Terminale Maths Complémentaires – Thème 08

EXPÉRIENCES RÉPÉTÉES, ÉCHANTILLONNAGE

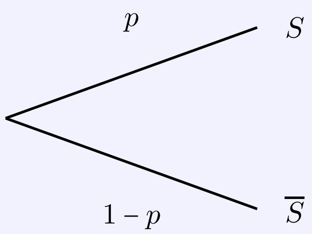


Table des matières

Dans ce chapitre, n et i désignent des entiers naturels.

I Variable aléatoire – rappels

1) Variable aléatoire et loi de probabilité

a) Variable aléatoire

DÉFINITION

Lorsqu'à chaque issue d'une expérience aléatoire, on associe un nombre réel, on dit que l'on définit une **variable aléatoire**.

REMARQUES

- Une variable aléatoire est généralement notée par une lettre majuscule : X, Y, Z, T, G...
- Lorsque x_1, x_2, ..., x_n sont les valeurs prises par une variable aléatoire X, on note $X = x_i$ l'événement « X prend la valeur x_i » (avec $1 \leqslant i \leqslant n$)

b) Loi de probabilité

DÉFINITION

Lorsqu'à chaque valeur x_i $(1 \leqslant i \leqslant n)$ prise par une variable aléatoire X, on associe la probabilité de l'événement $X = x_i$, on dit que l'on définit la **loi de probabilité de** X.
On représente généralement cette loi à l'aide d'un tableau :

Valeur x_i	x_1	x_2	...	x_n
$P(X = x_i)$	p_1	p_2	...	p_n

REMARQUE

On a alors :

$$\sum_{i=1}^{n} p_i = p_1 + p_2 + \ldots + p_n = 1$$

c) Exemple complet

Énoncé :
On lance un dé cubique, non pipé, dont les faces sont numérotées 1, 1, 1, 2, 3 et 4.
Soit X la variable aléatoire donnant le numéro apparu. Déterminer la loi de probabilité de X.

Correction :
Les valeurs prises par X sont 1, 2, 3 et 4.
Le dé étant non pipé, chaque face a la même probabilité d'être obtenue. 3 faces ayant le chiffre 1, on a donc $P(X = 1) = \dfrac{3}{6} = \dfrac{1}{2}$. De même, $P(X = 2) = \dfrac{1}{6}$, $P(X = 3) = \dfrac{1}{6}$ et $P(X = 4) = \dfrac{1}{6}$.
La loi de probabilité de X est donc :

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

2) Espérance, variance et écart-type d'une variable aléatoire

DÉFINITION

Soit n un entier naturel et soit X une variable aléatoire qui prend les valeurs x_1, x_2, ..., x_n.
On pose, pour tout entier naturel i compris entre 1 et n, $p_i = P(X = x_i)$.

- L'espérance mathématique de X est $E(X) = \sum_{i=1}^{n} p_i x_i = p_1x_1 + p_2x_2 + \ldots + p_nx_n$.
- La variance de X est $V(X) = \sum_{i=1}^{n} p_i(x_i - E(X))^2$.
- L'écart-type de X est $\sigma(X) = \sqrt{V(X)}$.

REMARQUE

La formule de König-Huyghens vue en classe de Première donne une autre expression de la variance : $V(X) = E(X^2) - E(X)^2$

PROPRIÉTÉS admises

Soient a et b deux réels et X une variable aléatoire. Alors :

$$E(aX + b) = aE(X) + b \quad ; \quad V(aX + b) = a^2V(X) \quad ; \quad \sigma(aX + b) = |a|\sigma(X)$$

II Loi uniforme discrète

1) Définition sur $\{1\,;2\,;3\,;\ldots;n\}$

DÉFINITION

Soit $n \in \mathbb{N}^*$, et soit X une variable aléatoire définie sur un univers Ω et à valeurs dans $\{1\,;2\,;3\,;\ldots;n\}$.
On dit que X suit la **loi uniforme discrète** sur $\{1\,;2\,;3\,;\ldots;n\}$ si et seulement si :

$$\forall k \in \{1\,;2\,;3\,;\ldots;n\},\ P(X = k) = \frac{1}{n}$$

EXEMPLE

On lance un dé équilibré à 6 faces et on note X le résultat donné par le lancer.
Alors X prend les valeurs entières de 1 à 6, et $P(X = 1) = P(X = 2) = \ldots = P(X = 6) = \dfrac{1}{6}$.
Donc X suit la loi uniforme discrète sur $\{1\,;2\,;3\,;\ldots;6\}$.

