

Plaque 5 – Simulation aléatoire et problèmes de synthèse

Python et calcul scientifique — Licence 1 (L1)

Exercices

Méthode

Simuler, c'est faire exécuter une expérience aléatoire un grand nombre de fois par l'ordinateur et mesurer des fréquences. Une simulation **ne démontre rien**, mais elle **contrôle** un calcul théorique (ou révèle une erreur) et donne des réponses là où le calcul exact est inaccessible. Deux réflexes : fixer la graine (*random.seed*) pour des résultats reproductibles, et toujours estimer la **précision** ($\approx \frac{1}{\sqrt{N}}$) avant d'interpréter un écart.

Exercice 1 – Tester un dé simulé

On exécute :

```
import random

random.seed(1)
n = 6000
compte = [0] * 7
for k in range(n):
    compte[random.randint(1, 6)] += 1
print(compte[1:])
```

- Le programme affiche [979, 1002, 1007, 995, 1045, 972]. Quelles valeurs attendait-on ?
- Calculer l'écart type du nombre de « 5 » et dire si l'écart 1045 – 1000 est anormal.
- Pourquoi la liste a-t-elle 7 cases ?
- Qu'aurait affiché le programme avec *randint(0, 6)* ? Comment le test aurait-il détecté l'erreur ?

Exercice 2 – Marche aléatoire : la loi en racine carrée

On simule $N = 10\,000$ marches aléatoires de $T = 100$ pas ± 1 et on calcule la moyenne de S_T^2 (*random.seed(2)*).

- Écrire le programme.
- Il affiche 99,48. Quelle valeur théorique attend-on ?
- En déduire l'ordre de grandeur de la distance à l'origine après T pas.
- Un marcheur fait 10 000 pas au hasard : à quelle distance typique se trouve-t-il ? Comparer à une marche en ligne droite.

Exercice 3 – Anniversaires : simuler pour contrôler

- Écrire un programme qui tire 23 anniversaires au hasard et teste s'il y a une coïncidence (utiliser *len(set(jours))*), répété 10 000 fois.
- Avec *random.seed(3)*, il affiche 0,5071. La valeur exacte est 0,5073. L'accord est-il satisfaisant ?

- c) Quelle taille d'échantillon N faudrait-il pour une précision de $\pm 0,001$ (deux écarts types) ?
- d) Pourquoi *set* rend-il le test simple ?

Exercice 4 – Estimer e avec des permutations

On mélange 100 000 fois la liste $[0, \dots, 9]$ (*random.shuffle*, graine 4) et on compte les permutations **sans point fixe**.

- a) Écrire le programme.
- b) La fréquence obtenue est 0,36625. Quelle valeur théorique attend-on (plaque de dénombrement) ?
- c) Le programme affiche aussi $\frac{N}{\text{nombre de dérangements}} \approx 2,730$. Pourquoi est-ce une estimation de e ?
- d) Évaluer la précision de cette estimation de e .

Exercice 5 – Volume des boules en grande dimension

On estime le volume de la boule unité de $\mathbb{R}^d$ en tirant $N = 100\,000$ points uniformes dans le cube $[-1, 1]^d$ (graine 5).

- a) Écrire la fonction *volume*(d, N).
- b) Résultats : 4,1872 ($d = 3$, exact $\frac{4}{3}\pi \approx 4,1888$), 5,2464 ($d = 5$, exact $\frac{8\pi^2}{15} \approx 5,2638$), 2,601 ($d = 10$, exact $\frac{\pi^5}{120} \approx 2,550$). Calculer les erreurs relatives.
- c) Pourquoi l'erreur relative augmente-t-elle avec d ? (Évaluer la proportion de points dans la boule pour $d = 10$.)
- d) Commenter : les points d'un cube de grande dimension sont-ils « près du centre » ?

Exercice 6 – Simuler une loi par inversion

On veut simuler X de densité $f(x) = 2x$ sur $[0, 1]$.

