1. Intervalle de fluctuation

On considère une population dans laquelle on suppose que la proportion d'un certain caractère C est p et dans laquelle on prélève au hasard et avec remise un échantillon n.

On considère la variable aléatoire X égale au nombre d'apparitions du caractère C dans l'échantillon et la variable aléatoire $F = \dfrac{X}{n}$ égale à la fréquence du caractère dans l'échantillon.

On détermine le plus petit entier a tel que $P(X \le a) > 0{,}025$ et le plus petit entier b tel que $P(X \le b) \ge 0{,}975$.

L'intervalle $I_n = \left[\dfrac{a}{n} ; \dfrac{b}{n}\right]$ est appelé intervalle de fluctuation de F au seuil de 95 %.

Remarque : Le tirage au hasard d'un individu dans une population qui peut présenter un caractère C avec une probabilité est assimilable à une épreuve de Bernoulli de paramètre p où le succès est « avoir le caractère C ».
Le prélèvement d'un échantillon de taille n, dans cette population s'assimile alors à un schéma de Bernoulli de paramètres $(n ; p)$ et la variable aléatoire X qui compte le nombre de succès suit la loi binomiale $B(n ; p)$.

La variable aléatoire $F = \dfrac{X}{n}$ représente alors la fréquence aléatoire du succès sur un échantillon de taille n.

Ainsi $p(\dfrac{a}{n} \le F \le \dfrac{b}{n}) = p(a \le X \le b)$. D'autres seuils sont possibles : 99 %, ...

2. Rappels de seconde

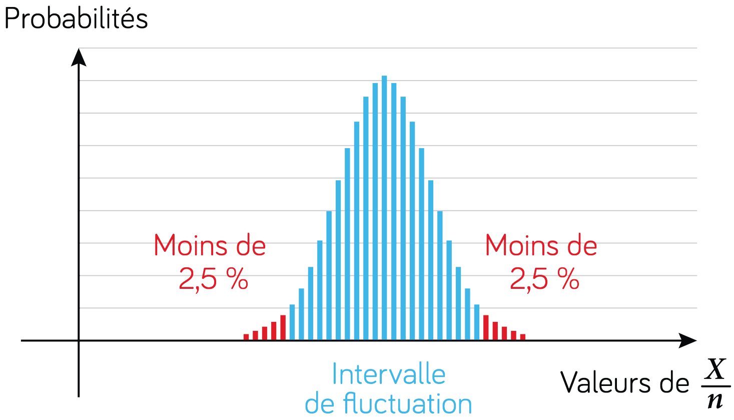
En classe de seconde, on a observé que sur un grand nombre d'échantillons de taille n, 95% au moins fournissent une fréquence f appartenant à l'intervalle $[p - \dfrac{1}{\sqrt{n}} ; p + \dfrac{1}{\sqrt{n}}]$ sous certaines conditions pour n et p :

On dispose donc en termes de probabilité du résultat suivant :

Pour $n \ge 25$ et $0{,}2 \le p \le 0{,}8$ lorsqu'on prélève un échantillon de taille n dans une population ou la probabilité du caractère est p, la fréquence f du caractère sur cet échantillon appartient à l'intervalle $[p - \dfrac{1}{\sqrt{n}} ; p + \dfrac{1}{\sqrt{n}}]$ avec une probabilité supérieure ou égale à 0,95 .

Soit : on a $P(p - \dfrac{1}{\sqrt{n}} \le F \le p + \dfrac{1}{\sqrt{n}}) \ge 0{,}95$, et on dit que $[p - \dfrac{1}{\sqrt{n}} ; p + \dfrac{1}{\sqrt{n}}]$ est un intervalle de fluctuation au seuil de 95 %.

3. Méthode de calculs



k	$p(X \le k)$
0	3,6562E-005
1	0,000524049
2	0,003611472
3	0,015961163
4	0,050951953
5	0,125598973
6	0,250010672
7	0,415892938
8	0,595598725
9	0,755337203
10	0,872478754
11	0,943473633
12	0,978971073
13	0,993534125
14	0,998388475
15	0,999682969
16	0,999952655
17	0,999994959
18	0,999999659
19	0,999999989
20	1

En pratique: on cherche l'intervalle $\left[\dfrac{a}{n} ; \dfrac{b}{n}\right]$ de plus faible amplitude ; pour cela, il suffit de chercher les plus petits entiers a et b tels que : $p(X \le a) \ge 0{,}025$ et $p(X \le b) \ge 0{,}975$ ce qui implique bien $p(a \le X \le b) \ge 0{,}95$.

