Statistiques Descriptives

L1 Économie

Seynabou Gueye, Mehdi Guelmamen, Emilien Macault

Faculté de Droit, Économie et Administration
Université de Lorraine

Chapitre 1 : Notions statistiques et collecte des données

Sommaire

Sommaire

Modalités pratiques

Ce cours mélangera cours magistraux et applications en classe $\rightarrow$ alternance de CM et de TD.

Vous êtes répartis en groupes avec différents professeurs mais le contenu couvert est le même $\rightarrow$ à la fin du semestre, le sujet d'examen est identique pour toute la promotion.

La présence est obligatoire $\rightarrow$ n'oubliez pas de remplir la feuille d'émargement.

La participation en classe est valorisée!

L'évaluation comprend une "colle" d'1h30 en cours de semestre (lors d'une séance) et un examen final.

Références bibliographiques

- Anderson, D. et al. (2015). Statistiques pour l'économie et la gestion. De Boeck Supérieur.
- Barnichon, D. (2008). Mathématiques et statistiques appliquées à l'économie: cours, méthodes, exercices corrigés. Bréal.
- Bailly, P. et al. (2007). Statistiques Descriptives : Exercices. Presses Universitaires de Grenoble.

Sommaire

Le mot "statistique" a été inventé au 18è siècle par l'économiste G. Achenwall, du mot italien "statista", personne d'Etat $\rightarrow$ à l'origine, les statistiques sont vues comme les outils des personnes d'Etat.

A l'origine, il désigne les **méthodes numériques** pour l'étude de **phénomènes sociaux** puis l'ensemble des techniques quantitatives pour analyser et prédire le comportement de **phénomènes observés**.

La statistique est un champ des **mathématiques appliquées**: c'est à la fois un ensemble de techniques et l'étude scientifique de celles-ci.

Introduction

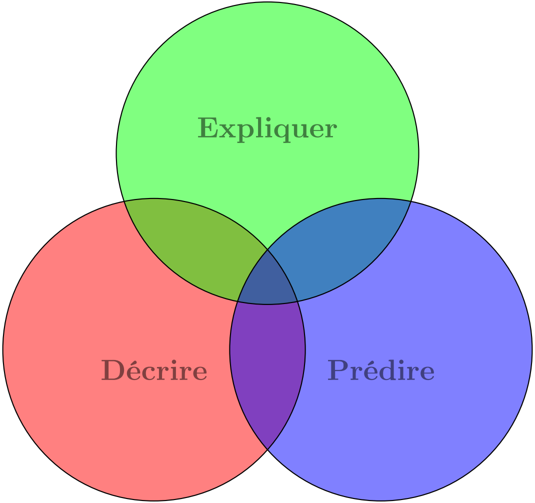
Les statistiques servent à :

1. Collecter des données (sondages, expériences);
2. Décrire des données, synthétiser l'information qu'elles contiennent;
3. Faire des prévisions, expliquer des phénomènes, prendre des décisions.

Elles sont utilisées dans tous les domaines:

- Prévisions météorologiques;
- Enquêtes sociologiques et démographiques;
- Marketing, finance, assurance;
- Expériences médicales, prédiction de pathologies;
- Mais aussi archéologie, géosciences, informatique, etc.

Les trois fonctions de la statistique



Introduction

L'idée fondamentale de la statistique est qu'en observant un **même** phénomène sur un "petit" nombre d'individus, on peut obtenir une information qui se **généralise** à ceux qu'on n'observe pas.

Chaque fonction de la statistique est un champ d'étude propre:

- Décrire $\rightarrow$ statistique descriptive;
- Prédire $\rightarrow$ statistique inférentielle;
- Expliquer $\rightarrow$ inférence causale.

Dans ce cours, on se concentrera sur la partie descriptive: comment synthétiser et visualiser l'information contenue par des observations ?

Sommaire

Introduction

Question: comment ne pas dire n'importe quoi en généralisant ce qu'on observe?

Il existe de nombreux pièges, ou **biais**, liés aux statistiques qui rendent les résultats inutilisables:

- Lors d'un projet, on demande à des étudiants de sonder les passagers des bus pour connaître le taux de fraudes sur les différentes lignes. Quel est le problème lié à ce type d'enquêtes ?

La description de données est également une étape délicate:

- Exemple : Que peut-on conclure suite à la divulgation d'une statistique mettant en avant que dans une ville donnée, le nombre de meurtres a augmenté de 100% entre 2021 et 2022 ?

Il faut donc faire attention tout au long du processus.

Exemples de biais

- Une application vous dit "Si vous appréciez notre service, laissez-nous une évaluation."

