Notions mathématiques et analytiques

Algèbre linéaire

L'algèbre linéaire est la **branche des mathématiques qui s'intéresse aux espaces vectoriels et aux transformations linéaires**, formalisation générale des théories des systèmes d'équation linéaires.

Scalaire

Les **nombres réels ou complexes qui multiplient les vecteurs dans un espace vectoriel** sont appelés des scalaires. Cette multiplication par un scalaire, permettant de multiplier un vecteur par un nombre pour produire un vecteur, correspond à la loi externe de l'espace vectoriel.

Etymologie

Le mot *scalaire* provient du mot anglais *scalar* dérivé lui-même du mot scale utilisé pour un **ensemble de nombres**. Ce dernier provient du latin *scala* désignant un **échelle**.

Définition

Un "vrai scalaire" est un **nombre qui est indépendant de la base choisie pour exprimer les vecteurs**, par opposition à un pseudo-scalaire(grandeur physique représenté par un nombre qui change de signe lorsque le système physique subit une symétrie ou une inversion polaire), qui est un nombre pouvant dépendre de la base.
Il est représenté soit par des lettres grecques, soit par des lettres en italiques.
Un scalaire est un tenseur d'ordre 0. Les quantités non scalaires sont dites pseudo-scalaires.

Exemples

- Un nombre qui mesure une température, une masse ou bien une hauteur est un scalaire.
- Une grandeur scalaire est un scalaire auquel est associé une unité (ex : masse en kg, température en °C)
- La coordonnée d'un vecteur dans une base est un réel (nombre sans unité).
- Un nombre issu d'une opération entre vecteurs peut être un pseudo ou un vrai scalaire selon que les vecteurs opérandes soient des pseudo-vecteurs ou des vecteurs vrais.
- La vélocité (vecteur vitesse) d'un objet ponctuel est un vecteur : elle est définie par un scalaire associé à une direction et un sens. De même pour l'accélération d'un objet ponctuel. La valeur de la vitesse ou de l'accélération est une grandeur scalaire (donc avec une unité).
- En mathématiques, le déterminant d'une matrice est un scalaire, mais pas celui d 'une famille de vecteurs : il dépend de la base choisie. Une permutation des vecteurs de base suffit à changer le signe du déterminant. C'est donc le cas pour le produit mixte de trois vecteurs puisqu'il est défini par un déterminant.

Espaces vectoriels

Un espace vectoriel est un **ensemble d'objets, appelés vecteurs, que l'on peut additionner entre eux, et que l'on peut multiplier par un scalaire** (pour les étirer ou les rétrécir, les tourner etc.). Autrement dit, c'est un **ensemble muni d'une structure permettant d'effectuer des combinaisons linéaires**.

Combinaisons linéaires

Une combinaison linéaire est une **expression construite à partir d'un ensemble de termes en multipliant chaque terme par une constante et en ajoutant le résultat**. Par exemple, une combinaison linéaire de x et y serait une expression de la forme ax + by où a et b sont des constantes (grandeur dont la valeur est fixée par convention ou par calcul, indépendamment du problème dans lequel elle est rencontrée ou bien nombre qui n'affecte aucune variable ou qui affecte un monôme d'ordre zéro dans une expression mathématique).

Transformation linéaires

Aussi appelée application linéaire ou opérateur linéaire, il s'agit d'une **application entre deux espaces vectoriels qui respecte l'addition des vecteurs et la multiplication des scalaires**, et préserve ainsi plus généralement les combinaisons linéaires.
L'expression peut s'utiliser aussi pour un morphisme (relative similitude d'objets considérés du point de vue de ce qu'ils partagent comme entités ou par leur relations) entre deux modules sur un anneau, avec une présentation semblable en dehors des notions de base et de dimension.

Cette notion étend celle de fonction linéaire et analyse réelle à des espaces vectoriels plus généraux.

Systèmes d'équations linéaires

Il s'agit d'un système d'équations constitué d'équations linéaires qui portent sur les mêmes inconnues.

