ANALYSE DE LA VARIANCE À DEUX FACTEURS (J3, L)
(12 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

L'**analyse de la variance à deux facteurs** est basée sur un **modèle d'analyse de la variance** associé à un plan d'expérience comportant deux **facteurs expérimentaux**.

(i) Soit F un **facteur** comportant I niveaux i = 1 ,..., I, G un facteur comportant J niveaux j = 1 ,..., J et N_{ij} le nombre d'**unités expérimentales** (ou **répétitions**) auxquelles est appliqué le **traitement** (i, j), avec $N = \Sigma_i \Sigma_j N_{ij}$ (nombre total d'unités). On observe sur ces unités les valeurs (résultats) $y_{ij,n(i,j)} \in \mathbf{R}$.

Dans ce cadre, un **modèle d'analyse de la variance** à deux facteurs, non équilibré et avec **intéractions**, peut s'écrire sous la forme :

(1) $$y_{ij,n(i,j)} = \beta^0 + \beta_i^1 + \beta_j^2 + \beta_{ij}^{12} + u_{ij,n(i,j)} ,$$

avec $n_{ij} \in \{1 ,..., N_{ij}\}$ et $(i, j) \in \{1,..., I\} \times \{1,..., J\}$.

Les hypothèses stochastiques les plus simples pour un tel modèle sont eg :

(2) $$\begin{aligned} &E\, u_{ij,n(i,j)} = 0, \\ &C\, (u_{ij,n(i,j)} , u_{kl,n(k,l)}) = \delta_{(i,j)(k,l)} \,.\, \sigma^2, \end{aligned}$$

pour tout (i, j, k, l), où $\delta_{(i,j)(k,l)}$ désigne le **symbole de KRONECKER**.

Enfin, les **conditions d'identification** du modèle imposent des **contraintes sur les paramètres**, eg :

(3) $$\begin{aligned} &\Sigma_{i=1}^{I} \beta_i^1 = \Sigma_{j=1}^{J} \beta_j^2 = 0, \\ &\Sigma_{i=1}^{I} \beta_{ij}^{12} = \Sigma_{j=1}^{J} \beta_{ij}^{12} = 0, \qquad \forall\ (i, j). \end{aligned}$$

(ii) Le modèle est équilibré si $N_{ij} = N_{00}$, $\forall$ (i, j), d'où $N = I\ J\ N_{00}$. L'estimation des paramètres β_i^1, β_j^2 et β_{ij}^{12} peut alors s'effectuer à l'aide de la **méthode des moindres carrés ordinaires**, contraints par (3) (cf **contrainte**, **régression contrainte**). Celle-ci conduit aux estimateurs suivants :

(4) $$\begin{aligned} &(\beta^0)^{\sim} = y_{...} , \\ &(\beta_i^1)^{\sim} = y_{i..} - y_{...} , \\ &(\beta_j^2)^{\sim} = y_{.j.} - y_{...} , \\ &(\beta_{ij}^{12})^{\sim} = y_{ij.} - y_{i..} - y_{.j.} + y_{...} , \end{aligned}$$

où l'on note :

$y_{...} = N^{-1} \Sigma_{i=1}^{I} \Sigma_{j=1}^{J} \Sigma_{n(00)=1}^{N(00)} y_{ij\ n(00)}$ (moyenne générale),

(5) $y_{i..} = (J\ N_{(00)})^{-1} \Sigma_{j=1}^{J} \Sigma_{n(00)=1}^{N(00)} y_{ij\ n(00)}$ (effet propre au niveau i de F),

$y_{.j.} = (I\ N_{(00)})^{-1} \Sigma_{i=1}^{I} \Sigma_{n(00)=1}^{N(00)} y_{ij\ n(00)}$ (effet propre au niveau j de G),

expressions dans lesquelles n(00) désigne n_{00} et N(00) désigne N_{00} .

Le plan étant équilibré, les **hypothèses** suivantes :

$H_1 : \beta_i^1 = 0, \forall\ i,$

(6) $H_2 : \beta_j^2 = 0, \forall\ j,$

$H_{12} : b_{ij}^{12} = 0, \forall\ (i, j),$

sont deux à deux orthogonales (ie $H_1 \perp H_2$, $H_2 \perp H_{12}$, $H_{12} \perp H_1$). Si l'on pose (somme des carrés des **perturbations**) :

(7) $q\ (u) = \Sigma_{i=1}^{I} \Sigma_{j=1}^{J} \Sigma_{n(00)=1}^{N(00)} u_{ij\ n(00)}^{\ 2},$

et si l'on admet l'hypothèse « stochastique » supplémentaire :

(8) $u_{ij\ n(00)} \sim \mathcal{N}(0, \sigma^2),\ \forall\ (i , j), \quad \forall\ n_{00}$,

on montre que, sous l'hypothèse H_1 :

(9) $I\ J\ (N_{00} - 1)\ (I - 1)^{-1} \{(S_1^2 - S^2) / S^2\} \sim \mathcal{F}(I\text{-}1, I\ J\ (N_{00} -1))$ (**loi de FISHER**),

où S^2 est la valeur minimum de q (u) dans le modèle initial {(1),(2)} (supposé équilibré), avec q (u) = S^2, et S_1^2 est la valeur minimum de q (u) le modèle {(1),(2)} (supposé équilibré) compte tenu de l'hypothèse H_1 .

Par ailleurs, sous l'hypothèse H_{12} , on montre que :

(10) $I\ J\ (N_{00} - 1)\ (I - 1)^{-1}(J - 1)^{-1}\ (S_{12}^2 - S^2) / S^2) \sim \mathcal{F}((I\text{-}1)(J\text{-}1), I\ J\ (N_{00} -1)),$

où S_{12}^2 est la valeur minimum de q (u) dans le modèle {(1),(2)} (supposé équilibré) compte tenu de l'hypothèse H_{12} .

(iii) Dans le cas où N_{00} = 1, les procédures précédentes ne s'appliquent plus. On doit généralement se restreindre à un **modèle additif**, ie tq $\beta_{ij}^{12} = 0, \forall\ (i, j)$. Ainsi en est-il dans le cas d'un **plan en blocs complets randomisés**.

(iv) Le modèle à effets fixes précédent (modèle de type I) peut être remplacé par un modèle à effets aléatoires (modèle de type II) ou par un modèle à effets mixtes (modèle de type III) (cf **modèle d'analyse de la variance**).