Exemples de biais

- Une application vous dit "Si vous appréciez notre service, laissez-nous une évaluation." → **biais de motivation:** les individus satisfaits répondront plus que les individus déçus;
- Un étudiant en stage fait circuler une étude de marché auprès de ses camarades de promo

Exemples de biais

- Une application vous dit "Si vous appréciez notre service, laissez-nous une évaluation." → **biais de motivation:** les individus satisfaits répondront plus que les individus déçus;
- Un étudiant en stage fait circuler une étude de marché auprès de ses camarades de promo → **biais d'endogamie:** les individus sondés sont tous des étudiants de la même promo;
- Une entreprise réalise un sondage en appelant des individus sur leur téléphone fixe

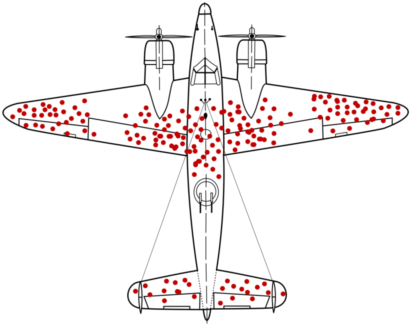
Exemples de biais

- Une application vous dit "Si vous appréciez notre service, laissez-nous une évaluation." → **biais de motivation:** les individus satisfaits répondront plus que les individus déçus;
- Un étudiant en stage fait circuler une étude de marché auprès de ses camarades de promo → **biais d'endogamie:** les individus sondés sont tous des étudiants de la même promo;
- Une entreprise réalise un sondage en appelant des individus sur leur téléphone fixe → **biais technologique:** les individus répondant sont susceptibles d'être des personnes âgées.

Ce sont des **biais de sélection**: les individus qu'on observe ne sont pas représentatifs car ils sont indirectement sélectionnés selon des critères précis.

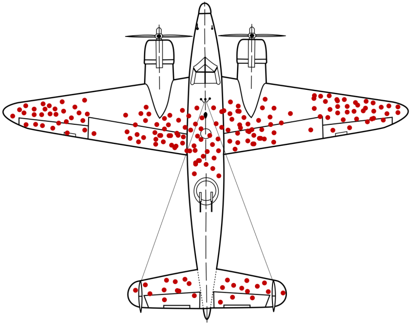
Le biais du survivant

Seconde Guerre mondiale: où renforcer le blindage des avions de l'US Air Force pour limiter les pertes?



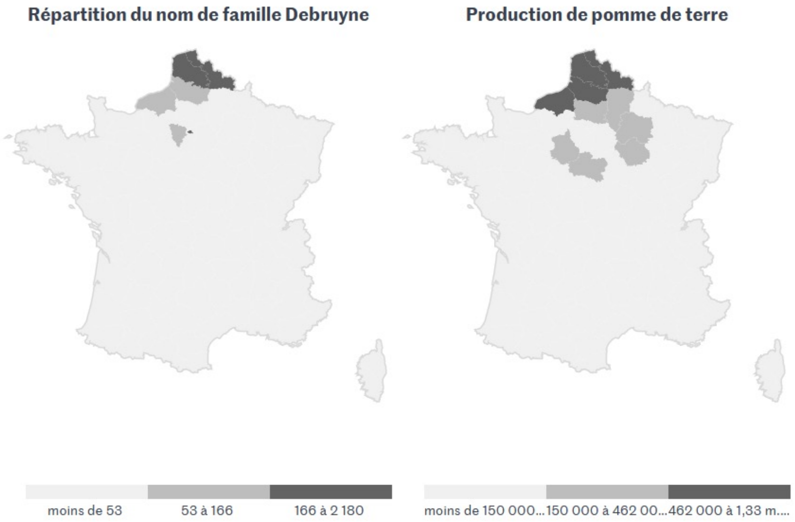
Le biais du survivant

Seconde Guerre mondiale: où renforcer le blindage des avions de l'US Air Force pour limiter les pertes?



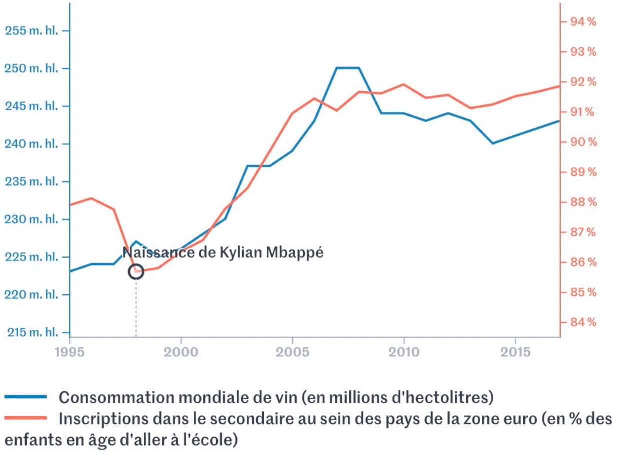
Réponse : là où il y a le moins d'impacts!

