ÉQUATION ESTIMANTE (H3, H7)
(01 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

Une **équation estimante** est une équation, ou un système d'équations, dont la résolution fournit un **estimateur** pour un **paramètre**.

(i) Soit ($\mathscr{X}$, $\mathscr{B}$, $(P_\theta^X)_{\theta \in \Theta}$) un **modèle image** dans lequel X représente une **variable observable** (**échantillon**), ou une **statistique** observable, et θ un **paramètre d'intérêt**. à valeurs dans un **ensemble** Θ (eg un **espace vectoriel** réel donné).

On appelle **équation estimante** de θ toute équation de la forme :

(1) ψ (X, θ) = 0,

dans laquelle $\psi : \mathscr{X} \times \Theta \mapsto \Theta$ est une **application mesurable** donnée, appelée **fonction estimante**.

Cette fonction définit (ou caractérise) la **méthode d'estimation** mise en oeuvre, et toute solution $\theta^{\sim}$ = t (X) de (1) est appelé **estimateur** de θ, ie estimateur de la **vraie valeur** θ^* du paramètre θ.

Des exemples importants d'équation estimante résultent des méthodes générales d'estimation suivantes :

(a) **méthode des moments**. L'équation d'identification des moments empiriques aux moments théoriques est une équation estimante ;

(b) **méthode du maximum de vraisemblance**. Le calcul de l'**estimateur du maximum de vraisemblance** est basé sur l'annulation du **gradient** de la **fonction de vraisemblance** : cette condition du premier ordre est un autre exemple d'équation estimante ;

(c) **méthode de moindre norme**, ou méthode de moindre distance, notamment associée à une distance quadratique (cf **forme quadratique**). Ainsi, la **méthode des moindres carrés généralisés**, appliquée à un **modèle de régression linéaire** (écrit dans un **espace d'observation** $\mathbf{R}^N$) y = z + u, avec z = X b, E u = 0 et V u = Σ, conduit à l'équation estimante suivante :

(2) $X'\, y \ = \ (X'\, \Sigma^{-1}\, X)^{-1}\, X'\, X\, b$.

Lorsque la **variable endogène** y est gaussienne, cette équation est souvent appelée **équation normale**, ou **système des équations normales**.

(ii) Soit y = z + u, avec z = X b, E u = 0 et V u = $\sigma^2 \,.\, I_N$, un **modèle de régression linéaire** observé dans $\mathbf{R}^N$. On suppose que :

(a) les vecteurs colonnes x_1 ,..., x_K composant la (N,K)-**matrice** X sont aléatoires, mais non asymptotiquement corrélés en probabilité avec u (ie $\text{plim}_N\, N^{-1}\, (X'\, u) = 0 \in \mathbf{R}^K$) ;

(b) les statistiques $N^{-1} X' y$ et $N^{-1} X' X$ convergent en probabilité (resp presque sûrement) vers les **moments** théoriques, notés resp M_{Xy} et M_{XX}, ie (cf **convergence en probabilité**, **convergence presque sûre**) :

$$\text{(3)} \quad \begin{cases} \mathrm{plim}_N\, N^{-1} (X' y) = M_{Xy} , \\ \mathrm{plim}_N\, N^{-1} (X' X) = M_{XX} . \end{cases}$$

On peut écrire (après prémultiplication par $N^{-1} X'$) :

(4) $N^{-1} X' y = N^{-1} X' z + N^{-1} X' u.$

Par suite :

(5) $\mathrm{plim}_N\, N^{-1} (X' y) = \mathrm{plim}_N\, N^{-1} (X' X)\, b,$

ie :

(6) $M_{Xy} = M_{XX}\, b.$

On appelle **équation estimante** (de b) associée à (6) l'équation empirique analogue à (4) dans laquelle on exclut (ou néglige) le second terme du second membre, ie :

(7) $N^{-1} X' y = N^{-1} X' X\, b$ (ou encore $X' y = X' X\, b$).

L'équation (7) n'est autre que l'équation normale résultant de la **méthode des moindres carrés ordinaires** appliquée au modèle de régression.

(iii) La procédure précédente peut s'interpréter comme une **méthode des moments**.

