TEST D'ANALYSE DE LA VARIANCE (J3, L)
(06 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé**, $(X^i) = (X_{i1}, ..., X_{i,N(i)})$ $(i = 1, ..., k)$ des **échantillons** donnés (cf **problème à plusieurs échantillons**), chacun issu d'une **variable parente** $\xi_i : \Omega \mapsto \mathbf{R}$ dont l'**espérance mathématique** $E\,\xi_i$ est notée μ_i .

On veut tester l'**hypothèse** d'**homogénéité** (stricte) (au premier ordre) des k **populations** $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, P^{\xi i})$ suivante :

(1) $H_0 : \mu_j = \mu_i , \quad \forall (i, j),$

contre l'**alternative** générale :

(2) H_1 : il existe (i, j) tq $\mu_j \neq \mu_i$,

on pose :

(3)
$$\bar{X}_i = N_i^{-1} \Sigma_{n(i)=1}^{N(i)} X_{in(i)} \qquad \text{(moyenne empirique de l'échantillon i, } \forall i \in Nk^*),$$
$$\bar{X}_N = N^{-1} \Sigma_{n=1}^{N} X_n \qquad \text{(moyenne empirique d'ensemble),}$$

où X est l'échantillon d'ensemble (ie le regroupement des échantillons X^i), n(i) et N(i) désignent resp n_i et N_i , et $N = \Sigma_{i=1}^k N_i$. On utilise alors une décomposition classique $Q_1 + Q_2$ de la **forme quadratique** Q associée à la **variance** totale suivante (cf **formule de KOENIG-HUYGENS**) :

(4) $\Sigma_{i=1}^k \{\Sigma_{n(i)=1}^{N(i)} (X_{in(i)} - \bar{X}_N)^2\} = Q_1 + Q_2 ,$

avec $Q_1 = \Sigma_{i=1}^k \{\Sigma_{n(i)=1}^{N(i)} (X_{in(i)} - \bar{X}_i)^2\}$ (**forme quadratique** intra-classes) et $Q_2 = \Sigma_{i=1}^k N_i (\bar{X}_i - \bar{X}_N)^2$ (forme quadratique inter-classes).

Par suite, si les **va** parentes ξ_i sont indépendantes, gaussiennes et homoscédastiques (ie leurs variances sont égales à une valeur commune) (cf **homoscédasticité**), on montre que la **statistique** suivante :

(5) $F_{N(1)...N(k)} = S_2^2 / S_1^2,$

dans laquelle $S_2^2 = (k - 1)^{-1} Q_2$ et $S_1^2 = (N - k)^{-1} Q_1$, suit une **loi de FISHER-SNEDECOR** à k - 1 et N - k **degrés de liberté**, ie :

(6) $F_{N(1)...N(k)} \sim \mathcal{F}(k - 1, N - k).$