

# Cours de Statistique Inférentielle

Jean Christophe meunier

## INTRODUCTION GENERALE

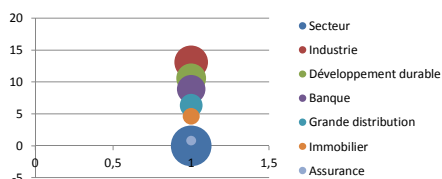
2<sup>ème</sup> Bac, Commerce Extérieur  
Année académique 2015-2016



Expérience pro	non	oui	TOTAL
0	+5 (50)	-13 (16)	66 (66)
1	-24 (77)	+57 (81)	15 (15)
TOTAL	57 (57)	24 (24)	81 (81)

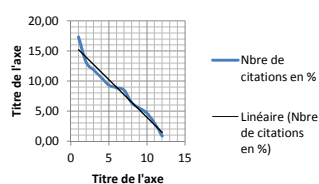
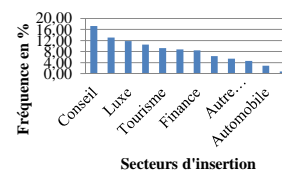
$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - m)^2}{n-1}}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

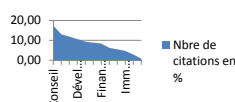


$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

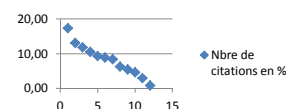
Préférence du secteur d'insertion



Nbre de citations en %



Nbre de citations en %



## Définitions

- Méthode scientifique
  - Consiste à réunir des données chiffrées sur des ensembles nombreux
  - Puis à analyser
  - Enfin à critiquer - tirer des conclusions sur – ces données (Dumas, 1967).
- Distinguer **LA** statistique de **UNE** statistique
  - LA : discipline/science
  - UNE : ensemble de données chiffrées sur un sujet précis

## A quoi servent les statistiques?

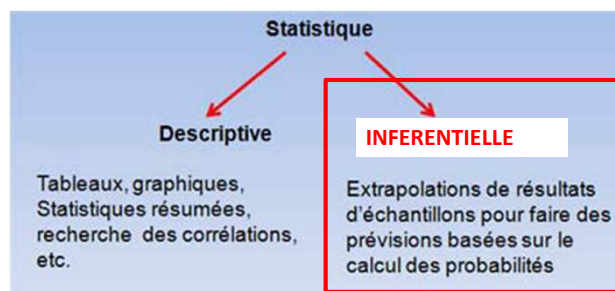
- Organiser, présenter et décrire des données
- Vérifier des relations entre variables
- Tester des hypothèses
- Généraliser à une population, des caractéristiques observées sur des échantillons
- Faire de “bonnes” prévisions

## Deux branches distinctes (1)

- Statistique descriptive
  - Organisation, présentation et analyses des données d'une population/d'un échantillon
  - En vue d'explorer les caractéristiques de cette population/cet échantillon
    - Ex: niveau moyen d'éducation, % allo/autochtones, taux de chômage
- Statistique inférentielle
  - Permet, à partir des résultats obtenus sur un sous-ensemble/échantillon, de généraliser les conclusions à l'ensemble de la population



## Deux branches distinctes (2)



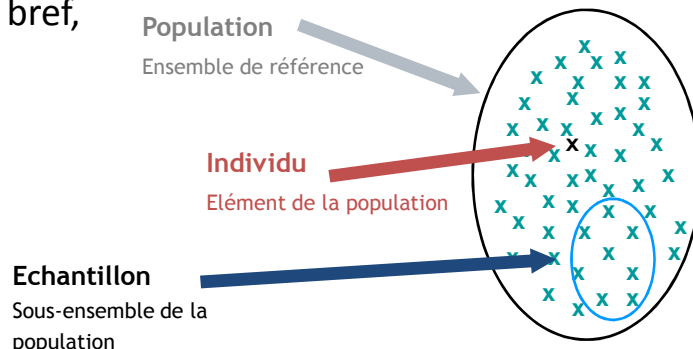
- La statistique, descriptive/inférentielle, est utilisée dans toutes les sciences :
  - Commerce, marketing, gestion, psychologie, agronomie, biologie, médecine, ...