Equations linéaires

Une équation à coefficients réels ou complexes est dite linéaire quand elle **peut être présentée sous la forme ax = b**

Vecteur

En mathématiques et en physique, un vecteur est un objet qui possède trois caractéristiques principales :

- Une **direction**
- Un **sens**
- Une **norme** (grandeur)

Il est généralement représenté par une flèche dont la longueur est sa norme.
Ils sont utilisés pour **modéliser des grandeurs** telles que les forces, les vitesses et les accélérations.
En géométrie, un vecteur peut **représenter la translation d'un point à un autre**.
En informatique, un vecteur, également nommé tableau dynamique ou table dynamique, désigne un **conteneur d'éléments ordonnés et accessibles par des indices, dont la taille est dynamique** : elle est mise à jour automatiquement lors d'ajouts ou de suppressions d'éléments.

Matrice

En mathématiques, une matrice est un **tableau rectangulaire de nombres, organisé en lignes et colonnes**.
Elles sont utilisées pour **représenter des données, résoudre des systèmes d'équations linéaires et effectuer des transformations géométriques.**
Les caractéristiques principales d'une matrice sont :

- Sa **Dimension** : (m x n), (m) lignes et (n) colonnes
- Ses **éléments / entrées / termes** : (a_{ij}) ou (i) représente le numéro de la ligne et (j) celui de la colonne.

Dimension

La dimension d'une matrice indique le **nombre de ses lignes et le nombre de ses colonnes dans cet ordre**. Par exemple, une matrice composée de 3 lignes et 5 colonnes est de dimension 3 par 5.

Types de matrices

Matrice carrée

Matrice ayant une dimension de (n x n), soit possédant le **même nombre de lignes et de colonnes**.

Matrice rectangulaires

Matrice ayant une dimension de (m x n), soit possédant un **nombre différent de lignes et de colonnes**.

Matrice diagonale

Matrice carrée ou tous les éléments en dehors de la diagonale principale sont nuls.
La diagonale principale est composée des éléments (a_{ij}).

Matrice triangulaires

Matrice carrée ou tous les éléments au-dessus (triangulaire inférieure) ou en dessous (triangulaire supérieure) de la diagonale principale sont nuls.

Matrice symétrique

Matrice carrée qui est égale à sa transposée, c'est-à-dire (A = A^T).
Cela signifie que (a_{ij} = a_{ji}) pour tout (i) et (j).

Matrice identité

Matrice carrée ou tous les éléments de la diagonale principale sont égaux à 1 et tous les autres sont nuls.
Elle est notée (I).

Matrice nulle

Matrice ou tous les éléments sont égaux à zéro.
Elle peut être de **n'importe quelle dimension**.

Matrice creuse (sparse)

Matrice dans laquelle la majorité des éléments sont nuls.
Elles sont couramment utilisées en informatique pour économiser de l'espace mémoire.

Matrice de permutation

Matrice carrée obtenue en permutant les lignes ou les colonnes de la matrice identité.
Elles sont utilisées pour représenter des permutations.

Il existe d'autres types de matrices, orthogonales, de Vandermonde, de Toeplitz, de Hankel…

Opérations sur les matrices

Addition

Pour additionner deux matrices, elles doivent avoir les **mêmes dimensions**.
On additionne les éléments correspondants de chaque matrice.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}$$

Multiplication par un scalaire

Pour multiplier une matrice par un scalaire, **on multiplie chaque élément de la matrice par ce scalaire**.

$$k \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} k \cdot a & k \cdot b \\ k \cdot c & k \cdot d \end{pmatrix}$$

Multiplication de matrices

Pour multiplier deux matrices, **le nombre de colonnes de la première matrice doit être égal au nombre de lignes de la deuxième matrice**.
Le produit de deux matrices (A) et (B) est une nouvelles matrice (C) où chaque éléments (a_{ij}) est calculé comme suit :

$$C = A \cdot B$$
$$c_{ij} = \sum_k a_{ik} \cdot b_{kj}$$

Par exemple, pour multiplier les matrices suivantes :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a \cdot e + b \cdot g & a \cdot f + b \cdot h \\ c \cdot e + d \cdot g & c \cdot f + d \cdot h \end{pmatrix}$$

Matrice inversible

Une matrice inversible est une **matrice carrée (A) pour laquelle il existe une matrice (B) de même taille n avec laquelle les produits (AB) et (BA) sont égaux à la matrice identité (I)**.
AB = BA = I_n

Dans ce cas la matrice B est unique, appelée matrice inverse de A notée B = A^{-1}
Il existe plusieurs solutions permettant de vérifier si une matrice carrée (M) est inversible (ou possède un inverse M^{-1}).
Par exemple si son déterminant △(M) ≠ 0.

