

Classe préparatoire B/L

Programme de mathématiques

Table des matières

Objectifs généraux de la formation	3
1 - Compétences développées	3
2 - Architecture des programmes	3
 Généralités	 5
1 - Logique	5
2 - Vocabulaire ensembliste	5
3 - Les nombres entiers	5
4 - La droite réelle	5
5 - Le plan complexe	6
 Première année	 6
I - Suites et séries de nombres réels	6
 II - Algèbre linéaire	 7
1 - L'espace \mathbf{R}^n	7
2 - Matrices et systèmes linéaires	7
3 - Matrices carrées inversibles	8
4 - Sous-espaces vectoriels de \mathbf{R}^n	8
5 - Applications linéaires entre sous-espaces vectoriels de \mathbf{R}^n	9
6 - Rang d'une matrice	9
7 - Espaces vectoriels	10
 III - Fonctions d'une variable réelle	 10
1 - Limites et continuité	10
2 - Dérivées	11
3 - Exemple d'étude de fonction : régression linéaire	12
4 - Intégration	12

IV - Probabilités	13
1 - Événements aléatoires	13
2 - Variables aléatoires discrètes	14
3 - Moments des variables aléatoires discrètes réelles positives	14
4 - Indépendance	15
5 - Processus de Bernoulli	16
Deuxième année	17
I - Algèbre et géométrie	17
1 - Somme directe, supplémentaire	17
2 - Valeurs propres des endomorphismes	18
3 - Produit scalaire	18
II - Étude locale des fonctions d'une variable réelle	19
1 - Fonctions polynomiales	19
2 - Développements limités	19
3 - Intégrales généralisées	20
III - Fonctions de deux variables réelles	20
1 - Exemples	20
2 - Dérivées partielles	21
3 - Fonctions quadratiques	21
4 - Retour sur la régression linéaire	21
5 - Étude des points critiques	22
IV - Probabilités	22
1 - Variables aléatoires à densité	22
2 - Loi normale, loi exponentielle	23
3 - Indépendance de variables à densité	23
4 - Statistiques	24

Objectifs généraux de la formation

Les mathématiques jouent un rôle important dans la société et ont une importance grandissante dans les sciences humaines et 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 ce programme est de permettre de manière équilibrée

- une formation par les mathématiques en tant que telles ;
- l'acquisition d'outils utiles notamment aux sciences sociales et à l'économie (probabilités et statistiques, introduction aux fonctions de deux variables par exemple).

L'objectif de la formation dans les classes préparatoires B/L n'est pas de former des professionnels des mathématiques. L'enseignement de mathématiques concourt à structurer la pensée des étudiants, à développer leurs capacités d'imagination et d'abstraction, et à les former à la rigueur et à la logique en insistant sur les divers types de raisonnement (par équivalence, implication, l'absurde, analyse-synthèse...). Il permet aux étudiants d'utiliser des outils mathématiques ou d'en comprendre l'usage dans diverses situations de leur parcours académique et professionnel. L'état de l'art en sciences sociales et économie a été un guide important pour donner aux étudiants de B/L les bases dont ils auront besoin pour aller plus loin.

Le programme définit les objectifs de l'enseignement de ces classes et décrit les connaissances et les capacités exigibles des étudiants. Il précise 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.

1 - Compétences développées

L'enseignement de mathématiques en classes préparatoires B/L permet de développer chez les étudiants les compétences générales 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 notamment issus de problèmes de sciences sociales ou économiques) 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.
- **Communiquer par écrit et oralement** : comprendre les énoncés mathématiques, savoir rédiger une solution rigoureuse, présenter une production mathématique.

2 - Architecture des programmes

Par rapport au programme précédent, le programme d'algèbre linéaire donne une place plus importante aux aspects matriciels et introduit des bases de géométrie euclidienne qui pourront être illustrées par

d'autres parties du programme. Il est important de mettre en valeur l'interaction entre les différentes parties du programme. À titre d'exemple, l'algèbre linéaire trouvera ainsi son application dans les problèmes d'optimisation, l'analyse et les probabilités dans les problèmes d'estimation.

Le programme a été rédigé sur deux années. Au sein de chaque année, aucun ordre particulier n'est imposé et chaque professeur conduit en toute liberté l'organisation de son enseignement. Le programme tient compte de l'évolution des programmes de Terminale tout en maintenant une exigence intellectuelle élevée adaptée aux étudiants de la filière B/L et à la place que prennent aujourd'hui les techniques quantitatives en sciences humaines et sociales.

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 et des exemples d'activités ou d'applications.

Généralités

Concernant cette partie, le vocabulaire doit être connu et un savoir-faire est attendu. Aucune difficulté théorique ne sera soulevée. Certaines des notions peuvent être introduites en situation sans faire l'objet de chapitres spécifiques.

1 - Logique

Connecteurs : et, ou, non, implication, réciproque, contraposée.

Quantificateurs \exists, \forall .

Raisonnement par l'absurde.

Notion de condition nécessaire et de condition suffisante.

Négation d'une phrase mathématique utilisant connecteurs et quantificateurs.

Introduit sur des exemples.

2 - Vocabulaire ensembliste

Appartenance, inclusion, notations \in, \subset .

Ensemble $\mathcal{P}(E)$ des parties de E .

Complémentaire, notation \bar{A} .

Union, intersection, notations \cup, \cap .

Distributivité, lois de Morgan.

Définition du produit cartésien d'ensembles.

Applications, composition, restriction.

Applications injectives, surjectives, bijectives.

Lien entre les opérations ensemblistes et les connecteurs logiques.

Exemples : $\mathbf{R}^2, \mathbf{R}^n$.

Introduire ces notions en situation. Le vocabulaire doit être connu.

3 - Les nombres entiers

Notations \mathbf{N} et \mathbf{Z} .

Raisonnement par récurrence.

Notations \sum, \prod . Définition de $n!$.

Formules

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}; \quad \sum_{k=0}^n x^k = \frac{x^{n+1} - 1}{x - 1}.$$

Il sera introduit sur des exemples.

Savoir les retrouver.

4 - La droite réelle

Propriétés élémentaires des opérations de $(\mathbf{R}, +, \times)$.

Manipulation d'inégalités.

Intervalles.

Valeur absolue, inégalité triangulaire.

Majorant, minorant, maximum, minimum, borne supérieure, borne inférieure.

La construction de \mathbf{R} n'est pas au programme.

L'existence d'une borne supérieure (inférieure) pour toute partie non vide majorée (minorée) de \mathbf{R} est admise.