EXERCICE

On considère une urne qui contient 4 boules **vertes**, 2 boules **bleues**, 2 boules **rouges**, 3 boules **orange** et 1 boule **jaune**. Toutes les boules sont indiscernables au toucher. On tire au hasard une boule dans l'urne.
- L'obtention d'une boule bleue ou d'une boule rouge rapporte 1 point.
- L'obtention d'une boule verte rapporte 2 points.
- L'obtention d'une boule orange ou d'une boule jaune rapporte 3 points.

Soit X la variable aléatoire qui compte le nombre de points obtenus après le tirage d'une boule de l'urne.

Montrer que X suit la loi uniforme discrète sur $\{1\,;2\,;3\}$.

2) Espérance de la loi uniforme discrète sur $\{1;2;3;\ldots;n\}$

PROPRIÉTÉ

Soit $n \in \mathbb{N}^*$ et soit X une variable aléatoire qui suit la loi uniforme discrète sur $\{1;2;3;\ldots;n\}$.
L'espérance mathématique de X vaut $E(X) = \dfrac{n+1}{2}$.

DÉMONSTRATION

X suit la loi uniforme discrète sur $\{1;2;3;\ldots;n\}$, donc les valeurs prises par X sont tous les entiers compris entre 1 et n, et pour tout entier k compris entre 1 et n, on a $P(X=k) = \dfrac{1}{n}$.
On peut dresser le tableau de la loi de X ainsi :

k	1	2	...	n
$P(X=k)$	$\frac{1}{n}$	$\frac{1}{n}$	...	$\frac{1}{n}$

Ainsi, par définition, $E(X) = 1\times\dfrac{1}{n} + 2\times\dfrac{1}{n} + 3\times\dfrac{1}{n} + \ldots + n\times\dfrac{1}{n} = \dfrac{1+2+3+\ldots+n}{n}$.

Or, pour tout entier naturel n non nul, $1+2+3+\ldots+n = \dfrac{n(n+1)}{2}$.

Donc $E(X) = \dfrac{\frac{n(n+1)}{2}}{n} = \dfrac{n(n+1)}{2n} = \dfrac{n+1}{2}$.

EXEMPLE

Dans l'exemple précédent du dé à 6 faces, on a $E(X) = \dfrac{6+1}{2} = 3{,}5$.
On peut l'interpréter ainsi : si on réalise un grand nombre de lancers de dé, dans des conditions identiques et indépendantes, la moyenne des résultats obtenus par la variable aléatoire X sera très proche de 3,5.

3) Loi uniforme sur un intervalle d'entiers $[\![a;b]\!]$

On peut généraliser la définition et la propriété précédentes dans le cas d'un intervalle d'entiers non plus de la forme $\{1;2;3;\ldots;n\}$, mais entre n'importe quel entier a jusqu'à un entier b :

DÉFINITION

Soient a et b deux entiers relatifs tels que $a<b$, et soit X une variable aléatoire définie sur un univers Ω et à valeurs dans $[\![a;b]\!]$ (c'est-à-dire tous les entiers compris entre a et b : $\{a;a+1;a+2;\ldots;b-1;b\}$).
On dit que X suit la **loi uniforme discrète** sur $[\![a;b]\!]$ si et seulement si :

$$\forall k \in [\![a;b]\!],\ P(X=k) = \frac{1}{b-a+1}$$

PROPRIÉTÉ

Soient a et b deux entiers relatifs tels que $a<b$, et soit X une variable aléatoire qui suit la loi uniforme discrète sur $[\![a;b]\!]$.
L'espérance mathématique de X vaut $E(X) = \dfrac{a+b}{2}$.

DÉMONSTRATION

La loi de X est :

k	a	$a+1$	...	b
$P(X=k)$	$\frac{1}{b-a+1}$	$\frac{1}{b-a+1}$	...	$\frac{1}{b-a+1}$

Alors $E(X) = a \times \frac{1}{b-a+1} + (a+1) \times \frac{1}{b-a+1} + ... + b \times \frac{1}{b-a+1} = \frac{1}{b-a+1}(a+(a+1)+...+b)$.

