Histoire 1

Au XVIII^{ième} siècle, l'Anglais Abraham Moivre et le Français Pierre Simon de Laplace ont étudié les distributions binomiales qui donnent par exemple la répartition des succès lorsqu'on répète plusieurs fois de façon identique une même expérience aléatoire à deux issues comme un lancer de pièce. Par exemple sur un échantillon de taille n, si on suppose que la probabilité de naissance d'un garçon est de $\dfrac{22}{21+22}$ (établie par mesure statistique), le nombre de naissances de garçons est une fonction dont la valeur dépend du hasard, on parle de variable aléatoire, et dont la loi suit une distribution binomiale de paramètres n et p. Par combinatoire, on peut déterminer les probabilités de tous les nombres de naissances de garçons possibles entre 0 et n. On peut en déduire la valeur moyenne et l'écart-type de la variable aléatoire comme pour une série statistique. À partir de ces indicateurs, Laplace a pu établir des intervalles de fluctuation du nombre de naissances de garçons et déterminer si la distribution sur une commune particulière était normale. Cette normalité, est attachée au théorème de Moivre-Laplace qui établit une convergence des distributions binomiales (après changement d'origine et d'échelle) vers la distribution en cloche d'une variable aléatoire universelle appelée loi normale centrée réduite. Plus généralement, cette universalité est atteinte lors de la superposition d'un grand nombre de causes aléatoires indépendante ce qui a permis à Gauss de développer une théorie de « l'incertitude de la mesure ».

1 Transformation affine de variables aléatoires

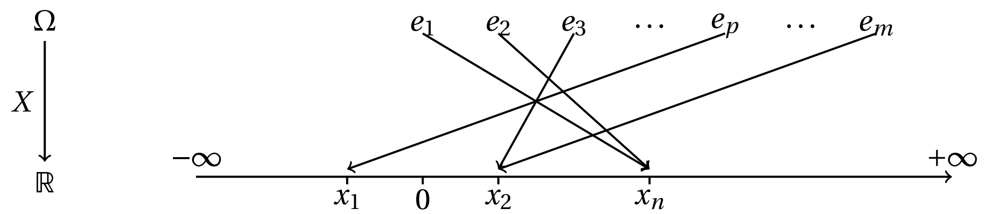
1.1 Rappels sur les variables aléatoires

Définition 1 ***Variable aléatoire et loi de probabilité***

Soit $\Omega = \{e_1, e_2, \cdots, e_p, \cdots, e_m\}$ **l'ensemble fini** décrivant l'univers d'une expérience aléatoire et $\mathbb{P}$ une loi de probabilité sur Ω.
A chaque issue on associe un nombre, on définit ainsi une fonction X de Ω dans $\mathbb{R}$ appelée **variable aléatoire**.
On note $X(\Omega) = \{x_1, x_2, \cdots, x_i, \cdots, x_n\}$ avec $n \leqslant m$ l'ensemble de ces nombres appelé **ensemble image** de X.

Définir la **loi de probabilité** de la variable aléatoire X, c'est associer à chaque valeur x_i le nombre $p_i = \mathbb{P}(X = x_i)$ qui est la probabilité de l'événement $\{X = x_i\}$ constitué des issues auxquelles x_i est associée.

valeur k	x_1	$\dots$	x_i	$\dots$	x_n
probabilité $\mathbb{P}(X=k)$	p_1	$\dots$	p_i	$\dots$	p_n



Définition 2

Soit X une variable aléatoire réelle définie sur l'univers fini Ω d'une expérience aléatoire muni d'une loi de probabilité $\mathbb{P}$.
On note $x_1, x_2, \ldots, x_n$ les valeurs prises par X avec les probabilités respectives $p_1, p_2, \ldots, p_n$.
On définit trois indicateurs pour caractériser la loi de la variable aléatoire X.

☞ L'**espérance E** de la variable aléatoire X, notée $\mathbb{E}(X)$, est définie par :

$$\boxed{\mathbb{E}(X) = \mathbb{P}(X = x_1) \times x_1 + \mathbb{P}(X = x_2) \times x_2 + \ldots + \mathbb{P}(X = x_n) \times x_n = p_1 x_1 + p_2 x_2 + \cdots + p_n x_n = \sum_{i=1}^{n} p_i x_i}$$

- $\mathbb{E}(X)$ s'interprète comme la **valeur moyenne** prise par X lorsque l'on répète un grand nombre de fois l'expérience (*loi des grands nombres*).
- $\mathbb{E}(X)$ s'exprime dans la même unité que les valeurs prises par X.
- Un jeu est **équitable** lorsque l'espérance de gain est nulle, il est **favorable au joueur** si cette espérance de gain est positive.