5 - Le plan complexe

Partie réelle, partie imaginaire, conjugué d'un nombre complexe.

Opérations, propriétés élémentaires de $(\mathbf{C}, +, \times)$, calcul du quotient en coordonnées cartésiennes.

Module, argument, notation exponentielle, calcul du produit et du quotient en coordonnées polaires.

Formules d'Euler et de Moivre.

Résolution des équations du second degré à coefficients réels.

On donnera l'interprétation géométrique de ces notions.

Lien avec la notion informelle de coordonnées polaires.

Lien avec les formules trigonométriques.

Les formules de résolution sont exigibles, ainsi que leur démonstration.

Première année

I - Suites et séries de nombres réels

Limite (finie ou infinie) d'une suite.

Unicité de la limite.

Toute suite convergente est bornée.

Opérations algébriques sur les suites convergentes.

Passage à la limite dans des inégalités.

Étude de convergence par encadrement.

Comparaison entre les suites $n!$, a^n , n^b , $(\ln n)^c$.

Suites monotones, limites des suites monotones.

On donnera la définition sans en faire un usage systématique.

Toute suite croissante admet une limite (finie ou infinie) qui est, si elle est majorée, la borne supérieure de ses termes (démonstration non exigible).

Aucune théorie générale n'est exigible sur les suites définies par récurrence.

Suites définies par récurrence, suites géométriques, suites arithmétiques, suites arithmético-géométriques.

Séries à termes positifs.

Somme (finie ou infinie).

On pourra montrer l'existence d'une valeur limite (finie ou infinie) par la croissance des sommes partielles.

Paradoxe de Zénon.

Résultat admis (les termes sont positifs).

$$\sum_k \sum_l a_{k,l} = \sum_l \sum_k a_{k,l}$$

Séries géométriques.

Convergence des séries de Riemann $\sum n^{-a}$.

Résultat admis, la démonstration par comparaison avec l'intégrale pourra être traitée en exercice le moment venu.

Si $0 \leq u_n \leq v_n$ pour tout n , $\sum_n u_n \leq \sum_n v_n$. En particulier, si $\sum_n v_n$ est finie, alors $\sum_n u_n$ aussi.

Sommation d'inégalités.

Admis.

À part ces exemples, on ne fera aucune théorie sur les séries à termes non positifs.

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}; \text{ pour } |x| < 1, \frac{1}{1-x} = \sum_{k=0}^{\infty} x^k.$$

II - Algèbre linéaire

L'accent sera mis dans la présentation de l'algèbre linéaire sur les sous-espaces de \mathbf{R}^n , avec de nombreux exemples en dimension 2 ou 3 visant à développer l'intuition géométrique. Représenter un sous-espace vectoriel comme noyau d'une matrice revient à en donner un système d'équations. Représenter un sous-espace vectoriel comme image d'une matrice revient à en donner une description paramétrique. On montrera notamment comment passer d'un point de vue à l'autre.

1 - L'espace \mathbf{R}^n

Définition de l'espace \mathbf{R}^n des n -uplets de réels, interprétation géométrique comme vecteurs.

Les opérations $+$: $\mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ et \cdot : $\mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n$.

Les propriétés :

commutativité $x + y = y + x$;

associativité $(x + y) + z = x + (y + z)$;

$(ab) \cdot x = a \cdot (b \cdot x)$;

opposé $x + (-1) \cdot x = 0$;

distributivité $(a+b) \cdot (x+y) = a \cdot x + a \cdot y + b \cdot x + b \cdot y$.

Notion de combinaison linéaire, définition d'un sous-espace vectoriel de \mathbf{R}^n .

Classification des sous-espaces vectoriels de \mathbf{R}^2 .

Représentation d'une droite vectorielle ou affine par équation ou par paramétrisation.

2 - Matrices et systèmes linéaires

Matrices à coefficients réels. Application associée de \mathbf{R}^m dans \mathbf{R}^n .

Définition d'une application linéaire de \mathbf{R}^m dans \mathbf{R}^n .

L'application associée à une matrice est linéaire.

Toute application linéaire est associée à une matrice.

Somme de matrices, produit par un réel, propriétés.

Produit de matrices et composition.

Systèmes linéaires, écriture sous la forme

$$A(x) = y.$$

Noyau d'une application linéaire. Une application linéaire est injective si et seulement si son noyau est réduit à $\{0\}$.

On définira les opérations par leur expression algébrique en en donnant l'interprétation géométrique.

Les propriétés pourront être démontrées à partir de la définition des opérations.

Si le sous-espace n'est pas $\{0\}$, il contient un vecteur non nul e . Si tous les éléments du sous-espace sont proportionnels à e , c'est une droite vectorielle, sinon, c'est \mathbf{R}^2 .

$$(x_i) \mapsto (\sum_j A_{ij}x_j)$$

Les opérations sur $\mathbf{M}_{n,m}(\mathbf{R})$ vérifient les mêmes propriétés que celles de \mathbf{R}^{mn} .

Algorithme de Gauss pour réduire un système linéaire (une matrice) à une forme échelonnée par opérations sur les lignes.

Le système $A(x) = 0$ a des solutions non triviales si A a moins de lignes que de colonnes.

Matrice transposée, produit des transposées.

3 - Matrices carrées inversibles

Dans cette section les matrices sont carrées (et réelles).

Matrice diagonale, triangulaire.

Trace, trace d'un produit.

Matrice inversible, équivalence entre :

- le système $A(x) = y$ a une unique solution pour tout y ;
- il existe une matrice carrée B telle que $AB = I = BA$.

Calcul de l'inverse par pivot de Gauss.

Inverse d'un produit.

Les opérations sur les lignes correspondent à des multiplications à gauche par des matrices inversibles, matrices des opérations élémentaires.

Déterminant des matrices 2×2 .

4 - Sous-espaces vectoriels de \mathbf{R}^n

Le noyau et l'image d'une matrice sont des sous-espaces vectoriels.

Sous-espace vectoriel engendré par une famille de vecteurs, famille génératrice.

Famille libre de vecteurs de \mathbf{R}^n , base d'un sous-espace vectoriel, coordonnées dans une base.

Base canonique de \mathbf{R}^n .

On donnera des exemples de résolution de systèmes linéaires (homogènes ou non) en utilisant l'algorithme de Gauss. Des exemples où il existe une solution unique, où il n'existe pas de solution, et où il existe plusieurs solutions seront traités. Dans ce dernier cas, on donnera une représentation paramétrique de l'ensemble des solutions.

