COURS

Bienaymé-Tchebychev — Concentration, loi des grands nombres

Tle

Programme de terminale générale, enseignement de spécialité — BO spécial n°8 du 25 juillet 2019

Sur 180 lancers de dé, on attend 30 six ; on en obtient rarement 30, mais presque jamais 10 ou 50. **L'écart type σ est l'unité naturelle des écarts à l'espérance : un écart d'au moins $k\sigma$ a une probabilité au plus $\frac{1}{k^2}$ (Bienaymé-Tchebychev). Pour la moyenne d'un échantillon de taille n, cette unité devient $\frac{\sigma}{\sqrt{n}}$: la moyenne se concentre autour de l'espérance, c'est la loi des grands nombres.**

1 L'inégalité de Bienaymé-Tchebychev

PROPRIÉTÉ

Inégalité de Bienaymé-Tchebychev (démontrée). Soit X une variable aléatoire d'espérance μ et de variance V. Pour tout réel $\delta > 0$:

$$P\big(|X-\mu| \geqslant \delta\big) \leqslant \frac{V}{\delta^2}$$

Preuve. X prend les valeurs x_i avec les probabilités p_i. Dans $V = \sum p_i\,(x_i-\mu)^2$, somme de termes positifs, on ne garde que les termes où $|x_i-\mu| \geqslant \delta$, qui valent chacun au moins $p_i\,\delta^2$: $V \geqslant \delta^2 \sum_{|x_i-\mu|\geqslant\delta} p_i = \delta^2\,P(|X-\mu| \geqslant \delta)$, et l'on divise par $\delta^2 > 0$.

PROPRIÉTÉ

Deux lectures. Par le contraire, $P(|X-\mu| < \delta) \geqslant 1-\frac{V}{\delta^2}$: X tombe dans $]\mu-\delta\,;\mu+\delta[$ avec une probabilité au moins $1-\frac{V}{\delta^2}$. Avec $\delta = k\sigma$ ($\sigma=\sqrt{V}$, $k>0$), la borne devient $\frac{1}{k^2}$: $\frac{1}{4}$ au-delà de 2σ, $\frac{1}{9}$ au-delà de 3σ. Des écarts de quelques σ sont **improbables** : $P(|X-\mu| \geqslant k\sigma) \leqslant \frac{1}{k^2}$. Pour $\delta \leqslant \sigma$, la borne vaut au moins 1 et n'apprend rien.

APPLIQUER L'INÉGALITÉ DE BIENAYMÉ-TCHEBYCHEV

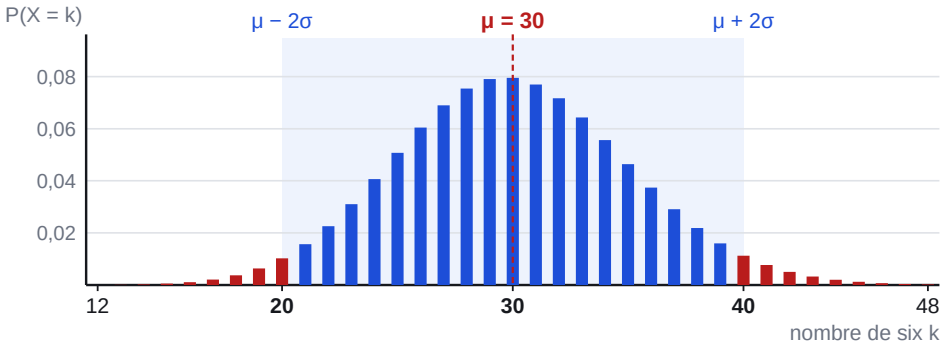
1. **Calculer μ et la variance V**, pas l'écart type. Rappel : pour la loi $\mathcal{B}(n\,;p)$, $\mu = np$ et $V = np(1-p)$.
2. **Écrire l'événement comme un écart à μ**, $\{|X-\mu| \geqslant \delta\}$ ou son contraire : un intervalle **centré en μ**. Si X et μ sont entiers, $|X-\mu| \leqslant 9$ équivaut à $|X-\mu| < 10$.
3. **Conclure par une inégalité**, jamais par une égalité : on obtient une borne, pas la probabilité.

