ESTIMABILITÉ (C5, G, H, J)
(10 / 06 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Soit ($\mathscr{X}$, $\mathscr{B}$, $P_\theta^X)_{\theta \in \Theta}$ un **modèle statistique** (**modèle image**), paramétré par un ensemble Θ, et g : $\Omega \mapsto \mathbf{R}^Q$ une fonction mesurable (cf **application mesurable**).

On dit que le **paramètre (vectoriel) transformé** τ = g (θ) est un **paramètre estimable**, ou encore que g est une **fonction estimable**, ssi il existe une **statistique** s : $\mathscr{X} \mapsto \mathbf{R}^Q$ tq :

(1) $E_\theta\, s\,(X) = g\,(\theta), \quad \forall\, \theta \in \Theta.$

Autrement dit, si S = s o X : $\Omega \mapsto \mathbf{R}^Q$ est la **va** (ou la statistique) associée à s, on dit que g (θ) (ou g) est **estimable** ssi S est centrée (au sens de l'**espérance**) en τ = g (θ), quelle que soit la valeur de θ.

(ii) Cette définition rencontre notamment une large application dans le **contexte** du **modèle linéaire**. En effet, si le modèle suivant (écrit dans un **espace d'observations**) :

(2) $y = X\,b + u, \qquad \text{avec } E\,u = 0,\ V\,u = \sigma^2 . I_N ,$

n'est pas de plein **rang** (ie si rg X < K), la **méthode des mco** ne permet pas d'estimer directement b $\in \mathbf{R}^K$, puisque la **matrice** X' X n'est pas inversible.

On peut cependant estimer certaines fonctions (linéaires) de b.

On dit ainsi que la valeur λ' b de la **forme linéaire** λ : b $\mapsto \lambda$ ' b est **estimable** (ou parfois **identifiable**) (R.C. BOSE) ssi :

(3) $X\,b_1 = X\,b_2 \;\Rightarrow\; \lambda'\,b_1 = \lambda'\,b_2 , \qquad \forall\,(b_1 , b_2) \in \mathbf{R}^K \times \mathbf{R}^K.$

En particulier, si $\lambda'\,e_K = 0$ (somme des coefficients nulle), avec $e_K = (1 ,..., 1)' \in \mathbf{R}^K$ (premier vecteur bissecteur), la valeur λ' b est appelée **contraste** (sur b).

On montre, dans ce cadre, que λ' b (ou la forme λ) est estimable ssi l'une des conditions suivantes est vraie :

(a) $\lambda \in$ Im X' (image de la transposée X' de X) ;

(b) il existe un **estimateur sans biais** de λ' b qui est linéaire pr à y, ie $\exists$ a $\in \mathbf{R}^N$ tq :

(4) $E\,a'\,y = \lambda'\,b, \qquad \forall\, b \in \mathbf{R}^K,$

ssi rg X' b = rg [X', λ], avec X' $\in M_{K,N}$ (**R**) et $\lambda \in \mathbf{R}^K$.

Plus généralement, pour estimer le **paramètre** :

(5) $c = L b$, avec $L \in M_{JK}(\mathbf{R})$,

on dit que c (ou L) est **estimable** ssi il existe une **matrice** $A \in M_{JN}(\mathbf{R})$ (dépendant en général de X) tq :

(6) $E A y = c$.

On montre que c est estimable ssi l'une des conditions suivantes est vérifiée :

(a) $\exists A \in M_{JN}(\mathbf{R})$ tq $L = A X$;

(b) rg [X /// L] = rg X, où [X /// L] désigne la matrice colonne dont le premier bloc est X et le second L (désigne un saut de ligne) ;

(c) il existe une g-inverse (cf **matrice inverse généralisée**) L^- de L tq : $\text{rg } X (I - L^- L) = \text{rg } X - \text{rg } L$;

(d) il existe une g-inverse X^- de X tq : $L X^- X = L$;

(e) rg [X' X /// A] = rg (X' X) ;

(f) il existe une g-inverse L^- de L tq : $\text{rg } (X' X)(I - L^- L) = \text{rg } (X' X) - \text{rg } L$;

(g) il existe une g-inverse $(X' X)^-$ de X' X tq : $L (X' X)^- X' X = L$.

(iii) Le **modèle d'analyse de la variance** et le **modèle d'analyse de la covariance** font souvent usage des concepts (estimabilité, constraste) et propriétés précédents.

(iv) La notion d’estimabilité ci-dessus est la plus courante, notamment à raison de la simplicité des calculs linéaires auxquels elle s’associe naturellement.

Elle peut, en principe, s’étendre et se définir au sens d'une autre **caractéristique** de **centralité** que l'espérance : eg au sens du **mode** ou de la **médiane**.