

# Sommes de variables aléatoires et loi des grands nombres

## Cours de maths - Terminale

Un sondeur interroge mille personnes, un casino fait tourner sa roulette des millions de fois, une usine contrôle un échantillon de pièces : à chaque fois, on **additionne ou on moyenne** beaucoup de résultats aléatoires, et pourtant le résultat final semble étrangement prévisible. Dans ce chapitre, tu apprends à calculer l'espérance et la variance d'une somme de variables aléatoires, tu découvres les **inégalités de concentration** qui mesurent l'écart à la moyenne, puis tu démontres la **loi des grands nombres**, qui justifie qu'une fréquence observée se rapproche d'une probabilité. Mathbot te propose même de simuler tout cela en Python !

## 1. Rappels : espérance, variance, indépendance

Une **variable aléatoire**  $X$  associe un nombre réel à chaque issue d'une expérience aléatoire. Si  $X$  prend les valeurs  $x_1, \dots, x_r$  avec les probabilités  $p_i = P(X = x_i)$ , on retrouve les trois grandeurs de la classe de 1ère.

### Espérance, variance, écart-type

- Espérance :  $E(X) = x_1 p_1 + \dots + x_r p_r$  (la valeur moyenne « à long terme »).
- Variance :  $V(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$  (la dispersion autour de la moyenne).
- Écart-type :  $\sigma(X) = \sqrt{V(X)}$ , exprimé dans la même unité que  $X$ .

### Variables indépendantes

Deux variables aléatoires  $X$  et  $Y$  sont **indépendantes** si, pour tous réels  $x$  et  $y$ ,  $P(X = x \text{ et } Y = y) = P(X = x) \times P(Y = y)$ . C'est le cas, par exemple, des résultats de deux lancers successifs d'un même dé.

### Exemple : le dé équilibré

Soit  $D$  le numéro obtenu avec un dé équilibré à six faces.  $E(D) = (1 + 2 + 3 + 4 + 5 + 6) \div 6 = 7/2$ .  $E(D^2) = (1 + 4 + 9 + 16 + 25 + 36) \div 6 = 91/6$ , donc  $V(D) = 91/6 - (7/2)^2 = 91/6 - 49/4 = 35/12 \approx 2,92$  et  $\sigma(D) \approx 1,71$ . Garde ces valeurs : elles serviront souvent.

## 2. Linéarité de l'espérance

### Propriété (linéarité de l'espérance)

Pour toutes variables aléatoires  $X$  et  $Y$  définies sur la même expérience et pour tous réels  $a$  et  $b$  :

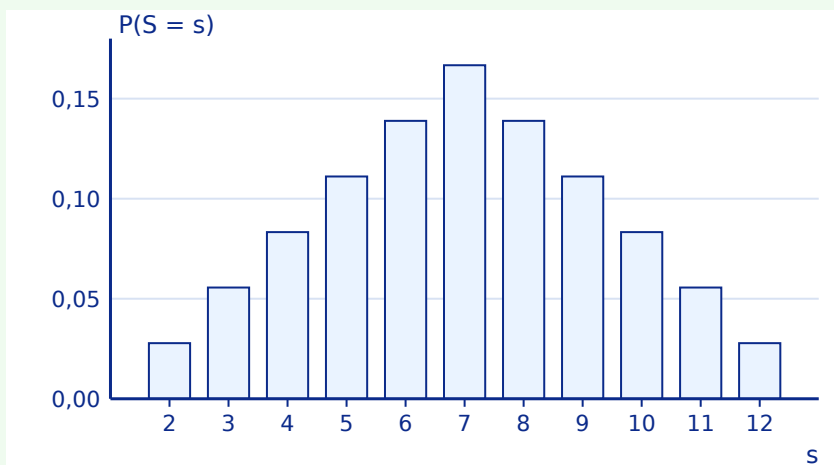
- $E(X + Y) = E(X) + E(Y)$  ;
- $E(aX + b) = a E(X) + b$ .

Ces égalités sont vraies **sans aucune hypothèse d'indépendance**.

**Démonstration.** Notons  $\Omega$  l'ensemble (fini) des issues. L'espérance de  $X$  vaut la somme, sur toutes les issues  $\omega$ , de  $X(\omega) \times P(\omega)$ . Donc  $E(X + Y) = \sum (X(\omega) + Y(\omega)) P(\omega) = \sum X(\omega)P(\omega) + \sum Y(\omega)P(\omega) = E(X) + E(Y)$ . De même  $E(aX + b) = \sum (aX(\omega) + b) P(\omega) = a \sum X(\omega)P(\omega) + b \sum P(\omega) = a E(X) + b$ , car la somme de toutes les probabilités vaut 1.