EXEMPLE

On lance 180 fois un dé équilibré ; X est le nombre de six. Majorer $P(|X-30| \geqslant 10)$, puis minorer $P(20 < X < 40)$.

1. Lancers indépendants : X suit $\mathcal{B}\left(180\,;\frac{1}{6}\right)$, $\mu = 30$ et
   $V = 180 \times \frac{1}{6} \times \frac{5}{6} = 25$
   $\sigma = 5$: un écart de 10 est un écart de 2σ.
2. Avec $\delta = 10$, $P(|X-30| \geqslant 10) \leqslant \frac{V}{\delta^2}$, et :
   $\frac{25}{10^2} = 0{,}25$
   $\{20 < X < 40\} = \{|X-30| < 10\}$ est l'événement contraire.

$P(|X-30| \geqslant 10) \leqslant 0{,}25$ et $P(20 < X < 40) \geqslant 1-0{,}25 = 0{,}75$. La loi binomiale donne la valeur exacte de l'écart, environ 0,057 : la borne est juste, mais large.



X suit B(180 ; 1/6)
μ = 30, σ = 5
en rouge : |X − 30| ≥ 10
probabilité exacte ≈ 0,057
B.-T. garantit ≤ 0,25

En rouge, les valeurs à au moins $2\sigma = 10$ de $\mu = 30$: probabilité totale environ 0,057, quand l'inégalité garantit 0,25 au plus.

Universelle, donc large. Valable pour **toute** variable aléatoire, l'inégalité est en contrepartie loin d'être optimale : elle borne par $\frac{1}{4}$ un écart d'au moins 2σ, probabilité souvent inférieure à 0,05 en simulation. Sans hypothèse de plus, on ne peut pas faire mieux : l'exercice 7 atteint la borne. **Algorithme.** `proba_ecart` ajoute les $\binom{n}{k}p^k(1-p)^{n-k}$ des k tels que $|k - pn| > \sqrt{n}$; avec $V = np(1-p)$ et $\delta = \sqrt{n}$, l'inégalité majore $P(|S_n - pn| \geqslant \sqrt{n})$, donc aussi cette probabilité, par $p(1-p)$, pour tout n. Pour $p = 0{,}5$, $\sqrt{n} = 2\sigma$.

P(|Sn − pn| > √n) pour Sn de loi B(n ; p)

```
from math import comb, sqrt
def proba_ecart(n, p):
    s = 0
    for k in range(n + 1):
        if abs(k - n*p) > sqrt(n):
            s = s + comb(n, k) * p**k * (1 - p)**(n - k)
    return s
for n in [25, 100, 400, 1000]:
    print(n, round(proba_ecart(n, 0.5), 4), 0.5 * (1 - 0.5))
```

sortie du programme, p = 0,5

n	exacte	B.-T.
25	0,0433	0,25
100	0,0352	0,25
400	0,0402	0,25
1000	0,0463	0,25

La probabilité exacte reste voisine de 0,04 quand n varie ; la borne vaut $p(1-p) = 0{,}25$ pour tout n.

PIÈGE

Écrire la borne $\frac{\sigma}{\delta^2}$, ou $\frac{V}{\delta}$. — c'est la **variance** sur le **carré** de l'écart. Avec $\sigma = 5$ et $\delta = 10$, $\frac{\sigma}{\delta^2} = 0{,}05$ est cinq fois trop petit, donc faux.

Ce que ça coûte : la borne, et toute taille d'échantillon qu'on en déduit.

Le réflexe : $\frac{V}{\delta^2} = \left(\frac{\sigma}{\delta}\right)^2$: avec $\delta = k\sigma$, on doit retrouver $\frac{1}{k^2}$.

