
Exemples de calcul des puissances n-ièmes d'une matrice. Cas d'une matrice diagonale.	
Formule du binôme pour les matrices qui commutent.	On se limitera à des exemples simples, par exemple lorsque l'une des matrices est nilpotente.

Écriture matricielle d'un système d'équations linéaires.	
Calcul de l'inverse d'une matrice par la méthode du pivot de Gauss. Calcul de l'inverse de la matrice A par la résolution du système $AX = Y$.	On se limitera à des matrices carrées d'ordre inférieur ou égal à 3.

II - Compléments d'intégration : propriétés de l'intégrale

Ce chapitre sera l'occasion de revenir sur les calculs d'intégrales introduits au S2.

Linéarité de l'intégrale.	
Intégration par parties.	Si u, v, u' et v' sont des fonctions continues sur $[a, b]$, alors : $\int_a^b u'(t)v(t)dt = [u(t)v(t)]_a^b - \int_a^b u(t)v'(t)dt$.
Positivité de l'intégrale. Comparaison d'intégrales.	

III - Compléments sur les sommes et séries numériques

1 - Compléments sur les sommes

Ce chapitre sera l'occasion de revenir sur les calculs de sommes traités en première année $\boxed{\blacktriangleright}$.

Somme télescopique.	
Décalage d'indice.	On se limitera, sur des exemples simples, à des décalages d'indice de type $k' = k + 1$.

2 - Séries

Les séries sont introduites exclusivement pour leurs applications au calcul des probabilités. Aucune difficulté ne sera soulevée.

Définition. Convergence d'une série. Somme d'une série convergente.	
Condition nécessaire de convergence.	Le terme général d'une série convergente tend vers 0.
Série géométrique. Convergence et somme.	La série $\sum x^n$ converge si et seulement si $\mid x \mid < 1$, et dans ce cas : $\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$. Les dérivées des séries géométriques ne font pas partie des attendus du programme.

IV - Probabilités et statistiques

Tout excès de technicité est exclu.

1 - Couples de variables aléatoires discrètes finies

Loi de probabilité d'un couple de variables aléatoires. Lois marginales, lois conditionnelles.	La loi de probabilité d'un couple de variables aléatoires discrètes est caractérisée par la donnée de $X(\Omega)$, $Y(\Omega)$ et pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, $P([X = x] \cap [Y = y])$.
Indépendance de deux variables aléatoires.	X et Y sont indépendantes si, pour tous intervalles réels I et J, les événements $[X \in I]$ et $[Y \in J]$ sont indépendants. On remarquera que si l'une des variables aléatoires X, Y est constante, X et Y sont indépendantes.
Espérance d'une somme de deux variables aléatoires, linéarité de l'espérance.	Résultat admis.
Espérance d'un produit de deux variables aléatoires.	$E(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP([X = x] \cap [Y = y])$. Résultat admis.
Cas de deux variables aléatoires X et Y indépendantes.	$E(XY) = E(X)E(Y)$. Résultat admis. La réciproque est fausse.
Covariance. Propriétés.	Notation $\mathrm{Cov}(X, Y)$. Linéarité à droite, à gauche. Symétrie. Si $a \in \mathbf{R}$, $\mathrm{Cov}(X, a) = 0$. $\mathrm{Cov}(X, X) = V(X)$.
Formule de Kœnig-Huygens.	$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$. Si X et Y sont indépendantes, leur covariance est nulle, la réciproque étant fausse.
Variance d'une somme de deux variables aléatoires.	
Coefficient de corrélation linéaire.	Notation $\rho(X, Y)$. Si $\sigma(X)\sigma(Y) \neq 0$, $\rho(X, Y) = \dfrac{\mathrm{Cov}\,(\mathrm{X}, \mathrm{Y})}{\sigma(X)\sigma(Y)}$.
Propriétés.	$\vert\rho(X, Y)\vert \leqslant 1$. Interprétation dans le cas où $\rho(X, Y) = \pm 1$.

2 - Variables aléatoires discrètes infinies

a) Généralités

On se limitera aux variables aléatoires positives dont l'image est indexée par $\mathbf{N}$*. Aucune difficulté théorique ne sera soulevée au moment de l'extension des propriétés.*

Notion d'espace probabilisé avec Ω non fini.
Extension des définitions et des propriétés des variables aléatoires discrètes au cas où l'image est un ensemble infini dénombrable : loi de probabilité, fonction de répartition, espérance, variance, écart-type.

b) Lois usuelles

Chacune des lois usuelles sera illustrée par un exemple concret d'une situation qu'elle modélise.

