DATA Camp M1

Mathématique

Algèbre Linéaire

Rappelons les notions essentielles à connaître sur les matrices.

Trace d'une matrice

La **trace** d'une matrice carrée A est la somme de ses éléments diagonaux :

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$$

Propriétés : - La trace est linéaire :

$$\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$$

- Invariance par similitude :

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

> **Note**
>
> La trace d'une matrice ne change pas lorsqu'on change de base.

Déterminant d'une matrice

Le déterminant d'une matrice $A \in \mathcal{M}_{n\times n}$ peut être calculé en utilisant le développement par rapport à une ligne ou une colonne :

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

où (A_{ij}) est la matrice obtenue en supprimant la (i)-ème ligne et la (j)-ème colonne de (A).

> **Tip**
>
> Dans certains cas, il peut être long de calculer le déterminant d'une matrice. Pour remédier à cela, il est utile de connaître par coeur le calcul d'un déterminant d'une matrice de taille 2×2 et d'une matrice diagonale.

Déterminant d'une matrice (2×2)

Pour une matrice $A \in \mathcal{M}_{2\times 2}$ de la forme :

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

le déterminant se calcule en utilisant la formule suivante :

$$\det(A) = ad - bc$$

Déterminant d'une matrice diagonale

Pour une matrice diagonale de la forme :

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

le déterminant est le produit des éléments diagonaux :

$$\det(D) = d_1 \times d_2 \times \cdots \times d_n$$

Les déterminants sont particulièrement utiles en algèbre linéaire pour déterminer si une matrice est inversible (une matrice est inversible si et seulement si son déterminant est non nul) et pour résoudre des systèmes d'équations linéaires. En algèbre, on définit un espace où les vecteurs

appartiennent et on utilise le plus souvent un espace euclidien. C'est un espace où est rattaché une distance qu'on appelle produit scalaire. ### Espace euclidien

Un **espace euclidien** est un espace vectoriel réel de dimension finie muni d'une forme bilinéaire symétrique définie positive, appelée **produit scalaire** :

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Quelques propriétés importantes du produit scalaire : - $\langle v, v \rangle = \|v\|_2^2 = \sum_{i=1}^{n} v_i^2$ - $\langle v, w \rangle = v^t w = \text{trace}(vw^t)$ - $\langle v, w \rangle = \langle w, v \rangle$ - $\langle v, w \rangle = 0 \iff v \perp w$

Quelques inégalités à connaître :

- Inégalité de Cauchy-Schwarz :
$$|\langle x, y \rangle| \leq \|x\| \|y\|$$
- Inégalité triangulaire :
$$\|x + y\| \leq \|x\| + \|y\|$$

Orthogonalité

Soient v et w, deux vecteurs appartenant à un espace vectoriel E de dimension n, muni du produit scalaire $\langle \cdot, \cdot \rangle$ associé à la norme euclidienne $\| \cdot \|_2$.On note $v = (v_1, \ldots, v_n)^t$ et $w = (w_1, \ldots, w_n)^t$ où $ v_i$ et w_j sont des scalaires (i.e. des réels).

Normalisation d'un vecteur

La normalisation d'un vecteur v est donnée par :

$$x = \frac{v}{\|v\|_2}$$

Ainsi, $\|x\|_2 = 1$. Un espace vectoriel engendré par $v_1, \ldots, v_p$ est identique à celui engendré par leurs versions normalisées.

Deux vecteurs x et y sont **orthogonaux** si leur produit scalaire est nul :

$$\langle x, y \rangle = 0$$

Une **base orthonormale** est une base où tous les vecteurs sont orthogonaux et de norme 1.

Matrice de Projection Orthogonale

Soit E un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ et W un sous-espace vectoriel de E. La projection orthogonale d'un élément B de E sur W est définie par :

$$\Pi_W B = \underset{a \in W}{argmin} \|B - a\|_2$$

La matrice de projection orthogonale de E sur W est notée Π_W.

Propriétés

- $\Pi_W^t = \Pi_W$
- $\Pi_W^2 = \Pi_W$
- $\text{trace}(\Pi_W) = \dim(\Pi_W)$
- $\Pi_W^\perp = I_E - \Pi_W$

Pour tout B, élément de E :

$$\Pi_{W^\perp} B = B - \Pi_W B$$

Construction de matrice de projection orthogonale

Soient $v_1, \dots, v_p$, p vecteurs de $\mathbb{R}^n$ avec $p \leq n$. Si $(v_1, \dots, v_p)$ est une base orthogonale de $\mathbb{R}^p$ et $W = \text{vect}(v_1, \dots, v_p)$, alors la matrice de projection W de $\mathbb{R}^n$ sur W est :

$$W = V(V^t V)^{-1} V^t$$

où V est la matrice dont les colonnes sont les vecteurs $v_1, \dots, v_p$.

Si $(v_1, \dots, v_p)$ est une base orthonormale de W, alors :

$$W = VV^t$$

Procédé d'orthogonalisation de Gram-Schmidt

Le but du procédé de Gram-Schmidt est de prendre un ensemble de vecteurs linéairement indépendants $(v_1, v_2, \dots, v_n)$ et de produire un ensemble de vecteurs orthogonaux $(u_1, u_2, \dots, u_n)$ qui engendrent le même sous-espace.

> **Tip**
>
> Avant de recourir au procédé de Gram-Schmidt, il faut s'assurer que les vecteurs ne soient pas déjà orthogonaux, cela serait une perte de temps de les recalculer. Ce calcul est long, il faut bien prendre son temps pour le faire et ne pas se précipiter.

Étapes du Procédé

Le premier vecteur orthogonal (u_1) est le premier vecteur de l'ensemble original :

$$u_1 = v_1$$

Pour chaque vecteur (v_k) (où $(k \geq 2)$), on soustrait les projections orthogonales de (v_k) sur les vecteurs orthogonaux précédemment calculés $(u_1, u_2, \ldots, u_{k-1})$:

$$u_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, u_j \rangle}{\langle u_j, u_j \rangle} u_j$$

Si l'on souhaite obtenir une base orthonormale, chaque vecteur (u_k) est normalisé pour obtenir (e_k) :

$$e_k = \frac{u_k}{||u_k||}$$

Base duale

Pour une base $\{e_1, e_2, \ldots, e_n\}$ d'un espace vectoriel E, la **base duale** $(e^1, e^2, \ldots, e^n)$ est définie par :

$$e^i(e_j) = \delta_{ij}$$

où δ_{ij} est le symbole de Kronecker.

La base duale permet de définir des formes linéaires et de travailler avec des espaces vectoriels de manière plus abstraite.

Diagonalisation d'une matrice

Une matrice carrée A est dite **diagonalisable** s'il existe une matrice diagonale D et une matrice inversible P telles que :

$$A = PDP^{-1}$$

D est une matrice diagonale contenant les valeurs propres de A

Diagonalisation d'une matrice symétrique réelle

Pour les matrices symétriques réelles, on peut toujours trouver une base orthonormale de vecteurs propres, ce qui permet de les diagonaliser par une matrice orthogonale Q :

$$A = QDQ^T$$

où Q est une matrice orthogonale.

> **Warning**
>
> Pour que la propriété soit vérifiée, il est essentiel que la matrice soit réelle, c'est-à-dire sans composantes complexes. En effet, la symétrie d'une matrice complexe n'implique pas nécessairement la propriété en question

Statistiques Descriptives Univariées

- **Moyenne** : La moyenne arithmétique d'une série de valeurs $x_1, x_2, \ldots, x_n$ est donnée par :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

  En théorie, on représente la moyenne comme l'espérance : Dans le cas discret :

$$E(X) = \sum k \cdot P(X = k)$$

$$k \in X(\Omega)$$

  Alors que dans le cas continu :

$$E(X) = \int x \cdot f_X(x)\, dx$$

$$x \in X(\Omega)$$

- **Médiane** : La médiane est la valeur qui sépare la série en deux parties égales. Pour une série ordonnée, si n est impair, la médiane est la valeur centrale. Si n est pair, c'est la moyenne des deux valeurs centrales.
- **Variance** : La variance mesure la dispersion des valeurs autour de la moyenne :

$$\text{Var}(X) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

  ou en théorie

$$Var(X) = E(X^2) - (E(X))^2 = E((x - E(X))^2)$$

- **Écart-type** : L'écart-type est la racine carrée de la variance :

$$\sigma = \sqrt{\text{Var}(X)}$$

> **Note**
>
> Plus l'écart-type est grand, plus les données sont dispersées autour de la moyenne.

Les statistiques descriptives univariées sont essentielles pour résumer et comprendre les caractéristiques principales d'une seule variable. Elles sont largement utilisées en analyse de données pour obtenir une vue d'ensemble rapide et efficace. Elles permettent aussi d'identifier rapidement des valeurs extrêmes.

