# Chapitre 4

# Estimation Ensembliste

## 4.1 Principe de la méthode

Soit un modèle statistique $\rho = (\varkappa, \mathcal{B}, \{P_\theta, \theta \in \Theta\})$ associé à une observation $x$ d'une v.a $X$ de loi $P_\theta$. La méthode consiste à donner un intervalle (resp.région), $C_X(\theta)$, vraisemblant la vrai valeur inconnu $\theta_0$ du paramètre $\theta$ à estimer (resp. g($\theta$)) de niveau de confiance $1-\alpha$ tel que :

$$\begin{cases} P(\theta \in C_X(\theta)) = 1-\alpha; \alpha \in [0,1]. \\ P[g(\theta) \in C_X(\theta)] = 1-\alpha \end{cases}$$

$\alpha = P[\theta \notin C_X(\theta)]$ est appelé le risque.

où les bornes de $C_X(\theta)$ dépendent de l'échantillon considéré.

Alors, pour trouver l'intervalle de confiance $C_X(\theta) = [a,b]$, on a :

$$\underbrace{P(\theta \in [a,b])}_{= P[a \leq \theta \leq b]} = \int_a^b f(\theta)\,d\theta = 1-\alpha \Leftrightarrow F(b) - F(a) = 1-\alpha, \text{ où } \alpha = \alpha_1 + \alpha_2 \text{ dont :}$$

.$\alpha_1 = P(\theta \leq a) = F(a) \to a = F^{-1}(\alpha_1)$ : fractile d'ordre $\alpha_1$ d'une certaine loi.

.$\alpha_2 = P(\theta \geq b) \to \underbrace{1-\alpha_2}_{= P[\theta \leq b]} = F(b) \to b = F^{-1}(1-\alpha_2)$ : fractile d'ordre $1-\alpha_2$ d'une certaine loi.

**Cas de $C_X(\theta)$ :**

•$\alpha_1 \neq 0, \alpha_2 \neq 0 :\Rightarrow$ Intervalle de confiance bilatéral de niveau $1-\alpha$.

•$\alpha_1 = 0 \to \alpha_2 = \alpha \Rightarrow$Intervalle de confiance unilatéral à droite. $= ]-\infty; k]$

•$\alpha_2 = 0 \to \alpha_1 = \alpha \Rightarrow$Intervalle de confiance unilatéral à gauche. $= [k, +\infty[$

**Remarque 24** *1.Dans la classe des lois symétriques telles que la loi normale centrée réduite et la loi de student, l'intervalle de confiance de niveau $1-\alpha$ est de longueur minimale pour $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$.*

*2. Afin de trouver l'intervalle (resp.région) de confiance, on doit passer par l'intermédiaire des fonctions pivotales.*

**Définition 25 (fonction pivotale)** *On appelle fonction pivotale pour $\theta$ toute application mesurable $\varphi$ qui est fonction des observations $x_1, ..., x_n$ et du paramètre $\theta$ dont sa loi ne dépend d'aucun paramètre.*

**Définition 26** *On appelle une fonction asymptotiquement pivotale ou quasi pivotale, toute fonction $Q$ converge en loi vers une loi de probabilité $p$ fixée.*

**Exemple 27** *Trouver des fonctions pivotales pour :*

*1/. n éch de $X \rightsquigarrow N(m, \sigma^2)$.*

*2/. $X \rightsquigarrow \mathcal{E}(\lambda)$* → détermine $IC_\alpha(\lambda)$ [comment construire un intervalle de confiance].

*2. 1 obs de $X \rightsquigarrow B(n,p)$.*

*3/. n éch de $X \rightsquigarrow \mathcal{P}(\lambda)$.*

## 4.2 Intervalles de confiance classiques

### 4.2.1 I.C de l'espérance de $N(m, \sigma^2)$

<u>$\sigma$ connu</u> :

$Q(x_1, ..., x_n, m) = \sqrt{n}\frac{\overline{x}-m}{\sigma}$ est une pivotale pour m de loi $N(0,1)$.

