

Plaquette 4 - Simulations en Python et lecture de résultats

Échantillonnage — Première spé maths

Corrections détaillées

Boîte à outils

Générer un succès de probabilité p .

```
from random import random
```

```
def succes(p):
    return random() < p
```

random() renvoie un flottant de $[0; 1[$ uniformément, donc $P(\text{random}() < p) = p$.

Simuler un échantillon et sa fréquence.

```
def frequence(n, p):
    s = 0
    for i in range(n):
        if random() < p:
            s = s + 1
    return s / n
```

Simuler plusieurs échantillons et compter les sorties d'intervalle.

```
def hors_intervalle(N, n, p, a, b):
    c = 0
    for k in range(N):
        f = frequence(n, p)
        if f < a or f > b:
            c = c + 1
    return c
```

Entiers aléatoires. *randint(a, b)* renvoie un entier de $\llbracket a; b \rrbracket$, **bornes incluses** — attention à ne pas confondre avec *range(a, b)*, qui exclut b .

Ce qu'une simulation montre et ne montre pas.

- Elle **illustre** la fluctuation d'échantillonnage et vérifie un ordre de grandeur.
- Elle ne **démontre** rien : une simulation est elle-même soumise au hasard.
- Plus N (le nombre d'échantillons) est grand, plus la fréquence observée des sorties d'intervalle approche les 5 % théoriques.

Contrôle de cohérence. La moyenne des fréquences simulées doit approcher p , avec un écart type de l'ordre de $\frac{\sigma(F)}{\sqrt{N}}$.

Correction de l'exercice 1 - Lecture d'un script de simulation

- **Simulation** : 10 000 épreuves de Bernoulli de paramètre 0,35.
- **Grandeur affichée** : la **fréquence** de succès observée.
- **Valeur attendue** : environ 0,35 ✓
- **Écart type de la fréquence** :

$$\sigma(F) = \sqrt{\frac{0,35 \times 0,65}{10\,000}} = \sqrt{2,275 \times 10^{-5}} \approx 0,0047697$$

- **Intervalle typique** ($\pm 2\sigma$) : $[0,3405 ; 0,3595]$ ✓
- **Conclusion** : le script affichera presque toujours une valeur entre 0,340 et 0,360.
- **Nombre de succès** : environ 3500, avec $\sigma(X) \approx 47,70$ ✓
- **Contrôle** : $\frac{47,70}{10\,000} = 0,00477$ ✓
- **Précision** : la fréquence donne les **deux premières décimales** de p de façon fiable, la troisième restant incertaine.
- **Pour gagner une décimale** : il faudrait $N = 10^6$ épreuves, soit **cent fois** plus ✓
- **Erreur d'interprétation** : deux exécutions du script donneront des valeurs **différentes** — c'est normal, ce n'est pas un bug.

Réponse. La fréquence de succès pour $p = 0,35$ sur 10 000 épreuves : environ 0,35 à $\pm 0,01$ près.

Correction de l'exercice 2 - Corriger un script

- **Erreur** : `randint(0, 6)` renvoie un entier de 0 à 6, soit **sept** valeurs possibles ✓
- **Conséquence** : la probabilité d'obtenir 6 vaut $\frac{1}{7} \approx 0,1429$ au lieu de $\frac{1}{6} \approx 0,1667$.
- **Correction** : remplacer par `randint(1, 6)`, qui donne bien six valeurs.
- **Écart attendu** : le script erroné affiche environ 0,143 au lieu de 0,167, soit une erreur relative de 14 % ✓
- **Détection statistique** : avec $n = 600$ et $p = \frac{1}{6}$, on a $E(X) = 100$ et $\sigma(X) = \sqrt{600 \times \frac{1}{6} \times \frac{5}{6}} \approx 9,1287$.
- **Le script erroné** donne en moyenne $\frac{600}{7} \approx 85,71$ succès, soit $\frac{85,71 - 100}{9,1287} \approx -1,565 \sigma$ — détectable, mais pas de façon certaine sur une seule exécution ✓
- **Avec** $n = 6000$: l'écart serait $\frac{857,1 - 1000}{\sqrt{6000 \times 5/36}} \approx \frac{-142,9}{28,87} \approx -4,949 \sigma$ — cette fois le bug est **flagrant** ✓
- **Bonne pratique** : toujours tester un simulateur sur un grand nombre de tirages et comparer à la valeur théorique.
- **Piège jumeau** : `range(1, 6)` exclut 6 et ne donne que cinq valeurs — la confusion entre `randint` (inclusif) et `range` (exclusif) est la source d'erreur la plus courante en Python.

