ESTIMATEUR DE PITMAN (G1, H)
(21 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

L'estimateur de PITMAN est un **estimateur** équivariant (cf **équivariance**) qui minimise l'écart quadratique moyen dans un **modèle statistique** dont les **lois** comportent un paramètre de position ou un paramètre d'échelle.

Soit $\{(\mathscr{X}, \mathscr{B}, (P_\theta^X)_{\theta \in \Theta}, D, L\}$ un **problème de décision** à **espace de décision** D et **fonction de perte** L. On suppose que $\mathscr{X} = \mathbf{R}^N$ et que $X : \Omega \mapsto \mathbf{R}^N$ est un **vecteur aléatoire** réel à coordonnées X_n iid comme $\xi \sim P_\theta^\xi$ (cf **échantillon iid**). La **loi de probabilité** de X est donc de la forme $P_\theta^X = \otimes_{n=1}^N P_\theta^{X(n)} = (P_\theta^\xi)^{\otimes N}$ (en notant X(n) pour désigner X_n). Enfin, on note $X^{(.)} = (X^{(1)}, ..., X^{(N)})$ la **statistique d'ordre** croissant associée à X (avec $X^{(n)} \leq X^{(n+1)}$, $\forall\, n \in \mathbf{N}_{N-1}{}^*$).

(i) Dans le cas d'un **paramètre de position** $\theta \in \Theta$ (avec $\Theta = \mathbf{R}$), il existe une application $a : \mathscr{X} \times \Theta \mapsto \mathbf{R}^N$ définie par :

(1) $\quad (x, \theta) \mapsto a(x, \theta) = x - \theta \,.\, e_N$,

où $e_N = (1, ..., 1)'$ (premier vecteur bissecteur de $\mathbf{R}^N$), et une **lp** P_0^X sur $\mathscr{B}$ tq :

(2) $\quad P_\theta^X = P_0^{a(X,\theta)} = P_0^{X - \theta\, e(N)}, \qquad \forall\, \theta \in \Theta,$

en notant e(N) pour désigner e_N . Ainsi, P_0^X est de la forme $(P_0^\xi)^{\otimes N}$.

On suppose que la fonction de perte $L : \Theta \times D \mapsto \mathbf{R}_+$ est quadratique en $x - \theta$, ie est définie par l'**écart quadratique moyen** :

(3) $\quad L(\theta, d) = l(d - \theta) = (d - \theta)^2.$

La **fonction de risque** associée s'écrit :

(4) $\quad R(\theta, \delta) = E_\theta\, l(\delta(X) - \theta) = E_0\, l(\delta(X)).$

On montre que cette fonction est minimum lorsqu'on adopte la **règle de décision pure** suivante, appelée **estimateur de E.J.G. PITMAN** de θ :

(5) $\quad \delta_N{}^*(X) = X_1 - E_0(X_1 / Y),$

où $Y = (Y_2, ..., Y_N)$, avec $Y_n = X_n - X_1$, $\forall\, n \in \{2, ..., N\}$, et où $E_0(. / Y)$ est l'**espérance conditionnelle** calculée avec la loi P_0^ξ (ie avec $\theta = 0$).

L'estimateur (5) s'explicite selon :

(6) $\quad \delta_N{}^*(X) = \int \theta\, f_0(a(X, \theta))\, d\theta \,/ \int f_0(a(X, \theta))\, d\theta,$

où $f_0 = dP_0^\xi / d\lambda_0$ et λ_0 est la **mesure de LEBESGUE** sur $\mathcal{B}_0 = \mathcal{B}_{\mathbf{R}}$.

Cet estimateur est invariant par le **groupe des translations** dans $\mathbf{R}^N$.

