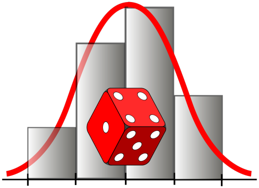
Estimation paramétrique par intervalle de confiance



Module : Techniques d'estimation pour l'ingénieur

Intervalle de confiance de la proportion

Construire un intervalle de confiance de la proportion est la détermination d'un intervalle pour le paramètre $p \in]0, 1[$ de la loi de Bernoulli au vu d'un échantillon $(X_1, \dots, X_n) \sim \mathcal{B}(p)$.

De ce fait, une proportion n'est que la fréquence de la valeur 1 dans l'échantillon.

On rappelle qu'on a déjà montré qu'un estimateur ponctuel de p est

$$\widehat{p}_n = \overline{X}_n = \frac{X}{n}$$

où X est le nombre de fois où le caractère apparaît dans l'échantillon de taille n.

Or pour un échantillon qui n'est pas normal, la loi de la statistique $\overline{X}_n$ n'est pas évident de la trouver et par la suite la détermination de l'intervalle de confiance n'est plus possible.
Mais lorsque n est suffisamment grand, en faveur du Théorème Central Limite (T.C.L) , on admet le résultat suivant:

Théorème 5

Si $np > 5$ et $n(1-p) > 5$ (ou n assez grand), alors l'intervalle de confiance de niveau de signification $1-\alpha$ pour une proportion p se présente comme suit:

$$IC(p) = \left[\widehat{p}_n - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}, \widehat{p}_n + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}\right]$$

Exemple

Sur 500 personnes interrogées, 274 ont déclaré qu'elles voteraient pour le candidat A.

1. Donner une estimation de p la proportion de personnes favorables au candidat A dans la population par intervalle de confiance au niveau de signification 95%.
2. Pour quel degré de confiance a-t-on la borne inférieure exactement égale à 50% ?

Intervalle de confiance de la proportion

Solution

On a $n = 500$, $\alpha = 1 - 0.95 = 0.05 \longrightarrow \frac{\alpha}{2} = 0.025$

La proportion dans l'échantillon est donnée par :

$$\widehat{p}_n = \frac{274}{500} = 0.548$$

Les hypothèses sont vérifiées, en effet on a:

$$n > 30 \ , \ n.p = 274 > 5 \ , \ \text{et } n(1-p) = 226 > 5.$$

Alors l'intervalle de confiance de la proportion p est donné par :

$$IC(p) = \left[\widehat{p}_n - z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}, \widehat{p}_n + z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} \right]$$

avec $z_{\frac{\alpha}{2}}$ est déterminée à partir d'une lecture inverse de la table normale $\mathcal{N}(0,1)$ tel que :

$$P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2} = 0.025$$

$$z_{\frac{\alpha}{2}} = 1.96$$

Intervalle de confiance de la proportion

$$\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} = \sqrt{\frac{0.548(1-0.548)}{500}} = 0,022$$

Par conséquent :

$$IC(p) = [0.548 - 1.96 \times 0.022; 0.548 + 1.96 \times 0.022] = [0.504; 0.591]$$

2. Pour quel degré de confiance a-t-on la borne inférieure exactement égale à 50% ?

   On a la borne inférieure de l'IC(p) est :

$$\widehat{p}_n - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}$$

Intervalle de confiance de la proportion

Par conséquent

$$
\begin{aligned}
\widehat{p}_n - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} = 0.5 \iff & \widehat{p}_n - 0.5 = z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} \\
\iff & z_{\frac{\alpha}{2}} = \frac{\widehat{p}_n - 0.5}{\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}} \\
\iff & z_{\frac{\alpha}{2}} = \frac{0.548 - 0.5}{0.022} \\
\iff & z_{\frac{\alpha}{2}} = 2.18
\end{aligned}
$$

Sachant que $z_{\frac{\alpha}{2}}$ vérifie

$$P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

Par une lecture directe de la table normale on aura $\frac{\alpha}{2} = 0.0146$, et par la suite :

$$1 - \alpha = 97\%$$