

Inégalités de Markov et de Bienaymé-Tchebychev

Plan du chapitre :

[I. Inégalité de Markov](#)

[II. Inégalité de Bienaymé-Tchebychev](#)

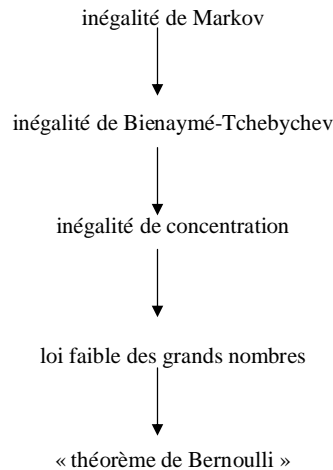
[III. Inégalité de concentration](#)

[IV. La loi faible des grands nombres](#)

[V. Échantillonnage et estimation](#)

[VI. Application au cas d'une loi binomiale](#)

Progression du chapitre



Dans tout le chapitre, on se place dans un espace probabilisé fini (Ω, P) .

I. Inégalité de Markov

1°) Propriété [inégalité de Markov]

Soit X une variable aléatoire réelle sur Ω à valeurs positives ou nulles.

Alors pour tout réel $a > 0$, on a $P(X \geq a) \leq \frac{E(X)}{a}$.

Cette inégalité a été nommée ainsi en l'honneur d'Andreï Markov. Andreï Andreïevitch Markov (1856-1922) est un mathématicien russe. Il est considéré comme le fondateur de la théorie des processus stochastiques.

Commentaires :

- La condition « X est à valeurs positives ou nulles » est essentielle.
- La loi de probabilité de X n'importe pas pour pouvoir appliquer cette inégalité. On sait que si X est une variable aléatoire à valeurs positives ou nulles, alors son espérance est aussi positive ou nulle.
- Si on applique l'inégalité de Markov pour un réel $a \leq E(X)$, alors $\frac{E(X)}{a} \geq 1$. On obtient donc une majoration de $P(X \geq a)$ par un réel supérieur ou égal à 1, ce qui ne présente pas d'intérêt puisque l'on sait qu'une probabilité est inférieure ou égale à 1. L'inégalité de Markov ne donne une majoration intéressante de $P(X \geq a)$ que lorsqu'on l'applique avec un réel $a > E(X)$.
- L'inégalité de Markov permet de démontrer que $P(X \geq a) \xrightarrow{a \rightarrow +\infty} 0$ (logique).

On applique pour cela le théorème des gendarmes car $0 \leq P(X \geq a) \leq \frac{E(X)}{a}$ pour tout réel $a > 0$ et

$$\frac{E(X)}{a} \xrightarrow{a \rightarrow +\infty} 0.$$

En fait, on a supposé que X est une variable aléatoire qui prend un nombre fini de valeurs (car l'univers de l'expérience aléatoire est fini) donc si a est supérieur à la plus grande valeur possible prise par X , alors $P(X \geq a) = 0$ (probabilité d'un événement impossible).

2°) Exemples

① Une usine produit en moyenne 35 pièces par semaine. On note X la variable aléatoire donnant le nombre de pièces produites par semaine. Que peut-on dire de la probabilité que l'usine produise plus de 70 pièces par semaine ?

D'après le contexte, on sait que X est à valeurs positives ou nulles et que $E(X) = 35$ (on assimile l'espérance à la moyenne).

D'après l'inégalité de Markov, $P(X \geq 70) \leq \frac{35}{70}$ soit $P(X \geq 70) \leq \frac{1}{2}$.

$\frac{1}{2}$ est un majorant de $P(X \geq 70)$.

② On lance n fois une pièce non truquée, n étant un entier naturel supérieur ou égal à 1. On note X la variable aléatoire qui compte le nombre de piles obtenus à l'issue des n lancers.

X suit la loi binomiale de paramètres n (nombre de répétitions) et $\frac{1}{2}$ (probabilité d'un succès).

D'après l'inégalité de Markov, $\forall \lambda \in \mathbb{R}_+^*$ $P(X \geq n\lambda) \leq \frac{E(X)}{n\lambda}$.

Or $E(X) = n \times \frac{1}{2} = \frac{n}{2}$ (formule de l'espérance d'une variable aléatoire qui suit une loi binomiale) donc

$$\frac{E(X)}{n\lambda} = \frac{1}{2\lambda}.$$

On en déduit que $\forall \lambda \in \mathbb{R}_+^*$ $P(X \geq n\lambda) \leq \frac{1}{2\lambda}$.

3°) Démonstration

On sait que $E(X) = \sum_{\omega \in \Omega} X(\omega) \times P(\{\omega\})$ (formule donnant l'espérance d'une variable aléatoire).

On note A l'événement constitué des éléments $\omega \in \Omega$ tels que $X(\omega) \geq a$.

Autrement dit, $A = \{X \geq a\}$.

Comme tous les termes de la somme sont positifs ou nuls, $E(X) \geq \sum_{\omega \in A} X(\omega) \times P(\{\omega\})$.

$$\text{On a } \sum_{\omega \in A} (X(\omega) \times P(\{\omega\})) \geq \sum_{\omega \in A} (a \times P(\{\omega\})).$$

$$\text{Or } \sum_{\omega \in A} (a \times P(\{\omega\})) = a \times \sum_{\omega \in A} P(\{\omega\}) = a \times P(A).$$

On a donc $E(X) \geq a \times P(A)$.

Comme a est strictement positif par hypothèse, cette dernière inégalité est équivalente à $P(A) \leq \frac{E(X)}{a}$ qui est l'inégalité de Markov.

II. Inégalité de Bienaymé-Tchebychev

On s'intéresse à l'écart entre une variable aléatoire X et son espérance.

1°) Théorème [inégalité de Bienaymé-Tchebychev]

Soit X une variable aléatoire réelle sur Ω .

Pour tout réel $a > 0$, on a : $P(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}$.

Irénée-Jules Bienaymé (1796-1878) est un mathématicien français qui a énoncé cette inégalité en 1853. C'est le mathématicien russe Pafnouti Tchebychev (1821-1894) qui l'a démontrée en 1867.

Bienaymé et Tchebychev (revue Tangente décembre 2020)

Pafnouti Lvovitch Tchebychev (1821-1894), le plus célèbre mathématicien russe du XIX^e siècle, ne se prénomait pas Bienaymé comme le pensent parfois certains étudiants. Si les noms de ces deux mathématiciens sont associés pour désigner une célèbre inégalité, ce n'est pas un hasard. Irénée-Jules Bienaymé (1796-1878) était un respectable inspecteur des finances lorsqu'il fut exclu de son poste en 1848 pour « manque de chaleur républicaine ». Il se tourne alors vers une brillante carrière de mathématicien qui le conduira à l'Académie des sciences. Il fait la connaissance de Tchebychev en octobre 1852 et une forte amitié s'installe entre les deux hommes. Le mathématicien russe séjourne plusieurs fois chez le savant français et celui-ci, féru de langue russe, traduit les écrits de son ami, ce qui permet à celui-ci de diffuser ses recherches. Dans le cadre d'un article pour défendre la méthode des moindres carrés de Laplace face aux critiques de Denis Poisson, Bienaymé énonce et démontre, en 1853, l'inégalité qui porte leurs deux noms. Tchebychev se rend compte de l'importance de ce résultat passé inaperçu par son ami français. Il la publie et l'utilise pour démontrer la loi des grands nombres dans un cadre général. La notoriété du Russe la fait connaître mais la véritable inégalité est sans doute qu'elle ne porte que le nom de Tchebychev, excepté dans la littérature française.

Autres notations :

On note m l'espérance de X et σ l'écart-type de X .

On a donc $E(X) = m$ et $V(X) = \sigma^2$.

L'inégalité de Bienaymé-Tchebychev s'écrit $P(|X - m| \geq a) \leq \frac{\sigma^2}{a^2}$.

Si on applique l'inégalité de Bienaymé-Tchebychev pour un réel $a \leq \sigma$, alors $\frac{\sigma}{a} \geq 1$ et par conséquent, $\frac{\sigma^2}{a^2} \geq 1$ (par élévation au carré des deux membres, sachant que tous les deux sont positifs).

On obtient donc une majoration de $P(|X - m| \geq a)$ par un réel supérieur ou égal à 1, ce qui ne présente pas d'intérêt puisque l'on sait qu'une probabilité est inférieure ou égale à 1.

L'inégalité de Bienaymé-Tchebychev ne donne une majoration intéressante de $P(|X - m| \geq a)$ que lorsqu'on l'applique avec un réel $a > \sigma$.

2°) Démonstration

On a l'équivalence fondamentale suivante : $|X - m| \geq a \Leftrightarrow |X - m|^2 \geq a^2$ (car les deux membres sont positifs ou nuls).

Or $Y = (X - m)^2$ est une variable aléatoire à valeurs positives ou nulles.

Donc par inégalité de Markov, $P(Y \geq a^2) \leq \frac{E(Y)}{a^2}$.

Or $E(Y) = V(X)$ d'où l'inégalité.

3°) Exemple

Le taux moyen de glycémie est de 1 g.L⁻¹ avec une variance de 0,1 (g.L⁻¹)².

Une personne présente un taux critique si son taux ne se situe pas dans l'intervalle]0,5 ; 1,5[.

Déterminer un majorant de la probabilité qu'une personne présente un taux critique.

On note X la variable aléatoire qui modélise la glycémie en $g.L^{-1}$.

