

# Échantillonnage et estimation

**Jaouad Madkour**

[jmadkour@uae.ac.ma](mailto:jmadkour@uae.ac.ma)

Faculté des sciences juridiques, économiques et sociales de Tanger  
Département des sciences économiques et de gestion

30 mars 2026



كلية العلوم القانونية والاقتصادية  
والاجتماعية - طنجة  
FSJES TANGER

## EXEMPLE INTRODUCTIF

### LA SOCIÉTÉ ELECTRONICS ASSOCIATES

On veut identifier le profil des 2500 employés de la société Electronics Associates. Les caractéristiques à identifier comprennent leur salaire annuel moyen et la proportion de ceux qui ont suivis la formation au management, mise en place par la société.

### CAS 1 : LES DONNÉES SONT DISPONIBLES ET À MOINDRE COÛT POUR TOUS LES EMPLOYÉS

Les caractéristiques sont **calculées**, avec exactitude, à l'aide de la **statistique descriptive** :

- Le salaire moyen des 2500 employés est  $\mu = 51800\$$ , avec un écart-type  $\sigma = 4000\$$ .
- La proportion des employés ayant suivi la formation est  $p = 0.60$ .

### CAS 2 : LES DONNÉES NE SONT PAS DISPONIBLES ET SONT COÛTEUSES À OBTENIR

Les caractéristiques sont **estimées**, de manière approximative, à partir d'un **échantillon** :

- Le salaire moyen de 30 employés est  $\bar{x} = 51814\$$ , avec un écart-type  $s = 3348\$$ .
- La proportion des employés ayant suivi la formation est  $\bar{p} = 0.63$ .

## COMPARAISON

### Population : 2500 employés

- Moyenne :  $\mu = 51800\$$
- Écart-type :  $\sigma = 4000\$$
- Proportion :  $p = 0.60$

### Échantillon : 30 employés

- Moyenne :  $\bar{x} = 51814\$$
- Écart-type :  $s = 3348\$$
- Proportion :  $\bar{p} = 0.63$

## REMARQUES

- Les caractéristiques de la population sont notées  $\mu$ ,  $\sigma$  et  $p$ , et celles de l'échantillon  $\bar{x}$ ,  $s$  et  $\bar{p}$ .
- Les caractéristiques de l'échantillon sont **légèrement** différentes de celles de la population.

## QUESTIONS

- Comment sélectionner un échantillon à partir d'une population ? et de quelle taille ?
- Comment estimer les paramètres d'une population à partir d'un échantillon ?
- Comment savoir si les estimations sont proches des vrais paramètres de la population ?

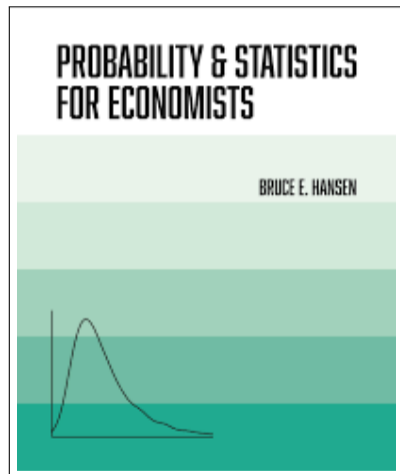
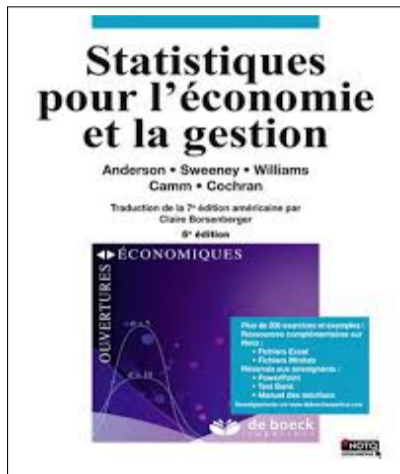
## OBJECTIFS DU COURS

- APPRENDRE À SÉLECTIONNER DES ÉCHANTILLONS
- APPRENDRE À DÉTERMINER DES DISTRIBUTIONS D'ÉCHANTILLONNAGE
- APPRENDRE À RÉALISER DES TESTS D'HYPOTHÈSES
- APPRENDRE À COMPARER DES MOYENNES ET DES PROPORTIONS
- APPRENDRE À RÉALISER DES TESTS D'INDÉPENDANCE