Statistiques Descriptives Bivariées

Définitions

- **Covariance** : La covariance entre deux variables X et Y mesure la manière dont deux variables varient ensemble.

1. Pour un échantillon de données :

$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

   où x_i et y_i sont les valeurs des échantillons, et $\bar{x}$ et $\bar{y}$ sont les moyennes des échantillons.

2. Pour des variables aléatoires :

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

   où $\mathbb{E}[X]$ et $\mathbb{E}[Y]$ sont les espérances (moyennes théoriques) des variables aléatoires X et Y.

- **Régression linéaire**:

La régression linéaire est une méthode statistique qui permet de modéliser la relation entre deux variables en ajustant une ligne droite à un ensemble de données. L'équation de la régression linéaire est :

$$y = a + bx$$

où y est la variable dépendante, x est la variable indépendante, a est l'ordonnée à l'origine (l'interception de la ligne avec l'axe des ordonnées), et b est la pente de la ligne (le taux de changement de y par rapport à x).

Probabilités

Soit E un ensemble muni d'une tribu T. On appelle probabilité toute application $P : T \to \mathbb{R}^+$ telle que : - $P(\emptyset) = 0$ - Si (A_n) est une suite d'éléments de T deux à deux disjoints alors :

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n).$$

Probabilités conditionnelles

En théorie des probabilités, nous nous intéressons souvent au comportement d'un aléa, sachant qu'un autre événement est déjà passé. C'est ce que nous appelons Les Probabilités Conditionnelles.

Considérant deux événements de probabilité non nulle, A et B, la probabilité conditionnelle de A sachant que B est réalisé (couramment dit A sachant B) est donnée par :

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

Par commutativité de l'intersection, nous avons :

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

En utilisant la formule ci-dessus, nous pouvons également exprimer la probabilité conditionnelle de B sachant A :

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

C'est ce que nous appelons la formule de Bayes.

Indépendance

Deux événements A et B sont dits indépendants si et seulement si :

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

En termes courants, deux événements sont indépendants si le résultat de l'un n'influence aucunement l'aboutissement de l'autre. Sous condition d'indépendance de A et B, la notion de la probabilité conditionnelle tombe à l'eau, car les événements évoluent l'un sans se soucier de l'autre. Ceci se traduit par :

$$P(A|B) = P(A)$$
$$P(B|A) = P(B)$$

Notons que si A est indépendant de B, il le sera par rapport à son complémentaire également, et vice versa. En général, pour une suite (A_n) d'événements indépendants, on a :

$$P\left(\bigcap A_i\right) = \prod P(A_i) = P(A_1) \cdot ... \cdot P(A_n)$$

Cette formule est largement utilisée en statistique.

> ⚠ Warning
>
> Ne pas confondre l'indépendance et l'incompatibilité des évènements

Variable aléatoire

Une variable aléatoire est un nombre qui dépend du résultat d'une expérience aléatoire. Chaque exécution de l'expérience génère une réalisation de la variable aléatoire.

Mathématiquement, on définit une variable aléatoire X comme une fonction $X : T \to \mathbb{R}$ qui associe à chaque événement s, un réel $X(s)$.

Par exemple, dans une queue pour la caisse d'un magasin, le nombre de clients est une variable aléatoire. La durée de traitement de chaque requête aussi. Remarquons que la première est un nombre entier. On dit qu'elle est à support discret. Alors que la deuxième est une durée (un nombre réel). On dit qu'elle est à support continu.

Qu'est- ce qui caractérise une variable aléatoire ?

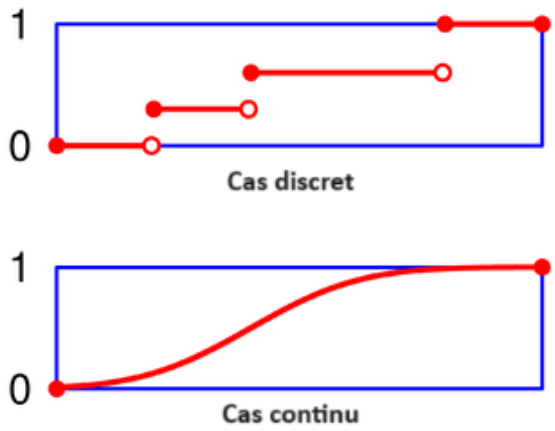
Fonction de répartition

Une variable aléatoire traduit le résultat d'une expérience aléatoire en nombre réel. La fonction de répartition transporte le calcul des probabilités concernant les réalisations de la variable aléatoire. C'est la fonction définie par :

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

Propriétés :

Pour tout x, $0 \leq F_X(x) \leq 1$ F_X est une fonction croissante. $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to \infty} F_X(x) = 1$



Probabilité ponctuelle / Densité

CAS DISCRET : Probabilité ponctuelle

La probabilité ponctuelle est la fonction qui décrit les sauts de la fonction de répartition :

$$P(X = k) = P(X \leq k) - P(X \leq k - 1) = p_k$$

$$\sum p_i = 1$$

CAS CONTINU : densité de probabilité La densité est la fonction qui décrit les variations de la fonction de répartition :

$$f(x) = \frac{dF_X}{dx}(x)$$

$$\int f(x)dx = 1$$

Fondements de probabilités : niveau basique

Centrage et réduction

Le centrage consiste à localiser la distribution autour de l'origine et la réduction consiste à normaliser la dispersion. La technique est simple :

$$Y = \frac{X - E(X)}{\sigma(X)}$$

Moments d'ordre r

Le moment d'ordre r est défini par :

$$\mu_r = E(X^r)$$

Le moment centré d'ordre r est défini ainsi :

$$\mu_r = E((X - E(X))^r)$$

Couples aléatoires

La fonction conjointe

$$F_{X,Y}(x, y) = P(X \leq x \cap Y \leq y)$$

est appelée la distribution conjointe de X et Y.

Dans le cas continu, la fonction définie par :

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}$$

est la densité conjointe du couple (X, Y). On a donc :

$$F_{X,Y}(x, y) = \int \int f_{X,Y}(t, u)\, dt\, du, \text{ où } -\infty < x, y < +\infty,$$

Dans le cas discret, on définit la fonction de probabilité conjointe :

$$P(X = x_i, Y = y_j) = p_{ij}$$

On a donc :

$$F_{X,Y}(x, y) = \sum \sum p_{ij}, \text{ où } x_i \leq x \text{ et } y_j \leq y$$

Loi mardinale

La loi marginale de X est définie comme suit :

$$f_X(x) = \int f_{X,Y}(x,y)\,dy, \text{ où } -\infty < x < \infty,$$

dans le cas continu, ou encore :

$$f_X(x_i) = \sum p_{ij}, \text{ où } j \text{ tel que } y_j \leq y$$

Si X et Y sont indépendants, alors :

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

COVARIANCE

La covariance que nous avons traité plus haut dans les statistiques bivariées mesure l'intensité de la relation linéaire entre deux variables aléatoires X et Y. Elle est définie comme suit :

$$Cov(X,Y) = E(XY) - E(X) \cdot E(Y)$$

Si X et Y sont indépendants, alors :

$$Cov(X,Y) = 0$$

> **Warning**
>
> Il est important de noter que la réciproque n'est pas vraie : la covariance n'implique pas nécessairement l'indépendance entre X et Y.

Propriétés

ESPÉRANCE

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$$

$$\mathbb{E}(a) = a$$

VARIANCE

$$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$$

$$\mathrm{Var}(a) = 0$$

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$$

$$\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X, Y)$$

COVARIANCE

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$$

$$\mathrm{Cov}(aX + b, cY + d) = ac\mathrm{Cov}(X, Y)$$

$$\mathrm{Cov}(aX + bY, U) = a\mathrm{Cov}(X, U) + b\mathrm{Cov}(Y, U)$$

$$\mathrm{Cov}(X, cU + dV) = c\mathrm{Cov}(X, U) + d\mathrm{Cov}(X, V)$$

$$\mathrm{Cov}(aX + bY, cU + dV) = ac\mathrm{Cov}(X, U) + ad\mathrm{Cov}(X, V) + bc\mathrm{Cov}(Y, U) + bd\mathrm{Cov}(Y, V)$$

Vecteurs aléatoires

Pour un vecteur aléatoire $(X_1, X_2, ..., X_n)$, l'espérance est toujours linéaire. Pour une suite $(a_i)_{i\in\{1,...,n\}}$ de réels, on a :

$$\mathbb{E}(a_1X_1 + a_2X_2 + ... + a_nX_n) = a_1\mathbb{E}(X_1) + a_2\mathbb{E}(X_2) + ... + a_n\mathbb{E}(X_n)$$

Si les variables aléatoires $X_1, X_2, ..., X_n$ sont indépendantes, alors la variance de leur somme est égale à la somme de leurs variances individuelles :

$$\mathrm{Var}(X_1 + X_2 + ... + X_n) = \mathrm{Var}(X_1) + ... + \mathrm{Var}(X_n)$$

Lois usuelles

Ces tableaux récapitulent les lois usuelles que vous pourrez rencontrer dans différents cours du master.