☞ La **variance V** de la variable aléatoire X, notée $\mathbb{V}(X)$, est définie par :

$$\boxed{\mathbb{V}(X) = p_1 (x_1 - \mathbb{E}(X))^2 + p_2 (x_2 - \mathbb{E}(X))^2 + \cdots + p_n (x_n - \mathbb{E}(X))^2 = \sum_{i=1}^{n} p_i (x_i - \mathbb{E}(X))^2}$$

- $\mathbb{V}(X)$ mesure la moyenne des carrés des écarts entre les valeurs prises par X et son espérance. On peut interpréter $\mathbb{V}(X)$ comme le carré de la distance euclidienne entre l'ensemble des valeurs prises par X et $\mathbb{E}(X)$.
- $\mathbb{V}(X)$ est un nombre positif qui s'exprime dans l'unité au carré des valeurs prises par X.
- Un jeu est d'autant plus **risqué** que sa variance est grande.
- $\mathbb{V}(X)$ se calcule en pratique avec la formule de König :

$$\boxed{\mathbb{V}(X) = \sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2}$$

☞ **L'écart-type** σ de la variable aléatoire X, noté $\sigma(X)$, est la racine carrée de $\mathbb{V}(X)$:

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

- $\sigma(X)$ mesure la distance moyenne des valeurs prises par X par rapport à sa valeur moyenne lorsque l'on répète un grand nombre de fois l'expérience.
- $\sigma(X)$ est un nombre positif qui s'exprime dans l'unité des valeurs prises par X.
- Un jeu est d'autant plus **risqué** que son écart-type est grand.

Capacité 1 ***Calculer une espérance, une variance, un écart type***

1. On considère la variable aléatoire Y dont la loi est donnée ci-dessous :

k	-5	1	2	10
$\mathbb{P}(Y=k)$	0,35	0,5	...	0,1

   a. Détailler les calculs de l'espérance $\mathbb{E}(Y)$ de Y, de sa variance $\mathbb{V}(Y)$ et de son écart-type $\sigma(Y)$.

   b. Retrouver ces résultats avec l'éditeur de listes de la calculatrice.

2. On lance un dé à 6 faces équilibré dont les faces sont numérotées de 1 à 6 et on note X le nombre porté par la face du dessus.

   Déterminer l'espérance, la variance et l'écart-type de la variable aléatoire X.

3. On lance un dé à n faces équilibré dont les faces sont numérotées de 1 à n et on note Y le nombre porté par la face du dessus.

   Déterminer l'espérance de la variable aléatoire Y.

1.2 Transformation affine d'une variable aléatoire

Définition 3

Soit X une variable aléatoire réelle définie sur l'univers fini Ω d'une expérience aléatoire muni d'une loi de probabilité $\mathbb{P}$.
Pour tous réel a et b, la variable aléatoire $aX+b$ associe à chaque issue ω le réel $a\mathrm{X}(\omega)+b$.
Si on note x_i avec $1 \leqslant i \leqslant n$ les valeurs prises par X alors variable aléatoire $aX+b$ prend les valeurs ax_i+b et pour tout indice i tel que $1 \leqslant i \leqslant n$, on a $\mathbb{P}(aX+b=ax_i+b)=\mathbb{P}(X=x_i)$.

Propriété 1

Soit X une variable aléatoire réelle définie sur l'univers fini Ω d'une expérience aléatoire muni d'une loi de probabilité $\mathbb{P}$.

Pour tous réel a et b, on a :

- $\mathbb{E}(aX+b)=a\mathbb{E}(X)+b$
- $\mathbb{V}(aX+b)=a^2\mathbb{V}(X)$
- $\sigma(a\mathrm{X}+b)=|a|\sigma(\mathrm{X})$

Démonstration

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

Capacité 2 ***Utiliser la notion d'espérance dans une résolution de problème***

1. Le nombre de spectateurs pour un festival de musique définit une variable aléatoire X d'espérance 12 000 et de variance 1 500. Chaque billet est vendu au tarif de 45 € et le coût global d'organisation du festival est de 100 000 €.

