MÉTHODE DES MOMENTS (H3)
(05 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

Avec la **méthode du maximum de vraisemblance** ou les **méthodes à distance minimale** (eg **méthode de moindre norme**, **méthode de moindres carrés**), la **méthode des moments** est une méthode générale d'estimation des paramètres d'une **loi de probabilité** (R.A. FISHER - K. PEARSON) (cf aussi **méthode des caractéristiques**).

(i) Soit $(\Omega, \mathcal{T}, P_\theta)_{\theta \in \Theta}$ un **modèle paramétrique** de base et $(\mathcal{X}, \mathcal{B}, P_\theta^X)_{\theta \in \Theta}$ son **modèle image** par une **statistique** (**échantillon**) X.

On suppose que :

(a) $\Theta \subset \mathbf{R}^Q$ et $\Theta \neq \emptyset$;

(b) le modèle est un **modèle d'échantillonnage** $\{\mathcal{X}_0, \mathcal{B}_0, (P_\theta^\xi)_{\theta \in \Theta}\}^{\otimes N}$, avec $\mathcal{X}_0 = \mathbf{R}$, $\xi : \Omega \mapsto \mathcal{X}_0$ est une **variable parente** de **loi** P_θ^ξ, et $X = (X_1, ..., X_N)$ est un **échantillon iid** selon P_θ^ξ ;

(c) les **moments algébriques** (théoriques) de ξ existent jusqu'à l'ordre p, ie il existe un entier $p \in \mathbf{N}^*$ tq $\xi \in L_{\mathbf{R}}^p (\Omega, \mathcal{T}, P_\theta)$, $\forall \theta \in \Theta$.

La **méthode des moments** consiste à estimer θ en identifiant resp les moments de la loi théorique P_θ^ξ aux moments de mêmes ordres de la **loi empirique** P_N associée à X.

En considérant des **moments centrés** (pr à l'**espérance**), on note resp :

(1) $\mu_j (\theta) = E_\theta (\xi - E_\theta \xi)^j = \int (\xi - E_\theta \xi)^j \, dP_\theta$, $\forall j \in N_p^*$,

les **moments** théoriques (**moments algébriques**) et :

(2) $m_j (N) = E_N (\xi - E_N \xi)^j = \int (\xi - E_N \xi)^j \, dP_N = N^{-1} \Sigma_{n=1}^N (X_n - \bar{X}_N)^j$, $\forall j \in N_p^*$,

les **moments empiriques** (cf **statistique naturelle**).

La méthode consiste alors à estimer le **paramètre** θ de la loi P_θ^ξ à l'aide de la solution (supposée exister et être unique) du système d'équations (en général non linéaire) :

(3) $\mu_j (\theta) = m_j (N)$, $\forall j \in N_p^*$.

En notant $\mu (\theta) = (\mu_1 (\theta), ..., \mu_p (\theta))'$ le vecteur des moments théoriques et $m (N) = (m_1 (N), ..., m_p (N))'$ celui des moments empiriques correspondants, on obtient le système vectoriel :

(4) $\mu(\theta) = m(N)$.

La fonction vectorielle $\mu : \Theta \mapsto \mathbf{R}^p$ est supposée régulière, ie vérifie des **conditions de régularité** permettant de calculer les moments théoriques en fonction de θ, ie de calculer l'**application inverse** μ^{-1} de μ (cf aussi **inverse**).

Le système (4) définit une **équation estimante** appelée **équation des moments**, ou **équation aux moments**.

La méthode suppose qu'il existe un nombre p d'équations au moins égal au nombre de paramètres θ_j à estimer (où $j \in N_p$*). Les conditions de régularité nécessitent notamment que rg $(D\,\mu(\theta)) = p$ afin d'assurer l'unicité locale de la solution, ie au **voisinage** de la **vraie valeur** $\theta^* \in \Theta$.

On note :

(5) $\theta_{m,N}$ ou $\theta_m^{\sim} = \mu^{-1}(m(N))$

la solution ainsi obtenue, appelée **estimateur des moments**, ou **estimateur par la méthode des moments**.

(ii) Sous certaines conditions, il existe ainsi une fonction mesurable $t_N : \mathbf{R}^N \mapsto \mathbf{R}^p$ définissant une **statistique** $T_N = t_N(X)$ tq $\theta_{m,N} = T_N$ (cf **application mesurable**).

Par suite :

(a) l'**estimateur** par les moments est un estimateur sans **biais** :

(6) $E_\theta\, \theta_{m,N} = \theta, \qquad \forall\, \theta \in \Theta$;

(b) si l'équation (4) admet une solution unique (celle donnée par (5)), et si μ est continue, on établit la **convergence presque sûre** :

(7) $\theta_{m,N} \to \theta$ (P_θ-p.s.), $\qquad \forall\, \theta \in \Theta$;

(c) sous la même condition d'unicité de $\theta_{m,N}$ ($\forall$ N $\in$ **N***) et si μ est différentiable (cf **différentiabilité**), on établit la **convergence en loi** gaussienne :

(8) $N^{1/2}(\theta_{m,N} - \theta) \to^{\mathscr{L}}{}_{N \to +\infty}\ \mathcal{N}_p(0, \Omega_\theta), \qquad \forall\, \theta \in \Theta,$

avec $\Omega_\theta = M_\theta A_\theta M_\theta'$, où $M_\theta \in M_p(\mathbf{R})$ est la **matrice** de terme général $(\partial \mu_j^{-1} / \partial \theta_k)(\theta)$, ie la **matrice jacobienne** $(D\mu^{-1} / D\theta)(\theta)$ de μ^{-1}, et où $A_\theta \in M_p(\mathbf{R})$ est une matrice qui dépend des moments théoriques ;

(d) on établit que :

(9) $\theta_{m,N} = \theta + O(N^{-1/2})$ (biais d'ordre $N^{-1/2}$) ;

(e) enfin, l'estimateur des moments n'est pas un **estimateur strict** de θ.