## PRÉ-REQUIS

1. STATISTIQUE DESCRIPTIVE
2. PROBABILITÉS

## RÉFÉRENCES BIBLIOGRAPHIQUES



## SOMMAIRE

CHAPITRE 0 : RAPPELS DE STATISTIQUE DESCRIPTIVE ET DE PROBABILITÉS

CHAPITRE 1 : ÉCHANTILLONNAGE ET DISTRIBUTIONS D'ÉCHANTILLONNAGE

CHAPITRE 2 : ESTIMATION PAR INTERVALLE

CHAPITRE 3 : TESTS D'HYPOTHÈSES

CHAPITRE 4 : COMPARAISONS DE MOYENNES ET ANALYSE DE LA VARIANCE

CHAPITRE 5 : COMPARAISONS DE PROPORTIONS ET TEST D'INDÉPENDANCE

## CHAPITRE 0

### RAPPELS DE STATISTIQUE DESCRIPTIVE ET DE PROBABILITÉS

SECTION 0.1 : RAPPELS DE STATISTIQUE DESCRIPTIVE

SECTION 0.2 : RAPPELS DE PROBABILITÉS

## SECTION 0.1

### RAPPELS DE STATISTIQUE DESCRIPTIVE

#### TERMINOLOGIE

- Les individus sont étudiés selon un caractère donné  $X$  (âge, profession, lieu de résidence, etc.)
- Ce caractère est appelé **variable statistique** et peut prendre plusieurs **modalités**  $x_1, x_2, \dots, x_k$
- Un nombre  $n_i$  d'individus possèdent la modalité  $x_i$ . Ainsi,  $n_i$  est l'**effectif** de  $x_i$ .
- $N$  est l'**effectif total** de tous les individus : i.e.  $N = n_1 + n_2 + \dots + n_k$ .
- Une proportion  $f_i = \frac{n_i}{N}$  d'individus possèdent la modalité  $x_i$ . Ainsi,  $f_i$  est la **fréquence** de  $x_i$ .

#### APPLICATION

$x_i$	20	21	22	23	24	$\Sigma$
$n_i$	10	20	40	20	10	100
$f_i$	0.1	0.2	0.4	0.2	0.1	1

## DESCRIPTION D'UNE VARIABLE QUANTITATIVE

### VARIABLE QUANTITATIVE DISCRÈTE

$x_i$	$x_1$	$x_2$	$\dots$	$x_k$	$\Sigma$
$n_i$	$n_1$	$n_2$	$\dots$	$n_k$	$N$

### VARIABLE QUANTITATIVE CONTINUE

$[x_i, x_{i+1}[$	$[x_1, x_2[$	$[x_2, x_3[$	$\dots$	$[x_k, x_{k+1}[$	$\Sigma$
$n_i$	$n_1$	$n_2$	$\dots$	$n_k$	$N$

- $[x_i, x_{i+1}[$  : la classe  $i$  de la variable statistique  $X$  ;
- $c_i = (x_i + x_{i+1})/2$  : le centre de la classe  $[x_i, x_{i+1}[$ .

## DESCRIPTION NUMÉRIQUE D'UNE VARIABLE QUANTITATIVE

- MOYENNE ARITHMÉTIQUE
- VARIANCE ET ÉCART-TYPE
- FRÉQUENCE

## MOYENNE ARITHMÉTIQUE

**Variable discrète**

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i x_i$$

**Variable continue**

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i c_i$$

## VARIANCE ET ÉCART-TYPE

**Variable discrète**

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^k n_i (x_i - \bar{x})^2 = \overline{x^2} - (\bar{x})^2$$

$$\sigma = \sqrt{\sigma^2}$$

**Variable continue**

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^k n_i (c_i - \bar{x})^2 = \overline{c^2} - (\bar{x})^2$$

$$\sigma = \sqrt{\sigma^2}$$

## FRÉQUENCE

**Variable discrète**

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

**Variable continue**

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

## APPLICATION

### Variable discrète

$x_i$	20	21	22	23	24	$\Sigma$
$n_i$	10	20	40	20	10	100
$f_i$	0.1	0.2	0.4	0.2	0.1	1

$$\bar{x} = \frac{1}{N} \sum_{i=1}^5 n_i x_i = 22$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^5 n_i (x_i - \bar{x})^2 = 1.2$$

$$\sigma = \sqrt{\sigma^2} \approx 1.1$$

### Variable continue

$x_i$	[20,21[	[21,22[	[22,23[	[23,24]	$\Sigma$
$c_i$	20.5	21.5	22.5	23.5	-
$n_i$	10	30	40	20	100
$f_i$	0.1	0.3	0.4	0.2	1

$$\bar{x} = \frac{1}{N} \sum_{i=1}^4 n_i c_i = 22.2$$

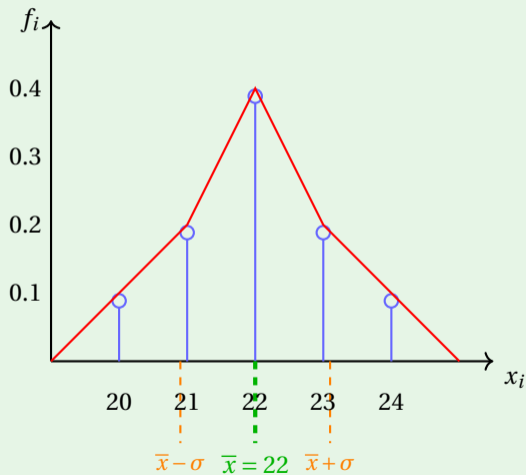
$$\sigma^2 = \frac{1}{N} \sum_{i=1}^4 n_i (c_i - \bar{x})^2 = 0.81$$