Alors : $P\left[-q_{1-\frac{\alpha}{2}} < Q < q_{1-\frac{\alpha}{2}}\right] = 1-\alpha$, où : $q_{1-\frac{\alpha}{2}}$ : fractile d'ordre $\left(1-\frac{\alpha}{2}\right)$ de $N(0,1)$.

$$\Leftrightarrow P\left[-q_{1-\frac{\alpha}{2}} < \sqrt{n}\frac{\overline{x}-m}{\sigma} < q_{1-\frac{\alpha}{2}}\right] = 1-\alpha$$

$$\Leftrightarrow P\left[\overline{x}-\frac{\sigma}{\sqrt{n}}q_{1-\frac{\alpha}{2}}<m<\overline{x}+\frac{\sigma}{\sqrt{n}}q_{1-\frac{\alpha}{2}}\right]=1-\alpha$$

D'où : $IC_{\alpha}(m)=\left[\overline{x}-\frac{\sigma}{\sqrt{n}}q_{1-\frac{\alpha}{2}},\overline{x}+\frac{\sigma}{\sqrt{n}}q_{1-\frac{\alpha}{2}}\right]$ ,où :

$h=\frac{\sigma}{\sqrt{n}}q_{1-\frac{\alpha}{2}}$ est appellée la précision.

$\underline{\sigma \text{ inconnu}}$ :

$Q(x_1,...,x_n,m)=\sqrt{n}\frac{\overline{x}-m}{S'}$ est une pivotale pour m de loi $st_{n-1}$.

Alors : $P\left[-st_{n-1}\left(1-\frac{\alpha}{2}\right)<Q<st_{n-1}\left(1-\frac{\alpha}{2}\right)\right]=1-\alpha$, où : $st_{n-1}\left(1-\frac{\alpha}{2}\right)$ : fractile d'ordre $\left(1-\frac{\alpha}{2}\right)$ de $st_{n-1}$.

On en déduit que $P\left[\overline{x}-st_{n-1}\left(1-\frac{\alpha}{2}\right)\frac{S'}{\sqrt{n}}<m<\overline{x}+st_{n-1}\left(1-\frac{\alpha}{2}\right)\frac{S'}{\sqrt{n}}\right]=1-\alpha$

D'où : $IC_{\alpha}(m)=\left[\overline{x}\mp st_{n-1}\left(1-\frac{\alpha}{2}\right)\frac{S'}{\sqrt{n}}\right]$.

### 4.2.2 I.C de la variance de $N(m,\sigma^2)$

$\underline{m \text{ connu}}$ :

$Q(x_1,...,x_n,\sigma^2)=\sum_{i=1}^{n}\frac{(x_i-m)^2}{\sigma^2}$ est une pivotale pour $\sigma^2$ de loi $\chi_n^2$.

Alors : $P\left[\chi_n^2(\alpha_1)<\sum_{i=1}^{n}\frac{(x_i-m)^2}{\sigma^2}<\chi_n^2(1-\alpha_2)\right]=1-\alpha$

$$\Leftrightarrow P\left[\frac{\sum_{i=1}^{n}(x_i-m)^2}{\chi_n^2(1-\alpha_2)}<\sigma^2<\frac{\sum_{i=1}^{n}(x_i-m)^2}{\chi_n^2(\alpha_1)}\right]=1-\alpha$$

**Cas particulier :** $\alpha_1=\alpha_2=\frac{\alpha}{2}$ : $\chi_n^2$ n'est pas symétrique.

<u>$m$ inconnu :</u>

$Q(x_1, ..., x_n, \sigma^2) = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{\sigma^2} = \frac{(n-1) S'^2}{\sigma^2}$ est une pivotale pour $\sigma^2$ de loi $\chi^2_{n-1}$.

$$\text{Alors : } IC_\alpha(\sigma^2) = \left[ \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{\chi^2_{n-1}(1-\alpha_2)}, \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{\chi^2_{n-1}(\alpha_1)} \right]$$

$$= \left[ \frac{(n-1) S'^2}{\chi^2_{n-1}(1-\alpha_2)}, \frac{(n-1) S'^2}{\chi^2_{n-1}(\alpha_1)} \right]$$

