Université de Parakou (UP)
Ecole Nationale de Statistique, de Planification
et de Démographie (ENSPD)

Master 1 Statistique, Planification et Démographie (Semestre 8)

ENSPD 2024-2025

Cours de Modèle Mixte

Enseignant :
Raymond M M AFFOSSOGBE
Docteur en Statistique-Probabilité
Email : rafmoqui@yahoofr
+229 96877413

Table des matières

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Introduction

Un modèle mixte, également connu sous le nom de modèle à effets mixtes ou modèle à mesures répétées, est un modèle statistique utilisée pour analyser des données qui présentent à la fois des facteurs fixes et aléatoires. Ce type de modèle est couramment utilisé dans les domaines de la recherche médicale, des sciences sociales, de la psychologie et d'autres disciplines où les données sont collectées à partir d'observations répétées sur les mêmes sujets ou unités expérimentales.

L'objectif principal d'un modèle mixte est de capturer à la fois les effets fixes et les effets aléatoires qui peuvent influencer les données. Voici les concepts clés associés aux modèles mixtes :

1. **Effets Fixes :** Ce sont des facteurs qui sont manipulés ou choisis délibérément dans une étude. Ils représentent les variables indépendantes qui sont d'intérêt pour le chercheur. Par exemple, dans une étude sur l'effet de différents médicaments, le type de médicament serait considéré comme un effet fixe.
2. **Effets Aléatoires :** Ce sont des facteurs qui ne sont pas contrôlés, mais qui peuvent contribuer à la variabilité observée dans les données. Ils représentent généralement des variables qui varient naturellement entre les sujets ou les unités expérimentales. Dans une étude impliquant plusieurs participants, l'effet individuel de chaque participant serait considéré comme un effet aléatoire.
3. **Effets Mixtes :** Les modèles mixtes combinent à la fois des effets fixes et des effets aléatoires. Ils permettent de prendre en compte la variabilité entre les sujets tout en analysant les effets des variables indépendantes d'intérêt. Ainsi, un modèle mixte peut aider à démêler les effets individuels des participants des effets dus aux manipulations expérimentales.
4. **Structure de Covariance :** Les modèles mixtes tiennent également compte de la corrélation potentielle entre les observations répétées d'un même sujet. La structure de covariance est importante pour gérer la dépendance entre les observations et pour obtenir des estimations statistiques précises.

5. **Estimation :** Les paramètres des modèles mixtes sont généralement estimés à l'aide de méthodes d'estimation comme la méthode des moindres carrés restreints (REML) ou la méthode des moindres carrés généralisés (GLS).

En résumé, les modèles mixtes offrent une approche puissante pour analyser des données complexes qui présentent à la fois des effets fixes et aléatoires. Ils sont utiles pour modéliser des situations où les données sont collectées à partir de sujets ou d'unités expérimentales répétées, tout en tenant compte des différences individuelles et de la variabilité inhérente.

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Chapitre un

Généralités sur le modèle linéaire

Un modèle de régression linéaire est un modèle statistique fondamentale utilisée pour décrire la relation entre une variable dépendante continue (ou variable réponse) et une ou plusieurs variables indépendantes (ou variables prédictives). L'objectif principal de la régression linéaire est de trouver la meilleure ligne droite (ou hyperplan dans le cas de régressions multiples) qui représente au mieux la relation entre les variables.

Les modèles de régression linéaire peuvent également être étendus à des cas plus complexes, tels que la régression linéaire multiple, où plusieurs variables indépendantes sont utilisées pour prédire la variable dépendante :

Le modèle de régression linéaire simple est défini par l'équation :

1.1 Ecriture du modèle

Pour un individu i

$$Y_i = \beta_0 + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \ldots + \beta_p \cdot X_{pi} + e_i$$

Où $X_{1i}, X_{2i}, \ldots, X_{pi}$ sont les variables indépendantes, $\beta = (\beta_0, \beta_1, \ldots, \beta_p)$ sont les coefficients correspondants, et e_i est toujours l'erreur aléatoire et sa représentation vectorielle s'écrit :

$$Y_i = (\mathbf{X}_i)^T \beta + e_i$$

. Pour N individus on a la représentation matricielle. Classiquement, on écrit :

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e} \tag{1.1}$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

où

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix}, \mathbf{X} = \begin{pmatrix} \mathbf{X}'_1 \\ \mathbf{X}'_2 \\ \vdots \\ \mathbf{X}'_N \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & \cdots & X_{1K} \\ 1 & X_{21} & \cdots & X_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{N1} & \cdots & X_{NK} \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_K \end{pmatrix}, \mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{pmatrix}$$

- **y** : est le vecteur $(N \times 1)$ des variables aléatoires dépendantes (observations) ;
- $\mathbf{X} = \left(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k, \ldots, \mathbf{X}_p\right)$: est la matrice $(N \times p)$ d'incidence des variables explicatives(dites aussi variables «indépendantes» ou covariables) qui peuvent être continues (régression) ou discrètes (ANOVA) ;
- $\boldsymbol{\beta}$: est le vecteur $(p\times 1)$ des coefficients dits de régression (covariables continues) ou «effets fixes) (covariables discrètes) ; a priori $\beta \in R^p$;
- **e** : est le vecteur $(N \times 1)$ de variables aléatoires résiduelles.

1.2 Hypothèses

Toutes ou partie des hypothèses suivantes peuvent être formulées ou requises selon les techniques statistiques employées :

- **Indépendance :** $\mathbf{e}_i$ est indépendante de $\mathbf{X}_i$.
- **Exogénéité :** On dit que les variables explicatives sont exogènes si elles ne sont pas corrélées au terme d'erreur. Ce qu'on note, pour le cas où la variable explicative est aléatoire, $\mathbb{E}\left(e_i \mid X_i\right) = 0$ en notation vectorielle et $\mathbb{E}(\mathbf{e} \mid \mathbf{X}) = 0$ en notation matricielle où $X = (\mathbb{I}_N, \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_K)$. Ceci implique que les erreurs sont centrées. Si les variables **X** sont constantes ceci est noté $\mathbb{E}\left(\varepsilon_i\right) = 0$[5].
- **spécification :** $E(y) = \mathbf{X}\beta$ si **X** est déterministe $E(y|\mathbf{X}) = \mathbf{X}\beta$ si **X** est une variable aléatoire.
- **indépendance des erreurs :** $\forall i \neq j \quad \operatorname{Cov}\left(e_i, e_j\right) = 0$
- **homoscédasticité :** Les termes d'erreurs sont supposés de variance constante, ce qui se traduit, si l'hypothèse précédente est vérifiée, par $\forall i = 1, \ldots, K \quad \mathbb{E}\left(e_i^2 \mid X_i\right) = \sigma^2$ si X est une variable aléatoire ou un ensemble de variables aléatoires, et par $\forall i = 1, \ldots, N \quad \mathbb{E}\left(e_i^2\right) = \sigma^2$ sinon. Si les deux précédentes hypothèses sont vérifiées, on peut l'écrire sous forme matricielle :

  $$\mathbb{V}(\mathbf{e} \mid \mathbf{X}) = \sigma^2 I_N$$

  avec I_N la matrice identité de taille N.
- **normalité :** Une hypothèse plus forte que les premières est celle consistant à dire que les termes d'erreurs suivent une loi normale, centrées, de variance σ^2 soit, $\varepsilon_i \mid x_i \sim \mathcal{N}(0, \sigma^2)$ en notation vectorielle et sous forme matricielle $\varepsilon \mid X \sim \mathcal{N}(0, \sigma^2 I_N)$.

1.3 Estimation

Soit $S(\beta) = ||\mathbf{y} - \mathbf{X}\beta||^2$ le carré de la distance euclidienne entre les observations et la partie explicative du modèle, considérée comme une fonction de β. L'estimateur des moindres carrés des paramètres β du modèle est donné par

$$\hat{\beta} = \arg\min_{\beta} S(\beta)$$

Or

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \tag{1.2}$$

$$\frac{\partial^2 S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} = 2\mathbf{X}'\mathbf{X} \text{ (définie non négative)} \tag{1.3}$$

La condition de convexité de la fonction $S(\beta)$ étant satisfaite, $\hat{\beta}$ est la solution de l'équation $\frac{\partial S(\beta)}{\partial \beta} = 0$ donnée par

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{y} \tag{1.4}$$

si la matrice $\mathbf{X}'\mathbf{X}$ est inversible ; sinon, on utilise en lieu et place l'inverse généralisé. Sous forme vectorielle, l'équation (1.4) s'écrit

$$\hat{\beta} = \left(\frac{1}{n}\sum \mathbf{X}_i\mathbf{X}_i'\right)^{-1}\left(\frac{1}{n}\sum \mathbf{X}_i y_i\right).$$

1.4 Modèle de régression linéaire mixte

Ce modèle s'écrit de la même manière que dans le cas précédent seulement

$$\mathbf{y} = \mathbf{X}\beta + \sum_{k=1}^{r} \mathbf{Z}_k\mathbf{u}_k + \mathbf{e}$$

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

tels que

— $\mathbf{y}$ est le vecteur $N \times 1$ défini précédemment,

— $\boldsymbol{\beta}$ est le vecteur $K \times 1$ inconnu de facteurs fixes à estimer.

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

- **X** est la matrice de taille $N \times K$ de données connu associés à β ; c'est souvent une matrice d'incidence (avec des éléments 0 et 1), mais peut aussi inclure des colonnes de covariables d'autres types.
- **u** est le vecteur des effets aléatoires. ;
- **Z** est la matrices d'incidence correspondant à **u**.
- **e** est le vecteur des termes résiduels.
- $\mathbf{Z}_k$ est la matrice des indicatrices (disposées en colonnes) des niveaux du k-ième facteur à effets aléatoires ($k = 1, \ldots, r$) ; on notera q_k le nombre de niveaux de ce facteur ; $\mathbf{Z}_k$ est donc de dimension $N \times q_k$.
- Nous allons noter $u_{k\ell}$ la v.a.r. associée au ℓ-ième niveau du k-ième facteur à effets aléatoires ($\ell = 1, \ldots, q_k$) ; pour tout ℓ, on a $u_{\mathrm{k}\,\ell} \sim \left(0, \sigma_k^2\right)$; pour un indice k donné (autrement dit, pour un facteur déterminé), les v.a.r. $u_{k\ell}$ sont supposées indépendantes (donc i.i.d.) ; bien entendu, deux observations de la v.a.r. réponse Y faites au même niveau ℓ du k-ième facteur sont corrélés, leur covariance comportant le terme σ_k^2 et, éventuellement, d'autres composantes de la variance ; par ailleurs, pour deux indices k et k' distincts, $u_{k\ell}$ et $u_{k'\ell'}$ sont indépendantes, pour tous les niveaux ℓ et ℓ'.
- Dans l'écriture matricielle ci-dessus, on a posé $\mathbf{u}_k = \begin{pmatrix} u_{k1} \\ \vdots \\ u_{kqk} \end{pmatrix}$, de sorte que l'on a $\mathbf{u}_k \sim \left(0, \sigma_k^2 \mathbf{I}_{q_k}\right)$, les vecteurs aléatoires $\mathbf{u}_k$ étant mutuellement indépendants.
- **e** est le vecteur aléatoire des erreurs du modèle ; il vérifie $\mathbf{e} \sim \mathcal{N}_n\left(0, \sigma^2 \mathbf{I}_n\right)$; de plus, il est supposé indépendant des $\mathbf{u}_k$.

Les propriétés habituellement attribuées à **u** et **e** sont les suivantes

$$\begin{array}{lll} \mathrm{E}(\mathbf{u}) = \mathbf{0} & \operatorname{var}(\mathbf{u}) = \mathbf{D} & \operatorname{cov}(\mathbf{u}, \mathbf{e}') = \mathbf{0} \\ \mathrm{E}(\mathbf{e}) = \mathbf{0} & \operatorname{var}(\mathbf{e}) = \mathbf{R}, \operatorname{var}(\mathbf{e}) = \sigma_e^2 \mathbf{I} \text{ dans la plupart des cas.} & \end{array}$$

u est tel que

$$\mathbf{u}' = \left[\begin{array}{cccccc} \mathbf{u}_1' & \mathbf{u}_2' & \cdots & \mathbf{u}_i' & \cdots & \mathbf{u}_r' \end{array}\right]$$
$$Z = \left[Z_1 Z_2 \cdots Z_i \cdots Z_r\right].$$

Si chaque $\mathbf{u}_i$ le i-ième facteur a q_i effets aléatoires alors

$$\operatorname{var}(\mathbf{u}_i) = \sigma_i^2 \mathbf{I}_{q_i} \quad \text{and} \quad \operatorname{cov}\left(\mathbf{u}_i, \mathbf{u}_{i'}'\right) = \mathbf{0} \quad \text{for i} \neq \text{i}'$$

and so

$$\mathrm{D} = \operatorname{diag}\left(\sigma_i^2 \mathbf{I}_{q_1}, \cdots, \sigma_r^2 \mathbf{I}_{q_r}\right)$$

Nous avons de ce fait :

$$
\begin{aligned}
\mathrm{E}(\mathbf{y} \mid \mathbf{u}) &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} \\
\mathrm{E}(\mathbf{y}) &= \mathbf{X}\boldsymbol{\beta} \\
\mathrm{var}(\mathbf{y}) &= \mathbf{V} = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R} = \sum_{i=i}^{r} \sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i' + \mathbf{R}.
\end{aligned}
$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Chapitre deux

Analyse des Variances (ANOVA)

Définition 2.1 (Plan équilibré, Plan déséquilibré) *Tout dispositif expérimental comportant au moins un facteur et comportant un nombre identique de répétitions dans chacune des modalités des facteurs est un plan équilibré. Un plan non équilibré est dit déséquilibré*

2.1 Analyse de la variance à un facteur

2.1.1 Modèles à un effet fixe

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Pour chacune des modalités nous effectuons $J \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J$ le nombre total de mesures ayant été effectuées. Nous introduisons le modèle :

$$Y_{i,j} = \mu + \alpha_i + \epsilon_{i,j}, \quad i = 1 \dots I, j = 1 \dots J$$
$$\text{avec la contrainte supplémentaire } \sum_{i=1}^{I} \alpha_i = 0,$$

où

— $Y_{i,j}$ est la valeur prise par la réponse Y dans la condition A_i lors de la j-ème répétition.
— μ est l'influence moyenne du facteur A sur Y
— α_i est l'effet de la classe i et est déterministe

Nous postulons les hypothèses classiques suivantes pour les erreurs :

— $\epsilon_{i,j} \perp\!\!\!\perp \epsilon_{k,l}$ si $(i,j) \neq (k,l)$ avec $1 \leqslant i,k \leqslant I,$ et $1 \leqslant j,l \leqslant J,$

— $\epsilon_{i,j} \sim \mathcal{N}(0, \sigma^2) \quad \forall (i,j), \quad 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, \quad .$

Nous regroupons les valeurs que peut prendre la réponse Y dans la condition A_i lors des J répétitions dans le tableau suivant :

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Facteur A	Y
A_1	$Y_{1,1}, \dots, Y_{1,J}$
$\vdots$	$\vdots$
A_i	$Y_{i,1}, \dots, Y_{i,J}$
$\vdots$	$\vdots$
A_I	$Y_{I,1}, \dots, Y_{I,J}$

Nous rappelons que la variation théorique due au facteur A est définie par :

$$SC_F = J \sum_{i=1}^{I} (Y_{i,\bullet} - Y_{\bullet,\bullet})^2$$

La variation résiduelle théorique est quant à elle définie par :

$$SC_R = \sum_{i=1}^{I} \left(\sum_{j=1}^{J} \left(Y_{i,j} - Y_{i,\bullet} \right)^2 \right)$$

Enfin la variation totale théorique est égale à :

$$SC_{TOT} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left(Y_{i,j} - Y_{\bullet,\bullet} \right)^2 .$$

où,

$$Y_{\bullet,j} = \frac{1}{I} \sum_{i=1}^{I} Y_{i,j} = \frac{1}{I} Y_{+,j}.$$

$$Y_{i,\bullet} = \frac{1}{J} \sum_{j=1}^{J} Y_{i,j} = \frac{1}{J} Y_{i,+}$$

$$Y_{\bullet,\bullet} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{i,j} = \frac{1}{IJ} Y_{+,+}.$$

Nous rappelons la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_F + SC_R$$

La liste $\boldsymbol{y}$ des données expérimentales $y_{1,1}, \dots, y_{1,J}, \dots, y_{2,1}, \dots, y_{2,J}, \dots, y_{I,J}$ permet de construire une réalisation du tableau précédent :

Facteur A	$\boldsymbol{y}$
A_1	$y_{1,1},\ldots,y_{1,J}$
$\vdots$	$\vdots$
A_i	$y_{i,1},\ldots,y_{i,J}$
$\vdots$	$\vdots$
A_I	$y_{I,1},\ldots,y_{I,J}$

et donc du tableau d'ANOVA des données éxpérimentales

Source	Variation	Degrés de liberté	Carré Moyen	F	Décision
Facteur	sc_F	$I-1$	$s_F^2=\frac{sc_F}{I-1}$	$f=\frac{s_F^2}{s_R^2}$	$\mathcal{H}_0$ ou $\mathcal{H}_1$
Résiduelle	sc_R	$n-I=I(J-1)$	$s_R^2=\frac{sc_R}{n-I}$		
Totale	sc_{TOT}	$n-1=IJ-1$			

Où $\mathcal{H}_0$ et $\mathcal{H}_1$ sont les hypothèses suivantes de test d'ANOVA.

$$\mathcal{H}_0 : \alpha_1=\alpha_2=\cdots=\alpha_I=0$$
$$\text{contre}$$
$$\mathcal{H}_1 : \text{ Il existe } i_0\in\{1,2,\ldots,I\} \text{ tel que } \alpha_{i_0}\neq 0.$$

$\mathcal{H}_0$ sous entend qu'il y a absence d'effet du traitement (facteur) A sur Y.
$F=\dfrac{Sc_F}{Sc_R}\sim$ Fisher de degré de liberté $(I-1,n-I)$

Règle de décision : On rejette $\mathcal{H}_0$ au risque α si la p-value $<\alpha$.