Réponse. `randint(0, 6)` donne sept valeurs : il faut écrire `randint(1, 6)`.

Correction de l'exercice 3 - Simuler une fréquence d'échantillon

Fonction.

```
from random import random

def frequence(n, p):
    s = 0
    for i in range(n):
        if random() < p:
            s = s + 1
    return s / n
```

- **Espérance** : $E(F) = 0,4$.
- **Écart type** :

$$\sigma(F) = \sqrt{\frac{0,4 \times 0,6}{50}} = \sqrt{0,0048} \approx 0,069282$$

- **Intervalle typique** ($\pm 2\sigma$) : $[0,2614; 0,5386]$ ✓
- **En effectifs** : entre 13 et 27 succès environ.
- **Valeurs possibles** : la fréquence est un multiple de $\frac{1}{50} = 0,02$ — elle ne peut **pas** valoir 0,41 par exemple ✓
- **Granularité** : c'est une limite intrinsèque des petits échantillons ; avec $n = 50$, la résolution est de 2 points.
- **Contrôles** : $E(X) = 20$ et $\sigma(X) = \sqrt{50 \times 0,24} = \sqrt{12} \approx 3,4641$ ✓
- **Contrôle de cohérence** : $\frac{3,4641}{50} = 0,069282$ ✓
- **Sur 1000 appels de la fonction** : la moyenne des fréquences devrait approcher 0,4 à $\frac{0,069282}{\sqrt{1000}} \approx 0,00219$ près ✓
- **Intervalle pour cette moyenne** : $[0,3956; 0,4044]$ — bien plus resserré que celui d'une fréquence isolée ✓

Réponse. $E(F) = 0,4$ et $\sigma(F) \approx 0,0693$: les fréquences se répartissent dans $[0,26; 0,54]$ environ.

Correction de l'exercice 4 - Compter les sorties d'intervalle

- **Écart type de la fréquence** :

$$\sigma(F) = \frac{0,5}{\sqrt{200}} \approx 0,035355$$

- **Demi-amplitude donnée** : 0,07.
- **En écarts types** : $\frac{0,07}{0,035355} \approx 1,980$ — presque exactement 2σ ✓
- **Proportion attendue hors intervalle** : environ 5 %.
- **Comptage attendu** : environ 25 échantillons sur 500 ✓
- **Calcul exact** : l'intervalle correspond à $86 \leq X \leq 114$ pour $X \sim B(200; 0,5)$.
- **Valeurs** : $P(X \leq 85) \approx 0,020018$ et, par symétrie, $P(X \geq 115) \approx 0,020018$.
- **Somme** : $\approx 0,040036$, soit un comptage attendu de $500 \times 0,040036 \approx 20$ ✓
- **Fluctuation de ce comptage** : il suit $B(500; 0,0400)$, d'espérance 20,0 et d'écart type $\sqrt{500 \times 0,04 \times 0,96} \approx 4,3818$ ✓
- **Intervalle attendu** : $[20 - 8,76; 20 + 8,76] \approx [11; 29]$ échantillons.
- **Conclusion** : tout comptage entre 11 et 29 serait normal ✓
- **Écart entre les deux prédictions** (20 exact contre 25 approché) : il vient de ce que l'approximation $\pm 2\sigma$ donne un intervalle **légèrement plus large** que le vrai intervalle à 95 % ✓

Réponse. Environ 20 échantillons (calcul exact) ou 25 (approximation à 5 %), avec une plage normale de $[11; 29]$.

