

Échantillonnage et estimation

I. Échantillonnage

- Un échantillon de taille n est une sélection de n individus choisis au hasard dans une population.
- Sur plusieurs échantillons de même taille, la fréquence d'un caractère observée varie d'un échantillon à l'autre. C'est ce qu'on appelle la **fluctuation d'échantillonnage**.

II. Intervalle de fluctuation

Au sein d'une population, on connaît la proportion p d'un certain caractère.

On prélève un échantillon de taille n et on note f la fréquence du caractère dans l'échantillon.

Si $0,2 \leq p \leq 0,8$ et $n \geq 25$ alors, dans au moins 95% des cas, f appartient à l'intervalle :

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

Cette intervalle est appelé **intervalle de fluctuation au seuil de 95%**.

Proportion p du caractère étudié

Échantillon de taille n

Fréquence f du caractère étudié

III. Intervalle de confiance

Au sein d'une population, la proportion p d'un certain caractère est inconnue.

On prélève un échantillon de taille n et on note f la fréquence du caractère dans l'échantillon.

Si $0,2 \leq f \leq 0,8$ et $n \geq 25$ alors, dans au moins 95% des cas, p appartient à l'intervalle :

$$\left[f - \frac{1}{\sqrt{n}}; f + \frac{1}{\sqrt{n}} \right]$$

Cet intervalle est appelé **intervalle de confiance**.

Méthode : utilisation d'un intervalle de confiance pour un sondage

Avant les élections, le candidat A commande un sondage effectué sur 250 personnes. 138 personnes interrogées déclarent avoir l'intention de voter pour le candidat A.

Le candidat A peut-il espérer être élu ?

Soit p la proportion théorique d'électeurs pour le candidat A.

La fréquence observée est égale à $f = \frac{138}{250} = 0,552$.

L'intervalle de confiance de p au seuil de 0,95 est $\left[0,552 - \frac{1}{\sqrt{250}}; 0,552 + \frac{1}{\sqrt{250}} \right]$, soit de façon approchée $[0,48; 0,62]$.

On a donc $0,48 < p < 0,62$. Or, pour être sûr que le candidat A soit élu, il faudrait que $p > 0,5$. Il est donc possible que le candidat A ne soit pas élu.