

Échantillonnage - Classe de 2nde

I - Échantillon

Définition :

On appelle **échantillon de taille n** la série statistique formée des résultats obtenus lorsqu'on répète n fois une expérience dans les mêmes conditions.

Exemple : Dans la population française, on interroge 100 personnes prises au hasard sur la nature de leur groupe sanguin.

Cette expérience est renouvelée dix fois. Le tableau ci-dessous indique la fréquence du nombre de personnes étant du groupe O positif pour chaque échantillon de taille 100 :

Échantillon n°	1	2	3	4	5	6	7	8	9	10
Fréquence d'apparition du groupe O^+	0,37	0,39	0,47	0,36	0,44	0,35	0,43	0,38	0,51	0,40

Remarques :

- les distributions des fréquences varient d'un échantillon à l'autre malgré le fait que les échantillons soient de même tailles, c'est ce qu'on appelle la fluctuation d'échantillonnage.

- lorsque la taille des échantillons augmentent, les distributions de fréquences ont tendance à se stabiliser comme nous allons le voir dans la partie suivante.

II - Fluctuation d'échantillonnage

Propriété : Considérons un caractère dont la proportion p dans une population donnée est comprise entre 0,2 et 0,8.

Alors, pour $n \geq 25$, 95% des échantillons de taille n auront une fréquence située dans l'intervalle :

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

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

Exemple : La proportion de personnes du groupe O^+ dans la population française est en réalité de 0,39.

Avec un échantillon de 100 personnes, on a donc 95% de chances d'obtenir une fréquence comprise entre 0,29 et 0,49 ce que semble corroborer le tableau précédent puisqu'on a obtenu qu'un échantillon dont la fréquence était en dehors de l'intervalle [0,29;0,49].

Sur la figure ci-dessous qui reprend les données de notre exemple, nous pouvons constater effectivement que plus la taille de l'échantillon augmente, plus la fréquence se stabilise et s'approche de 0,39 :



