CONCORDANCE (C5, D2, F3, I)
(14 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé** et $\zeta = (\xi, \eta) : \Omega \mapsto \mathbf{R}^2$ un **couple aléatoire** réel dont la **loi** est $P^{(\xi, \eta)}$.

Si (ξ', η') et (ξ'', η'') sont deux **copies** iid selon $P^{(\xi, \eta)}$, on dit que ce sont :

(a) des **copies concordantes** ssi : $(\xi' - \xi'').(\eta' - \eta'') > 0$;

(b) des **copies discordantes** ssi : $(\xi' - \xi'').(\eta' - \eta'') < 0$).

On appelle alors :

(a) **coefficient de concordance** (théorique) le nombre :

$(1)_a \quad \gamma = P([(\xi' - \xi'').(\eta' - \eta'') > 0])$.

(b) **coefficient de discordance** (théorique) le nombre :

$(1)_b \quad \gamma = P([(\xi' - \xi'').(\eta' - \eta'') < 0])$.

(ii) On appelle **coefficient de concordance (« tau ») de M.G. KENDALL** théorique (cf **coefficient de KENDALL**) le nombre :

$$(2) \qquad \tau = P([(\xi' - \xi'').(\eta' - \eta'') > 0]) - P([(\xi' - \xi'').(\eta' - \eta'') < 0]).$$

Par suite :

$$(3) \qquad \tau = 2\gamma - 1, \quad \text{ou encore} \quad \gamma = (\tau + 1) / 2.$$

(iii) Soit $Z = (Z_1, ..., Z_N)$ un N-**échantillon iid** selon $\zeta = (\xi, \eta)$, avec $Z_n = (X_n, Y_n)$, $\forall n \in \mathbf{N}_N^*$.

On appelle **coefficient de concordance de M.G. KENDALL** (empirique) la **statistique** :

$$(4) \qquad K_N = \Sigma_{\alpha=1}^{N-1} \Sigma_{\beta=\alpha+1}^{N} \delta_{\alpha\beta} ,$$

dans laquelle :

$$\delta_{\alpha\beta} = \begin{cases} -1 & \text{ssi } (X_\alpha - X_\beta).(Y_\alpha - Y_\beta) < 0, \\ 0 & \text{ssi } (X_\alpha - X_\beta).(Y_\alpha - Y_\beta) = 0, \\ +1 & \text{ssi } (X_\alpha - X_\beta).(Y_\alpha - Y_\beta) > 0. \end{cases}$$

Cette statistique s'écrit aussi :

(4)' $K_N = \Sigma_{\alpha=1}^{N-1} \Sigma_{\beta=\alpha+1}^{N} w(Z_\alpha, Z_\beta)$,

où :

$$w((a,b),(c,d)) = \begin{cases} -1 & \text{si } D < 0, \\ 0 & \text{si } D = 0, \\ +1 & \text{si } D > 0, \end{cases}$$

avec $D = (a - c).(b - d)$.

(iv) Pour tester l'**hypothèse statistique** d'**indépendance** $H_0 : P^{(\xi, \eta)} = P^\xi \otimes P^\eta$ (produit des **lois marginales** de ξ et de η), on peut tester l'hypothèse $H_0 : \gamma = 1 / 2$ (ou $\tau = 0$). Sous H_0, la statistique :

(5) $K_N' = K_N / C_N^2 = 2 . N^{-1} (N - 1)^{-1} K_N$

vérifie :

(7)
$$E_0 K_N' = 0,$$
$$V_0 K_N' = N^{-1} (N - 1)^{-1} \{4 . (N - 2) . a_1 + 2 . a_2\},$$

où $a_1 = C(\delta_{\alpha\beta}, \delta_{\alpha\gamma})$, $\forall (\alpha, \beta, \gamma)$ tq $\gamma \neq \beta \neq \alpha$, et $a_2 = V \delta_{\alpha\beta}$, $\forall (\alpha, \beta)$ tq $\beta \neq \alpha$.

A distance finie (avec $N << +\infty$), et sous H_0, la loi de K_N est tabulée.

Si $a_1 > 0$, on établit la **normalité asymptotique** :

(8) $\mathscr{L}\{(K_N' - E_0 K_N') / \sigma_0 (K_N')\} \to^{H0}_{N \to +\infty} \mathscr{N}_1 (0, 1)$ (**loi normale**),

ce qui permet de construire divers tests d'indépendance ou de concordance.