On a $E(X) = 1$ et $V(X) = 0,1$.

L'intervalle $]0,5; 1,5[$ a pour centre 1 et pour rayon 0,5. On peut écrire $]0,5; 1,5[=]1 - 0,5; 1 + 0,5[$.

D'après l'inégalité de Bienaymé-Tchebychev, $P(|X - 1| \geq 0,5) \leq \frac{0,1}{0,5^2}$ soit $P(|X - 1| \geq 0,5) \leq 0,4$.

La probabilité qu'une personne présente un taux critique est inférieure ou égale à 0,4.

Exercice :

On lance 100 fois une pièce équilibrée.

Déterminer à l'aide de l'inégalité de Bienaymé-Tchebychev un minorant de la probabilité que le nombre de piles soit compris dans l'intervalle $I =]40; 60[$.

On note X la variable aléatoire qui compte le nombre de piles obtenus à l'issue des 100 lancers.

On cherche à majorer $P(X \in I)$.

X suit la loi binomiale de paramètres $n = 100$ (nombre de répétitions) et $p = \frac{1}{2}$ (probabilité d'obtenir pile en un lancer).

On calcule l'espérance et la variance de X .

$E(X) = 100 \times \frac{1}{2} = 50$ (formule de l'espérance d'une variable aléatoire qui suit une loi binomiale)

$V(X) = 100 \times \frac{1}{2} \times \frac{1}{2} = 25$ (formule de la variance d'une variable aléatoire qui suit une loi binomiale)

L'intervalle I a pour centre 50 et pour rayon 10.

Le fait que X appartienne à I se traduit par $|X - 50| < 10$.

Nous allons transformer en utilisant l'événement contraire.

Pour cela, on écrit : $P(|X - 50| < 10) = 1 - P(|X - 50| \geq 10)$.

D'après l'inégalité de Bienaymé-Tchebychev, $P(|X - 50| \geq 10) \leq \frac{25}{10^2}$ soit $P(|X - 50| \geq 10) \leq \frac{1}{4}$.

On multiplie les deux membres de cette dernière inégalité par -1 .

On obtient $-P(|X - 50| \geq 10) \geq -\frac{1}{4}$.

En ajoutant ensuite 1 à chaque membre de l'inégalité, on obtient $1 - P(|X - 50| \geq 10) \geq 1 - \frac{1}{4}$ ce qui donne

finalement $P(|X - 50| < 10) \geq \frac{3}{4}$ soit $P(X \in I) \geq \frac{3}{4}$.

La probabilité que le nombre de piles soit dans l'intervalle I est supérieure ou égale à 0,75.

Complément :

La probabilité que X soit compris dans I est égale à 0,943... (utilisation de la commande de la calculatrice permettant de calculer des probabilités avec une loi binomiale).

$$P(X \in I) = P(40 < X < 60) = P(41 \leq X \leq 59) = 0,9431120663590...$$

Pour effectuer le calcul sur calculatrice, on peut éventuellement écrire $P(X \in I) = P(X \leq 59) - P(X \leq 40)$.

4°) Situation particulière de l'inégalité de Bienaymé-Tchebychev

On considère une variable aléatoire X d'espérance m et d'écart type σ .

Nous allons majorer grâce à l'inégalité de Bienaymé Tchebychev les probabilité que $|X - m| \geq \sigma$,

$|X - m| \geq 2\sigma$, $|X - m| \geq 3\sigma$.

L'inégalité de Bienaymé-Tchebychev donne $P(|X - m| \geq \sigma) \leq \frac{\sigma^2}{\sigma^2}$ soit $P(|X - m| \geq \sigma) \leq 1$.

Cette inégalité ne présente pas d'intérêt car elle toujours vraie (une probabilité est toujours inférieure ou égale à 1).

L'inégalité de Bienaymé-Tchebychev donne $P(|X - m| \geq 2\sigma) \leq \frac{\sigma^2}{(2\sigma)^2}$ soit $P(|X - m| \geq 2\sigma) \leq \frac{1}{4}$.

Cette inégalité est déjà plus intéressante.

On peut en déduire que $P(|X - m| < 2\sigma) \geq \frac{3}{4}$ ou encore $P(X \in]m - 2\sigma; m + 2\sigma[) \geq \frac{3}{4}$.

Par conséquent, on a la minoration suivante $P(X \in [m - 2\sigma; m + 2\sigma]) \geq \frac{3}{4}$.

Cette minoration permet de mieux comprendre le rôle l'écart-type comme mesure de dispersion des valeurs autour de l'espérance.

L'intervalle $[m - 2\sigma; m + 2\sigma]$ joue un rôle important en probabilités.

L'inégalité de Bienaymé-Tchebychev donne $P(|X - m| \geq 3\sigma) \leq \frac{\sigma^2}{(3\sigma)^2}$ soit $P(|X - m| \geq 3\sigma) \leq \frac{1}{9}$.

Cette inégalité est déjà plus intéressante.

L'inégalité de Bienaymé-Tchebychev montre donc que des écarts de X à $E(X)$ de quelques σ deviennent improbables.

Le caractère universel de l'inégalité de Bienaymé-Tchebychev a pour contrepartie le fait qu'elle est loin d'être optimale. Par exemple, on a démontré que $P(|X - m| \geq 2\sigma) \leq \frac{1}{4}$ mais on constate par simulation que cette probabilité est souvent majorée par 0,05 (travail fait en première).

III. Inégalité de concentration

1°) Propriété [inégalité de concentration]

Soit (X_1, X_2, \dots, X_n) un échantillon de variables aléatoires d'espérance m et d'écart-type σ .

On pose $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ (moyenne de l'échantillon).

Alors, pour tout réel a strictement positif, $P(|M_n - m| \geq a) \leq \frac{\sigma^2}{na^2}$.

2°) Démonstration

On sait que $E(M_n) = m$ et $V(M_n) = \frac{\sigma^2}{n}$ (propriété vue dans le chapitre précédent).

On applique ensuite l'inégalité de Bienaymé-Tchebychev.

On obtient directement : $P(|M_n - m| \geq a) \leq \frac{\sigma^2}{na^2}$.

IV. La loi faible des grands nombres

1°) Théorème [loi faible des grands nombres]

Soit (X_1, X_2, \dots, X_n) un échantillon de variables aléatoires d'espérance m et d'écart-type σ .

On pose $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ (moyenne de l'échantillon).

Alors, pour tout réel a strictement positif, $P(|M_n - m| \geq a) \xrightarrow{n \rightarrow +\infty} 0$.

La loi faible des grands nombres illustre d'une certaine manière le fait que la moyenne de l'échantillon se rapproche de l'espérance des variables aléatoires quand la taille de l'échantillon « devient grande », comme cela a été vu en 1^{ère}.

Il existe une loi forte des grands nombres qui n'est pas au programme.

2°) Démonstration

On utilise l'inégalité de concentration déduite de l'inégalité de Bienaymé-Tchebychev.

Cette inégalité s'écrit : $P(|M_n - m| \geq a) \leq \frac{\sigma^2}{na^2}$.

On peut écrire $0 \leq P(|M_n - m| \geq a) \leq \frac{\sigma^2}{na^2}$.

Comme $\frac{\sigma^2}{na^2} \xrightarrow{n \rightarrow +\infty} 0$, par théorème des gendarmes, on en déduit que $P(|M_n - m| \geq a) \xrightarrow{n \rightarrow +\infty} 0$.

V. Échantillonnage et estimation

Le samedi 3 septembre 2022

T spé 22 h 37

Pour chaque échantillon, la fréquence observée est plus ou moins proche de p .

On peut évaluer la distance entre f et p en calculant $|f - p|$.

Attention au sens du mot échantillon (différente de celui d'échantillon d'une variable aléatoire).

1°) Intervalle de fluctuation d'une fréquence

Cette notion, historiquement, apparaît dans les premiers travaux de Jacques Bernoulli (1654-1703).

Après sa mort, on retrouva dans ses papiers des calculs extrêmement poussés qu'il n'avait osé publier de son vivant.

Exemple :

On lance 100 fois une pièce de monnaie équilibrée. Il est très peu probable d'obtenir 50 piles et 50 faces. En revanche, il est intéressant de déterminer quelles sont les valeurs les plus probables du nombre de piles. Il est intéressant de déterminer une « fourchette » dans laquelle va se trouver le nombre de piles dans la plupart des cas

On note X le nombre de piles obtenus à l'issue des 100 lancers.

On est dans le cas d'un schéma de Bernoulli.

La variable X suit donc la loi binomiale de paramètres $n = 100$ (nombre de répétitions) et $p = \frac{1}{2}$ (probabilité d'obtenir pile, puisque la pièce est équilibrée).

On peut démontrer que la probabilité que $X \in \llbracket 40 ; 60 \rrbracket$ (il s'agit de l'intervalle d'entiers allant de 40 à 60 c'est-à-dire de l'ensemble des entiers naturels de 40 à 60 inclus : 40, 41, 42, ..., 59, 60) avec une probabilité proche de 0,95 (la probabilité est égale à 0,964799...).

On dit que l'intervalle d'entiers $\llbracket 40 ; 60 \rrbracket$ est un intervalle de fluctuation du nombre de piles au seuil approximatif de 95 %.

Formule générale donnant un intervalle de fluctuation d'une fréquence

$$\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$$

Il s'agit d'un intervalle de fluctuation centré de centre p .

Les conditions d'application sont ...