Ce résultat sera utilisé en théorie de la dimension.

Détermination d'un système d'équations pour l'image d'une matrice (écriture de l'image d'une matrice comme noyau d'une autre matrice).

$$(AB)^T = B^T A^T.$$

Résolution des systèmes linéaires associés.

La matrice B est alors unique, c'est l'inverse de A , notée A^{-1} .

La solution du système $A(x) = y$ est $x = A^{-1}(y)$.

Une matrice 2×2 est inversible si et seulement si son déterminant est non nul.

La théorie générale des déterminants est hors programme.

Le noyau est l'ensemble des solutions du système $A(x) = 0$, l'image est l'ensemble des seconds membres y tels que le système $A(x) = y$ a une solution. L'image est aussi le sous-espace vectoriel engendré par les colonnes.

Dans un sous-espace vectoriel de \mathbf{R}^n , une famille libre a un nombre d'éléments inférieur ou égal à celui d'une famille génératrice. Toutes les bases ont le même nombre d'éléments.

Toute famille libre de \mathbf{R}^n a au plus n éléments.

Existence d'une base pour tout sous-espace vectoriel de \mathbf{R}^n , théorème de la base incomplète.

Dimension d'un sous-espace vectoriel de \mathbf{R}^n .

Dans un sous-espace vectoriel de dimension d , une famille libre à d éléments est une base, une famille génératrice à d éléments est une base.

5 - Applications linéaires entre sous-espaces vectoriels de \mathbf{R}^n .

Noyau et image d'une application linéaire, rang.

Représentation par une matrice dans des bases.

Changement de bases, formule $A' = Q^{-1}AP$.

Toute application linéaire de rang r peut être représentée dans des bases appropriées par la matrice

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Théorème du rang.

Isomorphismes.

Une application linéaire entre sous-espaces vectoriels de même dimension est un isomorphisme si elle est injective, si elle est surjective.

Deux sous-espaces vectoriels isomorphes ont même dimension.

6 - Rang d'une matrice

Le rang en lignes est égal au rang en colonnes.

Le rang est le nombre de lignes non nulles dans les formes échelonnées.

Toute matrice carrée inversible à droite (à gauche) est inversible.

Pour la démonstration, on pourra utiliser qu'un système linéaire homogène ayant moins d'équations que d'inconnues a des solutions non triviales.

Trouver une base du noyau d'une matrice est une compétence exigible.

Si $E \subset F \subset \mathbf{R}^n$ sont des sous-espaces vectoriels, alors la dimension de E est inférieure ou égale à celle de F .

Cet énoncé pourra être démontré en utilisant le théorème de la base incomplète ou en utilisant les opérations élémentaires. Il conduit au théorème du rang.

En particulier \mathbf{R}^n et \mathbf{R}^m ne sont pas isomorphes pour $n \neq m$.

On pourra se ramener à la matrice équivalente de la forme $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

7 - Espaces vectoriels

Un espace vectoriel de dimension n est un ensemble E muni d'une opération interne $+$, d'une opération externe \cdot , et d'une bijection f de E dans \mathbf{R}^n qui préserve les combinaisons linéaires.

Applications linéaires entre espaces vectoriels, endomorphismes, isomorphismes.

Exemples : \mathbf{R}^n , $\mathbf{R}_n[x]$, $\mathbf{M}_{n,m}(\mathbf{R})$, $\mathcal{L}(E, F)$.

Sous-espaces vectoriels et restrictions d'applications linéaires.

Bases d'espaces vectoriels, la dimension est aussi le cardinal des bases.

On ne considérera donc que des espaces vectoriels de dimension finie. Cette définition a pour but de permettre de discuter d'espaces autres que \mathbf{R}^n , mais aucune difficulté abstraite ne sera soulevée.

La notion d'indéterminée ne sera pas introduite.

La bijection structurelle f permet de se ramener à la théorie de \mathbf{R}^n .

III - Fonctions d'une variable réelle

Un des objectifs principaux est de savoir étudier une fonction, tracer et interpréter son graphe.

Domaine de définition, tableau de variations, graphe d'une fonction.

Graphe des fonctions x^2 , x^3 , \sqrt{x} , $\sqrt[3]{x}$, $\ln x$, e^x , $\sin x$, $\cos x$, $\tan x$, $|x|$.

Fonctions périodiques, paires, impaires, majorées, minorées, bornées.

Fonctions monotones, strictement monotones.

On notera indifféremment une fonction de la variable x , quand le contexte est sans ambiguïté, par $x \mapsto f(x)$ ou simplement $f(x)$.

Ces graphes serviront d'illustrations aux concepts introduits dans cette section.

1 - Limites et continuité

Limite (finie ou infinie) d'une fonction en un point, limite à gauche, limite à droite, limites en $\pm\infty$.

Opérations algébriques sur les limites.

Limites et composition (de deux fonctions ou d'une fonction et d'une suite).

Unicité de la limite, passage à la limite dans des inégalités.

Étude de convergence par encadrement.

$\lim_0 x^a |\ln x|^b$, $\lim_\infty x^a |\ln x|^b$, $\lim_{\pm\infty} |x|^a e^x$.

Continuité, opérations sur les fonctions continues.

Les fonctions classiques sont continues sur leur domaine de définition : \ln , \exp , \sin , \cos , x^a , $|x|$.

Exemples de prolongement par continuité.

Les définitions seront interprétées graphiquement et illustrées par des exemples.

On remarquera, sans s'y attarder formellement, que la notion de limite en un point s'étend au cas d'une fonction non définie en ce point.

La démonstration de l'existence et de la valeur de ces limites n'est pas exigible.

La continuité est vraie par définition pour \exp , elle est admise pour \sin , \cos . Elle sera démontrée pour \ln en tant que fonction réciproque de \exp .

L'image d'un intervalle fermé borné par une fonction continue est un intervalle fermé borné. Théorème des valeurs intermédiaires.

Toute fonction continue sur un intervalle fermé borné atteint son maximum et son minimum.

Toute fonction continue et strictement monotone sur un intervalle admet une fonction réciproque sur l'intervalle image, qui est continue et strictement monotone.

Graphes de la fonction réciproque par symétrie.

2 - Dérivées

Dérivée, dérivée à gauche, dérivée à droite, interprétation graphique.

Équation de la droite tangente au graphe en un point.

Fonction dérivée. Cas des fonctions classiques : \ln , \exp , \sin , \cos , x^a .

Méthodes de calcul (linéarité, produit, quotient, composition).

Dérivée de la fonction réciproque.

Développement limité d'ordre 1.