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},\ det(M) = (a \times d) - (c \times b) \neq 0$$

On définit alors M^{-1} comme suit :

$$M^{-1} = \frac{1}{det(M)} \times \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Les différents types d'analyses

L'analyse est l'**opération intellectuelle consistant à décomposer un tout en ses éléments constituants afin d'en établir des relations**.
Chaque type d'analyse a son propre rôle crucial dans la prise de décision et la compréhension des données.

Descriptive (Qu'est-ce qui s'est passé ?)

- **But : Résumer et décrire** les caractéristiques principales des données.
- **Exemples :** Statistiques telles que les moyennes, médianes, écart-types, fréquences, pourcentages.
- **Utilisation :** Fournir un aperçu rapide de ce qui s'est passé ou de l'état actuel.
- **Application :** Rapports de performance, récapitulatifs de ventes, sondages d'opinion.

Diagnostique (Pourquoi cela s'est-il passé ?)

- **But : Comprendre** les causes ou les facteurs sous-jacents d'un phénomène.
- **Exemples :** Analyse des causes profondes, corrélations, tests statistiques.
- **Utilisation :** Identifier pourquoi quelque chose s'est produit.
- **Application :** Enquêtes sur les causes d'une baisse de ventes, analyse des incidents de sécurité, étude de marché.

Prédictive (Que va-t'il se passer ?)

- **But** : **Prévoir** les tendances ou les résultats futurs.
- **Exemples :** Modèles de régression, algorithmes d'apprentissage automatique, séries temporelles.
- **Utilisation :** Prédire ce qui est susceptible de se produire.
- **Application :** Prévisions de ventes, analyse de risque, scoring de crédit, prévisions météorologiques.

Prescriptive (Que devrions-nous faire ?)

- **But : Recommander** des actions pour atteindre des résultats souhaités.
- **Exemples :** Algorithmes d'optimisation, simulations, scénarios de décision.
- **Utilisation :** Conseiller ce qui doit être fait et pourquoi.
- **Application :** Planification stratégique, recommandation de produits, gestion de la chaîne d'approvisionnement.

Approches analytiques

Théorique

- **Définition :** L'analyse théorique **repose sur des concepts, des principes et des modèles abstraits pour comprendre et prédire des phénomènes**. Elle utilise les mathématiques, la logique et les hypothèses pour formuler des explications.
- **Exemple :** Utiliser des équations de la physique pour prédire le mouvement d'un objet, modéliser le comportement du climat avec des équations différentielles.
- **Nature :** Abstraite et souvent basée sur des hypothèses idéales.

Expérimentale

- **Définition :** L'analyse expérimentale **repose sur des expériences pratiques et des observations empiriques pour comprendre des phénomènes**. Elle implique la collecte et l'analyse de données par le biais d'expérimentations contrôlées.
- **Exemple :** Réaliser des expériences de laboratoire pour observer la réaction chimique, tester un nouveau médicament sur un groupe pour évaluer le coup bénéfice/risque.
- **Nature :** Concrète et basée sur des données réelles.

Différences clés :

- **Base de travail :**
  - **Théorique :** Modèles et formules abstraites.
  - **Expérimentale :** Observations et données empiriques.
- **Fiabilité :**
  - **Théorique :** Peut être limitée par les simplifications et approximations.
  - **Expérimentale :** Dépend de la précision et de la rigueur des expérimentations.
- **Objectif :**
  - **Théorique :** Comprendre les principes fondamentaux et faire des prédictions générales.
  - **Expérimentale :** Vérifier les hypothèses théoriques et observer des résultats concrets.

Notions de métrologie

La métrologie est la **discipline scientifique de la mesure**. Elle définit les principes et méthodes permettant de **garantir et maintenir la confiance** envers les mesures résultant des processus de mesure.

Valeur vraie

Valeur effective d'une grandeur que l'on cherche à mesurer (mesurande), **par essence impossible à connaître exactement** vu les **incertitudes inévitables**. Un bon système de mesure fournit un résultat plus proche de la valeur vraie.