Plus n est grand, plus l'intervalle est resserré.

Plus on lance, plus se concentre.

2°) Intervalle de fluctuation d'une moyenne

En 1^{ère}, on a étudié les problèmes d'échantillonnage de moyenne (« moyenne d'un échantillon ») en utilisant des simulations.

Propriété (admise sans démonstration)

On considère une population avec un caractère qualitatif C.
On note μ la moyenne et σ l'écart-type du caractère C dans la population.
On constitue N échantillons aléatoires de taille n.
Si m désigne la moyenne d'un échantillon, la proportion des cas où l'écart entre m et μ est inférieur ou égal à $\frac{2\sigma}{\sqrt{n}}$ est proche de 95 %.

Cela se traduit la proportion d'échantillons pour lesquels on a $|m - \mu| \leq \frac{2\sigma}{\sqrt{n}}$ (utilisation de la valeur absolue pour traduire l'écart) ou encore $m \in \left[\mu - \frac{2\sigma}{\sqrt{n}}; \mu + \frac{2\sigma}{\sqrt{n}} \right]$ est proche de 95 %.

On dit que l'intervalle $\left[\mu - \frac{2\sigma}{\sqrt{n}}; \mu + \frac{2\sigma}{\sqrt{n}} \right]$ est un intervalle de fluctuation de la moyenne au seuil approximatif de 95 % pour des échantillons aléatoires de taille n.

3°) Estimer une proportion

Dans de nombreuses situations, on est amené à estimer la proportion d'un caractère observé sur une population à partir d'un échantillon, **par exemple les sondages**. On dispose d'outils mathématiques permettant d'estimer la marge d'erreur de procédés.

Les estimations sont essentielles dans l'étude de populations, en sciences sociales ou en santé publique.

Lors de l'étude d'un caractère sur un échantillon prélever dans une population.

Un intervalle de confiance permet d'estimer entre quelles bornes la proportion réelle sur toute la population pourrait se trouver.

Constituer un échantillon pour estimer une proportion revient à étudier des sommes de variables aléatoires.

La taille n de cet échantillon amène à obtenir une suite de variables aléatoires sommes.

Plus la taille de l'échantillon est grande, plus l'estimation est précise : c'est la loi des grands nombres.

Certains échantillons « biaisés » ne reflètent pas la réalité de la population : des exemples existent, notamment lors de sondages électoraux.

Les instituts de sondages préfèrent la méthode des quotas afin de sélectionner des échantillons plus fidèles à la structure d'une population ; dans certains cas, cependant, on peut utiliser la méthode aléatoire.

Le lundi 30 mai 2022

En enseignement scientifique, en lien avec les populations, les élèves ont étudié :
intervalle de fluctuation
intervalle de confiance

fluctuare : mot latin lié aux bateaux (être battu par les flots) qui vient lui-même de fluctus (qui a donné les mots flots et flux en français)
fluctuation

Au premier trimestre, les notes d'un élève ont fluctué entre 0 et 5.

Fluctuations du cours de la bourse

Opinion fluctuante

Humeur fluctuante, états d'âme fluctuants

Il existe des méthodes pour déterminer des intervalles de fluctuation (par calcul ou autre).

Estimation d'une espérance

Si on fait 10 expériences et que l'on note T_1, T_2, \dots, T_{10} les nombres de réels qu'il a fallu choisir pour avoir une somme qui dépasse 1, $\frac{T_1 + T_2 + \dots + T_{10}}{10}$ donne une estimation du nombre moyen de nombres choisis.

Le 4-6-2022

Estimation d'une espérance

Loi des grands nombres

10. Utiliser un (bon) simulateur aléatoire pour simuler N réalisations $(x_i)_{0 \leq i < N}$ d'une variable aléatoire X revient en fait à simuler une réalisation d'un échantillon de N variables aléatoires $(X_i)_{0 \leq i < N}$ indépendantes de même loi que X.

11. D'après la Loi des grands nombres, si X est une variable d'espérance finie et si $(X_i)_{i \in \mathbb{N}}$

Est une suite de variables aléatoires indépendantes de même loi que X, alors la variable aléatoire $M_n = \sum_{0 \leq i < n} X_i$ converge presque sûrement vers la moyenne théorique $E(X)$ lorsque n tend vers $+\infty$.

11.1 Si l'entier N est assez grand, la moyenne des N réalisations simulées est donc assez proche de $E(X)$: $\frac{1}{N} \sum_{0 \leq i < N} x_i \approx E(X)$.

11.2 Si X est une variable aléatoire à valeurs dans N, alors la loi des grands nombres peut s'appliquer aux variables aléatoires de Bernoulli $[X_i=k]$ pour tout $k \in \mathbb{N}$, car il s'agit de variables aléatoires indépendantes, de même loi que $1[X=k]$.

On en déduit que, si l'entier N est assez grand, la fréquence empirique est assez proche de la probabilité théorique : $\frac{\# \{0 \leq i < N : x_i = k\}}{N} \approx P(X=k)$

12. De cette manière, la simulation d'un grand nombre de réalisations d'une variable aléatoire discrète X (au moyen d'un bon simulateur aléatoire) permet d'estimer la loi et l'espérance de cette variable.

12.1 Si la loi de X est connue, on peut ainsi tester la qualité du simulateur aléatoire.

12.2 Si la loi de X est difficile à calculer, on peut s'en faire une idée, soit en calculant des valeurs approchées des probabilités $P(X=k)$, soit en résumant les valeurs simulées sous la forme d'un histogramme.

Principe d'estimation d'une espérance à partir d'un échantillon

Soit (t_1, t_2, \dots, t_p) un échantillon de taille p de réalisations de la variable X.

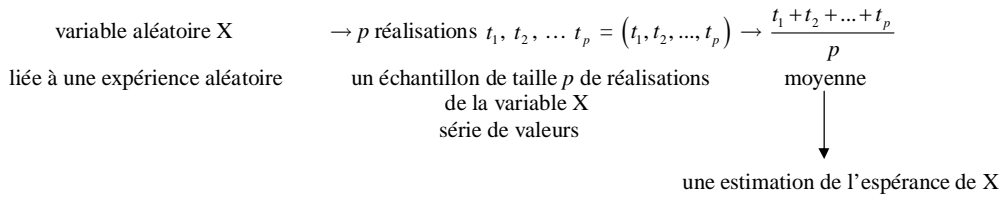
$\frac{t_1 + t_2 + \dots + t_p}{p}$ est une estimation de l'espérance de X.

Plus p est grand, plus cette estimation est précise.

Simulation

Programmes de simulation

Dans ce chapitre, nous allons reprendre l'interprétation de l'espérance d'une variable aléatoire.



Autrement dit, quand on calcule la moyenne des réalisations on obtient une valeur plus ou moins proche de l'espérance.

L'estimation est d'autant meilleure que p est grand.

Cela nous donne un lien entre probabilités et statistiques.

Principe de calcul d'une moyenne avec Python

Simulation d'expériences aléatoires

Commencer par importer la bibliothèque correspondante : `>>> import numpy.random as rd`

5.1 Simulations d'expériences

Pour simuler la majorité des expériences aléatoires étudiées dans votre programme, il suffit de connaître deux commandes :

— `rd.randint(a,b)`, avec a et b des entiers : choisit au hasard un entier entre a et b-1.

Comme pour la syntaxe range, l'entier b est exclu.

— `rd.random()` (ne rien mettre entre les parenthèses !) : choisit au hasard un réel entre 0 et 1

5.2 Échantillons de lois usuelles

Pour obtenir des simulations de lois usuelles discrètes, voici les commandes à connaître :

— N simulations de $B(n, p)$: `rd.binomial(n,p,N)`

— N simulations de $U([a, b])$: `rd.randint(a,b,N)`

— N simulations de $G(p)$: `rd.geometric(p,N)`

— N simulations de $P(\lambda)$: `rd.poisson(lambda,N)`

Enfin, pour obtenir N réels entre 0 et 1 (l'an prochain, cette loi sera appelée loi uniforme sur $[0,1[$):

`rd.random(N)`.

Pour faire des probabilités avec Python, on utilise la bibliothèque `numpy.random` que l'on importe à l'aide de la commande

`1 import numpy . random as rd`

1 La commande `rd.randint`

2 Simulation d'une épreuve de Bernoulli : commande `rd.binomial`

3 La commande `rd.random`

Deux syntaxes sont à connaître : `rd.random()` et `rd.randint(a,b)` cf. cours section 5.

La bibliothèque à importer est : `>>> import numpy.random as rd`

Le 31-5-2023

Fonction Python qui simule une réalisation d'une variable aléatoire X

sans argument

Fonction Python qui simule un échantillon de taille n de X prend pour argument un entier naturel $n \geq 1$ renvoie une liste de taille n de réalisations de X

boucle for // liste

Fonction Python qui simule une réalisation de la variable aléatoire moyenne

$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ où (X_1, X_2, \dots, X_n) est un échantillon de X

Modèle type pour un échantillon de taille n : `[X() for _ in range(n)]`

Voir document de Frédéric Junier

VI. Application au cas d'une loi binomiale

Dans ce paragraphe, nous allons donner une application historique importante de l'inégalité de Bienaymé-Tchebychev.

On considère une pièce truquée, telle que la probabilité d'obtenir pile en un lancer soit égale à p.

On effectue n lancers indépendants de cette pièce (n est un entier naturel non nul).

On note X_n la variable aléatoire égale au nombre de piles obtenus.

