Chapitre V

Analyse du plan $S \times A \times B$

El Methni M.

V-1. *A* et *B* sont deux facteurs à effets fixes

V-1-1 Généralités

On considère le plan complet $S \times A \times B$ défini par le croisement de trois facteurs S, A et B.
S est le facteur sujet (aléatoire) à n modalités. A et B sont deux facteurs à effets fixes, A possède r modalités et B c modalités. Chacun des n sujets est observé dans chacun des $r \times c$ croisements des modalités de A et B.

On dispose donc de $N=n \times r \times c$ observations y_{sij} que l'on présente sous la forme d'un tableau :

Sujets	a_1					…	a_i					…	a_r				
	b_1	…	b_j	…	b_c	…	b_1	…	b_j	…	b_c	…	b_1	…	b_j	…	b_c
1	y_{111}	…	y_{11j}	…	y_{11c}	…	y_{1i1}	…	y_{1ij}	…	y_{1ic}	…	y_{1r1}	…	y_{1rj}	…	y_{1rc}
2	y_{211}	…	y_{21j}	…	y_{21c}	…	y_{2i1}	…	y_{2ij}	…	y_{2ic}	…	y_{2r1}	…	y_{2rj}	…	y_{2rc}
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
s	y_{s11}	…	y_{s1j}	…	y_{s1c}	…	y_{si1}	…	y_{sij}	…	y_{sic}	…	y_{sr1}	…	y_{srj}	…	y_{src}
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
n	y_{n11}	…	y_{n1j}	…	y_{n1c}	…	y_{ni1}	…	y_{nij}	…	y_{nic}	…	y_{nr1}	…	y_{nrj}	…	y_{nrc}

V-1-2 Modèle univarié :

Les données peuvent être d'une part, considérées comme les $N = n \times r \times c$ observations d'une seule variable aléatoire Y dans les $N = n \times r \times c$ conditions expérimentales décrites par le croisement des trois facteurs S, A et B. Elles peuvent d'autre part, être regardées comme les n observations d'un vecteur de $r \times c$ variables aléatoires. Ces deux points de vue conduisent à l'élaboration de deux modèles, le <u>modèle mixte univarié</u> dans le premier cas, et le <u>modèle multivarié</u> dans le second. Comme pour le plan $S \times O$, nous allons étudier le modèle mixte univarié dans lequel nous considérons que les modalités du facteur sujet ont été obtenues par échantillonnage, le facteur sujet est donc un facteur aléatoire. Les facteurs A et B sont des facteurs à effets fixes. Chaque donnée y_{sij} correspond alors à l'observation d'une variable aléatoire Y_{sij} et on pose le modèle mixte suivant :

$$Y_{sij} = \mu_{ij} + \pi_s + (\alpha\pi)_{is} + (\beta\pi)_{js} + (\alpha\beta\pi)_{sij} + e_{sij}$$

Pour chaque croisement des modalités s, i et j nous ne disposons que d'une seule observation, l'interaction $(\alpha\beta\pi)_{sij}$ sera confondue avec le résidu et on pose : $\varepsilon_{sij}=(\alpha\beta\pi)_{sij} + e_{sij}$

$$\text{On a donc : } Y_{sij} = \mu_{ij} + \pi_s + (\alpha\pi)_{is} + (\beta\pi)_{js} + \varepsilon_{sij}$$

Où

Les μ_{ij} (i =1, 2, …, r, j=1, 2, …, c) sont des constantes qui mesurent les effets fixes des modalités (i,j) du croisement $A \times B$

Les π_s (s = 1, 2, …, n) sont des variables aléatoires indépendantes et identiquement distribuées (*iid*) de loi $\mathcal{N}(0\,;\sigma^2_\pi)$ qui mesurent les effets aléatoires des modalités s du facteur S

Les $(\alpha\pi)_{si}$ (s = 1, 2, …, n, i = 1, 2, …, r) sont des variables aléatoires indépendantes et identiquement distribuées (*iid*) de loi $\mathcal{N}(0\,;\sigma^2_{\alpha\pi})$ qui mesurent les effets aléatoires d'interactions entre le facteur S et le facteur A

