THÉORÈME DE GAUSS - MARKOV (H3, G4, J)

Le **théorème de GAUSS - MARKOV** constitue l'un des grands résultats de la théorie de la **régression**.

(i) Dans le cadre du **problème linéaire**, on note $\mathcal{E}(\mathcal{L})$ l'ensemble des **estimateurs** t définis par les propriétés suivantes :

(a) t est un estimateur linéaire (pr à y) de $c_P = A\, m_P$;

(b) t vérifie le **principe de réduction** suivant :

(1) $t\,(\mathcal{L}) = A\,(\mathcal{L})$.

Deux critères de comparaison entre estimateurs appartenant à $\mathcal{E}(\mathcal{L})$ sont les suivants. On préfère t_1 à t_2 (notation $t_1 \geq t_2$) ssi :

(a) **premier critère** (critère de **dispersion** scalaire quadratique, ou de perte en **norme** quadratique) :

(2) $E_P\,||t_1\,(y) - A\,m_P||^2 \leq E_P\,||t_2\,(y) - A\,m_P||^2, \quad \forall\, P \in \mathcal{P}$;

(b) **second critère** (critère de dispersion vectorielle quadratique, ou de perte vectorielle quadratique) :

(3) $D\,(t_1\,(y) - A\,m_P) \leq D\,(t_2\,(y) - A\,m_P), \quad \forall\, P \in \mathcal{P}$,

qui s'explicite selon :

(3)' $h'\,\{D\,(t_1\,(y) - A\,m_P)\} \leq h'\,\{D\,(t_2\,(y) - A\,m_P)\}, \quad \forall\, h \in \mathbf{R}^Q, \forall\, P \in \mathcal{P}$,

ou encore ;

(3)'' $E_P\,\{h'\,(t_1\,(y) - A\,m_P)\}^2 \leq E_P\,\{h'\,(t_2\,(y) - A\,m_P)\}^2, \quad \forall\, h \in \mathbf{R}^Q, \forall\, P \in \mathcal{P}$.

(ii) C'est souvent le second critère qui est choisi. Dans ce cas, le **théorème de C.F. GAUSS - A.A. MARKOV** se présente sous plusieurs formes :

(a) **forme géométrique**. Soit Π le **projecteur** orthogonal de $\mathbf{R}^N$ sur $\mathcal{L}$ et $t \in \mathcal{E}(\mathcal{L})$. Alors les propriétés suivantes sont équivalentes :

(a_1) $t = A\,\Pi$;

(a_2) $D\,(t\,(y) - A\,m_P) \leq D\,(s\,(y) - A\,m_P), \forall\, s \in \mathcal{E}\,(\mathcal{L}), \forall\, P \in \mathcal{P}$;

(a_3) $E_P \|t(y) - A\, m_P\|^2 \leq E_P \|s(y) - A\, m_P\|^2$, $\forall\, s \in \mathcal{E}(\mathcal{L})$, $\forall\, P \in \mathcal{P}$;

(b) **forme paramétrique**. Soit $X \in M_{NK}(\mathbf{R})$ et $B \in M_{QK}(\mathbf{R})$. On admet, outre les hypothèses usuelles du modèle de régression linéaire multiple (cf **méthode des moindres carrés ordinaires**), les suivantes :

(b_1) il existe $b_P \in \mathbf{R}^K$ tq :

(4) $\quad m_P = E_P\, y - X\, b_P$, $\forall\, P \in \mathcal{P}$;

(b_2) le nouveau **paramètre** $B\, b_P = d_P$ est estimable, ie il existe une matrice $A \in M_{QN}(\mathbf{R})$ tq :

(5) $\quad B = A\, X$;

(b_3) on se restreint à l'ensemble $\mathcal{E}_A(\mathcal{L})$ des estimateurs tq :

(6) $\quad t \in \mathcal{E}_A(\mathcal{L}) \Leftrightarrow \{T \in M_{QN}(\mathbf{R})$ et $T\, \Pi = A\, \Pi\}$,

autrement dit T et A coïncident sur Im X = $\mathcal{L}$ (où T désigne la **matrice** associée à t).

Alors, pour tout estimateur $t \in \mathcal{E}_A(\mathcal{L})$, les propriétés suivantes sont équivalentes :

(b_4) $T = A\, \Pi$;

(b_5) $V_P(T\, y) \leq V_P(S\, y)$, $\forall\, s \in \mathcal{E}_A(\mathcal{L})$ de matrice associée S et $\forall\, P \in \mathcal{P}$;

(c) **forme classique**. On suppose, outre ce qui précède, que X est injective (ie $X \in \text{Mono}(\mathbf{R}^K, \mathbf{R}^N)$), donc que $X'\, X \in R_K(\mathbf{R})$ (**matrice régulière**), et l'on pose :

(c_1) $T^\wedge = (X'\, X)^{-1}\, X'$;

(c_2) $t \in \mathcal{E} \Leftrightarrow T \in M_{KN}(\mathbf{R})$ et $T\, X = I_K$ (où T désigne la matrice de t).

Alors, t (resp T) vérifie les trois propriétés équivalentes suivantes :

(c_3) $T = T^\wedge$;

(c_4) $V_P(t(y)) = V_P(s(y))$, $\forall\, s \in \mathcal{E}$, $\forall\, P \in \mathcal{P}$;

(c_5) pour tout $y \in \mathbf{R}^N$, la **propriété des moindres carrés ordinaires** suivante est vérifiée :

(7) $\quad \|y - X\, T\, y\|^2 = \inf_{b \in \mathbf{R}^K} \|y - X\, b\|^2$.

(iii) Dans la forme classique (c) du théorème on a donc :

(8) $\quad t \in \mathcal{E} \Leftrightarrow t(y)$ estime b_P sans biais (ie $E\, t(y) = b_P$).

On ajoute parfois, à cette forme du théorème, la propriété suivante : si dim $\mathscr{L}$ = L est donnée (en particulier, si dim $\mathscr{L}$ = dim (Im X) = K), alors :

(9) $(N-L)^{-1} . ||(I_N - \Pi) y||^2$ est un **estimateur sans biais** de σ_P^2.

(iv) Le modèle de régression linéaire multiple, écrit dans l'**espace des observations** selon y = X b + u (où E u = 0, V y = V u = $\sigma_P^2 I_N$), est un exemple d'application dans lequel on estime b par la **méthode des mco** (ie b est estimé par $b^\wedge = (X' X)^{-1} X' y$). On cherche alors à estimer le nouveau paramètre $d_P = B b_P$.

Le théorème précédent indique que, dans la classe des estimateurs linéaires (pr à y) $d^\sim$ de d_P , supposés sans biais, l'unique estimateur de d_P tq la matrice $V_P d^\sim - V_P d$ est semi-définie positive (cf **matrice définie positive**) n'est autre que $d^\wedge = B b^\wedge$, où $d^\sim$ désigne un estimateur linéaire sans biais quelconque.

Il constitue un cas particulier, très utilisé, du **théorème de AITKEN-GAUSS-MARKOV**, dans lequel la dispersion V u ≠ $\sigma^2 I_N$.