

Echantillonnage

Échantillonnage et prise de décision

Échantillonnage : on considère une population pour laquelle on pense que la proportion d'un certain caractère est p . On prend un échantillon en choisissant simultanément n personnes dans la population. Le nombre d'individus X de l'échantillon possédant le caractère étudié suit une loi binomiale de paramètres n et p . On note f la fréquence constatée du caractère dans l'échantillon. On cherche un critère permettant de valider l'hypothèse faite sur la proportion.

Intervalle de fluctuation asymptotique au seuil 0,95 de la fréquence f : c'est l'intervalle :

$$I = \left[p - 1,96 \times \frac{\sqrt{p(1-p)}}{\sqrt{n}}; p + 1,96 \times \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$$

Si f est dans I , on est sûr à 95% que l'hypothèse était valide.

Estimation et intervalle de confiance

On considère une population pour laquelle on ne connaît *pas* la proportion p d'un certain caractère. On prend un échantillon en choisissant simultanément n personnes dans la population. Le nombre d'individus X de l'échantillon possédant le caractère étudié suit une loi binomiale de paramètres n et p . On note f la fréquence constatée du caractère dans l'échantillon. On cherche une estimation de p à partir de la fréquence f observée dans l'échantillon.

Théorème : on est sûr à 95% que p se trouve dans l'intervalle (dit « *intervalle de confiance de p à 95 %* ») :

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