Exercices sur les intervalles de fluctuation

Exercice 1

Un candidat lors une élection souhaite savoir s'il pourra être élu dès le premier tour (c'est à dire récolter plus de 50% des voix). Il organise un sondage portant sur un échantillon représentatif comportant 500 votants.

1. En supposant que 50% de la population souhaite voter pour ce candidat, donner l'intervalle de fluctuation au seuil de 95% pour un échantillon de 500 personnes.
2. Sur les 500 personnes interrogées, 223 disent qu'elles voteront pour ce candidat. Peut-il espérer être élu dès le premier tour?

Exercice 2

On cherche à savoir si une pièce est truquée à partir d’un échantillon de lancers de pièces : on obtient 2 050 fois piles en lançant 4 000 fois cette pièce et on veut tester l’hypothèse $p = 0{,}5$.

Exercice 3

Dans la réserve indienne d’Aamjiwnaag, située au Canada il est né entre 1999 et 2003, 132 enfants dont 46 garçons. Est-ce le fruit du hasard ? On considèrera que la proportion théorique est $p = 0{,}5$.

Exercice 4

Les entreprises sont sensées ne pas faire de discrimination quant au sexe des personnes employées.
Deux entreprises A et B ont respectivement 41 femmes pour 100 employés et 4 850 femmes sur 10 000 employés. Pour chacune des entreprises, la sélection est-elle équitable ?

Exercice 5

Dans la ville F, on considère qu'il y a 265 jours de soleil par an.
Déterminer l'intervalle de fluctuation de la fréquence de jours de soleil sur une période de 6 mois.
On considèrera qu’une année comporte 365 jours et 6 mois comportent 183 jours.

Exercice 6

Si on lance un dé, la proportion d'avoir une valeur supérieure ou égale à 5 est de $\frac{1}{3}$.

1. Déterminer les intervalles de fluctuation à 95% si on lance le dé :
   - 50 fois, - 250 fois, - 1000 fois ?
2. Combien de fois faudrait-il lancer le dé pour que l'intervalle de fluctuation correspondant à la sortie d'un nombre supérieur ou égal à 5 ait une amplitude inférieure à 0,01?

Exercice 7

Au Royaume Uni, 31% de collégiens souffrent d'asthme soit une proportion $p = 0{,}31$.

Dans un collège de 284 élèves, 81 ont mentionné « asthme » soit une fréquence de $\frac{81}{284} \simeq 0{,}285$.

Ce collège présente-t-il des statistiques inquiétantes par rapport à l’ensemble de la population ?

Exercice 8

Deux groupes de malades ont été soignés avec deux médicaments différents.
Dans le groupe A, il y a eu 50% de guérison.
Dans le groupe B, 31% de guérison.
Peut-on s'autoriser à penser, au vu de ces résultats, que le médicament administré au groupe A est plus efficace que le médicament administré au groupe B dans les cas suivants :

1. Les groupes A et B sont constitués de 30 personnes chacun,
2. Les groupes A et B sont constitués de 180 personnes chacun.

Exercice 9

La fréquence des yeux bleus en France est d’environ 0,31.
On a prélevé un échantillon de 50 individus dont 15 ont les yeux bleus.

a) Quel est l’intervalle de fluctuation au seuil de 95 % ?

b) Cet échantillon est-il représentatif de la population pour ce caractère ?
c) Votre classe est elle un échantillon représentatif de la population ?

Exercice 10

Depuis 1996, l'accès au Musée du Louvres est gratuit le premier dimanche de chaque mois.
La proportion des visiteurs français ces jours-là est de 0,59.
Une toute nouvelle exposition est proposée. On souhaite connaître son impact sur la fréquentation des visiteurs les jours de gratuité. Sur un échantillon de visiteurs un jour de gratuité, 67 % sont français.
Décider si la nouvelle exposition a eu un impact si l'échantillon comporte :

a) 50 visiteurs ?
b) 500 visiteurs ?

Exercice 11

La fréquence d'utilisation du diesel dans les moteurs de voiture est aujourd'hui de 0,52 en France.
On a étudié des échantillons dans différents départements.

- Echantillon 1 : Taille : 30 - Fréquence : 0,6.
- Echantillon 2 : Taille : 35 - Fréquence : 0,3.

Pour chaque échantillon, indiquer si la fréquence observée appartient ou non à l'intervalle de fluctuation au seuil de 95 %.