$$\sigma = \sqrt{\sigma^2} = 0.9$$

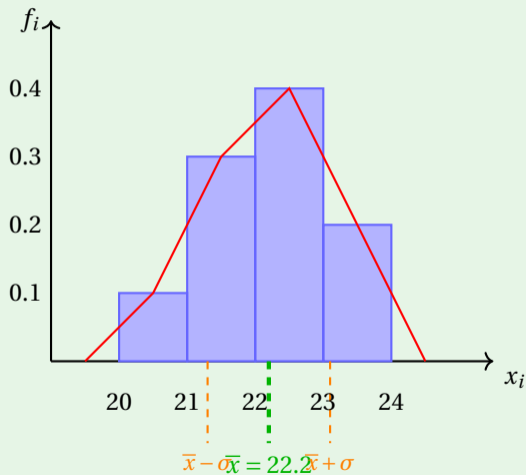
## RÉSUMÉ GRAPHIQUE D'UNE VARIABLE QUANTITATIVE

- DIAGRAMME EN BÂTONS
- HISTOGRAMME
- POLYGONE DES FRÉQUENCES

## DIAGRAMME EN BÂTONS - POLYGONE DES FRÉQUENCES



## HISTOGRAMME - POLYGONE DES FRÉQUENCES



## DESCRIPTION D'UNE VARIABLE QUALITATIVE

### VARIABLE QUALITATIVE NOMINALE

Catégories	Catégorie 1	Catégorie 2	...	Catégorie $k$	$\Sigma$
$n_i$	$n_1$	$n_2$	...	$n_k$	$N$

### VARIABLE QUALITATIVE ORDINALE

Catégories	1 <sup>ère</sup> catégorie	2 <sup>ème</sup> catégorie	...	$k^{\text{ème}}$ catégorie	$\Sigma$
$n_i$	$n_1$	$n_2$	...	$n_k$	$N$

## DESCRIPTION NUMÉRIQUE D'UNE VARIABLE QUALITATIVE

- MOYENNE ARITHMÉTIQUE
- VARIANCE ET ÉCART-TYPE
- FRÉQUENCE

## APPLICATION

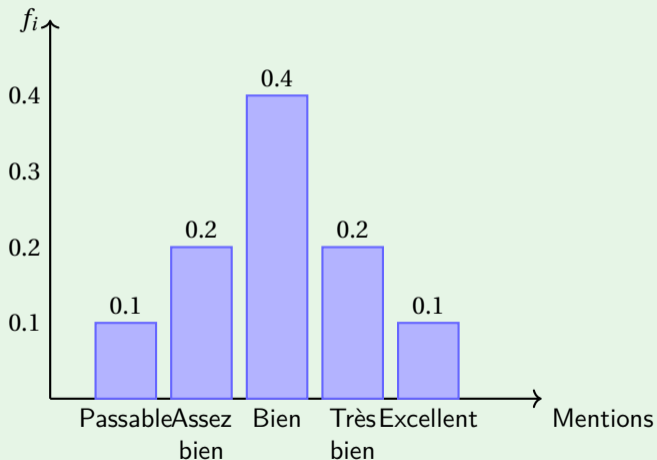
Mentions	Passable	Assez bien	Bien	Très bien	Excellent	$\Sigma$
$n_i$	10	20	40	20	10	100
$f_i$	0.1	0.2	0.4	0.2	0.1	1

Spécialités	Économie	Finance	Marketing	Logistique	Comptabilité	$\Sigma$
$n_i$	10	20	40	20	10	100
$f_i$	0.1	0.2	0.4	0.2	0.1	1

