FLUCTUATION D'ÉCHANTILLONNAGE

I) VOCABULAIRE

1) Échantillonnage

On s'intéresse à la proportion p d'individus ayant une caractéristique donnée dans une population très importante.
On est alors souvent dans l'impossibilité d'interroger tous les individus et on se contente de déterminer la proportion étudiée dans un <u>échantillon</u> de la population.
Cette proportion dans l'échantillon est notée f (parfois $\hat{p}$)

Ex :

Population :	Proportion étudiée :	Échantillon :
43 millions de Français en âge de voter	proportion de ceux qui vont voter pour Monsieur X	1000 électeurs choisis au hasard
Infinité de lancer de dés	probabilité d'obtenir un 6	100 lancers de dés
Baguettes de 250 g achetées dans ma boulangerie	proportion des baguettes qui pèsent moins de 230 g	100 dernières baguettes achetées

2) Fluctuation

Lorsque l'on compare plusieurs échantillons de même taille, on constate que les valeurs de f observées varient autour de la proportion p obtenue pour la population totale.
Ce phénomène est appelé « fluctuation d'échantillonnage ».
On observe que plus les échantillons sont grands, plus les fluctuations de f sont petites et plus les valeurs de f tendent vers p.

3) Expérience aléatoire

Une grande partie des situations concrètes sur lesquelles nous allons travailler sont issues du hasard : lancers de dés, d'une pièce de monnaie, choix d'une carte dans un jeu…
On parle « d'expérience aléatoire ».

4) Simulation

L'intérêt des expériences aléatoires est que l'on peut les simuler avec une calculatrice ou un ordinateur.

	Texas	Casio	Tableur
$x \in [0\ ;\ 1[$	NbrAléat	Ran#	=ALEA()
$n \in \{1;2;3;4;5;6\}$	NbrAléatEnt(1,6)	Int(6×Ran#+1)	=ALEA.ENTRE. BORNES(1;6)

II) INTERVALLE DE FLUCTUATION / DE CONFIANCE

Nous avons observé que plus les échantillons sont grands, plus la fluctuation entre les échantillons est faible.
Les statisticiens ont démontré ce résultat et ont établi un moyen simple pour prévoir l'amplitude de cette fluctuation en fonction de la taille de l'échantillon.

Dans les exercices, nous rencontrerons deux cas :

1) Intervalle de fluctuation

On connaît la fréquence p d'un résultat sur l'ensemble d'une population et on veut prévoir dans quelle fourchette devrait être la fréquence f du même résultat dans un échantillon de taille n.

Propriété 1 :
Si $p \in [0,2 ; 0,8]$ et $n > 25$ alors il y a environ 95 % de chances que l'on ait : $f \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$
On appelle cet intervalle : « intervalle de fluctuation ».

Ex 1 :
Pierre lance une pièce de monnaie 100 fois et obtient 41 fois pile seulement. Il en déduit que cette pièce est truquée. Qu'en pensez-vous ?

Rédaction :
La probabilité d'avoir pile est $p = 0,5$
La fréquence du pile dans l'échantillon est $f = 0,41$
La taille de l'échantillon est $n = 100$
On a : $p \in [0,2 ; 0,8]$ et $n > 25$

$$p - \frac{1}{\sqrt{n}} = 0,5 - \frac{1}{10} = 0,4$$

$$p + \frac{1}{\sqrt{n}} = 0,5 + \frac{1}{10} = 0,6$$

Il y a donc environ 95 % de chances que l'on ait entre 40 et 60 fois pile.
Cet échantillon ne permet donc pas de conclure que la pièce soit truquée.

2) Intervalle de confiance

On connaît la fréquence f d'un résultat pour un échantillon de taille n et on veut prévoir dans quelle fourchette devrait être la fréquence p du même résultat sur l'ensemble de la population.

D'après la propriété 1, on peut écrire :

$$f \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right] \Leftrightarrow p - \frac{1}{\sqrt{n}} \leqslant f \text{ et } f \leqslant p + \frac{1}{\sqrt{n}}$$

$$\Leftrightarrow p \leqslant f + \frac{1}{\sqrt{n}} \text{ et } f - \frac{1}{\sqrt{n}} \leqslant p$$

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

Propriété 2 :
Si $f \in [0,2 ; 0,8]$ et $n > 25$ alors il y a environ 95 % de chances que l'on ait : $p \in \left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}}\right]$
On appelle cet intervalle : « intervalle de confiance ».

Ex 2 :
A Joliville, il y a 85842 électeurs. Sur 1068 personnes interrogées, 550 déclarent qu'elles vont voter pour Monsieur Jolicœur. Monsieur Jolicœur va-t-il gagner ? (on suppose qu'il n'y a ni abstention, ni votes blancs)

Rédaction :
Parmi les personnes interrogées, la fréquence de ceux qui vont voter pour Monsieur Jolicœur est : $f = \frac{550}{1068} \approx 0,515$
La taille de l'échantillon est $n = 1068$
Appelons p la proportion de ceux qui vont voter pour Monsieur Jolicœur parmi l'ensemble des électeurs.

On a : $f \in [0,2 ; 0,8]$ et $n > 25$

$$f-\frac{1}{\sqrt{n}}=\frac{550}{1068}-\frac{1}{\sqrt{1068}}\approx 0,484$$

$$f+\frac{1}{\sqrt{n}}=\frac{550}{1068}+\frac{1}{\sqrt{1068}}\approx 0,546$$

Il y a donc environ 95 % de chances que le pourcentage de vote en faveur de Monsieur Jolicœur soit compris entre 48,4 % et 54,6 %.
Ce dernier n'est pas sûr de gagner, mais il a de sérieuses chances.

p153: 48
p154: 49, 51, 52, 53, 54, 55, 56
p155: 58, 59
p156: 61, 64
p200: 72, 73, 74

Pas dans [0,2 ; 0,8] :
p155: 60

Algo :
p154: 57
p156: 63
p187: 11
p196: 58
p198: 63, 64, 65, 66
p199: 69