Le B.O. N°6 16 SEPT. 2004 HORS-SÉRIE

Annexe 1

VOIE SCIENTIFIQUE

MATHÉMATIQUES ET INFORMATIQUE :

Objectifs de formation

I. OBJECTIFS GÉNÉRAUX DE LA FORMATION

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 enfin certains points de terminologie. Toutefois chaque professeur conduit en toute liberté l'organisation de son enseignement et le choix de ses méthodes.

Comme rien ne nuit plus à l'esprit critique que les connaissances superficielles, le parti retenu par ces programmes est plutôt celui du « peu et bien » que d'une surcharge excessive qui entrainerait certains étudiants à faire semblant d'avoir compris.

On a voulu maintenir un volume global raisonnable : les limites du programme sont clairement précisées. Il convient de souligner les nécessités impérieuses de les respecter, aussi bien au niveau de l'enseignement que des épreuves d'évaluation. Il faut se garder de donner cours à toute ambition théorique concernant les points du programme précisés comme de premières approches. Les commentaires du programme permettent de dégager les objectifs essentiels, indiquent les résultats admis sans démonstration et comportent des mises en garde afin d'éviter des sujets trop théoriques ou présentant une technicité excessive.

II. ARCHITECTURE ET CONTENUS DES PROGRAMMES

Le programme de mathématiques des classes préparatoires économiques et commerciales est organisé sur deux années. Il fait suite à celui du baccalauréat et précède des enseignements spécialisés de calcul économique et de calcul de gestion. C'est dire qu'il prolonge les programmes de l'enseignement secondaire dans un même esprit de rigueur, en choisissant ses points forts dans l'orientation visée ; il ne s'agit donc ni d'un recueil de recettes utiles ni d'un cours sur des fondements de mathématiques générales.

Les mathématiques jouent un rôle important dans les sciences de l'économie et de la gestion, rôle encore renforcé par le développement de la micro-informatique. Le raisonnement sur les faits économiques s'accommodant mal de la structure des entiers, la démarche consiste à conceptualiser un phénomène en passant du discret au continu, puis à développer des algorithmes et des programmes informatiques qui ramènent le continu au discret. Ainsi en est-il de l'analyse des données statistiques, du contrôle du traitement et du mouvement des marchandises, de l'analyse des risques financiers et de la gestion des fichiers d'ordinateurs, tous domaines faisant appel à des méthodes de calcul élaborées.

L'orientation du programme vers les sciences de l'économie et de la gestion s'organise autour de cinq points forts :

- En algèbre linéaire, la méthode du pivot de Gauss, l'exploitation des structures euclidiennes, notamment pour les problèmes de moindres carrés, la réduction des matrices carrées (matrices stochastiques, matrices de covariance, étude d'extremums,...) ;
- en analyse, la mise en évidence des relations entre les phénomènes discrets, décrits par des suites, et les phénomènes continus, décrits par des fonctions, l'emploi de représentations graphiques pour l'étude qualitative et quantitative de ces phénomènes, la recherche d'extremums en liaison avec des problèmes simples d'optimisation, et la maîtrise des fonctions usuelles, notamment les fonctions exponentielles ;

- en probabilités et en statistique, la consolidation des acquis de enseignement secondaire, l'initiation aux phénomènes aléatoires, notamment l'emploi des lois usuelles, et une première approche du lien entre le modèle probabiliste et les séries statistiques ;
- une valorisation des aspects numériques et graphiques dans l'ensemble du programme ;
- en relation avec le programme d'infonnatique, l'étude de quelques algorithmes numériques (recherche et mise en forme d'algorithmes, comparaison de leurs performances).

Ces cinq points forts trouveront leurs prolongements dans la scolarité des écoles : analyse des données multidimensionnelles, programmation linéaire, optimisation, modèles économiques et financiers, processus aléatoires, informatique.

III. UTILISATION DES TICE

L'emploi des calculatrices est défini par la réglementation en vigueur. Les étudiants doivent savoir utiliser une calculatrice programmable dans les situations liées au programme de la classe considérée. Cet emploi combine les capacités suivantes, qui constituent un savoir-faire de base et sont seules exigibles :

- Savoir effectuer des opérations arithmétiques sur les nombres réels et savoir comparer deux nombres réels ;
- savoir utiliser les touches des fonctions qui figurent au programme de la classe considérée et savoir programmer le calcul des valeurs d'une fonction d'une ou plusieurs variables permis par ces touches ;
- savoir programmer une instruction séquentielle, une instruction conditionnelle et une instruction itérative comportant éventuellement un test d'arrêt.

Les notions d'algorithmique et de programmation figurant au programme d'infonnatique peuvent intervenir dans les épreuves écrites ou orales de mathématiques. Lors des épreuves écrites, les étudiants peuvent rédiger des algorithmes, soit en français, soit en Pascal (tout autre langage de programmation est exclu) ; l'énoncé peut imposer qu'un programme (ou certaines de ses procédures) soit rédigé en Pascal. Toute rédaction de programmes relatifs à l'algèbre linéaire est exclue. L'étude des algorithmes ne porte que sur des exemples ; toute théorie générale des algorithmes est hors programme.

Voie scientifique : deuxième année

I – Algèbre linéaire

L'objectif de ce chapitre est d'approfondir l'étude de la réduction des endomorphismes et des matrices carrées abordée en première année (en se limitant au cas diagonalisable) et d'exploiter les résultats obtenus pour le calcul matriciel.

Les espaces vectoriels considérés dans ce chapitre sont de dimension finie sur $\mathbf{K} = \mathbf{R}$ *ou* $\mathbf{C}$.

