ÉCHANTILLONNAGE

I.Modélisation de la situation :

Définitions :

Lorsqu'on étudie une partie de la population, on dit qu'on étudie un **échantillon**.

Le nombre d'individus formant l'échantillon est appelé **taille de l'échantillon**. Nous le noterons n.

Notations :

On note p la probabilité associée au critère étudié sur une population.

On note f la fréquence observée du critère sur l'échantillon de taille n.

Remarques :

En fait, p est la proportion d'individus associés au critère dans la population totale, alors que f est la proportion d'individus associés au critère dans l'échantillon.

Plus la taille n de l'échantillon est grande, plus f se rapproche de p.

II. Intervalle de fluctuation :

Dans cette partie, on connaît la probabilité p associée au critère étudié sur une population et on veut estimer la fréquence f associée à ce critère dans un échantillon.

Définition :

Si $n \geqslant 25$ et $p \in [0{,}2 ; 0{,}8]$ dans plus de 95% des cas, f appartient à l'intervalle $\left[p-\frac{1}{\sqrt{n}} ; p+\frac{1}{\sqrt{n}}\right]$, appelé **intervalle de fluctuation**.

Remarque :

On parle aussi d'intervalle de fluctuation au seuil de 95% ou 0,95.

Exemple :

On sait que la proportion de foyers français possédant un chien est de 26,4 %. Lorsqu'on choisit un foyer au hasard, la probabilité qu'il possède un chien est donc de 26,4 %.

Un village compte 86 foyers, on veut estimer le nombre de chiens.

On a bien $n \geqslant 25$ et $p \in [0{,}2 ; 0{,}8]$, il y a donc 95 % de chance que la proportion de foyers ayant un chien appartienne à l'intervalle $\left[0{,}264-\frac{1}{\sqrt{86}} ; 0{,}264+\frac{1}{\sqrt{86}}\right] \approx [0{,}156 ; 0{,}372]$. Or $0{,}156 \times 86 \approx 13$ et $0{,}372 \times 86 \approx 32$. On peut donc estimer, au seuil des 95 %, qu'il y a entre 13 et 32 chiens au village.

Remarque :

Plus la taille de l'échantillon est grande, plus les estimations sont précises. Par exemple, si on avait pris une ville de 10 000 foyers, on aurait eu une proportion de foyers ayant un chien dans l'intervalle [0,256 ; 0,276].