Estimateur des paramètres du modèle : Les estimateurs $\widehat{\mu},\widehat{\alpha_1},\ldots,\widehat{\alpha_I},\widehat{\sigma^2}$ des paramètres $\mu,\alpha_1,\ldots,\alpha_I,\sigma^2$ du modèle sont donnés par les formules suivantes :

$$\begin{aligned}
\widehat{\mu} &= Y_{\bullet,\bullet}=\bar{Y},\\
\widehat{\alpha_i} &= Y_{i,\bullet}-\widehat{\mu}, 1\leqslant i\leqslant I,\\
\widehat{\sigma^2} &= \frac{SC_R}{n-I}=\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(Y_{i,j}-Y_{i,\bullet}\right)^2}{n-I}=S_R^2.
\end{aligned}$$

2.1.2 Modèles à un effet aléatoire

Exemple 2.1 *On a relevé les durées de gestation de 16 filles de 30 taureaux qui avaient été tirés au sort dans la population devant être étudiée. On voudrait savoir dans quelle mesure la* *durée de gestation est un caractère héréditaire - caractère se transmettant par les pères. Il s'agit de répondre, grâce à un échantillon comportant peu de taureaux, à une question concernant toute la population. Pour pouvoir étendre les résultats obtenus sur l'échantillon, il faut que celui-ci soit représentatif de toute la population et donc qu'il ait été obtenu par tir age au sort (indépendants et équiprobables). Il en découle que les taureaux de l'échantillon sont aléatoires et leurs effets sur leurs descendants sont a fortiori aléatoires.*

Le modèle s'écrira

$$y_{ij} = \mu + a_i + e_{ij} \quad j = 1, \ldots 16 \quad i = 1, \ldots 30$$

où y_{ij} représente la durée de gestation de la fille j du père i, μ est la moyenne générale, a_i est l'effet du père i, supposé alé atoire puisque le père est un individu tiré aléatoirement, e_{ij} est le résidu. On suppose les distributions suivantes :

$$a_i \sim N\left(0, \sigma_a^2\right)$$
$$e_{ij} \sim N\left(0, \sigma_e^2\right)$$

Les a_i et les e_{ij} sont supposés mutuellement indépendants.

Exemple 2.2 *En génétique quantitative, il n'est pas raisonnable de supposer que les effets de certains facteurs, comme l'effet de la mère sur les performances des descendants, est fixe. Il est supposé aléatoire, car cet effet peut être assimilé à un tirage aléatoire dans le génome de la mère. De façon plus générale, on considère que les effets d'un facteur sont aléatoires si les niveaux de ce facteur ne sont pas choisis mais tirés aléatoirement dans un ensemble de niveaux.*

Dans ce cas les A_i représentent un échantillon de taille I prélevé dans une population importante. Nous admettrons que les effets des A_i, les α_i, sont distribués suivant une loi normale centrée de variance σ_A^2. Le modèle ne dépend plus que de trois paramètres μ, σ^2 et σ_A^2. Pour chacune des modalités nous effectuons $J \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j} = \mu + \alpha_i + \epsilon_{i,j}, \quad i = 1 \ldots I, j = 1 \ldots J$$

où $Y_{i,j}$ est la valeur prise par la réponse Y dans la condition A_i lors de la j-ème répétition. Nous supposons que :

— $\mathcal{L}(\alpha_i) = \mathcal{N}\left(0, \sigma_A^2\right), \forall i, 1 \leqslant i \leqslant I,$

— $\alpha_i \perp\!\!\!\perp \alpha_j$ si $i \neq j$ et $1 \leqslant i, j \leqslant I$.

Nous postulons les hypothèses supplémentaires pour les erreurs :
— $\epsilon_{i,j} \perp\!\!\!\perp \epsilon_{k,l}$ si $(i,j) \neq (k,l)$ avec $1 \leqslant i,k \leqslant I,$ et $1 \leqslant j,l \leqslant J,$
— $\epsilon_{i,j} \sim \mathcal{N}(0,\sigma^2)$ $\forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J.$
— $\alpha_i \perp\!\!\!\perp \epsilon_{j,k}$ si $1 \leqslant i, j \leqslant I$, et $1 \leqslant k \leqslant J$.
Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Degrés de liberté	Carré Moyen	F	Décision
Facteur	sc_F	$I-1$	$s_F^2 = \frac{sc_F}{I-1}$	$f = \frac{s_F^2}{s_R^2}$	$\mathcal{H}_0$ ou $\mathcal{H}_1$
Résiduelle	sc_R	$n-I = I(J-1)$	$s_R^2 = \frac{sc_R}{n-I}$		
	sc_{TOT}	$n-1 = IJ-1$			

Où $\mathcal{H}_0$ et $\mathcal{H}_1$ sont les hypothèses suivantes de test d'ANOVA.

$$\mathcal{H}_0 : \sigma_A^2 = 0$$

contre

$$\mathcal{H}_1 : \sigma_A^2 \neq 0.$$

$\mathcal{H}_0$ sous entend qu'il y a absence d'effet du traitement (facteur) A sur Y.

$$F = \frac{Sc_F}{Sc_R} \sim \text{Fisher de degré de liberté } (I-1, n-I)$$

Règle de décision : On rejette $\mathcal{H}_0$ au risque α si la p-value $< \alpha$.

Estimateur des paramètres du modèle : Les estimateurs $\widehat{\mu}, \widehat{\sigma_A^2}, \widehat{\sigma^2}$ des paramètres $\mu, \sigma_A^2, \sigma^2$ du modèle sont donnés par les formules suivantes :

— $\widehat{\mu} = Y_{\bullet,\bullet} = \bar{Y}$,

— $\widehat{\sigma_A^2} = \dfrac{1}{J}\left(S_F^2 - S_R^2\right)$,

— $\widehat{\sigma^2} = \dfrac{SC_R}{n-I} = S_R^2$

— où $S_F^2 = \dfrac{SC_F}{I-1}$ et $S_R^2 = \dfrac{SC_R}{n-I}$.

Ce sont des estimateurs sans biais 2 .

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

2.2 Analyse de la variance à deux facteurs

2.2.1 Modèles à effets fixes

Sans répétition :

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Un facteur contrôlé B se présente sous J modalités, chacune d'entre elles étant notée B_j. Pour chacun des couples de modalités $\left(A_i, B_j\right)$ nous effectuons une mesure d'une réponse Y qui est une variable continue. Nous notons $n = I \times J$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j} = \mu + \alpha_i + \beta_j + \epsilon_{i,j}, \quad i = 1 \ldots I, j = 1 \ldots J$$

$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0 \text{ et } \sum_{j=1}^{J} \beta_j = 0,$$

où $Y_{i,j}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j\right)$. Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, \mathcal{L}\left(\epsilon_{i,j}\right) = \mathcal{N}\left(0, \sigma^2\right),$$

$$\epsilon_{i,j} \perp\!\!\!\perp \epsilon_{k,l} \text{ si } (i,j) \neq (k,l) \text{ avec } 1 \leqslant i,k \leqslant I \text{ et } 1 \leqslant j,l \leqslant J.$$

Nous regroupons les valeurs que peut prendre la réponse Y dans les conditions $\left(A_i, B_j\right)$ dans le tableau suivant :

Facteur A	B_1	$\ldots$	B_j	$\ldots$	B_J
	$Y_{1,1}$	$\ldots$	$Y_{1,j}$	$\ldots$	$Y_{1,J}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_i	$Y_{i,1}$	$\ldots$	$Y_{i,j}$	$\ldots$	$Y_{i,J}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_I	$Y_{I,1}$	$\ldots$	$Y_{1,j}$	$\ldots$	$Y_{I,J}$

Nous rappelons que la variation théorique due au facteur A est définie par :

$$SC_A = J \sum_{i=1}^{I} \left(Y_{i,\bullet} - Y_{\bullet,\bullet}\right)^2$$

Nous rappelons que la variation théorique due au facteur B est définie par :

$$SC_B = I \sum_{j=1}^{J} \left(Y_{\bullet,j} - Y_{\bullet,\bullet}\right)^2.$$

La variation résiduelle théorique est quant à elle définie par :

$$SC_R = \sum_{i=1}^{I} \sum_{j=1}^{J} \left(Y_{i,j} - Y_{i,\bullet} - Y_{\bullet,j} + Y_{\bullet,\bullet}\right)^2.$$

Enfin la variation totale théorique est égale à :

$$SC_{TOT} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left(Y_{i,j} - Y_{\bullet,\bullet}\right)^2.$$

Nous rappelons la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_B + SC_R$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I - 1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_R^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J - 1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_R^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Résiduelle	sc_R	$n_R = (I-1)(J-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJ - 1$			

Où $\mathcal{H}_0'$ et $\mathcal{H}_1'$ sont les hypothèses suivantes de test d'ANOVA.

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$

contre

$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

$F_A = \dfrac{Sc_A}{Sc_R} \sim$ Fisher de degré de liberté $[I-1, (I-1)(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0$ au risque α si la p-value $< \alpha$.

et $\mathcal{H}_0''$ et $\mathcal{H}_1''$ sont les hypothèses suivantes de test d'ANOVA.

$$\mathcal{H}_0'' : \beta_1 = \beta_2 = \cdots = \beta_J = 0$$

contre

$$\mathcal{H}_1'' : \text{ Il existe } j_0 \in \{1, 2, \ldots, J\} \text{ tel que } \beta_{j_0} \neq 0.$$

$F_B = \dfrac{Sc_B}{Sc_R} \sim$ Fisher de degré de liberté $[J-1,(I-1)(J-1)]$
Règle de décision : On rejette $\mathcal{H}_0$ au risque α si la p-value $< \alpha$.
Estimateur des paramètres du modèle : Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\beta_1}, \ldots, \widehat{\beta_J}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \beta_1, \ldots, \beta_J$, σ^2 du modèle sont donnés par les formules suivantes :

— $\widehat{\mu} = Y_{\bullet,\bullet} = \bar{Y}$,

— $\widehat{\alpha_i} = Y_{i,\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I$,

— $\widehat{\beta_j} = Y_{\bullet,j} - \widehat{\mu}, 1 \leqslant j \leqslant J$,

— $\widehat{\sigma^2} = \frac{SC_R}{(I-1)(J-1)} = s_R^2$.

Ce sont des estimateurs sans biais 1 .

Avec répétitions :

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Un facteur contrôlé B se présente sous J modalités, chacune d'entre elles étant notée B_j. Pour chacun des couples de modalités $\left(A_i, B_j\right)$ nous effectuons $K \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \epsilon_{i,j,k}, \quad i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K,$$

$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0, \sum_{j=1}^{J} \beta_j = 0,$$

$$\sum_{i=1}^{I} (\alpha\beta)_{i,j} = 0, \forall j \in \{1, \ldots, J\} \text{ et } \sum_{j=1}^{J} (\alpha\beta)_{i,j} = 0, \forall i \in \{1, \ldots, I\},$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j\right)$ lors du k-ème essai. Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\text{et } \epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec}$$
$$1 \leqslant i, l \leqslant I, 1 \leqslant j, m \leqslant J \text{ et } 1 \leqslant k, n \leqslant K.$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Nous regroupons les valeurs que peut prendre la réponse Y dans les conditions $\left(A_i, B_j\right)$ lors des K répétitions dans le tableau suivant :

Facteur A	Facteur B				
	B_1	$\ldots$	B_j	$\ldots$	B_J
A_1	$Y_{1,1,1}\ldots Y_{1,1,K}$	$\ldots$	$Y_{1,j,1}\ldots Y_{1,j,K}$	$\ldots$	$Y_{1,J,1}\ldots Y_{1,J,K}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_i	$Y_{i,1,1}\ldots Y_{i,1,K}$	$\ldots$	$Y_{i,j,1}\ldots Y_{i,j,K}$	$\ldots$	$Y_{i,J,1}\ldots Y_{i,J,K}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_I	$Y_{I,1,1}\ldots Y_{I,1,K}$	$\ldots$	$Y_{I,j,1}\ldots Y_{I,j,K}$	$\ldots$	$Y_{I,J,1}\ldots Y_{I,J,K}$

la variation théorique due au facteur A est définie par :

$$SC_A = JK\sum_{i=1}^{I}\left(Y_{i,\bullet,\bullet} - Y_{\bullet,\bullet,\bullet}\right)^2.$$

Nous rappelons que la variation théorique due au facteur B est définie par :

$$SC_B = IK\sum_{j=1}^{J}\left(Y_{\bullet,j,\bullet} - Y_{\bullet,\bullet,\bullet}\right)^2.$$

Nous rappelons que la variation théorique due à l'interaction des facteurs A et B est définie par :

$$SC_{AB} = K\sum_{i=1}^{I}\sum_{j=1}^{J}\left(Y_{i,j,\bullet} - Y_{i,\bullet,\bullet} - Y_{\bullet,j,\bullet} + Y_{\bullet,\bullet,\bullet}\right)^2.$$

La variation résiduelle théorique est quant à elle définie par :

$$SC_R = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(Y_{i,j,k} - Y_{i,j,\bullet}\right)^2.$$

où

$$Y_{\bullet,\bullet,\bullet} = \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{i,j,k} = \frac{1}{IJK}Y_{+,+,+}.$$

$$Y_{i,\bullet,\bullet} = \frac{1}{JK}\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{i,j,k} = \frac{1}{JK}Y_{i,+,+}.$$

$$Y_{\bullet,j,\bullet} = \frac{1}{IK}\sum_{i=1}^{I}\sum_{k=1}^{K} Y_{i,j,k} = \frac{1}{IK}Y_{+,j,+}.$$

$$Y_{\bullet\bullet\bullet,k} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{i,j,k} = \frac{1}{IJ}Y_{+,+,k}.$$

$$Y_{i,j,\bullet} = \frac{1}{K}\sum_{k=1}^{K} Y_{i,j,k} = \frac{1}{K}Y_{i,j,+}.$$

$$Y_{i,\bullet,k} = \frac{1}{J}\sum_{j=1}^{J} Y_{i,j,k} = \frac{1}{J}Y_{i,+,k}.$$

$$Y_{\bullet,j,k} = \frac{1}{I}\sum_{i=1}^{I} Y_{i,j,k} = \frac{1}{I}Y_{+,j,k}.$$