1) Sommes directes - Sous-espaces stables

Sous-espaces vectoriels supplémentaires, projecteurs associés. Caractérisation des projecteurs par la relation $p^2 = p$.	$E = \operatorname{Im} p \oplus \operatorname{Ker} p$
Somme directe d'une famille finie de sous-espaces vectoriels ; dimension d'une somme directe. Sous-espaces stables par un endomorphisme.	
Cas où l'espace est somme directe de sous-espaces stables.	Caractérisation matricielle.

2) Réduction des endomorphismes

a) Éléments propres d'un endomorphisme

Valeurs propres, vecteurs propres, sous-espaces propres d'un endomorphisme.	
Polynômes d'un endomorphisme.	$(aP+Q)(f) = a(P(f)) + Q(f)$ $(P \times Q)(f) = P(f) \circ Q(f)$ Si $f(x) = \lambda x$ alors $P(f)(x) = P(\lambda)x$
Polynômes annulateurs d'un endomorphisme.	Exemples des homothéties, des projecteurs et des symétries.
Si P est un polynôme annulateur de f et λ une valeur propre de f, alors λ est racine de P. Tout endomorphisme d'un espace de dimension finie admet au moins un polynôme annulateur non nul.	
Un endomorphisme d'un espace de dimension finie admet un nombre fini de valeurs propres et ses sous-espaces propres sont en somme directe.	Aucune autre connaissance sur les polynômes annulateurs ne figure au programme.

b) Endomorphisme diagonalisable

Un endomorphisme est diagonalisable si et seulement si l'espace est somme directe des sous-espaces propres.	
Caractérisation des endomorphismes diagonalisables à l'aide des dimensions des sous-espaces propres.	Si $\dim(E) = n$, tout endomorphisme de E admettant n valeurs propres distinctes est diagonalisable.

3) Réduction des matrices carrées

Changement de bases, matrice de passage.	Rappels.
Matrices semblables.	
Valeurs propres, vecteurs propres, sous-espaces propres d'une matrice carrée.	Valeurs propres des matrices triangulaires.
Polynômes d'une matrice. Polynômes annulateurs d'une matrice.	
Matrices diagonalisables, diagonalisation d'une matrice carrée.	Interprétation matricielle des résultats du paragraphe précédent. Application au calcul des puissances d'un endomorphisme ou d'une matrice.

II – Algèbre bilinéaire

Les espaces vectoriels considérés dans ce chapitre sont des espaces sur $\mathbf{R}$.

1) Produit scalaire

Produit scalaire, norme euclidienne sur un espace vectoriel réel.

Un produit scalaire est une forme bilinéaire symétrique, définie positive.
Produit scalaire canonique sur $\mathbf{R}^n$; exemples de produits scalaires.

Inégalité de Cauchy-Schwarz.

Cas de l'égalité.

Vecteurs orthogonaux, sous-espaces orthogonaux.

Familles orthogonales, familles orthonormales ou orthonormées.

On ne considèrera que des familles finies.
Toute famille orthogonale ne contenant pas le vecteur nul est libre.

Théorème de Pythagore.

2) Espaces euclidiens

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

Existence de bases orthonormées.

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

$$x = \sum_i < x, e_i > e_i \ , \ \|x\|^2 = \sum_i < x, e_i >^2.$$

Orthonormalisation de Schmidt, complétion d'une famille orthonormée en une base orthonormée.

Supplémentaire orthogonal d'un sous-espace vectoriel.

Notation $F^{\perp}$.

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

$$< x, y >= {}^tXY \ ; \ \|x\|^2 = {}^tXX$$

Changement de bases orthonormées.

La matrice de passage est orthogonale : $P^{-1} = {}^tP$.
Aucune autre connaissance sur les matrices orthogonales n'est au programme.

Projection orthogonale sur un sous-espace vectoriel.

Caractérisation par minimisation de la norme.

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

Application aux problèmes de moindres carrés :

- minimisation de $\|AX - B\|$

$A \in M_{n,p}(\mathbf{R})$ de rang p, $B \in M_{n,1}(\mathbf{R})$ et $X \in M_{p,1}(\mathbf{R})$

- ajustement affine d'une série statistique à deux variables par la méthode des moindres carrés, droites de régression.

3) 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.

Tout endomorphisme symétrique est diagonalisable et ses sous-espaces propres sont orthogonaux.
Toute matrice symétrique réelle est $\mathbf{R}$-diagonalisable.

Si A est symétrique réelle, il existe une matrice orthogonale P et une matrice diagonale D telles que $D = {}^tPAP$.
$A = \sum_i \lambda_i X_i {}^tX_i$ (où (X_i) est une base orthonormée de vecteurs-colonnes propres de A.)

Application à l'étude du signe d'une forme quadratique q sur $\mathbf{R}^n$, associée à un endomorphisme ou une matrice symétriques.

$\forall x \in \mathbf{R}^n$, $q(x) =< \varphi(x), x >= {}^tXAX$.
On ramènera cette étude à celle des valeurs propres de φ ou de A.

III–Intégrales sur un intervalle quelconque

Intégration sur un intervalle semi-ouvert.
Convergence de l'intégrale d'une fonction continue sur $[a, b[$ ($-\infty < a < b \leqslant +\infty$).

L'intégrale (dite *impropre*) est convergente si les primitives de la fonction à intégrer admettent une limite finie en b.
On pose alors $\displaystyle\int_a^b f(t)\,\mathrm{d}\,t = \lim_{x\to b}\int_a^x f(t)\,\mathrm{d}\,t$.
Convergence des intégrales $\displaystyle\int_1^{+\infty}\frac{\mathrm{d}\,t}{t^\alpha}, \int_a^b \frac{\mathrm{d}\,t}{(b-t)^\alpha}$.