### Exemple : deux dés

On lance deux dés équilibrés et on note  $S$  la somme des deux numéros. Plutôt que de dresser la loi de  $S$ , on écrit  $S = D_1 + D_2$  :  $E(S) = E(D_1) + E(D_2) = 7/2 + 7/2 = 7$ . La loi de  $S$ , tracée ci-dessous, est bien symétrique autour de 7.



### Astuce : décomposer pour calculer

Pour trouver l'espérance d'une quantité compliquée (nombre de succès, gain total...), cherche à l'écrire comme **somme de variables simples** dont tu connais l'espérance. La linéarité fait le reste, même si ces variables dépendent les unes des autres.

## 3. Variance d'une somme de variables indépendantes

### Propriété (variance et indépendance)

- Pour tous réels  $a$  et  $b$  :  $V(aX + b) = a^2 V(X)$  (ajouter  $b$  ne change pas la dispersion).
- Si  $X$  et  $Y$  sont **indépendantes** :  $V(X + Y) = V(X) + V(Y)$ .

**Démonstration de la seconde égalité.** On a  $V(X + Y) = E((X + Y)^2) - (E(X) + E(Y))^2$ . En développant,  $E((X + Y)^2) = E(X^2) + 2E(XY) + E(Y^2)$  et  $(E(X) + E(Y))^2 = (E(X))^2 + 2E(X)E(Y) + (E(Y))^2$ . Donc  $V(X + Y) = V(X) + V(Y) + 2(E(XY) - E(X)E(Y))$ . Pour des variables indépendantes,  $E(XY) = E(X)E(Y)$  (la probabilité de chaque couple se factorise), le dernier terme est nul et  $V(X + Y) = V(X) + V(Y)$ .

### Exemple : suite des deux dés

Les deux dés sont indépendants, donc  $V(S) = V(D_1) + V(D_2) = \mathbf{35/12} + \mathbf{35/12} = \mathbf{35/6} \approx 5,83$ . On peut le confirmer avec la loi de  $S$  :  $E(S^2) = \mathbf{329/6}$ , donc  $V(S) = \mathbf{329/6} - 49 = \mathbf{35/6}$ .

### Pièges classiques

- $V(X + Y) = V(X) + V(Y)$  **exige l'indépendance**. Contre-exemple : si  $Y = X$ , alors  $V(X + Y) = V(2X) = 4V(X) \neq 2V(X)$ .
- Pour une différence :  $V(X - Y) = V(X) + V(-Y) = V(X) + V(Y)$ . Les variances **s'ajoutent toujours**, jamais elles ne se retranchent.
- $V(aX) = a^2 V(X)$  et non  $a V(X)$  ; l'écart-type, lui, est multiplié par  $|a|$ .

### Application : retrouver la loi binomiale

Une variable de Bernoulli  $B$  de paramètre  $p$  a pour espérance  $p$  et variance  $p(1 - p)$ . Une variable  $X$  qui suit la loi binomiale  $B(n ; p)$  est la somme de  $n$  variables de Bernoulli indépendantes. Par linéarité  $E(X) = np$ , et par indépendance  $V(X) = np(1 - p)$ . Pour  $n = 60$  et  $p = \mathbf{1/6}$ , on obtient  $E(X) = 10$  et  $V(X) = 60 \times \mathbf{1/6} \times \mathbf{5/6} = \mathbf{25/3}$ .

## 4. Échantillon, somme et moyenne d'un échantillon

### Échantillon de variables aléatoires

Un **échantillon de taille  $n$**  d'une loi de probabilité est une liste  $(X_1 ; X_2 ; \dots ; X_n)$  de  $n$  variables aléatoires **indépendantes** qui suivent toutes **cette même loi**. On note  $\mu$  leur espérance commune,  $V$  leur variance commune et  $\sigma$  leur écart-type commun. On définit :

- la somme  $S_n = X_1 + X_2 + \dots + X_n$  ;
- la **moyenne**  $M_n = S_n \div n = (X_1 + X_2 + \dots + X_n) \div n$ .

