Intervalles de confiance

I Exemple introductif

On lance 100 fois de suite une même pièce de monnaie. On s'intéresse à la fréquence des « Pile » obtenue ; on trouve 0,42.
La pièce est-elle truquée ? (voir paragraphe II)

II Estimation d'une proportion, intervalle de confiance

Définition

Une expérience aléatoire est une expérience renouvelable dont les résultats possibles sont connus sans qu'on puisse déterminer lequel sera réalisé.
Un échantillon de taille n est constitué des résultats de n répétitions identiques et indépendantes de la même expérience.

On étudie un caractère statistique. On suppose que l'on ne connaît pas la proportion **réelle** p de celui-ci et l'on cherche à l'**estimer**.

Propriété

On considère un échantillon de taille n ($n \geqslant 30$) tel que $np \geqslant 5$ et $n(1-p) \geqslant 5$. (f est la fréquence **observée** dans l'échantillon).
Alors, la proportion **réelle** p du caractère appartient à l'intervalle de confiance $\left[f-\dfrac{1}{\sqrt{n}}\ ;\ f+\dfrac{1}{\sqrt{n}}\right]$ au **seuil de confiance** de 95 %, c'est-à-dire avec une probabilité de 95 %.
Cet intervalle est dit **intervalle de confiance** de p au seuil de confiance 0,95 (95 %).

Retour sur l'exemple introductif :
$n = 100 \geqslant 30$; $f = 0,52$ donc $nf = 42 \geqslant 5$ et $n(1-f) = 0,58 \geqslant 5$.
On peut utiliser intervalle de confiance au seuil 0,95.

$$I\left[\frac{1}{\sqrt{n}}\ ;\ f+\frac{1}{\sqrt{n}}\right] = \left[0,52-\frac{1}{100}\ ;\ 0,42+\frac{1}{10}\right) = [0,32\ ;\ 0,52].$$

0,5 appartient à cet intervalle donc on peut penser que le dé n'est pas truqué (avec un risque d'erreur de 5 %)

Exemple 1 Lors d'une élection mettant en lice deux candidats, le parti d'un candidat commande un sondage réalisé à partir de 1 600 personnes à l'issue duquel il est donné gagnant avec 52 % des voix.
A-t-il des raisons d'être confiant ?

Réponse :
Soit p la proportion de gens votant pour lui ; il est élu si $p \geqslant 0,5$.
D'après l'intervalle de confiance, $p \in \left[f-\dfrac{1}{\sqrt{1\,600}}\ ;\ f+\dfrac{1}{\sqrt{1\,600}}\right] = [0,495\ ;\ 0,545]$ au seuil de 95%, donc il n'est pas sûr d'être élu.
En effet, pour être élu, il doit obtenir 50 % des voix, donc $p = 0,5$; or la borne inférieure de l'intervalle de confiance est inférieure à 0,5, donc il n'est pas complètement sûr d'âtre élu.
Pour être plus spur, il faudrait interroger davantage de candidats. La largeur (amplitude) de l'intervalle de confiance est $\dfrac{2}{\sqrt{n}}$, donc diminie si n augmente (mais le sondage coûte alors plus cher ...)