A l'aide de la calculatrice :
T.I. : Dans f(x), entrer Y1 = binomFrép(n, p, X) et Table avec DebTable = 0, Pas = 1.
CASIO : Menu TABLE → OPTN → STAT → DIST → BINM → Bcd
A l'écran : BinominalCD(X,n,p)
(n et p étant les paramètres de la loi binomiale de X) , SET permet de régler les valeurs de k.
Dans la liste des valeurs de Y1, on cherche a et b vérifiant
$p(X \le a) \ge 0{,}025$ et $p(X \le b) \ge 0{,}975$.

Dans le tableau ci-contre, les valeurs de $p(X \le k)$ pour $X \sim B(20, 0.4)$.
dans ce cas, $a = 4$ et $b = 12$.

4. Prise de décision à partir d'un échantillon

Exemple : On cherche à savoir si une pièce est bien équilibrée.
On fait l'hypothèse que la pièce est bien équilibrée, donc la probabilité d'obtenir PILE est $p = 0{,}5$.
On lance n fois cette pièce et on détermine la fréquence f de PILE obtenue.
On se fixe le seuil 95 % et on détermine l'intervalle de fluctuation à l'aide de la loi $B(n\ ;\ 0{,}5)$.
On prend une décision :

• Si f n'est pas dans l'intervalle de fluctuation, on rejette l'hypothèse que la pièce est bien équilibrée avec un risque de se tromper de 5 %;
• Si f est dans l'intervalle de fluctuation, on ne rejette pas l'hypothèse que la pièce est bien équilibrée (on ne dit pas que l'on accepte cette hypothèse car le risque de se tromper en l'acceptant est inconnu).

Remarque :
C'est avec ce type de prise de décision (mais avec des méthodes beaucoup plus compliquées et beaucoup plus précises) qu'on détermine l'efficacité de certains médicaments ou les effets secondaires de ces médicaments.

EXERCICES PREMIÈRE S ESTIMATION

EXERCICE 1

On sait que dans la population française les naissances de garçons ont pour probabilité $p = 0{,}512$.
On prélève un échantillon de 100 enfants.
a) Donner un intervalle donnant la fréquence probable au seuil de 95%.
b) Quel est le nombre de garçons probables dans cet échantillon ?

EXERCICE 2

On lance un dé 200 fois et dans les résultats on constate que l'on a obtenu 60 fois la face 6.
Au seuil de 95 %, que peut-on dire de cette série de lancers ?

EXERCICE 3

On considère un paquet de cartes contenant 3 cœurs et 7 piques, on effectue 100 tirages d'une carte en remettant à chaque fois la carte dans le paquet. Déterminer un intervalle de fluctuation au seuil de 95 % de la fréquence d'une carte de cœur dans l'échantillon prélevé.

EXERCICE 4

Monsieur Z, chef du gouvernement d'un pays lointain, affirme que 52 % des électeurs lui font confiance. On interroge 100 électeurs au hasard (la population est suffisamment grande pour considérer qu'il s'agit de tirages avec remise). Déterminer à partir de quelles fréquences, au seuil de 95 %, on peut mettre en doute le pourcentage annoncé par Monsieur Z, dans un sens, ou dans l'autre.

EXERCICE 5

Un laboratoire pharmaceutique annonce qu'un médicament sauve 40 % des patients atteints d'une maladie rare.Pour contrôler cette affirmation, on teste ce médicament sur 100 patients atteints de cette maladie.
Soit X la variable aléatoire qui, à chaque échantillon de 100 malades choisis au hasard, associe le nombre de malades sauvés par ce médicament.
1. Quelle loi suit X ?
2. Déterminer les plus petits entiers a et b tels que $p(X \leq a) \geq 0{,}025$ et $p(X \leq b) \geq 0{,}975$.
3. Énoncer la règle de décision permettant de rejeter ou non l'hypothèse $p = 0{,}4$, selon la fréquence f observée dans l'échantillon.
4. Sur les 100 malades auxquels on a administré ce médicament, on en a sauvé 30. Au seuil de risque 5 %, que peut-on dire de l'annonce faite par le laboratoire ?

EXERCICE 6

En France, 22 % des sportifs licenciés ont une licence de football. Soit X la variable aléatoire qui, à chaque échantillon de 100 sportifs licenciés choisis au hasard, associe le nombre de licenciés de football.
Déterminer un intervalle de fluctuation au seuil de 95 %, puis au seuil de 99 %.
Comparer cet intervalle de fluctuation avec celui obtenu en seconde.