ESTIMATEUR EFFICACE (G4, H1)
(12 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un **estimateur efficace** est un estimateur possédant la plus petite **dispersion** (**variance** ou **matrice de covariance**) dans la classe des **estimateurs sans biais**.

(i) Soit $(\mathcal{X}, \mathcal{B}, P_\theta^X)_{\theta \in \Theta}$ un **modèle image**, $g : \Theta \mapsto \mathbf{R}^Q$ une fonction mesurable donnée (cf **application mesurable**), $t : \mathcal{X} \mapsto g(\Theta)$ un **estimateur** défini par une **statistique** T = t (X). On suppose T de carré intégrable, et l'on note $E_\theta\, t(X) = \int t(x)\, dP_\theta^X(x)$ son **espérance**, $V_\theta\, t(X) = E_\theta\, (T - g(\theta))\, (T - g(\theta))'$ sa **matrice de dispersion** et $\mathcal{C}$ la **classe des estimateurs sans biais** de g (θ), ie :

$$(1) \qquad t \in \mathcal{C} \;\Leftrightarrow\; E_\theta\, t(X) = g(\theta), \qquad \forall\, \theta \in \Theta.$$

On dit alors que T (resp t) est un **estimateur efficace** de $\tau = g(\theta)$ ssi :

$$(2) \qquad \begin{cases} t \in \mathcal{C}, \\ V_\theta\, t(X) \leq V_\theta\, s(X), \qquad \forall\, s \in \mathcal{C}, \end{cases}$$

où s définit une statistique $S : \Omega \mapsto g(\Theta)$, supposée de carré intégrable.

(ii) Si X est un **échantillon** de taille N, à valeurs dans un espace produit $\mathcal{X}$ (cf **produit d'espaces probabilisés**), et si N est fixé, il peut exister plusieurs estimateurs efficaces de g (θ).

(iii) Dans un **problème d'estimation**, un estimateur efficace est généralement (ie sous des hypothèses générales) un estimateur dont la variance est égale à la borne inférieure de l'**inégalité de CRAMER-DARMOIS-FRÉCHET-RAO**. C'est donc un estimateur sans biais de plus faible dispersion (pr aux hypothèses retenues). Ceci suppose le **modèle dominé**.

Plus précisément, on suppose que le modèle précédent est un **modèle régulier**, que T = t (X) est un **estimateur régulier** sans biais et que :

(a) la matrice J_θ associée à la **dérivée** D g (θ) est une (Q,Q)-**matrice régulière**, $\forall\, \theta \in \Theta$;

(b) $D_2 f = \mathrm{Grad}_\theta f$ est une **application continue** sur $\mathcal{X} \times \Theta$ (cf **gradient**), où $f = dP_\theta^X / d\mu$ désigne la **densité** de P_θ^X pr à une **mesure** μ définie sur $\mathcal{B}$.

Alors, pour que T soit un **estimateur efficace**, ie pour que la variance de T soit égale à la borne inférieure de l'inégalité de CRAMER-DARMOIS-FRÉCHET-RAO :

(3) $V_\theta T = J_\theta (I(\theta))^{-1} J_\theta'$, $\forall \theta \in \Theta$,

il faut et il suffit que la famille $(P_\theta^X)_{\theta \in \Theta}$ soit une **famille exponentielle**, ie dont les **dérivées de NIKODYM-RADON** (**densités** ou **fonctions de vraisemblance**) associées aux $(P_\theta^X)_{\theta \in \Theta}$ sont de la forme :

(4) $f(x, \theta) = (dP_\theta^X / d\mu)(x) = e^{< a(\theta), t(x) > + b(\theta) + c(x)}$,

dans laquelle :

(a) $a : \Theta \mapsto \mathbf{R}^Q$ et $b : \Theta \mapsto \mathbf{R}$ sont deux **fonctions numériques** ;

(b) $t : \mathscr{X} \mapsto \mathbf{R}^Q$ et $c : \mathscr{X} \mapsto \mathbf{R}$ sont deux fonctions mesurables (cf **application mesurable**).

Les propriétés des familles exponentielles impliquent que $T = t(X)$ est une **statistique exhaustive minimale** pour $\tau = g(\theta)$, et qu'elle est aussi une **statistique totale**.

Lorsque $Q = 1$, on dit que $T = t(X)$ est un **estimateur efficace** ssi T est sans biais et $V_\theta T = g'(\theta) / I(\theta)$.

(iv) La notion d'efficacité précédente est aussi appelée **efficacité au sens de R.A. FISHER** : c'est une **efficacité absolue** et **à distance finie** ($N << \infty$).

(iv) Deux extensions ont été définies : l'**efficacité relative** et l'**efficacité asymptotique** (relative).