Piège récurrent: la corrélation fallacieuse



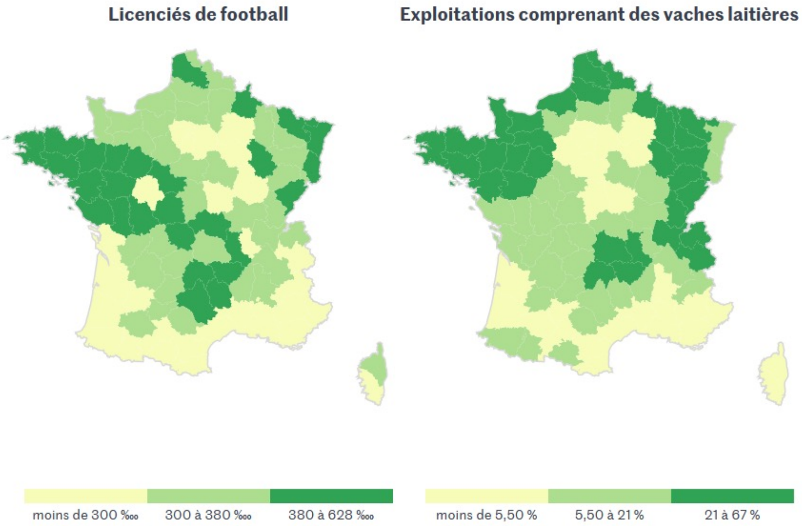
Source: Le Monde

Piège récurrent: la corrélation fallacieuse



Source: Le Monde. Coefficient de corrélation de 90,1%

Piège récurrent: la corrélation fallacieuse



Source: Le Monde

Piège récurrent: la corrélation fallacieuse

On dit que deux phénomènes sont **corrélés** s'ils évoluent conjointement, soit dans le même sens, soit en sens inverse.

On a tendance à croire que deux phénomènes corrélés ont un **lien de cause à effet** → **C'EST FAUX!!**

Exemple: en France, 57% des décès ont lieu à l'hôpital donc pour rester en bonne santé il vaut mieux éviter les hôpitaux.

Une corrélation peut être le fruit du hasard, d'une cause commune (ici, la maladie), etc. → **corrélation n'implique pas causalité!**

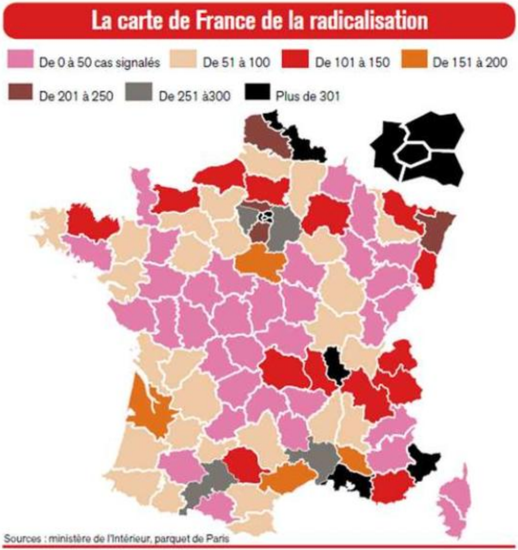
Visualiser des données

On peut aussi induire des personnes en erreur en présentant mal (volontairement ou non) les données dont on dispose.

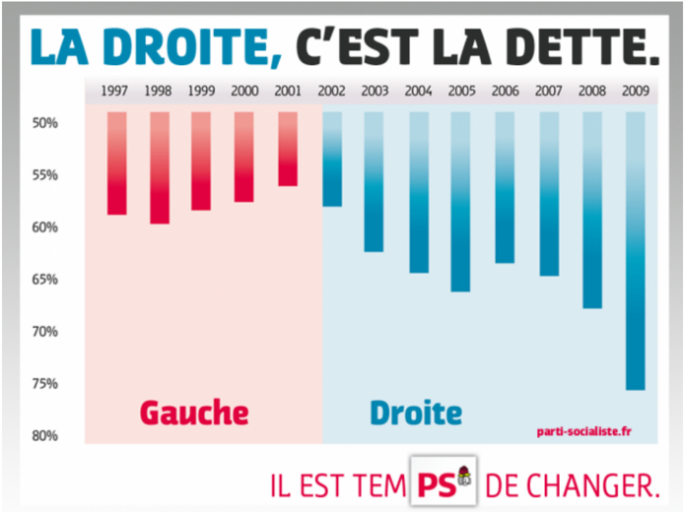
Par exemple:

- En prenant des échelles qui déforment la réalité;
- Avec des pourcentages dont la somme dépasse 100%;
- En comparant des quantités qui ne sont pas comparables;
- En mettant trop d'informations pour "noyer le poisson";
- En manipulant les données pour exclure des individus, changer des valeurs, etc.

