

Chapitre 16 : Échantillonnage.

Eliot ROY

2022-2023

Points du programme.

Thème : Statistiques et probabilité.

Contenus : En liaison avec la partie « Algorithme et programmation », on définit la notion d'échantillon. L'objectif est de faire percevoir, sous une forme expérimentale, la loi des grands nombres, la fluctuation d'échantillonnage et le principe de l'estimation d'une probabilité par une fréquence observée sur un échantillon.

- Échantillon aléatoire de taille n pour une expérience à deux issues.
- Version vulgarisée de la loi des grands nombres : « Lorsque n est grand, sauf exception, la fréquence observée est proche de la probabilité ».
- Principe de l'estimation d'une probabilité ou d'une proportion dans une population, par une fréquence observée sur un échantillon.

Capacités attendues.

- Lire et comprendre une fonction Python qui renvoie le nombre ou la fréquence de succès dans un échantillon de taille n pour une expérience aléatoire à deux issues.
- Observer la loi des grands nombres à l'aide d'une simulation Python ou tableur.

Activité spaghetti.

I Vocabulaire.

Définition : 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 joue à pile ou face dix fois de suite.

- On obtient : pile-face-face-pile-pile-pile-face-pile-face-face. C'est un échantillon de taille 10. La fréquence de « face » est $\frac{5}{10} = \frac{1}{2}$.
- Pile-pile-face-pile-pile-pile-face-pile-face-pile est un autre échantillon de taille 10. La fréquence de « face » est $\frac{3}{10}$.

Dans ces deux échantillons, la fréquence n'est pas la même : cela illustre le phénomène de fluctuation d'échantillonnage.

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

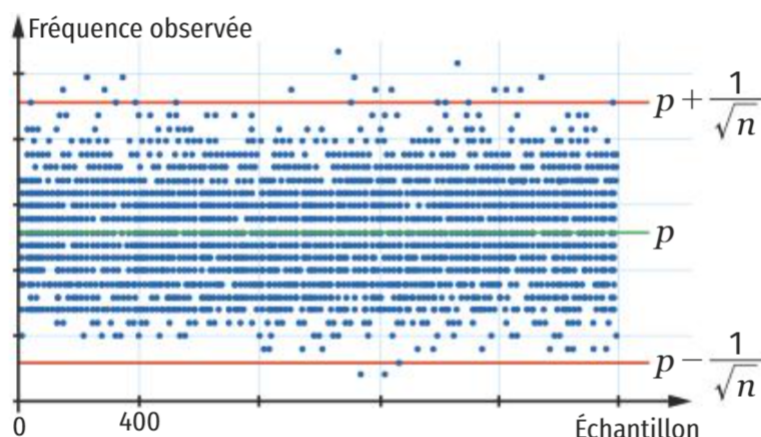
On s'intéresse à l'apparition d'un certain caractère dans une population. On note p la proportion théorique d'individus présentant ce caractère dans la population totale et on cherche à donner une estimation de p en minimisant le risque d'erreur.

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

Exemple : Une urne contient 5 000 boules : certaines sont rouges mais on ne sait pas dans quelle proportion.

Pour estimer la proportion de boules rouges dans cette urne, on prélève plusieurs échantillons

de taille 100 dans cette urne, et on observe la fréquence de boules rouges. On a représenté dans le graphique ci-dessous les fréquences observées pour 2 000 échantillons de taille 100. On constate que, dans une grande majorité des cas, cette fréquence est comprise entre $p - \frac{1}{\sqrt{n}}$ et $p + \frac{1}{\sqrt{n}}$.



Objectif : Observer la loi des grands nombre à l'aide d'une simulation (Python ou tableur).

Communs : Exercices 12, 13 page 325, 57, 58 page 329, 60 page 330.

Objectif : Lire et comprendre une fonction Python qui renvoie le nombre ou la fréquence de succès dans un échantillon de taille n pour une expérience aléatoire à deux issues.

Communs : Exercices 52, 53, 54, 55, 56 page 328.