   Soit B la variable aléatoire associée au bénéfice réalisé par l'organisateur du spectacle.

   Déterminer $\mathbb{E}(B)$ et $\sigma(B)$.

2. On considère que pour la session 2 020 d'un concours, la note X sur 10 attribuée à un candidat pris au hasard, aura pour espérance $\mathbb{E}(X) = 5,4$ et pour écart-type $\sigma(X) = 2$

   Le responsable du concours veut obtenir une moyenne de 5 avec un écart-type de 1,5. Ainsi, il veut appliquer une transformation affine à X en lui associant $aX + b$ avec a et b des réels et $a > 0$.

   a. Exprimer $\mathbb{E}(aX+b)$ et $\sigma(aX+b)$ en fonction de a et b.

   b. En déduire le calcul de a et b.

2 Somme de variables aléatoires

2.1 Activités

Activité 1

Faire les activités 1 et 2 p.400 du manuel Indice.

2.2 Définition

Définition 4

Soit X et Y deux variables aléatoires définies sur un univers fini Ω d'expérience aléatoire, qui prennent respectivement pour valeurs les réels x_i tels que $1 \leqslant i \leqslant n$ et y_j tels que $1 \leqslant j \leqslant m$, avec n et m entiers naturels.

- La **variable aléatoire somme** $X+Y$ prend toutes les valeurs possibles x_i+y_j avec :

$$1 \leqslant i \leqslant n \quad \text{et} \quad 1 \leqslant j \leqslant m$$

- Pour déterminer la loi de probabilité de $X+Y$ on détermine pour chaque valeur w prise par $X+Y$ la somme des $\mathbb{P}\big((X=x_i) \text{ et } (Y=y_j)\big)$ tels que $x_i+y_j=w$ avec $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$.

  En pratique, on peut utiliser un tableau à double entrée comme dans l'exemple ci-dessous où les valeurs de $X+Y$ sont notées en gras :

X \ Y	1	2	3	Loi de X
1	0,08 (**2**)	0,17 (**3**)	0,14 (**4**)	0,39
2	0,25 (**3**)	0,18 (**4**)	0,18 (**5**)	0,61
Loi de Y	0,33	0,35	0,32	1

  On en déduit la loi de probabilité de la somme $X+Y$:

k	2	3	4	5
$\mathbb{P}(X+Y=k)$	$0,08$	$0,17+0,25=0,42$	$0,18+0,14=0,32$	$0,18$

Capacité 3 ***Déterminer la loi d'une somme de variables aléatoires plus simples, capacité 1 p. 403 du manuel indice***

Agathe lance deux pièces de monnaie, l'une de 1 € et l'autre de 2 €
X est la variable aléatoire qui vaut 1 si elle obtient Pile avec la pièce de 1 € et 0 sinon.
Y est la variable aléatoire qui vaut 2 si elle obtient Pile avec la pièce de 2 € et 1 sinon.

1. Représenter cette situation à l'aide d'un arbre pondéré.
2. Compléter le tableau ci-contre et en déduire la loi de probabilité de $X+Y$.

X \ Y	1	2	Loi de X
0	...	...	...
1	...	...	...
Loi de Y	...	...	...

2.3 Propriétés

Définition 5

1. Soit X et Y deux variables aléatoires définies sur un univers fini Ω d'expérience aléatoire, qui prennent respectivement pour valeurs les réels x_i tels que $1 \leqslant i \leqslant n$ et y_j tels que $1 \leqslant j \leqslant m$, avec n et m entiers naturels.