Or $a+(a+1)+...+b = (1+2+...+b) - (1+2+...+(a-1)) = \frac{b\times(b+1)}{2} - \frac{(a-1)\times a}{2} = \frac{b^2+b-a^2+a}{2}$
$= \frac{(b+a)(b-a)+(b+a)}{2} = \frac{(b+a)(b-a+1)}{2}$.

Ainsi, $E(X) = \frac{1}{b-a+1} \times \frac{(b+a)(b-a+1)}{2} = \frac{a+b}{2}$.

EXEMPLE

Dans un groupe de 20 amis, 5 amis ont 16 ans, 5 ont 17 ans, 5 ont 18 ans et les 5 derniers ont 19 ans.
On choisit au hasard une personne du groupe. Alors la variable aléatoire X égale à l'age de la personne choisie au hasard sur la loi uniforme discrète sur $[\![16\,;19]\!]$.
Pour tout entier k compris entre 16 et 19, on a alors $P(X=k) = \frac{1}{19-16+1} = \frac{1}{4}$.

III Épreuve, loi et schéma de Bernoulli

1) Problématique

On lance vingt fois de suite, dans les mêmes conditions, un dé bien équilibré à 6 faces.

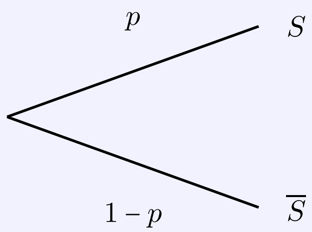
1. Quelle est la probabilité d'obtenir 20 fois la face 6 sur les 20 lancers ?
2. Quelle est la probabilité d'obtenir 0 fois la face 6 sur les 20 lancers ?
3. Quelle est la probabilité d'obtenir 4 fois la face 6 sur les 20 lancers ?

Correction :

1. Soit X la variable aléatoire égale au nombre de fois qu'apparait la face 6 sur les 20 lancers.
   Un seul chemin donne 20 fois la face 6 : celui du haut. Donc $P(X=20) = \left(\frac{1}{6}\right)^{20}$.
2. Un seul chemin donne 0 fois la face 6 : celui du bas. Donc $P(X=0) = \left(\frac{5}{6}\right)^{20}$.
3. Tous les chemins dans l'arbre donnant exactement 4 fois la face 6 sur les 20 lancers ont la même probabilité, car ils contiennent autant de branches qui vont vers le haut (4 branches car 4 succès) que de branches qui vont vers le bas (16 branches car 16 échecs) : cette probabilité vaut $\left(\frac{1}{6}\right)^{4} \times \left(\frac{5}{6}\right)^{16}$.
   Il reste à pouvoir déterminer le nombre de chemins dans l'arbre contenant exactement 4 branches vers le haut et 16 branches vers le bas. Ce nombre de chemins, qui correspond au nombre de **combinaisons** de 4 succès parmi 20 épreuves, est appelé un **coefficient binomial**. Il se note $\binom{20}{4}$ et la calculatrice (**nCr(20,4)**) donne 4845.
   Ainsi, $P(X=4) = \binom{20}{4} \times \left(\frac{1}{6}\right)^{4} \times \left(\frac{5}{6}\right)^{16} = 4845 \times \left(\frac{1}{6}\right)^{4} \times \left(\frac{5}{6}\right)^{16} \approx 0{,}2022$.

2) Épreuve de Bernoulli

DÉFINITION



Soit p un réel appartenant à l'intervalle $[0;1]$.

On appelle **épreuve de Bernoulli** toute expérience aléatoire n'admettant que deux issues, appelées généralement **succès** S et **échec** $\overline{S}$, et de probabilités respectives p et $q = 1-p$.

EXEMPLES

• Lancer une pièce de monnaie équilibrée et savoir si PILE est obtenu est une épreuve de Bernoulli de succès S : « Pile a été obtenu », de probabilité $p = 0{,}5$. L'échec est donc $\overline{S}$: « Face a été obtenu ».
• Interroger une personne dans la rue en France et lui demander si elle est gauchère est une épreuve de Bernoulli de succès S : « La personne est gauchère », de probabilité $p \approx 0{,}13$.