Loi géométrique.	Notation $X \hookrightarrow \mathcal{G}(p)$. La reconnaissance de la loi géométrique comme loi du premier succès est exigible.
Espérance et variance.	Résultats admis.
Loi de Poisson.	Notation $X \hookrightarrow \mathcal{P}(\lambda)$.
Espérance et variance.	Résultats admis.

ENSEIGNEMENT DE MATHÉMATIQUES DU QUATRIÈME SEMESTRE

I - Réduction des matrices carrées

L'objectif est l'introduction de la notion de valeurs propres et de vecteurs propres d'une matrice. La notion de polynôme minimal, la résolution générale des systèmes $AX = \lambda X$ (avec λ paramètre quelconque) et toute théorie sur la réduction sont hors programme.

Dans tout ce paragraphe, on évitera les méthodes trop calculatoires pour la recherche des éléments propres d'une matrice. En particulier, la résolution de systèmes à paramètres est à proscrire. Dans la pratique, on se limitera à des matrices carrées d'ordre inférieur ou égal à 3.

Polynôme d'une matrice. Polynôme annulateur.	Sur des exemples, utilisation d'un polynôme annulateur pour la détermination de l'inverse d'une matrice carrée. Toutes les indications devront être données aux candidats pour l'obtention d'un polynôme annulateur. On pourra vérifier que le polynôme $X^2 - (a+d)\,X + (ad - bc)$ est un polynôme annulateur de la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.
Matrices carrées diagonalisables.	Une matrice carrée A est diagonalisable s'il existe une matrice D, diagonale, et une matrice carrée P, inversible, telles que $D = P^{-1}AP$.
Valeur propre, vecteur propre d'une matrice carrée.	Avec les notations de la définition précédente, on remarquera que la matrice P est construite à partir de vecteurs propres de A et la matrice D des valeurs propres correspondantes, mais leur construction n'est pas exigible.
Si Q est un polynôme annulateur de A, toute valeur propre de A est racine de Q.	Résultat admis.

Recherche de valeurs propres.

Pour ce faire, on utilisera un polynôme annulateur.
La recherche de vecteurs propres ne pourra être demandée que dans le cas de valeurs propres de multiplicité 1. Dans les autres cas, les vecteurs propres devront être donnés.

Sur des exemples, diagonalisation d'une matrice carrée d'ordre inférieur ou égal à 3.
Application au calcul des puissances de A.

Sur des exemples, étude de suites linéaires récurrentes d'ordre 2 et de systèmes de suites récurrentes.
La méthode générale de résolution est hors-programme.

II - **Compléments d'analyse**

Les notions introduites dans ce chapitre le sont exclusivement pour leurs applications au calcul des probabilités. Aucune difficulté ne sera soulevée.

Le calcul des intégrales généralisées est effectué par des recherches de primitives sur des intervalles du type $[a, b]$, l'application de la relation de Chasles, et des passages à la limite en $-\infty$ et/ou $+\infty$.

Les intégrales généralisées en un point réel sont hors-programme.

Intégrale $\displaystyle\int_a^{+\infty} f(t)\mathrm{d}t$ où f est une fonction continue sur $[a, +\infty[$. Convergence et définition.

L'intégrale $\displaystyle\int_a^{+\infty} f(t)\mathrm{d}t$ converge si $\displaystyle\lim_{x\to+\infty}\int_a^x f(t)\mathrm{d}t$ existe et est finie, et dans ce cas, $\displaystyle\int_a^{+\infty} f(t)\mathrm{d}t = \lim_{x\to+\infty}\int_a^x f(t)\mathrm{d}t$.

Intégrale $\displaystyle\int_{-\infty}^{b} f(t)\mathrm{d}t$ où f est une fonction continue sur $]-\infty, b]$.

Extension aux intégrales $\displaystyle\int_{-\infty}^{+\infty} f(t)\mathrm{d}t$.

Aucune difficulté théorique ne sera soulevée lors du passage du calcul des intégrales des fonctions continues à celui des intégrales des fonctions continues sauf en un nombre fini de points.

III - **Probabilités et statistiques**

1 - Variables aléatoires à densité continue par morceaux

Ce paragraphe généralise l'étude de la loi uniforme effectuée en première année.

Le passage du cas discret au cas continu n'est pas explicité. On se limitera à des calculs de probabilités du type $P([X \in I])$, où I est un intervalle de $\mathbf{R}$.