Nom	Notation	$X(\Omega)$	$P(X = k)$	$E[X]$	$V(X)$
Uniforme	$X \sim U(\{1, 2, ..., n\})$	$\{1, 2, ..., n\}$	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Bernouilli	$X \sim B(p), 0 < p < 1$	$\{0, 1\}$	$P(X = 1) = p$ $P(X = 0) = 1 - p$	p	$p(1-p)$
Binomiale	$X \sim B(n, p), 0 < p < 1$	$\{1, 2, ..., n\}$	$C_k^n p^k (1-p)^{n-k}$	np	$np(1-p)$
Géométrique	$X \sim G(p), 0 < p < 1$	$\mathbb{N}$	$p(1-p)^k$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
Poisson	$X \sim P(\lambda), \lambda > 0$	$\mathbb{N}$	$\frac{\lambda^k}{k!}e^{-\lambda}$	λ	λ

Nom	Notation	$X(\Omega)$	$f_X(x)$	$E[X]$	$V(X$
Uniforme	$X \sim U([a,b]), a<b$	$[a,b]$	$\frac{1}{b-a}1_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(a-b)}{12}$
Exponentielle	$X \sim \mathcal{E}(\lambda), \lambda > 0$ $\mathcal{E}(\lambda) = \gamma(1,\lambda)$	$\mathbb{R}^+$	$\lambda e^{-\lambda x}1_{\mathbb{R}^+}(x)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normale ou Gaussienne	$X \sim N(m,\sigma^2), \sigma > 0$	$\mathbb{R}$	$\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$	m	σ^2
Gamma	$X \sim \gamma(\alpha,\theta), \alpha>0, \theta>0$ $\Gamma(\alpha) = \int_0^\infty e^{-x}x^{\alpha-1}\,dx$	$\mathbb{R}^+$	$\frac{\theta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\theta x}1_{\mathbb{R}^+}(x)$	$\frac{\alpha}{\theta}$	$\frac{\alpha}{\theta^2}$
Khi-2	$X \sim \chi^2(n), n \in \mathbb{N}^+$ $Y_1, Y_2, \dots, Y_n$ indépendantes, $Y_i \sim \mathcal{N}(0,1), \quad X = \sum_{i=1}^n Y_i^2$	$\mathbb{R}^+$	$\gamma\left(\frac{n}{2}, \frac{1}{2}\right)$	n	$2n$
Bêta	$X \sim B(\alpha,\theta), \alpha>0, \theta>0$ $B(\alpha,\theta) = \int_0^1 x^{\alpha-1}(1-x)^{\theta-1}\,dx$ $B(\alpha,\theta) = \int_0^1 x^{\alpha-1}(1-x)^{\theta-1}\,dx$ $X = \frac{Z}{1+Z}, \quad Z \sim B'(\alpha,\theta)$	$[0,1]$	$\frac{x^{\alpha-1}(1-x)^{\theta-1}}{B(\alpha,\theta)}1_{[0,1]}(x)$	$\frac{\alpha}{\alpha+\theta}$	$\frac{\alpha\theta}{(\alpha+\theta)^2(\alpha$
Bêta (prime)	$Z \sim B'(\alpha,\theta), \alpha>0, \theta>0$ $X \sim \gamma(\alpha,1), Y \sim \gamma(\theta,1), X \perp\!\!\!\perp Y$ $Z = \frac{X}{Y}$ $B(\alpha,\theta) = \int_0^\infty \frac{x^{\alpha-1}}{(1+x)^{\alpha+\theta}}\,dx$	$\mathbb{R}^+$	$\frac{z^{\alpha-1}}{B(\alpha,\theta)\cdot(1+z)^{\alpha+\theta}} \cdot 1_{\mathbb{R}^+}(z)$	$\frac{\alpha}{\theta-1}$ $\theta > 1$	$\frac{\alpha(\alpha+\theta}{(\theta-1)^2(}$ $\theta >$
Student T	$T \sim T(n), n \in \mathbb{N}^*$ $X \sim \mathcal{N}(0,1), Y \sim \chi^2(n), X \perp\!\!\!\perp Y$ $T = \frac{X}{\sqrt{Y/n}}$ $T^2/n \sim B'(1/2, n/2)$	$\mathbb{R}$	$\frac{\left(1+\frac{t^2}{n}\right)^{-(n+\frac{1}{2})}}{\sqrt{n}\cdot B(\frac{1}{2},\frac{n}{2})}$	0	$\frac{n}{n-2}$ $n >$
Fisher	$X \sim F(n,m)$ $N \sim \chi^2(n), n \in \mathbb{N}^*$ $M \sim \chi^2(m), \quad m \in \mathbb{N}^*$ $N \perp\!\!\!\perp M, \quad X = \frac{N}{n}/\frac{M}{m}$ $\frac{n}{m}X \sim B'\left(\frac{n}{2}, \frac{m}{2}\right)$	$\mathbb{R}^+$	$\frac{Kx^{n/2-1}}{(m+nx)^{(n+m)/2}}1_{\mathbb{R}^+}(x)$ $K = \frac{n^{\frac{n}{2}}\cdot m^{\frac{m}{2}}}{B(\frac{n}{2},\frac{m}{2})}$	$\frac{m}{m-2}$ $m > 2$	$\frac{2m^2(n+}{n(m-2)^2(}$ $m >$

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\, dt : \text{ désigne la fonction Gamma d'Euler}$$

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} : \text{ désigne la fonction Bêta}$$

Nous allons souvent rencontrer les lois grisées dans les Tests statistiques. Là encore, connaître les densités ne servirait pas à grand chose, mais ceci nous évitera de parler de lois dont nous ne connaissons pas la tête.

Fondements de probabilités

Vecteurs aléatoires

INDEPENDANCE DEUX À DEUX

Les variables $X_1, \dots, X_n$ sont deux à deux indépendantes si et seulement si : $\forall i \neq j, X_i$ et X_j sont indépendantes.

INDEPENDANCE MUTUELLE

Les variables $X_1, \dots, X_n$ sont mutuellement indépendantes si et seulement si :

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

L'INDÉPENDANCE MUTUELLE IMPLIQUE L'INDÉPENDANCE DEUX À DEUX. LA RÉCIROQUE EST FAUSSE.

Si $X_1, \dots, X_n$ sont mutuellement indépendantes, alors pour toute famille de fonctions réelles f_i, on a : $(f_1(X_1), \dots, f_n(X_n))$ sont indépendantes.

ESPÉRANCE ET VARIANCE-COVARIANCE

Soit $X = (X_1, \dots, X_n)^T$ un vecteur aléatoire. Dans le cas multidimensionnel, l'espérance scalaire est remplacée par un vecteur espérance.

$$E(X) = (E(X_1), \dots, E(X_n))^T$$

La variance unidimensionnelle est remplacée par la matrice symétrique de variance-covariance. Elle contient les variances en diagonale et les covariances ailleurs. On la note généralement Σ_X.

$$\Sigma_X = \begin{bmatrix} V(X_1) & \dots & \mathrm{Cov}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n, X_1) & \dots & V(X_n) \end{bmatrix}$$

Notions de convergence

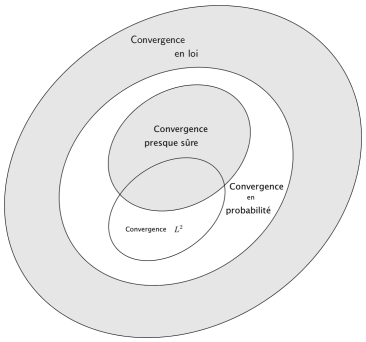
Si l'on pense à des données, vues comme réalisation de variables aléatoires $X_1, \dots, X_n$, il serait intéressant de se poser la question de savoir comment évolue cette suite lorsque n tend vers l'infini.

On dit que (X_n) converge presque sûrement vers X et on note $X_n \xrightarrow{\text{p.s.}} X$ si et seulement si : $P\left(\lim_{n \to +\infty} X_n = X\right) = 1$

On dit que (X_n) converge en probabilité vers X et on note $X_n \xrightarrow{\text{p}} X$ si et seulement si : $\forall \varepsilon > 0, \quad P(|X_n - X| > \varepsilon) \to 0$

On dit que (X_n) converge en loi vers X et on note $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si : $F_{X_n} \xrightarrow{n\to+\infty} F_X$ Où F_X dénote la fonction de répartition de X.