- a) Calculer la fonction de répartition F et sa réciproque ; en déduire que $\sqrt{U}$ convient, U uniforme sur $[0, 1]$.
- b) Le programme (graine 6, $N = 10^5$) donne une moyenne 0,6675 et $\mathbb{P}(X \leq 0,5) \approx 0,2489$. Calculer les valeurs exactes.
- c) Montrer qu'on obtient la même loi avec $\max(U_1, U_2)$, U_1, U_2 indépendantes uniformes.
- d) Écrire le programme de simulation.

Exercice 7 – Collectionneur : théorie et simulation

- a) Écrire *collection*(n) qui tire des numéros uniformes dans $\{1, \dots, n\}$ jusqu'à les avoir tous vus, et renvoie le nombre de tirages.
- b) Avec la graine 7, la moyenne sur 10 000 essais pour $n = 6$ vaut 14,74. Comparer à la valeur théorique.

- c) La variance théorique est $\sum_{k=1}^n \frac{1-p_k}{p_k^2}$ avec $p_k = \frac{n-k+1}{n}$; pour $n = 6$ elle vaut $\approx 38,99$. L'écart observé est-il normal ?
- d) Pourquoi utiliser un *set* plutôt qu'une liste ?

Exercice 8 – Taux de couverture d'un intervalle

Pour une proportion inconnue p , on utilise l'intervalle $\left[F_n - \frac{1}{\sqrt{n}}, F_n + \frac{1}{\sqrt{n}}\right]$ (F_n fréquence sur n tirages).

- a) Écrire un programme qui, pour $p = 0,3$ et $n = 400$, répète 1000 fois l'expérience et compte la proportion d'intervalles contenant p .
- b) Avec la graine 9, le programme affiche 0,981. Expliquer pourquoi on s'attendait à un taux d'au moins 95 %.
- c) Calculer le demi-intervalle en nombre d'écarts types pour $p = 0,3$ et en déduire un taux théorique proche de 97 %.
- d) Que signifie « intervalle de confiance à 95 % » ? Est-ce la probabilité que p soit dans **cet** intervalle ?

Exercice 9 – Monty Hall par simulation

- a) Écrire une simulation du jeu de Monty Hall (voiture et choix au hasard, l'animateur ouvre une porte perdante différente du choix) comptant les victoires de la stratégie « changer ».
- b) Avec la graine 10 et 30 000 parties, le programme affiche 0,6636. Comparer à la théorie.
- c) Modifier une ligne pour que l'animateur ouvre une porte **au hasard** parmi les deux non choisies, quitte à révéler la voiture (partie alors annulée). Que devient la probabilité de gagner en changeant, parmi les parties non annulées ?
- d) Quel est l'intérêt d'une simulation pour un problème aussi contre-intuitif ?

Exercice 10 – Attendre PP ou PF : vérification numérique

- a) Écrire *attente(motif)* qui lance une pièce jusqu'à ce que les deux derniers résultats forment *motif* et renvoie le nombre de lancers.
- b) Sur 100 000 essais (graine 11), les moyennes sont 6,006 pour *PP* et 3,988 pour *PF*. Comparer aux valeurs théoriques 6 et 4.
- c) Sachant que $V = 22$ pour *PP* et 4 pour *PF*, calculer l'écart type de chaque moyenne et conclure.
- d) Pourquoi le programme ne garde-t-il que le dernier caractère avant d'en ajouter un ?

Exercice 11 – Ruine du joueur : simulation

Un joueur part de 1 euro, gagne ou perd 1 euro à chaque partie (gain avec probabilité $p = \frac{2}{3}$), et s'arrête à 0 ou 3 euros.

- a) Écrire *partie(k, p)* renvoyant *True* si le joueur atteint 3.

- b) Sur 100 000 parties (graine 12), la fréquence de succès vaut 0,56928. Comparer à la valeur exacte $\frac{4}{7}$.
- c) Justifier la terminaison de la boucle avec probabilité 1 (sans variant déterministe).
- d) Modifier le programme pour estimer la durée moyenne du jeu. (Valeur théorique depuis 1 : $\frac{15}{7}$.)