3) Loi de Bernoulli

DÉFINITION

Soit p un réel de l'intervalle $[0;1]$ et X une variable aléatoire.
On réalise une épreuve de Bernoulli dont le succès S a pour probabilité p.
On dit que X est une **variable aléatoire de Bernoulli** lorsqu'elle est à valeurs dans $\{0;1\}$, où la valeur 1 est attribuée au succès.
On dit alors que X suit la **loi de Bernoulli** de paramètre p.
Autrement dit, on a $P(X=1)=p$ et $P(X=0)=1-p$.
On peut résumer la loi de Bernoulli par le tableau suivant :

x_i	1	0
$P(X=x_i)$	p	$1-p$

PROPRIÉTÉ

Soit p un réel de $[0;1]$, et soit X une variable aléatoire qui suit la loi de Bernoulli de paramètre p.
L'espérance mathématiques de X est $E(X)=p$.
La variance de X est $V(X)=p(1-p)$.

DÉMONSTRATION

$$\begin{aligned} E(X) &= 1\times P(X=1)+0\times P(X=0)\\ &= 1\times p+0\times(1-p)\\ &= p. \end{aligned}$$

$$\begin{aligned} V(X) &= P(X=1)\times(1-E(X))^2+P(X=0)\times(0-E(X))^2\\ &= p(1-p)^2+(1-p)(0-p)^2\\ &= p(1)p)^2+p^2(1-p)\\ &= p(1-p)(1-p+p)\\ &= p(1-p). \end{aligned}$$

REMARQUE

L'écart-type de X est alors $\sigma(X) = \sqrt{p(1-p)}$.

EXEMPLE

Un jeu de dé est tel que le joueur gagne lorsque le 6 sort et perd dans le cas contraire.
Soit S l'événement « le 6 sort » ; alors si le dé n'est pas pipé, $P(S) = \dfrac{1}{6}$ et $P(\overline{S}) = 1 - \dfrac{1}{6} = \dfrac{5}{6}$.
La variable aléatoire qui prend la valeur 1 si le 6 sort et la valeur 0 dans les cinq autres cas suit une loi de Bernoulli :

x_i	1	2
$P(X = x_i)$	$\frac{1}{6}$	$\frac{5}{6}$

4) Schéma de Bernoulli

DÉFINITION

Soit n un entier naturel non nul.
Un **schéma de Bernoulli** d'ordre n est la répétition de n épreuves de Bernoulli **identiques** et **indépendantes**.

EXEMPLE

L'expérience présentée en introduction du chapitre est un schéma de Bernoulli d'ordre 20.

5) Coefficients binomiaux

DÉFINITION

Soit n un entier naturel et p un réel appartenant à l'intervalle $[0\,;1]$.
On considère un schéma de Bernoulli de paramètres n et p.
Pour tout entier naturel k compris entre 1 et n, le nombre de façons d'obtenir exactement k succès parmi n épreuves s'appelle un **coefficient binomial**. Il est noté $\binom{n}{k}$ et se lit « k parmi n ».

EXEMPLE

Dans un schéma de Bernoulli d'ordre $n = 20$, le nombre de façons d'obtenir 4 succès parmi les 20 épreuves est $\binom{20}{4} = 4845$.

REMARQUE

Dans le cadre de cette option « Maths Complémentaires », les coefficients binomiaux se calculent à la calculatrice, sauf pour les premières valeurs de n (voir plus loin). Il existe bien sûr une façon plus rigoureuse de les calculer : $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$, avec $n! = n \times (n-1) \times \ldots \times 3 \times 2 \times 1$. (Voir le cours de Spécialité Maths)

PROPRIÉTÉS admises

- Soient n et k deux entiers naturels tels que $0 \leqslant k \leqslant n$. Alors :

$$\binom{n}{k} = \binom{n}{n-k}$$

- Soit n un entier naturel. Alors :

$$\binom{n}{0} = 1$$

- Soit n un entier naturel non nul. Alors :

$$\binom{n}{1} = n \qquad \text{et} \qquad \binom{n}{n} = 1$$

- Soient n et k deux entiers naturels tels que $0 \leqslant k \leqslant n-1$. Alors la **relation de Pascal** est :