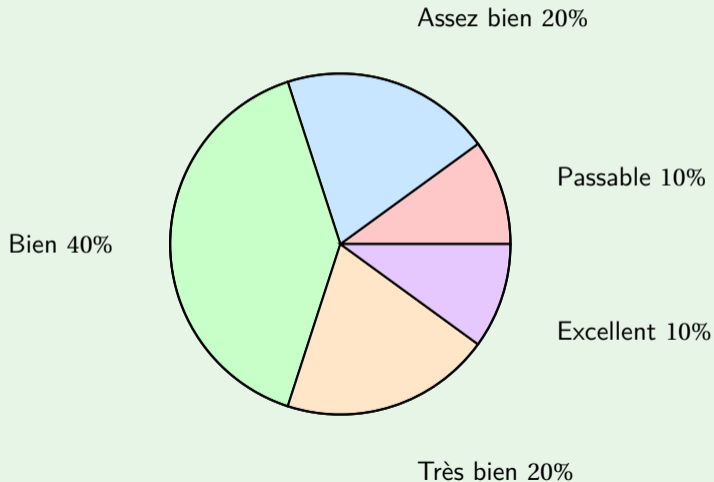
## RÉSUMÉ GRAPHIQUE D'UNE VARIABLE QUALITATIVE

- DIAGRAMME EN BARRES
- DIAGRAMME EN SECTEURS

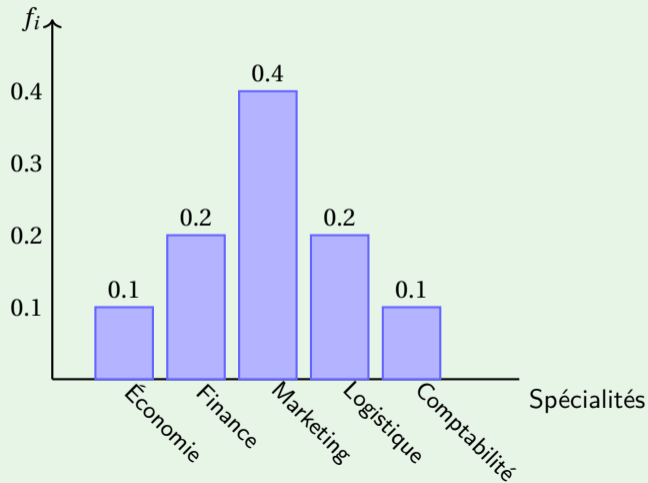
## DIAGRAMME EN BARRES



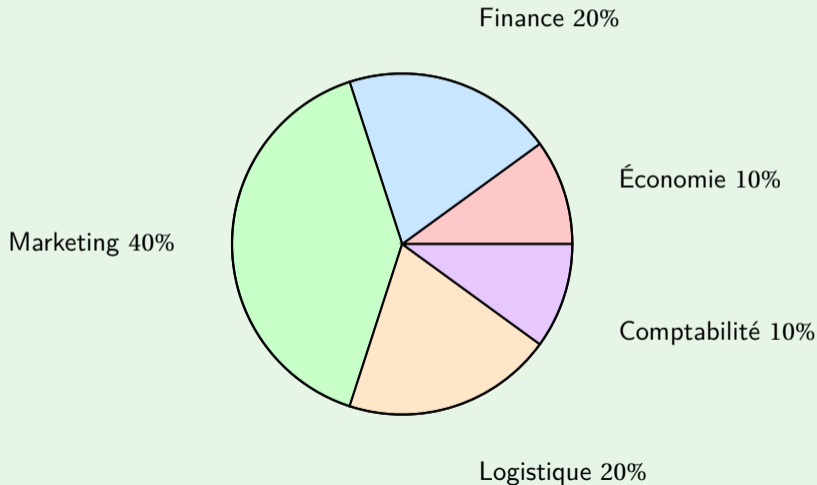
## DIAGRAMME EN SECTEURS



## DIAGRAMME EN BARRES



## DIAGRAMME EN SECTEURS



## SECTION 0.2

### RAPPELS DE PROBABILITÉS

#### TERMINOLOGIE

- **Expérience aléatoire** : expérience dont on ne peut prévoir le résultat avec certitude.
- **Ensemble fondamental**  $\Omega$  : ensemble des résultats possibles d'une expérience aléatoire.
- **Variable aléatoire**  $X$  : sa valeur est déterminée par le résultat  $\omega$  d'une expérience aléatoire.
- **Variable aléatoire discrète** : si  $X(\Omega)$  est dénombrable.
- **Variable aléatoire continue** : si  $X(\Omega)$  est non dénombrable de  $\mathbb{R}$ .

#### EXEMPLE

- **Expérience aléatoire** : Jeu de pile ou face - **Ensemble fondamental** :  $\Omega = \{\text{pile, face}\}$
- **Variable aléatoire**  $X$  : "gagner 5 dh si pile, perdre 3 dh si face" ( $X(\text{pile}) = 5$  et  $X(\text{face}) = -3$ ).
- $X(\Omega) = \{5, -3\}$  est dénombrable.  $X$  est une variable aléatoire discrète.