   X et Y sont **indépendantes** si pour tout couple (x_i, y_j) de valeurs possibles pour X et Y, on a :

$$\mathbb{P}\big((X=x_i) \text{ et } (Y=y_j)\big) = \mathbb{P}(X=x_i) \times \mathbb{P}(Y=y_j)$$

2. Des variables aléatoires X_1, X_2, ..., X_n définies sur une série d'épreuves indépendantes (le résultat d'une épreuve ne dépend pas des autres), sont **mutuellement indépendantes** si pour toute liste de valeurs $(x_1, x_2, \ldots, x_n)$ de valeurs possibles pour la liste de variables aléatoires $(X_1, X_2, \ldots, X_n)$, on a :

$$\mathbb{P}\big((X_1=x_1) \text{ et } (X_2=x_2) \text{ et } \ldots \text{ et } (X_n=x_n)\big) = \mathbb{P}(X_1=x_1) \times \mathbb{P}(X_2=x_2) \times \ldots \times \mathbb{P}(X_n=x_n)$$

⚠ L'*indépendance deux à deux* de variables aléatoires n'entraîne pas leur *indépendance mutuelle* (pour $n \geqslant 3$) mais dans le cadre du programme de terminale nous ne ferons pas de distinction et utiliserons parfois *indépendants* pour *mutuellement indépendants*. La justification de l'indépendance de deux variables aléatoires et de l'indépendance mutuelle de n variables sont hors programme.

Propriété 2 ***Linéarité de l'espérance et de la variance***

Soit X et Y deux variables aléatoires définies sur un univers fini Ω d'expérience aléatoire, qui prennent respectivement pour valeurs les réels x_i tels que $1 \leqslant i \leqslant n$ et y_j tels que $1 \leqslant j \leqslant m$, avec n et m entiers naturels.

1. L'espérance d'une somme de variables aléatoires $X+Y$ est toujours égale à la somme de leurs espérances, que X et Y soient indépendantes ou non :

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

2. Si on a deux variables aléatoires X et Y **indépendantes** alors la variance de leur somme $X+Y$ est égale à la somme de leurs variances :

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$$

Capacité 4 ***Calculer l'espérance et d'une variable aléatoire avec la propriété de linéarité, capacité 2 p. 403 du manuel Indice***

1. Soit X et Y deux variables aléatoires définies sur univers fini Ω d'expérience aléatoire.

   On sait que $\mathbb{E}(X) = 3$ et $\mathbb{E}(Y) = 4$.

   Déterminer l'espérance des variables aléatoires :

   $X+Y$, $X+Y+1$, $X-Y$, $X+X$, $\dfrac{1}{3}X + \dfrac{1}{2}Y$ et $aX + (1-a)Y$ où a est un réel.

2. Soit S_n la variable aléatoire qui représente le nombre de piles obtenus lorsqu'on lance une pièce équilibrée n fois.

   a. Compléter la fonction Python ci-dessous pour qu'elle renvoie une réalisation de S_n :

```python
from random import randint

def lancer():
    #renvoie 1 pour pile et 0 pour face (pièce équilibrée)
    return randint(0, 1)

def S(n):
    compteur = 0
    for k in range(n):
        compteur = ......
    return compteur
```

   b. Déterminer l'espérance de S_n.

Capacité 5 ***Calculer la variance d'une variable aléatoire en l'exprimant comme somme de variables aléatoires indépendantes***

1. Soit X et Y deux variables aléatoires indépendantes définies sur un univers fini Ω d'expérience aléatoire.

   On admet que si X et Y sont indépendantes alors pour tous réels a, b, c et d, les variables aléatoires $aX+b$ et $cY+d$ sont indépendantes.

   On sait que $\mathbb{V}(X)=3$ et $\mathbb{V}(Y)=4$.

   Déterminer la variance des variables aléatoires :

   $X+Y$, $X+Y+1$, $X-Y$, $X+X$, $\frac{1}{3}X+\frac{1}{2}Y$ et $aX+(1-a)Y$ où a est un réel.

2. Soit S_n la variable aléatoire qui représente le nombre de piles obtenus lorsqu'on lance une pièce équilibrée n fois.

   Déterminer la variance de S_n.

Capacité 6 ***Représenter une variable comme somme de variables aléatoires plus simples, capacité 3 p. 405 du manuel Indice***

Extrait du sujet Polynésie J2 Juin 2024.

Un sac opaque contient huit jetons numérotés de 1 à 8, indiscernables au toucher.
À trois reprises, un joueur pioche un jeton dans ce sac, note son numéro, puis le remet dans le sac.
Dans ce contexte, on appelle « tirage » la liste ordonnée des trois numéros obtenus.
Par exemple, si le joueur pioche le jeton numéro 4, puis le jeton numéro 5, puis le jeton numéro 1, alors le tirage correspondant est $(4\,;5\,;1)$.