Visualiser des données


La carte de France de la radicalisation
De 0 à 50 cas signalés
De 51 à 100
De 101 à 150
De 151 à 200
De 201 à 250
De 251 à300
Plus de 301
Sources : ministère de l'Intérieur, parquet de Paris

Visualiser des données



Conclusion

Appréhender les statistiques – que ce soit pour en produire ou en lire – requiert donc d'être:

- **Rigoureux** dans les méthodes et les calculs;
- **Attentif** aux différents biais qui pourraient perturber les données;
- **Objectif** dans la présentation des résultats;
- **Synthétique** dans la quantité d'information présentée.

Bref, faire des statistiques, ça s'apprend!

Sommaire

Sommaire

La population

En statistiques, une **population** représente l'ensemble des individus sur lequel porte l'étude. On la notera ici Ω.

Un **individu** – ou unité statistique – ω est un élément d'une population: $\omega \in \Omega$.

On observe rarement toute la population, et plutôt un **échantillon** $O \subseteq \Omega$, qui représente donc un sous-ensemble de la population.

Un individu appartient donc à un échantillon, lui-même appartenant à une population: $\omega \in O \subset \Omega$.

On note $\mathrm{Card}(O)$ ou n **l'effectif** (le nombre d'individus) de O. On appelle **effectif cumulé** l'effectif de Ω. Exemple :
$\Omega = \{a, b, c, d\}$ $\mathrm{Card}(\Omega) = 4$, donc n=4

La population

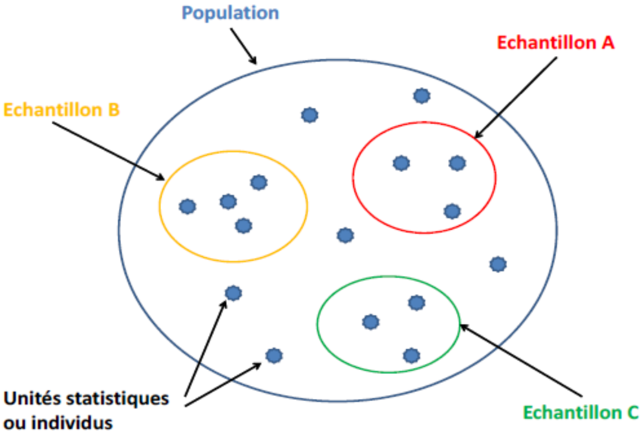


Figure: Illustration de la notion de population

La population

Pour récolter des données, deux méthodes sont généralement utilisées:

1. Le **sondage**, qui vise à interroger un échantillon représentatif de la population – donc un sous-ensemble de la population – pour extrapoler des résultats. C'est notamment comme cela que sont calculés les parts d'audience à la télévision, via un panel représentatif de 5000 foyers.
2. Le **recensement**, réalisé sur l'ensemble de la population. Par exemple, les communes de moins de 10 000 habitants réalisent une fois tous les cinq ans une enquête de recensement portant sur toute la population, à raison d'une commune sur cinq chaque année.

Sommaire

Variables et modalités

Lors d'études statistiques, on étudie des **caractères** – ou **variables statistiques** – possédés par tous les individus de la population.

On notera une variable par X, et on a:

$$\begin{aligned} X : \Omega &\to C \\ \omega &\mapsto X(\omega) \end{aligned} \tag{1}$$

Une variable X est donc une application de Ω dans C, l'ensemble des **valeurs** ou **modalités** de X.

Exemple : si X est la note d'un étudiant à un examen, généralement

- $C = \{0, 1, 2, 3, ..., 20\}$
- Pour chaque variables, on peut définir des modalités différentes selon les situations: notes avec demi-points, note supérieure ou inférieure à 10, etc.

Variables et modalités

Toutes les modalités d'un même caractère respectent 2 propriétés:

- **Incompatibles**: un individu ne peut pas correspondre à plusieurs modalités d'un même caractère.
  - Exemple: Lors d'un match de football à élimination directe, les modalités sont soit la victoire, soit la défaite. Les deux sont incompatibles.
- **Exhaustives**: chaque individu doit appartenir à une modalité. Sinon, un ou plusieurs individus ne sont pas considérés.

Exercice:

- Une entreprise emploie 500 salariés sur 3 usines (A, B et C). Pour étudier le statut conjugal des employés, on interroge les 150 salariés de l'usine A. Indiquer la population, l'échantillon, le caractère, les individus et les modalités:

Variables et modalités

Exercice:

- Une entreprise emploie 500 salariés sur 3 usines (A, B et C). Pour étudier le statut conjugal des employés, on interroge les 150 salariés de l'usine A. Indiquer la population, l'échantillon, le caractère, les individus et les modalités:
  - Population: ensemble des 500 salariés.
  - Échantillon: salariés travaillant dans l'usine A.
  - Caractère: statut conjugal.
  - Individus: salariés.
  - Modalités: célibataire, concubinage, union civile, marié(e), divorcé(e), veuf(ve).
- A partir de cette enquête, peut-on en déduire que la situation conjugale de l'ensemble des salariés suit la même répartition?