Correction de l'exercice 5 - Simuler une loi binomiale

Fonctions.

```
from random import random
```

```
def binomiale(n, p):
    s = 0
```

```

for i in range(n):
    if random() < p:
        s = s + 1
return s

def estime(N, n, p, k):
    c = 0
    for i in range(N):
        if binomiale(n, p) == k:
            c = c + 1
    return c / N

```

- **Valeur théorique** : $P(X = 5) = \frac{252}{1024} \approx 0,24609$.
- **Précision demandée** : $\pm 0,01$, soit $2\sqrt{\frac{0,24609 \times 0,75391}{N}} \leq 0,01$.
- **Calcul** : $\sqrt{0,24609 \times 0,75391} \approx 0,43077$, donc la condition devient $\frac{0,86154}{\sqrt{N}} \leq 0,01$.
- **Résolution** : $\sqrt{N} \geq 86,154$, donc

$$N \geq 7422$$
- **Réponse** : environ 7500 simulations ✓
- **Contrôle** : pour $N = 7500$, la demi-amplitude vaut $\frac{0,86154}{86,603} \approx 0,0099$ ✓
- **Coût en calculs** : $7500 \times 10 = 75\,000$ tirages aléatoires — instantané pour un ordinateur.
- **Précision relative** : $\frac{0,01}{0,24609} \approx 4,1\%$ ✓
- **Pour $\pm 0,001$** : il faudrait $N \geq 742\,119$, soit **cent fois** plus ✓
- **Comparaison avec le calcul exact** : $\frac{252}{1024}$ est **immédiat** et **exact** — la simulation n'a d'intérêt que quand le calcul théorique est hors de portée.

Réponse. $P(X = 5) \approx 0,246$; il faut $N \geq 7422$ simulations pour une précision de $\pm 0,01$.

Correction de l'exercice 6 - Simulation d'un test de décision

Protocole.

1. Simuler N échantillons de taille 100 avec $p = 0,5$. 2. Pour chacun, calculer la fréquence. 3. Compter combien sortent de $[0,4 ; 0,6]$. 4. Diviser par N : on doit trouver environ 0,05.

- **Calcul exact** : l'intervalle correspond à $40 \leq X \leq 60$ pour $X \sim B(100 ; 0,5)$.
- **Valeurs** : $P(X \leq 39) \approx 0,017600$ et $P(X \geq 61) \approx 0,017600$.
- **Somme** : $\approx 0,035200$ ✓
- **Résultat attendu** : environ 3,5 % et non 5 % !
- **Explication** : l'intervalle $[40 ; 60]$ est **plus large** que le vrai intervalle à 95 %, à cause de l'arrondi aux entiers ✓
- **Vrai intervalle à 95 %** : $[41 ; 59]$, avec $P(X \leq 40) + P(X \geq 60) \approx 2 \times 0,028444 \approx 0,056888$ — un peu **plus** de 5 % ✓
- **Conclusion** : à cause de la granularité, aucun intervalle en entiers ne vaut exactement 95 % ; on encadre entre $[41 ; 59]$ (94,3 %) et $[40 ; 60]$ (96,5 %) ✓
- **Précision de la simulation** : avec $N = 10\,000$, l'estimation de 0,0352 aurait un écart type de $\sqrt{\frac{0,0352 \times 0,9648}{10\,000}} \approx 0,0018$ ✓
- **Intervalle attendu** : $[0,0316 ; 0,0388]$ — largement distinguable de 0,05 ✓

— **Leçon** : la simulation ne « valide » pas le 5 % ici, elle **révèle** que la convention $\pm 2\sigma$ est approximative.

Réponse. Environ 3,5 % et non 5 % : l'intervalle $[0,4 ; 0,6]$ est légèrement trop large (granularité).

Correction de l'exercice 7 - Estimer π par simulation

— **Principe** : on tire un point uniformément dans le carré $[0 ; 1]^2$; il tombe dans le quart de disque avec la probabilité $\frac{\pi}{4}$.

