VARIABLES ALÉATOIRES

En 1654, *Blaise Pascal* (1623 ; 1662) entretient avec *Pierre de Fermat* (1601 ; 1665) des correspondances sur le thème des jeux de hasard et d'espérance de gain qui les mènent à exposer une théorie nouvelle : les calculs de probabilités.
Ils s'intéressent à la résolution de problèmes de dénombrement comme par exemple celui du *Chevalier de Méré* :
« Comment distribuer équitablement la mise à un jeu de hasard interrompu avant la fin ? »

I. Variable aléatoire et loi de probabilité

1) Variable aléatoire

Exemple :
Soit l'expérience aléatoire : "On lance un dé à six faces et on regarde le résultat."
L'ensemble de toutes les issues possibles $\Omega = \{1 ; 2 ; 3 ; 4 ; 5 ; 6\}$ s'appelle l'univers des possibles.
On considère l'événement A : "On obtient un résultat pair."
On a donc : $A = \{2 ; 4 ; 6\}$.
On considère l'événement élémentaire E : "On obtient un 3".
On a donc : $E = \{3\}$.

Définitions :
- Chaque résultat d'une expérience aléatoire s'appelle une issue.
- L'**univers des possibles** est l'ensemble des issues d'une expérience aléatoire.
- Un **événement** est un sous-ensemble de l'univers des possibles.
- Un **événement élémentaire** est un événement contenant une seule issue.

Exemple :
Dans l'expérience précédente, on considère le jeu suivant :
- Si le résultat est pair, on gagne 2€.
- Si le résultat est 1, on gagne 3€.
- Si le résultat est 3 ou 5, on perd 4€.

On a défini ainsi une variable aléatoire X sur $\Omega = \{1 ; 2 ; 3 ; 4 ; 5 ; 6\}$ qui peut prendre les valeurs 2, 3 ou –4.
On a donc : $X(1) = 3$, $X(2) = 2$, $X(3) = -4$, $X(4) = 2$, $X(5) = -4$, $X(6) = 2$

Définition : Une **variable aléatoire** X est une fonction définie sur un univers Ω et à valeur dans $\mathbb{R}$.

2) Loi de probabilité

Exemple :

Vidéo https://youtu.be/lBqkrg8pxQ4
Vidéo https://youtu.be/OnD_Ym95Px4

On considère la variable aléatoire X définie dans l'exemple précédent.
Chaque issue du lancer de dé est équiprobable et égale à $\frac{1}{6}$.
La probabilité que la variable aléatoire prenne la valeur 2 est égale à $\frac{3}{6}=\frac{1}{2}$.
On note : $P(X = 2) = \frac{1}{2}$.
De même : $P(X = 3) = \frac{1}{6}$ et $P(X = -4) = \frac{2}{6} = \frac{1}{3}$.

On peut résumer les résultats dans un tableau :

x_i	-4	2	3
$P(X = x_i)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

Ce tableau résume la loi de probabilité de la variable aléatoire X.

Définition : Soit une variable aléatoire X définie sur un univers Ω et prenant les valeurs $x_1, x_2, \ldots, x_n$.
La **loi de probabilité** de X associe à toute valeur x_i la probabilité $P(X = x_i)$.

Remarques :
- $P(X = x_i)$ peut se noter p_i.
- $p_1 + p_2 + \ldots + p_n = 1$

Exemple :
Dans l'exemple traité plus haut : $p_1 + p_2 + p_3 = \frac{1}{3}+\frac{1}{2}+\frac{1}{6} = 1$.

Méthode : Déterminer une loi de probabilité

Vidéo https://youtu.be/2Ge_4hclPnI

Soit l'expérience aléatoire : "On tire une carte dans un jeu de 32 cartes."
On considère le jeu suivant :
- Si on tire un cœur, on gagne 2 €.
- Si on tire un roi, on gagne 5 €.
- Si on tire une autre carte, on perd 1 €.

On appelle X la variable aléatoire qui à une carte tirée associe un gain ou une perte.
Déterminer la loi de probabilité de X.

La variable aléatoire X peut prendre les valeurs 2, 5, –1 mais aussi 7.