Alors $\forall \varepsilon \in \mathbb{R}_+^*$ $P\left(\left|\frac{X_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2}$ et $\forall \varepsilon \in \mathbb{R}_+^*$ $P\left(\left|\frac{X_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{4n\varepsilon^2}$.

C'est Jacques Bernoulli (1654-1705) qui publie l'une des premières versions de ce résultat dans son ouvrage posthume *Ars Conjectandi* en 1713. Il le démontre dans le cas particulier de la loi binomiale.

$\frac{X_n}{n}$ représente la fréquence de piles parmi les n lancers.

Démonstration :

X_n suit la loi binomiale de paramètres n (nombre d'épreuves) et p (probabilité d'obtenir pile en un lancer).

On pose $Y_n = \frac{X_n}{n}$.

On sait que l'espérance et la variance de X_n sont données par $E(X_n) = np$ et $V(X_n) = np(1-p)$.

$$E(Y_n) = \frac{E(X_n)}{n} = \frac{np}{n} = p$$

$$V(Y_n) = \frac{V(X_n)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

On applique l'inégalité de Bienaymé-Tchebychev à Y_n :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad P\left(\left|Y_n - E(Y_n)\right| \geq \varepsilon\right) \leq \frac{V(Y_n)}{\varepsilon^2} \text{ qui donne immédiatement}$$

$$\forall \varepsilon \in \mathbb{R}_+^* \quad P\left(\left|\frac{X_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2} \quad (1).$$

Nous avons la première majoration.

Pour obtenir la seconde majoration, il suffit d'observer que le produit $p(1-p)$ est maximal pour $p = \frac{1}{2}$ et le

maximum vaut $\frac{1}{4}$.

Il y a plusieurs moyens de justifier ce résultat. On peut, par exemple, considérer la fonction $f: x \mapsto x(1-x)$.

C'est une fonction polynôme du second degré.

On a donc $p(1-p) \leq \frac{1}{4}$.

$$(1) \text{ donne donc } \forall \varepsilon > 0 \quad P\left(\left|\frac{X_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{4n\varepsilon^2}.$$

Cette inégalité a une importance historique puisqu'elle apparaît dans les travaux de Jacques Bernoulli (meurt en 1705).

T spé

Exercices sur inégalités de Markov et de Bienaymé-Tchebychev

Inégalité de Markov

1 La moyenne d'une classe en mathématiques est de 12.

À l'aide de l'inégalité de Markov, majorer la probabilité qu'un élève choisi au hasard dans cette classe ait une moyenne supérieure à 14.

2 Dans un immeuble, l'ascenseur reste en moyenne 2 minutes au rez-de-chaussée avant d'être à nouveau sollicité.

À l'aide de l'inégalité de Markov, majorer la probabilité que l'ascenseur reste au rez-de-chaussée plus de 5 minutes.

3 La température moyenne aux Maldives est de 28,4°C. On suppose que la température n'est jamais négative. À l'aide de l'inégalité de Markov, majorer la probabilité que la température soit supérieure ou égale à 34°C un jour donné.

4 Sur une autoroute la vitesse moyenne des voitures est de 120 km·h⁻¹.

1°) À l'aide de l'inégalité de Markov, majorer la probabilité qu'un automobiliste roule à une vitesse supérieure ou égale à 150 km·h⁻¹.

2°) À l'aide de l'inégalité de Markov, minorer la probabilité qu'un automobiliste roule à une vitesse strictement inférieure à 130 km·h⁻¹.

Inégalité de Bienaymé-Tchebychev

5 On lance 3600 fois une pièce de monnaie non truquée.

1°) À l'aide de l'inégalité de Bienaymé-Tchebychev, minorer la probabilité que le nombre d'apparitions de « Pile » soit strictement compris entre 1600 et 2000.

2°) Calculer la valeur exacte de cette probabilité.

6 Sur les vingt matchs précédents, une équipe de rugby a marqué 60 essais. On note X la variable aléatoire donnant le nombre d'essais marqués au cours d'un match.

1°) Que vaut l'espérance de X ?

2°) On suppose que la variance de X est égale à 0,67.

Majorer la probabilité qu'au cours du prochain match l'écart entre le nombre d'essais marqués et la moyenne soit supérieur ou égal à 1.

3°) Minorer la probabilité que l'écart entre le nombre d'essais marqués et la moyenne soit strictement inférieur à 2.

7 Une classe présente les caractéristiques suivantes à un devoir : une moyenne de 12,4 et un écart-type de 1,2. Majorer la probabilité qu'un élève ait une note à ce devoir écartée d'au moins 2,5 points de la moyenne.

8 Dans une gare, le nombre moyen de passagers par jour est évalué à 5000 avec une variance de 2500.

Majorer la probabilité que l'écart entre le nombre de visiteurs enregistré lors d'une journée et la moyenne soit supérieur ou égal à 100.

9 Des clients estiment que leur temps d'attente en magasin varie trop fortement selon les périodes. Le gérant décide donc de vérifier cette impression. Il mesure que le temps d'attente moyen de ses clients est de douze minutes et que la probabilité qu'un client attende strictement entre neuf et quinze minutes est de 0,55. Que peut-il en déduire sur la valeur minimale de l'écart type du temps d'attente de ses clients ?

10 100 personnes jouent indépendamment à un même jeu dont la variable aléatoire associée au gain (en euros) a pour espérance 10 et pour variance 2.

Donner une minoration de la probabilité que la moyenne des gains de ces 100 personnes soit comprise strictement entre 7 euros et 13 euros.

11 On lance un dé tétraédrique équilibré dont les faces sont numérotées de 1 à 4.

On note chaque fois le numéro de la face sur laquelle repose le dé.

Combien de lancers doit-on effectuer pour être sûr au seuil de 99 % que la moyenne des résultats de ces lancers est comprise entre 2 et 3 exclus ?

12 Nicolas est facteur et distribue le courrier dans les logements d'un certain secteur géographique, ayant fait des études de mathématiques assez poussées, il a constaté que la probabilité d'un logement d'avoir du courrier un jour donné est égale à 0,6. Pour les besoins d'une enquête, Nicolas relève pendant 200 tournées supposées deux à deux indépendantes le nombre de logements pour lesquels il a distribué du courrier.

Majorer la probabilité que le nombre de personnes ayant reçu du courrier ne soit pas dans l'intervalle $I =]100; 140[$.

13 Dans une société de démarchage par téléphone, on estime que 40 % des personnes appelées répondent effectivement. Lorsque l'on appelle 1000 personnes, majorer la probabilité que le nombre de personnes répondant soit en dehors de l'intervalle $I =]300; 500[$.

14 On considère un jeu dont le gain algébrique en euro est donné par une variable aléatoire X de loi de probabilité suivante.

x_i	-2	-1	5
$P(X = x_i)$	0,8	0,15	0,05

1°) Calculer l'espérance et la variance de X .

2°) On considère 1000 personnes ayant joué à ce jeu durant une semaine.

Minorer la probabilité que le gain algébrique moyen sur ces 1000 personnes soit dans l'intervalle

$I =]-1,7; -1,3[$.

Inégalité de concentration

15 On considère une marche aléatoire sur \mathbb{Z} définie de la façon suivante : on part de 0 et, à chaque étape, indépendamment, on a une probabilité $\frac{1}{2}$ de faire un pas vers la droite et une probabilité $\frac{1}{2}$ de faire un pas vers la gauche.

Pour tout entier naturel i non nul, on note X_i la variable aléatoire dont la loi de probabilité est donnée par

$$P(X_i = 1) = P(X_i = -1) = \frac{1}{2}.$$

Pour tout entier naturel non nul n , on pose $S_n = X_1 + \dots + X_n$.

On représente une « trajectoire » par une ligne brisée joignant les points de coordonnées $(n, S_n(\omega))$.

Les valeurs de n sont portées en abscisse.

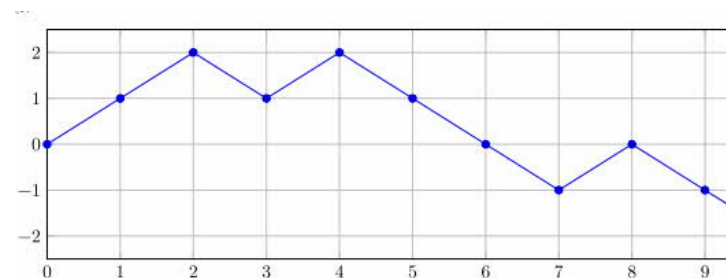
Voici un exemple de trajectoire :

i	0	1	2	3	4	5	6	7	8	9
Valeur de X_i	0	1	1	-1	1	-1	-1	-1	1	-1

X_0 prend initialement la valeur 0 puisque l'on part de 0.

On peut calculer S_0, S_1, S_2 etc.

On peut représenter la « trajectoire » sur un graphique par une ligne brisée.



1°) Que représente S_1 ? S_2 ? S_n pour n entier naturel quelconque non nul ?

Il s'agit de somme.

2°) Déterminer l'espérance et la variance X_i pour i entier naturel supérieur ou égal à 1.

En déduire l'espérance et la variance de S_n .

3°) Déterminer pour x réel strictement positif, une majoration de $P\left(\left|\frac{S_n}{n}\right| \geq x\right)$.

Loi faible des grands nombres et simulations

16 Pour une mise de 3 euros, un joueur peut lancer un dé cubique équilibré dont les faces sont numérotées de 1 à 6.

Avec un 6, il gagne 2 euros.