Les $(\beta\pi)_{sj}$ (s = 1, 2, …, n, j = 1, 2, …, c) sont des variables aléatoires indépendantes et identiquement distribuées (iid) de loi $\mathcal{N}(0\,;\sigma^2_{\beta\pi})$ qui mesurent les effets aléatoires d'interactions entre le facteur S et le facteur B
Les résidus ε_{sij} sont des variables aléatoires indépendantes et identiquement distribuées (iid) de loi $\mathcal{N}(0\,;\sigma^2)$
De plus, on suppose que les résidus ε_{sij} sont indépendantes des π_s, des $(\alpha\pi)_{si}$ et des $(\beta\pi)_{sj}$.
Nous pouvons réécrire l'effet fixe de la modalité (i,j) sous la forme suivante :

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

Où :
Le paramètre μ s'interprète comme un niveau général de réponse commun pour l'ensemble des observations,
Le paramètre $\alpha_i = \mu_i - \mu$ s'interprète comme l'effet de la modalité i du facteur A
Le paramètre $\beta_j = \mu_j - \mu$ s'interprète comme l'effet de la modalité j du facteur B
Le paramètre $(\alpha\beta)_{ij} = \mu_{ij} - \mu_i - \mu_j + \mu$ s'interprète comme l'effet de la modalité (i,j) du croisement $A\times B$ (l'effet d'interaction de A et B).
Le modèle mixte univarié peut finalement s'écrire :

$$Y_{sij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \pi_s + (\alpha\pi)_{is} + (\beta\pi)_{js} + \varepsilon_{sij}$$

Où : les Y_{sij} sont des variables aléatoires de loi : $\mathcal{N}(\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}\,, \sigma^2_\pi + \sigma^2_{\alpha\pi} + \sigma^2_{\beta\pi} + \sigma^2)$

Une telle paramétrisation nécessite de rajouter les contraintes d'identifiabilité suivantes :

$$\sum_{i=1}^{r}\alpha_i = 0 \quad \sum_{j=1}^{c}\beta_i = 0 \quad \sum_{i=1}^{r}(\alpha\beta)_{ij} = 0 \quad \sum_{j=1}^{c}(\alpha\beta)_{ij} = 0$$

V-1-3 Décomposition de la variation :
On commence par décomposer la variation totale en variation inter-sujets et variation intra-sujets :
$\mathbf{SCT}_{\text{obs}} = \mathbf{SC}_{\text{inter-sujets}} + \mathbf{SC}_{\text{intra-sujets}}$

$$\mathbf{SC}_{\text{inter-sujet}} = rc\sum_{s=1}^{n}(\bar{y}_{s00} - \bar{y})^2 = N \times \text{Variance des moyennes des sujets}$$

$$\mathbf{SC}_{\text{intra-sujet}} = \sum_{s,i,j}(y_{sij} - \bar{y}_{s00})^2$$

La variation inter-sujets est due au facteur sujet et la variation intra-sujets est due aux effets du croisement $A\times B$ et à l'effet des autres facteurs non contrôlés. L'effet du facteur $A\times B$ se fait à deux niveaux : effet principal et effet d'interaction.

$$\mathbf{SC}_{\text{intra-sujets}} = \mathbf{SC}(A\times B)_{\text{obs}} + \mathbf{SC}(A\times B)S_{\text{obs}} + \mathbf{SCR}_{\text{obs}}$$

Avec : $\mathbf{SC}(A\times B)_{\text{obs}} = \mathbf{SC}(A)_{\text{obs}} + \mathbf{SC}(B)_{\text{obs}} + \mathbf{SC}(AB)_{\text{obs}}$
et : $\mathbf{SC}(A\times B)S_{\text{obs}} = \mathbf{SC}(AS)_{\text{obs}} + \mathbf{SC}(BS)_{\text{obs}} + \mathbf{SC}(ABS)_{\text{obs}}$
Ne disposant que d'une seule observation pour chaque croisement (i,j,s) on ne peut donc séparer l'interaction SAB du résidu on a donc : $\mathbf{SC}(ABS)_{\text{obs}} = \mathbf{SCR}$
Conclusion : Dans le cas du plan complet $S\times A\times B$, la variation totale se décompose de façon additive en :

$$\boxed{\mathbf{SCT}_{\text{obs}} = \mathbf{SCS}_{\text{obs}} + \mathbf{SC}A_{\text{obs}} + \mathbf{SC}B_{\text{obs}} + \mathbf{SC}AB_{\text{obs}} + \mathbf{SC}AS_{\text{obs}} + \mathbf{SC}BS_{\text{obs}} + \mathbf{SCR}_{\text{obs}}}$$