En effet, si on tire le roi de cœur, on gagne 5(roi) + 2(cœur) = 7 €.

- Si la carte tirée est un cœur (autre que le roi de cœur), X = 2.
P(X = 2) = $\frac{7}{32}$.
- Si la carte tirée est un roi (autre que le roi de cœur), X = 5.
P(X = 5) = $\frac{3}{32}$.
- Si la carte tirée est le roi de cœur, X = 7.
P(X = 7) = $\frac{1}{32}$.
- Si la carte tirée n'est ni un cœur, ni un roi, X = –1.
P(X = –1) = $\frac{21}{32}$.

La loi de probabilité de X est :

x_i	–1	2	5	7
P(X = x_i)	$\frac{21}{32}$	$\frac{7}{32}$	$\frac{3}{32}$	$\frac{1}{32}$

On constate que : $p_1 + p_2 + p_3 + p_4 = \frac{21}{32} + \frac{7}{32} + \frac{3}{32} + \frac{1}{32} = 1$

II. Espérance, variance, écart-type

Définitions : Soit une variable aléatoire X définie sur un univers Ω et prenant les valeurs $x_1, x_2, \ldots, x_n$.
La loi de probabilité de X associe à toute valeur x_i la probabilité p_i = P(X = x_i).

- L'**espérance mathématique** de la loi de probabilité de X est :
  $E(X) = p_1 x_1 + p_2 x_2 + \ldots + p_n x_n = \sum_{i=1}^{n} p_i x_i$

- La **variance** de la loi de probabilité de X est :
  $V(X) = p_1(x_1 - E(X))^2 + p_2(x_2 - E(X))^2 + \ldots + p_n(x_n - E(X))^2 = \sum_{i=1}^{n} p_i (x_i - E(X))^2$

- L'**écart-type** de la loi de probabilité de X est : $\sigma(X) = \sqrt{V(X)}$

Méthode : Calculer l'espérance, la variance et l'écart-type d'une loi de probabilité

Vidéo https://youtu.be/AcWVxHgtWp4
Vidéo https://youtu.be/elpgMDSU5t8

Dans le jeu de la "Méthode" du paragraphe précédent, calculer l'espérance, la variance et l'écart-type de la loi de probabilité de X et interpréter les résultats pour l'espérance et l'écart-type.

E(X) = $\frac{21}{32}\times(-1)+\frac{7}{32}\times 2+\frac{3}{32}\times 5+\frac{1}{32}\times 7=\frac{15}{32}$

V(X) = $\frac{21}{32}\times\left(-1-\frac{15}{32}\right)^2+\frac{7}{32}\times\left(2-\frac{15}{32}\right)^2+\frac{3}{32}\times\left(5-\frac{15}{32}\right)^2+\frac{1}{32}\times\left(7-\frac{15}{32}\right)^2\approx 5{,}1865$

$\sigma(X)\approx\sqrt{5{,}1865}\approx 2{,}28$

L'espérance est égale à $\frac{15}{32}\approx 0{,}5$ signifie qu'en jouant un grand nombre de fois à ce jeu, on peut espérer en moyenne gagner environ 0,50 €.
L'écart-type est environ égal à 2,28 signifie qu'avec une espérance proche de 0,50 le risque de perdre de l'argent est important.

Remarques :
- L'espérance est la moyenne de la série des x_i pondérés par les probabilités p_i.
En effet :
E(X) = $p_1 x_1 + p_2 x_2 + \ldots + p_n x_n = \frac{p_1x_1+p_2x_2+\cdots+p_nx_n}{1}=\frac{p_1x_1+p_2x_2+\cdots+p_nx_n}{p_1+p_2+\cdots+p_n}$

En répétant un grand nombre de fois l'expérience, la loi des grands nombres nous permet d'affirmer que les fréquences se rapprochent des probabilités théoriques.
La moyenne des résultats se rapprochent donc de l'espérance de la loi de probabilité.
L'espérance est donc la moyenne que l'on peut espérer si l'on répète l'expérience un grand nombre de fois.

- La variance (respectivement l'écart-type) est la variance (respectivement l'écart-type) de la série des x_i pondérés par les probabilités p_i.
L'écart-type est donc une caractéristique de dispersion "espérée" pour la loi de probabilité de la variable aléatoire.