Dérivée et extremums.

Théorème de Rolle, égalité des accroissements finis.

Inégalité des accroissements finis.

Utilisation de la dérivée pour l'étude des variations.

Une fonction dont la dérivée est strictement positive est strictement croissante.

Dérivées d'ordre supérieur, fonctions \mathcal{C}^k , \mathcal{C}^∞ .

Résultat admis.

La fonction arctan sera introduite comme exemple.

La dérivabilité implique la continuité.

Application au calcul de la dérivée de arctan.

Une fonction dérivable sur $[a, b]$ atteint son minimum en un point x_0 . Si $x_0 \in]a, b[$ alors $f'(x_0) = 0$, si $x_0 = a$ alors $f'(a^+) \geq 0$, si $x_0 = b$ alors $f'(b^-) \leq 0$. Démonstration à partir du développement limité. Exemples.

On fera remarquer qu'un point critique (c'est-à-dire un point où la fonction dérivée s'annule) n'est pas forcément un extremum.

Démonstration ce qui précède.

On donnera les versions de l'inégalité pour $m \leq f' \leq M$ et pour $|f'| \leq k$.

Exemples d'application à la convergence de suites récurrentes.

Une fonction est constante sur un intervalle si et seulement si sa dérivée est identiquement nulle.

Une fonction est croissante sur un intervalle si et seulement si sa dérivée est positive ou nulle.

On donnera la démonstration. On remarquera, sans formalisation, que le résultat reste vrai si la dérivée s'annule en un nombre fini de points.

On insistera plus sur les conclusions et l'utilisation des résultats que sur leurs hypothèses de régularité.

3 - Exemple d'étude de fonction : régression linéaire

On considère des données se présentant comme des couples de variables (x_i, y_i) , $1 \leq i \leq n$, où x_i est vue comme une variable explicative de y_i . En notant \bar{x} et \bar{y} les moyennes, on cherche un coefficient a tel que $\bar{y} + a(x - \bar{x})$ soit une bonne approximation de y . On peut pour cela minimiser la somme des écarts quadratiques

$$\sum_i (y_i - \bar{y} - a(x_i - \bar{x}))^2.$$

En étudiant la fonction de a ci-dessus, on montre que la valeur optimale de a est

$$a = \frac{\sum_i ((x_i - \bar{x})(y_i - \bar{y}))}{\sum_i (x_i - \bar{x})^2}.$$

C'est le britannique Francis Galton qui a introduit la méthode au 19^e siècle pour étudier la taille des enfants y_i en fonction de la taille des parents x_i , obtenant un coefficient $a \approx 2/3$. Le fait que $a > 0$ signifie que les enfants sont (en moyenne) plus grands que la moyenne lorsque les parents le sont. Le fait que $a < 1$ signifie un retour vers la moyenne, « regression towards the mean » en anglais, d'où le nom de la méthode. Les explications génétiques que Galton donna de ce phénomène sont aujourd'hui considérées comme incorrectes.

On pourra interpréter cette valeur comme

$$a = \frac{\text{Cov}(x, y)}{\text{Var}(x)},$$

en lien avec le cours de probabilités.

4 - Intégration

Définition informelle de l'intégrale $\int_a^b f$ comme aire algébrique.

Propriétés de l'intégrale : linéarité, relation de Chasles, monotonie.

Inégalité de la moyenne.

La fonction $\int^x f$ est une primitive de f .

Les primitives de f diffèrent d'une constante additive.

Relation $\int_a^b f = F(b) - F(a)$.

Primitives de fonctions usuelles : e^x , x^a , $\ln x$, $\sin x$, $\cos x$.

Calcul d'intégrales : intégration par parties, changements de variable.

On ne soulèvera pas de difficulté sur la régularité de f , on limitera la discussion aux fonctions continues.

Les propriétés ne seront pas démontrées, mais interprétées en termes géométriques.

Démonstration à partir des propriétés de l'intégrale.

Vraie pour toute primitive F de f .

Il n'est pas attendu des étudiants qu'ils sachent trouver eux-mêmes les bons changements de variable, sauf dans quelques cas simples comme les changements affines.

IV - Probabilités

L'esprit du programme de probabilités est de familiariser les étudiants au concept de variables aléatoires et de leur indépendance, dont la partie statistique, en deuxième année, peut être vue comme un aboutissement. Les variables aléatoires finies ou discrètes sont plus à envisager comme un contexte dans lequel certaines des propriétés importantes peuvent être démontrées de manière simple que comme une source d'exercices de dénombrement.

1 - Événements aléatoires

Univers Ω , ensemble \mathcal{E} des événements.

Événement A et B , événement A ou B , événement contraire, événements incompatibles, famille complète d'événements.

Probabilité $P : \mathcal{E} \rightarrow [0, 1]$, axiomes :

$$P(\Omega) = 1, P(\emptyset) = 0;$$

$$P(A) \leq P(B) \text{ si } A \subset B;$$

$$P(A \cup B) = P(A) + P(B) \text{ si } A \cap B = \emptyset;$$

$$P(\cup_i A_i) = \sum_i P(A_i) \text{ si } A_i \text{ est une suite d'événements deux à deux disjoints.}$$

Probabilité conditionnelle sachant un événement B de probabilité non nulle :

$$P(A|B) = P_B(A) = \frac{P(A \cap B)}{P(B)}.$$

Formule des probabilités composées :

$$P(A_1 \cap A_2 \cap \dots \cap A_n) =$$

$$P(A_1)P(A_2|A_1) \dots P(A_n|A_1 \cap \dots \cap A_{n-1}).$$

Formule des probabilités totales :

$$P(A) = \sum_i P(A|B_i)P(B_i).$$

Formules de Bayes :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_i P(B|A_i)P(A_i)}$$

\mathcal{E} est un ensemble de parties de Ω . On ne soulèvera pas de difficultés sur l'ensemble \mathcal{E} .

Lien avec les opérations ensemblistes. On mentionnera que la réunion d'une suite d'événements est un événement.

On remarque que $P_B : \mathcal{E} \rightarrow [0, 1]$ est une probabilité.

(B_i) est une famille complète d'événements de probabilité strictement positive.

On donnera des applications concrètes de ces formules.

(A_i) est un système complet d'événements, tous les événements sont de probabilité non nulle. On donnera des applications concrètes.

2 - Variables aléatoires discrètes

On considère ici une variable aléatoire à valeurs dans un ensemble $\mathcal{X} \subset \mathbf{R}$ de la forme $\mathcal{X} = \{x_i, i \in I\}$, où I est soit \mathbf{N} , soit \mathbf{Z} , soit un ensemble fini.