Exercice:

- Une entreprise emploie 500 salariés sur 3 usines (A, B et C). Pour étudier le statut conjugal des employés, on interroge les 150 salariés de l'usine A. Indiquer la population, l'échantillon, le caractère, les individus et les modalités:
  - Population: ensemble des 500 salariés.
  - Échantillon: salariés travaillant dans l'usine A.
  - Caractère: statut conjugal.
  - Individus: salariés.
  - Modalités: célibataire, concubinage, union civile, marié(e), divorcé(e), veuf(ve).
- A partir de cette enquête, peut-on en déduire que la situation conjugale de l'ensemble des salariés suit la même répartition? → A priori non!

Variables qualitatives

Une variable X est **qualitative** lorsque ses modalités ne sont pas des quantités mesurables.

- Lorsque les modalités sont sujettes à une relation d'ordre, la variable est dite **ordinale**: par exemple les appréciations lors d'un sondage (satisfait, moyennement satisfait, peu satisfait);
- Lorsque ce n'est pas le cas elle est dite **nominale**: par exemple département de résidence, catégories socio-professionnelles, nom des journaux, signes astrologiques.

Ensembles discrets et continus

On dit qu'un ensemble de nombres est

- **Discret** lorsque ses éléments sont séparés les uns des autres: par exemple $\mathbb{N} = \{1, 2, 3, 4, \dots\} \rightarrow$ entre 1 et 3 il n'y a que 2;
- **Continu** si entre deux éléments distincts, il y a toujours d'autres éléments: par exemple l'intervalle $[0, 1] \rightarrow$ entre $0,1$ et $0,11$ il y a une infinité de valeurs : $0,105$, $0,100001$, etc.

ATTENTION: un ensemble contenant un nombre fini de nombres est toujours discret, mais l'inverse n'est pas vrai!

Variables quantitatives

Une variable X est **quantitative** lorsqu'elle représente une quantité mesurable.

- On dit que X est **discrète** lorsque ses valeurs appartiennent à un ensemble discret: il est possible de les énumérer : le nombre d'items d'une liste, le nombre de personnes dans une salle.
- On dit que X est **continue** lorsque ses valeurs appartiennent à un ensemble continu: il est impossible de les énumérer, par exemple la taille en mètres, une note à un examen.

Dans la pratique, une modalité est continue lorsqu'elle peut prendre un grand nombre de valeurs, notamment avec des décimales: par exemple le PIB d'un pays.

On peut transformer une variable continue en variable discrète en la regroupant en classes: par exemple le PIB arrondi au million d'euros.

Exercice

Exercice: Indiquer si les variables suivantes sont qualitatives (ordinales ou nominales), quantitatives discrètes ou continues: catégories socio professionnelles, poids en kg, nombre de parts fiscales par foyer, nombre de fautes par ligne, niveau de satisfaction, salaire.

Exercice

Exercice: Indiquer si les variables suivantes sont qualitatives (ordinales ou nominales), quantitatives discrètes ou continues: catégories socio professionnelles, poids en kg, nombre de parts fiscales par foyer, nombre de fautes par ligne, niveau de satisfaction, salaire.

- Catégories socio-professionnelles : qualitatif nominal.
- Poids: quantitatif continu.
- Nombre de parts fiscales par foyer: quantitatif discret.
- Nombre de fautes par ligne: quantitatif discret.
- Niveau de satisfaction: qualitatif ordinal.

Les effectifs

L'effectif n_i qui représente le nombre d'individus ayant la même modalité x_i est appelé **effectif partiel** ou **fréquence absolue**.

$$n_i = \text{Card}\left\{\omega \in \Omega : X(\omega) = x_i\right\}$$

Dans l'exemple de l'usine de 150 salariés, on peut calculer l'effectif des célibataires, des marié(e)s ou des veuf(ve)s.

Dans l'exemple des notes, on peut calculer l'effectif des élèves ayant eu moins de 10 et de ceux ayant eu 10 ou plus.

Les effectifs

L'effectif **cumulé** est la somme des k effectifs partiels. Il peut être **croissant** ou **décroissant**.

$$n = \text{Card}(\Omega) = n_1 + n_2 + n_3 + ... + n_k = \sum_{i=1}^{k} n_i$$

L'effectif cumulé sur toutes les modalités d'une variable est égal au nombre d'éléments de la population.

Par exemple:

- L'effectif des élèves ayant eu moins de 10 à un examen est l'effectif cumulé (croissant) des élèves ayant eu $1, 2, \ldots, 9$:

$$n_{<10} = n_1 + n_2 + \cdots + n_9$$

- L'effectif des élèves ayant eu 10 ou plus est correspond au nombre d'élèves moins l'effectif des élèves ayant eu moins de 10:

$$n_{\geq 10} = n - n_{<10} = n - n_9 - n_8 - \cdots - n_1$$

Les effectifs

Remarque :

Calculer des effectifs cumulés n'a de sens uniquement si l'on traite de variables ordinales. De tels calculs pour une variable nominale n'a aucun sens.