Propriétés : Soit une variable aléatoire X définie sur un univers Ω.
Soit a et b deux nombres réels.
On a : E(aX+b) = aE(X)+b V(aX+b) = a^2V(X)

Démonstrations :

$$\begin{aligned}E(aX+b)&=\sum_{i=1}^{n}p_i(ax_i+b)\\&=a\sum_{i=1}^{n}p_ix_i+b\sum_{i=1}^{n}p_i\\&=a\sum_{i=1}^{n}p_ix_i+b\times 1\\&=aE(X)+b\end{aligned}$$

$$\begin{aligned}V(aX+b)&=\sum_{i=1}^{n}p_i\big(ax_i+b-(aE(X)+b)\big)^2\\&=\sum_{i=1}^{n}p_i\big(ax_i-aE(X)\big)^2\\&=a^2\sum_{i=1}^{n}p_i\big(x_i-E(X)\big)^2\\&=a^2V(X)\end{aligned}$$

Méthode : Simplifier les calculs d'espérance et de variance à l'aide d'une variable aléatoire de transition

Vidéo https://youtu.be/ljITvCBExVY

Une entreprise qui fabrique des roulements à bille fait une étude sur une gamme de billes produites. Le diamètre théorique doit être égal à 1,3 cm mais cette mesure peut être légèrement erronée.
L'expérience consiste à tirer au hasard une bille d'un lot de la production et à mesurer son diamètre.
On considère la variable aléatoire X qui, à une bille choisie au hasard, associe son diamètre.
La loi de probabilité de X est résumée dans le tableau suivant :

x_i	1,298	1,299	1,3	1,301	1,302
P($X = x_i$)	0,2	0,1	0,2	0,4	0,1

Calculer l'espérance et l'écart-type de la loi de probabilité de X.

Pour simplifier les calculs, on définit la variable aléatoire $Y = 1000X - 1300$.
La loi de probabilité de Y est alors :

x_i	–2	–1	0	1	2
P($Y = x_i$)	0,2	0,1	0,2	0,4	0,1

Calculons l'espérance et la variance de la loi de probabilité de Y :
$E(Y) = -2\text{x}0,2 + (-1)\text{x}0,1 + 1\text{x}0,4 + 2\text{x}0,1 = 0,1$
$V(Y) = 0,2\text{x}(-2 - 0,1)^2 + 0,1\text{x}(-1 - 0,1)^2 + 0,2\text{x}(0 - 0,1)^2 + 0,4\text{x}(1 - 0,1)^2$
$+ 0,1\text{x}(2 - 0,1)^2 = 1,69$

On en déduit l'espérance et la variance de la loi de probabilité de X :
$E(Y) = E(1000X - 1300) = 1000\ E(X) - 1300$
Donc : $E(X) = \frac{E(Y)+1300}{1000} = \frac{0,1+1300}{1000} = 1,3001$

$V(Y) = V(1000X - 1300) = 1000^2\ V(X)$
Donc : $V(X) = \frac{V(Y)}{1000^2} = \frac{1,69}{1000^2}$
Et donc : $\sigma(X) = \sqrt{\frac{1,69}{1000^2}} = \frac{1,3}{1000} = 0,0013$
Conclusion : $E(X) = 1,3001$ cm et $\sigma(X) = 0,0013$ cm.

III. Somme de variables aléatoires

Exemple :
On considère deux jeux dont les gains sont donnés :
- pour le premier jeu, par la variable aléatoire X qui prend les valeurs 1 et 2.
- pour le second jeu, par la variable aléatoire Y qui prend les valeurs –2, 3 et 4.
Par exemple, l'évènement $(X=1)\cap(Y=-2)$ signifie qu'on a gagné 1 € au premier jeu et perdu 2 € au deuxième jeu.
Considérons la variable aléatoire somme $X+Y$ donnant le gain total cumulé aux deux jeux.
Alors la variable aléatoire $X+Y$ peut prendre les valeurs :
-1, 0, 4, 5 et 6.
En effet, on a par exemple $X+Y=0$ avec $(X=2)\cap(Y=-2)$.