2 Concentration de la moyenne, loi des grands nombres

DÉFINITION

Rappel du chapitre « Sommes de variables aléatoires ». Un **échantillon de taille** n d'une loi est une liste $(X_1, \ldots, X_n)$ de variables aléatoires **indépendantes** suivant cette loi, d'espérance μ et de variance V. Avec $S_n = X_1 + \cdots + X_n$ et $M_n = \frac{S_n}{n}$, la linéarité de l'espérance et l'additivité de la variance donnent $E(S_n) = n\mu$, $V(S_n) = nV$, puis :

$$E(M_n) = \mu \qquad V(M_n) = \frac{1}{n^2} \times nV = \frac{V}{n} \qquad \sigma(M_n) = \frac{\sigma}{\sqrt{n}}$$

PROPRIÉTÉ

Inégalité de concentration (démontrée). Si M_n est la moyenne d'un échantillon de taille n d'une variable aléatoire d'espérance μ et de variance V, alors pour tout réel $\delta > 0$, l'inégalité ci-dessous a lieu. *Preuve* : Bienaymé-Tchebychev appliquée à M_n, d'espérance μ et de variance $\frac{V}{n}$, donne la borne $\frac{V/n}{\delta^2}$.

$$P\big(|M_n - \mu| \geqslant \delta\big) \leqslant \frac{V}{n\delta^2}$$

PROPRIÉTÉ

Loi des grands nombres (démontrée). Pour tout réel $\delta > 0$, $P(|M_n - \mu| \geqslant \delta)$ tend vers 0 quand n tend vers $+\infty$. *Preuve* : $0 \leqslant P(|M_n - \mu| \geqslant \delta) \leqslant \frac{V}{\delta^2} \times \frac{1}{n}$, et $\frac{1}{n}$ tend vers 0 : le théorème des gendarmes conclut. Quelle que soit la précision δ fixée **à l'avance**, un grand échantillon a une moyenne qui ne s'écarte de μ d'au moins δ qu'avec une probabilité aussi petite qu'on veut.

PROPRIÉTÉ

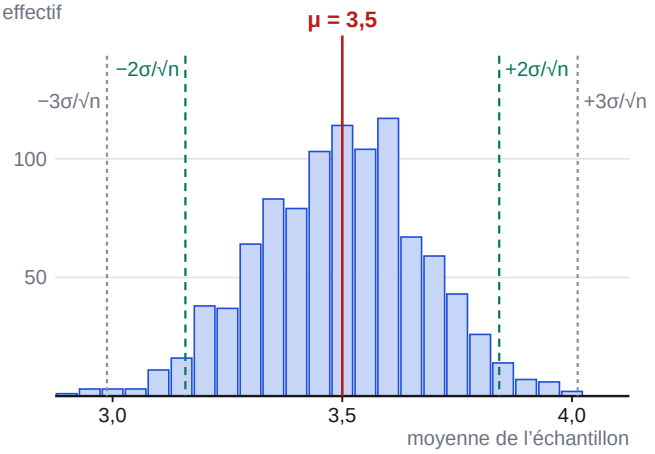
Fréquence et probabilité. Dans un schéma de Bernoulli, avec $X_i = 1$ en cas de succès et 0 sinon, $\mu = p$, $V = p(1-p)$ et $M_n = F_n$ est la **fréquence** des succès. À δ près, F_n est proche de p avec une probabilité qui tend vers 1 : c'est le « théorème d'or » de Jacques Bernoulli (1713), qui justifie d'estimer une probabilité par une fréquence. Comme $p(1-p) = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \leqslant \frac{1}{4}$, on a, quel que soit p, $P(|F_n - p| \geqslant \delta) \leqslant \frac{1}{4n\delta^2}$.

