MÉTHODE DES QUANTILES (H3, J1)
(17 / 01 / 2020, © Monfort, Dicostat2005, 2005-2020)

Deux signification peuvent être attribuées à cette espression :

(a) une méthode analogue à la **méthode des moments**, dans laquelle les **moments algébriques** sont remplacés par des **quantiles**. L'**équation estimante** se déduit par identification de chaque **quantile théorique** avec le **quantile empirique** correspondant ;

(b) une méthode d'**estimation** des paramètres d'un **modèle de régression** qui généralise la **méthode des moindres carrés**. Cette méthode contient, en particulier, la méthode d'estimation « médiane » d'un **paramètre de position**.

(i) Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé** et $\zeta = (\xi, \eta) : \Omega \mapsto \mathbf{R}^K \times \mathbf{R}$ un **couple aléatoire** constitué d'un vecteur (ou « liste ») ξ constitué(e) de K **variables exogènes** scalaires et d'une **variable endogène** scalaire η. On note P^ζ ou $P^{(\xi,\eta)}$ la **loi de probabilité** de $\zeta = (\xi, \eta)$, $P(. / \xi)$ la **loi conditionnelle** de η sachant ξ et $G(. / \xi)$ la **fonction de répartition** associée à $P(. / \xi)$.

Etant donnée une **fonction de régression** $f : \mathbf{R}^K \times \mathbf{R}^Q \mapsto \mathbf{R}$ tq (cf **modèle non linéaire**) :

(1) $\quad \eta = f(\xi, b) + \varepsilon,$

on peut écrire :

(2) $\quad G(t / \xi) = P(\eta < t / \xi) = P(\varepsilon < t - f(\xi, b)) = F_\varepsilon(t - f(\xi, b)), \qquad \forall\, t \in \mathbf{R},$

où F_ε est la **fr** propre de ε.

Soit $Z = (X, y)$ un échantillon d'observations de $\zeta = (\xi, \eta)$, dans lequel $X \in M_{NK}(\mathbf{R})$ est la **matrice d'observation** constituée des N observations des K variables exogènes ξ et y est le vecteur des N observations de η.

Pour toute observation (X_n, y_n) de (ξ, η), l'équation (1) est « observée » selon :

(3) $\quad y_n = f(X_n, b) + u_n, \forall\, n \in N_N^*,$

soit, de façon compacte :

(4) $\quad y = F(b) + u,$

où $F : \mathbf{R}^Q \mapsto \mathbf{R}^N$ est une fonction numérique vectorielle dépendant de X. On a donc aussi :

(5) $\quad G(t / \xi = X_n) = F_\varepsilon(t - f(X_n, b)), \qquad \forall\, n \in N_N^*.$

Etant donné $p \in]0, 1[$, on appelle **estimateur quantilaire**, ou **estimateur par quantile**, de $b \in \mathbf{R}^Q$ l'estimateur b_p^* associé au p-**quantile empirique** $q_p(u)$ calculé à partir des variables (inobservables) u_n qui sont les coordonnées de la **perturbation** vectorielle u.

Une méthode d'estimation associée à ce contexte est appelée **méthode des quantiles**, ou **méthode quantilaire**, ou **méthode par quantile**. Elle revient à résoudre le problème de **programmation mathématique** suivant :

(6) $\min_{b \in \mathbf{R}Q} \phi_N(b)$

(où **R**Q désigne $\mathbf{R}^Q$), dans lequel $\phi_N(b)$ est la **combinaison linéaire convexe** suivante :

(7) $\phi_N(b) = \Sigma_{n \in N+}\ p \,.\, |y_n - f(X_n, b)| + \Sigma_{n \in N-} (1 - p) \,.\, |y_n - f(X_n, b)|$,

avec $N^+ = \{n \in N_N^* : y_n \geq f(X_n, b)\}$ et $N^- = \{n \in N_N^* : y_n < f(X_n, b)\}$.

L'ensemble $E_Q(p)$ des solutions $b_p^\#$ ainsi obtenues ne se réduit pas nécessairement à un point (toujours noté $b_p^\#$ ou aussi $B_p^\#$) de $\mathbf{R}^Q$.