On dit que (X_n) converge en moyenne quadratique vers X et on note $X_n \xrightarrow{m.q.} X$ si et seulement si : $\mathbb{E}((X_n - X)^2) \to 0$ Cette définition peut se généraliser jusqu'à l'ordre n, mais nous n'en aurons pas besoin.



Loi faible des grands nombres

Soit $X_1, \ldots, X_n$ une suite de variables aléatoires indépendantes et de même loi telles que : $\mathbb{E}(X_i) = \mu$ et $\mathrm{Var}(X_i) = \sigma^2$ alors :

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu$$

Loi forte des grands nombres

Soit $X_1, \ldots, X_n$ une suite de variables aléatoires indépendantes et de même loi telles que : $\mathbb{E}(X_i) = \mu$ et $\mathrm{Var}(X_i) = \sigma^2$, alors :

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} \mu$$

Théorème Central Limite

Soit $X_1, \ldots, X_n$ une suite de variables aléatoires indépendantes et de même loi telles que : $\mathbb{E}(X_i) = \mu$ et $\mathrm{Var}(X_i) = \sigma^2$, alors :

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{\mathcal{L}oi} \mathcal{N}(0,1)$$

Statistique inférentielle : niveau basique

Echantillon / Estimateur

Le point de départ est un vecteur (ou un tableau dans le cas multidimensionnel) de données. Ces données peuvent être vues comme les réalisations $(x_1, x_2, \dots, x_n)$ d'une variable aléatoire X qui dépend d'un certain paramètre θ que nous allons chercher à estimer. Pour ce faire, nous allons construire un échantillon de cette variable. Un échantillon $(X_1, X_2, \dots, X_n)$ est un n-uplet de variables aléatoires indépendantes qui suivent toutes la même loi (celle de X). Un estimateur de θ est une fonction $\hat{\theta} = f(X_1, X_2, \dots, X_n)$ de notre échantillon, qui possède une loi de probabilité. Lorsque l'aléa est réalisé, $\hat{\theta}(\omega) = f(x_1, x_2, \dots, x_n)$ est une estimation de θ. Le but de ce cours est de construire le meilleur estimateur possible de θ.

Estimateur sans biais

Pour que l'estimation soit bonne, il faut que $\hat{\theta}$ soit proche de θ. Comme $\hat{\theta} = f(X_1, X_2, \dots, X_n)$ est une variable aléatoire, on ne peut imposer de condition qu'à sa valeur moyenne.

On définit ainsi le biais :

$$b_n(\hat{\theta}, \theta) = \mathbb{E}(\hat{\theta}_n) - \theta$$

Un estimateur est dit sans biais si $b_n(\hat{\theta}, \theta) = 0$, c'est-à-dire :

$$\mathbb{E}(\hat{\theta}_n) = \theta$$

Estimateur convergent

Un estimateur est dit convergent s'il converge en probabilité vers le paramètre à estimer :

$$\hat{\theta}_n \xrightarrow{P} \theta$$

En pratique, tout estimateur sans biais et dont la variance tend vers 0 est convergent.

Estimateur optimal

La qualité d'un estimateur est mesurée à travers son erreur quadratique moyenne définie par :

$$EQM(\hat{\theta}_n) = (b_n(\hat{\theta}, \theta))^2 + V(\hat{\theta}_n)$$

Comme nous cherchons tout le temps (presque) des estimateurs sans biais, il reste à comparer les variances.

Un estimateur ˆ1 est meilleur que ˆ2 si :

$$V(\hat{\theta}_1) < V(\hat{\theta}_2)$$

Inégalité de Rao-Cramer/ Efficacité

On définit la quantité d'information apportée par l'estimateur par :

$$I(\hat{\theta}_n) = -\left(\mathbb{E}\left(\frac{\partial L}{\partial \theta}\right)\right)^2$$

Où (,) = ( ) (nous reviendrons sur sa définition)

L'inégalité de Rao-Cramer postule que la variance d'un estimateur ne peut pas aller en delà d'un certain seuil :

$$V(\hat{\theta}_n) \geq \frac{1}{I(\hat{\theta}_n)}$$

Un estimateur est optimal (ou efficace) si sa variance vérifie le cas d'égalité.

Construction d'un estimateur

Un estimateur est une valeur qu'on ne peut pas obtenir en théorie donc on essaye de l'estimer. Par exemple, on ne peut pas calculer l'espérence d'une série de données et donc pour essayer d'obtenir une valeur on calculer une estimation. On recours à plusieurs estimateurs tels que le maximum de vraisemblance ou bien la méthode des moments.

> **Note**
>
> La méthode du maximum de vraisemblance est la plus souvent utilisé dans les modèles de prédiction.

La méthode du maximum de vraisemblance consiste à affecter θ la valeur qui maximise la probabilité d'observer $(x_1, x_2, ..., x_n)$ lorsque l'aléa du vecteur $(X_1, X_2, ..., X_n)$ tombe. Sans trop rentrer dans la théorie de la vraisemblance, nous allons présenter un algorithme en cinq étapes pour calculer cet estimateur (qui présente des propriétés assez séduisantes) :

Etape 1 : Calculer la fonction de vraisemblance

Dans le cas continu :

$$L(\mathbf{x}, \theta) = \prod_{i=1}^{n} f(x_i)$$

Dans le cas discret :

$$L(\mathbf{x}, \theta) = \prod_{i=1}^{n} P(X_i = x_i)$$

Etape 2 : Calculer le log-vraisemblance Il s'agit de calculer un maximum, ce qui revient à dériver. Il s'agit ici d'un produit de n facteurs, ce qui rend la dérivation assez coriace. La fonction logarithmique présente des propriétés assez sympas pour faciliter cette tâche.

Etape 3 : Calculer la dérivée de la log-vraisemblance

Etape 4 : Résoudre l'équation d'inconnue $\boldsymbol{\theta}\}$

$$\frac{\partial(\ln(L))}{\partial\theta} = 0 \Rightarrow \theta = \theta_0$$

Etape 5: Vérifier qu'il s'agit d'un maximum.

En s'assurant que :

$$\frac{\partial^2(\ln(L))}{\partial\theta^2} < 0$$

La méthode des moments consiste à égaliser les moments théoriques (espérance, variance) à leurs équivalents empiriques et à en dégager une estimation ponctuelle.

En pratique, il faut résoudre l'(les) équation(s) :

$$\mathbb{E}(X) = \overline{X} \text{ et } \mathrm{Var}(X) = S_n^2$$

avec :

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Il existe une autre méthode d'estimateurs qu'est la méthode des moindres carrés ordinaires.

Lorsqu'il s'agit de prendre une mesure avec un appareil doté d'une imprécision ε, alors le problème d'estimation peut s'écrire : $X = \theta + \varepsilon$. La méthode des moindres-carrés ordinaires consiste à trouver le paramètre θ qui minimise la somme des carrées des erreurs :

$$\theta_{MCO} = \arg\min\left(\sum_{i=0}^{n}\varepsilon_i^2\right) = \arg\min\left(\sum_{i=0}^{n}(X_i - \theta)^2\right)$$

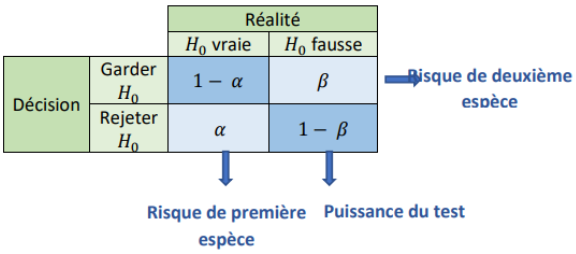
Intervalles de confiance

Un intervalle de confiance $[A, B]$ de niveau $1-\alpha$ est un intervalle aléatoire qui a la probabilité $1-\alpha$ de contenir le paramètre à estimer θ. Formellement, on écrit : $P(t_1(\theta) \leq f(X_1, \ldots, X_n) \leq t_2(\theta)) = P(A \leq \theta \leq B) = 1 - \alpha$

Test d'hypothèses

Dans le cadre d'un test d'hypothèse, nous cherchons à faire valoir une hypothèse en dépit d'une autre, qui lui est contradictoire.

On appellera la première (celle dont le rejet à tort sera le plus préjudiciable) « Hypothèse nulle » et la deuxième « Hypothèse alternative ».