Incertitudes

Manque ou absence de précision, description partielle d'un objet ou d'un événement donnant lieu à des cas limites, des zones troubles.
Plus précisément, l'imprécision d'un système de mesure qualifie la **dispersion des résultats** qu'il produit. L'imprécision est une source d'incertitude.
Il s'agit d'une notion traduisant un **état de connaissance incomplet** à propos d'un état du monde, passé, présent ou futur.
En particulier, on peut concevoir l'incertitude comme " **tout écart par rapport à la norme idéale et inatteignable d'une connaissance purement déterministe d'un système** ", c'est à dire toute divergence d'une description exhaustive d'un système, permettant d'en prédire la trajectoire d'évolution, de façon unique et certaine.

Relatives

L'incertitude relative est exprimée en pourcentage ou en fraction de la valeur mesurée. Elle **donne une idée de la proportion de l'incertitude par rapport à la valeur totale**.
Par exemple, en mesurant une longueur de 100 cm avec un incertitude de $\pm$ 2 cm, l'incertitude relative est donc de (2/100) * 100% = 2%.
Elle est **utile pour comparer la précision** de différentes mesures de différentes tailles ou unités.

Absolues

L'incertitude absolue est exprimée dans les mêmes unités que la valeur mesurée. Elle **indique la quantité de variation ou d'erreur attendue dans la mesure**.
Par exemple, en mesurant une longueur de 100 cm avec un incertitude de $\pm$ 2 cm, l'incertitude absolue est simplement $\pm$ 2 cm.
Elle **donne une idée directe et concrète de l'erreur ou de la variation** dans une mesure spécifique.

Différences clés :

L'**incertitude relative** fournit une **perspective normalisée**, facilitant les comparaisons entre différentes mesures.
L'**incertitude absolue** donne une **indication directe** de l'erreur potentielle.
L'une **donne une idée de l'ampleur proportionnelle de l'incertitude**, l'autre **montre l'erreur absolue potentielle**.

Type d'incertitude	Nature	Exemple	Caractéristique	Relative ou Absolue	Type
Statistique	Variabilité des données	Marges d'erreur des sondages	Quantifiable, peut être réduite avec plus de données	Relative	A
Épistémiques	Limitation des connaissances	Modèles scientifiques incomplets	Dépendante de l'état des connaissances	Relative	A
Aléatoire	Variabilité intrinsèque	Tirage aléatoire	Décrite par des distributions de probabilité	Relative	A
Ontologique	Complexité inhérente	Système climatique	Inhérente à la nature des choses	Absolue	B
Modélisation	Représentation des systèmes complexes	Modèles climatiques	Améliorée par de meilleurs modèles	Relative	A
Subjective	Jugement et perception humaine	Décision médicale, biais de lecture…	Difficile à quantifier	Relative	A
Méthodologique	Technique et approche méthodologiques	Collecte de données	Dépendante de l'amélioration des techniques	Relative	A
Prédictive	Prévisions et projections futures	Prévisions météorologiques	Dépendante de la précision des modèles	Relative	A
Spécification	Définition du problème	Modèles économiques	Dépendantes de la clarté de la spécifications	Relative	A
Données	Qualité et précision des données	Données manquantes ou erronées	Dépendante de la qualité des données	Relative	A

Petites notions fondamentales d’épistémologie

Biais

- **Définition :** Le biais est une **déviation systématique des résultats ou des prévisions par rapport à la réalité**. Cela provient souvent de préférences ou de méthodes inappropriées.
- **Exemple :** Biais d’échantillonnage ou certaines populations sont sous-représentées dans une étude.
- **Nature :** Peut être conscient ou inconscient, introduit par des préjugés personnels, des erreurs méthodologiques, ou des influences contextuelles.
- **Impact :** Peut conduire à des conclusions incorrectes ou trompeuses.

Bruit

- **Définition :** Le bruit est une **variation aléatoire ou une erreur sans direction systématique qui affecte les résultats de manière imprévisible**.
- **Exemple :** Fluctuation dans les mesures de température dues à des interférences environnementales.
- **Nature :** Non systématique, introduit par des variations aléatoires ou des imperfections dans les processus de mesure.
- **Impact :** Rend les données plus difficiles à interpréter en masquant les véritables tendances ou signaux.

Biais vs Bruit

- **Biais : Systématique**, **directionnel**, souvent lié à des **erreurs méthodologiques** ou des **préjugés**.
- **Bruit** : **Aléatoire**, **non directionnel**, lié à des **variations imprévisibles** ou des **imperfections de mesure**.

