Intervalles de fluctuation et de confiance

INTERVALLE DE FLUCTUATION

Propriété (Rappel de première et seconde

On se place dans la situation suivante :
La proportion d'individus possédant un certain caractère dans une population est supposée connue.
On note cette proportion : $\boxed{p}$

Dans 95% des échantillons de taille n choisis au hasard la proportion observée (notée $f_{échantillon}$ sera comprise entre $p - \frac{1}{\sqrt{n}}$ et $p + \frac{1}{\sqrt{n}}$.

$$\boxed{f_{échantillon} \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]}$$

Cette propriété n'est applicable que lorsque les conditions suivantes sont vérifiées :
$n \geq 30$; $n \times p \geq 5$ et $n \times (1 - p) \geq 5$
Cet intervalle est appelé **intervalle de fluctuation au seuil 95%**

Exemple d'application

On estime en France que 65% de la population part en vacance d'été .

1 On choisit 1000 Français au hasard .
Donner l'intervalle de fluctuation au seuil 95% associé à cette problématique

2 On a interrogé cette fois mille habitants de Paris.
Dans cet échantillon, la proportion de personnes partant en vacances d'été est $0,72$ (ou 72%.
Interpréter ce résultat.

Propriété (Conclusion

Il n'y a que 5% de chance que le hasard soit seul responsable de ce résultat.
En effet seuls 5% des échantillons de la population française donnerait de tels résultats.
Il semble que la population parisienne ne soit pas significative de la population française.
On peut affirmer :
La proportion de parisiens partant en vacances d'été est sans doute (5 % de risque) plus importante que dans la population en général

Prise de décision

Dans cette partie la proportion du caractère étudié n'est pas connue.
Mais on peut établir des conjectures sur cette proportion. On se sert alors de l'intervalle de fluctuation correspondant à cette hypothèse de proportion pour valider ou invalider cette hypothèse.

Exemple

Un fabricant d'alarme commande auprès de son fournisseur deux types de composants électroniques : C_1 et C_2.
Il commande 9000 composants de chaque sorte.

Au moment de la réception de la livraison, le service de contrôle prélève au hasard 50 composants et constate que 19 sont des modèles C_1.

Le responsable s'inquiète et se demande si la proportion 9000-900 est respectée ?
A-t-il raison de s'inquiéter et doit-il renvoyer la commande.

Raisonnement :
Étape 1 : **Hypothèse :** On veut valider ou invalider l'hypothèse suivante :
La commande est respectée par le fournisseur , c'est à dire $p = 0,5$
(il y a autant de composants de chaque sorte)
Étape 1 : **Intervalle de fluctuation :** On a $p = 0,5$ et $n = 50$.
L'intervalle de fluctuation est : $\dots\dots\dots\dots\dots\dots$
Étape 3 : **Fréquence observée** La proportion observée dans notre échantillon est : $\dots\dots\dots$
Étape 4 : **Prise de décision** .

Propriété (Prise de décision

f est la fréquence **observée** d'un échantillon de taille n.
I est l'intervalle de fluctuation asymptotique au seuil 95%.
On fait l'hypothèse : *La proportion est* p

- Si $p \in I$, alors on accepte l'hypothèse.
- Si $p \notin I$, alors on rejette l'hypothèse.

Exercice

Exercice 1

Lors d'un conseil municipal, le maire de la ville affirme je pense que :
« *70 % des habitants sont favorables à la construction d'une médiathèque* »
Une enquête est alors effectuée afin de vérifier l'affirmation du maire.
On interroge 200 habitants de la ville .
124 de ces habitants se déclarent favorable à la construction de cette médiathèque.

1 Quelle est ici la valeur de p ? $p = \cdots\cdots\cdots$

2 L'intervalle de fluctuation est donc : $I = \cdots\cdots\cdots$

3 La proportion observée dans notre échantillon est : $f = \cdots\cdots\cdots$

4 Doit on valider ou invalider l'affirmation du maire ?

Exercice 2

Un publicitaire se réjouit :
« *Après notre campagne publicitaire ,une personne sur 2 connaîtra ce nouveau produit* ».
La campagne publicitaire étant terminée, l'entreprise souhaite vérifier l'affirmation du publicitaire.
On choisit au hasard un échantillon de 624 personnes.
47% des personnes interrogées déclarent connaître le produit.
Á partir de cet échantillon peut-on accepter au risque 5%l'affirmation du publicitaire.

Exercice 3

Dans une entreprise, lorsque le processus de fabrication est bien mis en œuvre 8% des objets fabriqués ont un défaut de couleur.
Un contrôleur de l'entreprise prélève, dans la journée, 250 objets fabriqués.
Parmi ceux ci 40 ont un défaut de couleur

1 Que peut on dire au risque 5% du processus de fabrication ?

2 Le chef d'atelier décide alors de modifier le réglage des machines mélangeuses de couleur.
Cette fois sur un nouvel échantillon de 250 seulement 30 objets ont un défaut de couleur.
Le réglage du chef d'atelier a-t-il été efficace ?

Exercice 4

Un producteur de poire affirme à une enseigne de grande surface que 80% de sa production. est de premier choix.
L'enseigne décide d e tester cette affirmation.
Sur un échantillon choisi au hasard de 320 poires 28% sont de premier choix.
AÁ partir de cet échantillon, peut-on accepter au risque 5% l'hypothèse qu 80% de la production est de premier choix.
Autrement dit, l'enseigne peut-elle exiger une baisse de prix au producteur ?

Intervalle de confiance

Rappel de première et seconde

On se place cette fois dans une situation dans laquelle on ne connaît pas la proportion réelle d'un certain caractère dans une population.
On ne peut par ailleurs pas faire d'hypothèse à priori sur cette proportion *On serait alors dans le cas précédent.*
On cherche alors à évaluer cette **proportion réelle** à partir de la **proportion observée** dans un échantillon.
Le meilleur exemple de cette situation est le sondage politique. La propriété précise est donnée ci-dessous

Propriété

On ne connaît pas la proportion p d'apparition d'un caractère dans une population.
On connaît par contre sa fréquence $f_{échantillon}$ d'apparition dans un échantillon de taille n .
Il y a 95% de chance pour que la valeur de p se trouve dans l'intervalle

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

Il s'agit cette fois de $\boxed{\text{l'intervalle de } \boxed{\text{confiance}} \text{ au niveau de confiance 95\%}}$.
Cette méthode est valable dès que n est suffisamment grand

Exemple

On effectue un sondage d'opinion avant une élection. La question posée est : "Voterez vous pour le candidat A ?"
On interroge 1000 personnes .

- 320 personnes répondent OUI
- 600 personnes répondent NON
- 80 personnes s'abstiennent .

La fréquence **observée** de réponses favorables au candidat A est : · · · · · · · · · · · ·
Que peut on dire de la **proportion réelle** de personnes favorables au candidat A dans la population totale ?