

MODÈLE À ERREURS COMPOSÉES (J9)

(18 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

De façon générale, un **modèle à erreurs composées** est un **modèle** dans lequel les « **erreurs** » (erreurs sur équations ou erreurs sur variables) sont décomposables. Leurs composantes sont des **variables aléatoires** qui affectent les « **variables principales** », ou « **variables d'intérêt** » (cf eg **modèle à erreurs sur les variables**).

(i) Ainsi, un **modèle de régression** non linéaire (cf **modèle non linéaire**), écrit dans l'**espace des variables** selon :

$$(1) \quad \eta^* = f(\xi^*, b) + \varepsilon, \quad \text{avec } E \varepsilon = 0 \text{ et } b \in \mathbf{R}^Q,$$

est appelé **modèle à erreurs composées** ssi :

(a) la **perturbation aléatoire** ε peut représenter sous la forme (additive) $\varepsilon = \alpha + \beta + \gamma + \dots$, avec $\alpha \sim \mathcal{L}(\alpha)$, $\beta \sim \mathcal{L}(\beta)$, $\gamma \sim \mathcal{L}(\gamma)$, ... (**lois de probabilité** resp des va $\alpha, \beta, \gamma, \dots$) ;

(b) ou encore ssi eg la « vraie » **variable exogène** ξ^* est observée avec **erreur** φ selon (forme additive) $\xi = \xi^* + \varphi$ et la vraie **variable endogène** η avec erreur ψ selon $\eta = \eta^* + \psi$, où eg $\varphi = \alpha_\varphi + \beta_\varphi + \gamma_\varphi$ et $\psi = \alpha_\psi + \beta_\psi + \gamma_\psi$ (formes additives).

La décomposition des erreurs peut ne pas être additive : eg la **lp** d'une erreur peut résulter d'un **mélange de lois** composantes, ou d'une **convolution des lois** qui la composent, etc.

(ii) Plus spécifiquement, on appelle **modèle à erreurs composées** un modèle décrivant des données multi-indices dont la partie aléatoire (perturbation) admet une décomposition additive en fonction des **indices** (cf aussi **modèle à variance composée**). C'est donc la (seule) **perturbation aléatoire** qui est décomposée dans ce type de modèles.

Le cas du **modèle bi-indicé**, dans lequel les observations sont des **séries temporelles** de **coupes instantanées**, est typique. Soit $(\xi, \eta) : \Omega \mapsto \mathbf{R}^K \times \mathbf{R}$ un **couple aléatoire** dans lequel la liste des variables exogènes est notée $\xi = (\xi_1, \dots, \xi_K)$ et la variable endogène η . On suppose ce couple lié par une **régression** non linéaire (cf aussi **fonction de régression**), écrite dans l'**espace des variables** selon :

$$(2) \quad \eta = f(\xi, b) + \varepsilon, \quad \text{avec } E \varepsilon = 0 \text{ et } b \in \mathbf{R}^Q,$$

et qu'il est observé selon deux « dimensions » : eg la dimension individuelle $N_N^* = \{1, \dots, N\}$ et la dimension temporelle $N_T^* = \{1, \dots, T\}$. L'équation (1) s'exprime (dans l'**espace des observations**) selon le **modèle de régression** non linéaire :

$$(3) \quad y_{nt} = f(X_{nt}, b) + u_{nt},$$

avec $E u_{nt} = 0, \forall (n, t) \in N_N^* \times N_T^*$. Autrement dit, les observations sont constituées de $1 + K$ **tableaux statistiques** à deux dimensions $Y = (y_{nt})_{(n,t)}$ et $X_k = (x_{k,nt})_{(n,t)}, \forall k \in N_K^*$. On pose, en conformité, $u = (u_{nt})_{(n,t)}$.

On dit que (3) est un **modèle à erreurs composées** ssi la **perturbation** u se décompose, pour chaque observation d'indice (n, t) , selon :

$$(4) \quad u_{nt} = a_n + b_t + c_{nt},$$

avec $E a_n = E b_t = E c_{nt} = 0, \forall (n, t)$. Le terme a_n est appelé **composante individuelle** de u_{nt} , le terme b_t sa **composante temporelle** et le terme c_{nt} sa **composante mixte** (perturbation résiduelle commune individu-temps).