Avec un résultat compris 3 et 5 tous les deux inclus, il gagne 1 euro.

Avec un résultat inférieur ou égal à 2, il ne gagne rien.

Le joueur doit ensuite lancer une pièce équilibrée. Si la face visible est pile, il multiplie ses gains par 3 ; sinon il les multiplie par 4.

1°) Écrire un programme Python de simulation permettant de calculer la moyenne du gain pour un échantillon de taille n .

2°) Réaliser ce programme.

En déduire une estimation de l'espérance de gain, mise déduite.

Comparer avec la valeur exacte de l'espérance.

17 En présence de trois joueurs A, B, C, on lance trois dés cubiques équilibrés (un rouge, un bleu et un jaune) dont les faces sont numérotées de 1 à 6.

Le joueur A gagne 1 € si le résultat du dé rouge est inférieur à celui des deux autres dés.

Le joueur B gagne 1 € si le résultat du dé bleu est supérieur à celui des deux autres dés.

Le joueur C gagne 1 € si le résultat du dé jaune est compris strictement entre celui des deux autres dés.

Dans les autres cas, personne ne gagne.

1°) Écrire une fonction Python qui renvoie sous la forme d'une liste le montant en euros des trois joueurs après n lancers de dés.

2°) À l'aide d'un échantillon de 10000 parties, donner une estimation du gain moyen.

18 Une urne contient quatre boules rouges et une boule bleue. On réalise jusqu'à 10 tirages avec remise dans cette urne. Les tirages s'arrêtent dès que la boule bleue est tirée et on note G le nombre de boules rouges tirées.

1°) Recopier et compléter la fonction Python d'en-tête `def simu_G :` suivante afin de simuler la variable aléatoire G .

```
from random import randint

def simu_G() :
    g=0
    for tirage in range(10) :
        if randint(1,5)..... :
            return .....
        g=g+1
    return g
```

On décide de numéroté les boules : 1, 2, 3, 4 pour les boules rouges et 5 la boule bleue.

2°) Coder une fonction `ech` qui renvoie une liste de N réalisations de la variable aléatoire moyenne M_n de la loi de probabilité G .

3°) La probabilité de tirer la boule bleue en un lancer est $p = 0,2$ et la probabilité de tirer une boule rouge en un lancer est $q = 0,8$.

Vérifier que la loi de probabilité de G est donnée dans le tableau ci-dessous.

g_i	0	1	2	3	4	5	6	7	8	9	10
$P(G = g_i)$	p	pq	pq^2	pq^3	pq^4	pq^5	pq^6	pq^7	pq^8	pq^9	q^{10}

Calculer l'espérance de G .

Vérifier que les résultats obtenus par simulation sont proches de cette valeur.

19 On souhaite estimer le nombre de lancers nécessaires pour faire apparaître toutes les faces d'un dé cubique équilibré dont les faces sont numérotées de 1 à 6. Pour cela, on utilise l'algorithme suivant :

$a, b, c, d, e, f \leftarrow 0, 0, 0, 0, 0, 0$

$N \leftarrow 0$

Tantque $abcdef = 0$ **Faire**

$z \leftarrow$ entier aléatoire entre 1 et 6

Si $z = 1$

$a \leftarrow 1$

FinSi

Si $z = 2$

$b \leftarrow 1$

FinSi

Si $z = 3$

$c \leftarrow 1$

FinSi

Si $z = 4$

$d \leftarrow 1$

FinSi

Si $z = 5$

$e \leftarrow 1$

FinSi

Si $z = 6$

$f \leftarrow 1$

FinSi

$N \leftarrow N + 1$

FinTantque

1°) Expliquer le fonctionnement de cet algorithme.

2°) Modifier cet algorithme pour qu'il simule n fois l'expérience qui consiste à compter le nombre de lancers nécessaires pour que toutes les faces soient sorties une fois, et qui calcule ensuite le nombre moyen de lancers nécessaires.

3°) Programmer cet algorithme en Python. En déduire une estimation du nombre moyen de lancers nécessaires pour faire apparaître toutes les faces du dé.

4°) On peut démontrer que la valeur théorique est $6 \times \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} \right) = 14,7$.

Vérifier que les résultats obtenus par simulation sont proches de cette valeur.

C'est le problème du collectionneur.

20 On choisit au hasard réel quelconque entre 0 et 1, et on répète cela jusqu'à ce que la somme des nombres choisis dépasse 1.

Exemple avec des nombres décimaux (théoriquement, on choisit des réels, pas forcément décimaux) :

1^{er} nombre choisi : 0,372044

Comme il n'est pas supérieur à 1, on choisit un nouveau nombre au hasard entre 0 et 1.

2^e nombre choisi : 0,613

Comme la somme des deux premiers nombres choisis 0,372044 + 0,613 est inférieure ou égale à 1, on choisit un nouveau nombre au hasard entre 0 et 1.

3^e nombre choisi : 0,001007

Comme la somme des trois premiers nombres choisis 0,372044 + 0,613 + 0,001007 est inférieure ou égale à 1, on choisit un nouveau nombre au hasard entre 0 et 1.

4^e nombre choisi : 0,828

Cette fois, la somme des quatre premiers nombres choisis 0,372044 + 0,613 + 0,001007 + 0,828 dépasse 1.

On s'arrête là.

Dans cet exemple, il a fallu choisir 4 nombres au hasard entre 0 et 1 pour que la somme dépasse 1.

Conjecturer à l'aide d'un programme Python la valeur exacte du nombre moyen de nombres choisis.

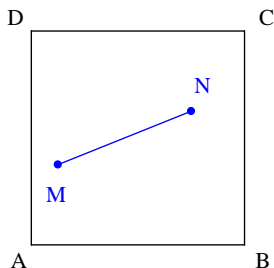
Ce problème est parfois désigné comme « paradoxe de la grenouille » ou « paradoxe des grenouilles » en référence à la formulation suivante du problème :

« Une grenouille veut rejoindre une mare située à un mètre devant elle. Pour l'atteindre, elle effectue des sauts supposés en ligne droite de longueur un nombre aléatoire de l'intervalle $[0; 1[$. »

21 Soit ABCD un carré de côté 1.

Soit M et N deux points au hasard situés à l'intérieur du carré.

On s'intéresse à la distance MN.



On peut noter d'emblée que la distance maximale de deux points quelconques situés à l'intérieur du carré est $\sqrt{2}$.

1°) Le but de cette question est de réaliser un programme Python permettant d'obtenir une estimation de la valeur moyenne de cette distance.

On se place dans un repère orthonormé (O, \vec{i}, \vec{j}) tel que O soit confondu avec A, $\overline{AB} = \vec{i}$ et $\overline{AD} = \vec{j}$.

a) Caractériser l'intérieur du carré ABCD par un système d'inéquations.

b) Rédiger le programme Python correspondant.

Rappel des deux fonctions de la bibliothèque random utilisables pour le programme :

- La fonction random renvoie un réel au hasard de l'intervalle $[0; 1[$;

- La fonction uniform(a, b) renvoie un réel au hasard dans l'intervalle $[a; b]$ (les arguments sont deux réels a et b tels que $a \leq b$).

On privilégiera la fonction uniform.

c) On peut démontrer que la distance moyenne est égale à $\frac{2 + \sqrt{2} + 5 \ln(\sqrt{2} + 1)}{15}$.

Comparer les estimations obtenues par simulation avec la valeur exacte.

3°) Faire le même travail pour deux points aléatoires :

- dans un cube d'arête 1 ;

- dans un disque de rayon 1.

22 On effectue une suite de lancers d'une pièce de monnaie. On suppose que les résultats des lancers sont indépendants et que, à chaque lancer, la pièce donne pile avec la probabilité p ($0 < p < 1$) et face avec la probabilité $q = 1 - p$.

1°) Réaliser un programme Python permettant d'estimer le nombre de lancers nécessaires pour obtenir deux piles de suite, c'est-à-dire lors de deux lancers consécutifs.

2°) Donner une estimation dans le cas où $p = \frac{1}{2}$ (cas d'une pièce équilibrée) et dans le cas où $p = \frac{1}{3}$.

3°) On peut démontrer que la valeur théorique est $\frac{1+p}{p^2}$.

Dans les deux cas, comparer les résultats obtenus par simulation avec la valeur théorique.

23 Soit A et B deux points quelconques du plan tels que $AB = 1$.

Soit M et N deux points au hasard appartenant au segment $[AB]$.

On s'intéresse à la distance MN.

Réaliser un programme Python permettant d'obtenir une estimation de la valeur moyenne de cette distance.

Conjecturer la valeur exacte de cette distance moyenne.

Propriété :

Une unité de longueur est fixée.

Soit Δ un axe de repère (O, \vec{i}) tel que $\|\vec{i}\| = 1$.

Soit M et N deux points de Δ d'abscisses respectives x et y.

La distance MN est égale à la valeur absolue de $x - y$.

24 On dispose d'une pièce non truquée. On étudie deux jeux.

Premier jeu :

On lance la pièce jusqu'à l'apparition du premier « Pile ».

On note X la variable aléatoire égale au nombre de fois où l'on a obtenu « Face ».

Rédiger une fonction Python permettant de simuler X.

Déterminer une estimation de l'espérance de X.

On admet que l'espérance de X est égale à $\frac{1}{p} - 1$.

Deuxième jeu :

On lance la pièce jusqu'à l'apparition du second « Pile ».