Par ailleurs, pour calculer par exemple, la probabilité de l'évènement $X+Y=5$, on pourra calculer : $P(X+Y=5)=P\big((X=1)\cap(Y=4)\big)+P\big((X=2)\cap(Y=3)\big)$
On cherche toutes les sommes $X+Y$ égales 5.

Si de plus, les évènements X et Y sont indépendants, alors on a :
$$P(X+Y=5)=P(X=1)P(Y=4)+P(X=2)P(Y=3)$$

Définition : Soit X et Y deux variables aléatoires. La **loi de probabilité** de la variable aléatoire somme $X+Y$ est donnée par :
$$P(X+Y=k)=\sum_{i+j=k}P((X=i)\cap(Y=j))$$
Si, de plus, les évènements $(X=i)$ et $(Y=j)$ sont indépendants, alors on a :
$$P(X+Y=k)=\sum_{i+j=k}P(X=i)\,P(Y=j)$$

Remarque : Le symbole Σ
Si par exemple, i et j sont des entiers naturels et $k=2$ alors :
$$P(X+Y=2)=\sum_{i+j=2}P((X=i)\cap(Y=j))$$
$$=P((X=0)\cap(Y=2))+P((X=1)\cap(Y=1))+P((X=2)\cap(Y=0))$$

Méthode : Déterminer la loi d'une somme de variables aléatoires

Vidéo https://youtu.be/0I7tz8oGh-s

On considère le jeu suivant qui se déroule en deux parties :
- La 1ère partie consiste à lancer une pièce de monnaie. Si on tombe sur « pile », on gagne 1 €, si on tombe sur « face », on gagne 2 €.
- La 2e partie consiste à lancer un dé à 6 faces. Si on tombe sur un chiffre pair, on gagne 1 €, si on tombe sur le « 3 » ou le « 5 », on gagne 2 €.
Si on tombe sur le « 1 », on perd 5 €.
La variable aléatoire X désigne les gains à la 1ère partie, la variable aléatoire Y désigne les gains à la 2e partie.

On considère que les variables aléatoires X et Y sont indépendantes.
Établir la loi de probabilité de la variable aléatoire somme $S = X + Y$ donnant le gain total cumulé à la fin des deux parties.

Dans le tableau ci-dessous, on présente toutes les sommes possibles :

Y \ X	1	2
-5	-4	-3
1	2	3
2	3	4

Ainsi, on a :
$P(S = -4) = P(X = 1)P(Y = -5)$ en effet, les variables X et Y sont indépendantes.
$= \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$

$P(S = -3) = P(X = 2)P(Y = -5)$
$= \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$

$P(S = 2) = P(X = 1)P(Y = 1)$
$= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

$P(S = 3) = P(X = 1)P(Y = 2) + P(X = 2)P(Y = 1)$
$= \frac{1}{2} \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{6} + \frac{1}{4} = \frac{5}{12}$

$P(S = 4) = P(X = 2)P(Y = 2)$
$= \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$

On peut présenter la loi de probabilité de S dans un tableau :

k	–4	–3	2	3	4
$P(S = k)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{5}{12}$	$\frac{1}{6}$

IV. Espérance et variance de combinaisons linéaires de variables aléatoires

1) Linéarité de l'espérance

Propriétés :
1) $E(aX + b) = aE(X) + b$ avec $a \in \mathbb{R}$ et $b \in \mathbb{R}$. 2) $E(X + Y) = E(X) + E(Y$

2) Variance

Propriété :
1) $V(aX+b)=a^2V(X)$ avec $a\in\mathbb{R}$ et $b\in\mathbb{R}$
2) Si X et Y sont deux variables aléatoires indépendantes : $V(X+Y)=V(X)+V(Y)$

Méthode : Simplifier les calculs d'espérance et de variance à l'aide d'une variable aléatoire de transition

Vidéo https://youtu.be/ljITvCBExVY

Une entreprise qui fabrique des roulements à bille fait une étude sur une gamme de billes produites. Le diamètre théorique doit être égal à 1,3 cm mais cette mesure peut être légèrement erronée.
L'expérience consiste à tirer au hasard une bille d'un lot de la production et à mesurer son diamètre.
On considère la variable aléatoire X qui, à une bille choisie au hasard, associe son diamètre.
La loi de probabilité de X est résumée dans le tableau suivant :

x_i	1,298	1,299	1,3	1,301	1,302
$P(X = x_i)$	0,2	0,1	0,2	0,4	0,1

Calculer l'espérance et l'écart-type de la loi de probabilité de X.

