Échantillonnage

I Échantillon

Définition I.1

On considère une expérience aléatoire à deux issues que l'on répète de manière indépendante c'est-à-dire que le résultat d'une expérience n'a pas d'influence sur la suivante.
Un échantillon de taille n de cette expérience est l'ensemble des résultats obtenus lorsqu'on répète n fois cette expérience aléatoire.

Exemple I.1 — On a lancé une pièce de monnaie 10 fois de suite et on a obtenu l'échantillon de taille 10 suivant :

Pile - Pile - Pile - Face - Pile - Face - Pile - Pile - Face - Pile

Déterminer la fréquence de « Pile » dans cet échantillon. → À rédiger

Exemple I.2 — On tire au sort 1000 personnes dans la population française et on observe si la personne est droitière ou non. Expliquer pourquoi on peut considérer ces tirages comme une expérience aléatoire répétée 1000 fois de manière indépendante. → À rédiger

II Fluctuation : version vulgarisée de la loi des grands nombres

Lorsqu'on réalise plusieurs échantillons de taille n, les fréquences observées ne sont pas toujours les mêmes. Ce phénomène s'appelle la **fluctuation d'échantillonnage**.

Proposition II.1

On considère une expérience aléatoire à deux issues que l'on répète de manière indépendante. On note p la probabilité d'une des issues et on considère un échantillon de taille n.
Lorsque n est grand, sauf exception, la fréquence observée f de cette issue est proche de la probabilité p.
Dans une grande majorité des cas, l'écart entre f et p est même inférieur à $\dfrac{1}{\sqrt{n}}$.

Exemple II.1 — On lance à dé à six faces bien équilibré et on s'intéresse à l'issue « Obtenir un 2 ». A l'aide d'un programme informatique, on a simulé le tirage de 10 échantillons de taille $n = 1000$. Pour chaque échantillon, voici la fréquence de nombre de fois où un 2 est sorti :

1. $f = 0,152$	2. $f = 0,144$	3. $f = 0,192$	4. $f = 0,171$	5. $f = 0,179$
6. $f = 0,164$	7. $f = 0,172$	8. $f = 0,165$	9. $f = 0,128$	10. $f = 0,165$

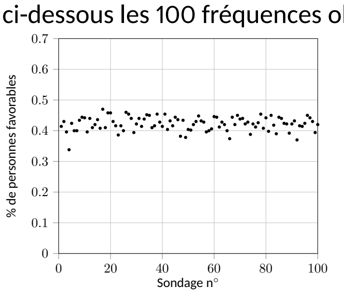
Déterminer la proportion des cas où l'écart entre f et la probabilité théorique p est inférieure à $\dfrac{1}{\sqrt{n}}$. → À rédiger

III Estimation d'une proportion dans une population

Si on ignore la proportion p d'un certain caractère dans une population, on peut essayer de l'estimer en tirant un échantillon de taille suffisamment grande.

Exemple III.1 — Dans une ville, un chercheur en médecine s'intéresse à la proportion de personnes qui sont atteintes de diabète de type 2. Il fait appel à un institut de sondage qui interroge 1153 personnes parmi lesquelles 70 se déclarent atteintes de cette maladie.
Quelle estimation, en pourcentage arrondi au centième, le chercheur obtient-il ainsi ? → À rédiger

Exemple III.2 — Avant un réferendum, un gouvernement souhaite estimer la proportion de votants qui répondront favorablement à la question posée. Il a réalisé 100 sondages d'opinions auprès de 500 personnes à chaque fois. Voici ci-dessous les 100 fréquences obtenues de personnes favorables :



Estimer la proportion de personnes favorables à ce réferendum. → À rédiger

Solutions

Exemple I.1
$f = \frac{7}{10} = 0,7$

Exemple I.2
Comme la population française est très grande par rapport au nombre de personnes tirées, ne pas « remettre » une personne tirée au sort dans la population ne va pas significativement changer la probabilité de choisir une personne droitière.

Exemple II.1
La probabilité théorique est $p = \frac{1}{6}$ et $\frac{1}{\sqrt{n}} \simeq 0,032$.
Il y a 9 fréquences telles que $|f - p| \leqslant 1/\sqrt{1000}$ ce qui fait une proportion de 90%.

Exemple III.1
$f = \frac{70}{1153} \simeq 0,0607$ donc on peut estimer à 6,07% la proportion de personnes atteintes de cette maladie.

Exemple III.2
On peut estimer la proportion de personnes favorables à $p = 42\%$.

Échantillonnage

A savoir faire à la fin du chapitre.

- Savoir déterminer la fréquence d'un caractère dans un échantillon aléatoire.
- Connaître la version vulgarisée de la loi des grands nombres.
- Savoir lire et comprendre une fonction Python renvoyant le nombre ou la fréquence de succès dans un échantillon de taille n pour une expérience aléatoire à deux issues.
- Savoir simuler N échantillons de taille n d'une expérience aléatoire à deux issues.
- Si p est la probabilité d'une issue et f sa fréquence observée dans un échantillon, savoir calculer la proportion des cas où l'écart entre p et f est inférieur ou égal à $1/\sqrt{n}$.
- Savoir estimer une proportion dans une population à partir d'un échantillon.

Échantillonnage

A savoir faire à la fin du chapitre.

- Savoir déterminer la fréquence d'un caractère dans un échantillon aléatoire.
- Connaître la version vulgarisée de la loi des grands nombres.
- Savoir lire et comprendre une fonction Python renvoyant le nombre ou la fréquence de succès dans un échantillon de taille n pour une expérience aléatoire à deux issues.
- Savoir simuler N échantillons de taille n d'une expérience aléatoire à deux issues.
- Si p est la probabilité d'une issue et f sa fréquence observée dans un échantillon, savoir calculer la proportion des cas où l'écart entre p et f est inférieur ou égal à $1/\sqrt{n}$.
- Savoir estimer une proportion dans une population à partir d'un échantillon.

Échantillonnage

A savoir faire à la fin du chapitre.

- Savoir déterminer la fréquence d'un caractère dans un échantillon aléatoire.
- Connaître la version vulgarisée de la loi des grands nombres.
- Savoir lire et comprendre une fonction Python renvoyant le nombre ou la fréquence de succès dans un échantillon de taille n pour une expérience aléatoire à deux issues.
- Savoir simuler N échantillons de taille n d'une expérience aléatoire à deux issues.
- Si p est la probabilité d'une issue et f sa fréquence observée dans un échantillon, savoir calculer la proportion des cas où l'écart entre p et f est inférieur ou égal à $1/\sqrt{n}$.
- Savoir estimer une proportion dans une population à partir d'un échantillon.