Reste d'une intégrale convergente.
Règles de calcul sur les intégrales convergentes, linéarité, relation de Chasles, positivité, inégalités.

Cas des fonctions positives.

L'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}\,t$ converge si et seulement si $x \mapsto \displaystyle\int_a^x f(t)\,\mathrm{d}\,t$ est majorée sur $[a, b[$.

Règles de comparaison dans les cas $f \leqslant g$, $f = o(g)$ et $f \underset{b}{\sim} g$ avec f et g positives au voisinage de b.
Convergence absolue.
La convergence absolue implique la convergence.

On pourra utiliser une des décompositions :
$f = (f + |f|) - |f|$ ou $f = \max(f, 0) - \max(-f, 0)$.

Si $\displaystyle\int_a^b f(t)\,\mathrm{d}\,t$ est absolument convergente alors

$$\left|\int_a^b f(t)\,\mathrm{d}\,t\right| \leqslant \int_a^b |f(t)|\,\mathrm{d}\,t.$$

Si f est une fonction continue positive et décroissante sur $[a, +\infty[$, la série de terme général $f(n)$ est de même nature que l'intégrale $\displaystyle\int_a^{+\infty} f(t)\,\mathrm{d}\,t$.

Exemple des intégrales et des séries de Riemann.

Extension des notions précédentes aux intégrales sur un intervalle quelconque.

Brève extension aux fonctions définies et continues sur $]a_1, a_2[\cup]a_2, a_3[\cup\cdots\cup]a_{p-1}, a_p[$.

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

On effectuera l'intégration par parties sur un segment et on fera un passage à la limite.

Changement de variable :
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 $\displaystyle\int_a^b f(u)\,\mathrm{d}\,u$ et $\displaystyle\int_\alpha^\beta f(\varphi(t))\varphi'(t)\,\mathrm{d}\,t$ sont de même nature et en cas de convergence sont égales.

Énoncé analogue dans le cas où φ est décroissante.
Les changements de variables non affines devront être indiqués aux candidats.

Définition de la fonction Γ sur $]0, +\infty[$.

$\Gamma(x) = \displaystyle\int_0^{+\infty} t^{x-1}e^{-t}\,\mathrm{d}\,t$.
Relations $\Gamma(x+1) = x\Gamma(x)$ et $\Gamma(n+1) = n!$
Seules ces propriétés sont exigibles.

IV – Fonctions numériques de plusieurs variables

L'objectif est d'arriver à une bonne maîtrise des problèmes d'extremums à partir d'un minimum d'outils théoriques.

La détermination de la nature topologique d'un ensemble n'est pas un objectif du programme ; elle devra toujours être précisée. Néanmoins, on y sensibilisera les élèves à l'aide d'exemples d'ensembles, en particulier ceux qui sont définis par des inégalités du type $\{x \in \Omega / f(x) < (\text{ou} >) a\}$ où f est continue sur l'ouvert Ω (idem avec $\leqslant$, $\geqslant$ et F fermé).

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 candidats.

L'accent est mis sur la structure euclidienne de $\mathbf{R}^n$ et la paramétrisation des droites et des segments.

On s'appuiera sur les notions étudiées en première année dans le cas de deux variables. L'espace $\mathbf{R}^n$ sera muni de la norme euclidienne usuelle. On pourra utiliser les inégalités $\max_{i\in\{1,\dots,n\}} |x_i| \leqslant \|x\| \leqslant \sqrt{n} \max_{i\in\{1,\dots,n\}} |x_i|$.

1) Droites affines de $\mathbf{R}^n$ - Eléments de topologie

Droite $d_{A,U}$ passant par A, de vecteur directeur U.	Paramétrisation : $t \mapsto A + tU,\ t \in \mathbf{R}$.
Segment $[A, B]$.	$t \mapsto A + t(B - A) = (1 - t)A + tB,\ t \in [0, 1]$.
Boules, parties ouvertes.	Intersection finie, union quelconque d'ouverts.
Parties fermées.	Un fermé est le complémentaire d'un ouvert.
Parties bornées, parties convexes de $\mathbf{R}^n$.	Aucune difficulté théorique ne sera soulevée.

2) Fonctions définies sur $\mathbf{R}^n$

Équation du graphe d'une fonction définie dans $\mathbf{R}^n$.	Cas des fonctions affines de n variables.
Lignes et ensembles de niveau.	On se limitera à des exemples simples.
Continuité d'une application d'une partie de $\mathbf{R}^n$ dans $\mathbf{R}$. Opérations sur les fonctions continues. Composition par une fonction continue de $\mathbf{R}$ dans $\mathbf{R}$.	Les applications coordonnées et la norme sont continues sur $\mathbf{R}^n$. Aucune difficulté théorique ne sera soulevée.
Une fonction continue sur une partie fermée bornée admet un maximum et un minimum.	Résultat admis.