Enfin la variation totale théorique est égale à :

$$SC_{TOT} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(Y_{i,j,k} - Y_{\bullet,\bullet,\bullet}\right)^2.$$

Nous rappelons la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_B + SC_{AB} + SC_R$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I - 1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_R^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J - 1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_R^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Interaction	sc_{AB}	$n_{AB} = (I-1)(J-1)$	$s_{AB}^2 = \frac{sc_{AB}}{n_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_R^2}$	$\mathcal{H}_0'''$ ou $\mathcal{H}_1'''$
Résiduelle	sc_R	$n_R = IJ(K-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJK - 1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$

contre

$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

$\mathcal{H}'_0$ sous entend l'absence d'effet du traitement A sur Y
$F_A = \dfrac{Sc_A}{Sc_R} \sim$ Fisher de degré de liberté $[I-1, IJ(K-1)]$
Règle de décision : On rejette $\mathcal{H}'_0$ au risque α si la p-value $< \alpha$.

et $\mathcal{H}''_0$ et $\mathcal{H}''_1$ sont les hypothèses suivantes de test d'ANOVA.

$$\mathcal{H}''_0 : \beta_1 = \beta_2 = \cdots = \beta_J = 0$$
$$\text{contre}$$
$$\mathcal{H}''_1 : \text{ Il existe } j_0 \in \{1, 2, \ldots, J\} \text{ tel que } \beta_{j_0} \neq 0.$$

$\mathcal{H}''_0$ sous entend l'absence d'effet du traitement A sur Y
$F_B = \dfrac{Sc_B}{Sc_R} \sim$ Fisher de degré de liberté $[J-1, IJ(K-1)]$
Règle de décision : On rejette $\mathcal{H}''_0$ au risque α si la p-value $< \alpha$.

et $\mathcal{H}'''_0$ et $\mathcal{H}'''_1$ sont les hypothèses suivantes de test d'ANOVA.

$$\mathcal{H}'''_0 : (\alpha\beta)_{1,1} = (\alpha\beta)_{1,2} = \cdots = (\alpha\beta)_{1,J} = (\alpha\beta)_{2,1} = \cdots = (\alpha\beta)_{I,J} = 0$$
$$\text{contre}$$
$$\mathcal{H}'''_1 : \text{ Il existe } (i_0, j_0) \in \{1, 2, \ldots, I\} \times \{1, 2, \ldots, J\} \text{ tel que } (\alpha\beta)_{i_0, j_0} \neq 0.$$

$\mathcal{H}'''_0$ sous entend l'absence d'interaction des effets des traitements A et B sur Y
$F_{AB} = \dfrac{Sc_{AB}}{Sc_R} \sim$ Fisher de degré de liberté $[J-1, IJ(K-1)]$
Règle de décision : On rejette $\mathcal{H}'''_0$ au risque α si la p-value $< \alpha$.

Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\beta_1}, \ldots, \widehat{\beta_J}, \widehat{(\alpha\beta)_{1,1}}, \widehat{(\alpha\beta)_{1,2}}, \ldots, \widehat{(\alpha\beta)_{1,J}}, \widehat{(\alpha\beta)_{2,1}}, \ldots, \widehat{(\alpha\beta)_{I,J}}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \beta_1, \ldots, \beta_J, (\alpha\beta)_{1,1}, (\alpha\beta)_{1,2}, \ldots, (\alpha\beta)_{1,J}, (\alpha\beta)_{2,1}, \ldots, (\alpha\beta)_{I,J}, \sigma^2$ du modèle sont donnés par les formules suivantes :

$$\widehat{\mu} = Y_{\bullet,\bullet\bullet\bullet} = \bar{Y}, \widehat{\alpha_i} = Y_{i,\bullet\bullet\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I, \widehat{\beta_j} = Y_{\bullet,j,\bullet} - \widehat{\mu}, 1 \leqslant j \leqslant J$$
$$\widehat{(\alpha\beta)_{i,j}} = Y_{i,j,\bullet} - Y_{i,\bullet,\bullet} - Y_{\bullet,j,\bullet} + Y_{\bullet,\bullet,\bullet}, 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J$$
$$\widehat{\sigma^2} = \frac{SC_R}{IJ(K-1)} = S_R^2.$$

Ce sont des estimateurs sans biais .

Emboité (ici c'est obligatoirement à répétition)

Nous sommes dans la situation particulière où les effets des niveaux du facteur B n'ont pas de signification concrète, par exemple ces niveaux dépendent du niveau

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

du facteur A considéré et une étude des effets principaux du facteur B n'a pas de pertinence.

Nous ne pouvons nous servir d'un modèle où les facteurs sont emboîtés \1 que si nous disposons de répétitions. Dans le cas contraire où les essais ne seraient pas répétés, l'effet dû au facteur B ne pourra être étudié et le modèle que nous devrons utiliser pour analyser les données sera l'un de ceux exposés au chapitre 2.

Ainsi par exemple un fabriquant de détergents alimente plusieurs chaînes de distribution : $A_1, A_2, \ldots, A_I$. Nous pensons que les boîtes de produit livrées à certaines chaînes de distribution contiennent une masse de détergent inférieure à celle des autres chaînes de distribution. Pour étudier cette situation, nous décidons de prélever K boîtes dans J magasins de chaque chaîne. Ainsi le second facteur B_j, associé au j-ème magasin dans la chaîne, est un repère qui n'a aucune signification réelle : il n'y a, par exemple aucune relation entre le magasin $n°3$ de la chaîne 1 et le magasin $n°3$ de la chaîne 4 . Il n'y a donc aucun intérêt à introduire un terme dans le modèle caractérisant l'effet principal du facteur B. Pour indiquer la dépendance des niveaux du second facteur B aux niveaux du premier facteur A nous notons les niveaux du second facteur B : $B_{j(i)}, 1 \leqslant i \leqslant I$ et $1 \leqslant j \leqslant J$.

1. Ces types de modèles sont également appelés des modèles hiérarchiques ou en anglais hierarchical ou nested models.

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Un facteur contrôlé B se présente sous J modalités, chacune d'entre elles dépendant du niveau A_i du facteur A et étant alors notée $B_{j(i)}$. Pour chacun des couples de modalités $\left(A_i, B_{j(i)}\right)$ nous effectuons $K \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{i,j,k}, \quad i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K$$
$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0, \sum_{j=1}^{J} \beta_{j(i)} = 0, \forall i \in \{1, \ldots, I\}$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_{j(i)}\right)$ lors du k-ème essai. Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\text{et } \epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec}$$
$$1 \leqslant i, l \leqslant I, 1 \leqslant j, m \leqslant J \text{ et } 1 \leqslant k, n \leqslant K.$$

Nous regroupons les valeurs que peut prendre la réponse Y dans les conditions $\left(A_i, B_{j(i)}\right)$ lors des K essais dans le tableau suivant :

A_1			$\dots$			A_I		
$B_{1(1)}$	$\dots$	$B_{J(1)}$	$\dots$	$\dots$	$\dots$	$B_{1(I)}$	$\dots$	$B_{J(I)}$
$Y_{1,1,1}$	$\dots$	$Y_{1,J,1}$	$\dots$	$\dots$	$\dots$	$Y_{I,1,1}$	$\dots$	$Y_{I,J,1}$
$\vdots$	$\vdots$	$\vdots$	$\dots$	$\dots$	$\dots$	$\vdots$	$\vdots$	$\vdots$
$Y_{1,1,K}$	$\dots$	$Y_{1,J,K}$	$\dots$	$\dots$	$\dots$	A_I	$\dots$	$Y_{I,J,K}$

la variation théorique due au facteur A est définie par :

$$SC_A = JK \sum_{i=1}^{I} \left(Y_{i,\bullet,\bullet} - Y_{\bullet,\bullet,\bullet}\right)^2 .$$

La variation théorique du facteur B dans le facteur A est définie par :

$$SC_{B|A} = K \sum_{i=1}^{I} \sum_{j=1}^{J} \left(Y_{i,j,\bullet} - Y_{i,\bullet,\bullet}\right)^2$$

La variation résiduelle théorique est quant à elle définie par :

$$SC_R = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(Y_{i,j,k} - Y_{i,j,\bullet}\right)^2 .$$

Enfin la variation totale théorique est égale à :

$$SC_{TOT} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(Y_{i,j,k} - Y_{\bullet,\bullet,\bullet}\right)^2 .$$

Nous rappelons la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_{B|A} + SC_R$$

On fait la remarque que

$$SC_{B|A} = SC_B + SC_{AB}.$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I-1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_R^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B dans A	$sc_{B\|A}$	$n_{B\|A} = J(I-1)$	$s_{B\|A}^2 = \frac{sc_{B\|A}}{n_{B\|A}}$	$f_{B\|A} = \frac{s_{B\|A}^2}{s_R^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Résiduelle	sc_R	$n_R = IJ(K-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJK-1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$
$$\text{contre}$$
$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

$F_A = \dfrac{S_A^2}{S_R^2} \sim$ Fisher de degré de liberté $[I-1, IJ(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0'' : \beta_{1(1)} = \beta_{2(1)} = \cdots = \beta_{J(1)} = \beta_{1(2)} = \cdots = \beta_{J(I)} = 0$$
$$\text{contre}$$
$$\mathcal{H}_1'' : \text{ Il existe } (i_0, j_0) \in \{1, 2, \ldots, I\} \times \{1, 2, \ldots, J\} \text{ tel que } \beta_{j_0(i_0)} \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet des facteurs B dans le facteur A.

$F_{B|A} = \dfrac{S_{B|A}^2}{S_R^2} \sim$ Fisher de degré de liberté $[I(J-1), IJ(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0''$ au risque α si la p-value $< \alpha$.

Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\beta_{1(1)}}, \widehat{\beta_{2(1)}}, \ldots, \widehat{\beta_{J(1)}}, \widehat{\beta_{1(2)}}, \ldots, \widehat{\beta_{J(I)}}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \beta_{1(1)}, \beta_{2(1)}$ du modèle sont donnés par les formules suivantes :

$$\widehat{\mu} = Y_{\bullet,\bullet,\bullet} = \bar{Y}, \widehat{\alpha_i} = Y_{i,\bullet,\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I,$$
$$\widehat{\beta_{j(i)}} = Y_{i,j,\bullet} - Y_{i,\bullet,\bullet}, 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,$$
$$\widehat{\sigma^2} = \frac{SC_R}{IJ(K-1)} = S_R^2.$$

Ce sont des estimateurs sans biais.

2.2.2 Modèles à effets aléatoires

Sans répétition

Dans ce cas les A_i représentent un échantillon de taille I prélevé dans une population importante. Nous admettrons que les effets des A_i, les α_i, sont distribués suivant une loi normale centrée de variance σ_A^2. les B_j représentent un échantillon de taille J prélevé dans une population importante. Nous admettrons que les effets des B_j, les β_j, sont distribués suivant une loi normale centrée de variance σ_B^2. Pour chacun des couples de modalités $\left(A_i, B_j\right)$ nous effectuons une mesure d'une réponse Y qui est une variable continue. Nous notons $n = I \times J$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j} = \mu + \alpha_i + \beta_j + \epsilon_{i,j}, \quad i = 1 \ldots I, j = 1 \ldots J$$

où $Y_{i,j}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j\right)$. Nous supposons que :

$$\mathcal{L}\left(\alpha_i\right) = \mathcal{N}\left(0, \sigma_A^2\right), \forall i, 1 \leqslant i \leqslant I,$$
$$\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J,$$

ainsi que l'indépendance des effets aléatoires :

$$\alpha_i \perp\!\!\!\perp \alpha_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant I, \quad \beta_i \perp\!\!\!\perp \beta_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant J,$$
$$\alpha_i \perp\!\!\!\perp \beta_j \text{ si } 1 \leqslant i \leqslant I \text{ et } 1 \leqslant j \leqslant J.$$

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i, j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, \mathcal{L}\left(\epsilon_{i,j}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\epsilon_{i,j} \perp\!\!\!\perp \epsilon_{k,l} \text{ si } (i, j) \neq (k, l) \text{ avec } 1 \leqslant i, k \leqslant I \text{ et } 1 \leqslant j, l \leqslant J,$$

ainsi que l'indépendance des effets aléatoires et des erreurs :

$$\alpha_i \perp\!\!\!\perp \epsilon_{j,k} \text{ si } 1 \leqslant i, j \leqslant I \text{ et } 1 \leqslant k \leqslant J,$$
$$\beta_i \perp\!\!\!\perp \epsilon_{j,k} \text{ si } 1 \leqslant j \leqslant I \text{ et } 1 \leqslant i, k \leqslant J.$$

Nous rappelons la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_B + SC_R$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I - 1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_R^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J - 1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_R^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Résiduelle	sc_R	$n_R = (I-1)(J-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJ - 1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \sigma_A^2 = 0$$
contre
$$\mathcal{H}_1' : \sigma_A^2 \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

$F_A = \dfrac{S_A^2}{S_R^2} \sim$ Fisher de degré de liberté $[I-1, (I-1)(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0'' : \sigma_B^2 = 0$$
contre
$$\mathcal{H}_1'' : \sigma_B^2 \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet du traitement B sur Y

$F_A = \dfrac{S_B^2}{S_R^2} \sim$ Fisher de degré de liberté $[J-1, (I-1)(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0''$ au risque α si la p-value $< \alpha$.

