Chapitre 3 :

19/12/18

Estimation par intervalle de confiance

1- Généralités :

On considère un modèle paramétrique induit par une v.a X où θ est un paramètre inconnu. Soit $\tilde{X}$ un échantillon extrait de X.

Contrairement à l'estimation ponctuelle qui consiste à fournir une valeur comme estimation du paramètre $g(\theta)$, on cherche un intervalle recouvrant très vraisemblablement la vraie valeur du paramètre $g(\theta)$.

Définition :

Soit $\alpha \in]0;1[$. On appelle intervalle de confiance de niveau de confiance $1-\alpha$ du paramètre $g(\theta)$ tout intervalle $C_n(\tilde{X})$ vérifiant :

$$\mathbb{P}\left[C_n(\tilde{X}) \ni g(\theta)\right] = 1-\alpha$$

Terminologie :

Si $\Theta \subset \mathbb{R}$, $C_n(\tilde{X})$ est de la forme $[A_n ; B_n]$ où A_n et B_n sont des variables aléatoires.

L'écart entre les bornes de l'intervalle est la marge d'erreur

Le réel α est appelé risque d'erreur. Il peut être réparti différemment de part et d'autres des bornes de l'intervalle de confiance.

Soit $\alpha = \alpha_1 + \alpha_2$ où α_1, α_2 mesurent respective-ment les risques à gauche et à droite.

- L'intervalle de confiance est bilateral quand $\alpha_1 \times \alpha_2 > 0$. L'intervalle est symetrique quand $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$. L'intervalle est disymetrique quand $\alpha_1 \neq \alpha_2$.
- Si $\alpha_1 \times \alpha_2 = 0$ $(\alpha_1 = 0$ ou $\alpha_2 = 0)$ alors l'in-tervalle est dit unilatéral. L'intervalle est unilaté-ral droit si $\alpha_1 = 0$ et $\alpha_2 = \alpha$. L'intervalle est unilatéral gauche si $\alpha_1 = \alpha$ et $\alpha_2 = 0$.

Le principe de la construction repose sur une variable pivotale.

Définition :

On dit que la statistique $\Pi(\underline{X};\theta)$ est une variable pivotale si sa loi est indépendante de θ.

Lorsque la recherche d'une telle variable n'est pas aisée, on utilise une variable asymptotiquement pivotale.

2- I.C pour une loi normale :

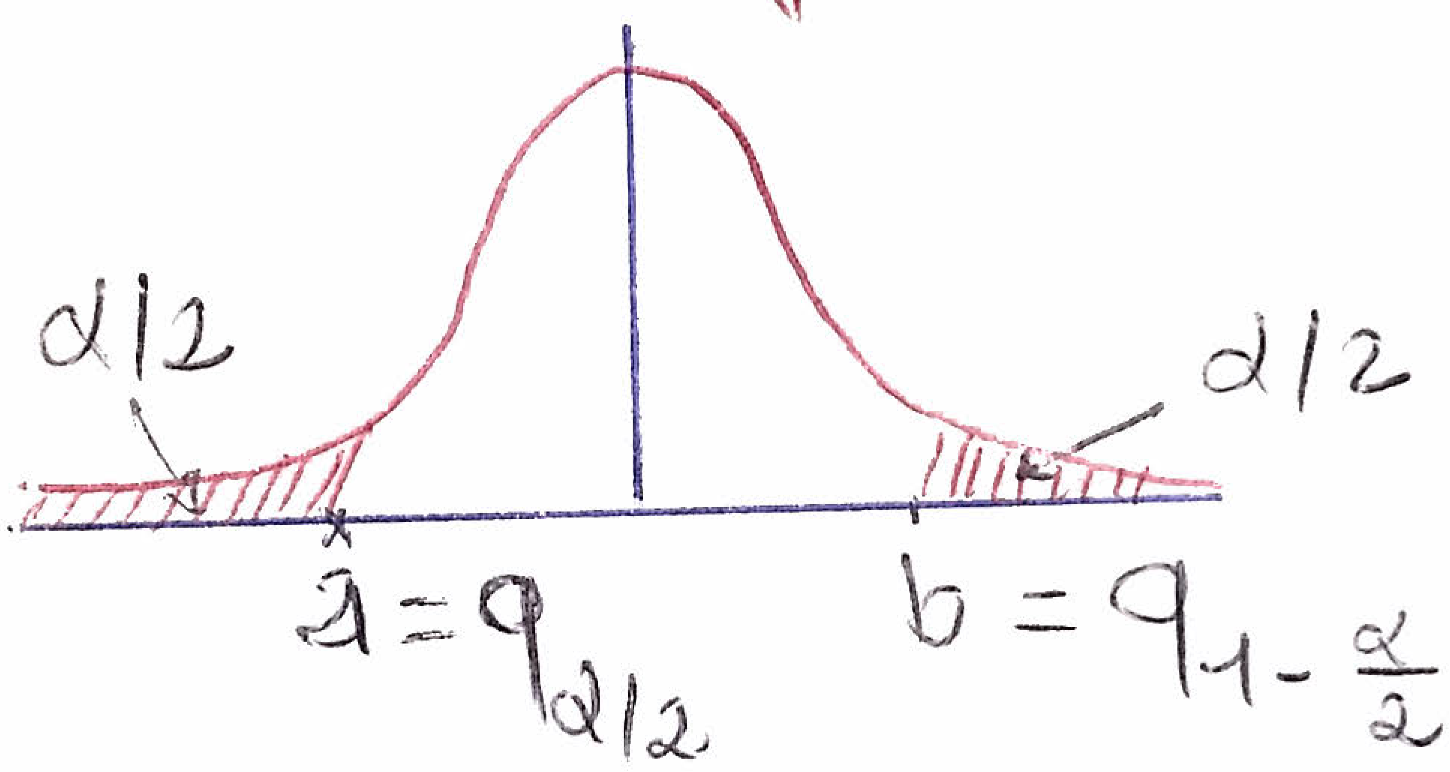
$X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$

