ESTIMATEUR SANS BIAIS (H1)
(17 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un **estimateur** T d'un paramètre θ (resp d'une fonction numérique g de ce paramètre) est appelé **estimateur sans biais** (au sens de l'espérance) ssi il est intégrable et si son **espérance mathématique** est égale au **paramètre** (resp à la valeur $\tau = g(\theta)$ d'une fonction d'intérêt g de θ), ie ssi son **biais** est nul (cf **estimateur ponctuel**) :

(0) $E_\theta T = \theta$ (resp $E_\theta T = g(\theta)$), $\forall \theta \in \Theta$.

(i) Dans un **problème d'estimation** « directe » $\{(\mathcal{X}, \mathcal{B}, P_\theta^X)_{\theta \in \Theta}, \Theta, L\}$, un estimateur $t : \mathcal{X} \mapsto \Theta$ du paramètre $\theta \in \Theta$ (avec $\Theta \subset \mathbf{R}^K$) est appelé **estimateur sans biais (ponctuel)** de θ ssi :

(1) $E_\theta T = \theta, \quad \forall \theta \in \Theta,$

où T = t (X) est la **statistique** définie par t et $E_\theta T = \int_\Theta t(x)\, dP_\theta^X(x)$.

On dit aussi que T est un **estimateur « absolument » sans biais**, ou un **estimateur « uniformément » sans biais**, puisque l'égalité précédente doit être vérifiée quelle que soit la « vraie » valeur θ du paramètre.

De façon analogue, dans le problème d'estimation « dérivé » $\{(\mathcal{X}, \mathcal{B}, P_\theta^X)_{\theta \in \Theta}, \mathbf{R}^Q, L\}$, un estimateur $t : X \mapsto g(\Theta)$ du paramètre dérivé $\tau = g(\theta)$ est appelé **estimateur sans biais (ponctuel)** de τ ssi :

(2) $E_\theta T = g(\theta) = \tau, \quad \forall \theta \in g(\Theta),$

où $g : \Theta \mapsto \mathbf{R}^Q$ est une fonction mesurable donnée et $g(\Theta) \in \mathcal{B}(\mathbf{R}^Q)$.

(ii) On dit souvent que θ (resp $\tau = g(\theta)$) est un **paramètre identifiable** (identification au premier ordre, ie au sens de l'espérance) lorsqu'il possède un estimateur sans biais (cf **identifiabilité**, **valeur identifiable d'un paramètre**).

(iii) Les définitions précédentes d'un estimateur sans biais correspondent à une notion d'**estimateur strict**.

Les définitions s'étendent directement au cas d'un **estimateur large**, ou **estimateur étendu**, en remplaçant Θ par $\mathbf{R}^K$ dans (1) et $g(\Theta)$ par $\mathbf{R}^Q$ dans (2) (cf **biais**). Si $V_\theta T = E_\theta (T - E_\theta T)(T - E_\theta T)'$ (**matrice de dispersion** de T), on peut définir une relation de préférence entre estimateurs sans biais de θ (resp de $g(\theta)$) en posant (cf **préordre**, **relation d'ordre**) :

(3) $T \succ S \Leftrightarrow V_\theta S \geq V_\theta T, \quad \forall \theta \in \Theta$

(ou $\forall \theta \in \mathbf{R}^K$ dans le cas large), l'inégalité étant au sens des **formes quadratiques**.

(iv) On peut « améliorer » un estimateur sans biais à l'aide d'une **statistique exhaustive** (cf **théorème de BLACKWELL-RAO**, **théorème de LEHMANN-SCHEFFÉ**).

(v) Si l'on considère un **espace d'échantillonnage** avec $\mathscr{X} = \mathbf{R}^N$, on dit que $T_N = t_N (X) = t_N (X_1 ,..., X_N)$ est un **estimateur asymptotiquement sans biais** du paramètre θ (resp de $g(\theta)$) ssi (cf **biais asymptotique**) :

(4) $E_\theta T_N \to_{N \to +\infty} \theta$ (resp $g(\theta)$), $\forall \theta \in \Theta$.

De façon équivalente, T_N est un estimateur asymptotiquement sans biais ssi :

(5) $B_\theta T_N = E_\theta (T_N - \theta) \to_{N \to +\infty} 0$, $\forall \theta \in \Theta$

resp :

(6) $B_\theta T_N = E_\theta (T_N - g(\theta)) \to_{N \to +\infty} 0$, $\forall \theta \in g(\Theta)$.