## VARIABLE ALÉATOIRE DISCRÈTE

Fonction de **probabilités**  $p(x)$  :

$$0 \leq p(x) \leq 1$$

$$\sum_{i=1}^k p(x_i) = 1$$

$x_i$	$x_1$	$x_2$	$\dots$	$x_k$	$\Sigma$
$p(x_i)$	$p(x_1)$	$p(x_2)$	$\dots$	$p(x_k)$	1

## VARIABLE ALÉATOIRE CONTINUE

Fonction de **densité**  $f(x)$  :

$$f(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

## MOMENTS D'UNE VARIABLE ALÉATOIRE

- ESPÉRANCE MATHÉMATIQUE
- VARIANCE ET ÉCART-TYPE

## VARIABLE STATISTIQUE VS VARIABLE ALÉATOIRE

### ILLUSTRATION : DÉ ÉQUILIBRÉ

#### Variable statistique

$x_i$	1	2	3	4	5	6	$\Sigma$
$f_i$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

$$\begin{aligned}\bar{x} &= \frac{1}{N} \sum_{i=1}^k n_i x_i \\ &= \sum_{i=1}^k \frac{n_i}{N} x_i \\ &= \sum_{i=1}^k f_i x_i\end{aligned}$$

#### Variable aléatoire

$x_i$	1	2	3	4	5	6	$\Sigma$
$p(x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

$$E(X) = \sum_{i=1}^k p(x_i) x_i$$

## MOMENTS D'UNE VARIABLE ALÉATOIRE

### ESPÉRANCE MATHÉMATIQUE

**Variable discrète**

$$E(X) = \sum_{i=1}^k x_i p(x_i)$$

**Variable continue**

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

### VARIANCE ET ÉCART-TYPE

**Variable discrète**

$$\sigma^2 = E[(X - E(X))^2] = E(X^2) - E^2(X)$$

$$\sigma = \sqrt{\sigma^2}$$

**Variable continue**

$$\sigma^2 = E[(X - E(X))^2] = E(X^2) - E^2(X)$$

$$\sigma = \sqrt{\sigma^2}$$

## LOIS DE PROBABILITÉS

### Variable discrète

- Loi uniforme discrète :  $X \sim \mathcal{U}\{x_1, x_2, \dots, x_n\}$

$$p(x) = 1/n$$

$$E(X) = \frac{n+1}{2}$$

$$V(X) = \frac{n^2 - 1}{12}$$

### Variable continue

- Loi uniforme continue :  $X \sim \mathcal{U}[a, b]$

$$f(x) = \frac{1}{b-a} \quad \text{pour } a \leq x \leq b$$

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$

## LOIS DE PROBABILITÉS

### Variable discrète

- Loi binomiale :  $X \sim \mathcal{B}(n, p)$

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$E(X) = np$$

$$V(X) = np(1-p)$$

### Variable continue

- Loi normale :  $X \sim \mathcal{N}(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

$$E(X) = \mu$$

$$V(X) = \sigma^2$$

## LOIS DE PROBABILITÉS

### Variable discrète

- Loi Poisson :  $X \sim \mathcal{P}(\mu)$

$$p(x) = \frac{e^{-\mu} \mu^x}{x!}$$

$$E(X) = \mu$$

$$V(X) = \mu$$

### Variable continue

- Loi exponentielle :  $X \sim \mathcal{E}(\lambda)$

$$f(x) = \lambda e^{-\lambda x} \quad \text{pour } x \geq 0$$

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$

## APPLICATION

### Variable discrète

$x_i$	20	21	22	23	24	$\Sigma$
$p(x_i)$	0.1	0.2	0.4	0.2	0.1	1

$$E(X) = \sum_{i=1}^5 x_i p(x_i) = 22$$

$$\sigma^2 = E[(X - E(X))^2] = E(X^2) - E^2(X) = 1.2$$

$$\sigma = \sqrt{\sigma^2} \approx 1.1$$

### Variable continue

$$f(x) = \frac{1}{0.9\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-22}{0.9}\right)^2\right]$$