Répéter  $n$  fois, de façon indépendante, une même expérience (lancer un dé, tirer avec remise, interroger au hasard une personne) fournit un échantillon de taille  $n$ .

### Propriété (somme et moyenne d'un échantillon)

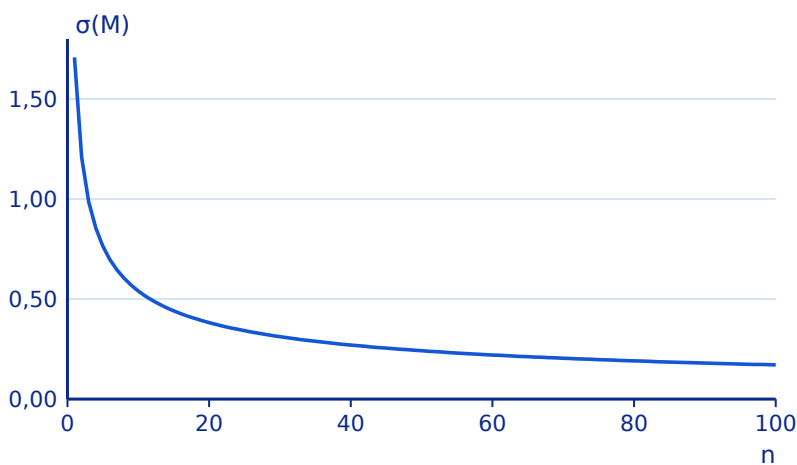
|               | Espérance | Variance                | Écart-type                          |
|---------------|-----------|-------------------------|-------------------------------------|
| Somme $S_n$   | $n\mu$    | $nV$                    | $\sigma\sqrt{n}$                    |
| Moyenne $M_n$ | $\mu$     | <b><math>V/n</math></b> | <b><math>\sigma/\sqrt{n}</math></b> |

**Justification.** Par linéarité,  $E(S_n) = \mu + \dots + \mu = n\mu$ . Par indépendance,  $V(S_n) = V + \dots + V = nV$ . Comme  $M_n = \frac{1}{n} S_n$ , on a  $E(M_n) = \frac{1}{n} \times n\mu = \mu$  et  $V(M_n) = (\frac{1}{n})^2 \times nV = \frac{V}{n}$ .

### Exemple : cent lancers de dé

Pour  $n = 100$  lancers ( $\mu = 7/2$ ,  $V = 35/12$ ) :  $E(S_{100}) = 350$ ,  $V(S_{100}) = 3500/12 \approx 291,7$  et  $\sigma(S_{100}) \approx 17,1$ . Pour la moyenne :  $E(M_{100}) = 3,5$ ,  $V(M_{100}) = 35/1200 = 7/240 \approx 0,0292$  et  $\sigma(M_{100}) \approx 0,171$ .

L'écart-type de la moyenne diminue comme  $1/\sqrt{n}$  : pour le dé, il passe de 1,71 (un seul lancer) à 0,171 (cent lancers). La courbe ci-dessous montre cette décroissance.



### Deux effets opposés

Quand  $n$  augmente, la **somme**  $S_n$  se disperse de plus en plus ( $\sigma\sqrt{n}$  grandit), alors que la **moyenne**  $M_n$  se concentre de plus en plus autour de  $\mu$  ( $\sigma \div \sqrt{n}$  diminue). Ne confonds jamais les deux.

## 5. Inégalité de Markov

### Inégalité de Markov

Soit  $X$  une variable aléatoire à **valeurs positives ou nulles** et a un réel strictement

positif. Alors  $P(X \geq a) \leq (E(X)) \div a$ .

**Démonstration.** Dans la somme qui définit  $E(X) = \sum x_i p_i$ , tous les termes sont positifs. En ne gardant que les termes pour lesquels  $x_i \geq a$ , on obtient  $E(X) \geq (\text{somme des } x_i p_i \text{ avec } x_i \geq a) \geq a \times (\text{somme des } p_i \text{ avec } x_i \geq a) = a P(X \geq a)$ . Il suffit de diviser par  $a > 0$ .

### Exemple

Un standard téléphonique reçoit en moyenne 4 appels par minute en soirée ; soit  $X$  le nombre d'appels reçus en une minute ( $X \geq 0$  et  $E(X) = 4$ ). Alors  $P(X \geq 20) \leq 4/20 = 0,2$  : il y a au plus 20 % de chances de recevoir 20 appels ou plus en une minute.