De même on a la décomposition des degrés de liberté :
N-1=ncr-1 = (n-1) + n(rc-1) = (n-1) + n(rc-1) - (c-1)(r-1)(n-1) + (c-1)(r-1)(n-1)

$$\boxed{N-1 = ncr-1 = (n-1)+(r-1)+(c-1)+(c-1)(r-1)+(r-1)(n-1)+(c-1)(n-1)}$$

Les différentes sommes de carrées se calculent de façon habituelles par :

$$\mathbf{SCT}_{\text{obs}} = \sum_{s=1}^{n}\sum_{j=1}^{c}\sum_{i=1}^{r}(y_{sij} - \bar{y})^2 = N\times\text{variance de toutes les observations}$$

$\mathbf{SCS}_{\text{obs}} = rc\sum_{s=1}^{n}(\overline{y}_{s00} - \overline{y})^2$ =N×variance des n moyennes des sujets

$\mathbf{SCA}_{\text{obs}} = nc\sum_{i=1}^{r}(\overline{y}_{0i0} - \overline{y})^2$ =N×variance des r moyennes par modalités de A

$\mathbf{SCB}_{\text{obs}} = nr\sum_{j=1}^{c}(\overline{y}_{00j} - \overline{y})^2$ =N×variance des c moyennes par modalités de B

$\mathbf{SC(}\text{A}\times\text{B}\mathbf{)}_{\text{obs}} = n\sum_{i=1}^{r}\sum_{j=1}^{c}(\overline{y}_{0ij} - \overline{y})^2$ =N×variance des r×c moyennes par croisement de A et B

SC(AB)$_{\text{obs}}$ = **SC**(A×B)$_{\text{obs}}$ - **SCA**$_{\text{obs}}$ - **SCB**$_{\text{obs}}$

$\mathbf{SC(}\text{A}\times\text{S}\mathbf{)}_{\text{obs}} = c\sum_{s=1}^{n}\sum_{i=1}^{r}(\overline{y}_{si0} - \overline{y})^2$ =N×variance des r×n moyennes par croisement de A et S

SC(AS)$_{\text{obs}}$ = **SC**(A×S)$_{\text{obs}}$ - **SCA**$_{\text{obs}}$ - **SCS**$_{\text{obs}}$

$\mathbf{SC(}\text{B}\times\text{S}\mathbf{)}_{\text{obs}} = r\sum_{s=1}^{n}\sum_{j=1}^{c}(\overline{y}_{s0j} - \overline{y})^2$ =N×variance des n×c moyennes par croisement de B et S

SC(BS)$_{\text{obs}}$ = **SC**(B×S)$_{\text{obs}}$ - **SCB**$_{\text{obs}}$ - **SCS**$_{\text{obs}}$
SCR$_{\text{obs}}$ = **SCT**$_{\text{obs}}$ - **SCS**$_{\text{obs}}$ - **SCA**$_{\text{obs}}$ - **SCB**$_{\text{obs}}$ - **SC**(AB)$_{\text{obs}}$ - **SC**(AS)$_{\text{obs}}$ - **SC**(BS)$_{\text{obs}}$

Dans le cadre du modèle statistique ces sommes de carrés sont des réalisations de variables aléatoires dont on calcule les espérances et plus généralement les distributions des probabilités.
On montre le théorème fondamental suivant :

Théorème : Sous les hypothèses du modèle, les statistiques **SCA**, **SCB**, **SCS**, **SCAB**, **SCAS**, **SCBS** et **SCR** sont indépendantes et d'espérances respectives :

$$E(\mathbf{SCS}) = (n-1)\sigma^2 + rc\sigma_\pi^2$$
$$E(\mathbf{SCA}) = (r-1)(\sigma^2 + c\sigma_{\alpha\pi}^2) + \sum_i nc\alpha_i^2$$
$$E(\mathbf{SCB}) = (c-1)(\sigma^2 + r\sigma_{\beta\pi}^2) + \sum_j nr\beta_j^2$$
$$E(\mathbf{SC(AB)}) = (r-1)(c-1)\sigma^2 + \sum_i\sum_j n(\alpha\beta)_{ij}^2$$
$$E(\mathbf{SC(AS)}) = (n-1)(r-1)(\sigma^2 + c\sigma_{\alpha\pi}^2)$$
$$E(\mathbf{SC(BS)}) = (n-1)(c-1)(\sigma^2 + r\sigma_{\beta\pi}^2)$$
$$E(\mathbf{SCR}) = (r-1)(c-1)(n-1)\sigma^2$$