SIMULER N ÉCHANTILLONS DE TAILLE N

1. **Distinguer n et N** : n est la taille d'un échantillon, N le nombre d'échantillons. Les N moyennes forment une série, d'écart type s.
2. **Comparer s à $\frac{\sigma}{\sqrt{n}}$**, et non à $\frac{\sigma}{\sqrt{N}}$: c'est n qui fixe la dispersion d'une moyenne. Compter les moyennes à au plus ks, ou $k\frac{\sigma}{\sqrt{n}}$, de μ ($k = 1$, 2, 3), et confronter à la garantie $1 - \frac{1}{k^2}$: 0,75 pour $k = 2$, environ 0,889 pour $k = 3$.

N échantillons de taille n d'un dé

```
from random import random, seed
from math import sqrt
seed(2026)
def de():
    return int(6 * random()) + 1
def moyenne(n):
    return sum(de() for i in range(n)) / n
N, n = 1000, 100
mu, sigma = 3.5, sqrt(35 / 12)
M = [moyenne(n) for j in range(N)]
m = sum(M) / N
s = sqrt(sum((x - m)**2 for x in M) / N)
t = sigma / sqrt(n)
print(round(s, 4), round(t, 4))
for k in [1, 2, 3]:
    a = len([x for x in M if abs(x - mu) <= k*s])
    b = len([x for x in M if abs(x - mu) <= k*t])
    print(k, a / N, b / N)
```



Graine 2026 : sortie 0.1796 0.1708, puis 1 0.667 0.667, 2 0.954 0.945, 3 0.996 0.995.

EXEMPLE

On lance $n = 100$ fois un dé équilibré ($\mu = 3{,}5$). Majorer $P(|M_{100} - 3{,}5| \geqslant 0{,}5)$, puis confronter la théorie à la simulation ci-dessus.

1. Pour un lancer, $E(X^2) = \frac{91}{6}$, donc $V = \frac{91}{6} - 3{,}5^2 = \frac{35}{12}$. Concentration avec $\delta = 0{,}5$:
$$\frac{35}{12} \div (100 \times 0{,}5^2) = \frac{7}{60} \approx 0{,}117$$
dans la simulation, 7 moyennes sur 1 000 seulement s'écartent autant de 3,5.
2. Écart type théorique de M_{100} : $\frac{\sigma}{\sqrt{100}} \approx 0{,}1708$, proche de $s \approx 0{,}1796$; à au plus $\frac{2\sigma}{\sqrt{n}}$ de μ, 94,5 % des moyennes, pour une garantie de 75 %.

$P(|M_{100} - 3{,}5| \geqslant 0{,}5) \leqslant 0{,}117$ **; la simulation confirme le facteur $\frac{1}{\sqrt{n}}$, et des bornes justes mais larges.**

DÉFINITION

Un pion part de 0 sur une droite graduée ; à chaque pas, il avance ou recule de 1, avec la probabilité $\frac{1}{2}$ chacun, indépendamment des pas précédents. Après n pas, il est en $S_n = X_1 + \cdots + X_n$, où X_i vaut 1 ou -1 : $E(X_i) = 0$, $V(X_i) = 1$, donc $E(S_n) = 0$ et $V(S_n) = n$. Ainsi $P(|S_n| \geqslant 2\sqrt{n}) \leqslant \frac{1}{4}$, et $\frac{S_n}{n}$ se concentre autour de 0, mais pas S_n.

une marche de 400 pas

```
from random import random, seed
seed(2026)
def marche(n):
    S = [0]
    for i in range(n):
        if random() < 0.5: S.append(S[-1] + 1)
        else: S.append(S[-1] - 1)
    return S
S = marche(400)
c = 0
for j in range(1000):
    if abs(marche(100)[-1]) >= 20:
        c = c + 1
print(S[400], S[400] / 400, c / 1000)
```



Graine 2026, sortie -34 -0.085 0.065 : S_{400} s'éloigne de 0, $\frac{S_{400}}{400}$ est petit, et 65 marches de 100 pas sur 1 000 finissent à au moins 20 de 0 (borne : 0,25).

