INTERVALLE DE CONFIANCE SUR LES MOYENNES

Source : Wannacott et Wannacott, Statistique : Economie, Gestion, Sciences , Médecine, Economica, 3° éd., 1984

Intervalle de confiance

Cadre de travail :
- on construit un échantillon aléatoire de n individus (soit la loi des X_i est gaussienne, soit n est grand)
- on calcule la moyenne $\overline{X}$ de cette échantillon
- on connaît l'écart-type s de la population

Objectif : on désire construire un intervalle de confiance, c'est-à-dire déterminer la plage de valeurs dans laquelle, à un niveau de confiance donné, se situe la moyenne m de la population compte tenu de l'observation effectuée.

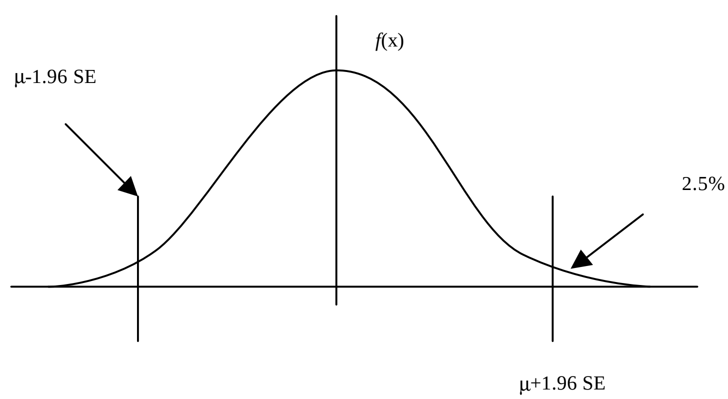
Solution :

$$\overline{X} \pm z_{0.025} SE$$

où :

- $SE = \dfrac{s}{\sqrt{n}}$
- le seuil de confiance est de 95%
- z est la loi normale standardisée

Pour un niveau de confiance de 95%, $z_{0.025}$ a pour valeur 1.96. Graphiquement, la situation est la suivante :



En général, σ est bien entendu inconnu. On doit alors l'estimer à partir de l'écart-type de l'échantillon.

Cadre de travail :
- on construit un échantillon aléatoire de n individus (soit la loi des X_i est gaussienne, soit n est grand)
- on calcule la moyenne $\bar{X}$ de cette échantillon
- on ne connaît pas l'écart-type s de la population et on l'estime à partir de l'échantillon (S).

Objectif : on désire construire un intervalle de confiance, c'est-à-dire déterminer la plage de valeurs dans laquelle, à un niveau de confiance donné, se situe la moyenne m de la population compte tenu de l'observation effectuée.

Solution :

$$\bar{X} \pm t_{0.025} SE$$

où :

- $SE = \frac{S}{\sqrt{n}}$ et $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
- le seuil de confiance est de 95%
- t est la valeur obtenue pour un niveau de 0.025 dans la loi de student avec n - 1 degré de liberté.

Pour un niveau de confiance de 95%, t a pour valeur 2.09 (20 ddl, si n=21).

NB : Intervalle de confiance sur une proportion

Pour de grands échantillons, sachant que l'espérance de P (la proportion calculée dans l'échantillon) est π (la proportion dans la population) et que son écart-type σ_p est $\sqrt{\frac{p(1-p)}{n}}$, on construit l'intervalle de confiance comme suit : $p = P \pm \sqrt{\frac{P(1-P)}{n}}$. Cette approche repose sur le fait que lorsque n devient grand, la distribution de P prend une forme de plus en plus proche de la normale.

Comparaison de deux moyennes

Cadre de travail :
- on construit deux échantillons aléatoires indépendants de n individus (soit la loi des X_i est gaussienne, soit n est grand)
- on calcule les moyennes $\bar{X}_1$ et $\bar{X}_2$ de ces échantillons
- on connaît les écart-types s_1 et s_2 des populations en question

Objectif : on désire construire un intervalle de confiance sur la différence entre les deux moyennes

Solution : $(m_1 - m_2) \in (\bar{X}_1 - \bar{X}_2) \pm z_{0.025}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ si l'on travaille à un niveau de confiance de 95%.

Si les écart-types des populations sont inconnus et si l'on peut supposer que les variances des populations sont égales, on utilise la loi de student avec ddl = $(n_1\text{-}1)+(n_2\text{-}1)$:

$$(m_1 - m_2) = (\overline{X}_1 - \overline{X}_2) \pm s_p t_{0.025} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\text{avec } s_p^2 = \frac{\sum (X_1 - \overline{X}_1)^2 + \sum (X_2 - \overline{X}_2)^2}{(n_1 - 1) + (n_2 - 1)}$$