— **Estimation** : la fréquence approche $\frac{\pi}{4}$, donc $4\frac{C}{N}$ approche π ✓

— **Valeur de référence** : $\frac{\pi}{4} \approx 0,78540$.

— **Écart type de la fréquence** :

$$\sigma(F) = \sqrt{\frac{0,78540 \times 0,21460}{10^5}} \approx \sqrt{1,6855 \times 10^{-6}} \approx 0,0012983$$

— **Écart type de l'estimation de π** : $4 \times 0,0012983 \approx 0,0051932$.

— **Intervalle typique ($\pm 2\sigma$)** : $[3,1312 ; 3,1520]$ ✓

— **Conclusion** : on obtient environ **deux décimales** de π — le script affichera typiquement 3,13 à 3,15.

— **Contrôle** : $\pi \approx 3,14159$, bien au centre de l'intervalle ✓

— **Pour trois décimales ($\pm 0,001$ à 2σ)** : il faut $4 \times 2\sqrt{\frac{0,78540 \times 0,21460}{N}} \leq 0,001$.

— **Calcul** : le numérateur vaut $8\sqrt{0,16856} \approx 3,2845$, donc $\sqrt{N} \geq 3284,5$ et $N \geq 1,08 \times 10^7$ tirages ✓

— **Comparaison** : la série $\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \dots\right)$ donne aussi une convergence lente, mais la méthode d'Archimède (polygones) atteint 10 décimales en quelques dizaines d'étapes.

— **Intérêt réel de Monte-Carlo** : non pas calculer π , mais estimer des intégrales en **grande dimension**, où les méthodes déterministes échouent.

Réponse. Estimation de π par Monte-Carlo : environ $\pm 0,005$ pour $N = 10^5$, soit deux décimales.

Correction de l'exercice 8 - Simuler un tirage sans remise

Script.

```
from random import shuffle

def tirage():
    urne = ["R"]*4 + ["N"]*6
    shuffle(urne)
    return urne[:3].count("R")
```

— **Principe** : on **mélange** l'urne, puis on prend les trois premières boules — équivalent à un tirage sans remise ✓

— **Loi théorique** (hypergéométrique) : $P(X = k) = \frac{\binom{4}{k}\binom{6}{3-k}}{\binom{10}{3}}$, avec $\binom{10}{3} = 120$.

— **Valeurs** :

- $P(X = 0) = \frac{1 \times 20}{120} = \frac{20}{120} \approx 0,16667$;
- $P(X = 1) = \frac{4 \times 15}{120} = \frac{60}{120} = 0,5$;
- $P(X = 2) = \frac{6 \times 6}{120} = \frac{36}{120} = 0,3$;
- $P(X = 3) = \frac{4 \times 1}{120} = \frac{4}{120} \approx 0,033333$.
- **Contrôle** : $\frac{20 + 60 + 36 + 4}{120} = 1$ ✓
- **Espérance** : $E(X) = 0,5 + 0,6 + 0,1 = 1,2$ ✓
- **Contrôle** : $3 \times \frac{4}{10} = 1,2$ — l'espérance est la **même** qu'avec remise ✓
- **Variance** : $E(X^2) = 0,5 + 1,2 + 0,3 = 2$, donc $V(X) = 2 - 1,44 = 0,56$ et $\sigma(X) \approx 0,74833$.
- **Comparaison avec remise** : $V = 3 \times 0,4 \times 0,6 = 0,72$ — la variance est **plus faible** sans remise ✓
- **Erreur à éviter** : utiliser *randint* trois fois simulerait un tirage **avec** remise, ce qui donnerait une variance de 0,72 au lieu de 0,56 — détectable sur un grand nombre de simulations ✓
- **Alternative** : *sample(urne, 3)* du module *random* fait directement un tirage sans remise.

Réponse. Mélanger puis prendre trois boules ; loi $\left(\frac{1}{6}; \frac{1}{2}; 0,3; \frac{1}{30}\right)$, $E(X) = 1,2$ et $V(X) = 0,56$.