On ramène toutes les sommes de carrés à des moyennes de carrés en divisant par les degrés de liberté correspondants.

$$\mathbf{MCA} = \frac{\mathbf{SCA}}{r-1} \quad \mathbf{MCB} = \frac{\mathbf{SCB}}{c-1} \quad \mathbf{MC(AS)} = \frac{\mathbf{SC(AS)}}{(r-1)(n-1)} \quad \mathbf{MCAB} = \frac{\mathbf{SC(AB)}}{(r-1)(c-1)}$$

$$\mathbf{MC(BS)} = \frac{\mathbf{SC(BS)}}{(c-1)(n-1)} \quad \mathbf{MCR} = \frac{\mathbf{SCR}}{(r-1)(c-1)(n-1)}$$

Ceci nous permet de réécrire le théorème précédent sous la forme suivante :

Théorème : Sous les hypothèses du modèle, les statistiques **MCA**, **MCB**, **MCS**, **MCAB**, **MCAS**, **MCBS** et **MCR** sont indépendantes et d'espérances respectives :

$$E(\mathbf{MCS}) = \sigma^2 + \frac{rc}{(n-1)}\sigma_\pi^2$$
$$E(\mathbf{MCA}) = \sigma^2 + c\sigma_{\alpha\pi}^2 + \frac{1}{r-1}\sum_i nc\alpha_i^2$$
$$E(\mathbf{MCB}) = \sigma^2 + r\sigma_{\beta\pi}^2 + \frac{1}{c-1}\sum_j nr\beta_j^2$$
$$E(\mathbf{MC(AB)}) = \sigma^2 + \frac{1}{(r-1)(c-1)}\sum_i\sum_j n(\alpha\beta)_{ij}^2$$
$$E(\mathbf{MC(AS)}) = \sigma^2 + c\sigma_{\alpha\pi}^2$$
$$E(\mathbf{MC(BS)}) = \sigma^2 + r\sigma_{\beta\pi}^2$$
$$E(\mathbf{SCR}) = \sigma^2$$

Ceci nous permet de tester l'existence des différents effets des facteurs. On peut construire des tests indépendants sur chacune des sources de variations :

Test 1 :
hypothèse nulle $\mathbf{H_0}_{\mathrm{A}}$: pas d'effet principal du facteur A $\Leftrightarrow (\forall i)\ \alpha_i = 0$ contre
l'hypothèse alternative : $\mathbf{H_1}_{\mathrm{A}}$: il existe un effet principal du facteur A $\Leftrightarrow (\exists i)\ \alpha_i \neq 0$

Théorème : sous l'hypothèse $\mathbf{H_0}_{\mathrm{A}}$, la statistique $\mathbf{F_A} = \dfrac{\mathbf{MCA}}{\mathbf{MC(AS)}}$ suit une loi de Fischer à r - 1 et (n - 1)(r - 1) degrés de liberté.

Le test1 est alors défini au seuil de signification α par la règle de décision suivante :
si $\mathbf{F}_{\mathrm{A\ obs}} \geq \lambda_\alpha$ alors on rejette l'hypothèse nulle
où λ_α est donné par l'équation : $\alpha = P(\mathbf{F}_{\mathrm{A}} \geq \lambda_\alpha)$.

Test 2 :
hypothèse nulle $\mathbf{H_0}_{\mathrm{B}}$: pas d'effet principal du facteur B $\Leftrightarrow (\forall j)\ \beta_j = 0$ contre
l'hypothèse alternative : $\mathbf{H_1}_{\mathrm{B}}$: il existe un effet principal du facteur B $\Leftrightarrow (\exists j)\ \beta_j \neq 0$

Théorème : sous l'hypothèse $\mathbf{H_0}_{\mathrm{B}}$, la statistique $\mathbf{F_B} = \dfrac{\mathbf{MCB}}{\mathbf{MC(BS)}}$ suit une loi de Fischer à c - 1 et (n - 1)(c - 1) degrés de liberté.

