PRÉCISION D'UNE RÉGRESSION (H5, J1)
(05 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

(i) Soit $y = F(b) + u$, avec $E\ u = 0$ et $V\ u = \sigma^2 . I_N$, un **modèle de régression** non linéaire (cf **modèle non linéaire**). On pose :

(1) $X_{(k)} = [x_1 ,..., x_k], \quad \forall k \in N_K^*,$

où x_k désigne le k-ième vecteur colonne des observations relatives à la **variable exogène** ξ_k . On note, conformément à (1) :

(2) $X_{(k)n} = [x_{n1} ,..., x_{nk}], \quad \forall (k , n) \in N_K^* \times N_N^*,$

puis :

(3) $z_{(k)n} = f_{(k)} (X_{(k)n} , b) = F_{(k)n} (b)$

et $z_{(k)} = F_{(k)} (b)$, où $F_{(k)} : \mathbf{R}^Q \mapsto \mathbf{R}^N$, $\forall k \in N_K^*$.

(ii) On considère alors la **suite** de modèles non linéaires suivants :

(4) $y = z_{(k)} + u_{(k)} = F_{(k)} (b) + u_{(k)} , \quad \forall k \in N_K^*,$

dans lesquels $u_{(k)}$ désigne un **vecteur aléatoire** à valeurs dans $\mathbf{R}^N$. On suppose que le **paramètre** $b \in \mathbf{R}^Q$ est estimé selon une méthode donnée (eg la **méthode des moindres carrés ordinaires**). On peut déduire de (4) une suite $(b_{(k)}\tilde{})_{k \in \{1 ,..., K\}}$ d'**estimateurs** de b associée à la suite de modèles précédente, et poser :

(5)
$y_{(k)}\tilde{} = F_{(k)} (b_{(k)}\tilde{})$ (**valeur ajustée** de y associée au modèle n° k),
$u_{(k)}\tilde{} = y - y_{(k)}\tilde{}$ (**résidu** relatif au modèle n° k).

On appelle alors parfois **précision de la régression**, ou **précision de l'ajustement**, (4) le carré de la **norme** du vecteur $u_{(k)}\tilde{}$, ie :

(6) S_k^2 ou $S_{1... k}^2 = ||u_{(k)}\tilde{}||_2^2 = ||y - F_{(k)} (b_{(k)}\tilde{})||_2^2.$

Ce nombre est donc le carré de la **distance** entre y et la sous-variété $V_{(k)} = \{z_{(k)} \in \mathbf{R}^N : z_{(k)} = F_{(k)} (\beta), \forall \beta \in \mathbf{R}^Q\}$ de $\mathbf{R}^N$:

(7) $S_k^2 = \inf_{z(k) \in V(k)} ||y - z^k||^2,$

où z(k) désigne $z_{(k)}$ et V(k) désigne $V_{(k)}$.

(iii) La notion de précision d'une régression permet alors de définir celle de **gain de précision**, ou **rendement partiel**, obtenu lorsqu'on passe du modèle n° k au modèle n° k + 1, $\forall k \in N_K^*$, eg :

(8) $G_k^2 = (S_{k+1}^2 - S_k^2) / S_k^2, \quad \forall\, k \in N_{K-1}^*.$