Densité de probabilité.	Une fonction f définie sur $\mathbf{R}$ est une densité de probabilité si elle est positive, continue sur $\mathbf{R}$ éventuellement privé d'un nombre fini de points et telle que $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$. On se limitera en pratique à des fonctions continues par morceaux, cette notion étant elle-même hors-programme.
Variable aléatoire à densité.	Une variable aléatoire X admet une densité si sa fonction de répartition F_X peut s'écrire sous la forme $x \mapsto \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t$ où f est une densité de probabilité. Sur des exemples, détermination d'une densité de $aX+b$ ou de X^2.
Espérance, variance et écart-type.	Aucune difficulté théorique ne sera soulevée.

2 - Variables aléatoires à densité usuelles

Chacune des lois usuelles sera illustrée par un exemple concret d'une situation qu'elle modélise.

Loi uniforme. Densité et fonction de répartition. Espérance et variance.	Notation $X \hookrightarrow \mathcal{U}[a,b]$.
Loi exponentielle. Densité et fonction de répartition. Espérance et variance.	Notation $X \hookrightarrow \mathcal{E}(\lambda)$.
Loi normale (ou de Laplace-Gauss) de paramètres m et σ^2, où $\sigma > 0$. Espérance et variance.	Densité. Notation $X \hookrightarrow \mathcal{N}(m,\sigma^2)$. Résultats admis.
Loi normale centrée réduite.	Densité. $X \hookrightarrow \mathcal{N}(m,\sigma^2)$ si et seulement si $X^* = \dfrac{X-m}{\sigma} \hookrightarrow \mathcal{N}(0,1)$. On attend des étudiants qu'ils sachent utiliser la fonction de répartition Φ de la loi normale centrée réduite. Pour tout réel x : $\Phi(-x) = 1 - \Phi(x)$.

3 - Convergences et approximations

a) Inégalité de Markov, inégalité de Bienaymé-Tchebychev.

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 à valeurs positives et admettant une espérance, $$\forall a > 0, \quad P([X \geqslant a]) \leqslant \frac{E(X)}{a}.$$ Résultat non exigible. On pourra appliquer cette inégalité à $Y = \lvert X\rvert^r, r \in \mathbf{N}^*$.
Inégalité de Bienaymé-Tchebychev.	Si X est une variable aléatoire admettant un moment d'ordre 2, $$\forall \varepsilon > 0, \quad P([\mid X - E(X) \mid \geqslant \varepsilon]) \leqslant \frac{V(X)}{\varepsilon^2}.$$ Résultat non exigible.

b) Suites de variables aléatoires discrètes finies

Indépendance mutuelle de n variables aléatoires.	Les variables aléatoires $X_1, \ldots, X_n$ sont mutuellement indépendantes si, pour tout choix de n intervalles réels $I_1, \ldots, I_n$, les événements $[X_1 \in I_1], \ldots, [X_n \in I_n]$ sont mutuellement indépendants.
Indépendance mutuelle d'une suite de variables aléatoires.	Les variables aléatoires de la suite $(X_n)_{n\in\mathbf{N}^*}$ sont dites mutuellement indépendantes si, pour tout entier $n \geqslant 1$, les variables aléatoires $X_1, \ldots, X_n$ sont mutuellement indépendantes.
Espérance de la somme de n variables aléatoires.	
Variance d'une somme finie de variables aléatoires indépendantes.	Application à la somme de n variables aléatoires de Bernoulli indépendantes.

c) Loi faible des grands nombres

Loi faible des grands nombres.	Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes admettant une même espérance m et une même variance et soit pour tout $n \in \mathbf{N}^*$, $\overline{X}_n = \dfrac{X_1 + \ldots + X_n}{n}$. Alors $\forall \varepsilon > 0, \lim\limits_{n\to+\infty} P([\lvert\overline{X}_n - m\rvert \geqslant \varepsilon]) = 0$. Illustrations. $\boxed{\blacktriangleright}$.

