Échantillonnage

Dans de nombreuses situations, l'étude complète d'un caractère doit être remplacée par une approche à l'aide d'un échantillon de taille *n*.
Pour déduire de cet échantillon des propriétés de la population totale, on étudie les propriétés de l'ensemble des échantillons de taille *n* appelé **échantillonnage de taille *n***.

1. Intervalle de fluctuation d'une proportion

La variable aléatoire F qui, à tout échantillon aléatoire prelevé avec remise et de taille n, associe la proportion du caractère f considéré dans l'échantillon suit approximativement la loi normale $N(m; \sigma)$ avec

$$m = p \quad \text{et} \quad \sigma = \sqrt{\frac{p(1-p)}{n}}$$

où p est la proportion du caractère dans la population.

Exemple. *Un joueur de tennis sert une première balle bonne six fois sur 10. Déterminer l'intervalle de fluctuation de la proportion de premières balles bonnes sur une série de 120 engagements, au seuil de 95 %.*

- Population : les premières balles servies
- Caractère : être bonne
- Proportion du caractère dans la population : $p = \frac{6}{10}$
- Taille des échantillons : $n = 120$

La variable aléatoire F qui, à toute série de 120 engagements, associe la proportion f de premières balles bonnes suit approximativement la loi normale $N(m;\sigma)$ avec

$$m = p = 0,6 \quad \text{et} \quad \sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0,6\times 0,4}{120}}$$

Il s'agit, connaissant la proportion p, de déterminer un intervalle dans lequel on a une probabilité connue de trouver la valeur du caractère f des échantillons.

L'intervalle de fluctuation de f au seuil $1-\alpha$ est :

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

avec u_α le réel tel que $P(-u_\alpha \leq Z \leq u_\alpha) = 1-\alpha$ où Z suit la loi normale $N(0;1)$.

Valeurs courantes :

si $1-\alpha = 0,99 \Rightarrow u_\alpha = 2,58$

si $1-\alpha = 0,95 \Rightarrow u_\alpha = 1,96$

si $1-\alpha = 0,9 \Rightarrow u_\alpha = 1,645$

Pour l'exemple du joueur de tennis on a :

- Seuil de 95% $\Rightarrow 1-\alpha = 0,95 \Rightarrow u_\alpha = 1,96$

L'intervalle de fluctuation de f au seuil de 95% est donc

$$I = \left[0,6 - 1,96 \sqrt{\frac{0,6\times 0,4}{120}} \; ; \; 0,6 + 1,96\sqrt{\frac{0,6\times 0,4}{120}} \right] = [0,51 \, ; \, 0,69]$$