

Exercices du Ch 15 « Échantillonnage »

Ex 15.1. Si on note P pour pile et F pour face, un exemple d'échantillon de taille 12 que l'on pourrait obtenir est :

P F F P P P F P F P P F

Ex 15.2 a) Recherche d'exemples selon les calculatrices :

Calculatrice TI 83 : Chaîne Youtube "Les Tutoriels",

Vidéo : "Simulation du jeu de pile ou face avec la TI-83
Première CE"

Calculatrice Casio : Chaîne Youtube "Frédéric Thebaud",

Vidéo : "Lancé de dé GRAPH 85+ avec"

Calculatrice Numworks : Chaîne Youtube "Maths chez nous"

Vidéo : "Calculatrice Numworks : Génération aléatoire de nombres"

(Pour tirer à pile ou face, générer aléatoirement un entier 0 pour pile, 1 pour face, avec $\text{randint}(0,1)$).

En cas de difficulté, nous reprendrons cette question lorsque nous serons ensemble en classe (amenez votre calculatrice!).

b) Un exemple de résultats :

En effectifs :

Pile	Face

En fréquences :

	Pile	Face
Fréquence :	0,48	0,52

Ex 15.3. i) On applique le théorème 15.1 pour déterminer l'intervalle de fluctuation.

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

Ici $p = 0,59$ (proportion "normale" de visiteurs français)

Dans cette question, $n = 50$

$$I = \left[0,59 - \frac{1}{\sqrt{50}} ; 0,59 + \frac{1}{\sqrt{50}} \right] = [0,469 ; 0,731].$$

La fréquence observée $f = 0,67$ est dans cet intervalle, donc l'hypothèse ne semble pas avoir eu d'impact.

$$2^o) I_{ci}, m = 500$$

$$I_2 = \left[0,59 - \frac{1}{\sqrt{500}} ; 0,59 + \frac{1}{\sqrt{500}} \right] \approx [0,565 ; 0,635]$$

La fréquence observée $f = 0,63$ n'est pas dans cet intervalle (elle est au-delà), donc finalement l'échantillon semble avoir eu un impact positif sur la fréquentation des usagers fauchés.

Cet impact n'était pas mesurable sur un échantillon trop petit.

Ex. 15.6. On a $p = 0,54$ et $m = 350$.

a) Intervalle de fluctuation au seuil des 95% :

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

$$\approx [0,487 ; 0,593]$$

b) On ne peut rien affirmer, car une élection au premier tour ne donne un résultat définitif que si le candidat a la majorité absolue ($> 0,5$).

Ainsi il y a 95% de chances que le résultat soit compris dans l'intervalle I , mais on n'en sait pas plus.

Ex. 15.5.

a) Pour $m = 80$

• À pied au $\bar{a}$ vélo, la fréquence d'usage est $p = 0,3$.

$$\text{Donc } I_1 = \left[p - \frac{1}{\sqrt{m}} ; p + \frac{1}{\sqrt{m}} \right] \approx [0,188 ; 0,412]$$

La fréquence $f_1 = 0,35$ de l'échantillon appartient à I_1 , donc l'échantillon est représentatif.

• En transports publics, $p = 0,3$; $I_2 \approx [0,188 ; 0,412]$.

$f_2 = 0,4$, et $f_2 \notin I_2$, donc l'échantillon est non représentatif.

• En voiture, $p = 0,4$, $I_3 \approx [0,288 ; 0,512]$

$f_3 = 0,25$ et $f_3 \notin I_3$, donc ici l'échantillon n'est pas représentatif.

b) Pour $m = 40$

• À pied au $\bar{a}$ vélo, $p = 0,4$, $I'_1 \approx [0,362 ; 0,438]$,

$f'_1 = 0,35 \in I'_1$, l'échantillon est représentatif.

• en transports en commun, $I'_2 \approx [0,442; 0,458]$ 1573

$f'_2 \approx 0,4 \in I'_2$, l'échantillon est représentatif.

• en ritue, $I'_3 \approx [0,242; 0,558]$

$f'_3 \approx 0,25 \in I'_3$, l'échantillon est représentatif.

Ex 15.6 Ici, la taille de l'échantillon est $n = 1362$.

Et la fréquence théorique (pourcentage d'individus A+ dans la population totale) est $p = 0,39$.