Une variable aléatoire sur \mathcal{X} est une fonction $X : \Omega \rightarrow \mathcal{X}$ telle que $X^{-1}(\{x\}) = \{X = x\}$ est un événement pour tout $x \in \mathcal{X}$.

La loi de la variable aléatoire est la fonction $x \mapsto p(x) := P(X = x)$.

$$\sum_{x \in \mathcal{X}} p(x) = 1$$

Pour toute partie $\mathcal{X}' \subset \mathcal{X}$,

$$P(X \in \mathcal{X}') = \sum_{x \in \mathcal{X}'} p(x).$$

Fonction de répartition, quantiles.

Variable de Bernoulli, loi de Bernoulli.

Loi uniforme sur un ensemble fini.

$(\{X = x\}, x \in \mathcal{X})$ est une famille complète d'événements.

$$F_X(x) = P(X \leq x).$$

L'indicatrice de l'événement A suit une loi de Bernoulli de paramètre $P(A)$.

3 - Moments des variables aléatoires discrètes réelles positives

Les variables aléatoires sont positives dans cette partie.

Espérance des variables aléatoires positives.

Moment d'ordre k (k -moment) $E(X^k) \in [0, \infty]$.

Inégalité $E(X)^2 \leq E(X^2)$.

Propriétés de l'espérance : linéarité, monotonie.

Variance des variables de moment d'ordre 2 fini :

$$V(X) = E((X - E(X))^2), \\ V(X) = E(X^2) - E(X)^2, V(aX + b) = a^2V(X).$$

Si $V(X) = 0$, la variable X est constante en dehors d'un événement de probabilité nulle.

Si X est une variable de Bernoulli de paramètre p , alors $E(X) = p$, $V(X) = p(1 - p)$.

Inégalité de Markov :

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

si X est une variable aléatoire à valeurs positives et $a > 0$.

On évoquera l'espérance des variables aléatoires finies de signe quelconque.

Formule de transfert.

Une variable prenant un nombre fini de valeurs a des moments finis.

Admise, mais on pourra par la suite faire le lien avec l'inégalité de Cauchy-Schwarz. Une variable de second moment fini est donc de premier moment fini.

La linéarité est admise.

Définition de l'écart type $\sigma(X) = \sqrt{V(X)}$.

Puisque $V(|X|) \geq 0$ on retrouve l'inégalité $E(|X|)^2 \leq E(X^2)$ dans le cas des variables de moment d'ordre 2 fini.

Inégalité de Bienaymé-Tchebychev :

$$P(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}$$

si X est une variable aléatoire de second moment fini et $a > 0$.

Covariance de deux variables aléatoires finies :

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

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

Coefficient de corrélation de variables aléatoires finies :

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}.$$

Invariance d'échelle :

$$\text{Cor}(X, Y) = \text{Cor}(aX + b, cY + d)$$

pour $a > 0, c > 0$.

4 - Indépendance

Indépendance de deux événements :

$$P(A \cap B) = P(A)P(B).$$

Indépendance de variables aléatoires discrètes.

Les événements A et B (avec $P(B) \neq 0$) sont indépendants si et seulement si $P(A|B) = P(A)$.

Les variables X_1, X_2, \dots, X_n , à valeurs dans $\mathcal{X}_1, \dots, \mathcal{X}_n$ sont dites indépendantes si, pour toutes parties $A_1 \subset \mathcal{X}_1, \dots, A_n \subset \mathcal{X}_n$,

$$P((X_1 \in A_1) \cap \dots \cap (X_n \in A_n)) = P(X_1 \in A_1) \dots P(X_n \in A_n).$$

Deux variables de Bernoulli B_1 et B_2 sont indépendantes si et seulement si les événements $\{B_1 = 1\}$ et $\{B_2 = 1\}$ sont indépendants.

Les variables discrètes X_1, \dots, X_n sont indépendantes si et seulement si

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2) \dots P(X_n = x_n)$$

pour tout $x_1 \in \mathcal{X}_1, \dots, x_n \in \mathcal{X}_n$.

La covariance de deux variables indépendantes est nulle.

Si $X_1, X_2, \dots, X_k, Y_1, \dots, Y_l$ sont des variables indépendantes, alors $f(X_1, \dots, X_k), Y_1, \dots, Y_l$ sont indépendantes pour toute fonction f .

Si les variables X_1, \dots, X_n sont indépendantes et de second moment fini, alors

$$E(X_1 \dots X_n) = E(X_1) \dots E(X_n)$$

$$\text{et } V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n).$$

Cette condition n'est pas suffisante.

On ne soulèvera aucune difficulté sur la notion de fonction de plusieurs variables.

Démonstration dans le cas fini.

5 - Processus de Bernoulli

On considère dans cette section une suite $X_i, i \in \mathbf{N}^*$ de variables indépendantes suivant une loi de Bernoulli de paramètre p . L'indépendance signifie ici que, pour tout n , les variables $X_i, 1 \leq i \leq n$ sont indépendantes.

Soit T l'indice du premier 1. Alors T est une variable aléatoire à valeurs dans $\mathbf{N}^* = \{1, 2, \dots\}$. Sa loi est la loi géométrique :

$$P(T = n) = p(1 - p)^{n-1}$$

$$P(T > n) = (1 - p)^n.$$

Les lois géométriques satisfont la propriété

$$P(T > j + k \mid T > j) = P(T > k).$$

Moments : $E(T) = 1/p, V(T) = (1 - p)/p^2$.

On pose $S_n = X_1 + \dots + X_n$. La loi de S_n est la loi binomiale de paramètres n et p :

$$P(S_n = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Relations :

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

$$\binom{n}{k} = \binom{n}{n-k}, \quad k \binom{n}{k} = n \binom{n-1}{k-1}.$$

Relation

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Moments :

$$E(S_n) = np, V(S_n) = nV(X_1) = np(1 - p).$$

La somme de deux variables binomiales indépendantes de paramètres (k, p) et (l, p) est une variable binomiale de paramètres $(k + l, p)$.

Loi de Poisson de paramètre λ :

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Moments d'une variable X suivant une loi de Poisson de paramètre λ :
 $E(X) = \lambda, \quad V(X) = \lambda.$

Notation $\mathcal{G}(p)$ de la loi géométrique de paramètre p .

On pourra remarquer que cette propriété caractérise les lois géométriques, mais la démonstration n'est pas exigible.

On présentera le calcul de l'espérance en admettant la dérivation sous le signe somme.