Exercice 12

On a programmé une fonction nommée **hasard()**, censée retourner le nombre 0 dans 50% des cas et le nombre 1 dans les autres cas.
Pour tester cette fonction, on utilise un programme basé sur l'algorithme suivant :

```
variable
    somme: nombre
début algorithme
    // initialisation
    somme ← 0
    // traitement
    pour i variant de 1 à 10 000
        somme ← somme + hasard()
    fin pour
    // sortie
    écrire "Le nombre 1 a été généré " somme " fois"
fin algorithme
```

1. Expliquer le fonctionnement de l'algorithme ci-dessus.
2. L'exécution de l'algorithme retourne le message "Le nombre 1 a été généré 4947 fois". Peut-on en déduire une anomalie pour la fonction **hasard()**?

CORRIGE des Exercices sur les Intervalles de fluctuation

Exercice 1

Un candidat lors une élection souhaite savoir s'il pourra être élu dès le premier tour (c'est à dire récolter plus de 50% des voix). Il organise un sondage portant sur un échantillon représentatif comportant 500 votants.

1. En supposant que 50% de la population souhaite voter pour ce candidat, donner l'intervalle de fluctuation au seuil de 95% pour un échantillon de 500 personnes.
2. Sur les 500 personnes interrogées, 223 disent qu'elles voteront pour ce candidat. Peut-il espérer être élu dès le premier tour?

CORRIGE :

1. On suppose que la proportion de la population qui votera pour ce candidat est CONNUE $p = 0,5$
   L'effectif de l'échantillon est n = 500.
   On a bien : $0,2 \leq p \leq 0,8$ et $n \geq 25$
   L'intervalle de fluctuation au seuil de 95% demandé est donc :
   $$I_F = \left[0,5 - \frac{1}{\sqrt{500}} ; 0,5 + \frac{1}{\sqrt{500}}\right] \text{ soit approximativement } I = [0,455 ; 0,545]$$
2. 223 personnes par rapport aux 500 interrogées représentent une fréquence de :
   $$f = \frac{223}{500} = 0,446$$
   Cette fréquence n'appartient pas à l'intervalle de fluctuation : Il est donc très peu vraisemblable que ce candidat soit élu dès le premier tour.

Exercice 2

On cherche à savoir si une pièce est truquée à partir d'un échantillon de lancers de pièces. Ainsi, si l'on obtient 2 050 fois piles en lançant 4 000 fois une pièce et que l'on veuille tester l'hypothèse $p = 0,5$, cela revient à tester si la pièce est truquée.

CORRIGE

On a bien : $0,2 \leq p \leq 0,8$ et $n \geq 25$
On sait que si n = 4 000 alors pour au moins 95% des expériences (qui consistent à lancer 4 000 pièces), les fréquences appartiendront à l'intervalle de fluctuation :

$$I_F = \left[0,5 - \frac{1}{\sqrt{4000}} ; 0,5 + \frac{1}{\sqrt{4000}}\right] = [0,484 ; 0,516]$$

Ici, on a : $f = \dfrac{2050}{4000} = 0,5125$ ainsi $f \in I_F$

Ceci signifie qu'on a 95% de chances de penser que la pièce n'est pas truquée mais aussi 5% de faire erreur.

Exercice 3

Dans la réserve indienne d'Aamjiwnaag, située au Canada il est né entre 1999 et 2003, 132 enfants dont 46 garçons. Est-ce le fruit du hasard ?

CORRIGE :

46 garçons sur 132 naissances alors que la proportion théorique de garçons est $p = 0,5$.
On a bien : $0,2 \leq p \leq 0,8$ et $n \geq 25$ donc l'intervalle de fluctuation au seuil de 95% est :

$$\left[0,5 - \frac{1}{\sqrt{132}} ; 0,5 + \frac{1}{\sqrt{132}}\right] = [0,413 ; 0,587]$$

Il n'y a donc que 5% de chances d'obtenir une valeur en dehors de cet intervalle.

La fréquence de l'échantillon est égale à $\dfrac{46}{132} = 0,348$: elle n'appartient pas à l'intervalle de fluctuation.

Il y a certainement un problème qu'il sera urgent de rectifier.

Exercice 4

Les entreprises sont sensées ne pas faire de discrimination quant au sexe des personnes employées.
Deux entreprises A et B ont respectivement 41 femmes pour 100 employés et 4 850 femmes sur 10 000 employés. Pour chacune des entreprises, la sélection est-elle équitable ?