Pour simplifier les calculs, on définit la variable aléatoire $Y = 1000X - 1300$.
La loi de probabilité de Y est alors :

y_i	–2	–1	0	1	2
$P(Y = y_i)$	0,2	0,1	0,2	0,4	0,1

Calculons l'espérance et la variance de la loi de probabilité de Y :
$E(Y) = -2\text{x}0,2 + (-1)\text{x}0,1 + 1\text{x}0,4 + 2\text{x}0,1 = 0,1$
$V(Y) = 0,2\text{x}(-2-0,1)^2 + 0,1\text{x}(-1-0,1)^2 + 0,2\text{x}(0-0,1)^2 + 0,4\text{x}(1-0,1)^2$
$+ 0,1\text{x}(2-0,1)^2 = 1,69$

On en déduit l'espérance et la variance de la loi de probabilité de X :
$E(Y) = E(1000X - 1300) = 1000\,E(X) - 1300$
Donc : $E(X)=\dfrac{E(Y)+1300}{1000}=\dfrac{0,1+1300}{1000}=1,3001$
$V(Y) = V(1000X - 1300) = 1000^2\,V(X)$
Donc : $V(X)=\dfrac{V(Y)}{1000^2}=\dfrac{1,69}{1000^2}$
Et donc : $\sigma(X)=\sqrt{\dfrac{1,69}{1000^2}}=\dfrac{1,3}{1000}=0,0013$
Conclusion : $E(X) = 1,3001$ cm et $\sigma(X)=0,0013$ cm.

V. Application à la loi binomiale

1) Échantillon d'une loi de probabilité

Exemple :
On étudie la fiabilité d'un composant électronique. On appelle X la variable aléatoire égale à 1 si le composant électronique ne se détériore pas suite aux tests effectués et 0 dans le cas contraire.
Le fabricant précise que le composant électronique ne subit pas de détériorations suite aux tests dans 99,8 % des cas.
Dans ce cas, la variable aléatoire X suit la loi de Bernoulli de paramètre 0,998.
On effectue les tests sur un échantillon de 100 composants électroniques prélevés au hasard dans le stock du fabricant.
On peut considérer alors que la liste $(X_1, X_2, X_3, \ldots, X_{100})$ forment un échantillon de taille 100 de variables aléatoires suivant la loi de Bernoulli de paramètre 0,998.

Définition : Un **échantillon de taille n d'une loi de probabilité** est une liste de n variables aléatoires indépendantes suivant cette loi.

Propriétés : Soit $(X_1, X_2, \ldots, X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant une même loi. On pose : $S = X_1 + X_2 + \cdots + X_n$, alors on a :
1) $E(S) = E(X_1) + E(X_2) + \cdots + E(X_n)$
2) $V(S) = V(X_1) + V(X_2) + \cdots + V(X_n)$

Méthode : Calculer l'espérance et la variance d'une variable aléatoire en l'exprimant comme somme de variables aléatoires indépendantes.

Vidéo https://youtu.be/fRYVMQk3bQQ

Un industriel fabrique des plats en verre qui doivent résister à de fortes températures afin de pouvoir être utilisés dans un four de cuisson. Pour vérifier la résistance du plat, on le soumet à une température de 350°C. On a constaté qu'en moyenne, sur un grand nombre de plats testés sortant de l'usine, 1,5 % des plats ne supportent pas une telle température et cassent.

On choisit au hasard 200 plats produits par l'usine et on effectue, pour chacun d'eux, le test de résistance. On admet qu'étant donné le grand nombre de plats produits par l'usine, ce choix peut être assimilé à un tirage fait de façon indépendante et avec remise.
On désigne par R la variable aléatoire comptant le nombre de plats résistants au test.
Calculer $E(R)$ et $\sigma(R)$. Interpréter les résultats.