3) Calcul différentiel

Les fonctions sont désormais supposées définies sur des ouverts de $\mathbf{R}^n$.

a) Premier ordre

Fonctions partielles en un point. Dérivées partielles d'ordre 1.	Notation $\dfrac{\partial f}{\partial x_i}(A)$.
Gradient en un point A.	Notation $\nabla f(A)$ ou ∇f_A.
Approximation locale par une fonction affine. Développement limité d'ordre 1 en un point d'une fonction de n variables. Fonctions de classe C^1 sur un ouvert. Opérations sur les fonctions de classe C^1. Composition par une fonction de $\mathbf{R}$ dans $\mathbf{R}$ de classe C^1.	La notion de différentielle est hors programme.
Existence d'un développement limité d'ordre 1 en un point pour une fonction de classe C^1 sur un ouvert. Fonction affine tangente.	$f(A+H) = f(A) + \langle \nabla f_A, H \rangle + \|H\|\varepsilon(H)$ où $\varepsilon(0) = 0$ et ε continue en 0. Résultat admis. Hyperplan affine tangent au graphe.
Dérivée d'une fonction de la forme $t \mapsto f \circ u(t)$ dans le cas où $u(t) = A + tU$, dérivée directionnelle.	On montrera que la dérivée en A dans la direction U est $\langle \nabla f_A, U \rangle$ et on en déduira une interprétation du gradient.
Accroissements finis.	Si f est de classe C^1 sur Ω et si $[A, A+H] \subset \Omega$ il existe $\theta \in]0,1[$ tel que $f(A+H) = f(A) + \langle \nabla f_{A+\theta H}, H \rangle$

b) Deuxième ordre

Dérivées partielles d'ordre 2.	Notations $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(A)$ et $\dfrac{\partial^2 f}{\partial x_i^2}(A)$.
Fonctions de classe C^2. Opérations sur les fonctions de classe C^2. Composition par une fonction de $\mathbf{R}$ dans $\mathbf{R}$ de classe C^2.	

Hessienne en un point A.

Notation $\nabla^2 f(A)$ ou $\nabla^2 f_A$

Théorème de Schwarz.

Ce résultat pourra être admis.

Développement limité d'ordre 2 d'une fonction de classe C^2.

$f(A+H) = f(A) + <\nabla f_A, H> + \frac{1}{2}q_A(H) + \|H\|^2\varepsilon(H)$ où $\varepsilon(0) = 0$, ε continue en 0 et q_A est la forme quadratique associée à la hessienne en A.
Ce résultat pourra être admis.

Dérivée seconde directionnelle en A.

La dérivée seconde en A dans la direction U est $q_A(U)$.

Égalité de Taylor-Lagrange à l'ordre 1.

Si f est de classe C^2 sur Ω et si $[A, A+H] \subset \Omega$ il existe $\theta \in]0, 1[$ tel que

$$f(A+H) = f(A) + <\nabla f_A, H> + \frac{1}{2}q_{A+\theta H}(H)$$

4) Extremums

a) Extremums sur un ouvert

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

Condition nécessaire du premier ordre : si une fonction de classe C^1 sur un ouvert admet un extremum local en un point, son gradient est nul en ce point.

Les points où le gradient s'annule sont appelés points critiques. Toutes les dérivées directionnelles en ces points sont nulles.

Condition suffisante d'extremum local pour une fonction f de classe C^2 sur un ouvert Ω : f admet un minimum (resp. maximun) en un point critique A si la forme quadratique q_A associée à la hessienne en A prend des valeurs strictement positives (resp. négatives) sur les vecteurs non nuls.
Point selle (ou col).

Notations de Monge et étude des cas $rt - s^2 > 0$, $rt - s^2 < 0$, $rt - s^2 = 0$ dans le cas $n = 2$.
Le signe de q_A pourra être déterminé à partir du signe des valeurs propres de la hessienne en A ou, sur des exemples, par une décomposition en carrés. La théorie de la décomposition de Gauss n'est pas au programme.

Exemples de recherche d'extremums globaux.
Exemple de recherche de la position du graphe d'une fonction de classe C^2 par rapport à l'hyperplan tangent.

On pourra faire une étude directe du signe sur Ω de $f(X) - f(A)$. Dans les situations qui s'y prêtent, on pourra étudier le cas où pour pour tout B de Ω, q_B est positive (resp. négative).

b) Recherche d'extremums sous contrainte d'égalités linéaires.

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

Condition nécessaire du premier ordre : si f est une fonction de classe C^1 sur un ouvert Ω, et si la restriction de f à $\mathcal{C}$ admet un extremum local en un point A, son gradient en A est orthogonal à $\mathcal{H}$.

On remarquera que $\mathcal{H}^\perp = \mathrm{vect}(\nabla g_i,\ 1 \leqslant i \leqslant p)$.
Pour tout H de $\mathcal{H}$, la dérivée de f en A dans la direction H est nulle.
Point critique pour l'optimisation sous contrainte.

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

On pourra faire une étude directe du signe sur Ω de $f(X) - f(A)$. On pourra dans les situations qui s'y prêtent, étudier le cas où pour tout B de $\mathcal{C}$ et tout H de $\mathcal{H}$ on a $q_B(H) \geqslant 0$ (resp. $q_B(H) \leqslant 0$).

V – Statistique descriptive bivariée

$\mathbf{R}^n$ sera muni du produit scalaire canonique. Conformément à l'usage, on notera $\overline{u}$ la moyenne arithmétique des coordonnées de tout élément u de $\mathbf{R}^n$.

Observation conjointe de deux caractères X et Y sur un échantillon de taille n de la population.
Cas de deux caractères quantitatifs.

L'observation est un n-uplet $((x_1, y_1), (x_2, y_2), \dots (x_n, y_n))$ d'éléments de $\mathbf{R}^2$.
On peut aussi la représenter par le couple (x, y) de vecteurs de $\mathbf{R}^n$ définis par $x = (x_1, x_2, \dots, x_n)$ et $y = (y_1, y_2, ..., y_n)$.

Nuage de points de $\mathbf{R}^2$ associé.
Point moyen $(\overline{x}, \overline{y})$ du nuage.

Groupement de termes et description d'une série statistique double : tableau des effectifs ou des fréquences. Fréquences marginales, fréquences conditionnelles.

