TEST DE WALD (I4)
(10 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

Le **test de WALD** recouvre des domaines variés de **Statistique**. Il suppose que le modèle considéré est un **modèle dominé**.

(i) Soit $(\mathscr{X}, \mathscr{B}, (P_\theta^X)_{\theta \in \Theta})$ un **modèle paramétrique** (**modèle image**) dans lequel $\mathscr{X} = \mathbf{R}^N$, $X = (X_1, ..., X_N)$ est l'**échantillon** observé et $\Theta \subset \mathbf{R}^Q$. On suppose que $\mathscr{P}^X = (P_\theta^X)_{\theta \in \Theta}$ est une **famille de lois dominée** uniformément par une **mesure positive** σ-finie μ (cf **mesure σ-finie**), et l'on note :

(1) $\quad f(. , \theta) = dP_\theta^X / d\mu$

la **dérivée de NIKODYM-RADON** (ie la **vraisemblance**) associée à toute loi P_θ^X (où $\theta \in \Theta$).

Etant donné une fonction mesurable $h : \Theta \mapsto \mathbf{R}^K$ (avec $K < Q$) (cf **application mesurable**), on définit l'ensemble $\Theta_0 \subset \Theta$ suivant :

(2)
$$\Theta_0 = \{\theta \in \Theta : h(\theta) = 0,$$
$$(\text{resp } \Theta_0 = \{\theta \in \Theta : h(\theta) = h(\theta_0)\}),$$

où $\theta_0 \in \Theta_0$ est donné, et l'on suppose que $\Theta_0 \neq \varnothing$ et $\Theta_0 \neq \Theta$.

On considère alors un **problème de test** portant sur l'**hypothèse de base** :

(3) $\quad H_0 : \theta \in \Theta_0$.

Le **test de A. WALD** est le test de H_0 fondé sur la **statistique de test** suivante, appelée **statistique de A. WALD** :

(4) $\quad W_N = h(\theta^\sim)' \{(D\, h(\theta^\sim)) (I(\theta^\sim))^{-1} (D\, h(\theta^\sim)'\}^{-1} h(\theta^\sim),$

dans laquelle :

(a) $\theta^\sim$ est l'estimateur du maximum de vraisemblance de θ, non contraint par h (ie $\theta^\sim \in \Theta$) ;

(b) D h est la (K, L)-**matrice** des **dérivées partielles** (**hessienne**) $\partial h_k / \partial \theta_k$ (k = 1 ,..., K ; q = 1,..., Q) ;

(c) I (θ) est la **matrice d'information** de FISHER (évaluée au point $\theta^\sim$).

Le test compare alors $h(\theta^\sim)$ à $0 \in \mathbf{R}^K$.

(ii) sous l'hypothèse H_0 , on établit la **propriété asymptotique** suivante (**convergence en loi**) :

(5) $\mathscr{L}(W_N) \to_{N \to +\infty} \mathscr{X}_K^2$ (**loi du chi-deux** à K **degrés de liberté**).

On peut ainsi tester H_0 contre l'**hypothèse alternative** $H_1 = H_0^c$ (ie $H_1 : \theta \in \Theta_1 = \Theta_0^c$), puisque (5) permet de définir une **région critique** asymptotique associée à tout seuil $\alpha \in]0,1[$. Cette région est de la forme :

(6) $w = [W_N > q_{1-\alpha}]$,

où $q_{1-\alpha}$ est le **quantile** d'ordre $1 - \alpha$ de la loi $\mathscr{X}_K^2$.

(iii) Le test de WALD ne nécessite pas, comme le **test du rapport des vraisemblances**, le calcul de l'**estimateur du mv** contraint par l'hypothèse H_0 (ie le calcul de θ^* tq $\theta^* \in \Theta_0$).

Les hypothèse de régularité habituellement faites sont les suivantes :

(a) $\Theta \in \mathscr{O}(\mathbf{R}^Q)$ (**ouvert** de $\mathbf{R}^Q$) ;

(b) le modèle considéré est un **modèle homogène**, ie $f > 0$ sur $\mathscr{X} \times \Theta$;

(c) f est de classe C^2, μ-p.p. pour tout $x \in \mathscr{X}$ (cf **dérivée**) ;

(d) le **support** Supp f ne dépend pas de θ et l'on peut dériver sous le signe d'**intégration** ;

(e) l'**information de FISHER** est une **matrice définie positive**.

(iv) Un exemple d'application est celui du **modèle de régression linéaire** multiple $y = X\, b + u$, dans lequel on suppose que $u \sim \mathscr{N}_N(0\ ,\ \Sigma)$ (où Σ est connue) (**loi normale multidimensionnelle** centrée), que $\Theta_0 = \{b \in \mathbf{R}^K : R\, b = r\}$, où $R \in M_{JK}(\mathbf{R})$ (matrice supposée de **rang** rg $\Sigma = J < K$) et $r \in \mathbf{R}^J$ sont connus, et où $\Theta_1 = \Theta_0^c$.

La statistique de WALD s'explicite ici selon :

(7) $W_N = (R\, b^\# - r)' \{R\, (X'\, \Sigma^{-1}\, X)^{-1}\, R'\}^{-1}\, (R\, b^\# - r)$,

où $b^\# = (X'\, \Sigma^{-1}\, X)^{-1}\, X'\, \Sigma^{-1}\, y$ est l'estimateur du maximum de vraisemblance (gaussien) non contraint de b (ie l'**estimateur des mcg** de b).

On remarque de W_N est égale à la **statistique des multiplicateurs de LAGRANGE** L_N et à celle du **rapport des vraisemblances** R_N .