On appelle X la variable aléatoire égale à 1 si le plat résiste au test et égale à 0 dans le cas contraire.
X suit une loi de Bernoulli de paramètre $p = 0{,}985$. En effet, d'après l'énoncé, $E(X) = 1 - 0{,}015 = 0{,}985$ et $E(X) = p$ dans le cas d'une loi de Bernoulli.
Par conséquent, $V(X) = 0{,}985 \times (1 - 0{,}985) = 0{,}014775$.

On peut considérer alors que la liste $(X_1, X_2, X_3, \dots X_{200})$ forment un échantillon de taille 200 de variables aléatoires suivant la loi de Bernoulli de paramètre $0{,}985$.
Et on a :
$E(R) = E(X_1) + E(X_2) + \cdots + E(X_{200}) = 200 \times 0{,}985 = 197$
$V(R) = V(X_1) + V(X_2) + \cdots + V(X_{200}) = 200 \times 0{,}014775 \approx 2{,}995$
car $X_1, X_2, \dots X_{200}$ sont indépendants.
Soit : $\sigma(R) = \sqrt{2{,}995} \approx 1{,}72$.
En moyenne, sur un grand nombre d'échantillons de 200 plats, on peut espérer trouver 197 plats résistants avec un écart-type proche de 1,72.

2) Échantillon de la loi de Bernoulli

Propriété : Soit $(X_1, X_2, \dots X_n)$ un échantillon de taille n de la loi de Bernoulli de paramètre p.
La variable aléatoire $S = X_1 + X_2 + \cdots + X_n$ suit la loi binomiale de paramètres n et p.

Exemple :
En reprenant l'exemple donné au début du paragraphe *III.1.*, la variable aléatoire $S = X_1 + X_2 + \cdots + X_{100}$ suit la loi binomiale de paramètres $n = 100$ et $p = 0{,}998$.

3) Espérance, variance et écart type de la loi binomiale

Propriété : Soit S une variable aléatoire qui suit la loi binomiale de paramètres n et p.
$E(S) = np$ $\qquad V(S) = np(1-p)$ $\qquad \sigma(S) = \sqrt{V(S)}$

Démonstration :
Vidéo https://youtu.be/ljWJfGLRgJE

Soit $(X_1, X_2, \dots X_n)$ un échantillon de taille n de la loi de Bernoulli de paramètre p.
On rappelle que pour une variable aléatoire X qui suit une loi de Bernoulli, on a :
$E(X) = p$ et $V(X) = p(1-p)$.
Donc, on a :
$E(S) = E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n) = p + p + \cdots + p = np$
$V(S) = V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n)$ car $X_1, X_2, \dots X_n$ sont indépendants.
Donc : $V(S) = p(1-p) + p(1-p) + \cdots = p(1-p) = np(1-p)$

Méthode : Calculer l'espérance, la variance et l'écart-type pour une loi binomiale

Vidéo https://youtu.be/95t19fznDOU
Vidéo https://youtu.be/MvCZw9XIZ4Q

On lance 5 fois de suite un dé à six faces.
On considère comme succès le fait d'obtenir 5 ou 6.
On considère la variable aléatoire S donnant le nombre de succès.
Calculer $E(S)$, $V(S)$ et $\sigma(S)$.

La variable aléatoire S suit la loi binomiale de paramètres $p = \frac{1}{3}$ et $n = 5$.
$E(S) = 5 \times \frac{1}{3} = \frac{5}{3} \approx 1{,}7$
$V(S) = 5 \times \frac{1}{3} \times \frac{2}{3} = \frac{10}{9}$
$\sigma(S) = \sqrt{\frac{10}{9}} = \frac{\sqrt{10}}{3}$
En moyenne, on peut espérer obtenir environ 1,7 fois un 5 ou un 6, en 5 lancers.