Correction de l'exercice 9 - Interpréter un histogramme simulé

- **Espérance de chaque fréquence** : 0,3.
- **Écart type de chaque fréquence** :

$$\sigma(F) = \sqrt{\frac{0,3 \times 0,7}{100}} \approx 0,045826$$
- **Écart type de la moyenne** sur 2000 échantillons :

$$\frac{0,045826}{\sqrt{2000}} \approx 0,0010247$$
- **Écart observé** : $0,3012 - 0,3 = 0,0012$.
- **En écarts types** : $\frac{0,0012}{0,0010247} \approx 1,171 \sigma$ ✓
- **Conclusion** : parfaitement **cohérent** avec la théorie ✓
- **Intervalle attendu** ($\pm 2\sigma$) : $[0,29795; 0,30205]$, qui contient 0,3012 ✓
- **Remarque d'échelle** : la moyenne de 2000 échantillons de taille 100 équivaut à une fréquence sur 200 000 épreuves, d'où la précision au millièmme ✓
- **Contrôle** : $\sqrt{\frac{0,21}{200\,000}} \approx 0,0010247$ ✓ — les deux calculs coïncident exactement.
- **Ce que la simulation confirme** : l'espérance 0,3, à environ 0,001 près.
- **Ce qu'elle ne confirme pas** : la valeur de $\sigma(F)$ pour un échantillon isolé — pour cela, il faudrait examiner la **dispersion** des 2000 fréquences, et non leur moyenne ✓
- **Contrôle à faire en plus** : l'écart type observé des 2000 fréquences devrait valoir environ 0,0458.

Réponse. Oui : l'écart de 0,0012 ne vaut que $1,17 \sigma$ (avec $\sigma \approx 0,00102$).

Correction de l'exercice 10 - Simuler une marche aléatoire

Script.

```
from random import randint
```

```
def marche(n):
    x = 0
    for i in range(n):
        if randint(0, 1) == 1:
            x = x + 1
        else:
            x = x - 1
    return x
```

- **Loi** : si k est le nombre de pas $+1$, alors $k \sim B(100; 0,5)$ et la position vaut $X = 2k - 100$.
- **Espérance** : $E(X) = 2 \times 50 - 100 = 0$ ✓
- **Variance** : $V(X) = 4 \times 25 = 100$, donc $\sigma(X) = 10$ ✓
- **Intervalle typique** ($\pm 2\sigma$) : $[-20; 20]$ ✓
- **Parité** : la position est **toujours paire** après 100 pas — elle ne peut jamais valoir 1 ou 7 ✓
- **Probabilité de revenir à 0** : $P(k = 50) = \binom{100}{50} 2^{-100} \approx 0,079589$, soit environ 8 % ✓
- **Probabilité de dépasser ± 20** : environ 5 %, soit 2,5 % de chaque côté ✓
- **Contrôle exact** : $P(k \leq 39) + P(k \geq 61) \approx 0,035200$ — un peu moins ✓
- **Piège du script** : `randint(0, 1)` donne bien deux valeurs équiprobables ✓ ; en revanche `randint(0, 2)` en donnerait trois et biaiserait la marche.
- **Contrôle d'un simulateur** : sur 10 000 marches, la moyenne des positions doit approcher 0 à $\frac{10}{100} = 0,1$ près, et l'écart type observé doit valoir environ 10 ✓
- **Distance moyenne à l'origine** : environ $\sigma\sqrt{\frac{2}{n}} \approx 7,98$ — une marche typique finit à environ 8 pas du départ, et non à 0.

Réponse. $E(X) = 0$, $\sigma(X) = 10$, positions paires dans $[-20; 20]$ le plus souvent.

Correction de l'exercice 11 - Fréquence et taille croissante

- **Écart type de la fréquence** : $\frac{0,5}{\sqrt{n}}$.