### Quand la borne est inutile

Si  $a \leq E(X)$ , la borne  $(E(X)) \div a$  est supérieure ou égale à 1 : l'inégalité est vraie mais ne dit rien. Markov n'est intéressante que pour  $a$  nettement plus grand que l'espérance. Et la positivité de  $X$  est indispensable.

## 6. Inégalité de Bienaymé-Tchebychev

### Inégalité de Bienaymé-Tchebychev

Soit  $X$  une variable aléatoire d'espérance  $\mu$  et de variance  $V$ . Pour tout réel  $\delta > 0$  :  $P(|X - \mu| \geq \delta) \leq V/\delta^2$ .

**Démonstration.** La variable  $Y = (X - \mu)^2$  est positive et  $E(Y) = V$  par définition de la variance. L'événement «  $|X - \mu| \geq \delta$  » est le même que «  $(X - \mu)^2 \geq \delta^2$  ». L'inégalité de Markov appliquée à  $Y$  avec  $a = \delta^2$  donne  $P(Y \geq \delta^2) \leq V/\delta^2$ .

Par passage à l'événement contraire :  $P(|X - \mu| < \delta) \geq 1 - V/\delta^2$ , ce qui signifie que  $X$  est dans l'intervalle  $]\mu - \delta ; \mu + \delta[$  avec une probabilité au moins égale à  $1 - V/\delta^2$ .

### Exemple : la somme de deux dés

Pour  $S$  de la partie 2 ( $\mu = 7$  et  $V = 35/6$ ) avec  $\delta = 4$  :  $P(|S - 7| \geq 4) \leq 35/6 \div 16 = 35/96 \approx 0,365$ . La valeur exacte est bien plus petite :  $|S - 7| \geq 4$  correspond aux sommes 2, 3, 11 et 12, soit  $(1 + 2 + 2 + 1) \div 36 = 1/6 \approx 0,167$ . L'inégalité est vraie, mais **pessimiste** : c'est le prix de sa généralité, elle s'applique à toute loi.

### Méthode : utiliser Bienaymé-Tchebychev

1. Identifie  $\mu = E(X)$  et  $V = V(X)$ .

- Traduis l'événement demandé sous la forme  $|X - \mu| \geq \delta$  (ou son contraire  $|X - \mu| < \delta$ ).
- Calcule la borne  $V/\delta^2$  ; vérifie qu'elle est inférieure à 1, sinon elle n'apporte rien.
- Conclus par une phrase : « la probabilité est inférieure ou égale à ... » ou, pour le contraire, « supérieure ou égale à ... ».

## 7. Inégalité de concentration

Appliquons Bienaymé-Tchebychev à la moyenne  $M_n$  d'un échantillon, dont on connaît l'espérance  $\mu$  et la variance  $V/n$ .

### Inégalité de concentration

Soit  $(X_1 ; \dots ; X_n)$  un échantillon de variables aléatoires d'espérance  $\mu$  et de variance  $V$ , et  $M_n$  sa moyenne. Pour tout réel  $\delta > 0$  :  $P(|M_n - \mu| \geq \delta) \leq V/n\delta^2$ .

**Démonstration.** D'après la partie 4,  $E(M_n) = \mu$  et  $V(M_n) = V/n$ . L'inégalité de Bienaymé-Tchebychev appliquée à  $M_n$  donne  $P(|M_n - \mu| \geq \delta) \leq (V \div n) \div \delta^2 = V/n\delta^2$ .

