MAXIMUM DE VRAISEMBLANCE GÉNÉRALISÉ (H3)
(21 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

La **méthode du maximum de vraisemblance généralisée** est une extension assez directe de la **méthode du maximum de vraisemblance**.

(i) On considère un **problème d'estimation** paramétrique basé sur un **modèle image** $(\mathcal{X}, \mathcal{B}, (P_\theta^X)_{\theta \in \Theta})$ dans lequel :

(a) $\mathcal{X} = \mathbf{R}^N$ et $\Theta = \mathbf{R}$;

(b) P_θ^X admet f pour **densité** (**vraisemblance**) pr à une **mesure positive** σ-finie μ donnée, avec $f(x, \theta) = (dP_\theta^X / d\mu)(x)$;

(c) la **Log-vraisemblance** $L(x, \theta) = \mathrm{Log}\, f(x, \theta)$, est définie $\forall (x, \theta) \in \mathcal{X} \times \Theta$;

(d) les coordonnées X_n ($n = 1,\dots, N$) de X sont des **copies** iid (cf **suite iid**) d'une **variable parente** $\xi : \Omega \mapsto \mathbf{R}$ (avec $\xi \sim P^\xi$).

La méthode du maximum de vraisemblance peut s'étendre s'il est possible de définir une fonction :

(1) $\varphi : \mathcal{X} \times \Theta \mapsto \mathbf{R}$

qui soit continue et dérivable pr à θ (cf **application continue**, **dérivabilité**, **différentiabilité**).

La **formule des accroissements finis** (partielle pr à θ) s'écrit :

(2) $\varphi(x, \tau) = \varphi(x, \theta) + D_2\, \varphi(x, \theta_0) \,.\, (\tau - \theta)$, avec $\theta_0 \in]\theta, \tau[$,

où $D_2\, \varphi(x, \theta) = d\varphi(x, \theta) / d\theta$.

En notant θ^* la **vraie valeur** du **paramètre** θ, on suppose que :

(a) l'**espérance** $E_{\theta^*}\, \varphi(X, \theta) = \int \varphi(x, \theta)\, dP_{\theta^*}^X(x)$ existe et définit une fonction $\theta \mapsto E_{\theta^*}\, \varphi(X, \theta)$ continue en θ et tq :

(3) $E_{\theta^*}\, \varphi(X, \theta^*) = 0$;

(b) $E_{\theta^*}\, (\varphi(X, \theta))^2 < +\infty$, d'où $V_{\theta^*}\, \varphi(X, \theta^*) = E_{\theta^*}\, (\varphi(X, \theta^*))^2$;

(c) $E_{\theta^*}\, (D_2\, \varphi(X, \theta))$ est finie et non nulle, $\forall \theta \in \Theta$.

La **méthode du maximum de vraisemblance généralisé** consiste à choisir pour estimateur de θ la valeur $\theta_{\varphi,N} = t_N(X_1, \dots, X_N)$ qui annule (pr à θ) la **moyenne empirique** :

(4) $\int \varphi (x, \theta)\, dP_N (x) = N^{-1} \Sigma_{n=1}^{N} \varphi (X_n , \theta)$,

où P_N désigne la **loi empirique** associée à X.

(ii) L'**estimateur du maximum de vraisemblance généralisé** précédent vérifie les propriétés suivantes :

(a) $\theta_{\varphi,N} \rightarrow_{N \rightarrow +\infty} \theta^*$ (**convergence en probabilité** et **convergence presque sûre**) ;

(b) $\mathscr{L} (N^{1/2} (\theta_{\varphi,N} - \theta^*)) \rightarrow_{N \rightarrow +\infty} \mathscr{N}_1 (0, \sigma_{\theta^*}{}^2)$ (**loi normale** centrée) (**convergence en loi**), avec $\sigma_{\theta^*}{}^2 = V_{\theta^*} \varphi (X, \theta^*) / \{E_{\theta^*} (D_2 \varphi (X, \theta^*))\}^2$.

On note parfois $\theta_{\varphi,N}$ selon $\theta_N (\varphi)$ pour exprimer que, par construction, cet estimateur dépend de φ. En particulier, si $\varphi (x, \theta) = D_2 L (x, \theta)$, on obtient l'**estimateur du mv** ordinaire.

(iii) La méthode se généralise directement au cas d'un paramètre vectoriel. En effet, si $\Theta \subset \mathbf{R}^Q$, on pose :

(a) $\varphi : \mathscr{X} \times \Theta \mapsto \mathbf{R}^Q$, fonction vectorielle définie par : $(x, \theta) \mapsto \varphi (x, \theta) = (\varphi_1 (x, \theta) \ldots \varphi_Q (x, \theta))'$, dans laquelle les fonctions φ_q (q = 1 ,..., Q) vérifient les mêmes hypothèses que dans le cas scalaire précédent ;

(b) $D_2 \varphi (x, \theta)$ est la **matrice jacobienne** des **dérivées partielles** pr à θ, définie par : $D_2 \varphi (x, \theta) = \{(\partial\varphi_q / \partial\theta_r)(x, \theta)\}$, où $(q, r) \in \{1,..., Q\}^2$;

(c) $\Sigma_\theta = V_\theta \varphi (X, \theta)$ est la (Q,Q)-**matrice de covariance** du **vecteur aléatoire** $\varphi (X, \theta)$.

L'estimateur obtenu possède des propriétés analogues à celles du cas scalaire, ie :

(a) $\theta_{\varphi,N} (\varphi) \rightarrow_{N \rightarrow +\infty} \theta^* \in \Theta \subset \mathbf{R}^Q$ (convergences en probabilité et presque sûre, $\forall\, \theta^* \in \Theta$) ;

(b) $\mathscr{L} (N^{1/2} (\theta_{\varphi,N} (\varphi) - \theta^*)) \rightarrow_{N \rightarrow +\infty} \mathscr{N}_Q (0, (\Sigma_{\theta^*} (\varphi))^{-1})$, avec $\Sigma_\theta (\varphi) = (E_\theta D_2 \varphi (X, \theta))\, \Sigma_\theta^{-1} (E_\theta D_2 \varphi (X, \theta))'$, $\forall\, \theta \in \Theta$.

(iv) D'autres généralisations de la méthode du mv prennent en compte :

(a) des **contraintes sur les variables** ou des **contraintes sur les observations** ;

(b) des **contraintes sur les paramètres** ;

(c) l'extension du **modèle paramétrique** au cas non paramétrique (cf **méthode non paramétrique**).