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Le **triangle de Pascal**, obtenu en utilisant la relation de Pascal ci-dessus, permet d'obtenir rapidement les valeurs de $\binom{n}{k}$ pour les premières valeurs de k et de n :

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$n=0$	1						
$n=1$	1	1					
$n=2$	1	2	1				
$n=3$	1	3	3	1			
$n=4$	1	4	6	4	1		
$n=5$	1	5	10	10	5	1	
$n=6$	1	6	15	20	15	6	1

IV Loi binomiale

1) Définition de la loi binomiale

DÉFINITION

Soient n un entier naturel non nul et p un réel de l'intervalle $[0\,;1]$.
Soit X la variable aléatoire comptant le nombre de succès obtenus lors la répétition de n épreuves de Bernoulli identiques et indépendantes et dont p est la probabilité du succès.
On dit alors que X suit la **loi binomiale** de paramètres n et p. On peut la noter $\mathcal{B}(n\,;p)$.

PROPRIÉTÉ

Soient n un entier naturel non nul et p un réel de l'intervalle $[0\,;1]$.
Soit X une variable aléatoire qui suit la loi binomiale de paramètres n et p.
Alors pour tout entier k compris entre 0 et n, on a :

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

DÉMONSTRATION

L'événement $X = k$ est associé à l'ensemble des chemins dans l'arbre pour lesquels il y a exactement k succès et donc $n - k$ échecs. Chacun de ces chemins a une probabilité égale au produit des probabilités inscrites sur les branches qui constituent ce chemin, c'est-à-dire $p^k(1-p)^{n-k}$. Or il y a $\binom{n}{k}$ chemins de ce type. D'où $P(X = k) = \binom{n}{k}p^k(1-p)^{n-k}$.

EXEMPLE

Une société spécialisée dans l'audience des médias estime que 19,8% des Français vont regarder le premier match de la France pour l'Euro 2016 de Football. Enzo contacte 20 de ses amis et on estime que le nombre d'amis d'Enzo est assez grand pour assimiler ce tirage à un tirage avec remise de 20 amis. Soit X la variable aléatoire qui donne parmi ces 20 personnes, le nombre de celles qui vont regarder le premier match de l'Euro 2016. (On donnera des valeurs approchées à 10^{-6} près)

1. Justifier que X suit une loi binomiale dont on précisera les paramètres.
2. Déterminer la probabilité qu'une personne contactée sur deux va regarder le match.
3. Déterminer la probabilité qu'au moins 2 personnes sur les 20 vont regarder le match.
4. Déterminer, à l'aide de la calculatrice, la probabilité qu'entre 5 et 15 personnes vont regarder le match.

2) Espérance et variance de la loi binomiale

PROPRIÉTÉ admise

Soient n un entier naturel non nul et p un réel de l'intervalle $[0;1]$.
Si X est une variable aléatoire qui suit la loi binomiale $\mathcal{B}(n;p)$, alors :

$$E(X) = np \qquad \text{et} \qquad Var(X) = np(1-p)$$

REMARQUE

L'écart-type de X est alors $\sigma(X) = \sqrt{np(1-p)}$.

EXEMPLE

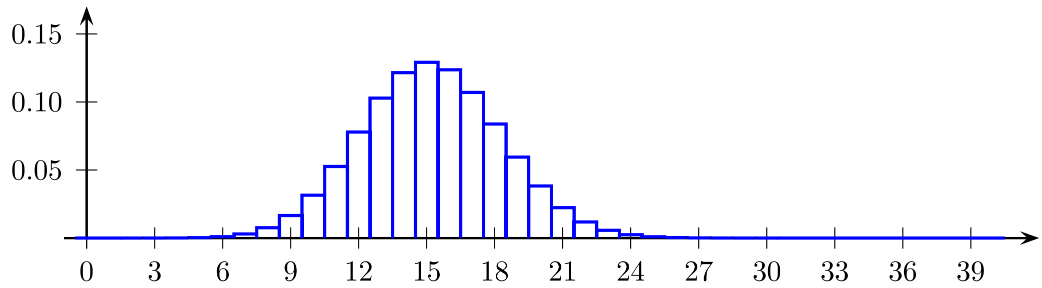
On reprend l'exemple précédent.
Calculer l'espérance de X et interpréter le résultat obtenu.