### 4.2.3 I.C pour le rapport des variances de deux lois $N(m_1, \sigma_1^2)$ et $N(m_2, \sigma_2^2)$

$n_1$ éch de $X_1 \rightsquigarrow N(m_1, \sigma_1^2)$

$n_2$ éch de $X_2 \rightsquigarrow N(m_2, \sigma_2^2)$

On sait que : $\frac{n_1 - 1}{\sigma_1^2} S_1'^2 \rightsquigarrow \chi^2_{n_1-1}$ et $\frac{n_2 - 1}{\sigma_2^2} S_2'^2 \rightsquigarrow \chi^2_{n_2-1}$

$X_1 \perp X_2 \Rightarrow S_1'^2 \perp S_2'^2$

Alors : $\frac{S_1'^2}{\sigma_1^2} \frac{\sigma_2^2}{S_2'^2} \rightsquigarrow F_{n_1-1, n_2-1}$

On en déduit que $Q = \frac{S_2'^2}{S_1'^2} \frac{\sigma_1^2}{\sigma_2^2}$ est une fonction pivotale pour $\frac{\sigma_1^2}{\sigma_2^2}$ de loi $F_{n_2-1, n_1-1}$.

D'où :

$P[F_{n_2-1, n_1-1}(\alpha_1) \leq Q \leq F_{n_2-1, n_1-1}(1-\alpha_2)] = 1 - \alpha$.

$$\Leftrightarrow IC_\alpha\left(\frac{\sigma_1^2}{\sigma_2^2}\right) = \left[ \frac{S_1'^2}{S_2'^2} F_{n_2-1, n_1-1}(\alpha_1), \frac{S_1'^2}{S_2'^2} F_{n_2-1, n_1-1}(1-\alpha_2) \right]; \alpha = \alpha_1 + \alpha_2.$$

### 4.2.4 I.C pour la différence des espérances de deux lois $N\left(m_1, \sigma_1^2\right)$ et $N\left(m_2, \sigma_2^2\right)$

Rappelons que $\overline{X}_1 \sim N\left(m_1, \frac{\sigma_1^2}{n_1}\right) \Longrightarrow \sqrt{n_1}\frac{\left(\overline{X}_1 - m_1\right)}{\sigma_1} \sim N(0,1)$

$$\overline{X}_2 \sim N\left(m_2, \frac{\sigma_2^2}{n_2}\right) \Longrightarrow \sqrt{n_2}\frac{\left(\overline{X}_2 - m_2\right)}{\sigma_2} \sim N(0,1)$$

On en déduit que $\overline{X}_1 - \overline{X}_2 \sim N\left(m_1 - m_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$

**i\Cas où $\sigma_1^2$ et $\sigma_2^2$ connus :**

$Q = \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - (m_1 - m_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ est une fonction pivotale pour $(m_1 - m_2)$ deloi $N(0,1)$

Alors : $P\left[-q_{1-\frac{\alpha}{2}} < Q < q_{1-\frac{\alpha}{2}}\right] = 1-\alpha \Leftrightarrow P\left[-q_{1-\frac{\alpha}{2}} < \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - (m_1 - m_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < q_{1-\frac{\alpha}{2}}\right] =$ $1-\alpha$.

D'où $IC_\alpha(m_1 - m_2) = \left[\left(\overline{X}_1 - \overline{X}_2\right) \pm q_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$

**ii\Cas où $\sigma_1^2 = \sigma_2^2 = \sigma^2$ inconnus :**

Notons que $\dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - (m_1 - m_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ n'est pas une fonction pivotale pour $(m_1 - m_2)$ car elle dépend de $\sigma$ inconnu. Ainsi, on pourrai estimer $\sigma^2$ par $S_1'^2$ ou bien $S_2'^2$ et faire apparaitre une loi de student à $n_1 - 1$ ou $n_2 - 1$ ddl. Mais, la statistique $S^2 = \dfrac{(n_1 - 1)S_1'^2 + (n_2 - 1)S_2'^2}{n_1 + n_2 - 2}$

est un ESB de $\sigma^2$ meilleur que $S_1'^2$ ou $S_2'^2$. Donc, on estime $\sigma^2$ par $S^{*2}$. Au quel cas, on a

