FEUILLE D'EXERCICES CHAPITRE 10

Exercice 1

Une urne contient des boules de différentes couleurs dont 75% de boules rouges.
Cyril tire une boule au hasard, note la couleur et la remet dans l'urne.
Il prétend avoir effectué cette expérience 60 fois et avoir obtenu 35 boules rouges.
Son frère Paulo affirme qu'il n'a pas fait l'expérience sérieusement.
On se propose de vérifier s'il a de bonnes raisons de l'affirmer.

1) Déterminer la proportion théorique p et la taille n de l'échantillon.
2) Calculer la fréquence observée f.
3) Calculer l'intervalle de fluctuation au seuil de 95% I_f.
4) Vérifier si la fréquence observée f appartient à l'intervalle de fluctuation I_f et conclure.

Organisons l'information

1) on souhaite soumettre à un test de conformité une proportion théorique $p = 0,75$.
Pour cela on constitue un échantillon de taille $n = 60$, sur lequel on observe une fréquence de boules rouges $f_{obs} = \frac{35}{60} = 0,583$

2) $f_{obs} = 0,583$

3) Si la proportion théorique est exacte, la fréquence observée pour un échantillon de taille n doit appartenir à l'intervalle de fluctuation d'échantillonnage. $I_f = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$ avec $p = 0,75$, $n = 60$

$I_f = [0,620 ; 0,880]$

4) $f_{obs} \notin I_f$ donc avec un risque d'erreur inférieur à 5% on peut remettre en cause la réalité de la proportion théorique de 75% de boules rouges. Il se peut également que l'échantillon ne soit pas représentatif, auquel cas, un test sur un échantillon de taille plus importante s'impose.

Exercice 2

La proportion de personnes aux cheveux châtains en France est d'environ 50%.
On a observé un échantillon de 150 personnes dont 89 ont les cheveux châtains.
Cet échantillon est-il représentatif de la population ?

Organisons l'information :

on a une proportion théorique de personnes chatains en France $p = 0,5$.
On souhaite vérifier cette donnée sur un échantillon de taille $n = 150$.
On obtient une fréquence observée $f_{obs} = \frac{89}{150} = 0,593$

Il ne s'agit pas ici de vérifier l'exactitude de la donnée de p.
En revanche on va évaluer la représentativité de l'échantillon test.

Si l'échantillon est représentatif de la population française alors f_{obs} appartient à l'intervalle de fluctuation de la fréquence observée pour 95% des échantillons de taille n.

$I_f = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$ On a donc $I_f = [0,418 ; \boxed{0,582}]$

$f_{obs} \notin I_f$: L'échantillon n'est pas représentatif ; les 59,3% observés sont significativement supérieurs aux 50% théoriques

Exercice 3

Dans une classe de 37 élèves, un délégué de classe a été élu avec 60% des voix.
Parmi les 17 filles, 11 d'entre elles ont voté pour ce délégué.
Les filles sont-elles représentatives des résultats des élections de délégués ?

Organisons l'information :

On a une population de taille 37 dont on extrait une sous-population de taille $n = 17$, constituant un échantillon de la population

Dans la population on a une proportion $p = 0,6$ d'opinion favorable au délégué élu.

Dans l'échantillon on a une fréquence observée d'opinion favorable $f_{obs} = \frac{11}{17} = 0,647$

f_{obs} est-il significativement de p ?
L'échantillon de filles est-il représentatif de la population de la classe ?

Si l'échantillon de taille $n = 17$ est représentatif de la population alors f_{obs} doit appartenir à l'intervalle de fluctuation d'échantillonnage au seuil de 0,95.

$I_f = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$ avec $p = 0,6$ et $n = 17$

Notons que $n < 25$ donc l'échantillon est trop petit pour que ce test de conformité soit fiable.

$I_f = [0,357 ; 0,843]$ $f_{obs} \in I_f$ donc on ne peut rejeter l'hypothèse selon laquelle l'échantillon de filles est représentatif de la population de la classe.

