PRÉVISION DES MOINDRES CARRÉS ORDINAIRES (H3, J1)
(05 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Soit :

(1) $\eta = \xi' b + \varepsilon$, avec $E\,\varepsilon = 0$, $V\,\varepsilon = \sigma^2$,

un **modèle de régression multiple** linéaire défini dans l'**espace des variables** (ξ, η), et « observé » dans l'**espace des observations** (X, y) selon :

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

où $X \in M_{NK}$ (**R**) et y prend ses valeurs dans $\mathbf{R}^N$.

On attribue au vecteur ξ des **variables exogènes** une valeur $\xi_0 \in \mathbf{R}^K$, on pose :

(3) $\eta_0 = \xi_0' b + \varepsilon_0$, avec $E\,\varepsilon_0 = 0$, $V\,\varepsilon_0^2 = \sigma^2$,

et l'on veut prévoir la valeur η_0 associée à ξ_0 .

On appelle alors :

(a) **prédicteur (conditionnel)** de η_0 sachant ξ_0 toute **vars** (ou toute **statistique** réelle scalaire) $\eta_0^* = t(X, y)$;

(b) **erreur de prévision** la vars $e_0^* = \eta_0^* - \eta_0$.

Comme η_0 est **inobservable**, on considère que sa prévision équivaut à l'**estimation** de sa **moyenne** (ou **espérance**) $E\,\eta_0 = \xi_0' b$.

On appelle alors :

(a) **prédicteur des moindres carrés ordinaires** (mco), le **prédicteur de C.F. GAUSS - A.A. MARKOV**, (conditionnel, ie sachant ξ_0) de $E\,\eta_0$, ie le « prédicteur » défini par :

(4) $\eta_0^{\wedge} = \xi_0' b^{\wedge}$,

où $b^{\wedge}$ est l'**estimateur des mco** de b dans le modèle (2) ;

(b) **erreur de prévision (des mco)** la va $\varepsilon_0^{\wedge} = \eta_0^{\wedge} - \eta_0$.

(ii) On montre que :

(a) l'erreur de prévision est nulle en moyenne (quel que soit ξ_0) :

(5) $E(\eta_0^{\wedge} - \eta_0) = E\,\varepsilon_0^{\wedge} = 0$;

(b) les variances de $\eta_0^\wedge$ et $\varepsilon_0^\wedge$ sont, resp :

$$V\,\eta_0^\wedge = \sigma^2\,\xi_0'\,(X'\,X)^{-1}\,\xi_0\,,$$

(6)

$$V\,\varepsilon_0^\wedge = \sigma^2\,\{1 + \xi_0'\,(X'\,X)^{-1}\,\xi_0\}\,;$$

(c) si $\mathscr{L} = \{\eta_0^\# = c'\,y : \forall\, c \in \mathbf{R}^N\}$ désigne la **classe des prédicteurs linéaires** (pr à y) de E η_0 et $\mathscr{C} = \{\eta_0^\# \in \mathscr{L} : E\,\eta_0^\# = E\,\eta\}$ la **classe des prédicteurs linéaires sans biais** de E η_0 , alors $\eta_0^\wedge$ est le prédicteur dans $\mathscr{C}$ dont la variance est minimale, ie tq :

(7) $\quad E\,(\eta_0^\wedge - \eta_0)^2 \leq E\,(\eta_0^\# - \eta_0)^2, \qquad \forall\, \eta_0^\# \in \mathscr{C}\,;$

(d) si $u \sim \mathscr{N}_N\,(0, \sigma^2\,I_N)$ (**loi normale multidimensionnelle** centrée), alors $\varepsilon_0^\wedge \sim \mathscr{N}_1\,(0, V\,\varepsilon_0^\wedge)$, où $V\,\varepsilon_0^\wedge$ est donnée en (6).

Un **intervalle de confiance** (conditionnel), bilatéral et symétrique, de η_0 au seuil 1 - $\alpha \in\,]0, 1[$ est fondé sur la propriété :

(8) $\quad (V\,\varepsilon_0^\wedge)^{-1/2}\,(\eta_0^\wedge - \eta_0) = \{(\eta_0^\wedge - \eta_0) / ((\sigma^2)^\wedge\,(1 + \xi_0'\,(X'\,X)^{-1}\,\xi_0)\}^{1/2} \sim \mathscr{S}_{N-K}$

(**loi de STUDENT** à N - K degrés de liberté).

(iii) Si le modèle considéré n'est pas linéaire, ie si :

(9) $\quad \eta = f\,(\xi, b) + \varepsilon, \qquad$ avec $E\,\varepsilon = 0$, $V\,\varepsilon = \sigma^2$,

est observé selon :

(10) $\quad y = F\,(b) + u, \qquad$ avec $E\,u = 0$ et $V\,u = \sigma^2\,.\,I_N$,

on peut suivre la même démarche que précédemment :

(a) on pose $\eta_0 = f\,(\xi_0\,, b) + \varepsilon_0$;

(b) on estime E η_0 à l'aide du prédicteur des mco (non linéaires) $\eta_0^\wedge = f\,(\xi_0\,, b_f^\wedge)$, où $b_f^\wedge$ est l'**estimateur des mco** non linéaires de b.

Cependant, comme $b_f^\wedge$ est en général biaisé ($E\,b_f^\wedge \neq b$), $\eta_0^\wedge$ est aussi biaisé.