TEST DE LEPAGE (F3, F6, I2)
(25 / 11 / 2019, © Monfort, Dicostat2005, 2005-2019)

(i) Dans le cadre du problème à deux échantillons (cf **problème à plusieurs échantillons**, avec k = 2), le **test de Y. LEPAGE** a pour but de tester l'**hypothèse de base** (égalité des **lois**) :

(1) $H_0 : F_2 = F_1$ (**homogénéité** stricte des populations)

contre une **alternative** de la forme suivante (**alternative d'échelle et de position**) :

(2) H_1 : il existe $(\alpha, \beta) \in \mathbf{R} \times \mathbf{R}_+^*$ tq $F_2 (x) = F_1 \{(x - \alpha) / \beta\}$.

Ceci revient à tester $H_0' : (\alpha, \beta) = (0, 1)$ contre $H_1' : (\alpha , \beta) \neq (0, 1)$.

Soit $X = (X^1 , X^2)$ l'échantillon d'ensemble (empilé) et R sa **statistique de rang**. On pose :

$$(3) \quad \begin{aligned} S_{N(1)} &= \Sigma_{n = 1}^{N(1)} R_n \\ T_{N(1)} &= \Sigma_{n = 1}^{N(1)} |R_n - (1/2) (N + 1)|. \end{aligned}$$

On définit la **statistique de Y. LEPAGE** comme **statistique de test** de H_0 suivante :

$$(4) \quad L_{N(1)N(2)} = (\beta_{S,0})^{-2} (S_{N(1)} - \alpha_{S,0})^2 + (\beta_{T,0})^{-2} (T_{N(1)} - \alpha_{T,0})^2,$$

avec les deux premiers **moments** (calculés sous H_0) suivants :

$$(5) \quad \begin{aligned} \alpha_{S,0} &= E_0\, S_{N(1)} = (1/2)\, N_2\, (N+1), \\ \beta_{S,0} &= V_0\, S_{N(1)} = (1/12)\, N_1\, N_2\, (N+1), \end{aligned}$$

$$(6) \quad \alpha_{T,0} = E_0\, T_{N(1)} = \begin{cases} (1/4)\, N_2\, (N+2), & \text{ssi } N \in 2\, \mathbf{N}^*, \\ (4\, N)^{-1}\, N_2\, (N+1)^2, & \text{ssi } N \in 2\, \mathbf{N}^* + 1, \end{cases}$$

$$(7) \quad \beta_{T,0} = V_0\, T_{N(1)} = \begin{cases} (1/48)\, (N - 1)^{-1}\, N_1\, N_2\, (N^2 - 4), & \text{ssi } N \in 2\, \mathbf{N}^*, \\ (1/48)\, N^{-2}\, N_1\, N_2\, (N+1)\, (N^2 + 3), & \text{ssi } N \in 2\, \mathbf{N}^* + 1. \end{cases}$$

(ii) Le **test de Y. LEPAGE**. est réalisé selon l'un des contextes suivants :

(a) à distance finie (ie lorsque $N << +\infty$), la loi de $L_{N(1)N(2)}$ est tabulée sous l'hypothèse H_0 , ce qui permet d'effectuer le test ;

(b) si l'échantillon est grand (ie lorsque $N >> 0$), le test se fonde sur les **propriétés asymptotiques** suivantes (**convergences en loi**), valables sous l'hypothèse H_0 :

(8) $\mathscr{L}(S_{N(1)}, T_{N(1)}) \to_{\min\{N(1),N(2)\} \to +\infty} \mathscr{N}_2(0, I_2)$

(**loi normale multidimensionnelle**), et :

(9) $\mathscr{L}(L_{N(1)N(2)}) \to_{\min\{N(1),N(2)\} \to +\infty} \mathscr{X}_2^2$ (**loi du chi-deux** à 2 **degrés de liberté**),

où l'on note, par commodité, N(i) pour désigner la taille N_i de l'échantillon n° i (avec i = 1, 2).