Estimateur des paramètres du modèle : Les estimateurs $\widehat{\mu}, \widehat{\sigma_A^2}, \widehat{\sigma_B^2}, \widehat{\sigma^2}$ des paramètres $\mu, \sigma_A^2, \sigma_B^2, \sigma^2$ du modèle sont donnés par les formules suivantes :

$$\begin{aligned}
\widehat{\mu} &= Y_{\bullet,\bullet} = \bar{Y}, \\
\widehat{\sigma_A^2} &= \frac{1}{J}\left(S_A^2 - S_R^2\right), \\
\widehat{\sigma_B^2} &= \frac{1}{I}\left(S_B^2 - S_R^2\right), \\
\widehat{\sigma^2} &= \frac{SC_R}{(I-1)(J-1)} = S_R^2
\end{aligned}$$

où $S_A^2 = \frac{SC_A}{n_A}, S_B^2 = \frac{SC_B}{n_B}$ et $S_R^2 = \frac{SC_R}{n_R}$.
Ce sont des estimateurs sans biais

rafmoqui@yahoo.fr ~ 96877413

Avec répétitions

Dans ce cas les A_i représentent un échantillon de taille I prélevé dans une population importante. Nous admettrons que les effets des A_i, les α_i, sont distribués suivant une loi normale centrée de variance σ_A^2. Les β_j représentent un échantillon de taille J prélevé dans une population importante. Nous admettrons que les effets des B_j, les β_j, sont distribués suivant une loi normale centrée de variance σ_B^2. Pour chacun des couples de modalités $\left(A_i, B_j\right)$ nous effectuons $K \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \epsilon_{i,j,k}, \quad i = 1 \dots I, j = 1 \dots J, k = 1 \dots K,$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j\right)$ lors du k-ème essai. Nous supposons que

$$\begin{aligned}
\mathcal{L}(\alpha_i) &= \mathcal{N}\left(0, \sigma_A^2\right), \forall i, 1 \leqslant i \leqslant I, \\
\mathcal{L}\left(\beta_j\right) &= \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J, \\
\mathcal{L}\left((\alpha\beta)_{i,j}\right) &= \mathcal{N}\left(0, \sigma_{AB}^2\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,
\end{aligned}$$

ainsi que l'indépendance des effets aléatoires :

$$\begin{gathered}
\alpha_i \perp\!\!\!\perp \alpha_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant I, \quad \beta_i \perp\!\!\!\perp \beta_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant J, \\
(\alpha\beta)_{i,j} \perp\!\!\!\perp (\alpha\beta)_{k,l} \text{ si } (i,j) \neq (k,l) \text{ avec } 1 \leqslant i, k \leqslant I \text{ et } 1 \leqslant j, l \leqslant J, \\
\alpha_i \perp\!\!\!\perp \beta_j \text{ si } 1 \leqslant i \leqslant I \text{ et } 1 \leqslant j \leqslant J, \\
\alpha_i \perp\!\!\!\perp (\alpha\beta)_{j,k} \text{ si } 1 \leqslant i, j \leqslant I \text{ et } 1 \leqslant k \leqslant J, \\
\beta_i \perp\!\!\!\perp (\alpha\beta)_{j,k} \text{ si } 1 \leqslant j \leqslant I \text{ et } 1 \leqslant i, k \leqslant J.
\end{gathered}$$

Nous postulons les hypothèses classiques pour les erreurs

$$\begin{gathered}
\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right), \\
\text{et } \epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec} \\
1 \leqslant i, l \leqslant I, 1 \leqslant j, m \leqslant J \text{ et } 1 \leqslant k, n \leqslant K,
\end{gathered}$$

ainsi que l'indépendance des effets aléatoires et des erreurs :

$$\begin{gathered}
\alpha_i \perp\!\!\!\perp \epsilon_{j,k,l} \text{ si } 1 \leqslant i, j \leqslant I, 1 \leqslant k \leqslant J, \text{ et } 1 \leqslant l \leqslant K, \\
\beta_i \perp\!\!\!\perp \epsilon_{j,k,l} \text{ si } 1 \leqslant j \leqslant I, 1 \leqslant i, k \leqslant J, \text{ et } 1 \leqslant l \leqslant K, \\
(\alpha\beta)_{i,j} \perp\!\!\!\perp \epsilon_{k,l,m} \text{ si } 1 \leqslant i, k \leqslant I, 1 \leqslant j, l \leqslant J, \text{ et } 1 \leqslant m \leqslant K.
\end{gathered}$$

On a la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_B + SC_{AB} + SC_R$$

et le tableau ANOVA qui s'écrit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A I - 1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_{AB}^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J - 1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_{AB}^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Interaction	sc_{AB}	$n_{AB} = (I-1)(J-1)$	$s_{AB}^2 = \frac{sc_{AB}}{n_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_R^2}$	$\mathcal{H}_0'''$ ou $\mathcal{H}_1'''$
Résiduelle	sc_R	$n_R = IJ(K-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJK - 1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \sigma_A^2 = 0$$

contre

$$\mathcal{H}_1' : \sigma_A^2 \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

$F_A = \dfrac{S_A^2}{S_{AB}^2} \sim$ Fisher de degré de liberté $[I-1, (I-1)(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0'' : \sigma_B^2 = 0$$

contre

$$\mathcal{H}_1'' : \sigma_B^2 \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet du traitement B sur Y

$F_B = \dfrac{S_B^2}{S_{AB}^2} \sim$ Fisher de degré de liberté $[J-1, (I-1)(J-1)]$

$$\mathcal{H}_0''' : \sigma_{AB}^2 = 0$$

contre

$$\mathcal{H}_1''' : \sigma_{AB}^2 \neq 0.$$

$\mathcal{H}_0'''$ sous entend l'absence d'effet du traitement A et B sur Y

$F_B = \dfrac{S_B^2}{S_{AB}^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1), IJ(K-1)]$

Estimateur des paramètres du modèle : Les estimateurs $\widehat{\mu}, \widehat{\sigma_A^2}, \widehat{\sigma_B^2}, \widehat{\sigma_{AB}^2}, \widehat{\sigma^2}$ des paramètres $\mu, \sigma_A^2, \sigma_B^2, \sigma_{AB}^2, \sigma^2$ du modèle sont donnés par les formules suivantes :

$$\begin{aligned}
\widehat{\mu} &= Y_{\bullet,\bullet,\bullet} = \bar{Y}, \\
\widehat{\sigma_A^2} &= \frac{1}{JK}\left(S_A^2 - S_{AB}^2\right), \\
\widehat{\sigma_B^2} &= \frac{1}{IK}\left(S_B^2 - S_{AB}^2\right), \\
\widehat{\sigma_{AB}^2} &= \frac{1}{K}\left(S_{AB}^2 - S_R^2\right), \\
\widehat{\sigma^2} &= \frac{SC_R}{(I-1)(J-1)} = S_R^2,
\end{aligned}$$

où $S_A^2 = \frac{SC_A}{n_A}, S_B^2 = \frac{SC_B}{n_B}, S_{AB}^2 = \frac{SC_{AB}}{n_{AB}}$ et $S_R^2 = \frac{SC_R}{n_R}$.
Ce sont des estimateurs sans biais.

Emboité (ici c'est toujours avec répétition)

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Les $\beta_{j(i)}$ représentent un échantillon de taille J prélevé dans une population importante dépendant du niveau A_i du facteur A. Nous admettrons que les effets des $B_{j(i)}$, les $\beta_{j(i)}$, sont distribués suivant une loi normale centrée de variance $\sigma_{B|A}^2$. Pour chacun des couples de modalités $\left(A_i, B_{j(i)}\right)$ nous effectuons $K \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{i,j,k}, \quad i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_{j(i)}\right)$ lors du k-ème essai. Nous supposons que

$$\begin{aligned}
\mathcal{L}(\alpha_i) &= \mathcal{N}\left(0, \sigma_A^2\right), \forall i, 1 \leqslant i \leqslant I, \\
\mathcal{L}\left(\beta_{j(i)}\right) &= \mathcal{N}\left(0, \sigma_{B|A}^2\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,
\end{aligned}$$

ainsi que l'indépendance des effets aléatoires :

$$\begin{aligned}
&\alpha_i \perp\!\!\!\perp \alpha_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant I, \\
&\beta_{j(i)} \perp\!\!\!\perp \beta_{l(k)} \text{ si } (i,j) \neq (k,l) \text{ avec } 1 \leqslant i,k \leqslant I \text{ et } 1 \leqslant j,l \leqslant J, \\
&\alpha_i \perp\!\!\!\perp \beta_{k(j)} \text{ si } 1 \leqslant i,j \leqslant I \text{ et } 1 \leqslant k \leqslant J.
\end{aligned}$$

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall(i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\text{et } \epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec}$$
$$1 \leqslant i,l \leqslant I, 1 \leqslant j,m \leqslant J \text{ et } 1 \leqslant k,n \leqslant K,$$

ainsi que l'indépendance des effets aléatoires et des erreurs :

$$\alpha_i \perp\!\!\!\perp \epsilon_{j,k,l} \text{ si } 1 \leqslant i,j \leqslant I, 1 \leqslant k \leqslant J, \text{ et } 1 \leqslant l \leqslant K,$$
$$\beta_{j(i)} \perp\!\!\!\perp \epsilon_{k,l,m} \text{ si } 1 \leqslant i,k \leqslant I, 1 \leqslant j,l \leqslant J, \text{ et } 1 \leqslant m \leqslant K.$$

Nous rappelons la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_{B|A} + SC_R$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I - 1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_{B\|A}^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B dans A	$sc_{B\|A}$	$n_{B\|A} = I(J-1)$	$s_{B\|A}^2 = \frac{sc_{B\|A}}{n_{B\|A}}$	$f_{B\|A} = \frac{s_{B\|A}^2}{s_R^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Résiduelle	sc_R	$n_R = IJ(K-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJK - 1$			

où les hypothèses sont les suivantes Nous souhaitons faire les tests d'hypothèse suivants :

$$\mathcal{H}_0' : \sigma_A^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1' : \sigma_A^2 \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

Sous $\mathcal{H}_0'$ $F_A = \dfrac{S_A^2}{S_{B|A}^2} \sim$ Fisher de degré de liberté $[I-1, I(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0'' : \sigma_{B|A}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1'' : \sigma_{B|A}^2 \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet du facteur B dans le facteur A.
Sous $\mathcal{H}_0''$ $F_{B|A} = \dfrac{S^2_{B|A}}{S_R} \sim$ Fisher de degré de liberté $[I(J-1), IJ(K-1)]$
Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

Les estimateurs $\widehat{\mu}, \widehat{\sigma^2_A}, \widehat{\sigma^2_{B|A}}, \widehat{\sigma^2}$ des paramètres $\mu, \sigma^2_A, \sigma^2_{B|A}, \sigma^2$ du modèle se déduisent des formules suivantes :

$$\begin{aligned}
\widehat{\mu} &= Y_{\bullet,\bullet,\bullet} = \bar{Y}, \\
\widehat{\sigma^2_A} &= \frac{1}{JK}\left(S^2_A - S^2_{B|A}\right), \\
\widehat{\sigma^2_{B|A}} &= \frac{1}{K}\left(S^2_{B|A} - S^2_R\right), \\
\widehat{\sigma^2} &= \frac{SC_R}{(I-1)(J-1)} = S^2_R,
\end{aligned}$$

où $S^2_A = \frac{SC_A}{n_A}, S^2_{B|A} = \frac{SC_{B|A}}{n_{B|A}}$ et $S^2_R = \frac{SC_R}{n_R}$. Ce sont des estimateurs sans biais

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Chapitre trois

Modèles à effets mixte

Exemple 3.1 *Le modèle linéaire mixte est largement utilisé dans le cadre de l'analyse de données longitudinales (plusieurs observations ont été mesurés sur le même individu au cours du temps) car il permet de prendre en compte des corrélations entre les observations. Si le nombre d'instants de mesures sur un individu est faible, on utilise le modèle mixte pour prendre en compte la dépendance entre observations issues du même individu. En revanche si le nombre d'instants de mesure n'est pas faible et si l'évolution dans le temps est le sujet d'intérêt, alors il faudra utiliser une autre catégorie de modèles, les modèles de séries chronologiques.*

3.1 Modèle à effet mixte à deux facteurs et sans répétitions

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Les β_j représentent un échantillon de taille J prélevé dans une population importante. Nous admettrons que les effets des B_j, les β_j, sont distribués suivant une loi normale centrée de variance σ_B^2. Pour chacun des couples de modalités $\left(A_i, B_j\right)$ nous effectuons une mesure d'une réponse Y qui est une variable continue. Nous notons $n = I \times J$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j} = \mu + \alpha_i + \beta_j + \epsilon_{i,j}, \quad i = 1 \ldots I, j = 1 \ldots J$$

$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0$$

où $Y_{i,j}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j\right)$. Nous supposons que :

$$\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J$$

ainsi que l'indépendance des effets aléatoires :

$$\beta_i \perp\!\!\!\perp \beta_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant J$$

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, \mathcal{L}\left(\epsilon_{i,j}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\epsilon_{i,j} \perp\!\!\!\perp \epsilon_{k,l} \text{ si } (i,j) \neq (k,l) \text{ avec } 1 \leqslant i,k \leqslant I \text{ et } 1 \leqslant j,l \leqslant J,$$

ainsi que l'indépendance des effets aléatoires et des erreurs :

$$\beta_i \perp\!\!\!\perp \epsilon_{j,k} \text{ si } 1 \leqslant j \leqslant I \text{ et } 1 \leqslant i,k \leqslant J.$$

La relation fondamentale de l'ANOVA s'écrit :

$$SC_{TOT} = SC_A + SC_B + SC_R$$

le tableau de l'ANOVA s'écrit comme ci-dessous :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	n_A	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_R^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	n_B	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_R^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Résiduelle	sc_R	n_R	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	n_{TOT}			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$
$$\text{contre}$$
$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

$F_A = \dfrac{S_A^2}{S_R^2} \sim$ Fisher de degré de liberté $[I-1, (I-1)(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0'' : \sigma_B^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1'' : \sigma_B^2 \neq 0.$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

$\mathcal{H}_0''$ sous entend l'absence d'effet du traitement A sur Y

$F_A = \dfrac{S_B^2}{S_R^2} \sim$ Fisher de degré de liberté $[J-1,(I-1)(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0''$ au risque α si la p-value $< \alpha$.

Estimateur des paramètres du modèle : Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\sigma_B^2}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \sigma_B^2, \sigma^2$ du modèle sont donnés par les formules suivantes :

$$\begin{aligned}
&\widehat{\mu} = Y_{\bullet,\bullet} = \bar{Y}, \widehat{\alpha_i} = Y_{i,\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I,\\
&\widehat{\sigma_B^2} = \frac{1}{I}\left(S_B^2 - S_R^2\right),\\
&\widehat{\sigma^2} = \frac{SC_R}{(I-1)(J-1)} = S_R^2,
\end{aligned}$$

où $S_B^2 = \frac{SC_B}{n_B}$ et $S_R^2 = \frac{SC_R}{n_R}$.

Ce sont des estimateurs sans biais.

3.2 Modèle à effet mixte à deux facteurs et avec répétitions

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Les β_j représentent un échantillon de taille J prélevé dans une population importante. Nous admettrons que les effets des B_j, les β_j, sont distribués suivant une loi normale centrée de variance σ_B^2. Pour chacun des couples de modalités $\left(A_i, B_j\right)$ nous effectuons $K \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle, dit restreint :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \epsilon_{i,j,k}, \quad i = 1\ldots I, j = 1\ldots J, k = 1\ldots K,$$

$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0,$$

$$\sum_{i=1}^{I} (\alpha\beta)_{i,j} = 0, \forall j \in \{1,\ldots,J\},$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j\right)$ lors du k-ème essai. Nous supposons que

$$\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J,$$
$$\mathcal{L}\left((\alpha\beta)_{i,j}\right) = \mathcal{N}\left(0, \sigma_{AB}^2\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,$$
$$\beta_i \perp\!\!\!\perp \beta_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant J,$$
$$\beta_i \perp\!\!\!\perp (\alpha\beta)_{j,k} \text{ si } 1 \leqslant j \leqslant I \text{ et } 1 \leqslant i, k \leqslant J.$$

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\text{et } \epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec}$$
$$1 \leqslant i, l \leqslant I, 1 \leqslant j, m \leqslant J \text{ et } 1 \leqslant k, n \leqslant K,$$

ainsi que l'indépendance des effets aléatoires et des erreurs :

$$\beta_i \perp\!\!\!\perp \epsilon_{j,k,l} \text{ si } 1 \leqslant j \leqslant I, 1 \leqslant i, k \leqslant J, \text{ et } 1 \leqslant l \leqslant K,$$
$$(\alpha\beta)_{i,j} \perp\!\!\!\perp \epsilon_{k,l,m} \text{ si } 1 \leqslant i, k \leqslant I, 1 \leqslant j, l \leqslant J, \text{ et } 1 \leqslant m \leqslant K.$$

Dans un modèle mixte restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici les $(\alpha\beta)_{i,j}$, ne sont pas mutuellement indépendants à cause des contraintes portant sur leur somme, $\sum_{i=1}^{I}(\alpha\beta)_{i,j} = 0, \forall j \in \{1, \ldots, J\}$. Ils le sont par contre dès que nous ne les considérons pas tous en même temps.

Certains logiciels travaillent par défaut avec un modèle non-restreint dont la définition est proche mais pas identique à celle du modèle précédent.

