Statistiques

1 Intervalle de fluctuation

Si la variable aléatoire X_n suit une loi binomiale $\mathscr{B}(n,p)$ et si l'on se trouve dans les conditions de l'approximation normale de la loi binomiale ($n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$), on définit alors l'**intervalle de fluctuation asymptotique au seuil de 95 %** par :

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

Cette intervalle peut éventuellement être simplifié par :

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

2 Prise de décision

Soit f_{obs} la fréquence d'un caractère observée d'un échantillon de taille n d'une population donnée. On suppose que les conditions de l'approximation normale de la loi binomiale sont remplies : $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$.

Hypothèse :
La proportion du caractère étudié dans la population est p.

Soit I_n l'intervalle de fluctuation asymptotique au seuil de 95 %.

- Si $f_{\text{obs}} \in I_n$; on ne peut rejeter l'hypothèse faite sur p.
- Si $f_{\text{obs}} \notin I_n$; on rejete l'hypothèse faite sur p.

3 Estimation - Intervalle de confiance

Pour des raisons de coût et de faisabilité, on ne peut étudier un certain caractère sur l'ensemble d'une population. La proportion p de ce caractère est donc inconnue.

On cherche alors à estimer p à partir d'un échantilon de taille n. On calcule alors la fréquence f_{obs} des individus de cet échantillon ayant ce caractère.

On observe la fréquence f_{obs} sur un échantillon de taille n. On appelle **intervalle de confiance de 95%** l'intervalle :

$$\left[f_{obs} - \frac{1}{\sqrt{n}} \, ; \, f_{obs} + \frac{1}{\sqrt{n}} \right]$$

Si l'on souhaitre encadrer p dans un intervalle de longueur a, on doit avoir : $n \geqslant \dfrac{4}{a^2}$