Chapitre 14
Loi des Grands Nombres

I. Introduction

À l'aide d'un ordinateur ou d'une calculatrice, on peut simuler le jet d'un dé à 6 faces équilibré. Dans l'expérience suivante, nous allons additionner les résultats en répétant l'expérience et diviser par le nombre de lancers, obtenant ainsi la moyenne des résultats. Pour cela, nous utilisons le logiciel Python en simulant successivement 10, 100, 1 000, 100 000 et enfin 1 000 000 de lancers !

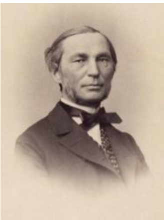
```
import random
def lan(n) :
    S=0
    for i in range(n) :
        S=S+random.randint(1,6)
    print(S/n)

lan(10)
lan(100)
lan(1000)
lan(100000)
lan(1000000)
```

```
3.2
3.56
3.53
3.49871
3.498358
```

On constate ainsi que les résultats se rapproche de la moyenne des résultats possibles

$$\frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3,5$$



Ce résultat n'a en fait rien de surprenant d'un point de vue mathématique et constitue une illustration de la **loi des grands nombres**. D'un point de vue historique, c'est Jacques Bernoulli (1654-1705) qui démontre le premier ce résultat dans le cadre de la loi binomiale. Il faut attendre 1867 pour que le mathématicien russe **Pafnouti Tchebychev** (1821-1894) démontre l'égalité du même nom et permette ainsi une démonstration plus générale de ce résultat.

II. Variables aléatoires réelles

1. Définition et premières propriétés

On considère un ensemble Ω muni d'une probabilité P. Une **variable aléatoire** X est une fonction de l'ensemble fondamental Ω à valeurs dans R.

Lorsque la variable X est à valeurs dans un ensemble discret de R (autrement dit, les valeurs de cet ensemble sont isolées dans R), on parle de **variable aléatoire discrète**.

Par la suite, on se place dans le cas discret. I est une partie de N et X désigne une variable aléatoire réelle tel que

$$X(\Omega) = \{x_i\}_{i\in I}$$

L'**espérance** d'une variable aléatoire (qui s'apparente à la moyenne statistique) est donnée par la formule

$$E(X) = \sum_{i\in I} x_i \cdot P(X = x_i)$$

Propriété 1
Soient a et b deux réels,
$E(aX + b) = a\cdot E(X) + b$

Démonstration

Posons $Y = aX + b$ et remarquons que $Y(\Omega) = \{ax_i + b\}_{i\in I}$

1er cas : $a = 0$

Y est constante égale à b donc $E(b) = b\cdot P(X = b) = b\times 1 = b$

2ème cas : $a\neq 0$

Remarquons dans ce cas que

$$\forall i\in I, P\left(Y = ax_i + b\right) = P\left(aX + b = ax_i + b\right) = P(X = x_i)$$

Ainsi,

$$E(Y) = \sum_{i\in I}\left(ax_i + b\right)\cdot P(Y = ax_i + b) = a\sum_{i\in I} x_i\cdot P(X = x_i) + b\underbrace{\sum_{i\in I} P(X = x_i)}_{=1} = a\cdot E(X) + b$$

La **variance** de la variable aléatoire X est donnée par la formule

$$V(X) = E\left((X - E(X))^2\right) = \sum_{i\in I} P(X = x_i)\cdot\left(x_i - E(X)\right)^2$$

L'**écart-type** d'une variable aléatoire X est donnée par la formule

$$\sigma(X) = \sqrt{V(X)}$$

Propriété 2
Soient a et b deux réels,

$$V(aX+b)=a^2\cdot V(X)$$
$$\sigma(aX+b)=|a|\sigma(X)$$

Démonstration

- Posons $Y=aX+b$ et remarquons que $Y(\Omega)=\{ax_i+b\}_{i\in I}$

1^er cas : $a=0$

Y est constante égale à b donc $Var(b)=P(X=b)\bullet(b-E(b))^2=1\times(b-b)^2=0$

2^ème cas : $a\neq0$

Remarquons dans ce cas que

$$\forall i\in I, P\left(Y=ax_i+b\right)=P\left(aX+b=ax_i+b\right)=P(X=x_i)$$

Ainsi,

$$Var(Y)=\sum_{i\in I}P(Y=ax_i+b)\cdot\left(ax_i+b-E(Y)\right)^2=\sum_{i\in I}P(X=x_i)\cdot\left(ax_i+b-aE(X)-b\right)^2$$

$$Var(Y)=\sum_{i\in I}P(X=x_i)\cdot\left(ax_i-aE(X)\right)^2=a^2\sum_{i\in I}P(X=x_i)\cdot\left(x_i-E(X)\right)^2=a^2\cdot Var(X)$$