Nous introduisons le modèle, dit non restreint :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \epsilon_{i,j,k}, \quad i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K$$
$$\text{avec la contrainte supplémentaire } \sum_{i=1}^{I} \alpha_i = 0,$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j\right)$. Nous supposons que

$$\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J$$
$$\mathcal{L}\left((\alpha\beta)_{i,j}\right) = \mathcal{N}\left(0, \sigma_{AB}^2\right), \quad \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J$$

ainsi que l'indépendance des effets aléatoires :

$$\beta_i \perp\!\!\!\perp \beta_j \text{ si } i \neq j \text{ et } 1 \leqslant i, j \leqslant J,$$
$$(\alpha\beta)_{i,j} \perp\!\!\!\perp (\alpha\beta)_{k,l} \text{ si } (i,j) \neq (k,l) \text{ avec } 1 \leqslant i, k \leqslant I \text{ et } 1 \leqslant j, l \leqslant J,$$
$$\beta_i \perp\!\!\!\perp (\alpha\beta)_{j,k} \text{ si } 1 \leqslant j \leqslant I \text{ et } 1 \leqslant i, k \leqslant J.$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall(i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0,\sigma^2\right),$$
$$\text{et } \epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec}$$
$$1 \leqslant i,l \leqslant I, 1 \leqslant j,m \leqslant J \text{ et } 1 \leqslant k,n \leqslant K,$$

ainsi que l'indépendance des effets aléatoires et des erreurs :

$$\beta_i \perp\!\!\!\perp \epsilon_{j,k,l} \text{ si } 1 \leqslant j \leqslant I, 1 \leqslant i,k \leqslant J, \text{ et } 1 \leqslant l \leqslant K,$$
$$(\alpha\beta)_{i,j} \perp\!\!\!\perp \epsilon_{k,l,m} \text{ si } 1 \leqslant i,k \leqslant I, 1 \leqslant j,l \leqslant J, \text{ et } 1 \leqslant m \leqslant K.$$

Dans un modèle mixte non restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici les $(\alpha\beta)_{i,j}$, sont mutuellement indépendants.

Il n'existe pas consensus sur une raison statistique qui permettrait de privilégier plutôt l'une ou l'autre de ces deux approches. Nous utiliserons toujours des modèles restreints. La relation fondamentale de l'ANOVA sécrit :

$$SC_{TOT} = SC_A + SC_B + SC_{AB} + SC_R$$

et son tableau :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I-1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_{AB}^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J-1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_R^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Interaction	sc_{AB}	$n_{AB} = (I-1)(J-1)$	$s_{AB}^2 = \frac{sc_{AB}}{n_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_R^2}$	$\mathcal{H}_0'''$ ou $\mathcal{H}_1'''$
Résiduelle	sc_R	$n_R = IJ(K-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJK-1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$
$$\textit{contre}$$
$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1,2,\ldots,I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

$F_A = \dfrac{S_A^2}{S_{AB}^2} \sim$ Fisher de degré de liberté $[I-1,(I-1)(J-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0'' : \sigma_B^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1'' : \sigma_B^2 \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet du traitement B sur Y

$F_B = \dfrac{S_B^2}{S_R^2} \sim$ Fisher de degré de liberté $[J-1, IJ(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0''$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0''' : \sigma_{AB}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1''' : \sigma_{AB}^2 \neq 0.$$

$\mathcal{H}_0'''$ sous entend l'absence d'effet du traitement B sur Y

$F_{AB} = \dfrac{S_{AB}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1), IJ(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0'''$ au risque α si la p-value $< \alpha$.

Estimateur des paramètres du modèle : Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\sigma_B^2}, \widehat{\sigma_{AB}^2}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \sigma_B^2, \sigma_{AB}^2, \sigma^2$du modèle sont donnés par les formules suivantes :

$$\begin{aligned}
\widehat{\mu} &= Y_{\bullet,\bullet} = \bar{Y}, \\
\widehat{\alpha_i} &= Y_{i,\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I, \\
\widehat{\sigma_B^2} &= \frac{1}{IK}\left(S_B^2 - S_R^2\right) \\
\widehat{\sigma_{AB}^2} &= \frac{1}{K}\left(S_{AB}^2 - S_R^2\right) \\
\widehat{\sigma^2} &= \frac{SC_R}{(I-1)(J-1)} = S_R^2,
\end{aligned}$$

où $S_B^2 = \frac{SC_B}{n_B}, S_{AB}^2 = \frac{SC_{AB}}{n_{AB}}$ et $S_R^2 = \frac{SC_R}{n_R}$.
Ce sont des estimateurs sans biais.

3.2.1 Modèle à effet mixte à deux facteurs emboités (ici il faut une répétition)

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Les β_j représentent un échantillon de taille J prélevé dans une population importante. Nous admettrons que les effets des $B_{j(i)}$, les $\beta_{j(i)}$, sont distribués suivant une loi normale centrée de variance $\sigma^2_{B|A}$. Pour chacun des couples de modalités $\left(A_i, B_{j(i)}\right)$ nous effectuons $K \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{i,j,k}, \quad i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K,$$
$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0,$$

Nous supposons que

$$\mathcal{L}\left(\beta_{j(i)}\right) = \mathcal{N}\left(0, \sigma^2_{B|A}\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J$$

ainsi que l'indépendance des effets aléatoires :

$$\beta_{j(i)} \perp\!\!\!\perp \beta_{l(k)} \text{ si } (i,j) \neq (k,l) \text{ avec } 1 \leqslant i,k \leqslant I \text{ et } 1 \leqslant j,l \leqslant J.$$

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right)$$
$$\text{et } \epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec}$$
$$1 \leqslant i,l \leqslant I, 1 \leqslant j,m \leqslant J \text{ et } 1 \leqslant k,n \leqslant K,$$

ainsi que l'indépendance des effets aléatoires et des erreurs :

$$\beta_{j(i)} \perp\!\!\!\perp \epsilon_{k,l,m} \text{ si } 1 \leqslant i,k \leqslant I, 1 \leqslant j,l \leqslant J, \text{ et } 1 \leqslant m \leqslant K.$$

la relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_{B|A} + SC_R$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I - 1$	$s^2_A = \frac{sc_A}{n_A}$	$f_A = \frac{s^2_A}{s^2_{B\|A}}$	$\mathcal{H}'_0$ ou $\mathcal{H}'_1$
Facteur B dans A	$sc_{B\|A}$	$n_{B\|A} = I(J-1)$	$s^2_{B\|A} = \frac{sc_{B\|A}}{n_{B\|A}}$	$f_{B\|A} = \frac{s^2_{B\|A}}{s^2_R}$	$\mathcal{H}''_0$ ou $\mathcal{H}''_1$
Résiduelle	sc_R	$n_R = IJ(K-1)$	$s^2_R = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJK - 1$			

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

où les hypothèses sont les suivantes

$$\mathcal{H}'_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$
$$\text{contre}$$
$$\mathcal{H}'_1 : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}'_0$ sous entend l'absence d'effet du traitement A sur Y

Sous $\mathcal{H}'_0$ $F_A = \dfrac{S_A^2}{S_{B|A}^2} \sim$ Fisher de degré de liberté $[I-1, I(J-1)]$

Règle de décision : On rejette $\mathcal{H}'_0$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}''_0 : \sigma^2_{B|A} = 0$$
$$\text{contre}$$
$$\mathcal{H}''_1 : \sigma^2_{B|A} \neq 0.$$

$\mathcal{H}''_0$ sous entend l'absence d'effet du traitement B dans A

Sous $\mathcal{H}''_0$ $F_{B|A} = \dfrac{S_{B|A}^2}{S_R^2} \sim$ Fisher de degré de liberté $[I-1, I(J-1)]$

Règle de décision : On rejette $\mathcal{H}'_0$ au risque α si la p-value $< \alpha$.

Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\sigma^2_{B|A}}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \sigma^2_{B|A}, \sigma^2$ du modèle sont donnés par les formules suivantes :

$$\begin{aligned}
\widehat{\mu} &= Y_{\bullet,\bullet,\bullet} = \bar{Y}, \\
\widehat{\alpha_i} &= Y_{i,\bullet,\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I, \\
\widehat{\sigma^2_{B|A}} &= \frac{1}{K}\left(S^2_{B|A} - S^2_R\right), \\
\widehat{\sigma^2} &= \frac{SC_R}{(I-1)(J-1)} = S^2_R,
\end{aligned}$$

où $S^2_{B|A} = \frac{SC_{B|A}}{n_{B|A}}$ et $S^2_R = \frac{SC_R}{n_R}$. Ce sont des estimateurs sans biais.

3.3 Modèle à effet mixte à trois facteurs et sans répétitions

Il existe deux possibilités : soit deux facteurs sont fixes et un est aléatoire, soit un facteur est fixe et deux sont aléatoires.

3.3.1 Premier cas : Deux facteurs sont à effets fixes et un facteur est à effets aléatoires.

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Un facteur contrôlé B se présente sous J modalités, chacune d'entre elles étant notée B_j. Les γ_k représentent un échantillon de taille K prélevé dans une population importante. Nous admettrons que les effets des C_k, les γ_k, sont distribués suivant une loi normale centrée de variance σ_C^2. Pour chacun des couples de modalités $\left(A_i, B_j, C_k\right)$ nous effectuons une mesure d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + \epsilon_{i,j,k},$$
$$i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K$$
$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0, \sum_{j=1}^{J} \beta_j = 0,$$
$$\sum_{i=1}^{I} (\alpha\beta)_{i,j} = 0, \forall j \in \{1, \ldots, J\}, \sum_{j=1}^{J} (\alpha\beta)_{i,j} = 0, \forall i \in \{1, \ldots, I\},$$
$$\sum_{i=1}^{I} (\alpha\gamma)_{i,k} = 0, \forall k \in \{1, \ldots, K\} \text{ et } \sum_{j=1}^{J} (\beta\gamma)_{j,k} = 0, \forall k \in \{1, \ldots, K\},$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$. Nous supposons que :

$$\mathcal{L}\left(\gamma_k\right) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\gamma)_{i,k}\right) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \quad \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$

ainsi que l'indépendance des effets aléatoires γ_k.

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec } 1 \leqslant i, l \leqslant I, 1 \leqslant j, m \leqslant J \text{ et } 1 \leqslant k, n \leqslant K,$$

ainsi que l'indépendance des effets aléatoires γ_k, $(\alpha\gamma)_{i,k}$ et $(\beta\gamma)_{j,k}$ et des erreurs $\epsilon_{i,j,k}$. Dans un modèle mixte restreint, les effets aléatoires croisant des facteurs à effets fixes

rafmoqui@yahoo.fr ~ 96877413

et à effets aléatoires, ici aussi bien les $(\alpha\gamma)_{i,k}$ que les $(\beta\gamma)_{j,k}$, ne sont pas mutuellement indépendants à cause des contraintes portant sur leur somme, $\sum_{i=1}^{I}(\alpha\gamma)_{i,k} = 0, \forall k \in \{1,\ldots,K\}$ et $\sum_{j=1}^{J}(\beta\gamma)_{j,k} = 0, \forall k \in \{1,\ldots,K\}$. Ils le sont par contre dès que nous ne les considérons pas tous en même temps.

Certains logiciels travaillent par défaut avec un modèle non-restreint dont la définition est proche mais pas identique à celle du modèle précédent.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + \epsilon_{i,j,k},$$
$$i = 1\ldots I, j = 1\ldots J, k = 1\ldots K$$

$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I}\alpha_i = 0, \sum_{j=1}^{J}\beta_j = 0,$$

$$\sum_{i=1}^{I}(\alpha\beta)_{i,j} = 0, \forall j \in \{1,\ldots,J\}, \sum_{j=1}^{J}(\alpha\beta)_{i,j} = 0, \forall i \in \{1,\ldots,I\},$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$. Nous supposons que :

$$\mathcal{L}(\gamma_k) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}((\alpha\gamma)_{i,k}) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \quad \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$

ainsi que l'indépendance des effets aléatoires γ_k, $(\alpha\gamma)_{i,k}$ et $(\beta\gamma)_{j,k}$.

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec } 1 \leqslant i,l \leqslant I, 1 \leqslant j,m \leqslant J \text{ et } 1 \leqslant k,n \leqslant K,$$

ainsi que l'indépendance des effets aléatoires γ_k, $(\alpha\gamma)_{i,k}$ et $(\beta\gamma)_{j,k}$ et des erreurs $\epsilon_{i,j,k}$. Dans un modèle mixte non restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici les $(\alpha\gamma)_{i,k}$ et les $(\beta\gamma)_{j,k}$, sont mutuellement indépendants.

Il n'existe pas consensus sur une raison statistique qui permettrait de privilégier plutôt l'une ou l'autre de ces deux approches. la relation fondamentale de l'ANOVA :

$$sc_{TOT} = sc_A + sc_B + sc_C + sc_{AB} + sc_{AC} + sc_{BC} + sc_R$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I-1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_{AC}^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J-1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_{BC}^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Facteur C	sc_C	$n_C = K-1$	$s_C^2 = \frac{sc_C}{n_C}$	$f_C = \frac{s_C^2}{s_R^2}$	$\mathcal{H}_0'''$ ou $\mathcal{H}_1'''$
Interaction AB	sc_{AB}	$n_{AB} = (I-1)(J-1)$	$s_{AB}^2 = \frac{sc_{AB}}{n_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_R^2}$	$\mathcal{H}_0^{(4)}$ ou $\mathcal{H}_1^{(4)}$
Interaction AC	sc_{AC}	$n_{AC} = (I-1)(K-1)$	$s_{AC}^2 = \frac{sc_{AC}}{n_{AC}}$	$f_{AC} = \frac{s_{AC}^2}{s_R^2}$	$\mathcal{H}_0^{(5)}$ ou $\mathcal{H}_1^{(5)}$
Interaction BC	sc_{BC}	$n_{BC} = (J-1)(K-1)$	$s_{BC}^2 = \frac{sc_{BC}}{n_{BC}}$	$f_{BC} = \frac{s_{BC}^2}{s_R^2}$	$\mathcal{H}_0^{(6)}$ ou $\mathcal{H}_1^{(6)}$
Résiduelle	sc_R	$n_R = (I-1)(I-1)(K-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	scToT	$n_{TOT} = IJK-1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$
contre
$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

Sous $\mathcal{H}_0'$ $F_A = \dfrac{S_A^2}{S_{AC}^2} \sim$ Fisher de degré de liberté $[I-1, (I-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0'' : \beta_1 = \beta_2 = \cdots = \beta_J = 0$$
contre
$$\mathcal{H}_1'' : \text{ Il existe } j_0 \in \{1, 2, \ldots, J\} \text{ tel que } \beta_{j_0} \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet du traitement C sur Y

Sous $\mathcal{H}_0''$ $F_C = \dfrac{S_B^2}{S_{BC}^2} \sim$ Fisher de degré de liberté $[J-1, (J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0''' : \sigma_C^2 = 0$$
contre
$$\mathcal{H}_1''' : \sigma_C^2 \neq 0.$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

$\mathcal{H}_0'''$ sous entend l'absence d'effet du traitement C sur Y

Sous $\mathcal{H}_0'''$ $F_C = \dfrac{S_C^2}{S_R^2} \sim$ Fisher de degré de liberté $[K-1,(I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0'''$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0^{(4)} : (\alpha\beta)_{1,1} = (\alpha\beta)_{1,2} = \cdots = (\alpha\beta)_{1,J} = (\alpha\beta)_{2,1} = \cdots = (\alpha\beta)_{I,J} = 0$$

contre

$$\mathcal{H}_1^{(4)} : \text{ Il existe } (i_0, j_0) \in \{1,2,\ldots,I\} \times \{1,2,\ldots,J\} \text{ tel que } (\alpha\beta)_{i_0,j_0} \neq 0.$$

$\mathcal{H}_0^{(4)}$ sous entend l'absence de l'interaction des traitement A et B sur Y

Sous $\mathcal{H}_0^{(4)}$ $F_{AB} = \dfrac{S_{AB}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1),(I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(4)}$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0^{(5)} : \sigma_{AC}^2 = 0$$

contre

$$\mathcal{H}_1^{(5)} : \sigma_{AC}^2 \neq 0.$$

$\mathcal{H}_0^{(5)}$ sous entend l'absence de l'interaction des traitement A et C sur Y

Sous $\mathcal{H}_0^{(5)}$ $F_{AC} = \dfrac{S_{AC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(K-1),(I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(5)}$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}_0^{(6)} : \sigma_{BC}^2 = 0$$

contre

$$\mathcal{H}_1^{(6)} : \sigma_{BC}^2 \neq 0.$$

$\mathcal{H}_0^{(6)}$ sous entend l'absence de l'interaction des traitement C et B sur Y

Sous $\mathcal{H}_0^{(6)}$ $F_{BC} = \dfrac{S_{BC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(K-1),(I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(6)}$ au risque α si la p-value $< \alpha$.

Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\beta_1}, \ldots, \widehat{\beta_J}, \widehat{\sigma_C^2}, \widehat{(\alpha\beta)_{1,1}}, \ldots, \widehat{(\alpha\beta)_{I,J}}, \widehat{\sigma_{AC}^2}, \widehat{\sigma_{BC}^2}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \beta_1, \ldots, \beta_J, \sigma_C^2, (\alpha\beta)_{1,1}, \ldots, (\alpha\beta)_{I,J}, \sigma_{AC}^2, \sigma_{BC}^2, \sigma^2$du modèle sont donnés par les formules suivantes :

$$\widehat{\mu} = Y_{\bullet,\bullet,\bullet} = \bar{Y}$$
$$\widehat{\alpha_i} = Y_{i,\bullet\bullet\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I,$$
$$\widehat{\beta_j} = Y_{\bullet,j,\bullet} - \widehat{\mu}, 1 \leqslant j \leqslant J,$$
$$\widehat{\sigma^2_C} = \frac{1}{IJ}\left(S^2_C - S^2_R\right)$$
$$\widehat{(\alpha\beta)_{i,j}} = Y_{i,j,\bullet} - Y_{i,\bullet,\bullet} - Y_{\bullet,j,\bullet} + \widehat{\mu}, \qquad 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,$$
$$\widehat{\sigma^2_{AC}} = \frac{1}{J}\left(S^2_{AC} - S^2_R\right)$$
$$\widehat{\sigma^2_{BC}} = \frac{1}{I}\left(S^2_{BC} - S^2_R\right)$$
$$\widehat{\sigma^2} = \frac{SC_R}{(I-1)(J-1)(K-1)} = S^2_R.$$

où $S^2_C = \frac{SC_C}{n_C}, S^2_{AC} = \frac{SC_{AC}}{n_{AC}}, S^2_{BC} = \frac{SC_{BC}}{n_{BC}}$ et $S^2_R = \frac{SC_R}{n_R}$. Ce sont des estimateurs sans biais.

3.3.2 Deuxième cas : Un facteur est à effets fixes et deux facteurs sont à effets aléatoires.

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Les β_j représentent un échantillon de taille J prélevé dans une population importante. Nous admettrons que les effets des B_j, les β_j, sont distribués suivant une loi normale centrée de variance σ^2_B. Les γ_k représentent un échantillon de taille K prélevé dans une population importante. Nous admettrons que les effets des C_k, les γ_k, sont distribués suivant une loi normale centrée de variance σ^2_C. Pour chacun des couples de modalités $\left(A_i, B_j, C_k\right)$ nous effectuons une mesure d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + \epsilon_{i,j,k},$$
$$i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K$$

avec les contraintes supplémentaires $\displaystyle\sum_{i=1}^{I} \alpha_i = 0,$

$$\sum_{i=1}^{I} (\alpha\beta)_{i,j} = 0, \forall j \in \{1, \ldots, J\} \text{ et } \sum_{i=1}^{I} (\alpha\gamma)_{i,k} = 0, \forall k \in \{1, \ldots, K\},$$

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$. Nous supposons que :

$$\begin{gathered}
\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J, \\
\mathcal{L}\left(\gamma_k\right) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K, \\
\mathcal{L}\left((\alpha\beta)_{i,j}\right) = \mathcal{N}\left(0, \sigma_{AB}^2\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, \\
\mathcal{L}\left((\alpha\gamma)_{i,k}\right) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K, \\
\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,
\end{gathered}$$

ainsi que l'indépendance des effets aléatoires β_j, γ_k, et $(\beta\gamma)_{j,k}$.

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\begin{gathered}
\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right), \\
\epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec } 1 \leqslant i, l \leqslant I, 1 \leqslant j, m \leqslant J \text{ et } 1 \leqslant k, n \leqslant K,
\end{gathered}$$

ainsi que l'indépendance des effets aléatoires $\beta_j, \gamma_k, (\alpha\beta)_{i,j}, (\alpha\gamma)_{i,k}$ et $(\beta\gamma)_{j,k}$ et des erreurs $\epsilon_{i,j,k}$.

Dans un modèle mixte restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici aussi bien les $(\alpha\beta)_{i,j}$ que les $(\alpha\gamma)_{i,k}$, ne sont pas mutuellement indépendants à cause des contraintes portant sur leur somme, $\sum_{i=1}^{I} (\alpha\beta)_{i,j} = 0, \forall j \in \{1, \ldots, J\}$ et $\sum_{i=1}^{I} (\alpha\gamma)_{i,k} = 0, \forall k \in \{1, \ldots, K\}$. Ils le sont par contre dès que nous ne les considérons pas tous en même temps.

Certains logiciels travaillent par défaut avec un modèle non-restreint dont la définition est proche mais pas identique à celle du modèle précédent.

Nous introduisons le modèle :

$$\begin{gathered}
Y_{i,j,k} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + \epsilon_{i,j,k} \\
i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K
\end{gathered}$$

avec les contraintes supplémentaires $\sum_{i=1}^{I} \alpha_i = 0$,

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$. Nous supposons que :

$$\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J,$$
$$\mathcal{L}\left(\gamma_k\right) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\beta)_{i,j}\right) = \mathcal{N}\left(0, \sigma_{AB}^2\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,$$
$$\mathcal{L}\left((\alpha\gamma)_{i,k}\right) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$

ainsi que l'indépendance des effets aléatoires $\beta_j, \gamma_k, (\alpha\beta)_{i,j}, (\alpha\gamma)_{i,k}$ et $(\beta\gamma)_{j,k}$. Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, \mathcal{L}\left(\epsilon_{i,j,k}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\epsilon_{i,j,k} \perp\!\!\!\perp \epsilon_{l,m,n} \text{ si } (i,j,k) \neq (l,m,n) \text{ avec } 1 \leqslant i,l \leqslant I, 1 \leqslant j,m \leqslant J \text{ et } 1 \leqslant k,n \leqslant K,$$

ainsi que l'indépendance des effets aléatoires $\beta_j, \gamma_k, (\alpha\beta)_{i,j}, (\alpha\gamma)_{i,k}$ et $(\beta\gamma)_{j,k}$ et des erreurs $\epsilon_{i,j,k}$.

Dans un modèle mixte non restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici les $(\alpha\beta)_{i,j}$ et les $(\alpha\gamma)_{i,k}$, sont mutuellement indépendants.

Il n'existe pas consensus sur une raison statistique qui permettrait de privilégier plutôt l'une ou l'autre de ces deux approches. la relation fondamentale de l'ANOVA :

$$sc_{TOT} = sc_A + sc_B + sc_C + sc_{AB} + sc_{AC} + sc_{BC} + sc_R.$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I-1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_{AB}^2+s_{AC}^2-s_R^2}$	$\mathcal{H}_0'$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J-1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_{BC}^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Facteur C	sc_C	$n_C = K-1$	$s_C^2 = \frac{sc_C}{n_C}$	$f_C = \frac{s_C^2}{s_{BC}^2}$	$\mathcal{H}_0'''$ ou $\mathcal{H}_1'''$
Interaction AB	sc_{AB}	$n_{AB} = (I-1)(J-1)$	$s_{AB}^2 = \frac{sc_{AB}}{n_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_R^2}$	$\mathcal{H}_0^{(4)}$ ou $\mathcal{H}_1^{(4)}$
Interaction AC	sc_{AC}	$n_{AC} = (I-1)(K-1)$	$s_{AC}^2 = \frac{sc_{AC}}{n_{AC}}$	$f_{AC} = \frac{s_{AC}^2}{s_R^2}$	$\mathcal{H}_0^{(5)}$ ou $\mathcal{H}_1^{(5)}$
Interaction BC	sc_{BC}	$n_{BC} = (J-1)(K-1)$	$s_{BC}^2 = \frac{sc_{BC}}{n_{BC}}$	$f_{BC} = \frac{s_{BC}^2}{s_R^2}$	$\mathcal{H}_0^{(6)}$ ou $\mathcal{H}_1^{(6)}$
Résiduelle	sc_R	$n_R = (I-1)(J-1)(K-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJK-1$			

où les hypothèses sont les suivantes

$$\mathcal{H}'_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$
$$\text{contre}$$
$$\mathcal{H}'_1 : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}'_0$ sous entend l'absence d'effet du traitement A sur Y

Sous $\mathcal{H}'_0$ $F_A = \dfrac{S_A^2}{S_{AB}^2 + S_{AC}^2 - S_R^2} \sim$ Fisher de degré de liberté $[I-1, n']$

$$n' = \frac{\left(s_{AB}^2 + s_{AC}^2 - s_R^2\right)^2}{\frac{\left(s_{AB}^2\right)^2}{(I-1)(J-1)} + \frac{\left(s_{AC}^2\right)^2}{(I-1)(K-1)} + \frac{\left(s_R^2\right)^2}{(I-1)(J-1)(K-1)}}.$$

Règle de décision : On rejette $\mathcal{H}'_0$ au risque α si la p-value $< \alpha$.

$$\mathcal{H}''_0 : \sigma_B^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}''_1 : \sigma_B^2 \neq 0.$$

$\mathcal{H}''_0$ sous entend l'absence d'effet du traitement B sur Y

Sous $\mathcal{H}''_0$ $F_B = \dfrac{S_B^2}{S_{BC}^2} \sim$ Fisher de degré de liberté $[J-1, (J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}'_0$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_B est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}'''_0 : \sigma_C^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}'''_1 : \sigma_C^2 \neq 0.$$

$\mathcal{H}'''_0$ sous entend l'absence d'effet du traitement C sur Y

Sous $\mathcal{H}'''_0$ $F_C = \dfrac{S_C^2}{S_{BC}^2} \sim$ Fisher de degré de liberté $[K-1, (J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}'''_0$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_C est supérieure ou égale à la valeur critique issue de la table.

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

$$\mathcal{H}_0^{(4)} : \sigma_{AB}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(4)} : \sigma_{AB}^2 \neq 0.$$

$\mathcal{H}_0^{(4)}$ sous entend l'absence d'effet de l'interaction des traitement A et B sur Y

Sous $\mathcal{H}_0^{(4)}$ $F_{AB} = \dfrac{S_{AB}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1),(I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(4)}$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_{AB} est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}_0^{(5)} : \sigma_{AC}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(5)} : \sigma_{AC}^2 \neq 0.$$

$\mathcal{H}_0^{(5)}$ sous entend l'absence d'effet de l'interaction des traitement A et C sur Y

Sous $\mathcal{H}_0^{(5)}$ $F_{AC} = \dfrac{S_{AC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(K-1),(I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(5)}$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_{AC} est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}_0^{(6)} : \sigma_{BC}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(6)} : \sigma_{BC}^2 \neq 0.$$

$\mathcal{H}_0^{(6)}$ sous entend l'absence d'effet de l'interaction des traitement B et C sur Y

Sous $\mathcal{H}_0^{(6)}$ $F_{BC} = \dfrac{S_{BC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(J-1)(K-1),(I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(6)}$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_{BC} est supérieure ou égale à la valeur critique issue de la table.

Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\sigma_B^2}, \widehat{\sigma_C^2}, \widehat{\sigma_{AB}^2}, \widehat{\sigma_{AC}^2}, \widehat{\sigma_{BC}^2}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I$, $\sigma_B^2, \sigma_C^2, \sigma_{AB}^2, \sigma_{AC}^2, \sigma_{BC}^2, \sigma^2$ du modèle sont donnés par les formules suivantes :

rafmoqui@yahoo.fr ~ 96877413

$$
\begin{aligned}
&\widehat{\mu} = Y_{\bullet,\bullet,\bullet} = \bar{y}\\
&\widehat{\alpha_i} = Y_{i,\bullet\bullet\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I,\\
&\widehat{\sigma_B^2} = \frac{1}{IK}\left(S_B^2 - S_{BC}^2\right),\\
&\widehat{\sigma_C^2} = \frac{1}{IJ}\left(S_C^2 - S_{BC}^2\right),\\
&\widehat{\sigma_{AB}^2} = \frac{1}{K}\left(S_{AB}^2 - S_R^2\right),\\
&\widehat{\sigma_{AC}^2} = \frac{1}{J}\left(S_{AC}^2 - S_R^2\right),\\
&\widehat{\sigma_{BC}^2} = \frac{1}{I}\left(S_{BC}^2 - S_R^2\right),\\
&\widehat{\sigma^2} = \frac{SC_R}{(I-1)(J-1)(K-1)} = S_R^2.
\end{aligned}
$$

où $S_B^2 = \frac{SC_B}{n_B}, S_C^2 = \frac{SC_C}{n_C}, S_{AB}^2 = \frac{SC_{AB}}{n_{AB}}, S_{AC}^2 = \frac{SC_{AC}}{n_{AC}}, S_{BC}^2 = \frac{SC_{BC}}{n_{BC}}$ et $S_R^2 = \frac{SC_R}{n_R}$.
Ce sont des estimateurs sans biais.

3.4 Modèle à effet mixte à trois facteurs et avec répétitions

Il existe deux possibilités : soit deux facteurs sont fixes et un est aléatoire, soit un facteur est fixe et deux sont aléatoires.

3.4.1 Premier cas : Deux facteurs sont à effets fixes et un facteur est à effets aléatoires.

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Un facteur contrôlé B se présente sous J modalités, chacune d'entre elles étant notée B_j. Les γ_k représentent un échantillon de taille K prélevé dans une population importante. Nous admettrons que les effets des C_k, les γ_k, sont distribués suivant une loi normale centrée de variance σ_C^2. Pour chacun des couples de modalités $\left(A_i, B_j, C_k\right)$ nous effectuons $L \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K \times L$ le nombre total de mesures ayant été effectuées.

Nous introduisons le modèle :

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

$$Y_{i,j,k,l} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + (\alpha\beta\gamma)_{i,j,k} + \epsilon_{i,j,k,l},$$
$$i = 1\ldots I, j = 1\ldots J, k = 1\ldots K, l = 1\ldots L,$$
$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I}\alpha_i = 0, \sum_{j=1}^{J}\beta_j = 0$$
$$\sum_{i=1}^{I}(\alpha\beta)_{i,j} = 0, \forall j \in \{1,\ldots,J\}, \sum_{j=1}^{J}(\alpha\beta)_{i,j} = 0, \forall i \in \{1,\ldots,I\}$$
$$\sum_{i=1}^{I}(\alpha\gamma)_{i,k} = 0, \forall k \in \{1,\ldots,K\} \text{ et } \sum_{j=1}^{J}(\beta\gamma)_{j,k} = 0, \forall k \in \{1,\ldots,K\}$$
$$\sum_{i=1}^{I}(\alpha\beta\gamma)_{i,j,k} = 0, \forall (j,k) \in \{1,\ldots,J\} \times \{1,\ldots,K\}$$
$$\sum_{j=1}^{J}(\alpha\beta\gamma)_{i,j,k} = 0, \forall (i,k) \in \{1,\ldots,I\} \times \{1,\ldots,K\}$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$ lors du l-ème essai. Nous supposons que :

$$\mathcal{L}(\gamma_k) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}((\alpha\gamma)_{i,k}) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\beta\gamma)_{i,j,k}\right) = \mathcal{N}\left(0, \sigma_{ABC}^2\right), \forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$

ainsi que l'indépendance des effets aléatoires γ_k.