On introduira à cette occasion les coefficients binomiaux $\binom{n}{k}$ de manière combinatoire et la notation $\mathcal{B}(n, p)$ pour la loi binomiale de paramètres n et p .

Triangle de Pascal.

Lien avec la loi binomiale de paramètres n et $p = a/(a + b)$ lorsque a et b sont positifs.

On justifiera ce résultat en interprétant cette somme comme une somme de $(k + l)$ variables de Bernoulli indépendantes.

Notation $\mathcal{P}(\lambda)$.

On présentera le calcul de l'espérance en admettant la dérivation sous le signe somme.

On démontrera la convergence quand $n \rightarrow \infty$ de la suite de lois binomiales de paramètres n et p_n vers la loi de Poisson de paramètre λ si $np_n \rightarrow \lambda$:

$$\forall k \quad \binom{n}{k} p_n^k (1 - p_n)^{n-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}.$$

Il s'agit ici d'une convergence pour chaque k . On ne discutera pas le concept de convergence en loi en général.

Interprétation, paradigme de Poisson :

La somme S_n d'un grand nombre de variables de Bernoulli indépendantes de petit paramètre suit approximativement la loi de Poisson de paramètre $E(S_n)$.

On illustrera ce paradigme par des exemples concrets.

Exemple : dans un texte, le nombre de coquilles par page peut être modélisé par une loi de Poisson. Si, dans une collection de qualité et de pagination homogènes, le nombre moyen de coquilles par page a été estimé à $1/2$, la probabilité qu'une page donnée ne contienne pas de coquille est de 60% environ.

Deuxième année

I - Algèbre et géométrie

Comme en première année, les scalaires sont réels, et le contexte est celui de \mathbf{R}^n (ou, éventuellement, d'espaces isomorphes à \mathbf{R}^n comme $\mathbf{R}_k[x]$ et $\mathbf{M}_{k,l}(\mathbf{R})$) et de ses sous-espaces.

1 - Somme directe, supplémentaire

Somme de sous-espaces vectoriels.

Somme directe de deux sous-espaces vectoriels, caractérisation par l'intersection.

Supplémentaire d'un sous-espace vectoriel dans un autre, dimension des supplémentaires.

Projection sur E parallèlement à F .

Tout endomorphisme P vérifiant $P^2 = P$ est une projection.

Symétrie par rapport à E le long de F .

Tout endomorphisme S vérifiant $S^2 = I$ est une symétrie.

$$E + F = \{x + y, x \in E, y \in F\}.$$

La réunion d'une base de E et d'une base de F est une base de $E \oplus F$.

Si F est un sous-espace vectoriel, et E un sous-espace vectoriel de F , alors E admet un supplémentaire dans F .

$$F = \ker P, E = \ker(P - I)$$

$$F = \ker(S + I), E = \ker(S - I)$$

Cette définition inclut le cas $S = I$.

2 - Valeurs propres des endomorphismes

Représentation d'un endomorphisme dans une base par une matrice carrée, changement de base, formule $A' = P^{-1}AP$.

Valeurs propres, vecteurs propres.

Des vecteurs propres dont les valeurs propres associées sont distinctes forment une famille libre. Endomorphisme diagonalisable. Matrice diagonalisable.

Une application linéaire d'un sous-espace vectoriel E dans lui-même est diagonalisable si et seulement s'il existe une base de E constituée de vecteurs propres de cette application linéaire.

Si une matrice carrée $n \times n$ admet n valeurs propres distinctes, alors elle est diagonalisable.

Diagonalisation pratique des matrices 2×2 .

Une matrice A est diagonalisable s'il existe une matrice inversible P telle que $P^{-1}AP$ soit diagonale. Un endomorphisme est diagonalisable s'il existe une base dans laquelle il est représenté par une matrice diagonale.

3 - Produit scalaire

Seul le produit scalaire usuel de \mathbf{R}^n sera étudié, aucun développement sur les espaces euclidiens plus généraux n'est au programme.

Définition du produit scalaire $\langle x, y \rangle = \sum x_i y_i$ et de la norme $\|x\| = \sqrt{\langle x, x \rangle}$.

Bilinéarité et symétrie du produit scalaire

Formule $\langle Ax, y \rangle = \langle x, A^T y \rangle$.

Orthogonalité entre deux vecteurs. Théorème de Pythagore : $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ si et seulement si x et y sont orthogonaux.

Familles orthogonales, orthonormées, base orthonormée.

Toute famille orthogonale de vecteurs non nuls est libre.

Toute famille orthonormée se complète en une base orthonormée.

Coordonnées dans une base orthonormée.

Inégalité de Cauchy-Schwarz :

$$\langle x, y \rangle \leq \|x\| \|y\|.$$

Inégalité triangulaire $\|x + y\| \leq \|x\| + \|y\|$.

Distance $\|x - y\|$.

La convergence de la suite $(\|x^{(m)}\|)$ vers 0 est équivalente à celle de chacune des suites de coordonnées $(x_i^{(m)})$.

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

Démonstration possible : Dans une base orthonormée telle que $\|x\| e_1 = x$, on a :

$$\|y\|^2 = \sum_i \langle e_i, y \rangle^2 \geq \langle e_1, y \rangle^2.$$

Boule, sphère de centre et de rayon donnés.

Orthogonal d'une partie, d'un sous-espace vectoriel.

L'orthogonal d'un sous-espace vectoriel en est un supplémentaire.

Pour tout sous-espace vectoriel E , $E^{\perp\perp} = E$.

Hyperplans.

Projection orthogonale P_E sur un sous-espace E muni d'une base orthonormée (e_i) (c'est la projection sur E parallèlement à E^\perp).

Une base de l'orthogonal donne un système d'équations du sous-espace.

L'orthogonal d'un vecteur non nul est un hyperplan.

Deux vecteurs orthogonaux à un même hyperplan sont colinéaires.

$$P_E(x) = \sum_i \langle e_i, x \rangle e_i$$

Le projeté $P_E(x)$ de x sur E est le point de E le plus proche de x . La distance au sous-espace E est $\|x - P_E(x)\|$.

II - Étude locale des fonctions d'une variable réelle

1 - Fonctions polynomiales

Étude de polynômes, limites en $\pm\infty$.

Racines (réelles), tout polynôme de degré impair admet une racine.

Factorisation d'un polynôme par $(x - x_0)$ si x_0 est une racine.

Multiplicité d'une racine : la racine x_0 est de multiplicité k si $f(x) = (x - x_0)^k g(x)$, avec $g(x_0) \neq 0$.