VI. Moyenne d'un échantillon

1) Définition

Exemple :
On lance un dé à six faces et on considère la variable aléatoire X qui prend la valeur 1 si le dé s'arrête sur un chiffre pair et la valeur 0 sinon.
X suit donc une loi de Bernoulli de paramètre $\frac{1}{2}$.
On répète deux fois de suite cette expérience. On considère alors l'échantillon (X_1, X_2) de taille 2 de variables aléatoires X_1 et X_2 suivant la même loi que X.
Il est ainsi possible d'évaluer le résultat d'une telle expérience en étudiant la variable aléatoire moyenne de X_1 et X_2.
On appelle M_2 la variable aléatoire moyenne de l'échantillon (X_1, X_2).
Alors M_2 peut prendre les valeurs suivantes :

Valeur de X_1	Probabilité de X_1	Valeur de X_2	Probabilité de X_2	Probabilité de (X_1, X_2)	Valeur de M_2	Probabilité de M_2
0	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$	$\frac{0+0}{2} = 0$	$\frac{1}{4}$
0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$	$\frac{0+1}{2} = \frac{1}{2}$	$\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$
1	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$	$\frac{1+0}{2} = \frac{1}{2}$	
1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$	$\frac{1+1}{2} = 1$	$\frac{1}{4}$

Et on a ainsi la loi de probabilité de M_2 :

k	0	$\frac{1}{2}$	1
$P(M_2 = k)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Définition : Soit $(X_1, X_2, \dots, X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant une même loi.
La variable aléatoire moyenne M_n de l'échantillon est donnée par :

$$M_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n).$$

2) Propriétés

Exemple :
On reprend l'exemple précédent.
- Calculons l'espérance de M_2 :

$$E(M_2) = 0 \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{2} + 1 \times \frac{1}{4} = \frac{1}{2}$$

On retrouve l'espérance de la variable X.
On comprend intuitivement que l'espérance de la variable aléatoire moyenne d'un échantillon $(X_1, X_2, \dots, X_n)$ est égale à l'espérance de la variable aléatoire X associée à cet échantillon.

- Calculons la variance de M_2 :

$$V(M_2) = \frac{1}{4} \times \left(0 - \frac{1}{2}\right)^2 + \frac{1}{2} \times \left(\frac{1}{2} - \frac{1}{2}\right)^2 + \frac{1}{4} \times \left(1 - \frac{1}{2}\right)^2 = \frac{1}{8}$$

Alors que :

$$V(X) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Ainsi la variance de la variable aléatoire moyenne est plus faible que la variance de la variable d'origine.
De plus, la dispersion de la variable aléatoire moyenne diminue au fur et à mesure que la taille de l'échantillon n augmente.
En effet, si l'échantillon devient plus grand, le nombre de situations pouvant donner des valeurs proches de l'espérance augmente.
Ainsi, les valeurs prises par la moyenne deviennent de plus en plus probables dans un voisinage de l'espérance.

Propriété : Soit une variable aléatoire X et soit un échantillon $(X_1, X_2, \dots, X_n)$ de taille n de variables aléatoires indépendantes suivant la même loi que X.

$$E(M_n) = E(X) \qquad V(M_n) = \frac{1}{n} V(X) \qquad \sigma(M_n) = \frac{1}{\sqrt{n}} \sigma(X)$$

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 :
$$E(M_{50})=E(X)=1{,}2$$
$$V(M_{50})=\frac{1}{50}V(X)=\frac{10{,}96}{50}=0{,}2192$$
$$\sigma(M_{50})=\frac{1}{\sqrt{50}}\sigma(X)\approx\frac{3{,}31}{\sqrt{50}}\approx 0{,}468$$

VII. Inégalité de Bienaymé-Tchebychev

Propriété : Soit une variable aléatoire X. Pour tout réel strictement positif δ, on a :
$$P(|X-E(X)|\geq\delta)\leq\frac{V(X)}{\delta^2}$$

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 $p=0{,}1$.
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) a) 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.
   b) Au regard des résultats obtenus par le programme, peut-on penser que l'inégalité de Bienaymé-Tchebychev a un caractère optimal ?

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 :

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

Ou encore : $P(|X - 2| \geq 2\sqrt{1{,}8}) \leq 0{,}25$

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

2) – <u>pour $\delta = 3\sigma(X)$:</u>

$$P(|X - E(X)| \geq 3\sigma(X)) \leq \frac{V(X)}{(3\sigma(X))^2}$$

Ou encore : $P(|X - 2| \geq 3\sqrt{1{,}8}) \leq \frac{1}{9}$

– <u>pour $\delta = 4\sigma(X)$:</u>

$$P(|X - E(X)| \geq 4\sigma(X)) \leq \frac{V(X)}{(4\sigma(X))^2}$$

Ou encore : $P(|X - 2| \geq 4\sqrt{1{,}8}) \leq 0{,}0625$

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

3) a)

```
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
```

b) 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.

