ESTIMATEUR DE JAMES-STEIN (H1)
(22 / 11 / 2019, © Monfort, Dicostat2005, 2005-2019)

Pour estimer un **paramètre de position**, un **estimateur** optimal n'est pas nécessairement un **estimateur admissible** : il peut exister un meilleur estimateur (au sens de la fonction de risque).

(i) Soit $\{(\mathcal{X}, \mathcal{B}, P_\theta^X)_{\theta \in \Theta}, D, L\}$ un **problème de décision** statistique tq (**problème d'estimation**) :

(a) $D = \Theta = \mathbf{R}^Q$;

(b) $X = (X_1, ..., X_N)$ est un **échantillon iid** comme la **variable parente** $\xi : \Omega \mapsto \mathcal{X}$, où $\mathcal{X} = \mathbf{R}^Q$ et tq $\xi \sim \mathcal{N}_Q(0, I_Q)$ (**loi gaussienne**) ;

(c) la **fonction de perte** est quadratique (scalaire), ie (cf **perte quadratique**, **écart quadratique moyen** généralisé) :

(1) $L(\theta, d) = ||\theta - d||^2 = \Sigma_{q=1}^Q (\theta_q - d_q)^2$.

On montre alors que :

(a) l'**estimateur du mv** $\overline{X}_N = N^{-1} \Sigma_{n=1}^N X_n$ (**moyenne empirique**) de θ est un estimateur « minimax » (cf **règle minimax**) mais ce n'est pas un **estimateur admissible** (pour le **risque** R associé à L) dès que $Q \geq 3$ (cf **fonction de risque**) ;

(b) l'**estimateur de W. JAMES - C. STEIN** de θ, défini selon :

(2) $T_N = \{1 - || \overline{X}_N ||^{-2} (Q - 2)\} \overline{X}_N$,

est meilleur que $\overline{X}_N$ (au sens de R).

(ii) Plus généralement, si $V \xi = \Sigma > 0$ est connue et tq $\Sigma \neq I_Q$, on se ramène au cas précédent en étudiant l'échantillon Y associé à la va $\eta = \Sigma^{-1/2} \xi$. L'estimateur de JAMES - STEIN s'écrit alors :

(3) $T_N' = \{1 - (\overline{X}_N' \Sigma^{-1} \overline{X}_N)^{-1} (Q - 2)\} . \overline{X}_N$.

C'est un **estimateur invariant** pour le groupe $\mathcal{G}$ des **applications linéaires** régulières de $\mathbf{R}^Q$ dans $\mathbf{R}^Q$ (cf aussi **matrice régulière**).