### Exemple : mille lancers de dé

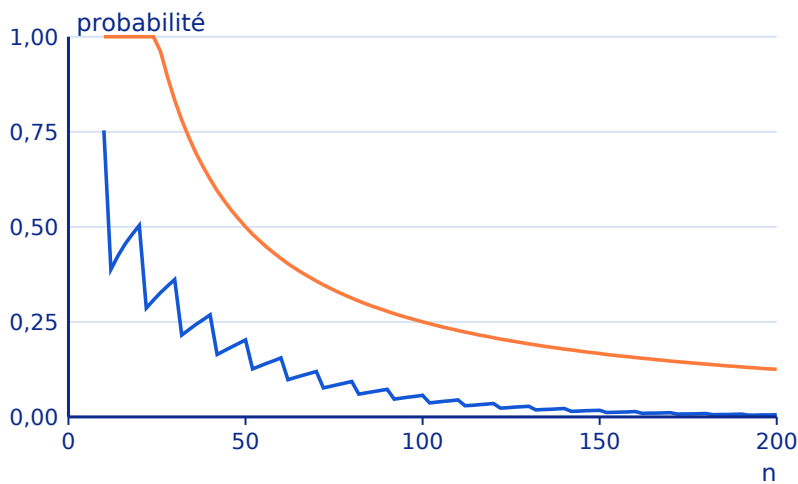
Avec  $V = 35/12$ ,  $\delta = 0,1$  et  $n = 1000$  :  $P(|M_{1000} - 3,5| \geq 0,1) \leq 35/12 \div (1000 \times 0,01) = 7/24 \approx 0,29$ . Autrement dit, avec une probabilité d'au moins 0,71, la moyenne des mille lancers est dans  $]3,4 ; 3,6[$ .

### Méthode : trouver la taille de l'échantillon

On cherche  $n$  pour que  $P(|M_n - \mu| \geq \delta)$  soit inférieure à un risque  $\alpha$  fixé. Il suffit que  $V/n\delta^2 \leq \alpha$ , c'est-à-dire  $n \geq V/\alpha\delta^2$ .

Pour le dé avec  $\delta = 0,1$  et  $\alpha = 0,05$  :  $n \geq 35/12 \div (0,01 \times 0,05) \approx 5833,3$ , donc il suffit de prendre  **$n = 5834$**  lancers.

Le graphique suivant compare, pour une pièce équilibrée ( $V = 0,25$ ) et  $\delta = 0,1$ , la borne  **$25/n$**  (en orange) à la vraie probabilité  $P(|F_n - 0,5| \geq 0,1)$  (en bleu), calculée avec la loi binomiale. La borne est toujours au-dessus, mais elle décroît elle aussi vers 0.



## 8. La loi des grands nombres

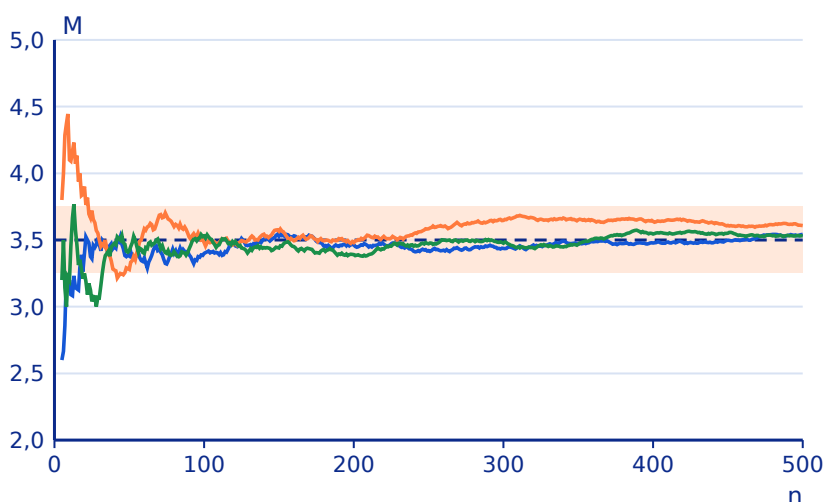
### Loi (faible) des grands nombres

Soit  $(X_1 ; \dots ; X_n)$  un échantillon de variables aléatoires d'espérance  $\mu$  et de moyenne  $M_n$ .  
Pour tout réel  $\delta > 0$  fixé :  $\lim_{n \rightarrow +\infty} P(|M_n - \mu| \geq \delta) = 0$ .

**Démonstration.** Pour tout  $n$ ,  $0 \leq P(|M_n - \mu| \geq \delta) \leq \mathbf{V/n\delta^2}$  d'après l'inégalité de concentration. Or  $\delta$  et  $V$  sont fixés, donc  $\mathbf{V/n\delta^2}$  tend vers 0 quand  $n$  tend vers  $+\infty$ . Le théorème des gendarmes donne le résultat.

**Interprétation.** Aussi petit que soit l'écart  $\delta$  choisi, la probabilité que la moyenne observée s'écarte de  $\mu$  d'au moins  $\delta$  devient aussi petite que l'on veut quand  $n$  est grand. Si les  $X_i$  sont des variables de Bernoulli de paramètre  $p$  (succès = 1, échec = 0), alors  $\mu = p$  et  $M_n$  est la **fréquence de succès** : la fréquence observée se rapproche de la probabilité  $p$ . C'est ce qui justifie l'estimation d'une probabilité par une fréquence dans une simulation ou un sondage.

