Intervalle de confiance

Othmane AJDOR

2018-2019

Table des matières

1 Intervalle de confiance

1.1 Sondage

Comment modéliser les réponses obtenues lors d'un sondage ?
Soit n= nombre de sondés à 2 choix -> généralisation

Pour la personne n_i, $X_i = 1$ ou 0 (p ou $1 - p$)

Supposant les elections entre Le Pen et Macron, si le sondage est effectué après un rally de Macron et aux personnes qui étaient présentes, le sondage n'aura pas de sens.

On fait l'hypothese que ce qu'une personne pense est different de ce pense une autre. De ce fait, elles sont indépendantes.

Le résultat obtenu après le sondage peut êtremodelisé par $\mathbf{X} = \sum_{i=1}^{n} X_i$ où X represente le nombre de X_i à 1 sur les n.
Cette variable aléatoire suit une loi binomiale $B(n; p)$.
Soit $n = 1000$ et $X = 600$, alors on estime p à l'aide de :

$E(X) = np$
$E(X_i) = np$
$E(X_i) \approx \frac{1}{n} \sum_{i=1}^{n} x_i$, x_i sont les valeurs des X_i.
$P_{est} = \frac{1}{n} \sum_{i=1}^{n} x_i$

1.2 definition

On appelle intervalle de confiance de seuil α pour un parametre θ inconnu un intervalle I dont au moins une des bornes est une variable aleatoire reelle telle que : $P(\theta \in I) = 1 - \alpha$.

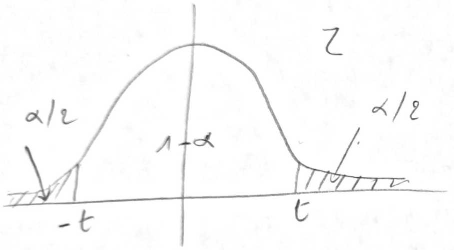
Plus le seuil est faible, plus l'intervalle est grand.
On prend souvent α à 5% ou alors 1% pour les risques sanitaires où on veut limiter les risques.

1.2.1 Exemple : IC pour la moyenne d'une loi normale

Soit $n = 20$ $\bar{x_n} = 64$ $s_n = 5 = \frac{1}{n} \sum_{i=1}^{n} x_i$.

IC est obtenu à l'aide d'une fonction pivotale. C'est une VA qui depend des choses qui nous interessent, ici n mais dont on connait parfaitement les lois. Elle nous permet de faire des calculs de probabilités.

Si les X_i $N(m, \sigma^2)$ et sont independantes, alors $Z = \sqrt{(n-1)}\frac{\bar{X_n}}{S_n}$ suit une loi de Student(n-1).



$I / \mathbb{P}(m \in I) = 1 - \alpha$

$\mathbb{P}(-t < Z < t) = 1 - \alpha$

$\mathbb{P}\left(\overline{X_m} - \frac{tS_n}{\sqrt{n-1}} < m < \overline{X_m} + \frac{tS_n}{\sqrt{n-1}}\right) = 1 - \alpha$

Soit $n = 20$, $\bar{x_n} = 64$, $s_n = 5$, $t_{5\%} = 2.1$, $t_{5\%} = 2.9$

$\mathbb{P}\left(\frac{tS_m}{\sqrt{m-1}} < \overline{X_n} - m < \frac{tS_n}{\sqrt{m-1}}\right) = 1 - \alpha$

$\mathbb{P}\left(\overline{X_m} - \frac{tS_n}{\sqrt{m-1}} < m < \overline{X_n} + \frac{tS_n}{\sqrt{n-1}}\right) = 1 - \alpha$

$S_m = 5$
$\bar{x}_m = 64$
$m = 20$

$t_{5\%} = 2,1$

$t_{1\%} = 2,9$

$64 - \frac{2,1 \times 5}{\sqrt{20-1}} = 61,59$

$64 - \frac{2,1 \times 5}{\sqrt{20-1}} = 66,4$

$IC_{5\%} = [61,59, 66,4]$

$IC_{1\%} = [60,67, 67,32]$

$IC_{1\%} = [60.7, 67.3]$ $\qquad$ $IC_{5\%} = [61, 66]$