— **Valeurs** :

n	$\sigma(F)$	intervalle typique
10	0,1581	[0,184 ; 0,816]
100	0,05	[0,400 ; 0,600]
1000	0,01581	[0,468 ; 0,532]
10 000	0,005	[0,490 ; 0,510]

- **Observation attendue** : les fréquences se **rapprochent** de 0,5 à mesure que n croît ✓
- **Attention** : la convergence n'est **pas monotone** — il est parfaitement possible d'observer 0,5 pile pour $n = 10$ et 0,507 pour $n = 10\,000$ ✓

- **Ce que garantit la théorie** : la **probabilité** de s'écarter de 0,5 de plus d'un δ fixé tend vers 0, pas la suite des écarts elle-même ✓
- **Granularité pour $n = 10$** : la fréquence est un multiple de 0,1 — elle ne peut valoir que 0 ; 0,1 ; ... ; 1 ✓
- **Écart absolu en effectifs** : $\sigma(X) = 0,5\sqrt{n}$ vaut 1,58 ; 5 ; 15,8 ; 50 — il **croît** ✓
- **Résolution du paradoxe** : l'écart sur le **nombre** de piles grandit, mais divisé par n il se réduit ✓
- **Contrôle** : pour $n = 10\,000$, l'intervalle en effectifs est [4900 ; 5100], soit ± 100 piles — bien plus qu'à $n = 10$ ($\pm 1,58$) ✓
- **Erreur d'interprétation à combattre** : il n'y a **aucune** « force de rappel » qui ramènerait les piles vers la moitié. C'est la division par n qui fait tout le travail.

Réponse. Fréquences convergeant vers 0,5 avec $\sigma(F) = \frac{0,5}{\sqrt{n}}$: 0,158 ; 0,05 ; 0,0158 ; 0,005.

Correction de l'exercice 12 - Simuler un intervalle de confiance

- **Demi-amplitude** : $\frac{1}{20} = 0,05$.
- **Condition** : l'intervalle contient 0,6 si et seulement si $|f - 0,6| \leq 0,05$ ✓
- **Écart type de la fréquence** :

$$\sigma(F) = \sqrt{\frac{0,6 \times 0,4}{400}} = \sqrt{0,0006} \approx 0,024495$$
- **En écarts types** : $\frac{0,05}{0,024495} \approx 2,041$ ✓
- **Proportion attendue** : un peu plus de 95 %.
- **Calcul exact** : $|f - 0,6| \leq 0,05$ équivaut à $220 \leq X \leq 260$ pour $X \sim B(400 ; 0,6)$.
- **Valeurs** : $P(X \leq 219) \approx 0,018652$ et $P(X \geq 261) \approx 0,017621$.
- **Somme** : $\approx 0,036273$, donc environ 96,4 % des intervalles contiennent 0,6 ✓
- **Comptage attendu** : environ 964 intervalles sur 1000 ✓
- **Fluctuation** : $\sqrt{1000 \times 0,0363 \times 0,9637} \approx 5,91$, donc un intervalle attendu de [952 ; 976] ✓
- **Interprétation du « niveau de confiance »** : ce n'est pas une propriété d'un intervalle, mais de la **méthode** — 96 % des intervalles construits ainsi attrapent la vraie valeur ✓
- **Pourquoi 96 % et non 95 %** : la version $\frac{1}{\sqrt{n}}$ utilise la majoration $p(1 - p) \leq \frac{1}{4}$, alors qu'ici $p(1 - p) = 0,24 < 0,25$ — l'intervalle est donc un peu **trop large** ✓
- **Avec la version affinée** : la proportion serait plus proche de 95 %.

Réponse. Environ 964 sur 1000 (soit 96,4 %, un peu plus que 95 %).

Correction de l'exercice 13 - Détecter un dé truqué par simulation

Protocole.

1. Construire l'intervalle de fluctuation sous l'hypothèse $p = \frac{1}{6}$. 2. Simuler de nombreux échantillons avec la **vraie** probabilité 0,22. 3. Compter la proportion d'échantillons **détectés** (hors de l'intervalle).

- **Intervalle sous $p = \frac{1}{6}$** : $E(X) = 50$ et $\sigma(X) = \sqrt{300 \times \frac{1}{6} \times \frac{5}{6}} \approx 6,4550$, donc [37,1 ; 62,9] — détection si $X \leq 37$ ou $X \geq 63$.
- **Loi réelle** : $X \sim B(300 ; 0,22)$, avec $E(X) = 66$ et $\sigma(X) = \sqrt{300 \times 0,22 \times 0,78} \approx 7,1750$.