- $\sigma(aX+b)=\sqrt{V(aX+b)}=\sqrt{a^2\cdot V(X)}=\sqrt{a^2}\sqrt{V(X)}=|a|\sigma(X)$

On démontrera au paragraphe suivant la propriété suivante qui permet de calculer plus simplement en pratique la variance de X

Propriété 3 (Formule de König-Huygens)

$$V(X)=E\left(X^2\right)-E^2(X)$$

Exercice

Une urne contient 3 sortes de boules de masses différentes : 7 boules de masse 1 kg, 5 boules de masse 3 kg et 3 boules de masse 5 kg. On tire au hasard une boule de l'urne et on note X sa masse.

1. *Déterminer la loi de la variable X.*

2. *Calculer l'espérance et la variance de X.*

Solution

1. Je détermine la loi de la variable X.

$$P(X=1)=\frac{7}{15}$$
$$P(X=3)=\frac{5}{15}$$
$$P(X=5)=\frac{3}{15}$$

2. Je calcule l'espérance et la variance de X.

$$E(X)=1\times\frac{7}{15}+3\times\frac{5}{15}+5\times\frac{3}{15}=\frac{37}{15}$$

$$V(X)=E\left(X^2\right)-E^2(X)=1^2\times\frac{7}{15}+3^2\times\frac{5}{15}+5^2\times\frac{3}{15}-\left(\frac{37}{15}\right)^2=\frac{127}{15}-\frac{1369}{225}=\frac{536}{225}$$

Remarque

On note souvent la loi d'une variable sous la forme d'un tableau.

x_i	**1**	**3**	**5**
$P(X=x_i)$	$\frac{7}{15}$	$\frac{5}{15}$	$\frac{3}{15}$

2. Opérations sur les variables aléatoires

On a vu dans le paragraphe précédent que, si $a\neq0$, la variable $Y=aX+b$ prenait les valeurs ax_i+b pour tout $i\in I$. On en a ensuite déduit que

$$E(Y)=E(aX+b)=aE(X)+b$$

Que se passe-t-il maintenant si l'on ajoute plusieurs variables aléatoires ?

Soient alors deux variables aléatoires X_1 et X_2 telles que $X_1(\Omega)=\{x_{1,i}\}_{i\in I}$ et $X_2(\Omega)=\{x_{2,j}\}_{i\in J}$ où I et J désignent des parties de N. La variable aléatoire X_1+X_2 est définie par

$$\left(X_1 + X_2\right)(\Omega) = \left\{x_{1,i} + x_{2,j}, (i,j)\in I\times J\right\}$$

Et

$$\forall s\in R, P\left(X_1 + X_2 = s\right) = \sum_{(i,j)\in I\times J\ x_{1,i}+x_{2,j}=s} P\left(X_1 = x_{1,i}\ et\ X_2 = x_{1,j}\right)$$

Si de plus X_1 et X_2 sont **indépendantes**,

$$\forall s\in R, P\left(X_1 + X_2 = s\right) = \sum_{(i,j)\in I\times J\ x_{1,i}+x_{2,j}=s} P\left(X_1 = x_{1,i}\right)\times P\left(X_2 = x_{1,j}\right)$$

Rappel

Deux variables aléatoires X_1 et X_2 sont **indépendantes** si, pour tous les intervalles A et B de R, l'égalité suivante est vérifiée

$$P\left(X_1\in A, X_2\in B\right) = P\left(X_1\in A\right)\cdot P\left(X_2\in B\right)$$

Dans le cas discret, si l'on considère les intervalles réduits à un seul nombre, on obtient l'égalité

$$\forall(x,y)\in R^2,\ P\left(X_1 = x, X_2 = y\right) = P\left(X_1 = x\right)\cdot P\left(X_2 = y\right)$$

Propriété 4

$$E\left(X_1 + X_2\right) = E\left(X_1\right) + E\left(X_2\right)$$

\- Admise –

Remarques

- Les deux variables X_1 et X_2 ne sont **pas nécessairement indépendantes**, contrairement à la propriété 5.
- À l'aide des propriétés 1 et 4, on peut montrer la propriété 3.

