Probabilités

VII] Fluctuation d'échantillonnage

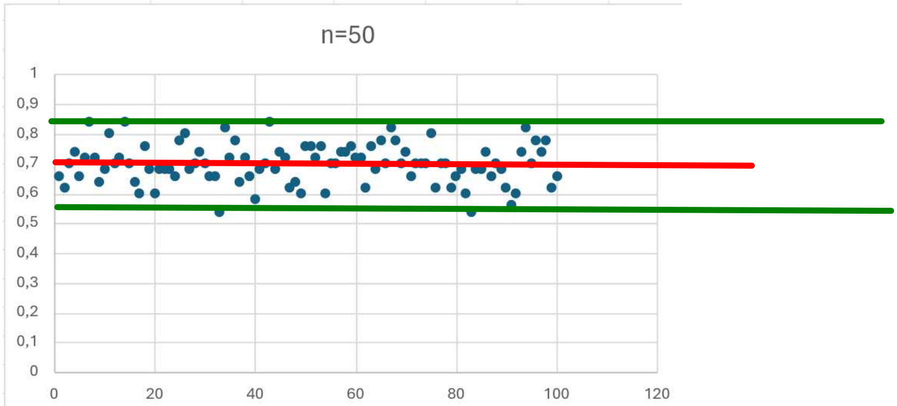
1. Vocabulaire

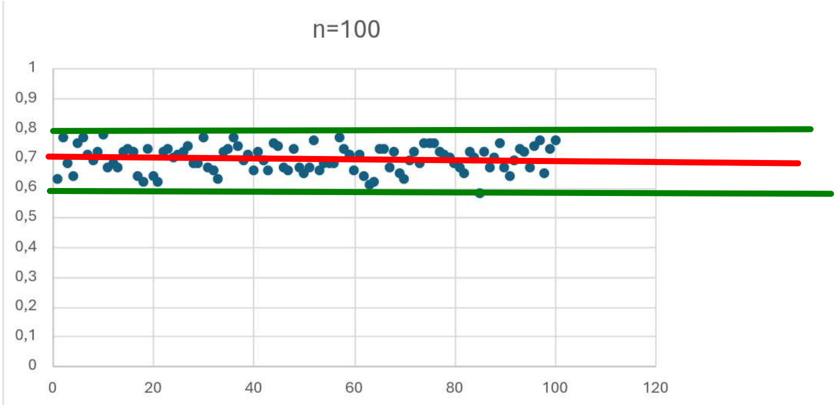
Un échantillon de taille n est la liste des résultats obtenus lorsqu'on répète n fois une même expérience aléatoire de façon indépendante. Sur plusieurs échantillons de même taille, la fréquence d'un caractère observé varie d'un échantillon à l'autre. C'est ce qu'on appelle la **fluctuation d'échantillonnage**.

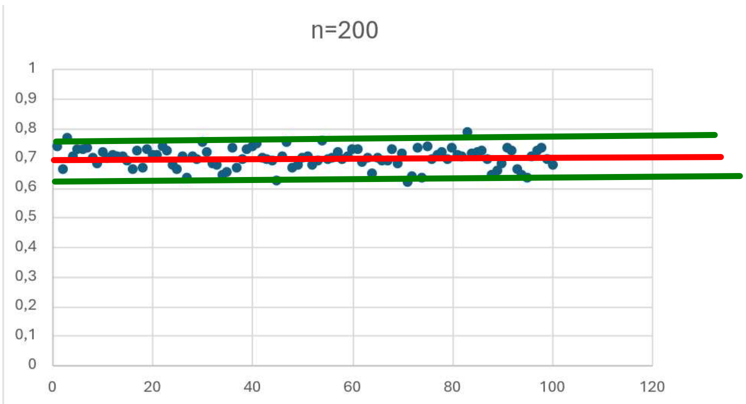
Exemple : On répète n fois une expérience de probabilité $0,7$.

Pour chacun des cas, tracer les droites d'équation $y = 0,7$, $y = 0,7 - \frac{1}{\sqrt{n}}$ et $y = 0,7 + \frac{1}{\sqrt{n}}$.

- $n = 50$ $\frac{1}{\sqrt{50}} = 0,14$ $0,56$ $0,84$
- $n = 100$ $\frac{1}{\sqrt{100}} = 0,1$ $0,6$ $0,8$
- $n = 200$ $\frac{1}{\sqrt{200}} = 0,07$ $0,63$ $0,77$







2. Estimation d'une proportion par une fréquence observée sur un échantillon

a. proportion dans un échantillon

On s'intéresse à l'apparition d'un certain caractère dans une population. On note p la proportion théorique d'individus présentant ce caractère dans la population totale et on cherche à donner une estimation de p en minimisant le risque d'erreur.

b. Théorème

Lorsque n est grand, la fréquence observée f d'individus présentant le caractère étudié dans un échantillon de taille n est telle que $|f-p| \leqslant \frac{1}{\sqrt{n}}$ dans une grande majorité des cas. (95%)

c. Formule

$$\boxed{p - \frac{1}{\sqrt{n}} \leq f \leq p + \frac{1}{\sqrt{n}}}$$

$$f - \frac{1}{\sqrt{n}} \leq p \leq f + \frac{1}{\sqrt{n}}$$

d. Exemple

Une urne contient 5000 boules dont 1500 sont noires et les autres rouges. On prélève un échantillons de taille 100 dans cette urne et on observe la fréquence de boules rouges.