Les calculs qui se cachent derrière le choix de l'hypothèse à garder sont compliqués. Mais BONNE NOUVELLE, la machine fera tour à notre place. Il suffit juste de suivre correctement la méthode :

Etape 1 : Choisir judicieusement les hypothèses à évaluer et fixer le risque α
Etape 2 : Choisir le test adapté à la procédure
Etape 3 : Rentrer la commande correspondante sur R et exécuter
Etape 4: Lire dans les sorties la p-value. si elle est supérieure à on **conserve** l'hypothèse nulle H0. Si elle lui est inférieure, on **rejette** H0 et on **accepte** l'hypothèse alternative H1.

Construction d'intervalles de confiance

Les intervalles de confiance sont des outils essentiels en statistique pour estimer des paramètres inconnus tout en mesurant l'incertitude associée à cette estimation. Ci-dessous, vous trouverez un tableau présentant la construction des intervalles de confiance pour différents paramètres.

<table>
<tr><th colspan="5">Construction d'intervalles de confiance usuels</th></tr>
<tr><th colspan="4">Pour les paramètres d'une loi $N(\mu,\sigma^2)$</th><th rowspan="3">Pour une proportion p</th></tr>
<tr><th colspan="2">Pour μ</th><th colspan="2">Pour σ^2</th></tr>
<tr><th>Avec σ connu</th><th>Avec σ inconnu</th><th>Avec μ connu</th><th>Avec μ inconnu</th></tr>
<tr>
<td>L'EMV de μ est :
$\overline{X_n} = \frac{1}{n}\sum_{i=0}^{n} X_i$, $\overline{X_n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
On a : $\sqrt{n}\,\frac{\overline{X_n}-\mu}{\sigma} \sim N(0,1)$
Nous allons donc utiliser les fractiles d'une loi normale centrée réduite.
$P\left(-u_{1-\frac{\alpha}{2}} \le \sqrt{n}\frac{\overline{X_n}-\mu}{\sigma} \le u_{1-\frac{\alpha}{2}}\right) = 1-\alpha$
Ceci donne :
$P\left(\overline{X_n} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X_n} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$
Et donc un intervalle de confiance pour μ de niveau de confiance $1-\alpha$ est :
$\left[\overline{X_n} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X_n} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$
Où $u_{1-\frac{\alpha}{2}}$ est le fractile d'ordre $1-\frac{\alpha}{2}$ d'une loi normale centrée réduite.</td>
<td>L'EMV de μ est :
$\overline{X_n} = \frac{1}{n}\sum_{i=0}^{n} X_i$, $\overline{X_n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
Comme σ inconnu, on l'estime avec son équivalent empirique.
$s_n'^2 = \frac{1}{n-1}\sum_{i=0}^{n}(X_i - \overline{X_n})^2$
On a: $\sqrt{n}\,\frac{\overline{X_n}-\mu}{s_n'} \sim St(n-1)$
Nous allons donc utiliser les fractiles d'une loi de Student à n-1 degrés de liberté.
$P\left(-t_{1-\frac{\alpha}{2}} \le \sqrt{n}\frac{\overline{X_n}-\mu}{s_n'} \le t_{1-\frac{\alpha}{2}}\right) = 1-\alpha$
Ceci donne :
$P\left(\overline{X_n} - t_{1-\frac{\alpha}{2}}\frac{s_n'}{\sqrt{n}} \le \mu \le \overline{X_n} + t_{1-\frac{\alpha}{2}}\frac{s_n'}{\sqrt{n}}\right) = 1-\alpha$
Et donc un intervalle de confiance pour μ de niveau de confiance $1-\alpha$ est :
$\left[\overline{X_n} - t_{1-\frac{\alpha}{2}}\frac{s_n'}{\sqrt{n}}, \overline{X_n} + t_{1-\frac{\alpha}{2}}\frac{s_n'}{\sqrt{n}}\right]$
Où $t_{1-\frac{\alpha}{2}}$ est le fractile d'ordre $1-\frac{\alpha}{2}$ d'une loi de Student à n-1 degrés de liberté.</td>
<td>L'EMV de σ^2 est :
$s_n^2 = \frac{1}{n}\sum_{i=0}^{n}(X_i - \mu)^2$
On a : $\frac{ns_n^2}{\sigma^2} \sim \chi_n^2$
Nous allons donc utiliser les fractiles d'une loi de Khi-deux à n degrés de liberté.
$P\left(c_1 \le \frac{ns_n^2}{\sigma^2} \le c_2\right) = 1-\alpha$
Ceci donne :
$P\left(\frac{ns_n^2}{c_2} \le \sigma^2 \le \frac{ns_n^2}{c_1}\right) = 1-\alpha$
Et donc un intervalle de confiance pour σ^2 de niveau de confiance $1-\alpha$ est :
$\left[\frac{ns_n^2}{c_2}, \frac{ns_n^2}{c_1}\right]$
Où c_1 (respectivement c_2) est le fractile d'ordre $1-\frac{\alpha}{2}$ (respectivement $\frac{\alpha}{2}$) d'une loi de Khi-deux à n degrés de liberté.</td>
<td>L'EMV de σ^2 est :
$s_n'^2 = \frac{1}{n-1}\sum_{i=0}^{n}(X_i - \overline{X_n})^2$
On a : $\frac{(n-1)s_n'^2}{\sigma^2} \sim \chi_{n-1}^2$
Nous allons donc utiliser les fractiles d'une loi de Khi-deux à n-1 degrés de liberté.
$P\left(c_1 \le \frac{(n-1)s_n'^2}{\sigma^2} \le c_2\right) = 1-\alpha$
Ceci donne :
$P\left(\frac{(n-1)s_n'^2}{c_2} \le \sigma^2 \le \frac{(n-1)s_n'^2}{c_1}\right) = 1-\alpha$
Et donc un intervalle de confiance pour σ^2 de niveau de confiance $1-\alpha$ est :
$\left[\frac{(n-1)s_n'^2}{c_2}, \frac{(n-1)s_n'^2}{c_1}\right]$
Où c_1 (respectivement c_2) est le fractile d'ordre $1-\frac{\alpha}{2}$ (respectivement $\frac{\alpha}{2}$) d'une loi de Khi-deux à n-1 degrés de liberté.</td>
<td>L'EMV de p est $\hat{p} = \frac{N}{n}$, où $N \sim B(n,p)$ (loi binomiale).
Pour n assez grand ($n \ge 30$), on a : $N \to N(np, np(1-p))$
Donc :
$\sqrt{n}\,\frac{\hat{p}-p}{\sqrt{p(1-p)}} \sim N(0,1)$
$P\left(-u_{1-\frac{\alpha}{2}} \le \sqrt{n}\frac{\hat{p}-p}{\sqrt{p(1-p)}} \le u_{1-\frac{\alpha}{2}}\right) = 1-\alpha$
Ceci donne :
$P\left(\hat{p} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \le p \le \hat{p} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right) = 1-\alpha$
Comme p (à l'intérieur de l'intervalle) est inconnu, on l'estime avec son équivalent empirique $\hat{p}$.
Un intervalle de confiance pour p de niveau de confiance $1-\alpha$ est :
$\left[\hat{p} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}, \hat{p} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right]$
Où $u_{1-\frac{\alpha}{2}}$ est le fractile d'ordre $1-\frac{\alpha}{2}$ d'une loi normale centrée réduite.</td>
</tr>
</table>

Economie

Concepts de base

Macroéconomie

La macroéconomie est l'étude économique d'un système ou de phénomènes à un niveau global de l'économie.

Microéconomie

La microéconomie se concentre sur l'observation et l'analyse des interactions à petite échelle.

Bien économique

"Chose utile à satisfaire un besoin, il faut que le bien soit disponible et en quantité limitée.

Un bien non économique est un bien qui s'obtient gratuitement, comme l'oxygène, contrairement à un bien économique qui s'obtient en payant."

Agent économique

"Un agent économique est un individu ou un groupe d'individus constituant un centre de décision économique indépendant."

Marché

"Le marché c'est une institution sociale qui permet l'échange entre l'offre et la demande."

Asymétrie d'information

"L'asymétrie d'information concerne les situations où les agents d'un marché ne possèdent pas de la même information sur un produit que ce soit au sujet de ses qualités ou de ses défauts"

Externalité

L'externalité désigne la conséquence (positive ou négative) d'une activité d'un agent économique sur un autre, sans qu'aucun des deux ne reçoive ou ne paye une compensation pour cet effet.