$$V(X) = E\left((X - E(X))^2\right) = E\left(X^2 - \underbrace{2E(X)}_{constante}X + \underbrace{E^2(X)}_{constante}\right) = E\left(X^2\right) - 2E(X)\times E(X) + E^2(X)$$

$$V(X) = E\left(X^2\right) - E^2(X)$$

Propriété 5

Soient deux variables X_1 et X_2 **indépendantes**,

$$V\left(X_1 + X_2\right) = V\left(X_1\right) + V\left(X_2\right)$$

\- Admise –

Exemple de somme de 2 variables aléatoires

On considère la variable aléatoire X_1 qui donne le résultat du lancer d'un dé à 6 faces équilibré et la variable aléatoire X_2 qui donne le résultat d'une pièce équilibrée marquée -1 sur une face et $+1$ sur l'autre. Ces deux variables sont évidemment **indépendantes**.

$X_1 + X_2$ peut prendre toutes les valeurs de 0 à 7

Par exemple,

$$P\left(X_1 + X_2 = 0\right) = P\left(X_1 = 1, X_2 = -1\right) = P\left(X_1 = 1\right)\times P\left(X_2 = -1\right) = \frac{1}{6}\times\frac{1}{2} = \frac{1}{12}$$

$$P\left(X_1 + X_2 = 1\right) = P\left(X_1 = 2, X_2 = -1\right) = P\left(X_1 = 2\right)\times P\left(X_2 = -1\right) = \frac{1}{6}\times\frac{1}{2} = \frac{1}{12}$$

$$P\left(X_1 + X_2 = 2\right) = P\left(X_1 = 3, X_2 = -1\right) + P\left(X_1 = 1, X_2 = +1\right) = \frac{1}{12} + \frac{1}{12} = \frac{2}{12}$$

…

On peut écrire la loi sous la forme d'un tableau.

k	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
$P\left(X_1 + X_2 = k\right)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Remarque

Les variables X_1 et X_2 sont uniformes mais pas leur somme $X_1 + X_2$

Par ailleurs,

$E(X_1) = 1\times\frac{1}{6} + \ldots + 6\times\frac{1}{6} = \frac{1+\ldots+6}{6} = \frac{21}{6} = 3,5$
$E(X_2) = -1\times\frac{1}{2} + 1\times\frac{1}{2} = 0$
Donc
$E(X_1 + X_2) = E(X_1) + E(X_2) = 3,5$

Remarque
On pourra vérifier que
$E(X_1 + X_2) = 0\times\frac{1}{12} + \ldots + 7\times\frac{1}{12} = \frac{0\times1+\ldots+7\times1}{12} = \frac{42}{12} = 3,5$

$V(X_1) = E(X_1^2) - E^2(X_1) = 1^2 \times \frac{1}{6} + \ldots + 6^2 \times \frac{1}{6} - \left(\frac{7}{2}\right)^2 = \frac{1+\ldots+36}{6} - \frac{49}{4} = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}$
$V(X_2) = E(X_2^2) - E^2(X_2) = 1\times1 - 0^2 = 1$
Donc
$V(X_1 + X_2) = V(X_1) + V(X_2) = \frac{47}{12}$
Remarque
On a gardé la même espérance mais augmenté la variance en ajoutant X_2 à X_1.

On peut généraliser les résultats précédents à n variables aléatoires (on le démontre par une simple récurrence…).

Propriété 6
Soient n variables $X_1, X_2, \ldots, X_n$
$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$$
Si de plus $X_1, X_2, \ldots, X_n$ sont **indépendantes deux à deux**,
$$V(X_1 + X_2 + \ldots + X_n) = V(X_1) + V(X_2) + \ldots + V(X_n)$$

3. Échantillon d'une loi de probabilité

Soit X une variable aléatoire de loi L. Soit $n\in N^*$, on appelle **échantillon de taille** n suivant loi L le vecteur aléatoire $(X_1, X_2, \ldots, X_n)$ où $X_1, X_2, \ldots, X_n$ sont des variables indépendantes et identiquement distribuées suivant la loi L.
On définit alors les variables aléatoires S_n et M_n respectivement **somme** et **moyenne** de l'échantillon $(X_1, X_2, \ldots, X_n)$ par
$$S_n = X_1 + X_2 + \ldots + X_n$$
Et
$$M_n = \frac{S_n}{n}$$
À l'aide des propriétés 1, 2, 4 et 5, on tire les propriétés suivantes.