CORRIGE :

La proportion théorique $p=0,5$ est connue : on a bien : $0,2 \le p \le 0,8$ et $n \ge 25$
Pour l'entreprise A l'intervalle de fluctuation de la fréquence au seuil de 95% est :

$$\left[0,5-\frac{1}{\sqrt{100}};0,5+\frac{1}{\sqrt{100}}\right]=[0,4;0,6] \quad \text{or} \quad f=\frac{41}{100}=0,41$$

→ $0,41 \in [0,4;0,6]$ donc l'échantillon peut représenter une situation de parité.

Pour l'entreprise B, l'intervalle de fluctuation est :

$$\left[0,5-\frac{1}{\sqrt{10000}};0,5+\frac{1}{\sqrt{10000}}\right]=[0,49;0,51] \quad \text{or f} = \frac{4850}{10000}=0,485$$

→ $0,485 \notin [0,49;0,51]$, donc l'échantillon n'est pas représentatif d'une situation de parité.

Exercice 5

Dans la ville F, on considère qu'il y a 265 jours de soleil par an.
Déterminer l'intervalle de fluctuation de la fréquence de jours de soleil sur une période de 6 mois.
On considèrera qu'une année comporte 365 jours et 6 mois comportent 183 jours.

CORRIGE

La proportion moyenne de jours de soleil est connue : $p=\frac{265}{365}\simeq 0,726$

Pour 6 mois : n > 25 et $p \in [0,2\ ;\ 0,8]$: l'intervalle de fluctuation à 95% pour 6 mois est donc :

$$\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]=\left[0,726-\frac{1}{\sqrt{183}};0,726+\frac{1}{\sqrt{183}}\right]=[0,652;0,800]$$

Pour 6 mois, soit 183 jours, l'intervalle à 95 % devient :

$$[0,652\times 183\ ;0,800\times 183]=[119\ \text{jours}\ ;146\ \text{jours}]$$

Exercice 6

Si on lance un dé, la proportion d'avoir une valeur supérieure ou égale à 5 est de $\frac{1}{3}$.

1. Déterminer les intervalles de fluctuation à 95% si on lance le dé :
   - 50 fois, - 250 fois, - 1000 fois ?
2. Combien de fois faudrait-il lancer le dé pour que l'intervalle de fluctuation correspondant à la sortie d'un nombre supérieur ou égal à 5 ait une amplitude inférieure à 0,01?

CORRIGE

1. Les conditions sont respectées : n > 25 et $p \in [0,2\ ;\ 0,8]$

   si on lance 50 fois le dé : $\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]=\left[\frac{1}{3}-\frac{1}{\sqrt{50}};\frac{1}{3}+\frac{1}{\sqrt{50}}\right]=[0,19;0,47]$

   si on lance 250 fois le dé : $\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]=\left[\frac{1}{3}-\frac{1}{\sqrt{250}};\frac{1}{3}+\frac{1}{\sqrt{250}}\right]=[0,27;0,40]$

   si on lance 1000 fois le dé : $\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]=\left[\frac{1}{3}-\frac{1}{\sqrt{1000}};\frac{1}{3}+\frac{1}{\sqrt{1000}}\right]=[0,30;0,36]$

2. **L'amplitude**, différence entre la borne inférieure et la borne supérieure, doit être inférieure à 0,01.

L'amplitude de l'intervalle $\left[p-\frac{1}{\sqrt{n}}\,;p+\frac{1}{\sqrt{n}}\right]$ est de : $\left(p+\frac{1}{\sqrt{n}}\right)-\left(p-\frac{1}{\sqrt{n}}\right)=\frac{2}{\sqrt{n}}$

$$\frac{2}{\sqrt{n}}\leq 0,01 \quad\Leftrightarrow\quad 2\leq 0,01\times\sqrt{n} \quad\Leftrightarrow\quad \frac{2}{0,01}\leq\sqrt{n} \quad\Leftrightarrow\quad 200\leq\sqrt{n} \quad\Leftrightarrow\quad 40\ 000\leq n$$

On devra donc lancer le dé plus de 40 000 fois

Exercice 7

Au Royaume Uni, 31% de collégiens souffrent d'asthme soit une proportion $p=0,31$.

Dans un collège de 284 élèves, 81 ont mentionné « asthme » soit une fréquence de $\frac{81}{284}\simeq 0,285$.

Ce collège présente-t-il des statistiques inquiétantes par rapport à l'ensemble de la population ?

CORRIGE

Pour n = 284 élèves : n > 25 et $p\in[0,2\ ;\ 0,8]$, donc l'intervalle de fluctuation 95% est :

$$\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]=\left[0,31-\frac{1}{\sqrt{284}};0,31+\frac{1}{\sqrt{284}}\right]=[0,25;0,37]$$

La fréquence f = 0,29 est dans cet intervalle donc la population du collège n'est pas significativement différente à 95% de la population du Royaume Uni.