1. Déterminer le nombre de tirages possibles.
2. **a.** Déterminer le nombre de tirages sans répétition de numéro.

   b. En déduire le nombre de tirages contenant au moins une répétition de numéro.

On note X_1 la variable aléatoire égale au numéro du premier jeton pioché, X_2 celle égale au numéro du deuxième jeton pioché et X_3 celle égale au numéro du troisième jeton pioché.
Puisqu'il s'agit d'un tirage avec remise, les variables aléatoires X_1, X_2, et X_3 sont indépendantes et suivent la même loi de probabilité.

3. Établir la loi de probabilité de la variable aléatoire X_1
4. Déterminer l'espérance de la variable aléatoire X_1

On note $S=X_1+X_2+X_3$ la variable aléatoire égale à la somme des numéros des trois jetons piochés.

5. Déterminer l'espérance de la variable aléatoire S.
6. Déterminer $P(S=24)$.

7. Si un joueur obtient une somme supérieure ou égale à 22, alors il gagne un lot.

   a. Justifier qu'il existe exactement 10 tirages permettant de gagner un lot.

   b. En déduire la probabilité de gagner un lot.

3 Échantillon de taille n d'une loi de probabilité

3.1 Définition

Définition 6

Soit n un entier naturel non nul.
Un **échantillon** d'une loi de probabilité est une liste $(X_1, X_2, \dots, X_n)$ de n variables aléatoires **indépendantes** et **identiques** qui suivent toutes cette loi.

3.2 Somme d'un échantillon et application à la loi binomiale

Définition 7

Soit une variable aléatoire X.
Soit n un entier naturel non nul et un **échantillon** de taille n, $(X_1, X_2, \dots, X_n)$ de la loi suivie par X.
La **somme** de cet échantillon est la variable aléatoire :

$$S_n = X_1 + X_2 + \dots + X_n$$

Propriété 3

Soit n un entier naturel non nul et S_n la somme d'un **échantillon** de taille n de la loi suivie par une variable aléatoire X.

1. $\mathbb{E}(S_n) = n\mathbb{E}(X)$.
2. $\mathbb{V}(S_n) = n\mathbb{V}(X)$.
3. $\sigma(S_n) = \sqrt{n} \times \sigma(X)$.

Démonstration ***Au programme***

....

....

....

....

....

....

....

....

....

....

....

....

Propriété 4

Soit n un entier naturel non nul et p un réel entre 0 et 1.
Une variable aléatoire S_n, qui suit une loi binomiale $\mathscr{B}(n\,;\,p)$ de paramètres n et p, est la somme $S_n = X_1 + X_2 + \ldots + X_n$ d'un échantillon de taille n d'une loi de Bernoulli de paramètre p.

1. $\mathbb{E}(S_n) = np$.
2. $\mathbb{V}(S_n) = np(1-p)$.
3. $\sigma(S_n) = \sqrt{np(1-p)}$.

Démonstration ***Au programme***

On admet qu'une variable aléatoire S_n suivant une loi binomiale $\mathscr{B}(n\,;\,p)$ est la somme de n variables aléatoires mutuellement indépendantes suivant une même loi de Bernoulli de paramètre p.
Une variable aléatoire X qui suit une loi de Bernoulli de paramètre p a :

- pour espérance $\mathbb{E}(X) = 1 \times p + 0 \times (1-p) = p$;
- pour variance $\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = 1^2 \times p + 0^2 \times (1-p) - p^2 = p - p^2 = p(1-p)$

En appliquant la propriété précédente, il vient : $\mathbb{E}(S_n) = np$, $\mathbb{V}(S_n) = np(1-p)$ et donc $\sigma(S_n) = \sqrt{np(1-p)}$.

Capacité 7 ***Espérance d'une loi binomiale***

Extrait du sujet Asie J1 juin 2024.

Dans la revue *Lancet Public Health*, les chercheurs affirment qu'au 11 mai 2020, 5,7 % des adultes français avaient déjà été infectés par la COVID 19.
Source : https ://www.thelancet.com/journals/lanpub/article/PIIS2468-2667 (21) 00064-5/fulltext
On se servira de cette donnée pour les parties A et B de cet exercice.