Caractéristiques d'une série statistique double : covariance, coefficient de corrélation.

On rappelle que ces paramètres empiriques sont calculés à partir de l'échantillon observé.

Droites de régression. Interprétation géométrique dans le plan $\mathbf{R}^2$.

On montrera le lien entre la recherche de la droite de régression de Y en X et la projection orthogonale, dans $\mathbf{R}^n$, de y sur le sous-espace vectoriel engendré par $(1,1,\ldots 1)$ et x.

VI – Probabilités

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.

Enfin, on établira 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 séries doubles absolument convergentes.

Convergence des séries doubles à termes positifs.

Si pour tout i, $\sum_j u_{i,j}$ converge vers un réel a_i et si $\sum_i a_i$ converge vers S, alors pour tout j, $\sum_i u_{i,j}$ converge vers un réel b_j et $\sum_j b_j$ converge vers S.

Théorème de comparaison dans le cas où, pour tous i, j, $0 \leqslant u_{i,j} \leqslant v_{i,j}$

Série double absolument convergente.

Toute série double absolument convergente est la différence de deux séries doubles à termes positifs convergentes (Cf. séries simples).

Propriété de Fubini, sommation par paquets.

On pourra utiliser la notation $\sum_{i,j} u_{i,j}$.

Aucun de ces résultats n'est exigible des candidats et tout exercice ou problème y faisant appel devra impérativement les rappeler.

Pour l'étude des lois à densité usuelles, on pourra mettre en évidence l'intérêt de l'utilisation des transformations affines pour diminuer le nombre de paramètres.

1) Variables aléatoires discrètes

a) Simulations informatiques

En liaison avec les activités informatiques, on établira des fonctions PASCAL pour simuler des variables aléatoires suivant les lois usuelles suivantes : loi uniforme, loi de Bernoulli, loi binomiale, loi géométrique.

Cette activité sera l'occasion de revoir les lois usuelles discrètes étudiées en première année.

On pourra étudier le problème plus délicat de la simulation d'une variable suivant une loi hypergéométrique.

On étudiera en exercice des simulations de variables aléatoires autres que celles considérées comme usuelles.

b) Espérance et conditionnement

Si X est une variable aléatoire discrète définie sur $(\Omega, \mathcal{A}, P)$ et si (A_n) est un système complet d'événements, X admet une espérance pour P si et seulement si elle admet une espérance pour toutes les P_{A_n} (où $P(A_n) \neq 0$) et

$$E(X) = \sum_{n, P(A_n)\neq 0} E(X|A_n)P(A_n)$$

Formule de l'espérance totale.

$E(X|A_n)$ est l'espérance de X pour la probabilité conditionnelle P_{A_n}.

2) Vecteurs aléatoires discrets

a) Compléments sur les couples de variables aléatoires réelles discrètes

Ce paragraphe sera l'occasion d'un rappel des propriétés des couples étudiées en première année.

Système complet et σ-algèbre $\mathcal{A}_{(X,Y)}$ associés au couple (X,Y).

Loi d'un couple, lois marginales, lois conditionnelles.

Loi d'une variable aléatoire $Z = g(X,Y)$ où g est une fonction définie sur l'ensemble $(X,Y)(\Omega)$.

Loi de la somme de deux variables indépendantes, produit de convolution.

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

Stabilité des ensembles des lois binomiales et de Poisson.

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

Sous réserve de convergence absolue :

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

Linéarité et croissance de l'espérance.

Variance de la somme de deux variables aléatoires discrètes, covariance et coefficient de corrélation linéaire, interprétation.

Notations $\mathrm{Cov}(X,Y)$ et ρ_{XY}.

Cas $\rho_{XY} = \pm 1$.

Formule de Huygens et calcul de la covariance.

Deux variables discrètes indépendantes admettant un moment d'ordre 2 sont non corrélées.

b) Vecteur aléatoire discret à valeurs dans $\mathbf{R}^n$.

Extension des notions du paragraphe précédent au cas des vecteurs discrets : système complet et σ-algèbre associés, théorème de transfert,...

Matrice de variance-covariance.

Indépendance mutuelle de n variables aléatoires réelles discrètes.

$P\left(\bigcap_{i=1}^{n}[X_i = x_i]\right) = \prod_{i=1}^{n} P([X_i = x_i])$ pour tout $(x_1,..,x_n)\in\mathbf{R}^n$.

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

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

Si X_1, X_2 ,..., X_n sont mutuellement indépendantes les σ-algèbres associées à $(X_1, X_2, \ldots, X_p)$ et à $(X_{p+1}, X_{p+2}, \ldots, X_n)$ sont indépendantes.

Les variables aléatoires $\varphi(X_1,\ldots,X_p)$ et $\psi(X_{p+1},\ldots,X_n)$ sont indépendantes (résultat admis).

Variance de la somme de n variables aléatoires.

Variance de la somme dans le cas de l'indépendance.

À titre d'exercice, on pourra calculer la variance dans le cas de la loi hypergéométrique.

3) Généralités sur les variables aléatoires réelles

On dit que X est une variable aléatoire définie sur $(\Omega,\mathcal{A})$ si X est une application de Ω dans $\mathbf{R}$ telle que pour tout intervalle I de $\mathbf{R}$, $X^{-1}(I)\in\mathcal{A}$.

Notations $[X\in I]$, $[X\leqslant x]$, *etc.*

σ-algèbre $\mathcal{B}$ des boréliens, σ-algèbre associée à une variable aléatoire.

Aucun développement théorique sur la tribu des boréliens n'est au programme.
On admettra que pour tout borélien B, $[X \in B]$ appartient à $\mathcal{A}$.