(iii) La méthode des moments est utilisée notamment :

(a) lorsque l'**estimateur du mv** n'est pas défini (ie n'existe pas) ;

(b) ou lorsque le calcul analytique de ce dernier est complexe ;

(c) ou encore lorsque les P_θ^ξ sont des lois sans **densité** (vraisemblances non définies)

Elle sert eg à estimer diverses **familles** de lois paramétriques (eg famille des **lois de PEARSON**).

Les **propriétés asymptotiques** de l'estimateur des moments précédent (convergences presque sûre et en loi) justifient cette méthode, compte tenu du fait que les moments empiriques tendent presque sûrement (ainsi qu'en loi) vers les moments théoriques correspondants (cf **fonction de répartition empirique**, **problème des moments** de T.J. STIELTJES).

L'**efficacité** (au sens de l'**inégalité de CRAMER-DARMOIS-FRÉCHET-RAO**) de la méthode est cependant souvent moindre que celle de la **méthode du mv** : ceci est le cas eg pour les lois de PEARSON de type III. A taille N d'échantillon fixée, la méthode est d'autant moins précise que l'ordre p de l'équation aux moments est plus élevé (cf **degré de liberté**).

(vii) A titre d'exemple, on considère l'estimation du paramètre (μ, σ^2) d'une **loi Log-normale** (généralisée) $\mathscr{L}(\mu, \sigma^2, y_0)$ à l'aide d'un N-**échantillon** $Z = (Z_1 ,..., Z_N)$.

(a) si le paramètre scalaire (borne inférieure du **support**) y_0 est connu, on peut centrer Z pr à y_0 , ie utiliser l'échantillon $X = (X_1 ,..., X_N)$, avec $X_n = Z_n - y_0$.

En notant $\bar{X}_N$ (**moyenne empirique**) et S_N^2 (**variance empirique**) les deux premiers moments de X, la méthode « identifie » les moments théoriques et empiriques, d'où le système (non linéaire) :

$$(12)\quad \begin{cases} \bar{X}_N = \exp\{\mu + (1/2)\,\sigma^2\}, \\ S_N^2 = \exp\{2\,(\mu + \sigma^2)\} - \exp\{2\,\mu + \sigma^2\}, \end{cases}$$

dont la solution en (μ, σ^2) s'écrit :

$$(13)\quad \begin{cases} \mu_{m,N} = \mathrm{Log}\,\{(S_N^2 + \bar{X}_N^2)^{-1/2}\ \bar{X}_N^2\}, \\ \sigma_{m,N}^2 = \mathrm{Log}\,\{1 + (S_N^2 / \bar{X}_N^2)\}\ ; \end{cases}$$

(b) si y_0 n'est pas connu, on peut l'estimer par $y_0^\sim = \min_n X_n = X^{(1)}$ (cf **statistique d'ordre**) et procéder comme ci-dessus.

(iv) La méthode des moments s'étend directement à un échantillon vectoriel, ie au cas où ξ est un **vecteur aléatoire** à valeurs dans $\mathbf{R}^K$ (au lieu de $\mathbf{R}$) : on utilise alors les **moments croisés** (théoriques et empiriques) entre les coordonnées de ξ.

Ainsi, les **moments croisés** (théoriques) d'ordre (α, β), avec $\alpha + \beta = j$, s'écrivent :

(10) $\mu_{\alpha\beta}(\theta) = E_\theta (\xi - E_\theta \xi)^\alpha (\xi - E_\theta \xi)^\beta = \int (\xi - E_\theta \xi)^\alpha (\xi - E_\theta \xi)^\beta \, dP_\theta , \quad \forall j \in N_p^*.$

(v) La **méthode des moments généralisés** constitue une extension de la méthode des moments.

La méthode des moments peut aussi s'étendre à d'autres types de moments (cf eg **moment absolu**, **moment équilibré**, **moment exponentiel**, **moment incomplet**). Ainsi, elle a été étendue aux **moments factoriels** : eg estimation d'un mélange (fini) de **lois discrètes** de même type (G. CAMMILLERI) (cf **mélange de lois**).

(vi) La **méthode des quantiles** est une méthode voisine dans laquelle on suppose données p valeurs distinctes $\alpha_j \in]0, 1[$, $\forall j \in N_p^*$.

La méthode consiste à résoudre l'équation estimante suivante, dite **équation aux quantiles** :

(11) $q_{N,\alpha(j)}(X) = F_N^{-1}(\alpha_j), \quad \forall j \in N_p^*,$

en notant $q_{N,\alpha(j)}(X)$ le **quantile empirique** d'ordre α_j , F_N la **fr empirique** associée à X et F_N^{-1} son inverse (ie la **fonction quantile** associée).