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k,l), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, 1 \leqslant l \leqslant L, \mathcal{L}\left(\epsilon_{i,j,k,l}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\epsilon_{i,j,k,l} \perp\!\!\!\perp \epsilon_{m,n,o,p} \text{ si } (i,j,k,l) \neq (m,n,o,p)$$
$$\text{avec } 1 \leqslant i, m \leqslant I, 1 \leqslant j, n \leqslant J, 1 \leqslant k, o \leqslant K \text{ et } 1 \leqslant l, p \leqslant L$$

ainsi que l'indépendance des effets aléatoires $\gamma_k, (\alpha\gamma)_{i,k}, (\beta\gamma)_{j,k}$ et $(\alpha\beta\gamma)_{i,j,k}$ et des erreurs $\epsilon_{i,j,k,l}$.

Dans un modèle mixte restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici aussi bien les $(\alpha\gamma)_{i,k}$ que les $(\beta\gamma)_{j,k}$ et que les $(\alpha\beta\gamma)_{i,j,k}$, ne sont pas mutuellement indépendants à cause des contraintes portant sur leur

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

somme, $\sum_{i=1}^{I}(\alpha\gamma)_{i,k} = 0, \forall k \in \{1,\ldots,K\}$ et $\sum_{j=1}^{J}(\beta\gamma)_{j,k} = 0, \forall k \in \{1,\ldots,K\}, \sum_{i=1}^{I}(\alpha\beta\gamma)_{i,j,k} = 0, \forall (j,k) \in \{1,\ldots,J\} \times \{1,\ldots,K\}, \sum_{j=1}^{J}(\alpha\beta\gamma)_{i,j,k} = 0, \forall (i,k) \in \{1,\ldots,I\} \times \{1,\ldots,K\}$. Ils le sont par contre dès que nous ne les considérons pas tous en même temps.

Certains logiciels travaillent par défaut avec un modèle non-restreint dont la définition est proche mais pas identique à celle du modèle précédent.

Nous introduisons le modèle :

$$Y_{i,j,k,l} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + (\alpha\beta\gamma)_{i,j,k} + \epsilon_{i,j,k,l}$$
$$i = 1 \ldots I, j = 1 \ldots J, k = 1 \ldots K, l = 1 \ldots L$$

avec les contraintes supplémentaires $\sum_{i=1}^{I} \alpha_i = 0, \sum_{j=1}^{J} \beta_j = 0$,

$$\sum_{i=1}^{I}(\alpha\beta)_{i,j} = 0, \forall j \in \{1,\ldots,J\}, \sum_{j=1}^{J}(\alpha\beta)_{i,j} = 0, \forall i \in \{1,\ldots,I\}$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$. Nous suppo- sons que :

$$\mathcal{L}(\gamma_k) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}((\alpha\gamma)_{i,k}) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\beta\gamma)_{i,j,k}\right) = \mathcal{N}\left(0, \sigma_{ABC}^2\right), \forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$

ainsi que l'indépendance des effets aléatoires $\gamma_k, (\alpha\gamma)_{i,k}, (\beta\gamma)_{j,k}, (\alpha\beta\gamma)_{i,j,k}$. Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k,l), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, 1 \leqslant l \leqslant L, \mathcal{L}\left(\epsilon_{i,j,k,l}\right) = \mathcal{N}\left(0, \sigma^2\right),$$
$$\epsilon_{i,j,k,l} \perp\!\!\!\perp \epsilon_{m,n,o,p} \text{ si } (i,j,k,l) \neq (m,n,o,p)$$
$$\text{avec } 1 \leqslant i, m \leqslant I, 1 \leqslant j, n \leqslant J, 1 \leqslant k, o \leqslant K \text{ et } 1 \leqslant l, p \leqslant L$$

ainsi que l'indépendance des effets aléatoires $\gamma_k, (\alpha\gamma)_{i,k}, (\beta\gamma)_{j,k}$ et $(\alpha\beta\gamma)_{i,j,k}$ et des erreurs $\epsilon_{i,j,k,l}$.

Dans un modèle mixte non restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici les $(\alpha\gamma)_{i,k}$, les $(\beta\gamma)_{j,k}$ et les $(\alpha\beta\gamma)_{i,j,k}$, sont mutuellement indépendants.

Il n'existe pas consensus sur une raison statistique qui permettrait de privilégier plutôt l'une ou l'autre de ces deux approches. Nous utiliserons toujours des modèles restreints.

$$sc_{TOT} = sc_A + sc_B + sc_C + sc_{AB} + sc_{AC} + sc_{BC} + sc_{ABC} + sc_R.$$

rafmoqui@yahoo.fr ~ 96877413

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	*Variation*	Ddl	Carré Moyen	F	Décision
Facteur A	sc_A	$n_A = I - 1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_{AC}^2}$	$\mathcal{H}_0$ ou $\mathcal{H}_1'$
Facteur B	sc_B	$n_B = J - 1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_{BC}^2}$	$\mathcal{H}_0''$ ou $\mathcal{H}_1''$
Facteur C	sc_C	$n_C = K - 1$	$sCB^2 = \frac{sc_C}{n_C}$	$f_C = \frac{s_C^2}{s_R^2}$	$\mathcal{H}_0'''$ ou $\mathcal{H}_1'''$
Interaction AB	sc_{AB}	$n_{AB} = (I-1)(J-1)$	$s_{AB}^2 = \frac{sc_{AB}}{n_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_{ABC}^2}$	$\mathcal{H}_0^{(4)}$ ou $\mathcal{H}_1^{(4)}$
Interaction AC	sc_{AC}	$n_{AC} = (I-1)(K-1)$	$s_{AC}^2 = \frac{sc_{AC}}{n_{AC}}$	$f_{AC} = \frac{s_{AC}^2}{s_R^2}$	$\mathcal{H}_0^{(5)}$ ou $\mathcal{H}_1^{(5)}$
Interaction BC	sc_{BC}	$n_{BC} = (J-1)(K-1)$	$s_{BC}^2 = \frac{sc_{BC}}{n_{BC}}$	$f_{BC} = \frac{s_{BC}^2}{s_R^2}$	$\mathcal{H}_0^{(6)}$ ou $\mathcal{H}_1^{(6)}$
Interaction ABC	sc_{ABC}	$n_{ABC} = (I-1)(J-1)(K-1)$	$s_{ABC}^2 = \frac{sc_{ABC}}{n_{ABC}}$	$f_{ABC} = \frac{s_{ABC}^2}{s_R^2}$	$\mathcal{H}_0^{(7)}$ ou $\mathcal{H}_1^{(7)}$
Résiduelle	sc_R	$n_R = IJK(L-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sC_{TOT}	$n_{TOT} = IJKL - 1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$
$$\text{contre}$$
$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

Sous $\mathcal{H}_0'$ $F_A = \dfrac{S_A^2}{S_{AC}^2} \sim$ Fisher de degré de liberté $[I-1, (I-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$ à l'aide d'une table, rejet si la valeur f_A est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle $\mathcal{H}_0'$ est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur .

$$\underline{\mathcal{H}_0'' : \beta_1 = \beta_2 = \cdots = \beta_J = 0}$$
$$\text{contre}$$
$$\mathcal{H}_1'' : \text{ Il existe } j_0 \in \{1, 2, \ldots, J\} \text{ tel que } \beta_{j_0} \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet du traitement B sur Y

Sous $\mathcal{H}_0''$ $F_B = \dfrac{S_B^2}{S_{BC}^2} \sim$ Fisher de degré de liberté $[J-1, (J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0''$ au risque α si la p-value $< \alpha$ à l'aide d'une table, rejet si la valeur f_B est supérieure ou égale à la valeur critique issue de la table.

Lorsque l'hypothèse nulle $\mathcal{H}_0''$ est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur .

$$\mathcal{H}_0''' : \sigma_C^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1''' : \sigma_C^2 \neq 0.$$

$\mathcal{H}_0'''$ sous entend l'absence d'effet du traitement C sur Y

Sous $\mathcal{H}_0'''$ $F_C = \dfrac{S_C^2}{S_R^2} \sim$ Fisher de degré de liberté $[K-1, IJK(L-1)]$

Règle de décision : On rejette $\mathcal{H}_0'''$ au risque α si la p-value $< \alpha$ à l'aide d'une table, rejet si la valeur f_C est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle $\mathcal{H}_0'''$ est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur .

$$\mathcal{H}_0^{(4)} : (\alpha\beta)_{1,1} = (\alpha\beta)_{1,2} = \cdots = (\alpha\beta)_{1,J} = (\alpha\beta)_{2,1} = \cdots = (\alpha\beta)_{I,J} = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(4)} : \text{ Il existe } (i_0, j_0) \in \{1, 2, \ldots, I\} \times \{1, 2, \ldots, J\} \text{ tel que } (\alpha\beta)_{i_0,j_0} \neq 0.$$

$\mathcal{H}_0^{(4)}$ sous entend l'absence d'effet l'interaction des facteurs A et B sur Y

Sous $\mathcal{H}_0^{(4)}$ $F_{AB} = \dfrac{S_{AB}^2}{S_{ABC}^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1), (I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(4)}$ au risque α si la p-value $< \alpha$ à l'aide d'une table, rejet si la valeur f_{AB} est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle $\mathcal{H}_0^{(4)}$ est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur .

$$\mathcal{H}_0^{(5)} : \sigma_{AC}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(5)} : \sigma_{AC}^2 \neq 0.$$

$\mathcal{H}_0^{(5)}$ sous entend l'absence d'effet l'interaction des facteurs A et C sur Y

Sous $\mathcal{H}_0^{(5)}$ $F_{AC} = \dfrac{S_{AC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(K-1), IJK(L-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(5)}$ au risque α si la p-value $< \alpha$ à l'aide d'une table, rejet si la valeur f_{AC} est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle $\mathcal{H}_0^{(5)}$ est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur .

rafmoqui@yahoo.fr ~ 96877413

$$\mathcal{H}_0^{(6)} : \sigma_{BC}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(6)} : \sigma_{BC}^2 \neq 0.$$

$\mathcal{H}_0^{(6)}$ sous entend l'absence d'effet l'interaction des facteurs B et C sur Y

Sous $\mathcal{H}_0^{(6)}$ $F_{BC} = \dfrac{S_{BC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(J-1)(K-1), IJK(L-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(6)}$ au risque α si la p-value $< \alpha$ à l'aide d'une table, rejet si la valeur f_{BC} est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle $\mathcal{H}_0^{(6)}$ est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur .

$$\mathcal{H}_0^{(7)} : \sigma_{ABC}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(7)} : \sigma_{ABC}^2 \neq 0.$$

$\mathcal{H}_0^{(7)}$ sous entend l'absence d'effet l'interaction d'ordre trois des facteurs A, B et C sur Y

Sous $\mathcal{H}_0^{(7)}$ $F_{ABC} = \dfrac{S_{ABC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1)(K-1), IJK(L-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(7)}$ au risque α si la p-value $< \alpha$ à l'aide d'une table, rejet si la valeur f_{ABC} est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle $\mathcal{H}_0^{(7)}$ est rejetée.

Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\beta_1}, \ldots, \widehat{\beta_J}, \widehat{\sigma_C^2}, \widehat{(\alpha\beta)_{1,1}}, \ldots, \widehat{(\alpha\beta)_{I,J}}, \widehat{\sigma_{AC}^2}, \widehat{\sigma_{BC}^2}, \widehat{\sigma_{ABC}^2}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots, \alpha_I, \beta_1, \ldots, \beta_J, \sigma_C^2, (\alpha\beta)_{1,1}, \ldots, (\alpha\beta)_{I,J}, \sigma_{AC}^2, \sigma_{BC}^2, \sigma_{ABC}^2, \sigma^2$ du modèle sont donnés par les formules suivantes :

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

$$
\begin{aligned}
&\widehat{\mu} = Y_{\bullet,\bullet,\bullet\bullet\bullet} = \bar{Y},\\
&\widehat{\alpha_i} = Y_{i,\bullet,\bullet,\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I,\\
&\widehat{\beta_j} = Y_{\bullet,j,\bullet,\bullet} - \widehat{\mu}, 1 \leqslant j \leqslant J,\\
&\widehat{\sigma^2_C} = \frac{1}{IJL}\left(S^2_C - S^2_R\right),\\
&\widehat{(\alpha\beta)_{i,j}} = Y_{i,j,\bullet,\bullet} - Y_{i,\bullet,\bullet,\bullet} - Y_{\bullet,j,\bullet,\bullet} + \widehat{\mu}, \qquad 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,\\
&\widehat{\sigma^2_{AC}} = \frac{1}{JL}\left(S^2_{AC} - S^2_R\right),\\
&\widehat{\sigma^2_{BC}} = \frac{1}{IL}\left(S^2_{BC} - S^2_R\right),\\
&\widehat{\sigma^2_{ABC}} = \frac{1}{L}\left(S^2_{ABC} - S^2_R\right),\\
&\widehat{\sigma^2} = \frac{SC_R}{IJK(L-1)} = S^2_R.
\end{aligned}
$$

où $S^2_C = \frac{SC_C}{n_C}, S^2_{AC} = \frac{SC_{AC}}{n_{AC}}, S^2_{BC} = \frac{SC_{BC}}{n_{BC}}, S^2_{ABC} = \frac{SC_{ABC}}{n_{ABC}}$ et $S^2_R = \frac{SC_R}{n_R}$. Ce sont des estimateurs sans biais.

3.4.2 Deuxième cas : Un facteur est à effets fixes et deux facteurs sont à effets aléatoires.

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i. Les β_j représentent un échantillon de taille J prélevé dans une population importante. Nous admettrons que les effets des B_j, les β_j, sont distribués suivant une loi normale centrée de variance σ^2_B. Les γ_k représentent un échantillon de taille K prélevé dans une population importante. Nous admettrons que les effets des C_k, les γ_k, sont distribués suivant une loi normale centrée de variance σ^2_C. Pour chacun des couples de modalités $\left(A_i, B_j, C_k\right)$ nous effectuons $L \geqslant 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K \times L$ le nombre total de mesures ayant été effectuées. Nous introduisons le modèle :

$$Y_{i,j,k,l} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + (\alpha\beta\gamma)_{i,j,k} + \epsilon_{i,j,k,l},$$
$$i = 1 \dots I, j = 1 \dots J, k = 1 \dots K, l = 1 \dots L,$$
$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0,$$
$$\sum_{i=1}^{I} (\alpha\beta)_{i,j} = 0, \forall j \in \{1, \dots, J\} \text{ et } \sum_{i=1}^{I} (\alpha\gamma)_{i,k} = 0, \forall k \in \{1, \dots, K\},$$
$$\sum_{i=1}^{I} (\alpha\beta\gamma)_{i,j,k} = 0, \forall (j,k) \in \{1, \dots, J\} \times \{1, \dots, K\},$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$ lors du l-ème essai. Nous supposons que :

$$\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J,$$
$$\mathcal{L}\left(\gamma_k\right) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\beta)_{i,j}\right) = \mathcal{N}\left(0, \sigma_{AB}^2\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,$$
$$\mathcal{L}\left((\alpha\gamma)_{i,k}\right) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\beta\gamma)_{i,j,k}\right) = \mathcal{N}\left(0, \sigma_{ABC}^2\right), \forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$

ainsi que l'indépendance des effets aléatoires β_j, γ_k et $(\beta\gamma)_{j,k}$.

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k,l), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, 1 \leqslant l \leqslant L, \mathcal{L}\left(\epsilon_{i,j,k,l}\right) = \mathcal{N}\left(0, \sigma^2\right)$$
$$\epsilon_{i,j,k,l} \perp\!\!\!\perp \epsilon_{m,n,o,p} \text{ si } (i,j,k,l) \neq (m,n,o,p)$$
$$\text{avec } 1 \leqslant i, m \leqslant I, 1 \leqslant j, n \leqslant J, 1 \leqslant k, o \leqslant K \text{ et } 1 \leqslant l, p \leqslant L$$

ainsi que l'indépendance des effets aléatoires β_j, γ_k, $(\alpha\beta)_{i,j}$, $(\alpha\gamma)_{i,k}$, $(\beta\gamma)_{j,k}$ et $(\alpha\beta\gamma)_{i,j,k}$ et des erreurs $\epsilon_{i,j,k,l}$.