On note Y la variable aléatoire égale au nombre de fois où l'on a obtenu « Face ».

Rédiger une fonction Python permettant de simuler Y.

Déterminer une estimation de l'espérance de Y.

On admet que l'espérance de Y est égale à $\frac{2}{p} - 2$.

Reprendre les questions précédentes en supposant que la pièce est truquée.

25 On fait tourner n fois une roulette à deux secteurs, l'un marqué 1 et l'autre 0. La probabilité d'obtenir le secteur 1 est notée p .

Dans la suite de 1 et de 0 obtenus, on s'intéresse aux « blocs » de 0 et 1. Par exemple, la suite 10010110101 contient 9 blocs : quatre blocs de 0 et cinq blocs de 1.

À chaque suite de n chiffres obtenus, la variable aléatoire S associe le nombre de blocs de la suite.

Par exemple, pour la suite définie plus haut, S prend la valeur $4 + 5 = 9$.

Pour tout entier naturel i tel que $1 \leq i \leq n$, on appelle X_i la variable aléatoire qui, au i -ième tour de roulette, associe le numéro du secteur obtenu.

Pour tout entier naturel i tel que $2 \leq i \leq n$, on appelle Y_i la variable aléatoire telle que $Y_i = 1$ si X_{i-1} est différent de X_i et 0 sinon.

1°) a) Donner les valeurs de $P(X_i = 0)$ et de $P(X_i = 1)$.

b) Justifier que : $P(Y_i = 1) = P(X_{i-1} = 0) \times P(X_i = 1) + P(X_{i-1} = 1) \times P(X_i = 0)$.

c) Déterminer $P(Y_i = 1)$ en fonction de p .

2°) a) Exprimer S en fonction de Y_2, Y_3, \dots, Y_n puis en déduire l'espérance de S en fonction de n et p .

b) Quel nombre moyen de blocs obtient-on pour $n = 1001$ et $p = 0,5$?

3°) On donne le script incomplet d'une fonction Python qui retourne le nombre de blocs dans le cas $n = 1001$ et $p = 0,5$ de la suite définie à la ligne $L = [\dots \text{for } k \text{ in range}(1001)]$.

On crée plusieurs fonctions.

Première fonction : compte le nombre de blocs dans une liste constituée de 0 et de 1 à compléter :
La liste contient au moins deux éléments.

```
def blocs01(L):
    t=(L[1] !=L[0])
    for i in range(1, len(L)):
        t=t+...
    return (t+1)
```

Solution :

```
def blocs01(L):
    t=(L[1] !=L[0])
    for i in range(1, len(L)):
        t=t+(L[i] !=L[i-1])
    return t
```

blocs01([0,1,1,0,1,1,1])

4

blocs01([0,1,1])

2

blocs01([0,1,0,1])

4

blocs01([0,1,1,1,1])

2

Deuxième fonction : tirage d'un nombre 0 ou 1

```
from random import random

def tirage():
    r=random
    if random<p:
        x=1
    else:
        X=0
    return x
```

Programme final :

```
from random import randint

def simuS(N):
    L=[tirage() for k in range(N)]
    s=blocs01(N)
    return s
```

a) Recopier ce programme et compléter la ligne $L = [\dots \text{for } k \text{ in range}(1001)]$.

b) Quel est le rôle de la ligne $s = (L[1] != L[0])$?

c) Compléter la ligne $s = s + \dots$ à l'aide d'un booléen.

Il s'agit du « nombre moyen de séries » dans une suite de pile ou face.

26 Pierre Rémond de Montmort rapporte dans son *Essai d'analyse sur les jeux de hasard* (1708) un jeu observé par un explorateur revenant du Canada.

On y joue avec 8 noyaux noirs d'un côté et blanc de l'autre ; on jette les noyaux en l'air : alors, si les noyaux noirs se trouvent impairs, celui qui a jeté les noyaux gagne ce que l'autre a mis au jeu.

S'ils se trouvent tous noirs ou tous blancs, il en gagne le double ; et hors de ces deux cas, il perd sa mise. Deux amis A et B jouent.

Le joueur A lance les noyaux.

Solutions

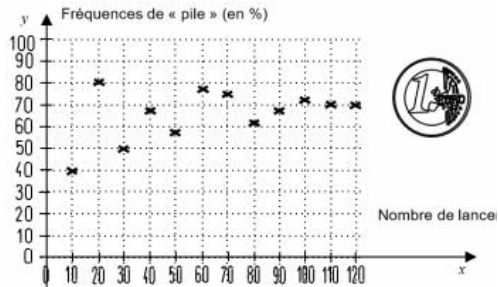
```
from random import random

def lancer():
    L=[]
    for i in range(8) :
        if random()<0.5:
            L.append('N')
        else:
            L.append('B')
    return L

def partie(mi se):
    gain=-mi se
    L=lancer()
    N=L.count('N')
    if N%2==1:
        gain=gain+mi se
    elif N==0 or N==8:
        gain=gain+2*mi se
    return
```

27 Les fréquences d'obtention de « Pile » lors de séries indépendantes de lancers d'une pièce « truquée » sont représentées sur le graphique ci-dessous.

Lorsque le nombre de lancers augmente, les fréquences se stabilisent.
Quelle est la probabilité d'obtenir « Pile » avec cette pièce ?



1

On note X la note d'un élève de la classe pris au hasard.

Comme la moyenne de la classe en mathématiques est 12, l'espérance de X est 12 (c'est-à-dire $E(X) = 12$).

X est une variable aléatoire dont les valeurs sont positives ou nulles.

D'après l'inégalité de Markov, on a $P(X \geq 14) \leq \frac{12}{14}$ soit $P(X \geq 14) \leq \frac{6}{7}$.

La calculatrice donne $\frac{6}{7} = 0,8571428\dots$

2

On note X le temps en minutes où l'ascenseur reste au rez-de-chaussée.

Comme l'ascenseur reste en moyenne 2 minutes au rez-de-chaussée avant d'être à nouveau sollicité, on en déduit que $E(X) = 2$.

X est une variable aléatoire dont les valeurs sont positives ou nulles.

D'après l'inégalité de Markov, on a $P(X \geq 5) \leq \frac{2}{5}$ soit $P(X \geq 5) \leq 0,4$.

3

On note X la température aux Maldives en °C.

Comme la température moyenne est de 28,4°C, on en déduit que $E(X) = 28,4$.

X est une variable aléatoire dont les valeurs sont positives ou nulles (vu la latitude des Maldives, on peut dire que la température y est toujours positive ou nulle).

D'après l'inégalité de Markov, on a $P(X \geq 34) \leq \frac{28,4}{34}$ soit $P(X \geq 34) \leq \frac{71}{85}$.

Il est possible de simplifier $\frac{28,4}{34}$ en $\frac{14,2}{17}$.

4

Soit X la vitesse d'un automobiliste en $\text{km} \cdot \text{h}^{-1}$.

Comme la vitesse moyenne des voitures est de $120 \text{ km} \cdot \text{h}^{-1}$, on en déduit que $E(X) = 120$.

X est une variable aléatoire dont les valeurs sont positives ou nulles.

1°)

D'après l'inégalité de Markov, $P(X \geq 150) \leq \frac{120}{150}$ soit $P(X \geq 150) \leq \frac{4}{5}$.

2°)

On écrit $P(X < 130) = 1 - P(X \geq 130)$.

D'après l'inégalité de Markov, on a $P(X \geq 130) \leq \frac{120}{130}$ soit $P(X \geq 130) \leq \frac{12}{13}$.

En multipliant les deux membres de cette inégalité par -1 , on obtient $-P(X \geq 130) \geq -\frac{12}{13}$.

En ajoutant ensuite 1 à chaque membre de cette inégalité, on obtient $1 - P(X \geq 130) \geq \frac{1}{13}$ soit

$$P(X < 130) \geq \frac{1}{13}.$$

5

1°) Soit X la variable aléatoire qui compte le nombre de piles obtenus à l'issue des 3600 lancers.

X suit la loi binomiale de paramètres $n = 3600$ (nombre de répétitions) et $p = \frac{1}{2}$ (probabilité d'un succès).

$$\text{On a } E(X) = 3600 \times \frac{1}{2} = 1800 \text{ et } V(X) = 3600 \times \frac{1}{2} \times \frac{1}{2} = 900.$$

On utilise les formules donnant l'espérance et la variance d'une variable aléatoire X qui suit la loi binomiale de paramètres n (nombre de répétitions) et p (probabilité d'un succès) rappelées dans l'encadré ci-dessous :

$$\begin{aligned} E(X) &= n \times p \\ V(X) &= n \times p \times (1 - p) \end{aligned}$$

On note A l'événement « Le nombre d'apparitions de piles est strictement compris entre 1600 et 2000 ».

$$A = (1600 < X < 2000) = (|X - 1800| < 200)$$

$$\text{On a } P(A) = 1 - P(|X - 1800| \geq 200).$$

D'après l'inégalité de Bienaymé-Tchebychev, $P(|X - 1800| \geq 200) \leq \frac{900}{200^2}$ soit

$$P(|X - 1800| \geq 200) \leq 0,0225.$$

$$\text{On a donc } -P(|X - 1800| \geq 200) \geq -0,0225.$$

En ajoutant 1 à chaque membre de cette inégalité, on obtient $P(A) \geq 0,9775$.

$$2^\circ) P(A) = P(1601 \leq X \leq 1999)$$

X suit la loi binomiale de paramètres $n = 3600$ (nombre de répétitions) et $p = \frac{1}{2}$ (probabilité d'un succès).