V Échantillonnage et estimation

1) Représentation graphique d'une loi binomiale

Soit X une variable aléatoire suivant la loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0;1]$.
On peut représenter la loi binomiale par un **diagramme en bâtons**, où les valeurs entières k prises par X sont représentées sur l'axe des abscisses, et les hauteurs de chaque baton est proportionnelle à la probabilité $P(X = k)$.

Par exemple, si X suit la loi binomiale de paramètres $n = 40$ et $p = 0{,}38$, on obtient le diagramme ci-dessous :



L'espérance de X est $E(X) = np = 40 \times 0{,}38 = 15{,}2$.
On remarque que les valeurs de X forment une représentation en forme de « cloche » qui est centrée sur l'espérance.
Plus l'écart-type est grand, plus les valeurs de X sont dispersées autour de $E(X)$: ainsi,le diagramme est plus étalé et « aplati ».

2) Intervalle de fluctuation et prise de décision

DÉFINITION

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p, où n est un entier naturel non nul et p un réel de $[0;1]$.
Un **intervalle de fluctuation** de X **au seuil de** 95% est l'intervalle $[a;b]$ où :
a est le plus petit entier tel que $P(X \leqslant a) > 0{,}025$;
b est le plus petit entier tel que $P(X \leqslant b) \geqslant 0{,}975$

REMARQUE

L'intervalle de fluctuation de la fréquence $F = \dfrac{X}{n}$ est alors $\left[\dfrac{a}{n};\dfrac{b}{n}\right]$.

EXEMPLE

Soit X une variable aléatoire suivant la loi binomiale de paramètres $n = 100$ et $p = 0{,}25$.

1. Tabuler sur la calculatrice la loi binomiale correspondante.
2. Déterminer l'intervalle de fluctuation au seuil de 95% de X.

La détermination d'un intervalle de fluctuation permet de décider si une affirmation faite sur une proportion dans une population est acceptable ou non.
En effet, en faisant une hypothèse sur la proportion p d'un caractère dans une population, on peut déterminer un intervalle de fluctuation I à 95% de la fréquence de ce caractère dans un échantillon de taille n.
Ainsi, on établira une règle de décision : si la fréquence observée dans l'échantillon n'appartient pas à I, comme cela n'a qu'une probabilité de 0,05 de se produire, alors on rejettera l'hypothèse faite sur p, avec un risque de se tromper de 5%.
On en déduit la propriété suivante :

PROPRIÉTÉ

On considère une population dans laquelle on **suppose** que la proportion d'un caractère est p. Après expérience, on **observe** f comme fréquence de ce caractère dans un échantillon de taille n.
Soit l'hypothèse : « La proportion de ce caractère dans la population est p ».
Si I est l'intervalle de fluctuation à 95% de dans un échantillon de taille n, alors :
- Si $f \notin I$: on rejette cette hypothèse au seuil de risque de 5%.
- Sinon, on ne rejette pas cette hypothèse au seuil de risque de 5%.

EXEMPLE

Dans un article, un journaliste affirme que 42% des jeunes qui aiment la lecture préfèrent lire le soir.
On interroge au hasard 140 jeunes qui aiment lire et on assimile le sondage à un tirage successif avec remise : on observe que 49 jeunes déclarent lire le soir.
Peut-on mettre en doute, au seuil de 5%, l'affirmation du journaliste ?

Correction :
On fait l'hypothèse que la proportion de jeunes aimant lire le soir est $p = 0{,}42$.
On interroge 140 jeunes au hasard. On note X la variable aléatoire égale au nombre de jeunes préférant lire le soir. X suit une loi binomiale de paramètres $n = 140$ et $p = 0{,}42$.

Dans la table des probabilités cumulées de X, on recherche :
- Le plus petit entier a tel que $p(X \leqslant a) > 0{,}025$. On trouve $a = 47$.
- Le plus petit entier b tel que $p(X \leqslant b) \geqslant 0{,}975$. On trouve $b = 70$.

Comme la taille de l'échantillon est $n = 140$, l'intervalle de fluctuation à 95% de la fréquence correspondant à X est $\left[\frac{47}{140}; \frac{70}{140}\right]$ soit environ $[0{,}33; 0{,}5]$.

