

TEST DE KENDALL (D1, I2)

(15 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

Il existe plusieurs tests du nom de **test de KENDALL**, notamment un **test non paramétrique** d'indépendance, dont on présente deux formes.

(i) Le **problème de l'indépendance** conduit à étudier un **couple aléatoire** $\zeta = (\xi, \eta) : \Omega \mapsto \mathbf{R}^2$ dont on observe un **échantillon** $((X_1, Y_1), \dots, (X_N, Y_N))$ (**copies**, indépendantes entre elles, du couple (ξ, η)). On note $P^\zeta = P^{(\xi, \eta)}$ leur loi commune.

Le test de l'hypothèse de l'**indépendance** :

(1) $H_0 : \xi$ est indépendante de η ,

s'écrit aussi :

(1) $H_0 : P^{(\xi, \eta)} = P^\xi \otimes P^\eta$,

où l'on note P^ξ (resp P^η) la **loi marginale** (ie la loi « propre ») de ξ (resp de η).

(ii) Pour tester H_0 contre une alternative de **régression négative** (resp de régression positive), on définit la **statistique de rang** (non linéaire en (R, S)) suivante, dite **statistique de M.G. KENDALL** :

(2) $K_N = K_N' - K_N''$,

dans laquelle (en notant u la **fonction de HEAVYSIDE**) :

(2a) $K_N' = \sum_{\alpha=1}^N \sum_{\beta=1, (\beta \neq \alpha)}^N u(X_\alpha - X_\beta) \cdot u(Y_\alpha - Y_\beta) = \sum \sum_{\alpha < \beta} u(R_\beta^* - R_\alpha^*)$

représente le nombre de paires de couples $((X_\alpha, Y_\alpha), (X_\beta, Y_\beta))$ qui varient dans le même sens ;

(2b) $K_N'' = \sum_{\alpha=1}^N \sum_{\beta=1, (\beta \neq \alpha)}^N u(X_\beta - X_\alpha) \cdot u(Y_\alpha - Y_\beta) = \sum \sum_{\alpha < \beta} u(R_\alpha^* - R_\beta^*)$

représente le nombre de paires de couples qui varient en sens contraire.

La statistique K_N permet d'estimer sans biais le **coefficient** suivant, appelé τ (ou « tau ») de **M.G. KENDALL** (théorique) (cf aussi **coefficient de KENDALL**) :

(3) $\tau = 4 \cdot \int F(x, y) dP^{(\xi, \eta)}(x, y) - 1$,

dans lequel F est la **fonction de répartition** de (ξ, η) .

Ce coefficient vérifie les propriétés suivantes :

(a) $|\tau| \leq 1$ (cf aussi **coefficient de corrélation**) ;

(b) si $\tau > 0$ (resp si $\tau < 0$), on parle de **dépendance positive** (resp de **dépendance négative**) entre ξ et η .

La **statistique** suivante, appelée τ (ou « **tau** ») de **M.G. KENDALL** (empirique) :

$$(4) \quad \tau_N = 2 \cdot N^{-1} (N-1)^{-1} K_N,$$

est un **estimateur sans biais** (ie $E \tau_N = \tau$).

D'autre part, sous l'hypothèse H_0 d'indépendance, on établit la **convergence en loi** suivante :

$$(5) \quad \mathcal{L}\{K_N / S_0(K_N)\} \rightarrow_{N \rightarrow +\infty} \mathcal{N}(0, 1) \quad (\text{loi normale réduite}),$$

avec pour premiers moments :

$$(6) \quad \begin{aligned} E_0(K_N) &= 0, \\ V_0(K_N) &= (N-1)N(2N+5)/18 = \{S_0(K_N)\}^2, \end{aligned}$$

où $\{S_0(K_N)\}^2$ désigne le carré de l'**écart-type** empirique $S_0(K_N)$.

Le **test de M.G. KENDALL** en résulte.

(iii) Si l'on définit le **coefficient de M.G. KENDALL** (théorique) associé à la loi $P^{(\xi, \eta)}$ du couple (ξ, η) de façon analogue à (2) :

$$(7) \quad \tau = 2 P([(\xi' - \xi'') (\eta' - \eta'') > 0]) - 1,$$

l'hypothèse d'indépendance peut aussi s'écrire :

$$(8) \quad H_0 : \tau = 0 \quad (\text{car } P^{(\xi, \eta)} = P^\xi \otimes P^\eta).$$

On définit, par suite, la **statistique de M.G. KENDALL** selon :

$$(9) \quad K_N = \sum_{\alpha=1}^{N-1} \sum_{\substack{\beta=\alpha+1 \\ (\beta \neq \alpha)}}^N v \{ (X_\alpha, X_\beta), (Y_\alpha, Y_\beta) \},$$

où la fonction $v : \mathbf{R}^2 \times \mathbf{R}^2 \mapsto \mathbf{R}$ vérifie la définition suivante :

$$(10) \quad v \{ (a, b), (c, d) \} = \begin{aligned} &+ 1 \quad \text{ssi } (a - b) (c - d) > 0, \\ &- 1 \quad \text{ssi } (a - b) (c - d) < 0. \end{aligned}$$

Le **test de M.G. KENDALL**, fondé sur K_N , est le test asymptotique basé sur la **convergence en loi** ci-après :

$$(11) \quad \mathcal{L}\{K_N / S_0(K_N)\} \xrightarrow{N \rightarrow +\infty} \mathcal{N}(0, 1),$$

dans laquelle l'écart-type $S_0(K_N)$ est défini (sous l'hypothèse H_0) en (6).