Un polynôme change de signe en une racine si et seulement si sa multiplicité est impaire.

Un polynôme f admet un extremum local en x_0 si et seulement si x_0 est une racine de f' de multiplicité impaire.

Les polynômes sont à coefficients réels et sont vus comme des fonctions d'une variable réelle.

L'existence d'une factorisation est admise, mais savoir factoriser en pratique est exigible.

Si x_0 est une racine de f de multiplicité $k \geq 2$, alors c'est une racine de f' de multiplicité $k - 1$.

Les étudiants doivent connaître le cas $f'(x_0) = 0$ et $f''(x_0) \neq 0$.

2 - Développements limités

On se limitera autant que possible à l'ordre 2 ou 3 et on ne cherchera aucune technicité.

Développements limités, formule de Taylor Young (admise).

Unicité du développement limité.

Développement limité de e^x et $(1 - x)^{-1}$ à tout ordre. Premiers termes de $(1 + x)^a$ et $\ln(1 + x)$. Allure locale du graphe d'une fonction admettant un développement limité du type

$$f(x) = a_0 + a_1x + a_kx^k + x^k\varepsilon(x),$$

où $k \geq 2$ et $a_k \neq 0$.

Quelques exemples simples de calcul de limite à l'aide de développements limités.

La forme du graphe en un point dépend principalement du premier terme non linéaire du développement limité. Exemples avec $k = 2$ ou $k = 3$.

Lien avec les extremums locaux et les points d'inflexion.

Détermination de l'asymptote oblique d'une fonction en l'infini. Position par rapport à l'asymptote.

Étude de la fonction e^{-x^2} .

3 - Intégrales généralisées

Notion d'intégrale généralisée pour des fonctions positives, du type

$$\int_a^b f \in [0, \infty]$$

où a ou b peuvent être infinis et où f est continue sur $]a, b[$.

Extension des propriétés de linéarité, de monotonie, et de la relation de Chasles à ce cadre.

$$\int_0^\infty e^{-t} dt = 1.$$

$$\int_{-\infty}^\infty e^{-t^2/2} dt = \sqrt{2\pi}.$$

Conditions de convergence de $\int_0^\infty t^{-a} dt$ et $\int_{-\infty}^\infty t^{-a} dt$.

III - Fonctions de deux variables réelles

Le niveau de formalisme de cette partie sera minimal. On ne cherchera pas à préciser les hypothèses générales des résultats. Aucune notion précise sur la classe de régularité d'une fonction de deux variables n'est exigible. Les fonctions seront le plus souvent définies sur le plan \mathbf{R}^2 tout entier. Dans le cas contraire, on ne soulèvera aucune difficulté liée au bord de l'ensemble de définition.

Notation $f(x) = f(x_1, x_2)$.

1 - Exemples

Graphe, lignes de niveau. Étude d'exemples, notamment les suivants (allure du graphe et des lignes de niveau) :

Fonctions coordonnées $(x_1, x_2) \mapsto x_1$ et $(x_1, x_2) \mapsto x_2$.

Fonctions linéaires $ax_1 + bx_2$.

En un minimum local à l'intérieur du domaine de définition, le coefficient a_2 d'ordre 2 du développement limité vérifie $a_2 \geq 0$.

Un point où $a_1 = 0$ et $a_2 > 0$ est un minimum local.

La forme du graphe doit être connue.

On donnera la définition comme limite tout en évoquant une notion informelle directe comme une aire (finie ou infinie).

Les deux terminologies pourront être employées : intégrale convergente (divergente) ; intégrale finie (infinie).

Admises.

Valeur admise, on pourra toutefois démontrer la convergence par comparaison.

Le vecteur (a, b) est orthogonal aux droites de niveau.

2 - Dérivées partielles

Dérivées partielles.

Définition d'un extremum local, condition nécessaire : $\partial_1 f = 0 = \partial_2 f$.

Point critique, tout minimum local est un point critique.

Notations $\partial_1 f, \partial_2 f$.

Le point x_0 est un minimum local s'il existe $\delta > 0$ tel que $f(x) \geq f(x_0)$ dans le disque $B(x_0, \delta)$. On ne discutera pas d'extremum atteint au bord du domaine de définition.

Démonstration des conditions d'optimalité à l'aide des fonctions $t \mapsto f(t, x_2)$ et $t \mapsto f(x_1, t)$.

3 - Fonctions quadratiques

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

La fonction quadratique a un extremum local (strict) en 0 si et seulement si le discriminant $\Delta = b^2 - ac$ est (strictement) négatif.

Détermination (dans les cas $\Delta < 0$) du type d'extremum en fonction du signe de a et c .

4 - Retour sur la régression linéaire

Il n'est pas demandé aux étudiants de connaître les résultats de cet exemple, mais de savoir les retrouver.

Dans la présentation classique de la régression linéaire, on cherche deux coefficients a et b tels que $b + ax$ soit une aussi bonne approximation que possible de y . On minimise pour cela la fonction

$$f(a, b) = \sum_i (y_i - b - ax_i)^2,$$

et on retrouve $a = \text{Cov}(x, y) / \text{Var}(x)$, $b = \bar{y} - a\bar{x}$.

On trouve le point critique en calculant les dérivées partielles. On remarque que c'est la même approximation que dans la première approche. On pourra justifier que c'est un minimum par l'interprétation géométrique suivante. Dans l'espace \mathbf{R}^n (n est le nombre de données), on considère $e = (1, \dots, 1)$, $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$. La fonction $f(a, b)$ s'interprète comme le carré de la distance entre les points y et $ax + be$. Les valeurs optimales de a et b correspondent donc à l'unique point $ax + be$ du sous-espace $E = \text{Vect}(e, x)$ tel que $y - (ax + be)$ est orthogonal à E . En écrivant le système

$$\langle y - (ax + be), x \rangle = 0, \langle y - (ax + be), e \rangle = 0,$$

on retrouve bien

$$a = \frac{\langle e, e \rangle \langle x, y \rangle - \langle x, e \rangle \langle y, e \rangle}{\langle e, e \rangle \langle x, x \rangle - \langle x, e \rangle^2}, \quad b = \frac{\langle y, e \rangle - a \langle x, e \rangle}{\langle e, e \rangle}.$$

5 - Étude des points critiques

Dérivées partielles d'ordre 2.

Égalité $\partial_{1,2}^2 f(x) = \partial_{2,1}^2 f(x)$.