Les effectifs

Exercice: Soit les notes de 30 étudiants sur 5. Calculer les effectifs cumulés croissants et décroissants.

Note x_i	0	1	2	3	4	5
Effectif n_i	4	6	7	7	4	2

Les effectifs (correction)

Note x_i	Effectif n_i	Effectif cumulé croissant	Effectif cumulé décroissant
0	4	4	30
1	6	10	26
2	7	17	20
3	7	24	13
4	4	28	6
5	2	30	2
Total	$n = 30$		

- Comment interpréter le chiffre 17?
  - 17 étudiants ont eu au plus 2 ou moins.
- Comment interpréter le chiffre 13?
  - 13 étudiants ont eu au moins 3 ou plus.

Sommaire

La fréquence simple

La **fréquence** f_i d'une modalité x_i est égale à l'effectif de la modalité (n_i) divisé par l'effectif total de l'échantillon (N).

Une fréquence est un nombre entre 0 et 1, mais pour qu'elle soit plus lisible, on peut la calculer en **pourcentage** en la multipliant par 100 :

$$f_i = \frac{n_i}{N} \times 100$$

Remarque

La somme des fréquences est toujours égale à 100%:

$$\sum_i f_i = \sum_i \frac{n_i}{N} = \frac{1}{N} \sum_i n_i = \frac{N}{N} = 1$$

La fréquence cumulée

La **fréquence cumulée** est obtenue en cumulant les fréquences simples. Elle peut être:

- **croissante** (fréquence cumulée croissante notée FCC);
- **décroissante** (fréquence cumulée décroissante notée FCD)

Remarque

On peut aussi calculer les fréquences cumulées en divisant les effectifs cumulés par l'effectif total.

Remarque 2 :

Calculer des fréquences cumulées n'a de sens uniquement si l'on traite de variables ordinales. De tels calculs pour une variable nominale n'a aucun sens.

La fréquence cumulée

Exercice : dans un ensemble de 20 individus, on a 5 personnes qui ont un salaire compris entre 1000€ et 1500€, 5 entre 1500€ et 2000€, et 10 au-dessus de 2000€. Calculez la fréquence simple ainsi que les fréquences cumulées croissantes et décroissantes.

La fréquence cumulée

Correction:

Salaire x_i	Effectif n_i	Fréquence simple (%)	FCC (%)	FCD (%)
[1000;1500[	5	25	25	100
[1500;2000[	5	25	50	75
[2000 et plus	10	50	100	50
Total	$n = 20$	100%		

- Que veut dire le 25% dans la colonne *FCC* ?

La fréquence cumulée

Correction:

Salaire x_i	Effectif n_i	Fréquence simple (%)	FCC (%)	FCD (%)
[1000;1500[	5	25	25	100
[1500;2000[	5	25	50	75
[2000 et plus	10	50	100	50
Total	$n = 20$	100%		

- Que veut dire le 25% dans la colonne *FCC* ?
    - Cela veut dire que 25% des individus gagnent 1500€ ou moins.
- Que veut dire le 75% dans la colonne *FCD* ?

La fréquence cumulée

Correction:

Salaire x_i	Effectif n_i	Fréquence simple (%)	FCC (%)	FCD (%)
[1000;1500[	5	25	25	100
[1500;2000[	5	25	50	75
[2000 et plus	10	50	100	50
Total	$n = 20$	100%		

- Que veut dire le 25% dans la colonne *FCC* ?
  - Cela veut dire que 25% des individus gagnent 1500€ ou moins.
- Que veut dire le 75% dans la colonne *FCD* ?
  - Cela veut dire que 75% des individus gagnent 1500€ ou plus.

Sommaire

Classes de valeurs

Il est possible de construire une nouvelle modalité appelée **classe** en regroupant plusieurs modalités d'une variable X.

Par exemple, pour une variable continue, on peut former des intervalles du type $[a, b[$. Par convention, on exclut b mais on inclut a.

Exemple: On observe les salaires avec les modalités suivantes: $x_1 \in [0; 500[$, $x_2 \in [500; 1000[$,$x_3 \in [1000; 1500[$, $x_4 \in [1500; 2000[$. On considère la classe $\tilde{x}$ des salaires entre 0 et 1000 euros.

- $\tilde{x}$ correspond au regroupement des modalités x_1 et x_2;
- Par convention, on exclut la borne supérieure: un individu avec un salaire exactement de 1000 euros n'appartient pas à $\tilde{x}$;
- Les nouvelles modalités de la variable X sont $(\tilde{x}, x_3, x_4)$.