$$\frac{\dfrac{(\overline{X}_1-\overline{X}_2)-(m_1-m_2)}{\sigma\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}}{\sqrt{\dfrac{\dfrac{(n_1+n_2-2)S^{*2}}{\sigma^2}}{(n_1+n_2-2)}}}\sim st_{n_1+n_2-2}$$

c'est-à-dire $Q=\dfrac{(\overline{X}_1-\overline{X}_2)-(m_1-m_2)}{S^*\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ est une fonction pivotale pour $(m_1-m_2)$ de loi $st_{n_1+n_2-2}$.D'où

$$IC_\alpha(m_1-m_2)=\left[(\overline{X}_1-\overline{X}_2)\pm st_{n_1+n_2-2}\left(\frac{\alpha}{2}\right)S^*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]$$

**Exercice :** Montrer que la statistique $S^{*2}$ est un ESB de $\sigma^2$ meilleur que $S_1'^2$ ou $S_2'^2$.

**iii\Cas où $\sigma_1^2\neq\sigma_2^2$ inconnus :**

En Supposant que $n_1=n_2=n$, alors $\sqrt{n}\dfrac{(\overline{X}_1-\overline{X}_2)-(m_1-m_2)}{\sqrt{\sigma_1^2+\sigma_2^2}}\xrightarrow{L}N(0,1)\ldots(1)$.

D'autre part, on sait que $\begin{cases}S_1^2\xrightarrow[P]{PS}\sigma_1^2\\ S_2^2\xrightarrow[P]{PS}\sigma_2^2\end{cases}\Longrightarrow\sqrt{S_1^2+S_2^2}\xrightarrow[P]{PS}\sqrt{\sigma_1^2+\sigma_2^2}$

Autrement dit, $\dfrac{\sqrt{S_1^2+S_2^2}}{\sqrt{\sigma_1^2+\sigma_2^2}}\xrightarrow[P]{PS}1\ldots(2)$.

$\dfrac{(1)}{(2)}\longrightarrow Q=\sqrt{n}\dfrac{(\overline{X}_1-\overline{X}_2)-(m_1-m_2)}{\sqrt{S_1^2+S_2^2}}\xrightarrow{L}N(0,1)$ $\sqrt{S_1^2+S_2^2}$

Donc, $Q$ est une fonction pivotale asymptotique pour $(m_1-m_2)$ de loi $N(0,1)$ et l'intervalle de confiance asymptotique pour $(m_1-m_2)$ au niveau $(1-\alpha)$ est tel que :

$$IC_\alpha(m_1-m_2)=\left[(\overline{X}_1-\overline{X}_2)\pm q_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2+S_2^2}{n}}\right]$$

## 4.3 Intervalle de confiance basé sur les estimateurs:

Lorsque il n'est pas easy de trouver une fonction pivotale pour $g(\theta)$ alors on pensera à construire un Intervalle de confiance de la forme: $\left[\hat{g}(\theta) - \delta_1 \; , \; \hat{g}(\theta) + \delta_2\right]$
où: $\hat{g}(\theta)$ est un estimateur ponctuelle de $g(\theta)$.

**Remarque:**

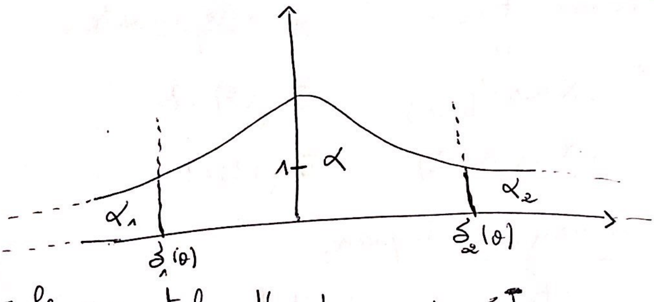
$$\boxed{\text{on obtiendera de meilleur intervalle de confiance } \underline{\underline{\text{si}}} \ \hat{g}(\theta) \text{ est un très bon estimateur de } g(\theta)}$$

- Soit $T$ un estimateur de $g(\theta)$ de fonction de répartition $F_\theta^T$ et de densité $f_\theta^T$

$$P\left(\delta_1(\theta) \leqslant T \leqslant \delta_2(\theta)\right) = 1-\alpha$$