Exercice 4

En 1976, dans un comté du Texas, Rodrigo Partida était condamné à huit ans de prison.
Il attaqua ce jugement au motif que la désignation des jurés de ce comté était discriminante à l'égard des Hispano-américains.
En effet, 79,1% de la population de ce comté était d'origine hispanique.
Alors que, sur les 870 personnes convoquées pour être juré, il n'y eut que 339 personnes d'origine hispanique. Qu'en pensez-vous ?

Organisons l'information

Il s'agit de vérifier si la fréquence observée $f_{obs} = \frac{339}{870} = \boxed{0,3897}$ est représentative de la proportion théorique $p = 0,791$

Si l'échantillon est représentatif de la population du comté, la fréquence observée de personnes d'origine hispanique doit appartenir à l'intervalle de fluctuation d'échantillonnage.

$I_f = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$ avec $p = 0,791$ et $n = 870$

on a donc $I_f = [\boxed{0,757} ; 0,825]$ $f_{obs} < 0,757$: f_{obs} est donc très significativement inférieure à p.

Exercice 7

Un fournisseur d'accès à Internet disposait de 25% de part de marché avant l'arrivée d'un nouveau concurrent.
Après l'arrivée de ce concurrent, il effectue une enquête sur un échantillon de 200 foyers et obtient 19% de part de marché.

Peut-il considérer que l'arrivée de ce nouveau concurrent lui a fait perdre des parts de marché.

Organisons l'information :

Suite à l'arrivée du concurrent, sur un échantillon de taille $n = 200$ le fournisseur d'accès internet souhaite estimer sa part de marché. Il obtient une fréquence observée de 19%. Dans quelle fourchette se situe le pourcentage réel de part de marché ?

On utilise l'intervalle de confiance au seuil de 95% :

$I_c = \left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}} \right]$ avec $f_{obs} = 0{,}19$, $n = 200$

$I_c = [0{,}119 ; 0{,}261]$

par défaut par excès

$0{,}25 \in I_c$ donc on ne peut pas affirmer que l'arrivée du concurrent a eu un impact sur la part de marché de ce FAI.
Pour le savoir il faudrait augmenter la taille de l'échantillon.

Exercice 10

Une urne contient un grand nombre des boules rouges et des boules blanches.
Sans les compter, on souhaiterait connaître la proportion p de boules rouges.
Pour cela, on effectue 100 tirages avec remise dans cette urne. On obtient 41 boules rouges.
Estimez p à l'aide de l'intervalle de confiance au niveau 0,95.

Organisons l'information :

En effectuant 100 tirages avec remise, on constitue un échantillon de taille $n = 100$.
On obtient une fréquence observée de boules rouges $f_{obs} = \frac{41}{100} = 0{,}41$

On utilise l'intervalle de confiance au niveau 0,95 pour estimer la proportion de boules rouges dans l'urne.

$I_c = \left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}} \right]$ avec $f_{obs} = 0{,}41$ et $n = 100$

$p \in [0{,}31 ; 0{,}51]$ Il y a entre 31% et 51% de boules rouges.

Exercice 5

Une maladie guérit naturellement dans 70% des cas. Un laboratoire souhaite tester l'efficacité d'un nouveau médicament.
Pour cela, on administre ce médicament à 500 personnes. Pour 77% d'entre elles, la guérison a eu lieu.
Que penser de l'efficacité de ce médicament ?

Organisons l'information :

on a une fréquence observée $f_{obs} = 0{,}77$ obtenue sur un échantillon de taille $n = 500$
on veut savoir si la proportion de guérison est significativement supérieure aux 70% obtenus naturellement.

On utilise l'intervalle de confiance pour estimer le taux théorique de guérison dans la population mondiale.

$I_c = \left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}} \right]$ avec $f_{obs} = 0{,}77$ $n = 500$

$I_c = [0{,}725 ; 0{,}815]$ La proportion théorique est supérieure à 70% donc le médicament est efficace

Exercice 6

Un centre commercial n'attire que 22% de clients hors de la communauté urbaine.
Souhaitant élargir sa clientèle, le centre commercial s'agrandit (nouveaux magasins, cinéma, restaurants, ...)
Après les travaux, voulant connaître l'impact de ses investissements, 300 clients sont interrogés : 72 d'entre eux habitent hors de la communauté urbaine.
Peut-on affirmer que l'agrandissement a eu un impact sur la fréquentation des clients hors communauté urbaine.