Soit x un point critique de f . Si le discriminant de la fonction quadratique

$$q(y_1, y_2) = \partial_{1,1}^2 f(x)y_1^2 + 2\partial_{1,2}^2 f(x)y_1y_2 + \partial_{2,2}^2 f(x)y_2^2$$

est strictement négatif, alors la fonction f a un extremum local en x , qui est de même nature que celui de la fonction q en 0.

Notations $\partial_{1,1}^2 f, \partial_{1,2}^2 f, \partial_{2,1}^2 f, \partial_{2,2}^2 f$

Admis.

Admis.

IV - Probabilités

1 - Variables aléatoires à densité

Soit ρ une fonction positive de \mathbf{R} dans \mathbf{R} . La fonction $X : \Omega \rightarrow \mathbf{R}$ est une variable aléatoire de densité ρ si, pour tout intervalle $]a, b[$ de \mathbf{R} , l'ensemble $\{X \in]a, b[\}$ est un événement et

$$P(X \in]a, b[) = \int_a^b \rho(x) dx.$$

Si X est une variable à densité, alors $P(X = x) = 0$ pour tout $x \in \mathbf{R}$

Fonction de répartition :

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x \rho.$$

Inverse (réciproque) de la fonction de répartition et quantiles.

Loi uniforme sur un intervalle borné.

Moments :

$$\int_{\mathbf{R}} |x|^k \rho(x) dx = \int_0^{\infty} x^k (\rho(x) + \rho(-x)) dx.$$

Espérance des variables à densité de premier moment fini :

$$E(X) = \int_0^{\infty} x \rho(x) dx - \int_0^{\infty} x \rho(-x) dx.$$

Propriétés de l'espérance : linéarité, monotonie.

Variance $V(X) = E((X - E(X))^2)$,

$$V(X) = E(X^2) - E(X)^2, V(aX + b) = a^2 V(X).$$

La variable

$$X^* := \frac{X - E(X)}{\sigma(X)}$$

est centrée réduite.

On se limitera au cas de densités ρ continues (sauf éventuellement en un nombre fini de points).

On remarquera que

$$\int_{-\infty}^{\infty} \rho = P(\Omega) = 1.$$

Un événement de probabilité nulle n'est pas forcément impossible.

La fonction de répartition caractérise la densité (si la densité est continue).

Exemples de calcul de fonctions de répartition et de densités de variables images $f \circ X$ avec f monotone.

Cette expression de l'espérance permet de ne manipuler que des intégrales de fonctions positives.

Admises.

Définition de l'écart type $\sigma(X) = \sqrt{V(X)}$.

Elle a pour densité

$$\rho^*(x) = \sigma(X) \rho(E(X) + \sigma(X)x).$$

Inégalité de Markov :

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

si X est une variable aléatoire à valeurs positives et $a > 0$.

Inégalité de Bienaymé-Tchebychev :

$$P(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}$$

si X est une variable aléatoire de second moment fini et $a > 0$.

2 - Loi normale, loi exponentielle

Une variable aléatoire gaussienne (ou normale) centrée réduite est une variable X admettant la densité

$$\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

On a $E(X) = 0$, $V(X) = 1$.

La variable $Y = \sigma X + E$ est alors une variable gaussienne (ou normale) de moyenne E et de variance σ^2 , elle admet la densité

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-E)^2/2\sigma^2}.$$

Loi exponentielle : $\rho(x) = \lambda e^{-\lambda x}$ sur \mathbf{R}_+

$E(X) = 1/\lambda$, $V(X) = 1/\lambda^2$.

Pour tout $x \geq 0$, $P(X > x) = e^{-\lambda x}$.

Les lois exponentielles satisfont la propriété

$$P(X > t + s | X > t) = P(X > s),$$

pour tous réels s et t positifs ou nuls.

Cette propriété caractérise les lois exponentielles (démonstration non exigible).

Les lois exponentielles s'interprètent comme les lois de durée de vie sans vieillissement. Ce sont des variantes continues des lois géométriques.

3 - Indépendance de variables à densité

Les variables aléatoires à densité X_1, \dots, X_n sont dites indépendantes si

$$\begin{aligned} P((X_1 \in I_1) \cap \dots \cap (X_n \in I_n)) \\ = P(X_1 \in I_1) \cdot \dots \cdot P(X_n \in I_n) \end{aligned}$$

pour tous intervalles I_k de \mathbf{R} .

Si X_1, \dots, X_n sont indépendantes :

$$E(X_1 \cdots X_n) = E(X_1) \cdots E(X_n),$$

$$V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n).$$

On limitera l'utilisation de cette définition.

Propriété admise. La loi d'une somme de variables aléatoires est hors programme.

On pourra, comme en première année, introduire la covariance et le coefficient de corrélation sans discuter les problèmes de définition des intégrales.

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

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

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}.$$

La covariance de deux variables indépendantes est nulle (condition non suffisante).

4 - Statistiques

Soit X une variable aléatoire (discrète ou à densité). Soient $X_i, i \in \mathbf{N}^*$, des variables aléatoires indépendantes de même loi que X . On considère les variables aléatoires $\bar{X}_n = (X_1 + \dots + X_n)/n$.

$$E(\bar{X}_n) = E(X), V(\bar{X}_n) = V(X)/n.$$

Loi faible des grands nombres :

$$P(|\bar{X}_n - E(X)| \geq \varepsilon) \rightarrow 0$$

pour tout $\varepsilon > 0$.

Intervalle de confiance :

La probabilité que l'intervalle

$$\left[\bar{X}_n - \sqrt{\frac{V}{na}}, \bar{X}_n + \sqrt{\frac{V}{na}} \right]$$

contienne $E(X)$ est supérieure ou égale à $1 - a$.

Démonstration dans le cas $V(X) < \infty$
par l'inégalité de Bienaymé-Tchebychev :

$$P(|\bar{X}_n - E(X)| \geq \varepsilon) \leq \frac{V(\bar{X}_n)}{\varepsilon^2} = \frac{V(X)}{n\varepsilon^2}.$$

On discutera la notion d'intervalle de confiance au niveau $1 - a$. On démontrera l'énoncé ci-contre à l'aide de l'inégalité de Bienaymé-Tchebychev.

En pratique :

La variance V est souvent inconnue, mais on peut la majorer :

dans le cas général d'une variable aléatoire bornée $|X| \leq M$, par M^2 ;

dans le cas d'une variable de Bernoulli, par $1/4$.

Application numérique :

Pour $n = 1000$, au seuil de confiance de 90%, l'incertitude est de 5% pour une variable de Bernoulli.