Avec la calculatrice, on obtient l'affichage 1.

Cela signifie que $P(A)$ est un nombre décimal très proche de 1.

6

$$1^\circ) E(X) = \frac{60}{20} = 3$$

$$2^\circ) V(X) = 0,67$$

D'après l'inégalité de Bienaymé-Tchebychev, on a $P(|X - 3| \geq 1) \leq \frac{0,67}{1^2}$ soit $P(|X - 3| \geq 1) \leq 0,67$.

3°)

$$\text{On a } P(|X - 3| < 2) = 1 - P(|X - 3| \geq 2).$$

D'après l'inégalité de Bienaymé-Tchebychev, on a $P(|X - 3| \geq 2) \leq \frac{0,67}{4}$.

$$\text{Donc } P(|X - 3| < 2) \geq 1 - \frac{0,67}{4} \text{ soit } P(|X - 3| < 2) \geq 0,8325.$$

7

Soit X la variable aléatoire égale à la note d'un élève au devoir.

$$\text{D'après l'énoncé, on a } E(X) = 12,4.$$

$$\sigma(X) = 1,2 \text{ donc } V(X) = [\sigma(X)]^2 = 1,2^2 = 1,44.$$

D'après l'inégalité de Bienaymé-Tchebychev, $P(|X - 12,4| \geq 2,5) \leq \frac{1,44}{2,5^2}$ soit

$$P(|X - 12,4| \geq 2,5) \leq 0,2304.$$

8

On note X le nombre de passagers par jour dans la gare.

$$\text{D'après l'énoncé, on a } E(X) = 5000 \text{ et } V(X) = 2500.$$

D'après l'inégalité de Bienaymé-Tchebychev, $P(|X - 5000| \geq 100) \leq \frac{2500}{100^2}$ soit $P(|X - 5000| \geq 100) \leq \frac{1}{4}$.

9

On note X la variable aléatoire qui compte le temps d'attente d'un client en minutes.

On a $E(X) = 12$.

On a $9 < X < 15 \Leftrightarrow |X - 12| < 3$.

On a $P(9 < X < 15) = 0,55$ donc $P(|X - 12| \geq 3) = 0,45$.

D'après l'inégalité de Bienaymé-Tchebychev, $P(|X - 12| \geq 3) \leq \frac{[\sigma(X)]^2}{3^2}$ donc

$$[\sigma(X)]^2 \geq 3^2 \times P(|X - 12| \geq 3)$$

On en déduit que $\sigma(X) \geq 3\sqrt{P(|X - 12| \geq 3)}$ soit $\sigma(X) \geq 3\sqrt{0,45}$

Avec la calculatrice, on obtient $3\sqrt{0,45} = 2,0124\dots$

10

On numérote les personnes de 1 à 100.

On note X_i le gain en euros de la personne numéro i .

$(X_1, X_2, \dots, X_{100})$ est un échantillon de variables aléatoires d'espérance 10 et de variance 2.

On note M la variable aléatoire moyenne de cet échantillon.

On cherche à minorer la probabilité que $7 < M < 13$ soit $|M - 10| < 3$ (on peut écrire $7 = 10 - 3$ et $13 = 10 + 3$).

On passe par l'événement contraire.

L'inégalité de concentration vue dans la propriété déduite de l'inégalité de Bienaymé-Tchebychev donne

$$P(|M - 10| \geq 3) \leq \frac{2}{100 \times 9} \text{ soit } P(|M - 10| \geq 3) \leq \frac{1}{450}.$$

On obtient $P(|M - 10| < 3) \geq \frac{449}{450}$.

11

On note n le nombre de lancers.

On note X_i la variable aléatoire donnant le résultat du lancer numéro i .

(X_1, X_2, \dots, X_n) est un échantillon de variables aléatoires d'espérance $\frac{5}{2}$ et de variance $\frac{5}{4}$.

On note M la variable aléatoire moyenne de cet échantillon.

On cherche n tel que $P(2 < M < 3) \geq 0,99$ (1).

$$(1) \Leftrightarrow P(|M - 2,5| < 0,5) \geq 0,99$$

$$\Leftrightarrow P(|M - 2,5| \geq 0,5) \leq 0,01$$

Or par l'inégalité de concentration déduite de l'inégalité de Bienaymé-Tchebychev,

$$P(|M - 2,5| \geq 0,5) \leq \frac{\frac{5}{4}}{n \times \frac{1}{4}} \text{ soit } P(|M - 2,5| \geq 0,5) \leq \frac{5}{n}.$$

Donc on est assuré que (1) est vérifiée dès lors que $\frac{5}{n} \leq 0,01$ soit $n \geq 500$.

On a effectué un raisonnement par condition suffisante.

Il suffit de choisir n supérieur ou égal à 500 pour garantir que (1) soit vérifiée.

12

On note X la variable aléatoire qui compte le nombre de logements pour lesquels Nicolas a déposé du courrier. X suit la loi binomiale de paramètres $n = 200$ (nombre de répétitions) et $p = 0,6$ (probabilité d'un succès).

On a $E(X) = 120$ et $V(X) = 48$.

On cherche à majorer la probabilité $P(X \notin I)$.

On observe que $100 = 120 - 20$ et $140 = 120 + 20$.

D'après l'inégalité de Bienaymé-Tchebychev, on a $P(X \notin I) = P(|X - 120| \geq 20) \leq \frac{48}{20^2}$ soit $P(X \notin I) \leq \frac{3}{25}$.

Il est possible de calculer $P(X \in I)$ et $P(X \notin I)$ avec la calculatrice.

Pour commencer, on peut observer que l'ensemble des entiers naturels appartenant à I est l'intervalle d'entiers $[101; 139]$.

Pour pouvoir utiliser la fonction de répartition de la loi binomiale, on écrit

$$P(X \in I) = P(X \leq 139) - P(X \leq 100).$$

On obtient l'affichage : 0,9952343369.

On en déduit que $P(X \notin I) = 0,004765663\dots$

Or on a $\frac{3}{25} = 0,12$ donc le résultat est bien en accord avec la majoration $P(X \notin I) \leq \frac{3}{25}$.

13

On note X la variable aléatoire qui compte le nombre de personnes répondant.

X suit la loi binomiale de paramètres $n = 1000$ (nombre de répétitions) et $p = 0,4$ (probabilité qu'une personne réponde).

On a $E(X) = 400$ et $V(X) = 240$.

On cherche la probabilité $P(X \notin I)$.

On observe que $300 = 400 - 100$ et $500 = 400 + 100$.

D'après l'inégalité de Bienaymé-Tchebychev, on a $P(X \notin I) = P(|X - 400| \geq 100) \leq \frac{240}{100^2}$ soit

$$P(X \notin I) \leq 0,024 \text{ (ou } P(X \notin I) \leq \frac{3}{125}).$$

Il est possible de calculer $P(X \in I)$ et $P(X \notin I)$ avec la calculatrice.

Pour commencer, on peut observer que l'ensemble des entiers naturels appartenant à I est l'intervalle d'entiers $[301; 499]$.

X suit la loi binomiale de paramètres $n = 1000$ et $p = 0,4$.

Pour pouvoir utiliser la fonction de répartition de la loi binomiale, on écrit

$$P(X \in I) = P(X \leq 499) - P(X \leq 300).$$

On obtient l'affichage : 0,999999999999.

Pour $P(X \notin I)$, on obtient l'affichage : $1,3298 \times 10^{-10}$.

Le résultat est donc bien en accord avec la majoration $P(X \notin I) \leq 0,024$.

14

1°)

$$\begin{aligned} E(X) &= -2 \times 0,8 - 1 \times 0,15 + 5 \times 0,05 \\ &= -1,5 \end{aligned}$$

$$\begin{aligned} V(X) &= (-2)^2 \times 0,8 + (-1)^2 \times 0,15 + 5^2 \times 0,05 - (-1,5)^2 \\ &= 2,35 \end{aligned}$$

2°) Les 1000 personnes constituent un échantillon $(X_1, X_2, \dots, X_{1000})$ de X .

On note M la moyenne de cet échantillon.

On cherche à minorer $P(M \in I)$.

On observe que $-1,7 = -1,5 - 0,2$ et $-1,3 = -1,5 + 0,2$.

$$M \in I \Leftrightarrow |M + 1,5| < 0,2$$

On utilise l'événement contraire $(M \notin I)$.

$$M \notin I \Leftrightarrow |M + 1,5| \geq 0,2$$

L'inégalité de concentration donne $P(M \notin I) \leq \frac{2,35}{1000 \times 0,2^2}$ soit $P(M \notin I) \leq \frac{2,35}{40}$ ce qui donne finalement

$$P(M \notin I) \leq 0,05875.$$

On en déduit facilement $P(M \in I) \geq 0,94125$.

15

On peut faire un graphique.

1°)

S_1 représente l'abscisse du point en lequel on se trouve après une étape.

S_2 représente l'abscisse du point en lequel on se trouve après deux étapes.

S_n représente l'abscisse du point en lequel on se trouve après n étapes.