PIÈGE

« Dix pile de suite : d'après la loi des grands nombres, face est devenu plus probable. » Ou : « $M_n = \mu$ pour n assez grand. » — les lancers sont indépendants : face garde la probabilité $\frac{1}{2}$. La loi ne compense rien, elle **dilue** l'écart en divisant par n ; et M_n reste aléatoire : seule la **probabilité** d'un écart d'au moins δ tend vers 0.

Ce que ça coûte : la justification, ou l'énoncé du théorème.

Le réflexe : énoncer la loi avec un δ fixé et une probabilité, jamais sur le prochain tirage.

3 Choisir une taille d'échantillon

TROUVER UNE TAILLE D'ÉCHANTILLON SUFFISANTE

1. **Fixer la précision δ et le risque α** : on veut $P(|M_n - \mu| \geqslant \delta) \leqslant \alpha$.
2. **Condition suffisante** : par concentration, il suffit que $\frac{V}{n\delta^2} \leqslant \alpha$, soit $n \geqslant \frac{V}{\alpha\,\delta^2}$. Si V est inconnue, on la majore : $\frac{1}{4}$ pour une fréquence.
3. **Arrondir à l'entier supérieur** et conclure « une taille de … suffit ».
   condition suffisante, pas nécessaire : une taille plus petite convient peut-être, sans que l'inégalité le prouve.

EXEMPLE

Un institut estime la proportion p de personnes favorables à un projet par la fréquence F_n observée sur n personnes choisies au hasard, de façon indépendante. Il veut $P(|F_n - p| \geqslant 0{,}05) \leqslant 0{,}05$. Quelle taille suffit ?

1. p est inconnue : $P(|F_n - p| \geqslant 0{,}05) \leqslant \frac{1}{4n \times 0{,}05^2}$, et il suffit que cette borne soit au plus 0,05, soit n au moins

$$\frac{1}{4 \times 0{,}05 \times 0{,}05^2} = 2000$$

2 000 personnes suffisent (contrôle : $\frac{1}{4 \times 2\,000 \times 0{,}05^2} = 0{,}05$). Il n'en faut pas forcément autant : pour $p = 0{,}5$ et $n = 400$, le calcul binomial exact donne déjà $P(|F_{400} - 0{,}5| \geqslant 0{,}05) \approx 0{,}051$.

PIÈGE

Résoudre $\frac{V}{n\delta^2} \leqslant \alpha$ en $n \leqslant \frac{V}{\alpha\delta^2}$, ou arrondir $1\,562{,}5$ à $1\,562$. — n est au dénominateur : ce sont les n **assez grands** qui conviennent. Arrondi par défaut, n ne vérifie plus la condition.

Ce que ça coûte : la taille demandée.

Le réflexe : remplacer n par la valeur trouvée : la borne doit être au plus α.

À VÉRIFIER

- Majorer $P(|X - \mu| \geqslant \delta)$ par $\frac{V}{\delta^2}$, avec la **variance**, et minorer le contraire.
- Ramener un intervalle à un écart **centré en** μ ; démontrer concentration et loi des grands nombres.
- Trouver une taille d'échantillon **suffisante**, arrondie par excès.
- Lire une simulation : n et N, s et $\frac{\sigma}{\sqrt{n}}$, garanties $1 - \frac{1}{k^2}$.

À RETENIR

- **Bienaymé-Tchebychev** : $P(|X - \mu| \geqslant \delta) \leqslant \dfrac{V}{\delta^2}$; avec $\delta = k\sigma$, la borne vaut $\dfrac{1}{k^2}$, universelle mais large.
- Moyenne d'un échantillon : $E(M_n) = \mu$, $V(M_n) = \dfrac{V}{n}$, $\sigma(M_n) = \dfrac{\sigma}{\sqrt{n}}$.
- **Concentration** : $P(|M_n - \mu| \geqslant \delta) \leqslant \dfrac{V}{n\delta^2}$, qui tend vers 0 : c'est la **loi des grands nombres**.
- Taille suffisante : $n \geqslant \dfrac{V}{\alpha\delta^2}$, arrondi par excès, avec $V \leqslant \dfrac{1}{4}$ pour une fréquence.