Propriété 7
Soit $n\in N^*$,
$E(S_n) = nE(X)$ $V(S_n) = nV(X)$ $\sigma(S_n) = \sqrt{n}\sigma(X)$ $E(M_n) = E(X)$

Exemple : schéma de Bernoulli et loi binomiale
Soit $p\in[0; 1]$, X suit une loi de **Bernoulli de paramètre** p (le succès $X = 1$ possède une probabilité de p)
$$X\sim B(p)$$
$P(X = 0) = 1 - p$ et $P(X = 1) = p$
$(X_1, X_2, \ldots, X_n)$ est un **schéma de Bernoulli**. S_n suit alors la loi **binomiale de paramètres** n **et** p (elle compte le

nombre de succès au bout de n tirages)

$$S_n \sim B(n,p)$$

$E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = 0 \times (1 - p) + 1 \times p = p$

$V(X) = E\left(X^2\right) - E^2(X) = 0^2 \times P(X = 0) + 1^2 \times P(X = 1) - p^2 = p - p^2 = p(1 - p)$

$\sigma(X) = \sqrt{V(X)} = \sqrt{p(1 - p)}$

On en déduit que

$E\left(S_n\right) = nE(X) = np$

$V\left(S_n\right) = nV(X) = np(1 - p)$

$\sigma\left(S_n\right) = \sqrt{n}\sigma(X) = \sqrt{np(1 - p)}$

Méthode

Calculer l'espérance, la variance et l'écart-type d'une variable aléatoire moyenne

🖥️ **Vidéo https://youtu.be/o67OOavrbHQ**

On considère la variable aléatoire X qui prend, de façon équiprobable, les valeurs –4, 0, 1, 3 et 6.

M_{50} est la variable aléatoire moyenne d'un échantillon de taille 50 de la loi de X.

Calculer l'espérance, la variance et l'écart type de M_{50}.

Par équiprobabilité, on établit le tableau de la loi de probabilité de X.

k	-4	0	1	3	6
$P(X = k)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

On a ainsi

$E(X) = \frac{1}{5} \times (-4) + \frac{1}{5} \times 0 + \frac{1}{5} \times 1 + \frac{1}{5} \times 3 + \frac{1}{5} \times 6 = 1,2$

$V(X) = \frac{1}{5} \times (-4 - 1,2)^2 + \frac{1}{5} \times (0 - 1,2)^2 + \frac{1}{5} \times (1 - 1,2)^2 + \frac{1}{5} \times (3 - 1,2)^2 + \frac{1}{5} \times (6 - 1,2)^2 = 10,96$

$\sigma(X) = \sqrt{10,96} \approx 3,31$

On en déduit que

$E\left(M_{50}\right) = E(X) = 1,2$

$V\left(M_{50}\right) = \frac{1}{50}V(X) = \frac{10,96}{50} = 0,2192$

$\sigma\left(M_{50}\right) = \frac{1}{\sqrt{50}}\sigma(X) \approx \frac{3,31}{\sqrt{50}} \approx 0,468$

III. Inégalité de Tchebychev et loi des grands nombres

L'inégalité de Bienaymé-Tchebychev permet de préciser l'écart à la moyenne d'une variable aléatoire.

Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire.

$$\forall \delta \in R, \delta > 0 \Longrightarrow P(|X - E(X)| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

Démonstration

Soit X une variable aléatoire. On reprend les notations de la partie II.

$$X(\Omega) = \left\{x_i\right\}_{i \in I}$$

Soit un réel $\delta > 0$, on considère l'événement $A = \{|X - E(X)| \geq \delta\}$

Il existe donc une partie $K \subset I$ telle que $A = \left\{x_i\right\}_{i \in K}$

Or,

$$\forall i \in I,\ P\left(X = x_i\right) \bullet \left(x_i - E(X)\right)^2 \geq 0$$

Donc,

$$V(X) = \sum_{i \in I} P(X = x_i) \cdot \left(x_i - E(X)\right)^2 \geq \sum_{i \in K} P(X = x_i) \cdot \left(x_i - E(X)\right)^2$$

Or, par définition de l'événement A,

$$\forall i \in K,\ P\left(X = x_i\right) \bullet \left(x_i - E(X)\right)^2 \geq \delta^2 P\left(X = x_i\right)$$

Donc, après factorisation par δ^2,

$$V(X) \geq \delta^2 \sum_{i \in K} P(X = x_i)$$

Or,

$$\sum_{i \in K} P(X = x_i) = P(X \in A) = P(|X - E(X)| \geq \delta)$$

On en déduit que

$$P(|X - E(X)| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

Méthode

Appliquer l'inégalité de Bienaymé-Tchebychev

🖥️ **Vidéo https://youtu.be/4XMvq1FnYwU**

Soit une variable aléatoire X qui suit la loi binomiale de paramètres $n = 20$ et $p = 0,1$.