(ii) Dans le cas d'un **paramètre d'échelle et de position** $\theta = (\alpha, \beta)$ (avec $\Theta = \mathbf{R} \times \mathbf{R}_+^*$), il existe une application b : $\mathcal{X} \times \Theta \mapsto \mathbf{R}^N$ définie par :

(7) $\quad (x, \theta) \mapsto b(x, \theta) = (x - \alpha \,.\, e_N) / \beta$

et une **loi de probabilité** $\mathcal{L}_{01}(.)$ sur $\mathcal{B}$ tq :

(8) $\quad P_\theta^X = \mathcal{L}_{01}(b(X, \theta)) = \mathcal{L}_{01}\{(X - \alpha \,.\, e_N) / \beta\}, \qquad \forall\, \theta \in \Theta.$

$\mathcal{L}_{01}(.)$ est donc de la forme $(\mathcal{L}_{01}(\xi))^{\otimes N}$).

Si la fonction de perte L : $\Theta \times D \mapsto \mathbf{R}_+$ est définie par l'écart quadratique moyen pr à $\theta = (\alpha, \beta)$, on montre que le risque associé à L est minimisé avec la règle de décision pure (vectorielle) suivante :

$$(9)\qquad \begin{aligned} \delta_{\alpha,N}^*(X) &= X^{(1)} - (X^{(2)} - X^{(1)})\,(A_N / B_N),\\ \delta_{\beta,N}^*(X) &= (X^{(2)} - X^{(1)})\,(B_N / C_N), \end{aligned}$$

dans laquelle :

$$(10)\qquad \begin{aligned} A_N &= \int \alpha\, \beta^{N-1} f_{01}(\alpha)\, f(\alpha+\beta) \prod_{n=3}^{N} f_{01}(\alpha + \beta Z_n)\, d\alpha\, d\beta,\\ B_N &= \int \beta^{N} f(\alpha)\, f_{01}(\alpha+\beta) \prod_{n=3}^{N} f_{01}(\alpha + \beta Z_n)\, d\alpha\, d\beta,\\ C_N &= \int \beta^{N+1} f_{01}(\alpha)\, f_{01}(\alpha+\beta) \prod_{n=3}^{N} f_{01}(\alpha + \beta Z_n)\, d\alpha\, d\beta. \end{aligned}$$

Dans (10), on note $Z = (Z_3, ..., Z_N)$, avec $Z_n = (X^{(n)} - X^{(1)}) / (X^{(2)} - X^{(1)})$, $\forall\, n \in \{3, ..., N\}$, $X^{(2)} - X^{(1)} > 0$, et f_{01} désigne la densité de $\mathcal{L}_{01}(\xi)$ pr à λ_0 (mesure de LEBESGUE sur $\mathcal{B}_0 = \mathcal{B}_{\mathbf{R}}$).

L'**estimateur de E.J.G. PITMAN** (9) du couple $\theta = (\alpha, \beta)$ s'explicite alors selon les formules conditionnelles :

$$(11)\qquad \begin{aligned} \delta_{\alpha,N}^*(X) &= X^{(1)} - (X^{(2)} - X^{(1)})\,\{E_{01}(X^{(1)}(X^{(2)} - X^{(1)}) / Z) / E_{01}((X^{(2)} - X^{(1)})^2 / Z)\},\\ \delta_{\beta,N}^*(X) &= (X^{(2)} - X^{(1)})\,\{E_{01}(U / Z) / E_{01}(U^2 / Z)\}, \end{aligned}$$

dans lesquelles E_{01} (. / Z) désigne l'espérance conditionnelle calculée avec la loi $\mathscr{L}_{01}$ (ξ) (ie pour la valeur (α, β) = (0,1)) et U est la va dont la loi $\mathscr{L}_{01}$ (υ) admet f_{01} pour densité pr à λ_0 (mesure de LEBESGUE sur $\mathscr{B}_0 = \mathscr{B}_{\mathbf{R}}$).

Cet estimateur est invariant par le **groupe de transformations** affines de $\mathbf{R}^N$ défini par $x \in \mathbf{R}^N \mapsto \alpha x + \beta$, $\forall (\alpha, \beta) \in \Theta$.

(iii) On peut définir, de façon analogue, un estimateur de PITMAN pour un paramètre de position α vectoriel. Si $\alpha \in \mathbf{R}^K$ (ie ξ est à valeurs dans $\mathbf{R}^K$), on montre que ce dernier estimateur est inadmissible lorsque $K \geq 3$ et admissible lorsque $K \leq 2$ pour une fonction de perte quadratique.