Amplitude et centre de classe

Lorsque les classes sont des intervalles, **l'amplitude** A d'une classe se définit comme la longueur de l'intervalle $[a, b[$

$$A = b - a \tag{2}$$

Dans l'exemple précédent, l'amplitude de $\tilde{x}$ est donc de 1000 euros.

Le centre C d'une classe est la valeur centrale de la classe $[a, b[$:

$$C = a + \frac{A}{2} = \frac{a+b}{2} \tag{3}$$

Pour la classe $X \in [1000; 1500[$, $C = 1250$.

Les regroupements en classes

Dans un échantillon, il existe des règles pour établir les classes ainsi que leur amplitude. Les classes obtenues ne doivent pas:

1. Avoir une amplitude trop large, ce qui affecte la précision des données.
2. Avoir une amplitude trop faible, ce qui donnerait un nombre de classes trop élevé et donc rendrait les données illisibles.

Pour déterminer le nombre de classes k d'un échantillon de taille N (soit l'effectif total), nous pouvons utiliser :

- La **règle empirique**: $2^k \geq N \Rightarrow k \geq \frac{\mathbf{ln}(N)}{\mathbf{ln}(2)}$
- La **règle de Sturge**: $k \sim 1 + 3,3 \times \mathbf{log}_{10}(N)$
- La **règle de Yule**: $k \sim 2,5 \times \sqrt[4]{N} = 2,5\sqrt{\sqrt{N}} = 2,5 \times N^{\frac{1}{4}}$

Exemple

Pour déterminer l'amplitude des classes, on utilise la formule suivante:

$$A = \frac{x_{max} - x_{min}}{k}$$

Avec x_{max} et x_{min} respectivement la plus grande et la plus petite valeur de x dans la série statistique, et k le nombre de classes.

Exemple : Représenter la variable statistique suivante sous forme de classes.

38	39	28	34	55	45	34	36	27	36
26	23	36	51	25	35	26	39	39	24
31	49	45	26	26	39	39	44	45	34
14	26	39	36	36	26	51	49	18	26

Exemple

On commence par **déterminer le nombre de classes**. On sait que l'effectif total est $N = 40$. On applique l'une des 3 règles :

- La règle empirique : $k \geq \frac{\ln(40)}{\ln(2)} = 5,322$
- La règle de Sturge : $k \sim 1 + 3,3 \times \log_{10}(40) = 6,287$
- La règle de Yule : $k \sim 2,5(40)^{\frac{1}{4}} = 6,29$

On choisit donc de regrouper les données en 6 classes.

Remarque

La **règle empirique** peut parfois manquer de précision.

L'amplitude est donnée par: $\frac{55-14}{6} \approx 6,83 \approx 7$

Exemple

On obtient le tableau suivant :

Classe (k)	Effectif (n_i)
$14 \geq X \geq 21$	2
$21 \geq X \geq 28$	10
$28 \geq X \geq 35$	5
$35 \geq X \geq 42$	14
$42 \geq X \geq 49$	4
$49 \geq X \geq 56$	5
Total	$N = 40$

Sommaire

1. Introduction
2. Notions statistiques et collecte des données
3. Les opérateurs somme et produit
   - L'opérateur somme Σ
   - L'opérateur produit $\prod$

Sommaire

L'opérateur somme Σ

Soit $x_1, x_2, x_3, ..., x_n \in \mathbb{R}$ une famille de nombres. Pour deux valeurs k et p avec $k \leq p$, on note :

$$\sum_{i=k}^{p} x_i = x_k + x_{k+1} + x_{k+2} + \cdots + x_p$$

Par exemple,

$$\sum_{i=2}^{4} x_i = x_2 + x_3 + x_4$$

Dans cette somme, i est un **indice** qui varie entre k et p.

La somme de k à p des x_i contient donc $p - k + 1$ **termes**.

L'opérateur somme Σ

On a les propriétés suivantes :

1. $\sum_{i=1}^{n}(x_i + y_i + z_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} z_i$
2. $\sum_{i=1}^{n}(ax_i) = a\sum_{i=1}^{n} x_i$
3. $\sum_{i=1}^{n} b = nb$

Conséquence :

$$\sum_{i=1}^{n}(ax_i + b) = a\sum_{i=1}^{n} x_i + nb$$

Et :

$$\sum_{i=1}^{n}(ax_i + b)^2 = a^2\sum_{i=1}^{n} x_i^2 + 2ab\sum_{i=1}^{n} x_i + nb^2$$

L'opérateur somme Σ

On a :

$$\sum_{i=1}^{n}(ax_i+b)^2=\sum_{i=1}^{n}(a^2x_i^2+2abx_i+b^2)$$

$$\text{car } (A+B)^2=A^2+2AB+B^2$$

Donc

$$\sum_{i=1}^{n}(ax_i+b)^2=(\sum_{i=1}^{n}a^2x_i^2)+(\sum_{i=1}^{n}2abx_i)+\sum_{i=1}^{n}(b^2)$$

$$\Rightarrow \sum_{i=1}^{n}(ax_i+b)^2=a^2\sum_{i=1}^{n}x_i^2+2ab\sum_{i=1}^{n}x_i+nb^2$$