Exercice 12 – Fabriquer du hasard : générateurs congruentiels

Un générateur congruentiel calcule $x_{n+1} = (ax_n + c) \bmod m$.

- a) Écrire `lcg(a, c, m, x0, n)` renvoyant les n premiers termes.
- b) Pour $(a, c, m) = (5, 3, 16)$ et $x_0 = 0$, le programme affiche 3, 2, 13, 4, 7, 6, 1, 8, 11, 10, 5, 12, 15, 14, 9, 0, 3, ... Quelle est la période ? Pourquoi ne peut-elle pas dépasser m ?
- c) Pour $(4, 1, 16)$: 1, 5, 5, 5, ... ; pour $(5, 0, 16)$ depuis 1 : 5, 9, 13, 1, ... Donner les périodes et commenter.
- d) Observer la parité des termes pour $(5, 3, 16)$; ce générateur est-il satisfaisant ?

Exercice 13 – Mélanger correctement une liste

On compare deux façons de mélanger $[1, 2, 3]$:

```
def fisher_yates(L):
    L = L[:]
    for i in range(len(L) - 1, 0, -1):
        j = random.randint(0, i)
        L[i], L[j] = L[j], L[i]
    return L

def naif(L):
    L = L[:]
    for i in range(len(L)):
        j = random.randint(0, len(L) - 1)
        L[i], L[j] = L[j], L[i]
    return L
```

- a) Sur 60 000 mélanges (graine 14), Fisher-Yates donne des effectifs entre 9828 et 10 150 pour les 6 permutations ; le mélange naïf donne environ 8900 pour trois d'entre elles et 11 100 pour les trois autres. Quels effectifs attendait-on ?
- b) Montrer que Fisher-Yates produit chaque permutation avec probabilité $\frac{1}{3!}$ (compter les suites de choix).
- c) Le mélange naïf fait $3^3 = 27$ suites de choix équiprobables. Pourquoi ne peut-il pas être uniforme sur 6 permutations ?
- d) L'énumération exacte des 27 cas donne les effectifs 4, 5, 5, 5, 4, 4. Retrouver les fréquences simulées.

Exercice 14 – Estimer π : comparaison générale des méthodes

On récapitule plusieurs façons de calculer π rencontrées dans ce chapitre.

- a) Monte-Carlo « succès/échec » : erreur en $\frac{1}{\sqrt{N}}$; combien de tirages pour 6 décimales ?
- b) Série de Leibniz : erreur $\approx \frac{1}{n}$; combien de termes pour 6 décimales ? Et avec la moyenne de deux sommes consécutives (erreur $\approx \frac{1}{2n^2}$) ?
- c) Formule de Machin $\frac{\pi}{4} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}$, avec $\arctan x = \sum_k \frac{(-1)^k x^{2k+1}}{2k+1}$: combien de termes du premier arctangente pour 6 décimales ?
- d) Conclure sur l'importance de la **vitesse de convergence**.

Exercice 15 - Synthèse : le problème de la secrétaire simulé

On reprend le problème de la secrétaire (plaquette de probabilités) avec $n = 100$ candidats : on refuse les k premiers, puis on retient le premier meilleur que tous les précédents (le dernier par défaut).

- a) Écrire *strategie*(n, k) qui simule une instance et renvoie *True* si le meilleur candidat est retenu.
- b) Sur 20 000 instances (graine 15), les fréquences de succès sont 0,227 ($k = 10$), 0,367 ($k = 37$), 0,309 ($k = 60$). Quelle valeur de k la théorie recommande-t-elle, et avec quelle probabilité de succès ?
- c) La probabilité exacte de succès pour la stratégie k vaut $\frac{k}{n} \sum_{j=k+1}^n \frac{1}{j-1}$. En l'approchant par une intégrale, montrer qu'elle est $\approx -x \ln x$ avec $x = \frac{k}{n}$, maximale en $x = \frac{1}{e}$.
- d) Calculer $-x \ln x$ pour $x = 0,1, 0,37, 0,6$ et comparer aux simulations.