



Comment estimer la proportion d'un caractère dans une population à partir d'un échantillon ?

L'intolérance au lactose est un problème de digestion du sucre contenu dans le lait et ses produits dérivés, le lactose. En effet, les personnes intolérantes au lactose ne fabriquent pas de lactase, l'enzyme capable de découper ce sucre. Il en résulte des symptômes digestifs bénins mais inconfortables. Chez tous les mammifères, la production de lactase disparaît après le sevrage. Mais chez l'homme, des mutations apparues au cours de l'évolution ont permis le maintien de la synthèse de cette enzyme à l'âge adulte, sachant que ces mutations sont assez hétérogènes selon les populations.

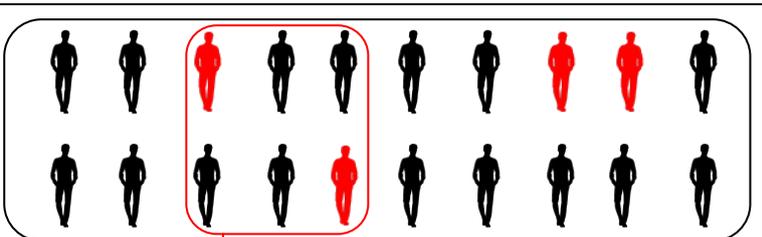


Léonard Hofstadter, un personnage de la série the big bang theory intolérant au lactose.

Problème : On cherche à connaître la proportion de Français intolérants au lactose avec l'intervalle de confiance le plus réduit possible.

On a commandé 4 études statistiques pour connaître la proportion de Français intolérants au lactose. Les résultats de ces études sont présentés dans le tableau ci-dessous.

Nom de l'étude	Z51	M12	R19	Z6PO
Population totale échantillonnée	199	302	412	538
Nombre de personnes intolérantes au lactose	41	81	121	132



Echantillon de n individus

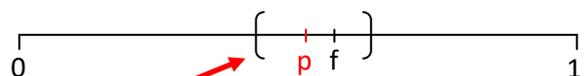
f = fréquence connue du caractère dans l'échantillon.

Population

p = proportion inconnue du caractère « intolérant » dans la population

La fréquence f d'un caractère dans un échantillon n peut être calculée à partir de cet échantillon de la population. La valeur de f permet ensuite d'estimer la proportion p de ce caractère dans la population avec une incertitude. On précise cette incertitude avec un intervalle de confiance (IC). Ainsi, p à 95% de chance d'être compris dans cet intervalle de confiance, exprimé selon le théorème de Moivre-Laplace :

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



Cet intervalle de confiance contiendra ainsi la valeur p dans 95% des cas.

Guide pour la résolution de l'activité.

Etape 1 : Calculer les fréquences d'apparition du caractère « intolérant au lactose » et l'intervalle de confiance pour les 4 études.

Etape 2 : Quelle étude présente le plus d'intérêt ? Pourquoi ?

Etape 3 : Conclure. Donner la proportion de personnes intolérantes au lactose dans la population française.

Si j'ai fini...

Donner un argument quant à l'avantage évolutif que fournit la capacité de digérer le lactose dans une société agricole.