Exercice :

1. Développer : $\sum_{i=1}^{3} x_i$ et $\sum_{j=0}^{2} x_j y_{2-j}$
2. Ecrire avec le symbole Σ : $x_4 + x_5 + x_6 + x_7$ et $z_0 + z_2 + z_4 + z_6$

Solutions :

1.
   - $\sum_{i=1}^{3} x_i = \underbrace{x_1}_{i=1} + \underbrace{x_2}_{i=2} + \underbrace{x_3}_{i=3}$
   - $\sum_{j=0}^{2} x_j y_{2-j} = \underbrace{x_0 y_2}_{j=0} + \underbrace{x_1 y_1}_{j=1} + \underbrace{x_2 y_0}_{j=2}$
2.
   - $x_4 + x_5 + x_6 + x_7 = \sum_{i=4}^{7} x_i$
   - $z_0 + z_2 + z_4 + z_6 = \sum_{p=0}^{3} z_{2p}$

L'opérateur somme Σ

Exercice :

Sachant que :

$$\sum_{i=1}^{4} n_i = N = 10, \sum_{i=1}^{4} n_i x_i = 6 \text{ et } \sum_{i=1}^{4} n_i x_i^2 = 30$$

Calculer :

- $A = \sum_{i=1}^{4} n_i(5x_i - 7)$
- $B = \sum_{i=1}^{4} n_i(2x_i + 3)^2$
- $C = \sum_{i=1}^{4} n_i(2x_i + 3)(x_i - 5)$

L'opérateur somme Σ

- $A = \sum_{i=1}^{4} n_i(5x_i - 7)$
  $A = \sum_{i=1}^{4} (5n_ix_i - 7n_i)$
  $A = 5(\sum_{i=1}^{4} n_ix_i) - 7(\sum_{i=1}^{4} n_i)$
  $A = 5 \times 6 - 7 \times 10 =$ **-40**

L'opérateur somme Σ

- $B = \sum_{i=1}^{4} n_i(2x_i + 3)^2$
  $B = \sum_{i=1}^{4} n_i(4x_i^2 + 12x_i + 9)$
  $B = 4(\sum_{i=1}^{4} n_i x_i^2) + 12(\sum_{i=1}^{4} n_i x_i) + 9(\sum_{i=1}^{4} n_i)$
  $B = 4 \times 30 + 12 \times 6 + 9 \times 10 = \mathbf{282}$

L'opérateur somme Σ

- $C = \sum_{i=1}^{4} n_i(2x_i + 3)(x_i - 5)$
  $C = \sum_{i=1}^{4} n_i(2x_i^2 - 10x_i + 3x_i - 15)$
  $C = 2(\sum_{i=1}^{4} n_i x_i^2) - 10(\sum_{i=1}^{4} n_i x_i) + 3(\sum_{i=1}^{4} n_i x_i) - 15(\sum_{i=1}^{4} n_i)$
  $C = 2 \times 30 - 10 \times 6 + 3 \times 6 - 15 \times 10 =$ **-132**

Sommaire

L'opérateur produit $\prod$

Soit $x_1, x_2, x_3, ..., x_n \in \mathbb{R}$ une famille de nombres. Pour deux valeurs k et p avec $k \leq p$, on note :

$$\prod_{i=k}^{p} x_i = x_k \times x_{k+1} \times ... \times x_p$$

L'opérateur produit $\prod$

On a les propriétés suivantes :

- Pour a réel: $\prod_{i=1}^{n}(ax_i) = a^n(\prod_{i=1}^{n} x_i)$
- $\prod_i(x_i y_i) = (\prod_i x_i)(\prod_i y_i)$
- Pour $x_i > 0$ et n_i entier, on a : $\mathbf{ln}(\prod_i(x_i^{n_i})) = \sum_i n_i \mathbf{ln} x_i$

Exercice :

1. Développer : $\prod_{i=1}^{3} i \times x_i$
2. Calculer : $\prod_{i=1}^{2}(1 + 3i + i^2)$

L'opérateur produit $\prod$ (solutions)

1. $\prod_{i=1}^{3} i \times x_i = \underbrace{(1 \times x_1)}_{i=1} \times \underbrace{(2 \times x_2)}_{i=2} \times \underbrace{(3 \times x_3)}_{i=3} = 6x_1x_2x_3$
2. $\prod_{i=1}^{2}(1+3i+i^2) = \underbrace{(1 + 3 \times 1 + 1^2)}_{i=1} \times \underbrace{(1 + 3 \times 2 + 2^2)}_{i=2} = 5 \times 11 = 55$