VIII. Inégalité de concentration

Propriété : Soit la variable aléatoire moyenne M_n d'un échantillon de taille n de la variable aléatoire X. Pour tout réel strictement positif δ, on a :

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

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.
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.

On cherche à calculer n tel que $P(0{,}03 < M_n < 0{,}37) \geq 0{,}95$
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 : $P(0{,}03 - 0{,}2 < M_n - 0{,}2 < 0{,}37 - 0{,}2) \geq 0{,}95$

Soit : $P(-0{,}17 < M_n - 0{,}2 < 0{,}17) \geq 0{,}95$

Soit encore : $P(|M_n - 0{,}2| < 0{,}17) \geq 0{,}95$

Et donc, en considérant l'évènement contraire :

$$1 - P(|M_n - 0{,}2| \geq 0{,}17) \geq 0{,}95$$
$$P(|M_n - 0{,}2| \geq 0{,}17) \leq 0{,}05$$

En prenant $\delta = 0{,}17$ dans l'inégalité de concentration, on a :

$$P(|M_n - E(X)| \geq \delta) \leq 0{,}05, \text{ avec } \frac{V(X)}{n\,\delta^2} = 0{,}05.$$

Or, $V(X) = p(1-p) = 0{,}2 \times 0{,}8 = 0{,}16$
On cherche donc un entier n tel que : $\dfrac{0{,}16}{n\,0{,}17^2} \leq 0{,}05$
Et donc : $n \geq \dfrac{0{,}16}{0{,}05 \times 0{,}17^2} \approx 110{,}7$
Pour $n \geq 111$, la probabilité que la moyenne M_n appartienne à l'intervalle $]0{,}03\,;0{,}37[$ est supérieure à 0,95.

IX. Loi des grands nombres

Propriété : Soit la variable aléatoire moyenne M_n d'un échantillon de taille n de la variable aléatoire X. Pour tout réel strictement positif δ, on a :

$$\lim_{n \to +\infty} P(|M_n - E(X)| \geq \delta) = 0$$

Remarque : La loi des grands nombres traduit le fait que plus la taille de l'échantillon d'une variable aléatoire X est grande, plus l'écart entre la moyenne de cet échantillon et l'espérance de la variable aléatoire X est faible.

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.
Et on nomme M_n la variable aléatoire moyenne d'un échantillon de taille n de la variable aléatoire X.
a) Simuler 500 valeurs de la variable aléatoire M_n par une fonction en Python dans le but d'estimer la probabilité $P(|M_n - E(X)| \geq \sigma(M_n))$.
Tester le programme pour des valeurs de n de plus en plus grande.
b) Que constate-t-on ?

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

$$V(X) = \frac{1}{5} \times (1-3)^2 + \frac{1}{5} \times (2-3)^2 + \frac{1}{5} \times (3-3)^2 + \frac{1}{5} \times (4-3)^2 + \frac{1}{5} \times (5-3)^2 = 2$$

$$\sigma(X) = \sqrt{2} \text{ donc } \sigma(M_n) = \frac{\sqrt{2}}{\sqrt{n}}$$

Ainsi : $P\big(|M_n - E(X)| \geq \sigma(M_n)\big) = P(|M_n - 3| \geq \frac{\sqrt{2}}{\sqrt{n}})$.

```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(n):
    echant=[simulMn(n) for i in range(500)]
    c=0
    d=math.sqrt(2/n)
    for e in echant:
        if abs(e-3)>=d:
            c=c+1
    return c/500
```

```
>>> echantMn(5)
0.31
>>> echantMn(5)
0.282
>>> echantMn(10)
0.036
>>> echantMn(20)
0.008
>>> echantMn(30)
0.002
>>> echantMn(50)
0.0
>>> echantMn(100)
0.0
>>> echantMn(1000)
0.0
```

b) 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.