2-1- IC de la moyenne μ

$\bar{X}$ est le meilleur estimateur

A- Si σ^2 connue :

$$\Pi(\tilde{X}; \theta) = \sqrt{n} \cdot \frac{\bar{X} - \mu}{\sigma} \longrightarrow \mathcal{N}(0;1)$$



$$\mathbb{P}\left(a \le \sqrt{n} \cdot \frac{\bar{X} - \mu}{\sigma} \le b\right) = 1 - \alpha$$

$$\mathbb{P}\left(\bar{X} - b\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} - a\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

or $b = 1 - \frac{\alpha}{2}$ et $a = q_{\frac{\alpha}{2}} = -q_{1-\frac{\alpha}{2}}$

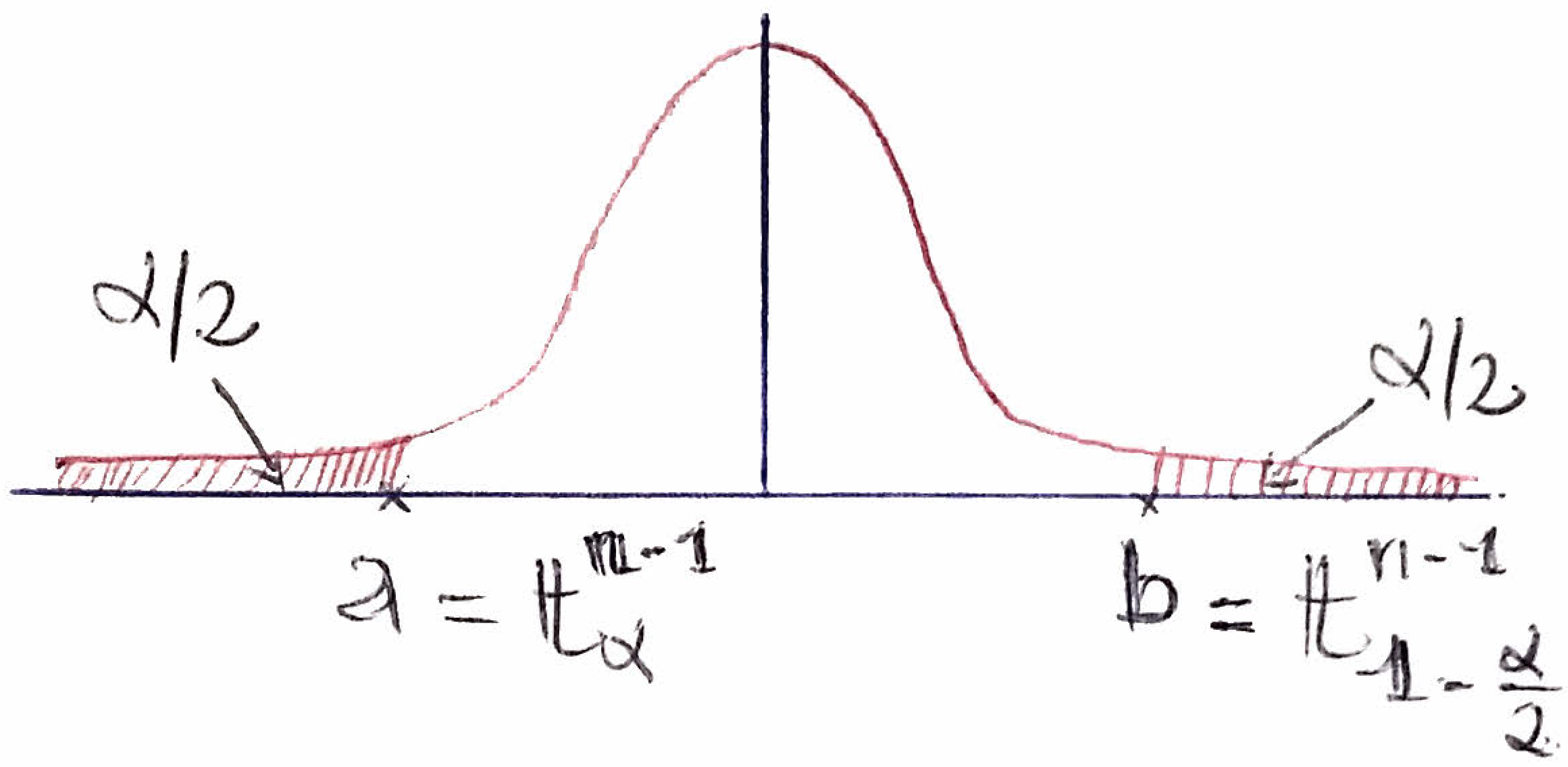
$$\mathbb{P}\left(\bar{X} - \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \le \mu \le \bar{X} + \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$\Rightarrow \boxed{IC_{(1-\alpha)\%}(\mu) = \left[\bar{X} - \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \,;\, \bar{X} + \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right]}$$

avec $q_{1-\frac{\alpha}{2}}$ quantile d'ordre $1-\frac{\alpha}{2}$ de la loi normale centrée réduite $\mathcal{N}(0;1)$.

B - Si σ^2 inconnue :

$$\Pi(\tilde{X};\mu) = \sqrt{n}\,\frac{\bar{X}-\mu}{S'} \longrightarrow \mathcal{S}t(n-1)$$



$$\boxed{IC_{(1-\alpha)\%}(\mu) = \left[\bar{X} \pm \frac{S'}{\sqrt{n}}\, t_{1-\frac{\alpha}{2}}^{n-1}\right]}$$

où $t_{1-\frac{\alpha}{2}}^{n-1}$ = quantile d'ordre $1-\frac{\alpha}{2}$ de la loi de Student $\mathcal{S}t(n-1)$.