La fréquence observée est $f = \frac{49}{140} = 0{,}35$, donc f **appartient** à l'intervalle $[0{,}33; 0{,}5]$.

On ne peut donc pas remettre en question l'hypothèse que la proportion des jeunes lecteurs préférant lire le soir est $p = 0{,}42$. On ne met donc pas en doute l'affirmation du journaliste.

EXERCICE

On a lancé 100 fois de suite une pièce de monnaie et le côté PILE est apparu 65 fois. Peut-on estimer que la pièce est équilibrée ?

Correction :
Tester si la pièce est équilibrée c'est tester l'hypothèse que la sortie de PILE a une probabilité égale à 0,5.
On observe dans un tableur (=binompdf(100,0.5)) que le plus petit entier tel que $p(X \leqslant a) > 0{,}025$ est $a = 40$, et le plus petit tel que $p(X \leqslant b) \geqslant 0{,}975$ est $b = 60$.
Sachant que $n = 100$, l'intervalle de fluctuation cherché est donc $I = [0{,}4; 0{,}6]$.
La fréquence d'apparition de piles sur les 100 lancers effectués est de 0,65 qui n'appartient pas à I donc on rejette l'hypothèse que la pièce soit équilibrée avec une probabilité inférieure à 5% de se tromper.

3) Intervalle de confiance et estimation

THÉORÈME — admis

Soit F_n la variable aléatoire fréquence qui à chacun des échantillons de taille n associe la fréquence du caractère dans cet échantillon.
La proportion inconnue p est telle que :

$$P\left(F_n - \frac{1}{\sqrt{n}} \leqslant p \leqslant F_n + \frac{1}{\sqrt{n}}\right) \geqslant 0{,}95$$

DÉFINITION

On observe la fréquence f_{obs} sur un échantillon de taille n et p désigne la proportion inconnue d'apparition du caractère dans la population entière. On appelle **intervalle de confiance** de p au niveau asymptotique de 95% l'intervalle :

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

PROPRIÉTÉ — admise

Cet intervalle de confiance a pour amplitude $\frac{2}{\sqrt{n}}$.

Ainsi, si l'on souhaite encadrer p dans un intervalle d'amplitude inférieure ou égale à un réel a, on doit avoir :

$$\frac{2}{\sqrt{n}} \leqslant a, \text{ c'est à dire : } n \geqslant \frac{4}{a^2}$$

REMARQUE

Pour un intervalle de confiance, puisque l'on veut que $P\left(F_n - \frac{1}{\sqrt{n}} \leqslant p \leqslant F_n + \frac{1}{\sqrt{n}}\right)$ soit **supérieur ou égal** à 0,95, il faut arrondir les bornes de l'intervalle de confiance ainsi :

- On arrondit la borne inférieure de l'intervalle de confiance par une valeur approchée **par défaut**.
- On arrondit la borne supérieure de l'intervalle de confiance par une valeur approchée **par excès**.

EXERCICE

Voici les résultats d'un sondage IPSOS réalisé avant l'élection présidentielle de 2002 pour *Le Figaro* et *Europe 1*, les 17 et 18 avril 2002 auprès de 989 personnes, constituant un échantillon national représentatif de la population française âgée de 18 ans et plus et inscrite sur les listes électorales.

On suppose cet échantillon constitué de manière aléatoire (même si en pratique ce n'est pas le cas). Les intentions de vote au premier tour pour les principaux candidats sont les suivantes :

Jacques Chirac : 20%
Lionel Jospin : 18%
Jean-Marie Le Pen : 14%

Les médias se préparent pour un second tour entre Jacques Chirac et Lionel Jospin.

1. Déterminer, pour chaque candidat, l'intervalle de confiance au niveau de confiance de 0,95 de la proportion inconnue d'électeurs ayant l'intention de voter pour lui.
2. Le 21 avril, les résultats du premier tour des élections sont les suivantes :
   Jacques Chirac : 19,88%
   Lionel Jospin : 16,18%
   Jean-Marie Le Pen : 16,86%
   Les pourcentages de voix recueillies par chaque candidat sont-ils bien dans les intervalles de confiance précédents ?
3. Pouvait-on, au vu de ce sondage, écarter avec un niveau de confiance de 0,95, l'un de ces trois candidats du second tour ?