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = 22$$

$$\sigma^2 = E[(X - E(X))^2] = E(X^2) - E^2(X) = 0.81$$

$$\sigma = \sqrt{\sigma^2} = 0.9$$

## CHAPITRE 1

### ÉCHANTILLONNAGE ET DISTRIBUTIONS D'ÉCHANTILLONNAGE

SECTION 1.1 : SÉLECTIONNER UN ÉCHANTILLON

SECTION 1.2 : ESTIMATION PONCTUELLE

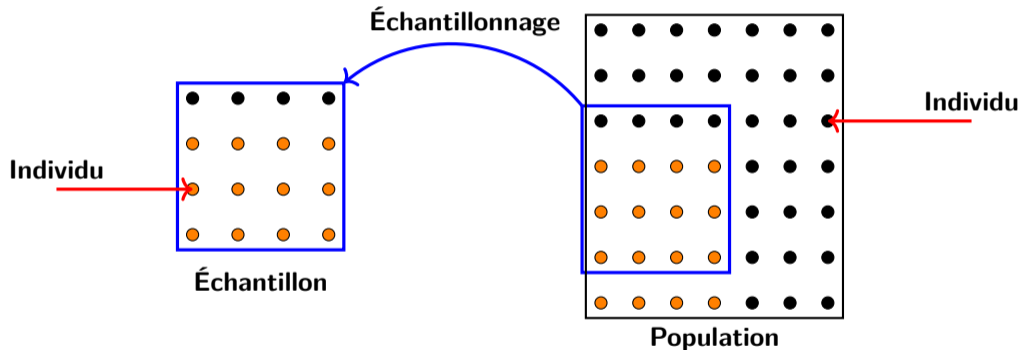
SECTION 1.3 : INTRODUCTION AUX DISTRIBUTIONS D'ÉCHANTILLONNAGE

SECTION 1.4 : DISTRIBUTION D'ÉCHANTILLONNAGE DE  $\bar{x}$

SECTION 1.5 : DISTRIBUTION D'ÉCHANTILLONNAGE DE  $\bar{p}$

SECTION 1.6 : AUTRES MÉTHODES D'ÉCHANTILLONNAGE

# Échantillonnage et distributions d'échantillonnage



## ÉLÉMENT OU INDIVIDU STATISTIQUE

Un **élément** est une entité pour laquelle des données sont collectées.

EXEMPLE : L'employé Dupond de la société Electronics Associates est un élément.

## POPULATION

Une **population** est l'ensemble de tous les éléments auxquels on s'intéresse.

EXEMPLE : Les 2500 employés de la société Electronics Associates constituent la population.

## ÉCHANTILLON

Un **échantillon** est un sous-ensemble de la population.

EXEMPLE : Les 30 employés sélectionnés forment un échantillon.

## REMARQUE

La constitution d'un **échantillon** permet de collecter des données pour répondre à une question concernant la **population**. Les résultats d'un échantillon ne fournissent que des **estimations** des caractéristiques de la population.

- La **moyenne d'échantillon**  $\bar{x}$  est une estimation de la **moyenne de la population**  $\mu$ .
- La **variance d'échantillon**  $s^2$  est une estimation de la **variance de la population**  $\sigma^2$ .
- La **proportion d'échantillon**  $\bar{p}$  est une estimation de la **proportion de la population**  $p$ .
- L'échantillon n'étant qu'une partie de la population, on s'attend à des **erreurs d'estimation**.
- Des méthodes d'**échantillonnage** appropriées permettent d'obtenir de "bonnes" estimations.
- Des **tests statistiques** permettent d'évaluer la justesse des résultats.

## EXEMPLE 1 : ÉLECTIONS DU SÉNAT AMÉRICAIN

Les membres d'un parti politique au Texas sont supposés soutenir un candidat particulier aux élections du Sénat américain, et les leaders du parti voudraient estimer la proportion d'électeurs favorables à leur candidat. Un échantillon de 400 électeurs texans a été sélectionné et 160 de ces 400 électeurs ont indiqué être favorables au candidat. Une **estimation** de la proportion d'électeurs favorables au candidat est donc  $160/400 = 0.40$ .

## EXEMPLE 2 : FABRICATION DE PNEUS

Un fabricant de pneus a conçu un nouveau type de pneu permettant d'accroître le kilométrage effectué, comparativement au nombre de kilomètres effectués avec les pneus actuellement fabriqués par l'entreprise. Pour estimer le nombre moyen de kilomètres effectués avec les nouveaux pneus, le fabricant a sélectionné un échantillon de 120 nouveaux pneus, dans le but de les tester. D'après les résultats du test, la moyenne de l'échantillon est égale à 36500 kilomètres. Par conséquent, une **estimation** du kilométrage moyen pour la population des nouveaux pneus est de 36500 kilomètres.

## DÉFINITIONS

### POPULATION ÉCHANTILLONNÉE

La **population échantillonnée** est la population à partir de laquelle l'échantillon est sélectionné.

### BASE D'ÉCHANTILLONNAGE

Le **base d'échantillonnage** est la liste des éléments à partir de laquelle on tire l'échantillon.

### EXEMPLE 1 : ÉLECTIONS DU SÉNAT AMÉRICAIN (POPULATION FINIE)

- Population échantillonnée : l'ensemble des électeurs du Texas.
- Base d'échantillonnage : la liste électorale officielle.

### EXEMPLE 2 : FABRICATION DE PNEUS (POPULATION INFINIE)

- Population échantillonnée : tous les pneus produits à partir du processus de production.
- Base d'échantillonnage : impossible d'énumérer les éléments d'une population infinie.

### SECTION 1.1

#### SÉLECTIONNER UN ÉCHANTILLON

- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION FINIE
- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION INFINIE

# Échantillonnage et distributions d'échantillonnage

Sélectionner un échantillon



## ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION FINIE

### ÉCHANTILLON PROBABILISTE

Un **échantillon probabiliste** est un échantillon permettant de faire de l'inférence statistique sur la population. Le type le plus simple d'échantillons probabilistes est l'**échantillon aléatoire simple**.