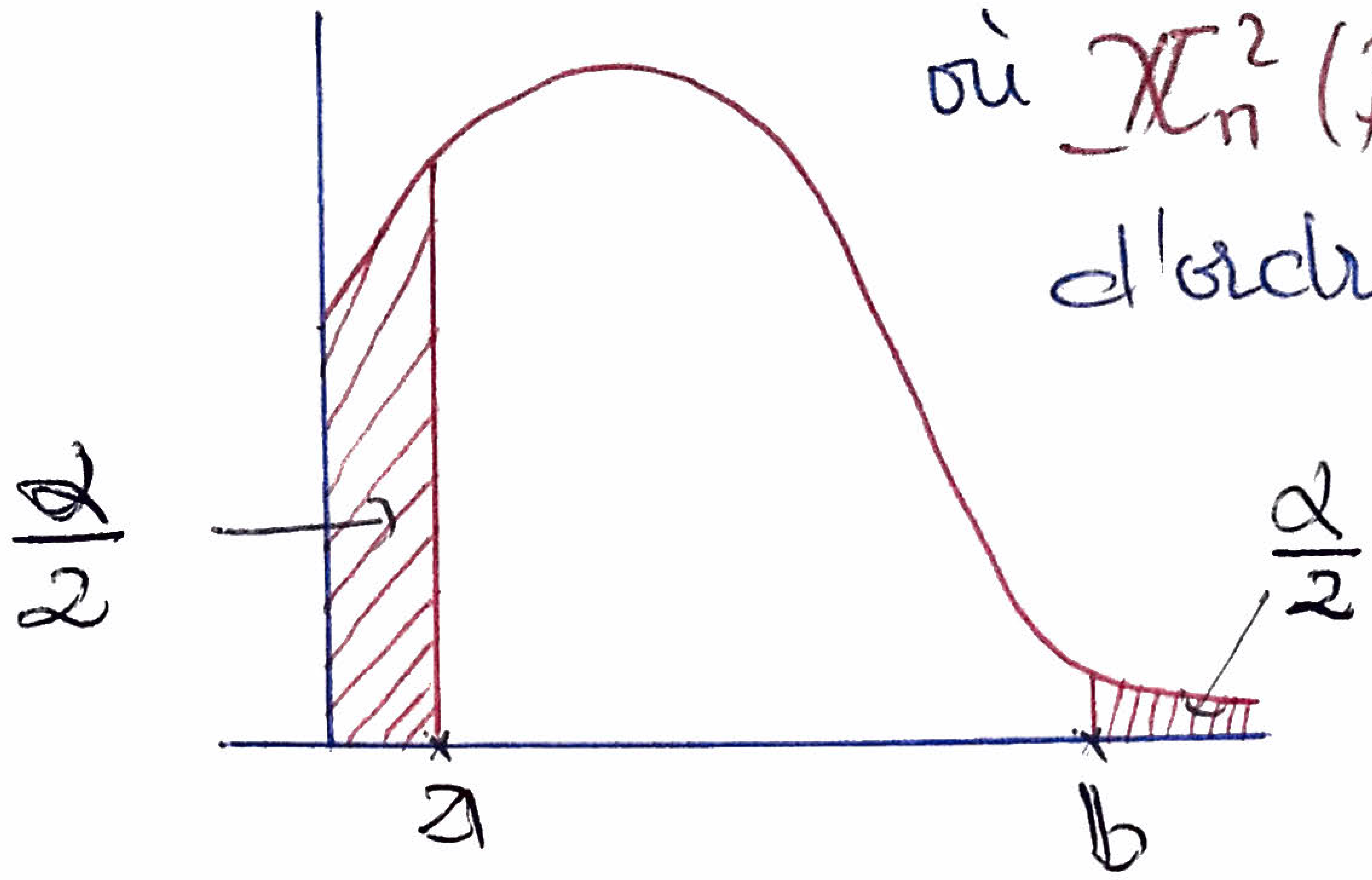
2-2- IC de la variance :

A - La moyenne μ connue :

$$\Pi(\tilde{X};\sigma^2) = \frac{n}{\sigma^2}\hat{\sigma}^2 \longrightarrow \chi^2_n$$

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$$

où $\chi^2_n(\beta)$ est le quantile d'ordre β de la loi χ^2_n



$$a = \chi^2_n(\alpha/2) \qquad b = \chi^2_n\left(1-\frac{\alpha}{2}\right)$$

$$\Rightarrow \quad a \leq \frac{n\hat{\sigma}^2}{\sigma^2} \leq b \quad \Rightarrow \quad \frac{1}{b} \leq \frac{\sigma^2}{n\hat{\sigma}^2} \leq \frac{1}{a}$$

$$\Rightarrow \quad \frac{n\hat{\sigma}^2}{b} \leq \sigma^2 \leq \frac{n\hat{\sigma}^2}{a}$$

$$\boxed{IC_{(1-\alpha)\%}(\sigma^2) = \left[\frac{n\hat{\sigma}^2}{\chi^2_n(1-\alpha/2)} ; \frac{n\hat{\sigma}^2}{\chi^2_n(\alpha/2)}\right]}$$

B- Si la moyenne μ est inconnue :

$$\boxed{IC_{(1-\alpha)\%}(\sigma^2) = \left[\frac{(n-1)S'^2}{\chi^2_{n-1}(1-\frac{\alpha}{2})} ; \frac{(n-1)S'^2}{\chi^2_{n-1}(\alpha/2)}\right]}$$

Remarque :

- Si on travaille avec une variable pivotale, on parle ~~de~~ d'intervalle de confiance exacte.
- Si on travaille avec une variable asymptotiquement pivotale, on parle d'intervalle de confiance asymptotique.