Concurrence Pure et Parfaite

La **CPP** repose sur cinq fondements :

- L'Atomicité du marché

Existence d'un grand nombre d'agent économique sur le marché, à tel un point que ni l'offre ni la demande ne peut exercer une action quelconque sur la production et les prix ;

- L'Homogénéité des produits

La préférence d'un produit à un autre du point de vue de l'acheteur se fait uniquement selon son prix ;

- Libre entrée et sortie sur le marché

Aucune firme ne peut s'opposer à l'arrivée d'un concurrent sur le marché, tout le monde est libre de l'intégrer ;

- Libre circulation des facteurs de production

Les facteurs de production (capital et travail) doivent être libre de se déplacer librement sans obstacle d'une industrie à l'autre ;

- La transparence de l'information

Offreurs et demandeurs sont parfaitement conscient des caractéristiques et prix des produits.

Monopole

"Le monopole est une situation dans un marché où un vendeur fait face aux multitudes vendeurs."

Segmentation de marché

"La segmentation de marché est un découpage du marché en groupes homogènes selon des critères spécifiques, que ce soit des critères démographiques ou bien géo-graphiques."

Discrimination par les prix

"La discrimination par les prix est le pouvoir de pratiquer des prix différents pour un même produit, peut s'appliquer sur la quantité ou bien selon la segmentation du marché."

Utilité

"L'utilité mesure le bien-être liée à la consommation d'un bien."

Actualisation

"L'Actualisation est un calcul permettant de transformer une valeur future en une valeur présente."

Que vaut aujourd'hui les X euros que j'aurais demain ?

$$V_a = \frac{V_f}{(1+i)^t}$$

$$V_a : Valeur\ Actuelle$$

$$V_f : Valeur\ future$$

$$i : Taux\ sans\ risque$$

$$t : Temps$$

Problèmes macroéconomiques

Il existe 4 grands problèmes macroéconomiques :

- Crises et récessions Ralentissement et/ou régression de l'activité économique ;
- Inflations Augmentation générale et durable du niveau des prix entraînant une perte du pouvoir d'achat de la monnaie ;
- Chômage Inactivité due au manque de travail ;
- Problème de l'équilibre extérieur Quand les importations sont plus importantes que les exportations, la balance commerciale est déséquilibrée.

La dissertation en économie

"On tient tout d'abord à remercier l'enseignant chercheur (en Philosophie économique, Théories économiques de la justice, Redistribution des revenus, Economie sociale, Economie publique), monsieur ***Jean-Sébastien Gharbi*** *pour cette rubrique d'aide à la dissertation."*

On pense souvent que la dissertation en économie est un exercice difficile et qui récompense mal le travail. C'est totalement faux. La dissertation est un exercice dans lequel il est facile d'obtenir la moyenne, et même (avec un peu d'entraînement) d'obtenir systématiquement de très bonnes notes. C'est un exercice relativement facile parce que c'est un exercice en très grande partie formel : tout est une question de méthode.

Faire une dissertation, c'est *montrer que vous êtes capable d'utiliser et de réorganiser vos connaissances pour répondre à une question de manière argumentée (c'est-à-dire sous la forme d'un raisonnement).* Autrement dit :

- ***Une dissertation n'est pas une question de cours.***

La première chose à faire, c'est de différencier question de cours et dissertation (qui sont souvent confondues). Une question de cours demande simplement de réciter un cours. Si vous ne faites que réciter votre cours dans un exercice de dissertation, vous aurez une mauvaise note. Pourquoi ? Parce que *l'exercice de dissertation suppose de montrer que vous êtes capable d'utiliser et de réorganiser vos connaissances* (dans un temps limité) – pas seulement de réciter une leçon apprise plus ou moins par cœur. Comme la question de cours, la dissertation

suppose donc que vous savez des choses sur le sujet, mais il est important de comprendre que la dissertation porte tout autant sur votre aptitude à organiser vos idées que sur vos connaissances.

- ***Dans une dissertation, la réponse donnée n'est pas importante !***

Une dissertation consiste toujours à répondre à une question. Les étudiants pensent parfois qu'il y a une « bonne » réponse à la question posée – qu'il s'agirait de trouver. D'ailleurs, cela contribue à l'idée (fausse) que la dissertation est un exercice aléatoire : si vous ne trouvez pas la bonne réponse, vous avez perdu. Cela aussi est faux : *il n'y a (en général) pas de « bonne » réponse à la question posée par la dissertation.* L'exercice de dissertation vient de la philosophie. Pensez-vous sérieusement que l'on puisse demander à un étudiant (ou à un professeur, d'ailleurs) de régler de façon définitive un débat philosophique qui a donné lieu à des controverses pendant des siècles en trois, quatre ou même sept heures ? La réponse évidente est « non ».

- ***Dans une dissertation, le plus important c'est l'argumentation !***

Si on ne s'intéresse pas à la réponse donnée. C'est tout simplement, parce que *ce qui intéresse votre lecteur, c'est la manière dont vous répondez* : votre aptitude à utiliser vos connaissances de manière argumentative pour défendre une conclusion. Sur le principe, il serait donc possible de défendre une conclusion choquante ou même offensante dans une dissertation, pour la bonne raison qu'on n'évalue pas la réponse que vous donnez, mais la manière dont vous amenez votre réponse. Votre réponse, à la limite, on ne s'y intéresse pas. Une fois cela dit, il est assez évident qu'il est beaucoup plus facile de défendre une position modérée et consensuelle, qu'une position offensante pour de nombreuses personnes. C'est la raison pour laquelle, il n'est pas du tout conseillé de chercher la provocation gratuite dans une dissertation.

Comment on fait une dissertation ?

Analyse du sujet

L'analyse du sujet est la première étape de la dissertation et l'une des plus importantes. Une dissertation se présente sous la forme d'un sujet. Il faut isoler la ou les deux notions principales du sujet.

Il y a quatre grands types de sujets : les sujets ne contenant qu'une seule notion (ex : « Les discriminations en France »), les couples de notions (ex : « Capitalisme et démocratie »), les citations (ex : « *Le système de production capitaliste est une démocratie économique dans laquelle chaque sou donne un droit de vote. Les consommateurs constituent le peuple souverain* », Ludwig von Mises) ou une question (ex : « La croissance économique s'oppose-t-elle à la préservation de l'environnement ? »).

Dans tous les cas, l'objectif est d'arriver à une question (donc les sujets les plus simples à traiter à ce stade, ce sont les questions : ils vous donnent immédiatement le problème à traiter). Mais la première chose à faire (même quand on a déjà la question), c'est de trouver le couple de notions impliquées dans le sujet. Souvent, c'est absolument évident, mais parfois il faut un peu chercher.

Il faut éviter à tout prix de faire un exposé quand on attend de vous une dissertation. Un hors-sujet, c'est de ne pas traiter le bon sujet. Si vous répondez de manière factuelle à un sujet de dissertation, vous faites pire : vous faites un hors-exercice.

Recherche des idées

Une fois qu'on a identifié un couple de notions, il faut (au brouillon) essayer de faire la liste des éléments du cours qui relient les deux notions. Il est important de ne noter que les éléments qui relient les deux notions (pour ne pas risquer de se perdre dans des éléments qui concernent seulement une seule des deux notions). Sur chacune des notions que vous aurez à traiter en dissertation, on a écrit des livres entiers. Il est impossible de tout dire dessus dans une dissertation. On se limite donc à ce qui relie les deux notions de notre sujet. Évidemment, si un élément pertinent vous vient en tête et qu'il ne se trouve pas dans votre cours, n'hésitez pas à le noter. Si cet élément peut être utilisé dans votre raisonnement, même comme exemple, ce sera un plus indiscutable.

Dans un premier temps, on note tout ce qui se présente à l'esprit. Ce n'est que dans un deuxième temps, quand on a un certain nombre d'éléments que l'on se pose la question : « Est-ce qu'il y a une manière qui saute aux yeux de relier tous ces éléments en répondant à la question (si elle a été posée de manière directe) ou pour répondre à une question comprenant le couple de notions (si la question n'a pas été formulée dans le sujet) ? ». Si la réponse est positive, on a trouvé la question qui structurera notre devoir. Si ce n'est pas le cas, il faut essayer de trouver une question qui relie le plus grand nombre des éléments que l'on a noté sur son brouillon – et donc laisser de côté les éléments qui ne servent pas. Il arrive souvent qu'une partie des éléments que l'on note sur son brouillon ne soit pas utilisée dans le devoir.

Bref, si notre sujet est un couple de notions ou une citation, il faut que l'on arrive à une question. Évidemment, quand notre sujet est déjà une question, on n'a pas autant de marge de manœuvre, mais en vérité, si on vous pose une question précise, c'est que vous avez les éléments pour y répondre dans le cours (donc la différence n'est pas très importante).

Mise en évidence d'un problème

Puisqu'elle ne doit pas être une question de fait, la question qui relie le plus d'éléments possibles parmi ceux qui associent les deux notions dans votre cours doit être une question conceptuelle. Pour le dire autrement, cette question doit être un problème (on parle souvent de « problématique » pour désigner ce problème dans le cadre d'une dissertation). Qu'est-ce qu'un problème ? C'est une question qui met en tension deux concepts et qui analyse les différents aspects de leur relation (conceptuelle).

