

Méthodes Statistiques

Corrigé de l'exercice 26

Sur un échantillon de 1000 amateurs de café, 300 personnes disent préférer le *robusta* à l'*arabica*.

26-1) Donner un intervalle de confiance au risque 3% pour la proportion d'individus préférant le *robusta*.

On estime la proportion p d'individus préférant le *robusta* dans la population des amateurs de café au moyen de la proportion empirique dans l'échantillon :

$$F_n = \frac{300}{1000} = 0.3$$

Comme on dispose d'un grand échantillon, on fait l'approximation suivante concernant la loi de probabilité de F_n :

$$\sqrt{n} \frac{F_n - p}{\sqrt{F_n(1 - F_n)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

Cela permet d'écrire l'intervalle de confiance :

$$IC(X_1, \dots, X_n) = \left[F_n \pm \sqrt{\frac{F_n(1 - F_n)}{n}} u_c \right]$$

où u_c est la valeur critique de la loi normale $\mathcal{N}(0, 1)$ au seuil α (ici $\alpha = 0.03$) :

$$u_c = Q_Z \left(1 - \frac{\alpha}{2} \right) = 2.17$$

On trouve finalement l'intervalle de confiance :

$$\left[0.3 \pm \sqrt{\frac{0.3(1 - 0.3)}{1000}} \times 2.17 \right] = \left[0.3 \pm 0.031 \right] = \left[0.269 ; 0.331 \right]$$

26-2) Donner un intervalle de confiance au risque 3% pour la proportion d'individus préférant l'*arabica*.

On refait les calculs en échangeant les amateurs de *robusta* et les amateurs d'*arabica*.

La proportion empirique des individus préférant l'*arabica* dans l'échantillon est :

$$F_n = \frac{700}{1000} = 0.7$$

On trouve finalement l'intervalle de confiance :

$$\left[0.7 \pm \sqrt{\frac{0.7(1 - 0.7)}{1000}} \times 2.17 \right] = \left[0.7 \pm 0.022 \right] = \left[0.678 ; 0.722 \right]$$