En pratique :

(a) les hypothèses du modèle supposent que les composantes sont sans **corrélacion**, ie :

$$(5) \quad C(a_n, b_t) = C(a_n, c_{nt}) = C(b_t, c_{nt}) = 0, \quad \forall (n, t).$$

On les suppose aussi homoscedastiques et sans autocorrélacion (cf **homoscedasticité, autocorrélacion**), ie :

$$C(a_\alpha, a_\beta) = \delta_{\alpha\beta} \cdot \sigma_a^2, \quad \forall (\alpha, \beta) \in (N_N^*)^2,$$

$$(6) \quad C(b_s, b_t) = \delta_{st} \cdot \sigma_b^2, \quad \forall (s, t) \in (N_T^*)^2,$$

$$C(c_{\alpha s}, c_{\beta t}) = \delta_{((\alpha,s),(\beta,t))} \cdot \sigma_c^2, \quad \forall ((\alpha, \beta), (p, t)) \in (N_N^*)^2 \times (N_T^*)^2.$$

On pose $F_{nt}(b) = f(X_{nt}, b)$ et l'on « empile » les observations dans l'ordre du produit cartésien $N_N^* \times N_T^*$, ie (cf **index des notations**, in fine) :

$$y = [y_{11} \text{ /// } y_{1T} \text{ /// } \dots \text{ /// } y_{N1} \text{ /// } y_{NT}],$$

$$(7) \quad F(b) = [F_{11}(b) \text{ /// } F_{1T}(b) \text{ /// } \dots \text{ /// } F_{N1}(b) \text{ /// } F_{NT}(b)],$$

$$u = [u_{11} \text{ /// } u_{1T} \text{ /// } \dots \text{ /// } u_{N1} \text{ /// } u_{NT}],$$

où le symbole /// dénote un saut de ligne.

La forme (3) se ramène à celle d'un **modèle de régression multiple** non linéaire :

$$(8) \quad y = F(b) + u, \quad \text{avec } E u = 0, \quad V u = \Sigma,$$

où la **dispersion** Σ , définie en $\{(4),(5),(6)\}$, s'écrit :

$$(9) \quad \Sigma = \sigma_a^2 A + \sigma_b^2 B + \sigma_c^2 I_{NT},$$

avec $A = I_N \otimes J_T$, $B = J_N \otimes I_T$ et $J_L = J_{LL}$, où $J_{LM} = e_L e_M'$ (**matrice des unités**, avec $(L, M) \in \mathbf{N}^* \times \mathbf{N}^*$);

(b) le modèle $\{(8),(9)\}$ peut être estimé par les méthodes habituelles (eg **méthode des moindres carrés**, **méthode du maximum de vraisemblance**), compte tenu de la formule :

$$(10) \quad \sigma_c^2 \Sigma^{-1} = I_{NT} - (a_T \cdot A + b_N \cdot B) + c_{NT} J_{NT},$$

dans laquelle :

$$(11)_a \quad \begin{aligned} a_T &= \sigma_a^2 / A_T, \\ b_N &= \sigma_b^2 / B_N, \\ c_{NT} &= (A_T B_N)^{-1} (A_T + B_N - \sigma_c^2)^{-1} \cdot \sigma_c^2 \sigma_b^2 (A_T + B_N), \end{aligned}$$

avec :

$$(11)_b \quad \begin{aligned} A_T &= T \sigma_a^2 + \sigma_c^2, \\ B_N &= N \sigma_b^2 + \sigma_c^2. \end{aligned}$$

(c) si le modèle (2) ou le modèle (3) sont linéaires (avec $Q = K$ et $F = X$), le modèle $\{(8),(9)\}$ est un **modèle linéaire**.

On peut alors définir l'**estimateur de la covariance** (de **A. HUSSAIN - T.D. WALLACE**) de b comme **estimateur des moindres carrés ordinaires** b_Q^\wedge calculé sur le modèle :

$$(12) \quad Q y = Q X b + Q u, \quad \text{avec } Q = I_{NT} - (T^{-1} A + N^{-1} B) + (NT)^{-1} J_{NT}.$$

La matrice Q admet pour **matrice inverse conditionnelle** la matrice $Q^c = \sigma_c^{-2} \Sigma$ (ie la matrice tq $Q Q^c Q = Q$).

L'estimateur cherché, $b_Q^\wedge = (X' Q X)^{-1} X' Q y$, est donc :

- (a) un **estimateur sans biais**, ie $E b_Q^\wedge = b$;
- (b) un estimateur de **dispersion** elliptique, ie $V b_Q^\wedge = \sigma_c^2 (X' Q X)^{-1}$;
- (c) un **estimateur convergent** en probabilité, ie $\text{plim } b_Q^\wedge = b$;
- (d) sous certaines hypothèses, un estimateur asymptotiquement gaussien (cf **normalité asymptotique**), ie :

$$(13) \quad \mathcal{L} \{(N \cdot T)^{1/2} (b_Q^\wedge - b)\} \rightarrow_{\min(N,T) \rightarrow +\infty} \mathcal{N}_K(0, \Xi),$$

avec $\Xi = \text{plim}_{(N,T)} (NT)^{-1} (X' Q X)$. On suppose notamment que $\Xi < \infty$ (matrice finie) et que le rapport N / T reste fini dans les calculs de limites précédents.