ESTIMATEUR D'ORDRE (F6, H1)
(15 / 06 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un **estimateur d'ordre** est un estimateur non paramétrique utilisé notamment pour l'**estimation robuste** d'un **paramètre de position** (cf aussi **estimation non paramétrique**, **estimateur de rang**, **robustesse**).

(i) Soit $(\Omega, \mathcal{T}, P_\theta)_{\theta \in \Theta}$ un **modèle statistique** et $\xi : \Omega \mapsto \mathbf{R}$ une **vars** de **loi** P_θ^ξ (où $\theta \in \Theta$). On considère que le **modèle image** est un **modèle d'échantillonnage** $(\mathbf{R}^N, \mathcal{B}(\mathbf{R}^N), (P_\theta^\xi)^{\otimes N})_{\theta \in \Theta}$. Par suite, X est un **échantillon iid** comme ξ. On suppose que $\Theta = \mathbf{R}$ et que $\theta \in \Theta$ est un **paramètre de position**, ie qu'il existe une **lp** P_0^ξ tq $P_\theta^\xi = P_0^{a(\xi, \theta)}$, où $a : \mathbf{R}^2 \mapsto \mathbf{R}$ est la fonction d'**écart** : $(x, \theta) \mapsto a(x, \theta) = x - \theta$. On note F_0 la fr (resp f_0 la **densité**) associée à P_0^ξ.

La **fonction de répartition** F_θ de ξ s'écrit :

(1) $$x \in \mathbf{R} \mapsto F_\theta(x) = F_0(x - \theta).$$

On appelle alors **estimateur d'ordre** (**linéaire**) de θ une **statistique** T_N tq :

(2) $$T_N = t_N(X) = \Sigma_{n=1}^N a_N(n) . X^{(n)},$$

dans laquelle $X^{(.)} = (X^{(1)}, ..., X^{(N)})$ est la **statistique d'ordre** associée à X et $(a_N(n))_{n=1,...,N}$ est une **suite** réelle finie, dont les éléments sont appelés **poids** ou **pondérations** (cf **score**).

(ii) Lorsque $N >> 0$, on définit souvent la suite $(a_N(n))_{n=1,...,N}$ à partir d'une **fonction de « score »** donnée $J : [0, 1] \mapsto \mathbf{R}_+$ selon (cf **fonction score**) :

(3) $$a_N(n) = J\{n / (N + 1)\} / \Sigma_{n=1}^N J\{n / (N + 1)\}.$$

On établit que la fonction J qui définit l'estimateur T_N le plus efficace est :

(4) $$J = \varphi \circ F_0^{-1}, \qquad \text{avec } \varphi = - D(f_0' / f_0),$$

où D symbolise l'opérateur de **dérivation**.

(iii) On suppose que l'estimateur d'ordre T_N est de la forme :

(5) $$T_N = N^{-1} . \Sigma_{n=1}^N c_N(n) . X^{(n)} \qquad (\text{ie } a_N(n) = c_N(n) / N),$$

que J est bornée et continue, que $\xi \in L_{\mathbf{R}}^2(\Omega, \mathcal{T}, P)$, $\forall \theta$, et que l'on choisit $c_N(n) = J\{n / (N + 1)\}$ (S.M. STIGLER).

On montre alors que T_N est asymptotiquement gaussien (cf **loi asymptotique**, **loi gaussienne**) :

(6) $\mathscr{L}\ \{(T_N - E_0\, T_N) / \sigma\ (T_N)\} \rightarrow_{N \rightarrow +\infty} \mathscr{N}(0, 1)$ (**loi normale réduite**),

dès que J et F_0 vérifient la condition :

(7) $\int \mathbf{1}(\mathbf{R}^2)\ .\ J\ (F_0\ (x))\ .\ J\ (F_0\ (y))\ .\ [F\ (\min\ (x, y)) - F\ (x)\ .\ F\ (y)]\ dx\ dy\ >\ 0.$

Cette propriété permet de fonder divers **tests non paramétriques** relatifs à θ.

T_N est aussi appelé **estimateur de type L**, ou **L-estimateur**, de θ.