

TRAVAUX PRATIQUES DE MATHÉMATIQUES AVEC SCILAB

En première année, les étudiants ont acquis les bases de manipulation du logiciel Scilab. L'objectif de l'enseignement d'informatique de seconde année est de leur permettre d'utiliser Scilab 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 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.

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, d'autres en groupes réduits.

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 du programme de 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, avec parcimonie, l'objectif principal de l'activité informatique restant la mise en pratique des 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 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 ou deux 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 :

cdfnorf, 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 issues de l'économie, du monde de l'entreprise ou de la finance, 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 bâton, 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**)

On introduira dans ce paragraphe la notion de produit scalaire dans \mathbf{R}^2 simplement dans l'objectif de l'étude de la droite de régression d'une série statistique à deux variables.

Notion de produit scalaire dans \mathbf{R}^2 . Orthogonalité. Série statistique à deux variables, nuage de points associé.
Point moyen (\bar{x}, \bar{y}) du nuage.
Covariance empirique, coefficient de corrélation empirique, 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 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 deux variables

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

Graphe d'une fonction de deux variables, lignes de niveau. Dérivées partielles, représentation du gradient.

Étude d'extrema locaux et globaux.

À cette occasion, on pourra mettre en évidence l'orthogonalité du gradient avec les lignes de niveau.

Programmation de fonctions variées permettant de mettre en évidence les notions d'extrema locaux ou globaux. On pourra prendre des exemples issus de l'économie ou de la finance : minimisation du risque, maximisation du profit...

5 - Simulation de lois

(Durée indicative : 6 heures. Compétences développées : **C1**, **C2** 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.

Simulation de la loi uniforme sur $[0, 1]$; sur $[a, b]$.

Utilisation du générateur `grand`.

Méthode d'inversion.

Application de la méthode d'inversion pour la simulation de la loi exponentielle de paramètre λ .

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 `rand`.

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.

6 - Estimation ponctuelle ou 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. On pense typiquement à des probabilités d'événements ou des 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 (cf. ci-dessous).

Comparaison de différents estimateurs ponctuels d'un paramètre.

On pourra utiliser des données issues de situations réelles (simple comparaison de valeurs numériques) ou créer plusieurs jeux de données par simulation grâce à la commande `rand`. Dans ce dernier cas, on pourra comparer les lois des estimateurs par exemple à l'aide d'histogrammes. Estimation par intervalle de confiance du paramètre d'une loi de Bernoulli et de l'espérance d'une loi normale.

Comparaison des intervalles de confiance d'un paramètre obtenus par différentes méthodes.

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