$$X \sim B(20; 0,1)$$

1. Appliquer l'inégalité de Bienaymé-Tchebychev avec $\delta = 2\sigma(X)$. Interpréter.

2. Recommencer avec $\delta = 3\sigma(X)$, puis $\delta = 4\sigma(X)$. Que constate-t-on ?

3. Simuler N valeurs de la variable aléatoire X par une fonction en Python dans le but d'estimer la probabilité $P(|X - 2| \geq 2\sigma(X))$. On testera le programme pour différentes valeurs de N. Au regard des résultats obtenus par le programme, peut-on penser que l'inégalité de Bienaymé-Tchebychev a un caractère optimal ?

Solution

1. $E(X) = 20 \times 0,1 = 2$ $\quad V(X) = 20 \times 0,1 \times 0,9 = 1,8$ $\quad \sigma(X) = \sqrt{1,8}$

Ainsi, on obtient $P(|X - E(X)| \geq 2\sigma(X)) \leq \frac{V(X)}{(2\sigma(X))^2} \Leftrightarrow P(|X - 2| \geq 2\sqrt{1,8}) \leq 0,25$

La probabilité que l'écart de X à $E(X)$ soit supérieur à $2\sigma(X)$ est majorée par 0,25.

2. $\delta = 3\sigma(X)$

$$P(|X - E(X)| \geq 3\sigma(X)) \leq \frac{V(X)}{(3\sigma(X))^2} \Leftrightarrow P(|X - 2| \geq 3\sqrt{1,8}) \leq \frac{1}{9}$$

$\delta = 4\sigma(X)$

$$P(|X - E(X)| \geq 4\sigma(X)) \leq \frac{V(X)}{(4\sigma(X))^2} \Leftrightarrow (|X - 2| \geq 4\sqrt{1,8}) \leq 0,0625$$

On peut en déduire que les écarts de X à $E(X)$ deviennent de plus en plus improbables.

3. On utilise le programme suivant

```python
import random as rd
import math

def simulX():
    a=0
    for expe in range(20):
        if rd.randint(1,100)<=10:
            a=a+1
    return a

def proba(N):
    echant=[simulX() for i in range(N)]
    c=0
    d=2*math.sqrt(1.8)
    for e in echant:
        if abs(e-2)>=d:
            c=c+1
    return c/N
```

```
>>> proba(1000)
0.038
>>> proba(10000)
0.0454
>>> proba(100000)
0.04178
>>> proba(100000)
0.04516
```

On constate qu'un écart à $E(X)$ supérieur à $2\sigma(X)$ est de probabilité souvent inférieure 0,05 (0,038 ; 0,0454 ; 0,04178 ; 0,04516) alors que l'inégalité de Bienaymé-Tchebychev nous donne pour cette même probabilité une majoration par 0,25. L'inégalité est donc loin d'être optimale.

Inégalité de concentration

Soit $n \in N^*$ et M_n la variable aléatoire moyenne d'un échantillon de taille n d'une variable aléatoire de moyenne μ et de variance V.

$$\forall \delta \in R, \delta > 0 \Longrightarrow P\left(\left|M_n - \mu\right| \geq \delta\right) \leq \frac{V}{n\delta^2}$$

Démonstration

On applique l'inégalité de Tchebychev à M_n. Soit un réel $\delta > 0$

$E\left(M_n\right) = \mu$ et $V\left(M_n\right) = \frac{1}{n}V$

Donc

$P\left(\left|M_n - E\left(M_n\right)\right| \geq \delta\right) = P\left(\left|M_n - \mu\right| \geq \delta\right) \leq \frac{V(M_n)}{\delta^2} \Leftrightarrow P\left(\left|M_n - \mu\right| \geq \delta\right) \leq \frac{V}{n\delta^2}$

Méthode

Appliquer l'inégalité de concentration pour déterminer la taille d'un échantillon

🖥️ **Vidéo https://youtu.be/7Nk9U-zwWOA**

Soit une variable aléatoire X qui suit la loi de Bernoulli de paramètre 0,2.

$$X \sim B(p)$$

On considère un échantillon de n variables aléatoires suivant la loi de X. On appelle M_n la variable aléatoire moyenne associée à cet échantillon. Déterminer la taille n de l'échantillon tel que la probabilité que la moyenne M_n appartienne à l'intervalle $]0,03 ; 0,37[$ soit supérieure à 0,95.

