

Chapitre 14 – Échantillonnage

Cours – Niveau Seconde

1 Population et échantillon

Une **population** est l'ensemble sur lequel on souhaite obtenir des informations.

Un **échantillon** est une partie de cette population.

Exemple

On souhaite connaître la proportion d'élèves d'un lycée qui utilisent les transports en commun.

La population est l'ensemble des élèves du lycée.

On interroge seulement 100 élèves : ces 100 élèves constituent un échantillon.

2 Proportion dans une population

Supposons qu'une population comporte :

$$N$$

individus.

Parmi eux :

$$M$$

possèdent une certaine caractéristique.

La proportion correspondante est :

$$p = \frac{M}{N}$$

avec :

$$0 \leq p \leq 1.$$

Exemple

Dans un lycée de 800 élèves, 520 utilisent les transports en commun.

La proportion est :

$$p = \frac{520}{800} = 0,65.$$

Donc :

$$65\%$$

des élèves utilisent les transports en commun.

3 Fréquence dans un échantillon

On prélève un échantillon de taille :

$$n.$$

On observe :

k

individus possédant la caractéristique étudiée.

La fréquence observée est :

$$f = \frac{k}{n}.$$

Exemple

On interroge 100 élèves.

Parmi eux, 68 utilisent les transports en commun.

La fréquence observée est :

$$f = \frac{68}{100} = 0,68.$$

Ainsi :

$$f = 68\%.$$

4 Proportion et fréquence

La proportion p concerne toute la population.

La fréquence f concerne un échantillon particulier.

En général :

$$f \neq p.$$

Mais si l'échantillon est suffisamment grand et correctement choisi, la fréquence observée est souvent proche de la proportion réelle.

5 Fluctuation d'échantillonnage

Si on prélève plusieurs échantillons de même taille dans une même population, on n'obtient pas exactement la même fréquence à chaque fois.

C'est ce qu'on appelle la **fluctuation d'échantillonnage**.

Exemple

Supposons que :

$$p = 0,5.$$

On réalise plusieurs échantillons de taille 100.

On peut obtenir :

$$f_1 = 0,47,$$

$$f_2 = 0,53,$$

$$f_3 = 0,49,$$

$$f_4 = 0,56.$$

Les fréquences fluctuent autour de :

$$0,5.$$

6 Effet de la taille de l'échantillon

Plus la taille de l'échantillon est grande, plus les fréquences ont tendance à se rapprocher de la proportion réelle.

Par exemple, pour une proportion :

$$p = 0,5,$$

on peut observer :

$$0,3, 0,6, 0,4$$

avec des petits échantillons,
mais plutôt :

$$0,49, 0,51, 0,50$$

avec de très grands échantillons.

7 Simulation

Une simulation permet de reproduire une expérience aléatoire un grand nombre de fois.
Elle peut être réalisée :

- avec Python ;
- avec un tableur ;
- avec une calculatrice.

Exemple

Pour simuler une population dans laquelle :

$$p = 0,6,$$

on peut associer :

- succès avec probabilité 0,6 ;
- échec avec probabilité 0,4.

8 Intervalle de fluctuation approché

Dans certains cas étudiés en Seconde, pour un échantillon de taille n et une proportion p , on peut utiliser un intervalle de fluctuation approché de la forme :

$$\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$$

sous certaines conditions.

Cet intervalle donne une zone dans laquelle la fréquence observée a de fortes chances de se trouver.

Exemple

Supposons :

$$p = 0,5$$

et :

$$n = 100.$$

Alors :

$$\frac{1}{\sqrt{100}} = 0,1.$$

Donc :

$$I = [0,5 - 0,1; 0,5 + 0,1].$$

Ainsi :

$$I = [0,4; 0,6].$$

Une fréquence comme :

$$0,52$$

est donc compatible avec cette proportion.

9 Tester une fréquence observée

Supposons qu'une entreprise affirme que :

$$p = 0,7$$

de ses clients sont satisfaits.

On interroge :

$$n = 100$$

clients.

L'intervalle approché est :

$$\left[0,7 - \frac{1}{10}; 0,7 + \frac{1}{10}\right].$$

Donc :

$$[0,6; 0,8].$$

Si la fréquence observée vaut :

$$f = 0,73,$$

elle appartient à l'intervalle.

Le résultat est donc compatible avec l'affirmation.

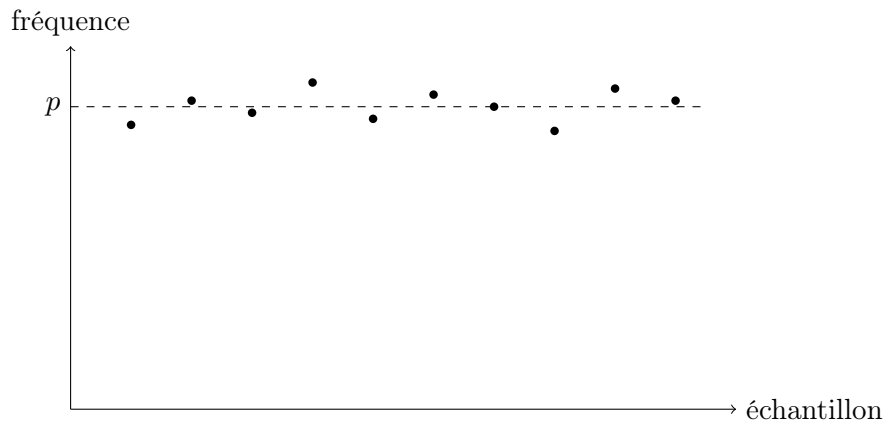
Si on observe :

$$f = 0,48,$$

le résultat semble difficilement compatible avec l'affirmation.

10 Représentation graphique

On peut représenter plusieurs fréquences obtenues lors de simulations.



Les fréquences varient autour de la proportion p .

11 Attention à l'échantillon

Pour qu'un échantillon fournisse des informations utiles, il doit être choisi de manière pertinente.

Un échantillon biaisé peut donner une vision incorrecte de la population.

Par exemple, interroger uniquement les élèves d'une seule classe ne permet pas forcément de représenter correctement tout un lycée.

12 Méthodes à connaître

À la fin du chapitre, il faut savoir :

1. distinguer population et échantillon ;
2. calculer une proportion ;
3. calculer une fréquence ;
4. comprendre la fluctuation d'échantillonnage ;
5. comprendre l'effet de la taille d'un échantillon ;
6. simuler des échantillons ;
7. calculer un intervalle de fluctuation approché simple ;
8. comparer une fréquence observée à un intervalle ;
9. interpréter les résultats d'une simulation ;
10. identifier un éventuel biais d'échantillonnage.

À retenir

$$p = \frac{\text{effectif dans la population}}{\text{effectif total}}$$

$$f = \frac{\text{effectif observé dans l'échantillon}}{n}$$

et, dans les situations où la formule est applicable :

$$I = \left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right].$$