Voici trois simulations indépendantes de la moyenne de  $n$  lancers d'un dé ; la bande orange représente l'intervalle  $]3,25 ; 3,75[$  autour de  $\mu = 3,5$ .



## Ce que la loi des grands nombres ne dit PAS

- Elle ne dit pas que le hasard « compense » : après dix « pile » de suite, la probabilité d'avoir « pile » au lancer suivant vaut toujours **1/2**, car les lancers sont indépendants.
- Elle ne dit pas que la **somme** se rapproche de  $n\mu$  : l'écart  $S_n - n\mu$  a un écart-type  $\sigma\sqrt{n}$  qui **grandit**. Seule la **moyenne** (ou la fréquence) se stabilise.
- Elle ne garantit rien pour un  $n$  précis : il s'agit d'une probabilité qui tend vers 0, pas d'une certitude.

## 9. Simulations avec Python

La loi des grands nombres se visualise très bien avec un programme. Le module **random** fournit `random.randint(a, b)` (entier aléatoire entre a et b inclus) et `random.random()` (nombre aléatoire de [0 ; 1[).

### Moyenne de n lancers de dé

```
import random

def moyenne(n):
    somme = 0
    for _ in range(n):
        somme = somme + random.randint(1, 6)
    return somme / n

print(moyenne(10))
print(moyenne(1000))
print(moyenne(100000))
```

Les trois affichages sont proches de 3,5, de plus en plus lorsque  $n$  grandit (les valeurs exactes changent à chaque exécution).

### Estimer une probabilité d'écart

Pour estimer  $P(|M_n - 3,5| \geq 0,1)$ , on répète l'expérience « calculer  $M_n$  » un grand nombre de fois et on compte la proportion d'écarts trop grands.

```
def proba_ecart(n, delta, essais):
    mauvais = 0
    for _ in range(essais):
```

```

        if abs(moyenne(n) - 3.5) >= delta:
            mauvais = mauvais + 1
        return mauvais / essais

print(proba_ecart(100, 0.1, 2000))
print(proba_ecart(1000, 0.1, 2000))

```

Les résultats seront nettement plus petits que les bornes de Bienaymé-Tchebychev (**7/24** pour  $n = 1000$ ) : la borne est valable, mais loin d'être optimale.

### Conseil de Mathbot

Avant de lancer une simulation, calcule la borne théorique : si ta simulation donne une proportion plus grande que la borne, c'est que ton programme (ou ton calcul de  $V$ ) contient une erreur !

## À retenir

- Linéarité :  $E(X + Y) = E(X) + E(Y)$  et  $E(aX + b) = aE(X) + b$ , toujours (sans indépendance).
- $V(aX + b) = a^2V(X)$  ; si  $X$  et  $Y$  sont indépendantes,  $V(X + Y) = V(X) + V(Y)$  (et  $V(X - Y) = V(X) + V(Y)$ ).
- Échantillon de taille  $n$  :  $n$  variables indépendantes de même loi  $(\mu, V, \sigma)$ .  $S_n$  : espérance  $n\mu$ , variance  $nV$ .  $M_n = S_n/n$  : espérance  $\mu$ , variance  $V/n$ , écart-type  $\sigma/\sqrt{n}$ .
- Markov :  $X \geq 0$  et  $a > 0 \Rightarrow P(X \geq a) \leq E(X)/a$ .
- Bienaymé-Tchebychev :  $P(|X - \mu| \geq \delta) \leq V/\delta^2$ , donc  $P(|X - \mu| < \delta) \geq 1 - V/\delta^2$ .
- Concentration :  $P(|M_n - \mu| \geq \delta) \leq V/(n\delta^2)$  ; pour un risque  $\alpha$ , il suffit de  $n \geq V/(\alpha\delta^2)$ .
- Loi des grands nombres : pour tout  $\delta > 0$ ,  $P(|M_n - \mu| \geq \delta)$  tend vers 0 quand  $n$  tend vers  $+\infty$  ; une fréquence se rapproche de la probabilité.
- La loi des grands nombres concerne la moyenne, pas la somme, et n'a rien d'une loi de compensation.