$p = \frac{3500}{5000} = 0,7$ $\quad \frac{1}{\sqrt{100}} = 0,1$

$0,6 \leq f \leq 0,8$ entre 60 et 80 boules rouges

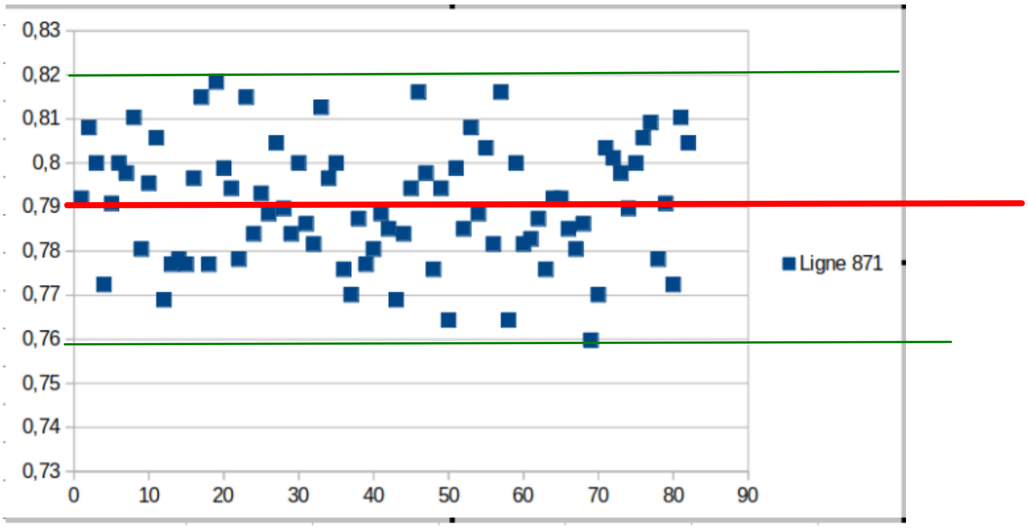
3. Applications

a. Exercice 1

«

En mars 1972, dans un comté du sud du Texas, Rodrigo Partida était inculpé puis condamné à huit ans de prison pour cambriolage d'une résidence privée dans la nuit avec tentative de viol.Après avoir épuisé toutes les voies de recours, il attaqua ce jugement le 9 novembre 1976 au motif que la désignation des jurés de ce comté était discriminante à l'égard des Américains d'origine mexicaine. Alors que 79,1% de la population du comté était d'origine mexicaine (recensement de 1970), sur les 870 personnes convoquées pour être jurés sur la période 1962-1972, seules 339 d'entre elles étaient d'origine mexicaine.Lors du procès, un statisticien produisit des arguments pour convaincre la Cour Suprême du bien fondé de la requête de l'accusé (dont les juges votèrent à 5 contre 4 en faveur de la requête).

»



$p = 0{,}791 \qquad n = 870$

$\frac{1}{\sqrt{870}} = 0{,}03$

$f \in [0{,}761 ; 0{,}821]$

$0{,}761 \leq f \leq 0{,}821$

b. Exercice 2

En vue d'une élection en Léonie, un institut de sondage veut estimer la proportion d'électeurs favorables au candidat sortant Léon. Pour ce faire, l'institut procède à un sondage aléatoire de taille 3000 et obtient 1560 intentions de vote pour Léon.

i. Calculer la fréquence f des intentions de vote en faveur du candidat, sur ce sondage.

$f = \frac{1560}{3000} = 0{,}52$

ii. Peut-on affirmer, en supposant que les électeurs maintiennent leur choix le jour du vote, que Léon sera élu ? Pourquoi ?

Non, ça dépend . . . .

iii. Expliquer pourquoi on peut dire que Léon est quand même sûr à 95% d'être élu.

$\frac{1}{\sqrt{3000}} = 0,01$

$0,52 - 0,01 \leq p \leq 0,52 + 0,01$

$0,51 \leq p \leq 0,53$

iv. Si la taille du sondage avait été de 1000, pour une fréquence obtenue de 0,52, la conclusion aurait-elle été la même ? NON !

$\frac{1}{\sqrt{1000}} = 0,03$

$0,49 \leq p \leq 0,55$

v. À partir de quelle taille de sondage (avec une fréquence de 0,52) Léon est-il sûr à 95% de gagner ?

challenge ...