

Probabilités et échantillonnage

L'essentiel sur une page

Ce chapitre apprend à calculer des probabilités (équiprobabilité, réunion, contraire, probabilité conditionnelle), à modéliser par un tableau ou un arbre pondéré, et à interpréter une fréquence observée grâce à la fluctuation d'échantillonnage et à la loi des grands nombres.

1 Définitions

- **Expérience aléatoire** : issues connues, résultat imprévisible. **Univers** Ω : ensemble des issues ; **événement** : ensemble d'issues.
- **Probabilité** : nombre de $[0 ; 1]$ attaché à chaque issue, somme égale à 1.
- **Contraire** $\overline{A}$; **intersection** $A \cap B$ (« et ») ; **réunion** $A \cup B$ (« ou ») ; **incompatibles** : $A \cap B = \emptyset$.
- **Probabilité conditionnelle** : $P_A(B)$, probabilité de B sachant A (avec $P(A) \neq 0$).
- **Échantillon de taille n** : n répétitions de l'expérience ; **fréquence observée** $f = \frac{\text{effectif}}{n}$.

2 Formules

- Équiprobabilité : $P(A) = \frac{\text{nombre d'issues de } A}{\text{nombre total d'issues}}$.
- $P(\overline{A}) = 1 - P(A)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$; si incompatibles, $P(A \cup B) = P(A) + P(B)$.
- $P_A(B) = \frac{P(A \cap B)}{P(A)}$ et donc $P(A \cap B) = P(A) \times P_A(B)$.
- Arbre : chemin = produit des branches ; événement = somme des chemins ; chaque nœud totalise 1.

3 Méthodes

1. **Tableau à double entrée** : compléter les totaux ; probabilité simple = effectif $\div$ total ; $P_A(B) = \frac{\text{case } A \cap B}{\text{total de } A}$.
2. **Arbre pondéré** : premier niveau = probabilités connues ; second niveau = conditionnelles ; multiplier le long d'un chemin, additionner les chemins ; pour renverser la condition, diviser $P(A \cap B)$ par $P(B)$.
3. **Estimer une probabilité** : calculer f ; si n est grand, $p \approx f$ (loi des grands nombres) ; si n est petit, invoquer la fluctuation.
4. **Simuler en Python** : boucle for, test `randint(a, b)` ou `random()` $\rightarrow$ p, compteur, puis `compteur / n`.

4 Échantillonnage

- **Fluctuation d'échantillonnage** : la fréquence varie d'un échantillon à l'autre, d'autant moins que n est grand.
- **Loi des grands nombres** (version vulgarisée) : quand n devient très grand, f se rapproche de p .
- Une fréquence n'est pas une probabilité : f dépend de l'échantillon, p du modèle.

5 Pièges

- $P_A(B) \neq P_B(A)$: même numérateur, dénominateurs différents (faux positifs d'un test).
- Ne pas oublier $-P(A \cap B)$ dans la formule de la réunion.
- Sans remise : au second tirage, le total et l'effectif de la sorte tirée diminuent de 1.
- Le hasard n'a pas de mémoire : un résultat « en retard » n'est pas plus probable.
- Un petit échantillon ne prouve rien ; un grand échantillon est un indice sérieux.

L'ESSENTIEL

— $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ et $P(\overline{A}) = 1 - P(A)$.

— $P_A(B) = \frac{P(A \cap B)}{P(A)}$: on divise par la probabilité de la condition.

— Sur un arbre : on multiplie le long d'un chemin, on additionne les chemins.

— En général $P_A(B) \neq P_B(A)$: un test fiable peut donner surtout des faux positifs si le caractère est rare.

— Loi des grands nombres : sur un grand échantillon, la fréquence observée se rapproche de la probabilité.