### ÉCHANTILLON ALÉATOIRE SIMPLE

Un **échantillon aléatoire simple** de taille  $n$ , issu d'une population finie de taille  $N$ , est un échantillon sélectionné de manière à ce que chaque combinaison possible de  $n$  éléments a la même chance d'être sélectionnée.

### REMARQUE

Les statisticiens recommandent de sélectionner un échantillon probabiliste lorsque l'on sélectionne un échantillon à partir d'une population finie.

# Échantillonnage et distributions d'échantillonnage

Sélectionner un échantillon



## PROCÉDURE DE SÉLECTION D'UN ÉCHANTILLON ALÉATOIRE SIMPLE

Choisir les  $n$  éléments de l'échantillon un par un, avec ou sans remise, de façon à ce que les éléments restants dans la population aient la même probabilité d'être sélectionnés.

## ÉCHANTILLONNAGE SANS REMISE

Un élément ne peut être sélectionné qu'une seule fois.

## ÉCHANTILLONNAGE AVEC REMISE

Un élément peut être sélectionné plusieurs fois.

## REMARQUE

L'échantillonnage sans remise est la procédure d'échantillonnage la plus utilisée. Lorsque l'on se réfère à un échantillonnage aléatoire simple, il est sous-entendu que l'échantillonnage est sans remise.

# Échantillonnage et distributions d'échantillonnage

Sélectionner un échantillon



## EXEMPLE : LA SOCIÉTÉ ELECTRONICS ASSOCIATES

Un identifiant unique a été affecté à chaque employé de la société Electronics Associates lors de son recrutement. Il s'agit d'un code composé des initiales de l'employé, suivies de deux chiffres indiquant l'année de son recrutement et de quatre chiffres indiquant son ordre de recrutement dans la société. Par exemple, l'identifiant de John Doe, le dix-neuvième employé de la société recruté en 2026 est *JD260019* où *J* et *D* sont ses initiales, 26 est l'année de son recrutement et 0019 son ordre de recrutement.

## PROCÉDURE DE SÉLECTION

Pour sélectionner un échantillon de 30 employés parmi les 2500 employés de la société, on peut se baser uniquement sur les quatre derniers chiffres de leurs identifiants (compris entre 0001 et 2500). Afin de garantir que tous les employés ont la même probabilité d'être sélectionnés, on génère 30 nombres aléatoires à partir d'une loi uniforme discrète  $U_{\{1,2,\dots,2500\}}$ . L'échantillon aléatoire simple sera constitué par les employés dont l'identifiant se termine par les numéros d'ordre tirés au hasard. Le tirage aléatoire se fait de manière informatisée.



### SÉLECTION INFORMATISÉE : CODE R

L'exécution de la commande suivante dans R permet de tirer (sans remise) 30 nombres aléatoires dans une loi uniforme discrète  $U_{\{1,2,\dots,2500\}}$  :

```
sample(1 :2500, 30, replace = FALSE)
```

### RÉSULTAT DE LA SÉLECTION

L'échantillon est formé des employés dont les quatre derniers chiffres des identifiants sont :

---

1842	0467	1923	0876	2104	1358	0921	0345	1769	0502
2381	1124	0678	2045	0193	1456	<b>0019</b>	2267	0412	1634
0975	2418	0563	1890	0327	1246	0605	2179	0831	1502

---

L'employé John Doe dont l'identifiant est *JD260019* fait partie de l'échantillon sélectionné.

### ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION INFINIE

#### POPULATION INFINIE

Une **population** est **infinie** si elle est infiniment grande ou si ses éléments sont générés par un processus pour lequel il n'y a pas de limite quant au nombre d'éléments qui peuvent être générés.

#### REMARQUE

Dans un tel cas, on ne peut pas sélectionner un échantillon aléatoire simple car on ne peut pas définir un cadre d'analyse contenant tous les éléments, mais on sélectionne un **échantillon aléatoire**.

#### ÉCHANTILLON ALÉATOIRE

Un **échantillon aléatoire** de taille  $n$  issu d'une population infinie est un échantillon tel que :

1. Chaque élément sélectionné est issu de la même population.
2. Chaque élément est sélectionné indépendamment des autres.

# Échantillonnage et distributions d'échantillonnage

Sélectionner un échantillon



## EXEMPLE 1 : REMPLISSAGE DE BOÎTES DE CÉRÉALES

Considérons une chaîne de production conçue pour remplir des boîtes de céréales d'un poids moyen de 24 onces. Des échantillons de 12 boîtes remplies via ce processus sont périodiquement sélectionnés par un inspecteur de la qualité pour déterminer si le processus fonctionne correctement ou non.

## CONDITION 1 : CHAQUE ÉLÉMENT SÉLECTIONNÉ EST ISSU DE LA MÊME POPULATION.

Pour s'assurer que la première condition est satisfaite, les boîtes doivent être sélectionnées à peu près au même moment dans le temps. De cette façon, on évite de sélectionner certaines boîtes lorsque la chaîne de production fonctionne correctement et d'autres boîtes lorsque le processus n'est plus sous contrôle et que les boîtes sont sur- ou sous-remplies.