- **Probabilité de détection** : $P(X \geq 63) \approx 0,68358$ (la branche basse est négligeable).
- **Puissance du test** : environ 68 % ✓
- **Interprétation** : le test détecte le truquage dans environ deux cas sur trois — c'est **insuffisant** pour une conclusion fiable.
- **En écarts types** : 63 est à $\frac{63 - 66}{7,1750} \approx -0,418\sigma$ de l'espérance réelle, d'où une probabilité de dépassement légèrement supérieure à $\frac{1}{2}$ ✓
- **Avec $n = 1000$ lancers** : l'intervalle sous $\frac{1}{6}$ devient $[166,7 - 23,57; 166,7 + 23,57] \approx [143,1; 190,2]$, tandis que la loi réelle a pour espérance 220 et pour écart type $\approx 13,10$.
- **Détection** : 190 est à $\frac{190 - 220}{13,10} \approx -2,290\sigma$, donc $P(X \geq 191) \approx 0,989$ — la puissance monte à 99 % ✓
- **Leçon** : la **puissance** d'un test (sa capacité à détecter un écart réel) croît avec n . Pour un écart de 0,22 contre 0,167, il faut environ 1000 lancers pour être tranquille ✓
- **Rôle de la simulation** : elle permet d'estimer cette puissance sans calcul théorique, ce qui est précieux dès que le modèle se complique.

Réponse. Puissance d'environ 68 % pour $n = 300$, contre 99 % pour $n = 1000$.

Correction de l'exercice 14 - Erreur dans une simulation

- **Erreur** : le test est $s > 3$, ce qui estime $P(X \geq 4)$ et non $P(X \geq 3)$ ✓
- **Correction** : écrire $s \geq 3$.
- **Valeurs théoriques** : $P(X \geq 3) \approx 0,61722$ et $P(X \geq 4) \approx 0,35039$.
- **Écart** : le script erroné affiche environ 0,35 au lieu de 0,62 — une erreur **énorme**, égale à $P(X = 3) \approx 0,26683$ ✓
- **Contrôle** : $0,61722 - 0,35039 = 0,26683$ ✓
- **Détectabilité** : avec $N = 10\,000$, l'écart type de l'estimation vaut environ $\sqrt{\frac{0,35 \times 0,65}{10\,000}} \approx 0,0048$ — l'erreur de 0,267 représente **55 écarts types**, donc elle est immédiatement visible ✓
- **Bonne pratique** : comparer systématiquement le résultat d'une simulation à une valeur théorique connue, même approximative.
- **Piège jumeau** : `range(10)` parcourt j de 0 à 9, soit bien **dix** épreuves ✓ — là, pas d'erreur.
- **Autre piège classique** : oublier de réinitialiser $s = 0$ à chaque échantillon ferait cumuler les succès sur toutes les itérations, donnant un résultat aberrant proche de 1 ✓
- **Ordre de grandeur de contrôle** : $E(X) = 3$, donc $P(X \geq 3)$ doit être **supérieur** à 0,5 — le 0,35 affiché aurait dû alerter immédiatement ✓

Réponse. Le test $s > 3$ estime $P(X \geq 4) \approx 0,350$ au lieu de $P(X \geq 3) \approx 0,617$: il faut $s \geq 3$.