Souvent les étudiants ont peur de ne pas trouver le « bon » problème. Pourtant, si on suit la méthode de dissertation, il n'y a pas de risque de se tromper. En effet, on ne doit pas choisir un problème d'abord (sans savoir si on a de quoi le traiter) et le traiter ensuite. Vous aurez noté qu'on procède exactement dans le sens inverse : on voit à quelle question on peut répondre avec les éléments qu'on a sur son brouillon et on pose précisément la question à laquelle on sait qu'on peut répondre.

Construction du plan

Pour la même raison qu'au-dessus, la construction du plan ne doit pas être très difficile : il s'agit de rassembler les différents éléments qui permettent de répondre à la question (au problème que l'on va poser) de façon à y apporter une réponse.

On va apporter la réponse que les éléments disponibles nous permettent d'atteindre. Il y a une seule contrainte : votre plan doit être suffisamment détaillé. Comme l'objectif de la dissertation, c'est de montrer que vous êtes capable d'utiliser et de réorganiser vos connaissances. Si vous ne faites que deux parties sans sous-parties à l'intérieur (il n'y aurait donc qu'une seule articulation logique), on trouvera que vous n'avez pas assez structuré votre devoir. Le découpage minimal, c'est d'avoir quatre éléments (en général, on fait deux grandes parties avec deux sous-parties chacune, donc on a trois articulations).

Vos parties et vos sous-parties doivent correspondre à des étapes de votre raisonnement (on dit souvent qu'il faut une idée par sous-partie), donc votre plan doit donner la structure du raisonnement grâce auquel vous allez répondre à la question posée. Le plan (qui est annoncé à la fin de l'introduction et qui doit être apparent dans le devoir, nous reviendrons sur ce point un peu plus loin) doit permettre de comprendre la structure de votre devoir d'un coup d'œil – simplement en lisant les titres.

Un élément qui permet de savoir si votre plan est bon, c'est de se demander si à la fin de la première partie, on est arrivé à un état de la réflexion différent de celui de la fin de l'introduction.

Qu'est-ce que la première partie a permis de comprendre ? Et est-ce que la seconde partie apporte quelque chose d'autre ? Si chaque partie représente une étape dans un raisonnement et que votre devoir complet est donc un raisonnement, votre plan est forcément bon : vu que c'est précisément ce qu'on attend de vous.

Il ne faut jamais faire deux sous-parties dans une partie sous la forme d'une seule phrase coupée par des points de suspension (ex : « A) L'organisation scientifique du travail a permis la croissance des trente glorieuses… », « B) mais, elle a aussi eu des conséquences négatives, notamment sur le plan social »). En effet, cela revient à pointer du doigt que vous opérez une coupure arbitraire (donc que vous n'articulez pas de manière assez nette les différentes parties de votre devoir). Sur le principe, les deux sous-parties sont reliées par les points de suspension et donc ne forment qu'une partie sans coupure. Préférez toujours les titres qui se succèdent sans être grammaticalement liés les uns aux autres. Dans l'exemple ci-dessus, il suffit de faire deux phrases pour découper les deux idées. Si on ne peut pas couper grammaticalement les deux titres, c'est la preuve que l'articulation pose problème.

Dans l'idéal, un plan est équilibré : chaque partie comprend le même nombre de sous-parties que l'autre et elles font à peu près la même longueur. Et si vous faites plus de sous-parties dans une partie que dans l'autre, il faut que les sous-parties soient un peu plus longues dans la partie qui contient moins de sous-parties. Remarquez que si vous faites un plan avec deux parties, deux sous-parties, ce dernier problème ne se pose pas.

Un point important et souvent totalement négligé par les étudiants : même quand c'est tentant, *on ne fait jamais de plan centré sur les auteurs*. Un plan par auteurs conduit très souvent à suivre l'ordre chronologique et à présenter les positions des auteurs sans les confronter réellement les unes aux autres . Ce qui doit structurer le plan, ce sont les concepts (c'est-à-dire les notions qui nous avaient permis de construire le problème à résoudre).

En réalité, il n'est pas rare qu'on suive plus ou moins l'ordre chronologie, mais il est essentiel de se focaliser sur les concepts, et pas sur les auteurs. Pourquoi ? Parce que l'enchaînement ou l'opposition de concepts constitue un raisonnement (ce que vous devez faire !), alors que l'enchaînement ou l'opposition d'auteurs constitue un exposé (ce que vous ne devez pas faire !). En réalité, c'est assez facile à faire il suffit de s'interdire de mentionner le nom des auteurs dans les titres de partie ou de sous-partie.

Vu que l'objectif du plan, c'est de répondre à une question qui met les deux notions du sujet en relation, *les parties (ou les sous-parties) qui se focalisent sur une seule des deux notions sont à éviter à tout prix : elles sont simplement hors-sujet*. Sur un sujet comme « capitalisme et démocratie », le plan « première partie : capitalisme », « deuxième partie : démocratie » est parmi les pires possibles.

Jusqu'à présent, nous n'avons encore rien écrit sur la copie elle-même. Nous n'avons travaillé que sur le brouillon. Nous avons deux notions clés, un problème et un plan. Ce sont

les éléments fondamentaux du devoir. Il faut à présent passer à la rédaction. Nous allons nous intéresser d'abord à l'introduction.

La rédaction

(Introduction, Conclusion, Développement)

L'introduction est la partie la plus importante de la dissertation. Elle permet de savoir pourquoi le problème se pose, comment il se pose et comment il va être résolu. A quoi sert l'introduction ?

Le rôle de l'introduction, sa raison d'être, c'est de construire et d'énoncer le problème (la problématique) auquel le reste du devoir va répondre. Il ne suffit donc pas de poser la question (pour cela deux lignes suffiraient) et de commencer le développement. L'introduction, comme son nom le dit très bien, va introduire le problème, c'est-à-dire qu'elle va nous y amener, rapidement, certes, mais en plusieurs étapes très codifiées.

Une introduction de dissertation comprend obligatoirement (au minimum) cinq éléments : une accroche, une définition des termes du sujets, la construction du problème, l'énoncé du problème et l'annonce du plan. Comme une introduction de dissertation fait entre 20 lignes et une page et demie (grand maximum), il faut être efficace.

- ***L'accroche***

Une introduction de dissertation suit des règles assez rigides. Elle commence toujours par une accroche.

Une « accroche », c'est une phrase ou deux qui vont contenir la ou les deux notion(s) du sujet. Son rôle est d'amener par étapes le lecteur vers la question que vous allez poser. Elle sert donc d'introduction à l'introduction. Une accroche peut être une citation (il y en a toujours dans un cours) ou un fait récent (le chômage a-t-il baissé récemment ? Un candidat à l'élection présidentielle a-t-il dit qu'il fallait juger sa politique en fonction de son impact sur le niveau de chômage ?). Si on n'a pas de citation ou de fait relevant de l'actualité, on peut amener le sujet de façon plus habituelle.

Il est important d'éviter un certain nombre de formules toutes faites et souvent utilisées comme « De tous temps… », « De tous temps, les hommes… » ou encore les affirmations très générales (et que vous ne justifierez pas) comme : « Le chômage est un phénomène économique important, c'est pourquoi il faut l'étudier ». Le défaut de tous ces débuts d'accroche, c'est qu'ils peuvent servir pour n'importe quel sujet et que cela se voit.

On met une accroche parce que cela permet de mentionner les termes du sujet sans commencer directement par une définition – ce qui constitue l'étape suivante.

- **Définition des termes du sujet**

L'accroche a introduit les notions, mais sans les définir – comme si tout le monde savait précisément de quoi il s'agit (ce qui n'est pas si surprenant, on ne passe pas son temps à définir tous les mots qu'on utilise). Mais, pour utiliser les deux notions du sujet de façon un peu plus précise, il faut les définir. Les définitions que l'on va donner dans une introduction n'ont pas pour objectif de définir les notions de manière exhaustive ou dans l'absolu. Elles doivent permettre de comprendre le lien (ou l'opposition) entre les deux notions et orienter l'introduction de façon à ce que l'on puisse construire le problème – avant de l'énoncer (autrement dit, elles doivent ouvrir la voie aux deux étapes suivantes de l'introduction).

Du coup, les définitions que l'on va donner vont dépendre du problème que l'on souhaite atteindre.

- **Construction du problème**

Comme on sait à quelle question on doit arriver (que cette question nous ait été donnée par le sujet ou que ce soit la question à laquelle on est le mieux armé pour apporter une réponse), il ne va pas être difficile de passer des définitions au problème. Cela suppose juste de montrer qu'avec les définitions que l'on vient de donner, il y a une question se pose avec force.