Le test2 est alors défini au seuil de signification α par la règle de décision suivante :
si $\mathbf{F}_{\mathrm{B\ obs}} \geq \lambda_\alpha$ alors on rejette l'hypothèse nulle
où λ_α est donné par l'équation : $\alpha = P(\mathbf{F}_{\mathrm{B}} \geq \lambda_\alpha)$.

Test 3 :
hypothèse nulle $\mathbf{H_0}_{\mathrm{AB}}$: pas d'effet d'interaction $\Leftrightarrow (\forall i,j)\ (\alpha\beta)_{ij} = 0$ contre
l'hypothèse alternative : $\mathbf{H_1}_{\mathrm{AB}}$: il existe une interaction $\Leftrightarrow (\exists i,j)\ (\alpha\beta)_{ij} \neq 0$

Théorème : sous l'hypothèse $\mathbf{H_0}_{\mathrm{AB}}$, la statistique $\mathbf{F_{AB}} = \dfrac{\mathbf{MC(AB)}}{\mathbf{MCR}}$ suit une loi de Fischer à (r - 1)(c - 1) et (n - 1)(c - 1) (r - 1) degrés de liberté.

Le test3 est alors défini au seuil de signification α par la règle de décision suivante :

si $\mathbf{F}_{AB\ obs} \geq \lambda_\alpha$ alors on rejette l'hypothèse nulle

où λ_α est donné par l'équation : $\alpha = P(\mathbf{F}_{AB} \geq \lambda_\alpha)$.

On présente l'étude dans le tableau d'analyse de la variance :

Source	$\mathbf{SC}_{obs}$	ddl	$\mathbf{MC}_{obs}$	$\mathbf{F}_{obs}$
Inter-sujets : S	$\mathbf{SCS}_{obs}$	n - 1		
Intra-sujets :				
A	$\mathbf{SCA}_{obs}$	r - 1	$\mathbf{MCA}_{obs}$	$\mathbf{F}_{A\ obs}$
AS	$\mathbf{SC(AS)}_{obs}$	(r - 1)(n - 1)	$\mathbf{MC(AS)}_{obs}$	
B	$\mathbf{SCB}_{obs}$	c - 1	$\mathbf{MCB}_{obs}$	$\mathbf{F}_{B\ obs}$
BS	$\mathbf{SC(BS)}_{obs}$	(c - 1)(n - 1)	$\mathbf{MC(BS)}_{obs}$	
AB	$\mathbf{SC(AB)}_{obs}$	(r - 1)(c - 1)	$\mathbf{MC(AB)}_{obs}$	$\mathbf{F}_{AB\ obs}$
R	$\mathbf{SCR}_{obs}$	(r - 1)(c - 1)(n - 1)	$\mathbf{MCR}_{obs}$	
Total	$\mathbf{SCT}_{obs}$	N - 1		

V-1-5 Condition de validation :

Le fait de mesurer plusieurs fois la variable réponse sur le même sujet introduit des corrélations entre les observations faites sur ce même sujet. Dans le cas du modèle mixte univarié on montre

$$\mathrm{cov}(Y_{sij}, Y_{s'i'j'}) = \begin{cases} 0 \quad \text{si } s \neq s' \\ \sigma_\pi^2 \quad \text{si } s = s' \; et \; i \neq i' \; et \; j \neq j' \\ \sigma_\pi^2 + \sigma_{\alpha\pi}^2 \quad \text{si } s = s' \; et \; i = i' \; et \; j \neq j' \\ \sigma_\pi^2 + \sigma_{\beta\pi}^2 \quad \text{si } s = s' \; et \; i \neq i' \; et \; j = j' \\ \sigma_\pi^2 + \sigma_{\alpha\pi}^2 + \sigma_{\beta\pi}^2 \quad \text{si } s = s' \; et \; i = i' \; et \; j = j' \end{cases}$$

que :

$$cor(Y_{sij}, Y_{s'i'j'}) = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_{\alpha\pi}^2 + \sigma_{\beta\pi}^2 + \sigma^2}$$

$$cor(Y_{sij}, Y_{s'i'j}) = \frac{\sigma_\pi^2 + \sigma_{\alpha\pi}^2}{\sigma_\pi^2 + \sigma_{\alpha\pi}^2 + \sigma_{\beta\pi}^2 + \sigma^2}$$

$$cor(Y_{sij}, Y_{s'ij'}) = \frac{\sigma_\pi^2 + \sigma_{\beta\pi}^2}{\sigma_\pi^2 + \sigma_{\alpha\pi}^2 + \sigma_{\beta\pi}^2 + \sigma^2}$$

