QUASI-VRAISEMBLANCE (C5, H3)
(16 / 06 / 2020, © Monfort, Dicostat2005, 2005-2020)

La notion de **quasi-vraisemblance** est notamment utilisée en matière de **spécification** des modèles, de **comparaison de modèles** ou de **robustesse** (cf aussi **pseudo-vraisemblance**).

(i) Soit E un **espace euclidien** réel de dimension finie (Dim E = Q), et $(\Omega, \mathcal{T}, \mathcal{P})$ un **modèle statistique** tq $\mathcal{P} = (P_\theta)_{\theta \in \Theta}$ (forme paramétrique), où $\Theta \subset E$ est une **partie compacte** de E, et une **famille** $\mathcal{Q} = (Q_\theta)_{\theta \in \Theta}$ (aussi indexée par Θ) de **probabilités** Q_θ définies sur $\mathcal{T}$ (cf aussi **indice**).

Soit un **espace d'observation** $(\mathcal{X}, \mathcal{B})$ et une **va** (**échantillon**) $X : \Omega \mapsto \mathcal{X}$. On note alors $(\mathcal{X}, \mathcal{B}, \mathcal{P}^X)$ le **modèle image** du précédent par X, avec la famille des lois $\mathcal{P}^X = (P_\theta^X)_{\theta \in \Theta}$, et $\mathcal{Q}^X = (Q_\theta^X)_{\theta \in \Theta}$ la famille des lois qui sont les images de $\mathcal{Q}$ par X. On admet que $\mathcal{P}^X$ est une **famille de lois dominée** par une mesure positive σ-finie μ (définie sur $\mathcal{B}$).

On suppose alors qu'une **loi de probabilité** (inconnue) Q_θ^X engendre effectivement l'observation X : Q_θ^X peut appartenir à $\mathcal{P}^X$ ou non.

On appelle alors **quasi-vraisemblance** (R.W.M. WEDDERBURN) fondée sur X l'application $L : \mathcal{X} \times \Theta \mapsto \mathbf{R}_+$ définie par la **densité** (**dérivée de NIKODYM-RADON**) suivante :

(1) $L(. , \theta)$ ou $f(. , \theta)$ ou $f_\theta = dP_\theta^X / d\mu$.

(ii) Toute solution, si elle existe, du problème de **programmation mathématique** :

(2) $\sup_{\theta \in \Theta} L(x, \theta)$

définit, en général, une **application mesurable** $t : \mathcal{X} \mapsto \Theta$, donc un **estimateur** T = t (X) du paramètre θ.

T est appelé **estimateur du maximum de quasi-vraisemblance** (ou, parfois, **estimateur du quasi-maximum de vraisemblance**) de θ.

Le préfixe qualificatif « quasi » signifie que T n'est pas nécessairement obtenu en maximisant la vraisemblance associée à la « vraie » famille des lois $\mathcal{Q}^X$ dont l'une d'elle engendre l'observation X.

Si $\mathcal{Q}^X \subset \mathcal{P}^X$, l'estimateur T se confond avec l'**estimateur du maximum de vraisemblance** usuel (**maximum de vraisemblance contraint** si $\Theta \neq \mathbf{R}^Q$).

Comme dans la **méthode du mv**, on a souvent $\mathcal{X} = \mathcal{X}_0^N$ (puissance d'un ensemble d'observation $\mathcal{X}_0$ donné) et $\mathcal{B} = \mathcal{B}_0^{\otimes N}$ (tribu produit). De plus, X est un **échantillon iid** selon la loi considérée P_θ^ξ, définie sur $\mathcal{B}_0$, et $P_\theta^X = (P_\theta^\xi)^{\otimes N}$. Par suite :

(3) $\quad L(x, \theta) = \prod_{n=1}^N f_0(x_n, \theta),$

où $x = (x_1, ..., x_N)$, $f_0(. , \theta) = dP_\theta^\xi / d\mu_0$ et μ_0 est la mesure positive σ-finie définie sur $\mathcal{B}_0$ et tq $\mu = \mu_0^{\otimes N}$.