Correction de l'exercice 15 - Nombre de simulations nécessaire

- **Précision absolue visée** : 10 % de 0,05, soit $\delta = 0,005$.
- **Condition** : $2\sqrt{\frac{0,05 \times 0,95}{N}} \leq 0,005$.
- **Calcul** : $\sqrt{0,0475} \approx 0,21794$, donc $\frac{0,43589}{\sqrt{N}} \leq 0,005$.
- **Résolution** : $\sqrt{N} \geq 87,178$, donc

$$N \geq 7600$$

- **Contrôle** : pour $N = 7600$, la demi-amplitude vaut $\frac{0,43589}{87,178} \approx 0,0050$ ✓
- **Nombre de succès attendus** : $7600 \times 0,05 = 380$ ✓
- **Règle pratique** : pour une précision relative de 10 %, il faut environ **400 succès** — donc $N \approx \frac{400}{p}$ ✓
- **Contrôle** : $\frac{400}{0,05} = 8000$, du bon ordre ✓
- **Pour** $p = 0,001$: il faudrait $N \approx 400\,000$ simulations — l'estimation des **événements rares** est très coûteuse ✓
- **Contrôle** : $2\sqrt{\frac{0,001 \times 0,999}{400\,000}} \approx 2 \times 0,00004998 \approx 0,0001$, soit 10 % de 0,001 ✓
- **Leçon** : c'est le **nombre de succès**, et non N , qui détermine la précision relative. D'où les méthodes spécialisées (échantillonnage préférentiel) pour les événements rares.

Réponse. $N \geq 7600$ simulations (règle : environ 400 succès attendus).

Correction de l'exercice 16 - Synthèse : protocole complet

a) Protocole.

1. **Hypothèse** : $p = 0,10$. 2. **Paramètres** : $E(X) = 50$ et $\sigma(X) = \sqrt{500 \times 0,1 \times 0,9} = \sqrt{45} \approx 6,7082$. 3. **Intervalle de fluctuation** ($\pm 2\sigma$) : $[50 - 13,42; 50 + 13,42] \approx [36,6; 63,4]$. 4. **Règle de décision** : rejet si $X \leq 36$ ou $X \geq 64$.

b) Observation : $X = 36$.

- $36 \notin [36,6; 63,4] \rightarrow$ on **rejette** l'annonce ✓
- **En écarts types** : $\frac{36 - 50}{6,7082} \approx -2,087 \sigma$ ✓
- **Probabilité exacte** : $P(X \leq 36) \approx 0,018633$, soit environ 1,9 % ✓
- **Fréquence observée** : $\frac{36}{500} = 0,072$ ✓
- **Intervalle de confiance affiné** : $2\sqrt{\frac{0,072 \times 0,928}{500}} \approx 2 \times 0,011557 \approx 0,023114$, soit $[0,0489; 0,0951]$ — il **exclut** 0,10 ✓
- **Conclusion** : le taux réel est plutôt autour de 7,2 %, et l'annonce de 10 % est réfutée.

c) Puissance si $p = 0,07$.

- **Loi réelle** : $X \sim B(500; 0,07)$, avec $E(X) = 35$ et $\sigma(X) = \sqrt{500 \times 0,07 \times 0,93} \approx 5,7054$.
- **Détection** : il faut $X \leq 36$ (la branche haute est hors de portée).
- **Probabilité** : $P(X \leq 36) \approx 0,61264$ ✓
- **Puissance** : environ 61 % — le test ne détecte le biais que **trois fois sur cinq** ✓
- **En écarts types** : 36 est à $\frac{36 - 35}{5,7054} \approx 0,175 \sigma$ au-dessus de l'espérance réelle, d'où une probabilité un peu supérieure à $\frac{1}{2}$ ✓
- **Avec 2000 parties** : l'intervalle sous $p = 0,1$ devient $[200 - 26,83; 200 + 26,83] \approx [173,2; 226,8]$, alors que la loi réelle a pour espérance 140 et pour écart type $\approx 11,41$.
- **Détection** : 173 est à $\frac{173 - 140}{11,41} \approx 2,892 \sigma$ au-dessus, donc la détection est quasi certaine (environ 99,8 %) ✓
- **Recommandation** : pour tester sérieusement une machine à sous, il faut au moins **2000 parties** — ce qui explique pourquoi ces contrôles sont menés par des

organismes disposant de gros volumes de données ✓

Réponse. a) rejet si $X \leq 36$ ou $X \geq 64$ · b) rejet (36, soit $-2,09 \sigma$) · c) puissance d'environ 61 % (environ 99,8 % sur 2000 parties).