

TEST DE GEARY (C5, D2, F3, I2, N2)

(19 / 11 / 2019, © Monfort, Dicostat2005, 2005-2019)

Le **test de GEARY** est un test alternatif au **test de DURBIN-WATSON** basé sur le signe des **résidus** d'une régression.

(i) Soit :

$$(1) \quad y = Xb + u, \quad \text{avec } E u = 0, \quad V u = \sigma_u^2 I_T,$$

un **modèle de régression linéaire** multiple exprimé dans l'**espace d'observation** R^N . On estime b par la **méthode des moindres carrés ordinaires**, ce qui conduit à l'**estimateur des mco** $\hat{b}$. On note $\hat{u} = (u_1^{\wedge}, \dots, u_T^{\wedge})'$ le vecteur des résidus des moindres carrés ordinaires qui s'en déduisent (ie $\hat{u} = y - X \hat{b}$).

On suppose que les **perturbations aléatoires** u_t ($t = 1, \dots, T$) sont susceptibles d'être autocorrélées au premier ordre, ie :

$$(2) \quad u_t = \rho \cdot u_{t-1} + v_t, \quad \text{avec } |\rho| \leq 1, \quad E v_t = 0 \quad \text{et} \quad V v_t = \sigma_v^2,$$

et que les v_t sont indépendantes et normales :

$$(3) \quad v_t \sim \mathcal{N}(0, \sigma_v^2), \quad \forall t \in N_T^*.$$

Par suite, $u \sim \mathcal{N}_T(0, \Sigma)$, où $\Sigma = (\sigma_{st})_{(s,t)}$ est une **matrice de dispersion** dont les termes sont : $\sigma_{st} = \sigma_v^2 \rho^{|s-t|}$.

Le **test de R.C. GEARY** concerne l'**hypothèse de base** :

$$(4) \quad H_0 : \rho = 0 \quad (\text{absence d'autocorrélation d'ordre 1})$$

et se fonde sur la **statistique de test** suivante, appelée **statistique de R.C. GEARY** :

$$(5) \quad G_T = \sum_{t=2}^T s_t,$$

avec :

$$(5)' \quad s_t = \begin{cases} 1 & \text{si } u_{t-1}^{\wedge} \cdot u_t^{\wedge} < 0, \\ 0 & \text{sinon,} \end{cases}$$

où l'inégalité $u_{t-1}^{\wedge} \cdot u_t^{\wedge} < 0$ traduit un changement de signe des résidus entre deux instants consécutifs ou entre deux périodes consécutives.

Ce test suppose les résidus $u_t^{\wedge}$ indépendants et de même **loi**, en sorte que :

$$(6) \quad \begin{aligned} E G_T &= (T - 1) / 2, \\ V G_T &= (T - 1) / 4. \end{aligned}$$

Les **valeurs critiques** sont donc obtenues à partir de la **loi binômiale** (ou de son **approximation** normale), dont la moyenne et la variance sont données par (6).

(ii) Si les **variables exogènes** du modèle (1) comportent une **constante** (eg si $x_1 = e = (1, \dots, 1)' \in \mathbf{R}^N$), la somme des résidus est nulle ($e_T' \hat{u} = 0$), donc les $\hat{u}_t$ sont négativement corrélés (même si les perturbations u n'étaient pas corrélées, ie même si $\rho = 0$). En effet, dans ce cas, $\hat{u} \sim \mathcal{N}_T(0, \Omega)$, avec $\Omega = M \Sigma M$ et $M = I_T - X(X'X)^{-1}X'$. Par suite, le calcul des deux premiers **moments** $E G_T$ et $V G_T$ est plus complexe.

Plusieurs approximations en ont été proposées, dont l'**approximation de D.K. GUILKEY - P. SCHMIDT** (cf aussi **approximation, correction**) :

$$(7) \quad \begin{aligned} E G_T &= 2^{-1} (T - 1) + \pi^{-1}, \\ V G_T &= 4^{-1} (T - 1) + \pi^{-2} (2 - \pi). \end{aligned}$$