Exercice 8

Deux groupes de malades ont été soignés avec deux médicaments différents.

Dans le groupe A, il y a eu 50% de guérison.

Dans le groupe B, 31% de guérison.

Peut-on s'autoriser à penser, au vu de ces résultats, que le médicament administré au groupe A est plus efficace que le médicament administré au groupe B dans les cas suivants :

1. Les groupes A et B sont constitués de 30 personnes chacun,
2. Les groupes A et B sont constitués de 180 personnes chacun.

CORRIGE

1. Les groupes A et B sont constitués de 30 personnes chacun :

   Les conditions sont respectées : n > 25 et les proportions connues $p\in[0,2\ ;\ 0,8]$

   L'intervalle de fluctuation pour le médicament A est : $\left[0,5-\frac{1}{\sqrt{30}};0,5+\frac{1}{\sqrt{30}}\right]=[0,32;0,68]$

   L'intervalle de fluctuation pour le médicament B est : $\left[0,31-\frac{1}{\sqrt{30}};0,31+\frac{1}{\sqrt{30}}\right]=[0,13;0,49]$

   Cet exercice pose un problème pour le médicament B car son efficacité (= 0,31) ne fait pas partie de l'intervalle de confiance à 95% de A $\left([0,32;0,68]\right)$ donc le résultat du traitement A est significativement différent du résultat du groupe B.

2. Les groupes A et B sont constitués de 180 personnes chacun :

   Les conditions sont respectées : n > 25 et les proportions connues $p\in[0,2\ ;\ 0,8]$

   L'intervalle de fluctuation pour le médicament A est : $I_A=\left[0,5-\frac{1}{\sqrt{180}};0,5+\frac{1}{\sqrt{180}}\right]=[0,425;0,575]$

   L'intervalle de fluctuation pour le médicament B est : $I_B=\left[0,31-\frac{1}{\sqrt{180}};0,31+\frac{1}{\sqrt{180}}\right]=[0,235;0,385]$

   Il n'y a pas chevauchement entre ces deux intervalles de confiance à 95%.

Dans cette partie, I_B ne fait pas partie de l'intervalle de confiance à 95% de I_A donc le résultat du traitement A est significativement différent du résultat du groupe B.

Exercice 9

La fréquence des yeux bleus en France est d'environ 0,31.
On a prélevé un échantillon de 50 individus dont 15 ont les yeux bleus.

a) Quel est l'intervalle de fluctuation au seuil de 95 % ?
b) Cet échantillon est-il représentatif de la population pour ce caractère ?
c) Votre classe est elle un échantillon représentatif de la population ?

CORRIGE :

a) La proportion d'yeux bleus est CONNUE $p = 0,31$.
L'effectif de l'échantillon est $n = 50$.
On a bien : $0,2 \leq p \leq 0,8$ et $n \geq 25$
L'intervalle de fluctuation demandé est donc :
$$\mathrm{I} = \left[0,31 - \frac{1}{\sqrt{50}} ; 0,31 + \frac{1}{\sqrt{50}}\right]$$ soit approximativement $I = [0,169 ; 0,451]$

b) L'effectif de l'échantillon est $n = 50$.
$\frac{15}{50} = 0,3$ donc $\frac{15}{50} \in [0,169 ; 0,451]$: cet échantillon est représentatif de la population.

c) !!!

Exercice 10

Depuis 1996, l'accès au Musée du Louvres est gratuit le premier dimanche de chaque mois.
La proportion des visiteurs français ces jours-là est de 0,59.
Une toute nouvelle exposition est proposée. On souhaite connaître son impact sur la fréquentation des visiteurs les jours de gratuité. Sur un échantillon de visiteurs un jour de gratuité, 67 % sont français.
Décider si la nouvelle exposition a eu un impact si l'échantillon comporte :

c) 50 visiteurs ?
d) 500 visiteurs ?