Epistémologie

- **Définition :** L’épistémologie est la **branche de la philosophie qui étudie la nature, l’origine et la portée de la connaissance**.
- **Lien avec biais et bruit :** L’épistémologie **s’intéresse à la manière dont nous acquérons des connaissances et comment le biais et le bruit peuvent affecter la validité et la fiabilité de cette connaissance**. Comprendre et gérer le biais et le bruit est une partie intégrante de l’épistémologie.

Probabilité et loi de probabilité

Probabilité

- **Définition :** La probabilité est une **mesure mathématique de la vraisemblance qu'un événement se produise**. Elle est représentée par un **nombre compris entre 0 et 1**, où 0 indique que l'événement ne se produira jamais et 1 qu'il se produira toujours.
- **Exemple :** La probabilité de tirer un as d'un jeu de cartes standard (52 cartes) est de 4/42, soit environ 0.077 ou 7.7%.

Loi de probabilité

- **Définition :** Une loi de probabilité (ou distribution de probabilité) décrit **comment les probabilités sont distribuées sur les différentes valeurs possibles d'une variable aléatoire**. Elle permet de **connaître la probabilité de chaque événement**.

Loi Normale (Distribution Normale) :

- **Définition :** Également connue sous le nom de distribution gaussienne, c'est une **distribution continue en forme de cloche symétrique autour de la moyenne.** Elle décrit **comment les valeurs d'une variable se répartissent de manière homogène autour de la moyenne**.
- **Caractéristiques :**
    - La courbe est symétrique par rapport à la moyenne.
    - Environ 68% des valeurs se trouvent à un écart-type de la moyenne, 95% à deux écart-types et 99.7% à trois écarts-types.
    - Utilisée pour modéliser de nombreux phénomènes naturels comme la taille, les notes d'examens…

Loi Binomiale :

- **Définition :** Il s'agit d'une **distribution discrète qui décrit le nombre de succès dans une séquence d'expériences indépendantes de Bernoulli identiques**, où chaque expérience a **deux résultats possibles** (succès ou échec).
- **Caractéristiques :**
    - Définie par deux paramètres : le nombre d'essais (n) et la probabilité de succès de chaque essai (p).
    - La probabilité de K succès dans n essais est donnée par la formule de probabilité binomiale.

En résumé

- **Probabilité : mesure la vraisemblance** d'un **événement unique**.
- **Loi de probabilité** : **décrit la distribution des probabilités de tous les résultats** possibles d'une **variable aléatoire**.

Différents types de variables

Une variable est un **symbole ou un nom représentant une valeur pouvant varier**. Elle est utilisée dans divers domaines tels que les mathématiques, la programmation et les sciences pour **stocker, manipuler et référencer des valeurs dynamiques**.

En programmation

Variable entières (int)

Représente des **nombres entiers** (1, 85, -9)

Variable flottantes (float)

Représente des **nombres à virgule flottante** (0.356, -5.16, 0.0001)

Variable de caractères (char)

Représente **un seul caractère** (‘a’, ‘Z’)

Variable de chaîne (string)

Présente une **séquence de caractères** (‘Bonjour’, ‘Variable’)

Variable booléennes (bool)

Représente des **valeurs vraies ou fausses** (True, False)

En statistique et mathématiques

Variables qualitatives

Classifient les données en **catégories distinctes** et **décrivent des qualités ou des caractéristiques** et ne sont **généralement pas numériques**.

Variable nominales

Variables qualitatives qui **catégorisent les données sans ordre particulier**. Elles permettent de **distinguer des groupes ou catégories sans hiérarchie entre elles**.

Variable ordinales

Variables qualitatives qui **classent les données dans un ordre ou une hiérarchie**. Elles permettent de **voir une progression ou un rang entre les catégories**.

Variables quantitatives

Mesurent une **quantité** et sont **numériques**.

Variable discrètes

Prend des **valeurs distinctes**. (nombre d'enfants dans une famille…)

Variable continues

Peut prendre n'importe quelle **valeur dans un intervalle**. (Taille en centimètre…)

Variables dépendantes

Dépend d'**une ou plusieurs autres variables** (y dans y = 2x = 3)

Variables indépendantes

Peut être **modifiée indépendamment** (x dans y = 2x +3)

Différents types de corrélations

La corrélation est une **mesure statistique** qui **indique la force et la direction d'une relation entre deux variables**. Elle **quantifie la mesure** dans laquelle les variations sont associées aux variations d'une autre variable.

Corrélation positive

Lorsque **deux variables augmentent ou diminuent ensemble**.