2-3 - IC pour deux échantillons gaussiens :

Soient $X_1, ---, X_n \xrightarrow{iid} \mathcal{N}(m_x, \sigma_x^2)$

$Y_1, ---, Y_n \xrightarrow{iid} \mathcal{N}(m_y ; \sigma_y^2)$

On suppose que les échantillons sont independants et de même variances (σ^2)

A - IC de $m_x - m_y$:

* Si σ^2 connue :

$$\bar{X} \longrightarrow \mathcal{N}\left(m_x ; \frac{\sigma^2}{n}\right)$$

$$\bar{Y} \longrightarrow \mathcal{N}\left(m_y ; \frac{\sigma^2}{m}\right)$$

Comme $X \perp Y$, $\bar{X} - \bar{Y} \longrightarrow \mathcal{N}\left(m_x - m_y ; \sigma^2\left(\frac{1}{n}+\frac{1}{m}\right)\right)$

$$\frac{(\bar{X}-\bar{Y}) - (m_x - m_y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} \longrightarrow \mathcal{N}(0;1)$$

Donc, $\Pi\left[(\tilde{X}, \tilde{Y}), m_x - m_y\right] = \dfrac{(\bar{X}-\bar{Y}) - (m_x - m_y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}$

* Si σ^2 inconnue :

On a : $\dfrac{(n-1) S_x'^2}{\sigma^2} \longrightarrow \chi^2_{n-1}$

$\dfrac{(m-1) S_Y'^2}{\sigma^2} \longrightarrow \chi^2_{m-1}$

Comme $X \perp\!\!\!\perp Y$ alors

$$\frac{(n-1)S_x'^2 + (m-1)S_y'^2}{\sigma^2} \longrightarrow \chi^2_{n+m-2}$$

$$\frac{\dfrac{(\bar{X}-\bar{Y})-(m_x-m_y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\dfrac{(n-1)S_x'^2+(m-1)S_y'^2}{\sigma^2(n+m-2)}}} = \frac{(\bar{X}-\bar{Y})-(m_x-m_y)}{S_p'\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

où $S_p' = \dfrac{(n-1)S_x'^2+(m-1)S_y'^2}{n+m-2}$

$$\Rightarrow \boxed{\frac{(\bar{X}-\bar{Y})-(m_x-m_y)}{S_p'\sqrt{\frac{1}{n}+\frac{1}{m}}} \longrightarrow \mathcal{S}t(n+m-2)}$$

B - IC de $\frac{\sigma_x^2}{\sigma_y^2}$

$$\left.\begin{array}{l} \dfrac{(n-1)S_x'^2}{\sigma_x^2} \longrightarrow \chi^2_{n-1} \\[2ex] \dfrac{(m-1)S_y'^2}{\sigma_y^2} \longrightarrow \chi_{m-1} \end{array}\right\} \Rightarrow \boxed{\frac{\dfrac{S_x'^2}{\sigma_x^2}}{\dfrac{S_y'^2}{\sigma_y^2}} \longrightarrow \mathcal{F}(n-1;m-1)}$$

3 - IC pour une loi de Bernouilli (proportion)

$$X_1, \dots, X_n \longrightarrow Ber(p)$$

$$\hat{p} = \frac{1}{n}\sum X_i \Rightarrow n\hat{p} = \sum X_i \rightsquigarrow \mathcal{B}(n;p)$$

Pour $n \nearrow$, $n\hat{p} \rightsquigarrow \mathcal{N}(np; np(1-p))$

$$\frac{n\hat{p} - np}{\sqrt{np(1-p)}} \longrightarrow \mathcal{N}(0;1)$$

$$\hat{p} \xrightarrow{P} p$$

$$\Rightarrow \frac{n\hat{p} - np}{\sqrt{n\hat{p}(1-\hat{p})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0;1)$$

$$\sqrt{n} \cdot \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0;1)$$

Variable asymtotiquement pivotale:

$$\boxed{IC(p) = \left[\hat{p} \pm q_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]}$$

$$IC(P_1 - P_2) = \left[(\hat{P}_1 - \hat{P}_2) \pm q_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n} + \frac{\hat{P}_2(1-\hat{P}_2)}{m}}\right]$$

(TAF)