Dans le cas du modèle multivarié on considère que les données sont les réalisations de n vecteurs aléatoires de dimension rc. $Y_s=(Y_{s1}, Y_{s2}, \ldots, Y_{src})$ indépendants et de même loi normale caractérisés par : $E(Y_{sk})=\mu_{ij}$ $\quad$ $\mathrm{Var}(Y_{sk})=\sigma^2_{ij}$ $\quad$ $\mathrm{cov}(Y_{sk}, Y_{sk'}) = \mathrm{cov}_{kk'}$

Le modèle mixte univarié est donc un cas particulier du modèle multivarié correspondant aux hypothèses de circularité de la <u>matrice de variance-covariance</u> Σ de la forme :

		A_1			…	A_i					…	A_r		
		B_1	…	B_c	…	B_1	…	B_j	…	B_c	…	B_1	…	B_c
A_1	B_1	$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$												
	…													
	B_c	$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$										
…														
A_i	B_1	$\sigma^2_\pi+\sigma^2_{\pi\beta}$		σ^2_π		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$								
	…													
	B_j	σ^2_π		σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$						
	…													
	B_c	σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$				
…														
A_r	B_1	$\sigma^2_\pi+\sigma^2_{\pi\beta}$		σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		σ^2_π		σ^2_π		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$		
	…													
	B_c	σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		σ^2_π		σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$

		B_1			…	B_j					…	B_c		
		A_1	…	A_r		A_1	…	A_i	…	A_r		A_1	…	A_r
B_1	A_1	$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$												
	…													
	A_r	$\sigma^2_\pi+\sigma^2_{\pi\beta}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$										
…														
B_j	A_1	$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		σ^2_π		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$								
	…													
	A_i	σ^2_π		σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$						
	…													
	A_r	σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$				
…														
B_c	A_1	$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		σ^2_π		σ^2_π		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$		
	…													
	A_r	σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		σ^2_π		σ^2_π		$\sigma^2_\pi+\sigma^2_{\pi\alpha}$		$\sigma^2_\pi+\sigma^2_{\pi\beta}$		$\sigma^2_\pi + \sigma^2_{\pi\alpha} + \sigma^2_{\pi\beta} + \sigma^2$

V-1-6 Exemple :

Un expérimentateur veut étudier l'effet de la consommation de lécithine sur les troubles de mémoire. Il choisit 4 sujets auxquels il administre un traitement quotidien. Au bout d'un mois, de deux mois et de six mois de traitements, il fait passer à chaque sujet deux tests. Le premier test (test 1) est le même chaque mois, le deuxième test (test 2) est une forme parallèle chaque mois.

	Test 1			Test 2		
Sujet	M_1	M_2	M_6	M_1	M_2	M_6
1	10	11	9	3	6	3
2	18	20	17	16	20	14
3	6	8	8	5	6	3
4	4	9	9	10	10	6

Pratique des calculs :

	Test 1				Test 2				
Sujet	M_1	M_2	M_6		M_1	M_2	M_6		
1	10	11	9	$\bar{y}_{110}=10$	3	6	3	$\bar{y}_{120}=4$	$\bar{y}_{100}=7$
2	18	20	17	$\bar{y}_{210}=18,33$	16	20	14	$\bar{y}_{220}=16,66$	$\bar{y}_{200}=17,5$
3	6	8	8	$\bar{y}_{310}=7,33$	5	6	3	$\bar{y}_{320}=4,66$	$\bar{y}_{300}=6$
4	4	9	9	$\bar{y}_{410}=7,33$	10	10	6	$\bar{y}_{420}=8,66$	$\bar{y}_{400}=8$
	$\bar{y}_{011}=9,5$	$\bar{y}_{012}=12$	$\bar{y}_{013}=10,75$		$\bar{y}_{021}=8,5$	$\bar{y}_{022}=10,5$	$\bar{y}_{023}=6,5$		
	$\bar{y}_{010}=10,75$				$\bar{y}_{020}=8,5$				