(ii) Dans le cas d'un **modèle linéaire**, Q = K et f est bilinéaire sur $\mathbf{R}^K$ (ie $f(\xi, b) = \xi' b$) (cf **forme multilinéaire**). Si l'on pose $\mathcal{N}_K^* = \{\sigma(1), ..., \sigma(K)\}$ (partie constituée de K éléments $\sigma(k) \in N_N^*$, avec $K < N$), $\mathcal{P}_K = \{\mathcal{N}_K^*\}$ et $(\mathcal{N}_K^*)^c = N_N^* \setminus \mathcal{N}_K^*$, on peut partitionner X et y selon :

(8)
$X = [X^1 /// X^2]$ (matrice à deux blocs),
$y = [y^1 /// y^2]$ (vecteur à deux blocs),

avec la (K,K)-**matrice** $X^1 = \{X_n : n \in \mathcal{N}_K^*\}$ et la (N-K,K)-matrice $X^2 = \{X_n : n \in (\mathcal{N}_K^*)^c\}$ (de même pour y^1 et y^2), où /// désigne un saut de lignes.

Par suite, on montre que :

(a) si l'on pose $\mathcal{N}(K) = \{\mathcal{N}_K^* : \text{rg } X^1 = K\}$ et si rg X = K, alors l'ensemble $E_Q(p)$ des estimateurs p-quantilaires possède au moins un élément de la forme :

(9) $b_p^\# = (X^1)^{-1} y^1$ pour un certain $\mathcal{N}_K^* \in \mathcal{N}(K)$.

De plus, $E_Q(p)$ est l'**enveloppe convexe** de l'ensemble $\{b_p^\#\}$ des solutions de la forme (9) ;

(b) si l'on note explicitement $b_p(X, y)$ au lieu de b_p et $E_Q(p, X, y)$ au lieu de $E_Q(p)$, on établit les propriétés suivantes :

$b_p(X, \lambda y) = \lambda b_p(X, y)$, $\forall \lambda \in \mathbf{R}_+$,

$$(10)\quad \begin{aligned} & b_{1-p}(X, \lambda y) = \lambda\, b_p(X, y), && \forall\, p \in]0, 1[\text{ et } \forall\, \lambda \in \mathbf{R}_- = -\mathbf{R}_+ , \\ & b_p(X, y + X b) = b_p(X, y) + b, && \forall\, b \in \mathbf{R}^K, \\ & b_p(X . R, y) = R^{-1}\, b_p(X, y), && \forall\, R \in R_K(\mathbf{R}). \end{aligned}$$

(iii) Si F_ε est continue, la solution $b_p^\#$ définie en (9) est P-p.s. l'unique solution du problème (5) ssi :

$$(10)\quad (p - 1)\, e_K' < \Sigma_{n \in \mathcal{N}K^*} \{(1 / 2) - (1 / 2) . \mathrm{sgn}\,(y_n - X_n\, b_p^\#) - p\} - X_n\, (X^1)^{-1} < p\, e_K',$$

où $e_K = (1,..., 1)' \in \mathbf{R}^K$ est le premier vecteur bissecteur de $\mathbf{R}^K$ et $\mathcal{N}K^*$ désigne par commodité $\mathcal{N}_K^*$.

(iv) Le programme (6) équivaut au programme suivant :

$$(11)\quad \min_{(\beta,\, u+ ,\, u-)} \{p\, e_N'\, u^+ + (1 - p)\, e_N'\, u^-\},$$

contraint par la condition :

$$(12)\quad y = F(\beta) + u^+ - u^-,$$

où $u = u^+ - u^-$ est la décomposition de u en sa **partie positive** $u^+ = (u_1^+ ,..., u_N^+)$, avec $u_n^+ = \sup\,(u_n , 0)$, $\forall\, n \in N_N^*$, et sa partie négative $u_n^- = (u_1^- ,..., u_N^-)$, avec $u_n^- = \sup\,(-u_n, 0)$, $\forall\, n \in N_N^*$.

(v) Ainsi, le cas particulier où p = 1/2 correspond à la **méthode de la médiane** et l'**estimateur de la médiane** $b_{1/2}^\#$ de b dans le modèle (1) n'est autre que l'estimateur des moindres écarts absolus, ie la solution du programme min $||u||_1$ pr à $b \in \mathbf{R}^Q$ (cf **méthode des moindres écarts absolus**).