ESTIMATEUR D'UNE LOI (H)
(15 / 12 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un estimateur « légal » est un estimateur qui prend ses valeurs dans une **famille** de **lois de probabilité**.

(i) Un exemple d'estimation légale est le suivant (I. OLKIN - C.H. SPIEGELMAN). On considère un **modèle statistique** $(\Omega, \mathscr{T}, \mathscr{R})$ dans lequel $\mathscr{R} = \mathscr{P} \cup \mathscr{Q}$, où $\mathscr{P} = (P_\theta)_{\theta \in \Theta}$ est de type paramétrique (avec $\Theta \subset \mathbf{R}^Q$) et $\mathscr{Q}$ une famille quelconque. Soit $(\mathscr{X}, \mathscr{B})$ un **espace d'observation** et $X : \Omega \mapsto \mathscr{X}$ une **va** ou une **statistique** donnée (**échantillon**), la **loi de probabilité** de X est de la forme :

(1) $\quad R^X(\theta, \alpha) = \alpha \,.\, P_\theta^X + (1 - \alpha) \,.\, Q^X \in \mathscr{R}^X,$

ie est une **combinaison linéaire convexe** d'une loi paramétrique et d'une loi non paramétrique, images resp de $P_\theta \in \mathscr{P}$ et de $Q \in \mathscr{Q}$ par X, où $\alpha \in [0, 1]$.

Pour estimer le mélange (1), l'approche « paramétrique » correspond à $\alpha = 1$, et l'approche « non paramétrique » à $\alpha = 0$.

(ii) Soit $X = (X_1, ..., X_N)$ un N-**échantillon** et d une **distance entre probabilités** sur $\mathscr{R}$. On peut estimer $R^X(\theta, \alpha)$ à partir de la **loi empirique** P_N (fondée sur X) en utilisant une méthode de distance minimum pr à (θ, α) lorsque Q^X est donnée (ou déjà estimée) (cf **méthodes à distance minimale**).

Lorsque $\alpha \in]0, 1[$, on parle parfois d'**approche semi-paramétrique**.

En particulier, si α est estimé par $\alpha^{\sim}$, on peut tester l'**hypothèse de base** $H_0 : \alpha = 0$ (approche non paramétrique) contre l'hypothèse $H_a : \alpha = 1$ (approche paramétrique) (ou inversement).

En pratique, on se limite à une **caractéristique** (eg une **densité** ou une **fr**) des lois en question, où $\gamma = dQ^X / d\mu$ est estimée par une **méthode non paramétrique** (cf **estimateur de la densité**).

(iii) Il existe aussi des méthodes « indirectes » d'estimation d'une **lp** : eg estimation de sa densité (cf **estimateur de la densité**) ou de sa fr (cf **estimateur de la fonction de répartition**), ou encore de sa **fc**, etc.

On doit alors vérifier la qualité de l'estimateur légal ainsi obtenu. Ainsi, si le modèle considéré $\mathscr{P}^\xi$ est un **modèle dominé** par une **mesure** μ (d'où une famille de densités tq f), et si $f^{\#}$ est un estimateur de f associé à un ensemble d'**observations** X, on peut écrire :

(2)

$$P^{\xi}(B) = \int_B f(x)\, d\mu, \qquad \forall B \in \mathscr{B} \text{ (}\textbf{tribu de parties}\text{)},$$

$$(P^{\xi})^{\#}(B) = \int_B f^{\#}(x)\, d\mu, \qquad \forall B \in \mathscr{B}.$$

L'estimateur indirect $(P^{\xi})^{\#}$, ainsi obtenu, de la loi P^{ξ} doit donc vérifier des propriétés d'**optimalité** : eg être sans **biais** pr à P^{ξ}, ou encore être « proche » de P^{ξ} dans un certain sens (cf eg **distance entre lois**, **proximité**, **adéquation**).