TEST DES SCORES NORMAUX (C7, F6, I2)
(05 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

Ce **test de rang** est fondé sur une **statistique linéaire de rang** et son objet est la vérification de l'hypothèse d'**homogénéité** des **lois** dans le cadre du problème à deux échantillons (cf **problème à plusieurs échantillons**, **test de FISHER-YATES**).

(i) Soit $U = (U_1 ,..., U_N) \sim \mathcal{N}_N(0 , I_N)$ un **vecteur aléatoire** normal, $U^{(.)}$ la **statistique d'ordre** associée à U et R sa **statistique de rang**.

On appelle **statistique des scores normaux** la **statistique** suivante :

(1) $$S_N = \Sigma_{n(1)=1}^{N(1)} a_N (R_n)$$

dans laquelle la suite $(a_N (n))_{n=1,...,N}$ est la suite des **scores normaux** $a_N (n) = E\, U^{(n)}$ (pour tout n = 1,...,N) et n(i) (resp N(i)) désigne l'indice n_i (resp la taille N_i de l'échantillon n° i), avec i = 1 , 2.

Cette suite vérifie la propriété suivante :

(2) $$a_N (N - n) = - a_N (n), \qquad \forall\, n = 1,...,N.$$

(ii) Sous l'hypothèse H_0 d'**homogénéité** (stricte) des deux populations dont sont extraits les échantillons X^1 et X^2 (ie égalité des **fonctions de répartition** F_1 et F_2 qui les génère respectivement) on montre que :

(3) $$E_0\, S_N = 0,$$
$$V_0\, S_N = \{\sigma_0 (S_N)\}^2 = N^{-1} (N-1)^{-1} N_1 \,.\, N_2 \,.\, \Sigma_{n=1}^{N} \{a_N (n)\}^2,$$

où $N = N_1 + N_2$ et $\sigma_0 (S_N)$ désigne l'**écart-type** de S_N .

La table « statistique » de la loi de S_N est définie pour de « petits » échantillons (ie $\max (N_1 , N_2) << +\infty$), ce qui permet de définir des **régions critiques** à distance finie, associées à tout seuil $\alpha \in]0,1[$.

Si la taille des échantillons est « grande » (ie $N_1 >> 0$ et $N_2 >> 0$), on utilise la **propriété asymptotique** suivante (**convergence légale**) :

(4) $$\mathcal{L}(U_N) \rightarrow_{\min (N(1), N(2)) \to \infty} \mathcal{N}(0, 1)$$ (**loi normale réduite**),

où U_N désigne la statistique centrée réduite $(S_N - E_0\, S_N) / \sigma_0 (S_N)$.

On peut donc alors définir des régions critiques (asymptotiques) à partir de la loi normale et le test ainsi mis en oeuvre est appelé **test des scores normaux**.

Ce test est souvent utilisé contre des **alternatives d'échelle**.