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

Résultat admis.

Fonction de répartition d'une variable aléatoire : $F_X(x) = P(X \leqslant x)$. Propriétés.

F_X est croissante et continue à droite, $\lim\limits_{-\infty} F_X = 0$, $\lim\limits_{+\infty} F_X = 1$.

Indépendance deux à deux et indépendance mutuelle d'un ensemble fini ou d'une suite de variables aléatoires.

Deux variables aléatoires X et Y sont indépendantes si et seulement si pour tout $(t, t') \in \mathbf{R}^2$,
$P([X \leqslant t] \cap [Y \leqslant t']) = P([X \leqslant t])P([Y \leqslant t'])$

4) Variables aléatoires à densité

a) Définition

On dit qu'une variable aléatoire X est à densité si sa fonction de répartition F_X est continue sur $\mathbf{R}$ et de classe C^1 sur $\mathbf{R}$ privé d'un ensemble fini de points.
Toute fonction f_X à valeurs positives, qui ne diffère de F'_X qu'en un nombre fini de points, est une densité de X.

Pour tout x de $\mathbf{R}$, $F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$.
Toute fonction f positive, continue sur $\mathbf{R}$ privé d'un nombre fini de points et telle que $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$ est la densité d'une variable aléatoire.

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

Les candidats devront savoir retrouver les densités de $aX + b$ $(a \neq 0)$, X^2 et $\varphi(X)$ avec φ de classe C^1 strictement monotone sur $X(\Omega)$.

b) Moments d'une variable aléatoire à densité.

Espérance d'une variable aléatoire à densité.

$E(X) = \displaystyle\int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t$ si cette intégrale converge.

Positivité de l'espérance.

Théorème de transfert lorsque $X(\Omega)$ est un intervalle I d'extrémités a et b $(-\infty \leqslant a < b \leqslant +\infty)$.

Si φ est définie et continue sur I éventuellement privé d'un nombre fini de points, $E(\varphi(X))$ existe et est égale à $\displaystyle\int_a^b \varphi(t) f_X(t)\,\mathrm{d}t$ si et seulement si cette intégrale converge absolument.
On pourra le démontrer dans le cas où φ est de classe C^1, avec φ' strictement positive (ou strictement négative) et le vérifier dans des cas simples.
Cette démonstration n'est pas exigible.

Moments d'ordre 2, variance, écart-type.
Variables aléatoires centrées, centrées réduites.

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

c) Somme de variables aléatoires à densité

Les résultats de ce paragraphe seront admis. En cas d'utilisation du produit de convolution, la preuve de sa légitimité n'est pas exigible des candidats.

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

Si la fonction h définie par la relation $h(x) = \displaystyle\int_{-\infty}^{\infty} f_X(t) f_Y(x - t)\,\mathrm{d}t$ est définie et continue sauf peut-être en un nombre fini de points, c'est une densité de Z.
C'est le cas si f_X (ou f_Y) est bornée.

d) Somme de variables aléatoires quelconques

Les résultats de ce paragraphe seront admis.

Espérance d'une somme de variables aléatoires. Croissance de l'espérance.

Si X et Y admettent une espérance, $X + Y$ admet une espérance et $E(X + Y) = E(X) + E(Y)$.
Si de plus $X \leqslant Y$ (ps) alors $E(X) \leqslant E(Y)$.

Variance d'une somme de variables aléatoires indépendantes.

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

e) Lois usuelles

Loi uniforme sur un intervalle, espérance et variance.

On pourra utiliser la fonction `random` du générateur aléatoire PASCAL pour illustrer cette loi.

$$X \hookrightarrow \mathcal{U}[0,1] \Longleftrightarrow Y = a + (b-a)X \hookrightarrow \mathcal{U}[a,b]$$

Loi exponentielle, caractérisation par l'absence de mémoire, espérance et variance.

$$X \hookrightarrow \mathcal{E}(1) \Longleftrightarrow Y = \frac{1}{\lambda}X \hookrightarrow \mathcal{E}(\lambda) \quad (\lambda > 0)$$

Lois γ et Γ. Espérance et variance.
Stabilité pour la somme.
Loi de la somme de n variables exponentielles indépendantes et de même paramètre.

$X \hookrightarrow \gamma(\nu)$, $\nu > 0$, si

$$\begin{cases} X(\Omega) = \mathbf{R}^+ \\ f_X(t) = \dfrac{1}{\Gamma(\nu)} t^{\nu-1} e^{-t} \quad \text{si } t > 0. \end{cases}$$

$$X \hookrightarrow \gamma(\nu) \Longleftrightarrow Y = bX \hookrightarrow \Gamma(b,\nu) \quad (b > 0)$$

Loi normale centrée réduite, loi normale (ou de Laplace-Gauss), espérance, variance.
Transformées affines de lois normales.
Stabilité pour la somme.

$$X \hookrightarrow \mathcal{N}(0,1) \Longleftrightarrow Y = \sigma X + m \hookrightarrow \mathcal{N}(m,\sigma^2),\ (\sigma > 0)$$

On attend des élèves qu'ils sachent représenter graphiquement les fonctions densités des lois normales et utiliser la fonction de répartition Φ de la loi normale centrée réduite.

5) Convergences et approximations

a) Convergence en probabilité

Inégalités de Markov et de Bienaymé-Tchebychev.

Si X est une variable aléatoire discrète ou à densité admettant un moment d'ordre 2, alors

$$P([|X| \geqslant \varepsilon]) \leqslant \frac{E(X^2)}{\varepsilon^2} \quad \text{d'où}$$

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

Convergence en probabilité : si (X_n) et X sont des variables aléatoires définies sur $(\Omega, \mathcal{A}, P)$, (X_n) converge en probabilité vers X si, pour tout $\varepsilon > 0$, $\lim\limits_{n\to\infty} P([|X_n - X| > \varepsilon]) = 0$.