Organisons l'information

on a une proportion connue de 22% de clients hors de la zone urbaine.
Suite à des travaux, on effectue un sondage sur un échantillon de taille $n = 300$
on obtient une fréquence observée de clients hors zone urbaine $f_{obs} = \frac{72}{300} = 0{,}24$

On veut estimer la nouvelle proportion de clients hors zone urbaine
Pour cela, on utilise l'intervalle de confiance au niveau 95.

$I_c = \left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}}\right]$ avec $f_{obs} = 0{,}24$ $n = 300$

$I_c = [0{,}182 ; 0{,}298]$

donc $18{,}2\% \leq p \leq 29{,}8\%$ $0{,}22 \in I_c$ on ne peut donc pas affirmer que l'agrandissement a eu un impact sur la fréquentation des clients hors communauté urbaine

Exercice 8

Un musée national connaît une proportion de visiteurs français égale à 71%.
Durant l'été, le musée propose une exposition temporaire sur le thème de l'Egypte ancienne.
On souhaite connaître son impact sur la fréquentation des visiteurs français.
Décider si la nouvelle exposition a eu un impact dans les cas suivants :
1) Sur un échantillon de 50 visiteurs, 82% sont français.
2) Sur un échantillon de 500 visiteurs, 77% sont français.

Organisons l'information :

on a une proportion connue de 71% de visiteurs français. Suite à une exposition temporaire, on veut estimer la nouvelle proportion de visiteurs français.
Pour cela on réalise un sondage sur deux échantillons de tailles $n_1 = 50$
$n_2 = 500$

on observe les fréquences observées $f_{obs_1} = 0{,}82$ $f_{obs_2} = 0{,}77$

Les 2 échantillons permettent d'estimer $p_{théo_1}$ et $p_{théo_2}$

$I_{c_1} = \left[f_{obs_1} - \frac{1}{\sqrt{n_1}} ; f_{obs_1} + \frac{1}{\sqrt{n_1}}\right]$ donc $p_{théo_1} \in [0{,}678 ; 0{,}962]$

$I_{c_2} = \left[f_{obs_2} - \frac{1}{\sqrt{n_2}} ; f_{obs_2} + \frac{1}{\sqrt{n_2}}\right]$ $p_{théo_2} \in [0{,}725 ; 0{,}815]$

Le premier échantillon ne permet pas de statuer sur une modification de la fréquentation des visiteurs français. En revanche, le second échantillon de taille supérieure met en évidence une augmentation de la fréquentation puisque $p_{théo\,2} \geq 72{,}5\%$. C'est mieux que les 71% obtenus avant l'exposition.

Exercice 9

Léanne a lancé une pièce de monnaie 50 fois. Elle a obtenu 32 fois « pile ».
A-t-elle de bonnes raisons de s'étonner du résultat.

Organisons l'information :

on a une proportion théorique $p_{théo} = 0,5$ si la pièce est équilibrée.

Si la pièce est équilibrée, sur un échantillon de taille $n = 50$ on doit obtenir une fréquence observée f_{obs} dans l'intervalle de fluctuation d'échantillonnage $I_f = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$

or $I_f = [0,358 ; 0,642]$ et $f_{obs} = \frac{32}{50} = 0,64$

$0,64 \in I_f$ donc le résultat obtenu ne peut être considéré comme anormal, avec un risque inférieur à 5% de se tromper en affirmant cela. Un échantillon de plus grande taille permettrait de répondre avec plus d'assurance en réduisant l'incertitude due à la fluctuation d'échantillonnage.

Exercice 11

On effectue un sondage auprès de 800 personnes pour leur demander leur intention de vote aux prochaines élections. 54% d'entre elles déclarent de façon ferme vouloir voter pour le candidat A.
Le candidat A peut-il espérer être élu ?