L'intervalle de fluctuation est donc $I = \left[0,39 - \frac{1}{\sqrt{1362}}; 0,39 + \frac{1}{\sqrt{1362}} \right]$

$$I \approx [0,3627; 0,4173]$$

La fréquence observée $f = 0,26$ n'est pas dans cet intervalle, donc l'échantillon n'est pas représentatif.

Ex 15.7

1°) Ici, $n = 500$

fréquence de femmes parmi les candidats:

$$p_1 = \frac{\text{nombre de femmes}}{\text{nombre total de candidats}} = \frac{706}{1638 + 706} \approx 0,3287.$$

Il s'agit de la "fréquence théorique" (notée p dans le cours) car théoriquement, si la pitié est respectée, on devrait retrouver cette proportion de femmes parmi les candidats admis.

L'intervalle de fluctuation est donc $I_1 \approx \left[0,3287 - \frac{1}{\sqrt{500}}; 0,3287 + \frac{1}{\sqrt{500}} \right]$

$$I_1 \approx [0,28398; 0,37342].$$

Or la fréquence de femmes parmi les reçus ("fréquence observée") est $f_1 = \frac{188}{500} \approx 0,376$

$f_1 \notin I_1$, donc on peut conclure que la pitié n'a pas été respectée ("trop" de femmes ont été reçues par rapport à la fréquence attendue).

2°) On a toujours $n = 500$

$$p_2 = \frac{1356}{1356 + 698} \approx 0,6602.$$

$$I_2 \approx \left[0,6602 - \frac{1}{\sqrt{500}}; 0,6602 + \frac{1}{\sqrt{500}} \right] \approx [0,615; 0,7069]$$

et $f_2 = \frac{361}{500} \approx 0,682 \in I_2$, donc la pitié a été respectée.

Ex. 15.8. Ici, $n = 200$, $p = 0,37$ d'où l'intervalle de fluctuation $I = \left[0,37 - \frac{1}{\sqrt{200}} ; 0,37 + \frac{1}{\sqrt{200}} \right]$

$$\approx [0,299 ; 0,441]$$

$$f = \frac{92}{200} \approx 0,46 \notin I$$

Cette fréquence est au-delà de l'intervalle de fluctuation normal des ventes : il semble donc que la campagne ait eu un impact positif sur les ventes.

Ex. 15.9.

$$1^{\circ}) I_A = \left[0,07 - \frac{1}{\sqrt{400}} ; 0,07 + \frac{1}{\sqrt{400}} \right] \approx [0,02 ; 0,12]$$

$$I_B = \left[0,07 - \frac{1}{\sqrt{650}} ; 0,07 + \frac{1}{\sqrt{650}} \right] \approx [0,031 ; 0,109]$$

$$2^{\circ}) f_A = \frac{33}{400} \approx 0,0825 \in I_A$$

$$f_B = \frac{78}{650} \approx 0,12 \notin I_B$$

Donc la production dans l'atelier A est de bonne qualité, mais pas celle dans l'atelier B.

Intervalle de fluctuation

(90% des exo de bac)

$$I = \left[\cancel{p} - 1,96 \frac{\sqrt{p(1-p)}}{\sqrt{n}} ; p + 1,96 \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$$

probabilité théorique, taille de l'échantillon, dans la population totale.

- Quand l'énoncé parle de "fluctuation". (103 p 606)
- Quand on suppose que la fréquence d'un caractère dans la population est p (au second degré) mais, dit que c'est p et que l'on veut vérifier comment évolue cette fréquence dans un échantillon (129) ou si la fréquence observée dans un échantillon est "normale" / par rapport à la fréquence "attendue" / compte tenu de la fréquence dans la population totale (p. de Bac)

⚠ Dans tous les cas, avant d'utiliser les formules qui donnent ces intervalles, il faut vérifier les conditions d'application du th. de Darw-lyfke:

$$n \geq 30 ; np \geq 5 \text{ et } n(1-p) \geq 5$$

- Dans les exercices: 103 p 606, 129 p 614, p. de Bac.

Intervalle de confiance

(10% des exo de bac)

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

fréquence observée dans l'échantillon, taille de l'échantillon

Quand l'utiliser?

- Quand l'énoncé parle d'"intervalle de confiance"
- Quand on a observé une fréquence sur un échantillon (de taille n), et que l'on veut estimer ce qui se passe dans la population totale (cas des sondages).

- Dans les exercices: 116 p 109, 118 p 109