EFFICACITÉ RELATIVE ASYMPTOTIQUE ENTRE ESTIMATEURS (G4, G11, H5)
(08 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

La notion d'**efficacité relative asymptotique entre estimateurs** est une extension de la définition de l'**efficacité relative entre estimateurs** (à distance finie) au **contexte** dans lequel le nombre d'observations augmente indéfiniment.

(i) On considère un **problème d'estimation** dans lequel l'**échantillon** $X = (X_1, ..., X_N)$ est de type discret et de taille (finie) $N \in \mathbf{N}^*$.

On appelle **efficacité (absolue) d'un estimateur** (ponctuel) $T_N = t_N(X)$ le rapport de la borne inférieure de l'**inégalité de CRAMER-DARMOIS-FRÉCHET-RAO** et de sa **variance** propre, ie :

(1) e_N ou $e_N(T) = \{N . I(\theta)\}^{-1} . V T_N \geq 1$,

où $I(\theta)$ est l'**information de FISHER** (scalaire) et $X = (X_1, ..., X_N)$ un **échantillon iid** qui suit la **loi** $P_\theta^X = (P_\theta^\xi)^{\otimes N}$, avec $\theta \in \Theta$ et $\Theta \subset \mathbf{R}$.

Si $e_N = 1$, ie si l'estimateur T_N de θ atteint sa borne inférieure, on dit que T_N est un **estimateur (absolument) efficace** de θ. On parle aussi d'**efficacité absolue**.

Si le **modèle d'échantillonnage** de taille N est plongé dans un modèle d'échantillonnage asymptotique, ainsi que la suite $T = (T_N)_{N \in \mathbf{N}^*}$ des estimateurs de θ correspondante, on appelle **efficacité (absolue) asymptotique** de T la limite (si elle existe) :

(2) $e_\infty = \lim_{N \in \mathbf{N}} e_N$.

(ii) Etant donné deux suites $S = (S_N)_{N \in \mathbf{N}^*}$ et $T = (T_N)_{N \in \mathbf{N}^*}$ d'estimateurs de θ, on appelle :

(a) **efficacité relative (à distance finie)** de T pr à S le rapport de leurs variances, ie :

(3) $e_N(T / S) = V_\theta T_N / V_\theta S_N$.

Elle se déduit donc du rapport de leurs efficacités relatives respectives $e_N(S)$ et $e_N(T)$, puisque la borne inférieure $N . I(\theta)$ leur est commune ;

(b) **efficacité relative asymptotique** de T pr à S la limite (si elle existe) du rapport précédent, ie :

(4) $e_\infty(T / S) = \lim_{N \to +\infty} e_N(T / S)$.

Si S et T sont des **estimateur sans biais** (ie si $E_\theta T_N = E_\theta S_N = \theta$ à distance finie, ou si, asymptotiquement, $\lim_N E_\theta T_N = \lim_N E_\theta S_N = \theta$, $\forall$ N et $\forall$ θ), alors $e_N(T / S)$

s'interprète comme le carré du rapport entre les **coefficients de variation** de T_N et de S_N. Si T_N est un **estimateur plus « précis »** que S_N, alors $e_N (T / S) < 1$.

En particulier, l'**efficacité relative asymptotique** d'une suite T pr à une suite S, toutes deux supposées constituées d'**estimateurs efficaces** de θ, est le rapport (s'il est fini) entre les tailles d'échantillon tq :

(5) $e_\infty (T / S) = \lim_{\min (N',N'') \to +\infty} \{N' / N'' : V_\theta T_{N''} = V_\theta S_{N'}\}$.

Par exemple, s'il existe $p \in$ **N*** et $q \in$ **N***, ainsi que $(a, b) \in \mathbf{R}^2$ tq :

(6)
$$V_\theta S_{N'} \sim a . (N')^{-p}$$
$$V_\theta T_{N''} \sim b . (N'')^{-q}$$

lorsque $\min (N', N'') \to +\infty$, alors l'égalité $V_\theta T_{N''} = V_\theta S_{N'}$ implique :

(7) $a / b = \lim_{\min (N',N'') \to +\infty} (N')^p / (N'')^q$, avec $p \geq q$.

Par suite :

(a) si $p > q$, la limite de N' / N'' est $+\infty$. On dit alors que T est d'**efficacité asymptotique nulle** ;

(b) si $p = q$, alors :

(8) $\lim_{\min (N',N'') \to +\infty} N' / N'' = (a / b)^{1/q} = \lim_{\min (N',N'') \to +\infty} (V_\theta T_{N''} / V S_{N'})^{1/q}$,

et l'efficacité relative asymptotique de T pr à S s'écrit :

(9) $e_\infty (T / S) = \lim_{\min (N',N'') \to +\infty} (V_\theta T_{N''} / V_\theta S_{N'})^{-1/q}$.

(iii) On adopte parfois une définition (équivalente) inverse de (4). Si S et T sont deux suites d'estimateurs sans biais (ou convergents en probabilité) de θ, on appelle alors **efficacité (relative) asymptotique** de T pr à S la limite suivante (si elle existe) :

(10) $e_\infty^2 = \lim_N (V_\theta S_N / V_\theta T_N)$.

Dans ce cas, T est dite plus efficace que S si $e_\infty^2 > 1$.

(iv) Les définitions précédentes s'étendent au cas où l'on estime une fonction scalaire $g (\theta)$ définie par une fonction mesurable donnée $g : \Theta \mapsto \mathbf{R}$. Les suites tq S ou T concernent alors des estimateurs de $g (\theta)$.

Elles s'étendent aussi au cas de paramètres vectoriels $\theta \in \Theta$ (avec $\Theta \subset \mathbf{R}^Q$) (cf **efficacité relative entre estimateurs**).

(v) Le choix de la **variance** (ou parfois de l'**écart quadratique moyen** lorsque les estimateurs considérés sont biaisés) comme critère de **dispersion** d'un estimateur, donc comme critère de son **efficacité**, tient surtout au fait que de nombreux estimateurs sont (asymptotiquement) normaux (cf **normalité asymptotique**).