**Remarque:**

$$\boxed{T \text{ n'est pas une fonction pivotale donc la loi de } T \text{ depend de } \theta \text{ et } \delta_1(\theta) \text{ et } \delta_2(\theta) \text{ depent aussi de } \theta}$$

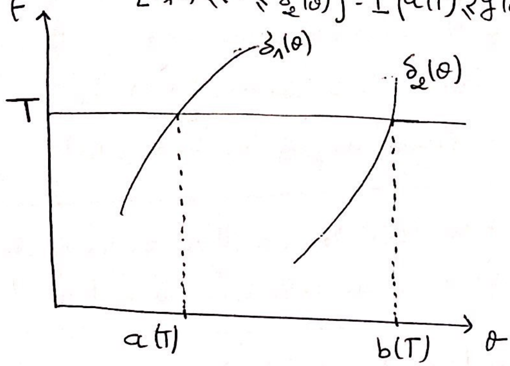



- $\delta_1(\theta)$ = le quantile d'ordre $\alpha_1$ de $F_\theta^T$
- $\delta_2(\theta)$ = le quantile d'ordre $(1-\alpha_2)$ de $F_\theta^T$

$$\alpha_1 + \alpha_2 = \alpha$$

Si on réussit à écrire : $\delta_1(\theta) \leq T \leq \delta_2(\theta) \Leftrightarrow a(T) \leq g(\theta) \leq b(T)$

On aura alors : $P\left[\delta_1(\theta) \leq T \leq \delta_2(\theta)\right] = P\left(a(T) \leq g(\theta) \leq b(T)\right] = 1-\alpha$.




calculer T(x) pour un échantillon fixé

Remarque : l'intervalle de confiance $[a(T), b(T)]$ sera plus petit (meilleur) quand les courbes $\delta_1(\theta)$ et $\delta_2(\theta)$ se rapproche l'une de l'autre.
c.à.d ; la distribution de T sera moins dispersée.
On prendra si c'est possible l'E.S.B.V.U.M qui revient en même en pratique à prendre une stat exhaustive et complète si elle existe.

Exemple : $\cdot X \sim U_{[0,\theta]}$ $\quad IC_\alpha(\theta) = ?$

$\cdot X \sim B(p)$ $\quad IC_\alpha(p) = ?$

4.4 Régions asymptotiques :

a) Si il n'existe pas une pivotale pour $g(\theta) \Rightarrow$ on détermine si possible une pivotale asymptotique
$Q(x_1, \dots, x_n, \theta)$ = fonction en $(x_1, \dots, x_n)$ et $\theta$ tq :
$Q(x_1, \dots, x_n, \theta) \underset{n \to \infty}{\sim}$ loi libre (indépendante de paramètre).

Exemple

$\cdot X \leadsto \mathcal{B}(p)$ . $IC_\alpha(p) = ?$

**b) Région de confiance asymptotique basée sur l'e.m.v**

soit $\hat{\theta}_n$ est l'e.m.v de $\theta$.

On sait que $\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{loi} N(0, I^{-1}(\theta))$

$\Rightarrow \sqrt{I(\theta)}\,\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{loi} N(0,1)$.

$\Rightarrow Q(x_1, .. x_n, \theta) = \sqrt{I(\theta)}\,\sqrt{n}\,(\hat{\theta}_n - \theta)$ est une fonction pivotale asymptotique pour $\theta$ de loi $N(0,1)$, [$\hat{\theta}_n$ est en fonction de $(x_1, .. x_n)$]

d'où

$$P\left[-q_{1-\frac{\alpha}{2}} \leq \sqrt{I(\theta)}\,\sqrt{n}\,(\hat{\theta}_n - \theta) \leq q_{1-\frac{\alpha}{2}}\right] = 1 - \alpha$$

dans la pratique $I(\theta)$ est remplacée par $I(\hat{\theta}_n)$

Exemple :

$X \leadsto \mathcal{P}(\lambda)$

Donner I.C asymptotique pour $\lambda$ de niveau $1-\alpha$.