MAXIMUM DE VRAISEMBLANCE CONTRAINT (A11, H3)
(18 / 06 / 2020, © Monfort, Dicostat2005, 2005-2020)

La **méthode du maximum de vraisemblance contraint** (ou sous contrainte) s'applique lorsque le **paramètre** d'un modèle est astreint à parcourir une partie stricte (eg une **partie bornée**) de l'espace des paramètres (cf **méthode du mv**).

(i) Soit $(\Omega, \mathcal{T}, P_\theta)_{\theta \in \Theta}$ un **modèle statistique**, $(\mathcal{X}, \mathcal{B}) = (\mathcal{X}_0^N, \mathcal{B}_0^{\otimes N})$ un **espace d'échantillonnage** et $X : \Omega \mapsto \mathcal{X}$ une **va** donnée (eg un **échantillon**). On note $(\mathcal{X}, \mathcal{B}, (P_\theta^X)_{\theta \in \Theta})$ le **modèle image**. On suppose que $\Theta = \mathbf{R}^Q$ et qu'il existe une **application mesurable** $h : \mathbf{R}^Q \mapsto \mathbf{R}^L$ (avec $L < Q$) tq $\theta \in \Lambda$, variété définie selon :

(1) $\Lambda = \{\theta \in \Theta : h(\theta) = 0\}$.

Si μ est une **mesure positive** σ-finie, définie sur $\mathcal{B}$, et dominant la famille $(P_\theta^X)_{\theta \in \Theta}$ (cf **famille de lois dominée**), alors $f : \mathcal{X} \times \Theta \mapsto \mathbf{R}_+$ désigne la **fonction de vraisemblance** du modèle, ie $f(x, \theta) = dP_\theta^X(x) / d\mu(x)$, $\forall (x, \theta)$.

On appelle **estimateur du maximum de vraisemblance contraint** de θ toute solution en τ du problème de **programmation mathématique** suivant :

(2) $\max_{\tau \in \Lambda} f(x, \tau)$.

(ii) Soit θ^* la **vraie valeur** du paramètre ($\theta^* \in \Lambda$). Sous certaines **conditions de régularité**, on montre que :

(a) si $I_N(\theta)$ désigne la **matrice d'information** de FISHER, et si la **matrice** :

(3) $M_N = I_N(\theta^*) + (D h(\theta^*))' (D h(\theta^*))$

est une **matrice définie positive**, alors il existe une solution locale $\theta_{\Lambda,N}$ au problème (2) ;

(b) s'il existe une telle solution, et si $\lim_N I_N(\theta) = I(\theta)$ existe, alors :

(1) $\theta_{\Lambda,N}$ est à la fois un **estimateur convergent** (en probabilité) vers θ^* et un **estimateur asymptotiquement efficace** de θ^* ;

(2) on a la **convergence en loi** :

(4) $\mathcal{L}(N^{1/2} (\theta_{\Lambda,N} - \theta^*)) \to_{N \to +\infty} \mathcal{N}_K(0, \Sigma^+)$

(**loi normale multidimensionnelle** centrée), dans laquelle Σ^+ désigne la **matrice pseudo-inverse** (au sens de MOORE-PENROSE) de la matrice suivante :

(5) $\Sigma = \{I_K - (D\,h\,(\theta^*)) + (D\,h\,(\theta^*))\,I_N\,(\theta^*)\} + I_K - \{D\,h\,(\theta^*) + D\,h\,(\theta^*)\}.$