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.

La loi faible des grands nombres permet une justification partielle, a posteriori, de la notion de probabilité d'un événement, introduite intuitivement.

b) Convergence en loi

Définition de la convergence en loi d'une suite $(X_n)_{n\in\mathbf{N}}$ de variables aléatoires vers X.
Cas où les X_n prennent leurs valeurs dans $\mathbf{N}$.
Convergence d'une suite de variables aléatoires suivant la loi binomiale $\mathcal{B}(n, \lambda/n)$ vers une variable suivant la loi de Poisson de paramètre λ.

Théorème de la limite centrée : étant données des variables aléatoires X_n indépendantes et de même loi, admettant une espérance et une variance, la variable aléatoire centrée réduite S_n^* associée à

$$S_n = \sum_{k=1}^{n} X_k$$

converge en loi vers la loi normale centrée réduite, c'est-à-dire :

$$\lim_{n\to+\infty} P([a \leqslant S_n^* \leqslant b]) = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

Résultat admis.
On utilisera ce résultat pour justifier l'approximation d'une loi binomiale et d'une loi de Poisson, dans certains cas, par une loi normale.

On étudiera en exercice le principe de la simulation informatique d'une loi normale en utilisant, par exemple, la variable centrée réduite associée à $X = \frac{1}{n}\sum_{k=1}^{n} Y_k$ où les Y_k sont des variables aléatoires indépendantes suivant toutes une loi uniforme à densité sur $[0, 1]$.

c) Approximations

Exemples d'approximation (cas de la loi hypergéométrique, de la loi binomiale, et de la loi de Poisson).

Toutes indications devront être fournies aux candidats quant à la justification de l'utilisation des approximations.

6) Estimation

L'objectif est d'initier les étudiants à la démarche de la statistique inférentielle en présentant le problème de l'estimation, ponctuelle ou par intervalle de confiance, de certaines valeurs caractéristiques de la loi de probabilité régissant l'observation d'un nombre réel associé à une expérience aléatoire reproductible dans des conditions identiques et indépendantes.

On se placera dans le cas où cette loi, non entièrement déterminée, appartient à une famille $(\mu_\theta)_{\theta\in\Theta}$ de lois définies sur une même partie $\mathcal{X}$ de $\mathbf{R}$ et dépendant d'un paramètre θ (scalaire ou vectoriel) appartenant à un sous-ensemble Θ de $\mathbf{R}^k$.

On limitera le problème à l'estimation d'un réel de la forme $g(\theta)$ où g désignera une application de $\Theta \subset \mathbf{R}^k$ dans $\mathbf{R}$. Bien entendu, pour tout θ de Θ, $g(\theta)$ représentera une valeur caractéristique de la loi μ_θ telle que son espérance, sa variance, son étendue...

Pour tout θ de Θ, l'ensemble Ω des suites $(x_n)_{n\geqslant 1}$ à valeurs dans $\mathcal{X}$ sera supposé muni d'une loi notée P_θ telle que pour tout $k \geqslant 1$, $X_k : (x_n)_{n\geqslant 1} \mapsto x_k$ soit une variable aléatoire de loi μ_θ et que les $(X_n)_{n\geqslant 1}$ soient mutuellement P_θ-indépendantes. Pour tout n de $\mathbf{N}^*$, on dit alors que $(X_1, ..., X_n)$ est pour tout θ de Θ un n-échantillon indépendant, identiquement distribué (en abrégé i.i.d) de la loi μ_θ.

On appelle statistique toute variable aléatoire réelle de la forme $\varphi_n(X_1, X_2, \ldots, X_n)$ où φ_n est une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$.

L'objectif de l'estimation est de localiser $g(\theta)$ grâce à l'unique donnée d'un échantillon $x_1, ..., x_n$ de réalisations des variables $X_1, X_2, \ldots, X_n$ obtenues après avoir observé n fois le phénomène dans des conditions identiques et indépendantes.

Toutes les notions qui suivent devront être abondamment illustrées à l'aide de simulations informatiques.

a) Estimation ponctuelle.

Estimer ponctuellement $g(\theta)$ par $\varphi_n(x_1, ..., x_n)$ où $\varphi_n(X_1, X_2, \ldots, X_n)$ est une statistique et $(x_1, ..., x_n)$ est la réalisation de l' échantillon $(X_1, ..., X_n)$ de la loi μ_θ, c'est décider d'accorder à $g(\theta)$ la valeur $\varphi_n(x_1, ..., x_n)$.

Un estimateur de $g(\theta)$ est une statistique $T_n = \varphi_n(X_1, ..., X_n)$.

La réalisation $\varphi_n(x_1, ..., x_n)$ de la variable aléatoire T_n est l'estimation de $g(\theta)$.
Cette estimation ne dépend que de l'échantillon $(x_1, x_2, ..., x_n)$ observé.

Exemples simples d'estimateurs.

Exemple de la moyenne empirique $\dfrac{X_1 + X_2 + \cdots + X_n}{n}$ lorsque μ_θ admet une espérance pour tout θ de Θ.

Biais d'un estimateur de $g(\theta)$: si pour tout θ de Θ, T_n admet une espérance, on appelle biais de T_n en θ le réel $E_\theta(T_n) - g(\theta)$.
Estimateur sans biais.