## Echantillon VS. Population (1)

- Population :
  - Ensemble des individus qui partagent les mêmes caractéristiques
  - Représente **TOUS** les individus qui partagent ces caractéristiques
  - Doit être définie avec précision
    - Ex : tous les humains, habitants de WSP, ensemble des étudiants en bac 2 'commerce extérieur' à l'EPFC,...
- Echantillon :
  - Sous-ensemble de la population (tiré au sort ou par méthode qui assure la représentativité de la population)
  - Public cible de la statistique inférentielle : plus facilement appréhendable

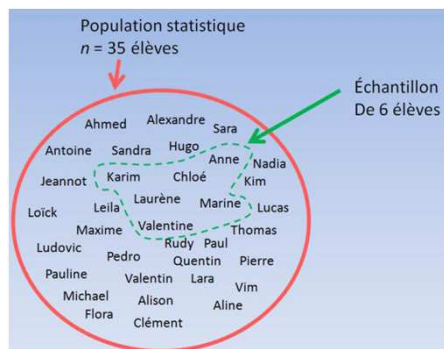
## Echantillon VS. Population (2)

- Individus :
  - Du latin : 'indivisible'
  - Synonyme : Sujets, observations, unités, ...
  - Être humains, êtres vivants, objets concrets ou abstraits
- En bref,



## Application concrète (1)

- Groupe-classe de 35 étudiants
  - Taille de la population :  $N=35$
  - Taille de l'échantillon :  $N=6$



## Application concrète (2)

- Organisation des données/variables
  - Variable = caractéristique observée sur les individus de la population/l'échantillon
  - Tableau à double entrée
    - Lignes : individus, sujets,...
    - Colonnes : variables

Sujets / Variables	V1	V2	V3	...
1				
2				
3				
...				

## Application concrète (3)

- Pour chacun des 35 élèves, différentes caractéristiques recueillies

Prénom	Sexe	Taille (cm)	Nombre de frères et soeurs	Résultat test d'anglais	
Lara	F	168	1	D	
Chloé	F	163	2	D	
Flora	F	161	4	B	
Sara	F	165	1	B	
Kim	G	165	0	A	
Leila	F	168	2	B	
Sandra	F	161	2	D	
Pauline	F	163	1	C	
Anne	F	165	3	A	
Laurene	F	165	1	C	
Lucas	G	170	3	C	
Quentin	G	175	1	D	
Valentine	F	168	1	A	
Clément	G	165	0	B	
Hugo	G	170	1	C	
Marine	F	170	0	C	
Valentin	G	170	1	B	
Aline	F	170	1	B	
Alison	F	165	1	C	
Nadia	F	163	0	D	
Rudy	G	163	1	A	
Jeannot	G	170	4	C	
Pedro	G	175	2	B	
Ahmed	G	175	4	A	
Pierre	G	180	2	C	
Michael	G	170	2	C	
Alexandre	G	175	5	B	
Loick	G	175	1	A	
Antoine	G	180	0	C	
Thomas	G	180	1	B	
Ludovic	G	180	1	A	
Karim	G	182	1	C	
Paul	G	175	3	B	
Maxime	G	182	2	C	
Vim	G	180	0	D	

Connaissance de l'anglais

A : Très bonne connaissance  
 B : bonne connaissance  
 C : Connaissance moyenne  
 D : Connaissance médiocre  
 E : Aucune connaissance

## Application concrète (4)

- Idem sur tableur Excel

	Prénom	Sexe	Taille (cm)	Nombre de frères et soeurs	Résultat test d'anglais		
1	Lara	F	168	1	D		
2	Chloé	F	163	2	D		
3	Flora	F	161	4	B		
4	Sara	F	165	1	B		
5	Kim	G	165	0	A		
6	Leila	F	168	2	B		
7	Sandra	F	161	2	D		
8	Pauline	F	163	1	C		
9	Anne	F	165	3	A		
10	Laurène	F	165	1	C		
11	Lucas	G	170	3	C		
12	Quentin	G	175	1	D		
13	Valentine	F	168	1	A		
14	Clément	G	165	0	B		
15	Hugo	G	170	1	C		
16	Marine	F	170	0	C		
17	Valentin	G	170	1	B		
18	Aline	F	170	1	B		
19	Alison	F	165	1	C		
20	...	...	...	...	...		