4 - Estimation

L'objectif de ce chapitre est, sans insister sur les aspects formels, de dégager la signification de la loi des grands nombres (approche fréquentiste) et de mettre en place la problématique de l'estimation. On introduit sur un exemple simple et concret (par exemple un sondage) cette problématique : on considère un phénomène aléatoire, qu'on a abstrait par une variable aléatoire réelle X dans une famille de lois *dépendant d'un paramètre inconnu θ (sur l'exemple du sondage, une loi de Bernoulli). Le problème de l'estimation consiste alors à déterminer une valeur approchée du paramètre θ à partir d'un échantillon de données $x_1, \ldots, x_n$ obtenues en observant n fois le phénomène.*

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_θ-indépendantes et de même loi que X pour tout θ. On pourra éventuellement introduire la notion d'estimateur, mais ce n'est pas un attendu du programme. Dans les cas considérés, le paramètre sera déterminé par la moyenne de la variable aléatoire. On s'appuie sur la loi faible des grands nombres pour justifier l'utilisation de l'estimateur $\overline{X}_n = \dfrac{X_1 + \ldots + X_n}{n}$ pour estimer l'espérance commune des variables aléatoires indépendantes X_i de même loi que X.

Echantillon.

a) Estimation ponctuelle.

La réalisation de $\overline{X}_n = \dfrac{X_1 + \ldots + X_n}{n}$ observée sur l'échantillon $x_1, \ldots, x_n$ est l'estimation du paramètre obtenue sur cet échantillon.

On donne les exemples de la loi de Bernoulli et de la loi de Poisson.

b) Estimation par intervalle de confiance.

La démarche consiste non plus à donner une estimation ponctuelle du paramètre θ mais à trouver un intervalle aléatoire, appelé intervalle de con

fiance, qui le contienne avec une probabilité minimale donnée. Ce paragraphe a uniquement pour but de préciser le vocabulaire employé. Les situations seront étudiées sous forme d'exercices dans des séances d'exercices et de travaux pratiques, aucune connaissance autre que ce vocabulaire n'est exigible sur les intervalles de confiance. On introduit l'intervalle de confiance obtenu à partir de l'inégalité de Bienaymé-Tchebychev. On en explique la signification. On remarque que la précision augmente avec la taille de l'échantillon. La démonstration n'est pas un attendu du programme.

Intervalle de confiance : La probabilité que l'intervalle $\left[\bar{X}_n - \sqrt{\frac{V(X)}{na}}, \bar{X}_n + \sqrt{\frac{V(X)}{na}}\right]$ contienne la moyenne $E(X)$ est supérieure à $1 - a$.

On se limitera au cas d'une variable de Bernoulli.
Résultat non exigible.
En pratique, la variance V est inconnue, mais on peut la majorer par $\frac{1}{4}$.
On particularise numériquement les intervalles de confiance au seuil de confiance de 90 % et de 95 %.

Intervalle de confiance de la moyenne d'une loi normale dont l'écart-type est connu.

On remarque que dans la pratique, l'écart-type n'est pas connu, ce qui conduit à utiliser l'écart-type de l'échantillon (écart-type empirique).

Enseignement annuel d'informatique et d'algorithmique

I - Éléments d'informatique et d'algorithmique

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

Le programme d'informatique s'articule autour de quatre thèmes : statistiques descriptives bivariées, études de suites et de fonctions, simulation de lois, estimation.

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.

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.

Pour certains thèmes, il sera nécessaire d'introduire de nouvelles notions mathématiques ; celles-ci seront introduites lors des séances d'informatique ; elles ne pourront en aucun cas être exigibles des étudiants, et toutes les précisions nécessaires seront données lors de leur utilisation.

Le langage informatique retenu pour la programmation dans ce programme des classes économiques et commerciales, option technologique, est Python.

Toute la richesse du langage Python ne peut pas être entièrement maîtrisée par un étudiant, aussi seules les fonctions et commandes exigibles du programme de première année sont exigibles, et leur syntaxe précise doit être rappelée. D'autres fonctions, par commodité, pourront être utilisées en classe, mais ceci ne pourra se faire qu'avec parcimonie. L'objectif principal de l'activité informatique reste la mise en pratique de connaissances mathématiques. Ces commandes supplémentaires devront être présentées en préambule et toutes les précisions nécessaires devront être 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.

1 - Liste des 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 : Porter un regard critique sur les méthodes d'estimation et de simulation.

C4 Stocker, organiser et extraire des données structurées volumineuses.

II - **Langage Python**

Les commandes exigibles ont été listées dans le programme de première année.

III - **Liste des thèmes**

1 - Statistiques descriptives bivariées

(Durée indicative : 4 heures. Compétences développées : **C1** et **C3**)

On s'appuiera sur les représentations graphiques pour montrer l'intérêt et les limites des indicateurs.