Solution

On cherche à calculer n tel que $P\left(0,03 < M_n < 0,37\right) \geq 0,95$

Dans l'idée d'appliquer l'inégalité de concentration, on fait apparaitre l'espérance de X dans l'inégalité.

Or,

$$E(X) = p = 0,2$$

Ainsi, on cherche n tel que

$P\left(0,03 - 0,2 < M_n - 0,2 < 0,37 - 0,2\right) \geq 0,95 \Leftrightarrow P\left(-0,17 < M_n - 0,2 < 0,17\right) \geq 0,95 \Leftrightarrow P\left(\left|M_n - 0,2\right| < 0,17\right) \geq$

Et donc, en considérant l'évènement contraire

$$1 - P\left(\left|M_n - 0,2\right| \geq 0,17\right) \geq 0,95 \Leftrightarrow P\left(\left|M_n - 0,2\right| \geq 0,17\right) \leq 0,05$$

En prenant $\delta = 0,17$ dans l'inégalité de concentration, on obtient

$$P\left(\left|M_n - E(X)\right| \geq \delta\right) \leq \frac{V(X)}{n\,\delta^2}$$

Or, $V(X) = p(1 - p) = 0,2 \times 0,8 = 0,16$

On cherche donc un entier n (non nul) tel que

$$\frac{0,16}{n \times 0,17^2} \leq 0,05 \Leftrightarrow n \geq \frac{0,16}{0,05 \times 0,17^2} \Leftrightarrow n \geq 111$$

Pour $n \geq 111$, la probabilité que la moyenne M_n appartienne à l'intervalle $]0,03 ; 0,37[$ est supérieure à 0,95.

Loi faible des grands nombres

Soit $n \in N$ et M_n la variable aléatoire moyenne d'un échantillon de taille n d'une variable aléatoire de moyenne μ et de variance V.

$$\forall \delta \in R, \delta > 0 \Longrightarrow P\left(\left|M_n - \mu\right| \geq \delta\right) = 0$$

Démonstration

On applique l'inégalité de concentration. Soit un réel $\delta > 0$

$$0 \leq P\left(\left|M_n - \mu\right| \geq \delta\right) \leq \frac{V}{n\delta^2}$$

Or,

$$\frac{V}{n\delta^2} = 0$$

Donc, d'après le théorème des gendarmes,

$$P\left(\left|M_n - \mu\right| \geq \delta\right) = 0$$

Méthode

Simuler des valeurs d'une variable aléatoire moyenne dans le but d'observer la loi des grands nombres

On considère la variable aléatoire X qui prend ses valeurs de manière équiprobable parmi les entiers 1 à 5. On nomme M_n la variable aléatoire moyenne d'un échantillon de taille n de la variable aléatoire X.

1. Simuler 500 valeurs de la variable aléatoire M_n par une fonction en Python dans le but d'estimer la probabilité $P(|M_n - E(X)| \geq \delta)$. Tester le programme pour différentes valeurs de δ et des valeurs de n de plus en plus grande.

2. Que constate-t-on ?

Solution

1. $E(X) = \frac{1}{5}\times 1 + \frac{1}{5}\times 2 + \frac{1}{5}\times 3 + \frac{1}{5}\times 4 + \frac{1}{5}\times 5 = 3$

Soit un réel $\delta > 0$,

$$P\left(\left|M_n - E(X)\right| \geq \delta\right) = P(|M_n - 3| \geq \delta)$$

```python
import random as rd
import math

def simulMn(n):
    S=[rd.randint(1,5) for i in range(n)]
    Mn=sum(S)/n
    return Mn

def echantMn(d,n):
    echant=[simulMn(n) for i in range(500)]
    c=0
    for e in echant:
        if abs(e-3)>=d:
            c=c+1
    return c/500
```

```
>>> echantMn(0.5,10)
0.304
>>> echantMn(0.5,100)
0.002
```

```
>>> echantMn(0.2,10)
0.712
>>> echantMn(0.2,100)
0.186
>>> echantMn(0.2,1000)
0.0
```

```
>>> echantMn(0.1,10)
0.924
>>> echantMn(0.1,100)
0.514
>>> echantMn(0.1,1000)
0.03
>>> echantMn(0.1,5000)
0.0
```

2. Plus la taille de l'échantillon est grande, plus l'écart entre la moyenne de cet échantillon et l'espérance de la variable aléatoire X est faible.