(v) Dans un problème d'**estimation non paramétrique**, basé sur un modèle de la forme non explicitement paramétrée $(\mathscr{X}, \mathscr{B}, \mathscr{P}^X)$, et si $(\Gamma, \mathscr{B}_\Gamma)$ est un **espace mesurable** de **caractéristiques** utiles associé à la **famille** $\mathscr{P}^X$, on définit, de façon analogue, la notion d'**estimateur sans biais d'une caractéristique** $\gamma \in \Gamma$. Ainsi :

(7) $E_P T = \gamma$, $\forall P \in \mathscr{P}$,

en notant $E_P T = E_P (t \circ X) = \int (t \circ X) dP = \int t(x) dP^X (x)$, où $\mathscr{P}$ est la **famille** des **probabilités** (cf aussi **mesures de probabilité**) P qui admettent les **lp** $P^X \in \mathscr{P}^X$ pour images par X.

(vi) Les estimateurs précédents ont été supposés intégrables. De plus, les définitions « standards » sont basées sur l'**espérance mathématique** pour définir le **biais** d'un estimateur.

Cependant, si l'on appelle **écart** (ou **centrage**) de T pr à θ (resp pr à $g(\theta)$) la « **statistique »** $D = T - \theta$ (resp $D = T - g(\theta)$), on peut aussi bien appeler :

(a) **biais au sens du mode** le mode S_D de D, ou de sa lp ;

(b) **biais au sens de la médiane** (ou, plus généralement, **biais au sens d'un quantile** d'ordre p) la médiane $Q_{1/2} D$ (ou, plus généralement, le p-quantile $Q_p D$) de D, ou de sa **lp** ;

(c) Si P^D est la loi de D, on peut, plus généralement, appeler **biais au sens d'une caractéristique de centralité** $\gamma = c(P^D)$ de P^D la **caractéristique de centralité** de cette loi (cf aussi **paramètre de position**).

Ces extensions peuvent être utiles lorsque les estimateurs ne sont pas eg intégrables.

En pratique, l'espérance mathématique présente souvent des avantages de calcul et possède des propriétés mathématiques qui expliquent la fréquence de son usage ainsi que la terminologie courante (par défaut, ou implicite).

(vii) Un estimateur dont le biais (resp le biais asymptotique) n'est pas nul est appelé **estimateur biaisé** (resp **estimateur asymptotiquement biaisé**) du paramètre considéré. Il est souvent possible de « réduire » ce biais (cf **méthode de QUENOUILLE**), ou encore de déduire un estimateur sans biais à partir d'un estimateur biaisé.

(viii) Ce dernier cas est illustré eg par le problème d'estimation de la **variance** d'une **population** (variance « théorique »).

Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé** et $\xi : \Omega \mapsto \mathbf{R}$ une **vars** de carré intégrable. On cherche à estimer $V\ \xi = \sigma^2$ (variance théorique de ξ) au vu d'un N-**échantillon iid** distribué comme la **variable parente** ξ.

La **variance empirique** $S_N^2 = N^{-1}\ \Sigma_{n=1}^N (X_n - \overline{X}_N)^2$ est un estimateur (naturel) biaisé de σ^2 : $E\ S_N^2 = N^{-1} . (N - 1)\ \sigma^2 \neq \sigma^2$ (cf **statistique naturelle**). Par suite, $S_N'^{\ 2} = (N / (N - 1))\ S_N^2$ est un estimateur sans biais de σ^2.

Cette méthode vaut, plus généralement, lorsque :

(8) $\quad E_\theta\ T = \alpha + \beta . g(\theta)$, $\quad$ ou $\quad T = \alpha + \beta . \tau$ $\quad$ (où α et β sont connus),

ce qui correspond à un estimateur T biaisé de $\tau = g(\theta)$. En effet, si $\beta \neq 0$, l'estimateur $\beta^{-1} . (T - \alpha)$ est sans biais. Elle n'est pas adaptée au cas d'une transformation non linéaire (cf eg **inégalité de JENSEN**).

(ix) Enfin, on appelle **estimateur conditionnellement sans biais** un estimateur dont l'**espérance conditionnelle** (pr à des **événements** ou pr à des **va** donné(e)s) est égale au paramètre (ou à la fonction du paramètre) considéré(e), eg (avec des notations évidentes) :

(9) $\quad E_\theta (T / S_1 ,..., S_L) = g(\theta)$ $\quad$ (P_θ -p.s.), $\forall\ \theta \in \Theta$.

Cette notion intervient, notamment, en **théorie des processus** (eg **prévision conditionnelle**) ou en théorie de la **régression**.