Organisons l'information

on souhaite estimer une proportion de votants pour le candidat A. Pour cela on effectue un sondage sur un échantillon de taille $n = 800$ et on relève une fréquence observée d'intentions favorables à A $f_{obs} = 0,54$

Estimons le pourcentage théorique de suffrages en faveur de A à l'aide de l'intervalle de confiance au seuil de 0,95.

$$I_c = \left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}}\right] = [0,504 ; 0,576]$$

L'intervalle de confiance met en évidence que le candidat A peut espérer un taux supérieur à 50% : il peut raisonnablement espérer être élu au 1er tour.

Exercice 12

Le conseil municipal d'une commune organise un sondage effectué au hasard sur 400 personnes pour leur demander s'ils sont favorables à de nouveaux investissements de voirie. 235 personnes donnent un avis favorables.
Donner un intervalle de confiance au seuil de 95% de la proportion p des personnes favorables aux investissements de voirie. Que peut conclure le conseil municipal ?

Organisons l'information

Sur un échantillon de taille $n = 400$ on recueille une fréquence observée d'opinion favorable $f_{obs} = \frac{235}{400} = 0,5875$

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

$$p \in [0,5375 ; 0,6375]$$

L'intervalle de confiance met en évidence que plus de 53,75% de la population est favorable à cet investissement donc le conseil municipal peut voter en faveur de ces travaux en sachant que la population y est favorable.

Exercice 13

Pour mesurer l'audience d'une émission télévisée (audimat), on installe des appareils électroniques chez certains téléspectateurs.
Une chaîne de télévision a mesuré à partir de 1000 appareils et a relevé une audience de 31% sur le créneau 19/20h.
Donner une fourchette de l'audimat p sur ce créneau au seuil de 95%.

Organisons l'information

on mesure sur un échantillon de taille $n = 1000$ une fréquence observée d'audience $f_{obs} = 0,31$

on souhaite estimer la proportion d'auditeurs en France sur ce créneau horaire.

$$p_{théo} \in I_c = \left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}} \right] \quad \text{avec } f_{obs} = 0,31 \quad n = 1000$$

$$p \in [0,278 ; 0,342]$$: l'audimat est compris entre 27,8% et 34,2%

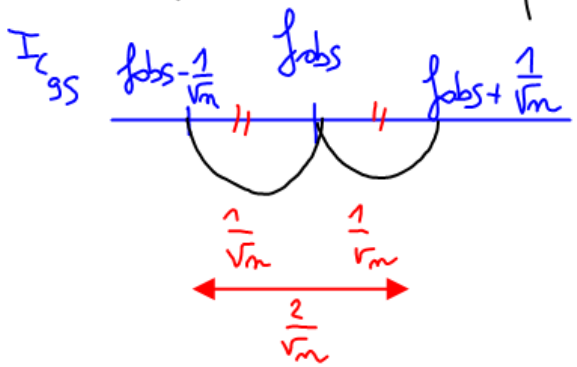
Exercice 14

On veut estimer la proportion p de jeunes de 10 à 15 ans disposant d'un forfait mobile.
On sait que p est compris entre 50% et 70%.
Quelle doit être la taille minimale de l'échantillon pour obtenir un résultat avec une précision de 1% au seuil de 0,95. Une telle enquête est-elle envisageable ?

Organisons l'information

on effectue un sondage sur un échantillon de taille n.
On recueille une fréquence observée f_{obs}.

On peut estimer la proportion théorique p à l'aide de l'intervalle de confiance. On sait que $p \in \left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}} \right]$



On cherche n tel que $\frac{1}{\sqrt{n}} \leq 1\%$

$$\frac{1}{\sqrt{n}} \leq 0,01 \iff \frac{1 \times \cancel{\sqrt{n}}}{\cancel{\sqrt{n}}} \leq 0,01 \times \sqrt{n}$$

$$\iff 1 \leq 0,01 \times \sqrt{n}$$

$$\iff \frac{1}{0,01} \leq \sqrt{n}$$

$$\iff 100 \leq \sqrt{n}$$

$$\iff 100^2 \leq n$$

d'où $n \geq 10000$

La taille de l'échantillon doit dépasser 10000.
C'est possible mais peu courant !

Méthode :
- L'intervalle de confiance est centré sur f_{obs} avec une précision $\pm \frac{1}{\sqrt{n}}$
- La largeur de l'intervalle de confiance est $\frac{2}{\sqrt{n}}$