CORRIGE :

a) La proportion de visiteurs français est CONNUE $p = 0,59$.
L'effectif de l'échantillon est n = 50.
On a bien : $0,2 \leq p \leq 0,8$ et $n \geq 25$
L'intervalle de fluctuation demandé est donc :
$$\mathrm{I} = \left[0,59 - \frac{1}{\sqrt{50}} ; 0,59 + \frac{1}{\sqrt{50}}\right]$$ soit approximativement I = $[0,449 ; 0,731]$
$0,67 \in [0,449 ; 0,731]$ donc cette expo n'a pas eu d'influence déterminante.

b) L'effectif de l'échantillon est n = 500.
$$\mathrm{I} = \left[0,59 - \frac{1}{\sqrt{500}} ; 0,59 + \frac{1}{\sqrt{500}}\right]$$ soit approximativement I = $[0,545 ; 0,635]$
$0,67 \notin [0,545 ; 0,635]$ donc cette exposition semble déterminante au regard de cet effectif.

Exercice 11

La fréquence d'utilisation du diesel dans les moteurs de voiture est aujourd'hui de 0,52 en France.
On a étudié des échantillons dans différents départements.

- Echantillon 1 : Taille : 30 - Fréquence : 0,6.
- Echantillon 2 : Taille : 35 - Fréquence : 0,3.

Pour chaque échantillon, indiquer si la fréquence observée appartient ou non à l'intervalle de fluctuation au seuil de 95 %.

CORRIGE :

a) La proportion de diesel est CONNUE $p = 0,52$.
L'effectif de l'échantillon 1 est n = 30.
On a bien : $0,2 \leq p \leq 0,8$ et $n \geq 25$
L'intervalle de fluctuation est donc :

$$I = \left[0,52 - \frac{1}{\sqrt{30}} ; 0,52 + \frac{1}{\sqrt{30}}\right]$$ soit approximativement $I = [0,337 ; 0,703]$

$0,6 \in [0,337 ; 0,703]$: l'échantillon 1 est représentatif de la population au seuil de 95%.

b) La proportion de diesel est CONNUE p = 0,52.
L'effectif de l'échantillon 1 est n = 35.
On a bien : $0,2 \leq p \leq 0,8$ et $n \geq 25$
L'intervalle de fluctuation est donc :

$$I = \left[0,52 - \frac{1}{\sqrt{35}} ; 0,52 + \frac{1}{\sqrt{35}}\right]$$ soit approximativement $I = [0,351 ; 0,689]$

$0,3 \notin [0,351 ; 0,689]$: l'échantillon 2 n'est pas représentatif de la population au seuil de 95%.

Exercice 12

On a programmé une fonction nommée **hasard()**, censée retourner le nombre 0 dans 50% des cas et le nombre 1 dans les autres cas.
Pour tester cette fonction, on utilise un programme basé sur l'algorithme suivant :

variable
  somme: nombre
début algorithme
  // initialisation
  somme ← 0
  // traitement
  pour i variant de 1 à 10 000
    somme ← somme + **hasard()**
  fin pour
  // sortie
  écrire "Le nombre 1 a été généré " somme " fois"
fin algorithme

1. Expliquer le fonctionnement de l'algorithme ci-dessus.
2. L'exécution de l'algorithme retourne le message "Le nombre 1 a été généré 4947 fois". Peut-on en déduire une anomalie pour la fonction **hasard()**?

CORRIGE :

1."somme ← 0" : initialise la variable somme à 0.
"pour i variant de 1 à 10 000" : on effectue une boucle 10 000 fois.
"somme ← somme + **hasard()**": on ajoute le résultat de la fonction **hasard()** à la variable somme. La variable somme ne sera pas modifiée si **hasard()** renvoie zéro. Elle sera incrémentée de 1 lorsque **hasard()** retourne 1.
La variable somme va donc compter le nombre de fois où la fonction **hasard()** retourne "1".
"écrire "Le nombre 1 a été généré " somme " fois" " : On affiche le résultat stocké dans la variable somme.
Si la fonction **hasard()** fonctionne correctement, le nombre affiché devrait avoisiner :

$$10\,000 \times \frac{50}{100} = 5\,000$$

2. On souhaite que la proportion de chiffres "1" retournés avoisine les 50% (soit une proportion de 0,5).
L'algorithme effectue 10 000 tests de la fonction **hasard()**.

On a bien : $0,2 \leq 0,5 \leq 0,8$ et $10\,000 \geq 25$
L'intervalle de fluctuation au seuil de 0,95 est donc :

$$I = \left[0,5 - \frac{1}{\sqrt{10\,000}} ; 0,5 + \frac{1}{\sqrt{10\,000}}\right] = [0,49 ; 0,51]$$

Le message retourné par l'algorithme indique une proportion de résultats "1" égale à :

$$\frac{4947}{10\,000} = 0,4947.$$

Ce nombre appartient bien à l'intervalle I.
Aucune anomalie n'a donc été détectée par l'algorithme.