1. On prélève un individu dans la population française adulte au 11 mai 2020.

   On note I l'évènement : « l'adulte a déjà été infecté par la COVID 19 »

   Quelle est la probabilité que cet individu prélevé ait déjà été infecté par la COVID 19 ?

2. On prélève un échantillon de 100 personnes de la population supposées choisies de façon indépendante les unes des autres.

   On assimile ce prélèvement à un tirage avec remise.

   On appelle X la variable aléatoire qui compte le nombre de personnes ayant déjà été infectées.

a. Justifiez que X suit une loi binomiale dont on donnera les paramètres.

b. Calculer son espérance mathématique. Interpréter ce résultat dans le cadre de l'exercice.

c. Quelle est la probabilité qu'il n'y ait aucune personne infectée dans l'échantillon ?
On donnera une valeur approchée à 10^{-4} près du résultat.

3.3 Moyenne d'un échantillon

Définition 8

Soit une variable aléatoire X.
Soit n un entier naturel non nul et un **échantillon** de taille n, $(X_1, X_2, \ldots, X_n)$ de la loi suivie par X.
La **moyenne** de cet échantillon est la variable aléatoire :

$$M_n = \frac{S_n}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Propriété 5

Soit n un entier naturel non nul et M_n la moyenne d'un **échantillon** de taille n de la loi suivie par une variable aléatoire X.

1. $\mathbb{E}(M_n) = \mathbb{E}(X)$.
2. $\mathbb{V}(M_n) = \dfrac{\mathbb{V}(X)}{n}$.
3. $\sigma(M_n) = \dfrac{\sigma(X)}{\sqrt{n}}$.

Démonstration

....

....

....

....

....

....

....

....

....

....

Corollaire

Soit n un entier naturel non nul, p un réel entre 0 et 1 et une variable aléatoire S_n, qui suit une loi binomiale $\mathscr{B}(n\,;\,p)$ de paramètres n et p.

1. $\mathbb{E}\left(\frac{S_n}{n}\right) = p$.

2. $\mathbb{V}\left(\frac{S_n}{n}\right) = \frac{p(1-p)}{n}$.

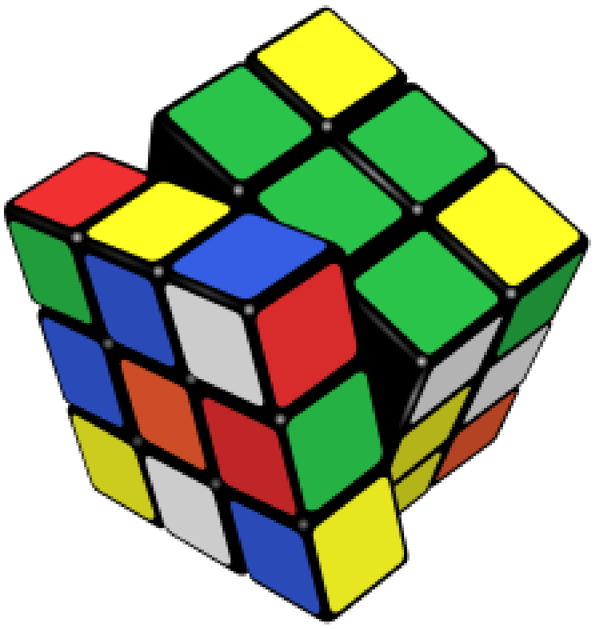
3. $\sigma\left(\frac{S_n}{n}\right) = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$.

Capacité 8 *Utiliser la somme et la moyenne d'un échantillon de taille n, capacité 4 p. 405 du manuel Indice*

Chaque jour, Erno résout une fois le Rubik's Cube. X est l'écart, en secondes, entre le temps qu'il réalise et son meilleur temps. Voici la loi de probabilité de X :

k	0	1	2	10	20
$\mathbb{P}(X=k)$	$0,5$	$0,25$	$0,1$	$0,1$	$0,05$

1. Déterminer l'espérance et l'écart-type de l'écart quotidien moyen sur une dizaine de jours.
2. Reprendre les calculs précédents mais sur vingt puis trente puis cinquante jours. Que peut-on remarquer ?



Auteur : Booyabazooka ; Licence : Creative Commons Attribution-Share Alike 3.0 Unported

Table des matières