2°)

x_i	-1	1
$P(X = x_i)$	$\frac{1}{2}$	$\frac{1}{2}$

$$E(X_i) = 0$$

$$V(X_i) = E((X_i - E(X_i))^2) = E((X_i - 0)^2) = E(X_i^2) = 1^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} = 1$$

$$V(X_i) = E(X_i^2) - [E(X_i)]^2 = 1^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} - 0^2 = 1$$

3°)

$$E(S_n) = \sum_{k=1}^{k=n} E(X_i) = 0$$

Les variables X_1, X_2, \dots, X_n sont deux à deux indépendantes [on dit que (X_1, X_2, \dots, X_n) est un échantillon de variables de même loi].

On peut donc appliquer la propriété sur la variance d'une somme.

$$V(S_n) = \sum_{k=1}^{k=n} V(X_i) = \sum_{k=1}^{k=n} 1 = n$$

$$4^\circ) \text{ On écrit } P\left(\left|\frac{S_n}{n}\right| \geq x\right) = P(|S_n| \geq nx) = P(|S_n - E(S_n)| \geq nx).$$

On applique ensuite l'inégalité de Bienaymé-Tchebychev.

$$P(|S_n - E(S_n)| \geq nx) \leq \frac{V(S_n)}{(nx)^2} \text{ soit } P(|S_n| \geq nx) \leq \frac{n}{n^2 x^2} \text{ ce qui donne finalement } P(|S_n| \geq nx) \leq \frac{1}{nx^2}.$$

16

1°)

```

from random import randint

def de() :
    r=randint(1 ;6)
    if r<=2
        return 0
    elif r>5 :
        return 2
    else :
        return 1

def piece() :
    s= randint(0,1)
    if g==0 :
        return 3
    else :
        return 9

def gain(n) :
    L= [pièce ()* dé () - 3 for i in range (n)]
    return sum(L)/n

```

2°)

On définit les événements suivants.

A : « obtenir un numéro inférieur ou égal à 2 »

B : « obtenir un numéro entre 3 et 5 »

C : « obtenir un 6 »

D : « obtenir pile »

E : « obtenir face »

On dresse un tableau de probabilités.

x_i	-3	0	1	3	5	
$P(X=x_i)$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{12}$	1

$$E(X) = -1 + \frac{1}{4} + \frac{1}{4} + \frac{5}{12}$$

$$= -\frac{1}{12}$$

17

```

from random import randint

def jeu() :
    a, b, c=0, 0, 0
    L=[randint(1,6) for i in range (2)]
    if L[1]<L[2] and L[1]<L[0]:
        a+=1
    if L[0]>L[1] and L[0]>L[2]:
        b+=1
    if L[2]<L[0] and L[2]>L[1] or L[2]<L[1] and L[2]>L[0]
        c+=1
    return a+b+c

def gain(n):
    G=[jeu() for i in range(n)]
    return sum(G)/n

```

18

1°)

```

from random import randint

def simu_G() :
    g=0
    for tirage in range(10) :
        if randint(1,5)..... :
            return .....
            g=g+1
    return g

```

Le programme suivant est faux.

```

from random import randint

def simu_G() :
    g=0
    while rand(1,5)!=5 :
        g=g+1
    return g

```

2°) Coder une fonction ech qui renvoie une liste de N réalisations de la variable aléatoire moyenne M_n de la loi de probabilité G.

3°)

$$\begin{aligned}
E(G) &= 0 \times p + 1 \times pq + 2pq^2 + 3pq^3 + \dots + 9pq^9 + 10q^{10} \\
&= p(q + 2q^2 + 3q^3 + \dots + 9q^9) + 10q^{10} \\
&= \frac{34868196}{9765625} \\
&= 3,5705032\dots
\end{aligned}$$

$$5(1 - 3 \times 0,8^{10}) = 3,38938\dots$$

19

20

On s'intéresse au nombre de nombres choisis.

Le problème peut s'implémenter en Python à l'aide de la fonction suivante :

```

from random import random

def moyenne(n):
    T = 0
    for k in range(n):
        s, i = 0, 0
        while s < 1:
            s += random()
            i += 1
        T += i
    return T / n

```

Autre proposition élève avec une liste le 25-5-2022 :

```

from random import uniform

def moyenne(n):
    L = []
    for i in range(n):
        S = 0
        N=0
        while S < 1:
            r=uniform(0, 1)
            S += r
            N+= 1
        L.append(N)
    return sum(L) / len(L)

```

```

On obtient alors :
>>> moyenne(10**7)
2.7186881
>>> moyenne(10**8)
2.71829368
>>> moyenne(10**9)
2.718286834

```

La moyenne semble être égale à e (nombre de Néper). Ce n'est qu'une conjecture.

On peut démontrer cette conjecture grâce à des outils qui dépassent le programme de terminale.

Ce problème s'appelle aussi « paradoxe des grenouilles » ou « paradoxe de la grenouille ».

21

Dans cet exercice, on cherche à estimer la distance moyenne entre deux points pris au hasard à l'intérieur d'un carré de côté 1.

1°)

$$a) \begin{cases} 0 \leq x \leq 1 \\ 0 \leq y \leq 1 \end{cases}$$

```

from random import random

def dist():
    # choisit 2 points au hasard et retourne la distance entre eux
    x1 = random()
    x2 = random()
    y1 = random()
    y2 = random()
    return ( (x1 - x2) ** 2 + (y1 - y2) ** 2 ) ** 0.5

f = 0
N = 10000
for k in range(1000):
    m = 0
    for k in range(N):
        m += dist()
    f += m/N
print(f / 1000)

```

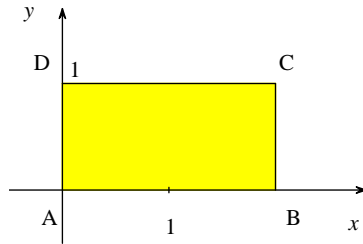
"0.5213943752132129" la première fois
"0.5208317781716778" une deuxième fois
"0.5213418328109504" une troisième fois

Je vais m'arrêter là car je vois que les trois valeurs ne sont pas éloignées du tout.

Je peux donc raisonnablement conclure qu'en moyenne, les deux points sont à une distance d'à peu près 0,521 l'un de l'autre.

Avec la calculatrice, $\frac{2 + \sqrt{2} + 5 \ln(\sqrt{2} + 1)}{15} = 0,521405433\dots$

3°) distance moyenne dans un cube d'arête 1 : 0,6620196846610... (valeur obtenue le 30-5-2022)
 Cas d'un rectangle de longueur 2 et de largeur 1 :



Programme Vincente Seixas le 30-5-2022 :

```
from math import
from random import uni form

def di stance2d(xa, xb, ya, yb):
    d=sqrt((xa-xb)**2+(ya-yb)**2)
    return d

d=0
for i in range(1000):
    r=di stance2d(uni form(0, 2), uni form(0, 1), uni form(0, 1))
    d=d+r
pri nt("moyenne: ", d/1000)
```

On trouve des valeurs autours de 0,79 – 0,81.

22

Le 30-5-2022

Proposition pour une probabilité du type $1/p$ (Vicente Seixas)

```
from random import random

def exo22(n, p):
    L=[]
    for i in range(n):
        x=0
        a=0
        b=randi nt(1, p)
        while a!=1 or b!=1:
            a=b
            b=randi nt(1, p)
            x+=1
        L.append(x)
    return sum(L)/len(L)
```

22

```
from random import uni form

def lancer(n, p):
    t=0
    compt=0
    tour=0
    for i in range(n):
        while compt!=2:
            res=uni form(0, 1)
            if res<=p:
                compt+=1
                tour+=1
            else:
                compt=0
                tour+=1

    compt=0
    t+=1
    return tour/n
```

Exercice 22 : Voir démonstration DS 2 15 octobre 2022 ECG 2^e année Anthony Mansuy

23

```
from math import abs
from random import uni form

def esti mdi st(n):
    d=0
    for i in range(n) :
        x,y= uni form(0, 1), uni form(0, 1)
        r=abs(x-y)
        d=d+r
    return d/n
```

24

Premier jeu :

```

from random import random

def simuX(p) :
    n=0
    while random()<p:
        n=n+1
    return n

```

Deuxième jeu :

```

from random import random

def simuY(p) :
    n=0
    while random()<p:
        n=n+1
    while random()<p:
        n=n+1
    return n

```

Autre proposition d'après EML 2018 corrigé Major-Prepa

```

from random import random

def simuY(p) :
    x=0
    n=0
    while n<2:
        if random()<p:
            n=n+1
        else :
            x=x+1
    return x

```

Source exercice 24 : Jérémy Legendre Concours blanc 14 mai 2022 exercice 2

Deux joueurs, A et B, décident de s'affronter dans un jeu de hasard.

• Le joueur A lance une pièce donnant PILE avec la probabilité

Sujets enregistrés le 11-7-2022 d'après sujets Ecricome 2014 voie économique

25

1°) a) Donner les valeurs de $P(X_i = 0)$ et de $P(X_i = 1)$.

On peut faire un tableau donnant la loi de probabilité.

X_i suit la loi binomiale de paramètre p .

b) Justifier que : $P(Y_i = 1) = P(X_{i-1} = 0) \times P(X_i = 1) + P(X_{i-1} = 1) \times P(X_i = 0)$.

c) Déterminer $P(Y_i = 1)$ en fonction de p .

$$P(Y_i = 1) = (1-p) \times p + p \times (1-p)$$

$$P(Y_i = 1) = 2p(1-p)$$

3°)

```

from random import randint

def b1ocs01():
    L=[rand(0, 1) for k in range(1001)]
    s=(L[1]!=L[0])
    for i in range(1, 1000):
        s=s+(L[i]==L[i+1])
    return (1+s)

```