Elle se généralise au cas où $V\, u = \Sigma > 0$ (positivité au sens des **formes quadratiques** associées), en écrivant $N^{-1} X' \Sigma^{-1} y = N^{-1} X' \Sigma^{-1} z + N^{-1} X' \Sigma^{-1} u$, et en posant :

$$\text{(3)'} \quad \begin{cases} M_{Xy} (\Sigma) = \mathrm{plim}\, (N^{-1} X' \Sigma^{-1} y), \\ M_{XX} (\Sigma) = \mathrm{plim}\, (N^{-1} X' \Sigma^{-1} X). \end{cases}$$

Par suite, les analogues de (6) et (7) sont :

(6)' $M_{Xy} (\Sigma) = M_{XX} (\Sigma)\, b,$

(7)' $N^{-1} X' \Sigma^{-1} y = N^{-1} X' \Sigma^{-1} X\, b + N^{-1} X' \Sigma^{-1} u,$

en supposant que Σ est connue (cf **méthode des mcg**).

(iv) Dans le modèle initial, où $V\, u = \sigma^2 . I_N$, si les vecteurs x_k sont supposés aléatoires mais asymptotiquement corrélés avec u (ie si $\mathrm{plim}_N\, N^{-1} X' u \neq 0$) et s'il existe des **variables instrumentales** v_l ($\forall\, l \in N_L^*$) formant une (N,L)-**matrice** V et non asymptotiquement corrélées avec u (ie tq $\mathrm{plim}_N\, N^{-1} V' u = 0$), on peut écrire :

(8) $N^{-1} H' y = N^{-1} V' z + N^{-1} V' u = N^{-1} V' X b + N^{-1} V' u.$

Par suite, en posant :

(9)
$$M_{Vy} = \text{plim } (N^{-1} V' y),$$
$$M_{VX} = \text{plim } (N^{-1} X' X),$$

on obtient :

(10) $M_{Vy} = M_{VX}\, b.$

On appelle alors **équation estimante**, ou ici **équation instrumentale**, ou encore **équation quasi-normale**, (de b) associée à (10) l'équation « empirique » suivante :

(11) $N^{-1} V' y = N^{-1} V' X b$ (ou encore $V' y = V' X b$).

En général, on admet l'hypothèse rg M_{VX} = rg $V' X = K$ (ce qui implique $K \leq L$). Dans la **méthode des variables instrumentales**, l'estimateur de b ainsi obtenu s'explicite selon : $b_V^{\sim} = (V' X)^{-1} V' y$.

(v) Ce qui précède s'étend au **modèle non linéaire** $y = z + u = F(b) + u$, avec $E\, u = 0$ et $V\, u = \sigma^2 . I_N$. S'il existe des **variables** « auxiliaires » g_q ($\forall\, q \in N_Q^*$) formant la (N,Q)-matrice G (qui dépend, en général, de X) non asymptotiquement corrélées avec u (ie tq $\text{plim}_N (N^{-1} G' u) = 0$), on peut écrire :

(12) $N^{-1} G' y = N^{-1} G' F(b) + N^{-1} G' u,$

ie, en posant $M_{Gy} = \text{plim } (N^{-1} G' y)$ et $M_{GF} = \text{plim } (N^{-1} G' F(b))$:

(13) $M_{Gy} = M_{GF}(b).$

On appelle alors **équation estimante** de b l'équation empirique analogue à (13), ie :

(14) $N^{-1} G' y = N^{-1} G' F(b)$ (ou aussi $G' y = G' F(b)$).

Pour résoudre (14), il est nécessaire, en posant $H(b) = G' F(b)$, que l'application dérivée $D H \in \mathscr{L}(\mathbf{R}^Q)$ soit de plein **rang** (rg (D H) = Q) dans un **voisinage** de la **vraie valeur** $b^* \in \mathbf{R}^Q$ de b. Dans ce cas, b est estimé par le vecteur aléatoire $b_G^{\sim} = H^{-1} (G' y)$.

En particulier, si $G = D F(b)$, on retrouve en (14) l'équation normale du modèle non linéaire. Si Q = K et si F est linéaire (ie si $F(b) = X b$), on retrouve le modèle linéaire (pr à b) initial.

(vi) Il existe d'autres **équations estimantes**, eg celles dérivées des estimations robustes (cf **robustesse**, etc) ou des méthodes avec pénalisation (cf eg **méthode des moindres carrés pénalisés**).

A titre d'exemple, si $X = (X_1, ..., X_N)$ est une **suite** (**échantillon aléatoire**) de taille N, si $\Theta \in \mathcal{O}(\mathbf{R}^Q)$ et si $\rho : \mathcal{X} \times \Theta \mapsto \mathbf{R}^Q$ est une fonction donnée (vérifiant des hypothèses ad hoc), un **estimateur de P.J. HUBER**, ou **M-estimateur**, est une solution en θ de l'équation estimante suivante :

(15) $\psi(X, \theta) = \Sigma_{n=1}^{N} \rho(X_n, \theta) = 0.$

Cette équation généralise la condition du premier ordre associée à la méthode du maximum de vraisemblance.