$E_\theta(T_n)$ désigne l'espérance de T_n pour la loi P_θ.
On appellera biais de T_n, la fonction b_{T_n} de Θ dans $\mathbf{R}$ définie par $b_{T_n}(\theta) = E_\theta(T_n) - g(\theta)$.

Risque quadratique d'un estimateur : si pour tout θ de Θ, T_n admet un moment d'ordre 2, le risque quadratique de T_n en θ est le réel $E_\theta\big((T_n - g(\theta))^2\big)$.

On appellera risque quadratique de T_n la fonction qui à tout θ de Θ associe le réel

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

On pourra de même appeler variance de T_n la fonction définie par $\forall \theta \in \Theta$, $V_\theta(T_n) = E_\theta\big((T_n - E_\theta(T_n))^2\big)$.

Le risque quadratique d'un estimateur est la somme de sa variance et du carré de son biais.

$$r_{T_n}(\theta) = b^2_{T_n}(\theta) + V_\theta(T_n).$$

Quand un estimateur est sans biais, son risque quadratique est sa variance.

Si T_n est sans biais, $r_{T_n}(\theta) = V_\theta(T_n)$.

Suite $(T_n)_{n \geqslant 1}$ d'estimateurs de $g(\theta)$.

Chaque T_n est de la forme $\varphi_n(X_1, X_2, \ldots, X_n)$.

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

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

Une suite d'estimateurs $(T_n)_{n\geqslant 1}$ de $g(\theta)$ est convergente si pour tout θ, la suite $(T_n)_{n\geqslant 1}$ converge en probabilité vers $g(\theta)$.
Si les μ_θ admettent une variance, la moyenne empirique est un estimateur sans biais et convergent de l'espérance de μ_θ.

Cette convergence pourra être étudiée à l'aide des inégalités de Bienaymé-Tchebychev ou de Markov.
Par abus de langage on dit aussi que l'estimateur est convergent (ou consistant).
Toutes les notions de ce paragraphe seront illustrées à l'aide de simulations informatiques.

c) Estimation par intervalle de confiance

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.

Dans tout ce paragraphe U_n et V_n désigneront des statistiques $\gamma_n(X_1, ..., X_n)$ et $\delta_n(X_1, ..., X_n)$.

$[U_n, V_n]$ est un intervalle de confiance de $g(\theta)$ au niveau de confiance $1 - \alpha$ si pour tout θ de Θ,
$P_\theta(U_n \leqslant g(\theta) \leqslant V_n) \geqslant 1 - \alpha$.

Sa réalisation est l'estimation de cet intervalle de confiance.
α est appelé le risque.
On distinguera probabilité et confiance et on éclairera la notion de risque à l'aide de simulations informatiques.

Estimation par intervalle du paramètre d'une variable de Bernoulli.
Estimation par intervalle de l'espérance d'une loi normale d'écart-type donné.

On pourra comparer, en majorant $p(1-p)$ par $1/4$, les intervalles de confiance obtenus par l'inégalité de Bienaymé-Tchebychev et par l'approximation normale de la loi binomiale.
Pour de grands échantillons, on pourra utiliser le théorème de la limite centrée pour donner un intervalle de confiance de l'espérance d'une variable de variance bornée.

VII – Éléments d'algorithmique

L'objectif est double :
– d'une part, consolider les acquis de première année concernant l'utilisation du langage PASCAL, notamment dans l'utilisation de fonctions et de procédures ;
– d'autre part, enrichir le champ des algorithmes rencontrés par l'étude de listes (en fait, ici, des tableaux à une dimension) et des situations aléatoires plus complexes.

1) L'environnement Pascal

a) Procédures et fonctions
Déclaration de procédures et de fonctions, structure, passage de paramètres en valeur et en variable. Notion de variables locales et globales.

b) Procédures et fonctions récursives.
Principe et utilisation.

2) Listes des savoir-faire supplémentaires exigibles en deuxième année

Écrire (ou comprendre) des outils de calcul élémentaire utilisant des algorithmes itératifs ou récursifs. Déterminer le nombre d'opérations élémentaires effectuées dans l'exécution de ces algorithmes.

On pourra tirer les exemples d'algorithmes mathématiques déjà étudiés en première année :
n^p, $n!$, $\binom{n}{p}$, suites récurrentes...
On écrira ces outils sous forme de fonctions Pascal.
On montrera le piège des appels récursifs multiples.

Écrire ou comprendre des algorithmes de calcul issus du cours d'analyse :
- calcul approché de la valeur d'une intégrale ;
- calculs utilisant les suites récurrentes du type $u_{n+1} = f(u_n)$.

Écrire des fonctions PASCAL, utilisant le générateur aléatoire PASCAL, simulant des variables aléatoires suivant les lois suivantes :
loi uniforme sur $[\![n_1, n_2]\!]$, loi de Bernoulli, loi binomiale, loi géométrique, loi uniforme à densité sur $[a, b]$.

On pourra étudier en exercice la simulation de la loi hypergéométrique (un algorithme récursif permet une écriture simple).

Utiliser les outils précédents, ou des algorithmes directs, pour écrire des fonctions informatiques simulant des variables aléatoires plus difficiles à étudier.

Exemples : somme de n variables aléatoires indépendantes suivant une loi uniforme, variable définie par un conditionnement, somme de variables aléatoires non indépendantes, maximum de plusieurs variables...

Utiliser ces modèles pour réaliser l'estimation de paramètres de phénomènes aléatoires.

Estimation de l'espérance d'une variable aléatoire.

Écrire ou comprendre des outils élémentaires de gestion de listes définies par un tableau à une dimension.
On étudiera en particulier :
- la recherche de la valeur et du rang des extremums de cette liste,
- la recherche dichotomique d'un élément dans une liste ordonnée.