Encore une fois, cela peut sembler très artificiel (et ça l'est). Toutefois, l'intérêt de cet aspect artificiel, c'est qu'il nous garantit que l'on ne va pas se perdre en chemin. Quand on fait une dissertation, on ne cherche pas son chemin : on sait où on va et on ne fait qu'expliquer pourquoi on y va. Le sujet que l'on construit ne tombe pas du ciel, il vient de notre cours. Les définitions ne tombent pas du ciel, elles donnent les éléments qui vont nous permettre de poser la question à laquelle on sait déjà comment on va répondre. Bref, l'étape de construction du problème est importante parce qu'elle montre que vous avez des aptitudes pour vous faire comprendre à l'écrit (et il ne faut surtout pas la négliger), mais elle n'est pas une étape difficile ou magique.

Vous pourriez être surpris que l'on construise le sujet, alors qu'il nous est parfois donné sous forme de question (dans les autres cas, on comprend mieux pourquoi il est nécessaire de construire le problème). En fait, c'est une manière de montrer que vous êtes capable de vous approprier le sujet. Vous ne traitez pas le sujet parce qu'on vous l'a donné (même si vous c'est une des raisons pour lesquelles vous faites une dissertation), mais parce que vous comprenez pourquoi la question se pose. Et comment mieux montrer qu'on comprend un problème qu'en montrant en quoi il est problématique ? Autrement dit, même quand votre sujet a la forme d'une question, vous devez passer par l'étape de construction du sujet dans l'introduction.

- **Enoncé du problème**

L'énoncé du problème doit prendre la forme d'une question. Il est le point final de l'étape juste précédente. Une fois qu'on a les éléments qui permettent de comprendre que le problème se pose, il faut explicitement exprimer le problème lui-même. *On exprime toujours le problème sous la forme d'une question (parce que c'est une manière de montrer qu'il appelle une réponse) et d'une question unique.* Poser deux, trois ou quatre questions ce serait soit redire plusieurs fois

la même chose (et si votre première question est claire, c'est inutile), soit poser (volontairement ou pas) plusieurs questions différentes. Or, vous ne pourrez pas répondre convenablement et dans les règles de la dissertation à plusieurs questions en un seul devoir. Vous devrez donc choisir entre ne pas traiter certaines des questions que vous avez explicitement posées (et dans ce cas pourquoi les poser explicitement) ou essayer de les traiter toutes (ce qui vous conduira à un devoir dont la ligne directrice sera au mieux difficile à suivre, au pire inexistante). Si on se rappelle du côté formel et rhétorique d'une dissertation, on comprend qu'il ne faut poser qu'une seule question : celle à laquelle vous apportez une réponse.

Lorsque le sujet est une question, faut-il répéter mot pour mot le sujet comme énoncé du problème ? Il y a deux écoles : la première dit qu'il faut reformuler la question pour montrer que vous la comprenez. Ainsi un sujet comme « Les dépenses publiques permettent-elles de réduire le chômage ? », on pourrait proposer une problématique comme « les dépenses publiques sont-elles efficaces à court et à long terme pour lutter contre le chômage ? ».

Si vous faites correctement votre travail de définition des termes et de construction du sujet (dans les deux étapes précédentes), je pense qu'aucun correcteur ne vous reprochera de reprendre le sujet mot pour mot dans votre énoncé du problème. Ce qui pose problème pour les partisans de la première façon de faire, c'est quand on peut se demander si l'étudiant comprend que la question qu'il pose est un problème conceptuel, c'est-à-dire qui vient d'une tension entre deux notions. Dans une introduction qui remplit correctement son rôle de construction du problème, le fait de répéter le sujet mot pour mot n'est pas un souci.

- **Annonce du plan**

Une introduction doit toujours se terminer par une annonce du plan (ce n'est pas une option, c'est une obligation). L'annonce de plan dit à votre lecteur comment vous allez répondre au problème que vous venez de poser. Dans une dissertation, on ne joue pas sur le suspens. On ne cherche pas à surprendre son correcteur. Il faut donc annoncer le plan de manière à ce qu'il comprenne que vous allez répondre au problème posé par un raisonnement et qu'il comprenne aussi quels vont être les principales étapes de votre raisonnement (c'est-à-dire de votre devoir).

Vous allez donc annoncer vos (deux ou trois) grandes parties. Il est conseillé fortement d'utiliser les formules (un peu lourdes en termes de style, mais très claires) « dans une première partie, nous montrerons que… », puis « dans une deuxième partie, nous verrons que … ». Quand vous ne le faites pas, il arrive trop souvent que votre lecteur ne sache pas si vous allez faire formellement deux ou trois parties – pour peu que vous utilisiez des mots comme « et », « puis » ou « ensuite », qui peuvent aussi bien marquer des étapes à l'intérieur d'une grande partie que le passage d'une partie à une autre.

La conclusion

Vous pourriez être surpris de voir la conclusion arriver aussi tôt dans le devoir. La raison, c'est qu'il est inconcevable de ne pas répondre à la question posée en introduction – si vous ne répondez pas le devoir n'aura, littéralement, servi à rien. Or, il est évident qu'en

partiel, on est souvent pris par le temps. On rédige donc la conclusion juste après avoir rédigé l'introduction au brouillon (on la rédige aussi au brouillon, d'ailleurs). Comme ça si on est pris par le temps, on pourra recopier la conclusion déjà prête avant de rendre le devoir. S'il faut couper quelque chose en raison du temps limité de l'épreuve, il vaut mieux couper un bout du développement que rendre une dissertation sans conclusion.

La première phrase de votre conclusion doit apporter la réponse à la question que vous avez posée en introduction. Elle doit le faire de façon absolument claire et donc il est conseillé de reprendre exactement la question en la tournant en une phrase affirmative ou en une phrase négative selon votre réponse. *Le rôle de la conclusion, c'est de répondre à la question.* Il ne faut pas qu'on relise la conclusion en se demandant quelle était la réponse – et même en se demandant si une réponse a été donnée. Cela ne vous empêche pas de donner une réponse nuancée, mais il faut une réponse claire.

Une conclusion de dissertation ne résume pas le devoir (on vient de le lire, c'est tout à fait inutile). Une conclusion n'introduit jamais un élément qui n'a pas été abordé dans le devoir, mais qui aurait pu y être discuté. Si jamais votre correcteur n'a pas vu que vous avez oublié de parler de quelque chose d'important, vous n'allez tout de même pas lui dire qu'il manque quelque chose dans votre devoir (chacun son boulot). La dissertation est un exercice de rhétorique, votre objectif, c'est de convaincre votre lecteur : ce n'est pas à vous de dire qu'il manque quelque chose, même si vous le savez.

On conseille parfois de finir sa dissertation sur une ouverture. Une ouverture est un nouveau problème qui se pose une fois que vous avez répondu au problème de votre devoir. Cela revient à suggérer une autre dissertation possible une fois qu'on considère votre réponse comme acceptée. Trop souvent, les étudiants finissent leurs devoirs de manière particulièrement maladroite parce qu'ils ne comprennent pas ce qu'est une ouverture. Mon conseil est d'éviter de faire une ouverture, au moins au début : ce n'est pas une obligation et cela peut donner une très mauvaise impression finale.

La rédaction du devoir

Une fois tout cela fait, on prend sa copie (totalement vierge à ce moment) et on commence à écrire dessus : on recopie l'introduction, on rédige le développement directement sur la copie (on ne rédige jamais son développement sur le brouillon, cela prend beaucoup trop de temps à recopier). *Le développement du devoir doit contenir des titres apparents pour les parties et les sous-parties.* Cela signifie que le titre de votre grande partie est marqué dans votre copie (précédé d'un « I) ») et qu'il est isolé du texte et souligné. Bref, on doit pouvoir voir apparaître d'un coup d'œil votre plan en survolant votre copie du regard.

Comme dit juste au-dessus, si on manque de temps, on coupe une partie du développement et on recopie la conclusion qui se trouve sur le brouillon. Attention : si vous ne rédigez pas tout le développement, mettez tout de même le plan apparent pour les parties et sous-parties non développées. C'est précisément parce qu'on a une idée de ce que vous auriez écrit qu'il est possible (en cas de gros manque de temps) de ne pas rédiger tout le développement. Si vous ne détaillez pas votre plan, c'est la trame de votre raisonnement qui manque et c'est beaucoup

plus ennuyeux. Si vous ne pouvez pas rédiger tout le développement, je vous conseille de mettre des éléments que vous auriez utilisé sous forme de liste de tirets (en plus des titres apparents qui sont obligatoires).