Corrélation négative

Lorsque **l'augmentation d'une variable correspond à la diminution de l'autre**.

Corrélation nulle

Lorsque les **variations d'une variable** ne sont **pas associées aux variations de l'autre**.

Corrélation linéaire

Indique une **relation proportionnelle entre deux variables**, représentable par une **ligne droite dans un graphique**. Elle peut être **positive ou négative**.

Coefficient de corrélation

Le **coefficient de corrélation** noté *r* (souvent calculé avec la formule de corrélation de Pearson pour les relations linéaires) **varient entre -1 et 1**.

- r = 1 : **Corrélation parfaite**.
- r = -1 : **Corrélation négative parfaite**.
- r = 0 **Aucune corrélation**.

Corrélation non linéaire

Indique une **relation entre deux variables** qui **ne peut pas être** représentée par **une ligne droite**. Elle **peut prendre de nombreuses formes différentes** (courbes …)

Coefficient de corrélation

Méthode comme la corrélation de Spearman ou de Kendall…

Corrélation partielle

Mesure la **force et la direction de la relation entre deux variables tout en contrôlant l'effet d'une ou plusieurs autres variables**.

Statistique descriptives

Définition

Les statistiques descriptives sont des outils mathématiques et statistiques qui permettent de **résumer et de décrire les caractéristiques principales d'un ensemble de données**.
Elles fournissent une **vue d'ensemble des données** en termes de tendances centrales, de dispersion et de forme.

Mesures de tendance centrale

Moyenne

La **somme de toutes les valeurs divisée par le nombre de valeurs**.

Médiane

Valeur centrale qui sépare la moitié supérieure de la moitié inférieure des données.

Mode

Valeur apparaissant le plus fréquemment dans un ensemble de données.

Mesures de position et dispersion

Maximum

La **plus grande valeur** d'un ensemble de données.

Minimum

La **plus petite valeur** d'un ensemble de données.

Quartiles

Valeur divisant un ensemble de données en **quatre parties égales**.

- **Premier quartile (Q1)** : **25%** des données sont en dessous de cette valeur.
- **Médiane (Q2**) : **50%** des données sont en dessous de cette valeur.
- **Troisième quartile (Q3)** : **75%** des données sont en dessous de cette valeur.

IQR (Intervalle interquartile)

Mesure la **dispersion des valeurs** au sein d'un ensemble de données.
(IQR) = (Q3) - (Q1)

Mesures de dispersion

Variance

La **moyenne des carrés des écarts par rapport à la moyenne**.

Ecart-type

Racine carrée de la variance.

Mesures de Forme

Asymétrie

Mesure la **symétrie des données par rapport à la moyenne**.

Kurtosis

Mesure l'**aplatissement des données**.

Représentation graphiques

Boxplot

Diagramme en boîte **représentant la distribution des données à travers leurs quartiles**. Il **montre la médiane**, **les quartiles**, ainsi que **les valeurs extrêmes** et les éventuels outliers.
Très utile pour **visualiser la dispersion des données** et **identifier des valeurs aberrantes**.

Histogramme

Graphique qui représente la **distribution d'un ensemble de donnée**s en utilisant des barres pour **montrer la fréquence des valeurs dans différents intervalles** (bins).
Très utile pour **visualiser la forme de la distribution des données**, **comprendre la densité des valeurs et repérer les tendances et les écarts**.

Théorème et Concepts Mathématiques

Espérance

L'espérance d'une variable aléatoire est une **mesure de la moyenne de ses valeurs possibles**, pondérées par leurs probabilités respectives.
C'est la **valeur que l'on attend en moyenne après un grand nombre de répétitions** de l'expérience aléatoire.
Par exemple, si on lance un dé à 6 faces, l'espérance des résultats est (1+2+3+4+5+6)/6 = 3.5.

Théorème central limite

Ce théorème stipule que, **pour un échantillon de taille suffisante, la distribution de la somme (ou de la moyenne) des variables aléatoires indépendantes et identiquement distribuées s'approche d'une distribution normale**, quelle que soit la forme de la distribution initiale des variables.
Il **justifie l'utilisation de la distribution normale** dans de nombreuses situations pratiques.

Dérivée

La dérive d'une fonction **mesure la vitesse à laquelle la fonction change par rapport à une variable indépendante**.
C'est une notion fondamentale en calcul différentiel.