Dans un modèle mixte restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici aussi bien les $(\alpha\beta)_{i,j}$ que les $(\alpha\gamma)_{i,k}$ et que les $(\alpha\beta\gamma)_{i,j,k}$, ne sont pas mutuellement indépendants à cause des contraintes portant sur leur somme, $\sum_{i=1}^{I}(\alpha\beta)_{i,j} = 0, \forall j \in \{1, \dots, J\}, \sum_{i=1}^{I}(\alpha\gamma)_{i,k} = 0, \forall k \in \{1, \dots, K\}$ et $\sum_{i=1}^{I}(\alpha\beta\gamma)_{i,j,k} = 0, \forall (j,k) \in \{1, \dots, J\} \times \{1, \dots, K\}$. Ils le sont par contre dès que nous ne les considérons pas tous en même temps.

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Certains logiciels travaillent par défaut avec un modèle non-restreint dont la définition est proche mais pas identique à celle du modèle précédent. Nous introduisons le modèle :

$$Y_{i,j,k,l} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{i,j} + (\alpha\gamma)_{i,k} + (\beta\gamma)_{j,k} + (\alpha\beta\gamma)_{i,j,k} + \epsilon_{i,j,k,l},$$
$$i = 1 \dots I, j = 1 \dots J, k = 1 \dots K, l = 1 \dots L$$
$$\text{avec les contraintes supplémentaires } \sum_{i=1}^{I} \alpha_i = 0$$

où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans les conditions $\left(A_i, B_j, C_k\right)$. Nous supposons que :

$$\mathcal{L}\left(\beta_j\right) = \mathcal{N}\left(0, \sigma_B^2\right), \forall j, 1 \leqslant j \leqslant J,$$
$$\mathcal{L}\left(\gamma_k\right) = \mathcal{N}\left(0, \sigma_C^2\right), \forall k, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\beta)_{i,j}\right) = \mathcal{N}\left(0, \sigma_{AB}^2\right), \forall (i,j), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J,$$
$$\mathcal{L}\left((\alpha\gamma)_{i,k}\right) = \mathcal{N}\left(0, \sigma_{AC}^2\right), \forall (i,k), 1 \leqslant i \leqslant I, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\beta\gamma)_{j,k}\right) = \mathcal{N}\left(0, \sigma_{BC}^2\right), \forall (j,k), 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$
$$\mathcal{L}\left((\alpha\beta\gamma)_{i,j,k}\right) = \mathcal{N}\left(0, \sigma_{ABC}^2\right), \forall (i,j,k), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K,$$

ainsi que l'indépendance des effets aléatoires $\beta_j, \gamma_k, (\alpha\beta)_{i,j}, (\alpha\gamma)_{i,k}, (\beta\gamma)_{j,k}$ et $(\alpha\beta\gamma)_{i,j,k}$. Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i,j,k,l), 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K, 1 \leqslant l \leqslant L, \mathcal{L}\left(\epsilon_{i,j,k,l}\right) = \mathcal{N}\left(0, \sigma^2\right)$$
$$\epsilon_{i,j,k,l} \perp\!\!\!\perp \epsilon_{m,n,o,p} \text{ si } (i,j,k,l) \neq (m,n,o,p)$$
$$\text{avec } 1 \leqslant i, m \leqslant I, 1 \leqslant j, n \leqslant J, 1 \leqslant k, o \leqslant K \text{ et } 1 \leqslant l, p \leqslant L$$

ainsi que l'indépendance des effets aléatoires $\beta_j, \gamma_k, (\alpha\beta)_{i,j}, (\alpha\gamma)_{i,k}, (\beta\gamma)_{j,k}$ et $(\alpha\beta\gamma)_{i,j,k}$ et des erreurs $\epsilon_{i,j,k,l}$.

Dans un modèle mixte non restreint, les effets aléatoires croisant des facteurs à effets fixes et à effets aléatoires, ici les $(\alpha\beta)_{i,j}, (\alpha\gamma)_{i,k}$ et les $(\alpha\beta\gamma)_{i,j,k}$, sont mutuellement indépendants.

Il n'existe pas consensus sur une raison statistique qui permettrait de privilégier plutôt l'une ou l'autre de ces deux approches. Nous utiliserons toujours des modèles restreints. Nous rappelons la relation fondamentale de l'ANOVA :

$$sc_{TOT} = sc_A + sc_B + sc_C + sc_{AB} + sc_{AC} + sc_{BC} + s_{ABC} + sc_R.$$

Pour des données expérimentales, nous aurons le tableau d'ANOVA comme suit :

Source	Variation	$\overline{\text{Ddl}}$	Carré Moyen	F	Décisi
Facteur A	sc_A	$n_A = I-1$	$s_A^2 = \frac{sc_A}{n_A}$	$f_A = \frac{s_A^2}{s_{AB}^2+s_{AC}^2-s_{ABC}^2}$	$\mathcal{H}_0'$ ou
Facteur B	sc_B	$n_B = J-1$	$s_B^2 = \frac{sc_B}{n_B}$	$f_B = \frac{s_B^2}{s_{BC}^2}$	$\mathcal{H}_0''$ ou
Facteur C	sc_C	$n_C = K-1$	$s_C^2 = \frac{sc_C}{n_C}$	$f_C = \frac{s_C^2}{s_{BC}^2}$	$\mathcal{H}_0'''$ ou
Interaction AB	sc_{AB}	$n_{AB} = (I-1)(J-1)$	$s_{AB}^2 = \frac{sc_{AB}}{n_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_{ABC}^2}$	$\mathcal{H}_0^{(4)}$ ou
Interaction AC	sc_{AC}	$n_{AC} = (I-1)(K-1)$	$s_{AC}^2 = \frac{sc_{AC}}{n_{AC}}$	$f_{AC} = \frac{s_{AC}^2}{s_{ABC}^2}$	$\mathcal{H}_0^{(5)}$ ou
Interaction BC	sc_{BC}	$n_{BC} = (J-1)(K-1)$	$s_{BC}^2 = \frac{sc_{BC}}{n_{BC}}$	$f_{BC} = \frac{s_{BC}^2}{s_R^2}$	$\mathcal{H}_0^{(6)}$ ou
Interaction ABC	sc_{ABC}	$n_{ABC} = (I-1)(J-1)(K-1)$	$s_{ABC}^2 = \frac{sc_{ABC}}{n_{ABC}}$	$f_{ABC} = \frac{s_{ABC}^2}{s_R^2}$	$\mathcal{H}_0^{(7)}$ ou
Résiduelle	sc_R	$n_R = IJK(L-1)$	$s_R^2 = \frac{sc_R}{n_R}$		
Totale	sc_{TOT}	$n_{TOT} = IJKL-1$			

où les hypothèses sont les suivantes

$$\mathcal{H}_0' : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$

contre

$$\mathcal{H}_1' : \text{ Il existe } i_0 \in \{1, 2, \ldots, I\} \text{ tel que } \alpha_{i_0} \neq 0.$$

$\mathcal{H}_0'$ sous entend l'absence d'effet du traitement A sur Y

Sous $\mathcal{H}_0'$ $F_A = \dfrac{S_A^2}{S_{AB}^2 + S_{AC}^2 - S_{ABC}^2} \sim$ Fisher de degré de liberté $[I-1, n']$

$$n' = \frac{\left(s_{AB}^2 + s_{AC}^2 - s_R^2\right)^2}{\frac{\left(s_{AB}^2\right)^2}{(I-1)(J-1)} + \frac{\left(s_{AC}^2\right)^2}{(I-1)(K-1)} + \frac{\left(s_R^2\right)^2}{(I-1)(J-1)(K-1)}}.$$

Règle de décision : On rejette $\mathcal{H}_0'$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_A est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle $\mathcal{H}_0'$ est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur.

$$\mathcal{H}_0'' : \sigma_B^2 = 0$$

contre

$$\mathcal{H}_1'' : \sigma_B^2 \neq 0.$$

$\mathcal{H}_0''$ sous entend l'absence d'effet du traitement B sur Y

Sous $\mathcal{H}_0''$ $F_B = \dfrac{S_B^2}{S_{BC}^2} \sim$ Fisher de degré de liberté $[J-1, (J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0''$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_B est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}_0''' : \sigma_C^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1''' : \sigma_C^2 \neq 0.$$

$\mathcal{H}_0'''$ sous entend l'absence d'effet du traitement C sur Y

Sous $\mathcal{H}_0'''$ $F_C = \dfrac{S_B^2}{S_{BC}^2} \sim$ Fisher de degré de liberté $[K-1, (J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0'''$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_C est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}_0^{(4)} : \sigma_{AB}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(4)} : \sigma_{AB}^2 \neq 0.$$

$\mathcal{H}_0^{(4)}$ sous entend l'absence d'effet l'interaction des traitements A et B sur Y.

Sous $\mathcal{H}_0^{(4)}$ $F_{AB} = \dfrac{S_{AB}^2}{S_{ABC}^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1), (I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(4)}$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_{AB} est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}_0^{(5)} : \sigma_{AC}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(5)} : \sigma_{AC}^2 \neq 0.$$

$\mathcal{H}_0^{(5)}$ sous entend l'absence d'effet l'interaction des traitements A et C sur Y.

Sous $\mathcal{H}_0^{(5)}$ $F_{AC} = \dfrac{S_{AC}^2}{S_{ABC}^2} \sim$ Fisher de degré de liberté $[(I-1)(K-1), (I-1)(J-1)(K-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{()}$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_{AC} est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}_0^{(6)} : \sigma_{BC}^2 = 0$$
$$\text{contre}$$
$$\mathcal{H}_1^{(6)} : \sigma_{BC}^2 \neq 0.$$

$\mathcal{H}_0^{(6)}$ sous entend l'absence d'effet l'interaction des traitements B et C sur Y.

Sous $\mathcal{H}_0^{(6)}$ $F_{BC} = \dfrac{S_{BC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(J-1)(K-1), IJK(L-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(6)}$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_{BC} est supérieure ou égale à la valeur critique issue de la table.

$$\mathcal{H}_0^{(7)} : \sigma_{ABC}^2 = 0$$
contre
$$\mathcal{H}_1^{(7)} : \sigma_{ABC}^2 \neq 0.$$

$\mathcal{H}_0^{(7)}$ sous entend l'absence d'effet de l'interaction des traitements A, B et C sur Y.

Sous $\mathcal{H}_0^{(7)}$ $F_{ABC} = \dfrac{S_{BC}^2}{S_R^2} \sim$ Fisher de degré de liberté $[(I-1)(J-1)(K-1), IJK(L-1)]$

Règle de décision : On rejette $\mathcal{H}_0^{(6)}$ au risque α si la p-value $< \alpha$ ou à l'aide d'une table, rejet si la valeur f_{ABC} est supérieure ou égale à la valeur critique issue de la table.

Les estimateurs $\widehat{\mu}, \widehat{\alpha_1}, \ldots, \widehat{\alpha_I}, \widehat{\sigma_B^2}, \widehat{\sigma_C^2}, \widehat{\sigma_{AB}^2}, \widehat{\sigma_{AC}^2}, \widehat{\sigma_{BC}^2}, \widehat{\sigma_{ABC}^2}, \widehat{\sigma^2}$ des paramètres $\mu, \alpha_1, \ldots,$ $\alpha_I, \sigma_B^2, \sigma_C^2, \sigma_{AB}^2, \sigma_{AC}^2, \sigma_{BC}^2, \sigma_{ABC}^2, \sigma^2$ du modèle sont donnés par les formules suivantes :

$$\begin{aligned}
\widehat{\mu} &= Y_{\bullet,\bullet,\bullet,\bullet} = \bar{Y} \\
\widehat{\alpha_i} &= Y_{i,\bullet,\bullet\bullet} - \widehat{\mu}, 1 \leqslant i \leqslant I, \\
\widehat{\sigma_B^2} &= \frac{1}{IKL}\left(S_B^2 - S_{BC}^2\right), \\
\widehat{\sigma_C^2} &= \frac{1}{IJL}\left(S_C^2 - S_{BC}^2\right), \\
\widehat{\sigma_{AB}^2} &= \frac{1}{KL}\left(S_{AB}^2 - S_{ABC}^2\right), \\
\widehat{\sigma_{AC}^2} &= \frac{1}{JL}\left(S_{AC}^2 - S_{ABC}^2\right), \\
\widehat{\sigma_{BC}^2} &= \frac{1}{IL}\left(S_{BC}^2 - S_R^2\right), \\
\widehat{\sigma_{ABC}^2} &= \frac{1}{L}\left(S_{ABC}^2 - S_R^2\right), \\
\widehat{\sigma^2} &= \frac{SC_R}{IJK(L-1)} = S_R^2.
\end{aligned}$$

où $S_B^2 = \frac{SC_B}{n_B}, S_C^2 = \frac{SC_C}{n_C}, S_{AB}^2 = \frac{SC_{AB}}{n_{AB}}, S_{AC}^2 = \frac{SC_{AC}}{n_{AC}}, S_{BC}^2 = \frac{SC_{BC}}{n_{BC}}, S_{ABC}^2 = \frac{SC_{ABC}}{n_{ABC}}$ et $S_R^2 = \frac{SC_R}{n_R}$.
Ce sont des estimateurs sans biais.

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Chapitre quatre

Bibliographie

Livres : 1. Pinheiro, J. C., & Bates, D. M. (2000). "Mixed-Effects Models in S and S-PLUS". Springer-Verlag. Ce livre couvre en détail l'utilisation des modèles linéaires mixtes dans le logiciel R.

2. Verbeke, G., & Molenberghs, G. (2009). "Linear Mixed Models for Longitudinal Data". Springer. Un livre de référence sur les modèles linéaires mixtes appliqués aux données longitudinales.

3. Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). "Applied Longitudinal Analysis". Wiley. Ce livre explore les modèles linéaires mixtes pour l'analyse de données longitudinales.

Articles et ressources en ligne : 4. Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (2006). "SAS for Mixed Models". SAS Institute. Ce livre explique comment utiliser SAS pour les modèles linéaires mixtes.

5. Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). "Fitting Linear Mixed-Effects Models Using lme4". Journal of Statistical Software, 67(1), 1-48. Cet article décrit l'utilisation du package R 'lme4' pour les modèles linéaires mixtes.

6. Raudenbush, S. W., & Bryk, A. S. (2002). "Hierarchical Linear Models : Applications and Data Analysis Methods". Sage Publications. Ce livre se concentre sur les modèles linéaires mixtes à plusieurs niveaux.

7. Singer, J. D., & Willett, J. B. (2003). "Applied Longitudinal Data Analysis : Modeling Change and Event Occurrence". Oxford University Press. Ce livre explore l'analyse des données longitudinales en utilisant des modèles linéaires mixtes.

8. Harville, D. A. (1977). "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems". Journal of the American Statistical Association, 72(358), 320-338. Cet article classique présente des méthodes d'estimation des composantes de variance dans les modèles linéaires mixtes.

9. Hox, J. J. (2010). "Multilevel Analysis : Techniques and Applications". Routledge. Ce livre couvre l'analyse de données à plusieurs niveaux, y compris les modèles linéaires mixtes.

10. Frédéric Bertrand, Cours de Master1 sur les éléments d'analyse de la variance, **9 septembre 2019.**

11. Christèle Robert-Granié, Bertrand Servin, cours de Modèle Linéaire mixte gaussien, **Mars 2012.**

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413

Chapitre cinq

Annexe

Dr. Raymond M M AFFOSSOGBE ©ENSPD-2024-2025
rafmoqui@yahoo.fr ~ 96877413