## CONDITION 2 : CHAQUE ÉLÉMENT EST SÉLECTIONNÉ INDÉPENDAMMENT DES AUTRES.

La seconde condition est satisfaite en définissant le processus de production de façon à ce que chaque boîte de céréales soit remplie de manière isolée.

# Échantillonnage et distributions d'échantillonnage

Sélectionner un échantillon



## EXEMPLE 2 : CLIENTS D'UN FAST-FOOD

On veut sélectionner et interviewer un échantillon des clients d'un fast-food afin de déterminer leur profil. Le processus d'arrivée des clients est permanent et il n'y a aucun moyen d'obtenir une liste de tous les clients formant la population.

## CONDITION 1 : CHAQUE ÉLÉMENT SÉLECTIONNÉ EST ISSU DE LA MÊME POPULATION.

On doit sélectionner l'échantillon à partir des personnes qui entrent dans le restaurant et font un achat pour garantir la première condition. Si une personne sélectionnée est entrée dans le restaurant sans faire d'achat, cette personne n'est pas un client et la première condition est violée.

## CONDITION 2 : CHAQUE ÉLÉMENT EST SÉLECTIONNÉ INDÉPENDAMMENT DES AUTRES.

L'objectif de la seconde condition est d'éviter un **biais de sélection**. Un biais de sélection survient par exemple en sélectionnant des clients d'une tranche d'âge particulière et pas d'autres tranches d'âge. Un biais de sélection peut être évité en s'assurant que les clients sont sélectionnés indépendamment les uns des autres.



### SECTION 1.2

#### ESTIMATION PONCTUELLE

#### PARAMÈTRE

Un **paramètre** est une caractéristique numérique d'une population, telle que sa moyenne  $\mu$ , son écart type  $\sigma$ , sa proportion  $p$  possédant un caractère donné, etc. La valeur d'un paramètre est généralement inconnue et doit être estimée à partir d'un échantillon à l'aide d'un **estimateur**.

#### ESTIMATEUR

Un **estimateur** est une formule mathématique qui utilise les données d'un échantillon pour calculer une **estimation** d'un paramètre inconnu d'une population.



### ESTIMATION

Une **estimation** d'un paramètre inconnu est une approximation de ce paramètre calculée à partir d'un échantillon en utilisant un estimateur.

### ESTIMATION PONCTUELLE

Une **estimation ponctuelle** d'un paramètre est la valeur numérique unique obtenue en appliquant un estimateur à un échantillon.

### ESTIMATION PAR INTERVALLE DE CONFIANCE

Une **estimation par intervalle de confiance** d'un paramètre est la fourchette des valeurs numériques obtenue en appliquant un estimateur à un échantillon et pouvant contenir la vraie valeur du paramètre avec un certain niveau de confiance.



### EXEMPLE INTRODUCTIF

#### LA SOCIÉTÉ ELECTRONICS ASSOCIATES

On veut identifier le profil des 2500 employés de la société Electronics Associates (EAI). Les caractéristiques à identifier comprennent le salaire annuel moyen des employés et la proportion ayant suivi le programme de formation au management, mis en place par la société.

En utilisant l'ensemble de données de EAI, on peut calculer les caractéristiques numériques, appelées **paramètres**, de toute la population :

- Moyenne de la population :  $\mu = 51800\$$
- Écart type de la population :  $\sigma = 4000\$$
- Proportion de la population :  $p = 0.6$

# Échantillonnage et distributions d'échantillonnage

Estimation ponctuelle



Supposons que les informations nécessaires sur les employés de la société EAI ne sont pas disponibles dans les bases de données de la société. Alors, comment peut-on obtenir des **estimations** des paramètres de la population, en utilisant un échantillon d'employés à la place de la population constituée de 2500 employés ?

Supposez que l'on utilise un échantillon de 30 employés.



### SECTION 1.3

#### INTRODUCTION AUX DISTRIBUTIONS D'ÉCHANTILLONNAGE

- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION FINIE
- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION INFINIE
- APPLICATIONS



## SECTION 1.4

### DISTRIBUTION D'ÉCHANTILLONNAGE DE $\bar{x}$

- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION FINIE
- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION INFINIE
- APPLICATIONS

## SECTION 1.5

### DISTRIBUTION D'ÉCHANTILLONNAGE DE $\bar{p}$

- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION FINIE
- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION INFINIE
- APPLICATIONS



### SECTION 1.6

#### AUTRES MÉTHODES D'ÉCHANTILLONNAGE

- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION FINIE
- ÉCHANTILLONNAGE À PARTIR D'UNE POPULATION INFINIE
- APPLICATIONS