Série statistique à deux variables, nuage de points associé. Point moyen $(\bar{x}, \bar{y})$ du nuage.	On tracera le nuage de points et on pourra effectuer des pré-transformations pour se ramener au cas linéaire.
Covariance empirique, coefficient de corrélation empirique, droites de régression.	Analyse de deux caractères quantitatifs : covariance empirique, corrélation linéaire empirique, ajustement affine par la méthode des moindres carrés. On différenciera les variables explicatives des variables à expliquer et on soulignera la distinction entre corrélation et causalité. On pourra donner des exemples d'utilisation de la droite de régression pour faire des prévisions dans le cadre de problèmes concrets. On pourra utiliser les commandes : `plt.scatter, np.polyfit, np.corrcoef` ou un tableur

2 - Simulation de lois, application au calcul d'espérances

(Durée indicative : 4 heures. Compétences développées : **C1**, **C2**, et **C3**)

Ces simulations de variables aléatoires seront introduites comme illustrations de problèmes concrets, et permettront d'en vérifier la compréhension par les étudiants. 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. On pourra utiliser les générateurs de nombres aléatoires selon les lois uniformes, binomiales, géométriques, normales, de la bibliothèque `numpy.random` *:* `rd.random, rd.binomial, rd.randint, rd.geometric, rd.poisson, rd.exponential, rd.normal` .

Simulation de la loi uniforme sur $[0, 1]$; sur $[a, b]$.	
Simulation de phénomènes aléatoires à partir de lois usuelles. Méthodes de simulation d'une loi géométrique.	Comparaison entre différentes méthodes : utilisation d'une loi de Bernoulli et d'une boucle `while`, utilisation du générateur `rd.random`.
Simulation de lois usuelles.	

3 - Bases de données

(Durée indicative : 4 heures. Compétences développées : **C4**)

Dans la continuité du programme de première année, on poursuit l'étude du langage SQL avec la création de table et l'interrogation avancée via l'instruction JOIN. On introduira ces concepts à l'aide d'exemples simples issus de contextes appropriés.

a) Commandes exigibles

"SELECT* FROM nom_de_table_1 INNER JOIN nom_de_table_2".	Réalisation d'une jointure. On pourra ajouter une condition " ONΦ" dans le cas où Φ est une conjonction d'égalités. Aucune autre notion de jointure n'est dans ce programme.
"CREATE TABLE nom_de_table".	Création d'une table.

b) Commandes non exigibles

Les commandes non exigibles ont été listées dans le programme de première année.

4 - Théorème limite central

(Durée indicative : 4 heures. Compétences développées : **C1**, **C2**, et **C3**)

L'objectif est ici de dégager des conséquences importantes du théorème limite central qui n'est pas au programme. On met en œuvre sur des exemples ce théorème, qu'on pourra énoncer sans formaliser la notion de convergence. On souhaite dégager la pertinence de l'utilisation de la loi normale pour modéliser les phénomènes résultant de nombreux phénomènes aléatoires indépendants et de l'intervalle de confiance asymptotique dont on pourra mettre en valeur la précision.

Etude de la distribution des moyennes empiriques par simulation informatique de la loi de $X = \dfrac{Y_1 + \ldots + Y_n}{n}$ où $Y_1, \ldots, Y_n$ sont des variables aléatoires indépendantes suivant toutes une même loi discrète d'espérance μ et d'écart type σ.	On cherche à visualiser la convergence vers la loi normale d'espérance μ et d'écart type σ. On pourra produire un échantillon de taille N des moyennes empiriques d'échantillons de taille n d'une variable aléatoire d'espérance μ et d'écart type σ, et le représenter sous forme d'un histogramme. On observera l'effet de l'augmentation de n sur la dispersion des moyennes.
Simulation informatique de la loi de $X = \dfrac{Y_1 + \ldots + Y_n}{n}$ où $Y_1, \ldots, Y_n$ sont des variables aléatoires indépendantes suivant toutes la loi uniforme à densité sur $[0, 1]$.	On remarquera que la variable aléatoire centrée réduite associée à X est une approximation de la loi normale centrée réduite et on sensibilisera les étudiants au théorème limite central, en testant cette simulation avec d'autres lois.
Intervalle de confiance asymptotique.	On compare l'intervalle de confiance obtenu avec l'inégalité de Bienaymé-Tchebychev avec l'intervalle de confiance asymptotique, qu'on présentera en invoquant le théorème limite central pour estimer le paramètre d'une loi de Bernoulli.