$\bar{y}_{001}=9$

$\bar{y}_{002}=11,25$

$\bar{y}_{003}=8,625$

$\bar{y}=8$

Pratique des calculs : autre méthode

A \ B	M_1	M_2	M_6	moyennes
T_1	9,50	12,00	10,75	10,75
T_2	8,50	10,50	6,50	8,50
moyennes	9,00	11,25	8,625	9,625

A \ S	1	2	3	4
T_1	10,00	18,33	7,33	7,33
T_2	4,00	16,67	4,67	8,67
moyennes	7,00	17,50	6,00	8,00

B \ S	1	2	3	4
M_1	6,50	17,00	5,50	7,00
M_2	8,50	20,00	7,00	9,50
M_6	6,00	15,50	5,50	7,50
moyennes	7,00	17,50	6,00	8,00

Calcul des sommes des carrées :

SCT=645,625 **SCS**=508,125 **SCA**=30,375 **SCB**=32,25

SC(A×B)=74,875 **SC**(AB)=12,25 **SC**(A×S)=579,5652 **SC**(AS)=41,133

SC(B×S)=546,125 **SC**(BS)=5,75 **SCR**=15,742

Tableau d'analyse de la variance :

Source	$\mathbf{SC}_{obs}$	ddl	$\mathbf{MC}_{obs}$	$\mathbf{F}_{obs}$
Inter S	$\mathbf{SCS}_{obs}$=508,125	n - 1=3		
Intra S				
A	$\mathbf{SCA}_{obs}$=30,375	r - 1=1	$\mathbf{MCA}_{obs}$=30,375	$\mathbf{F}_{A\ obs}$=2,215
AS	$\mathbf{SC(AS)}_{obs}$=41,133	$(r$ - 1$)(n$ - 1$)$=3	$\mathbf{MC(AS)}_{obs}$=13,711	
B	$\mathbf{SCB}_{obs}$=32,25	c - 1=2	$\mathbf{MCB}_{obs}$=16,125	$\mathbf{F}_{B\ obs}$=16,827
BS	$\mathbf{SC(BS)}_{obs}$=5,75	$(c$ - 1$)(n$ - 1$)$=6	$\mathbf{MC(BS)}_{obs}$=0,9583	
AB	$\mathbf{SC(AB)}_{obs}$=12,25	$(r$ - 1$)(c$ - 1$)$=2	$\mathbf{MC(AB)}_{obs}$=6,125	$\mathbf{F}_{AB\ obs}$=2,334
R	$\mathbf{SCR}_{obs}$=15,742	$(r$ - 1$)(c$ - 1$)(n$ - 1$)$=6	$\mathbf{MCR}_{obs}$=2,624	
Total	$\mathbf{SCT}_{obs}$=645,625	N - 1=23		

Test de l'existence des différents effets :

Test 1 : Effet principal du facteur *A*
hypothèse nulle $\mathbf{H}_{\mathbf{0}\mathrm{A}}$: pas d'effet principal du facteur A $\Leftrightarrow (\forall i)\ \alpha_i = 0$ contre
l'hypothèse alternative : $\mathbf{H}_{\mathbf{1}\mathrm{A}}$: il existe un effet principal du facteur A $\Leftrightarrow (\exists i)\ \alpha_i \neq 0$

Théorème : sous l'hypothèse $\mathbf{H}_{\mathbf{0}\mathrm{A}}$, la statistique $\mathbf{F_A} = \dfrac{\mathbf{MCA}}{\mathbf{MC(AS)}}$ suit une loi de Fischer à r – 1=1 et $(n$ - 1$)(r$ - 1$)$=3 degrés de liberté.

Le test1 est alors défini au seuil de signification α=5% par la règle de décision suivante :
si $\mathbf{F}_{A\ obs} \geq \lambda_\alpha$=10,1 alors on rejette l'hypothèse nulle
où λ_α est donné par l'équation : $\alpha = P(\mathbf{F}_A \geq \lambda_\alpha)$.
Or $\mathbf{F}_{A\ obs}$ =2,215 < λ_α=10,1 on en conclue qu'il n'y a pas d'effet principal du facteur *A*

Test 2 : Effet principal du facteur *B*
hypothèse nulle $\mathbf{H}_{\mathbf{0}\mathrm{B}}$: pas d'effet principal du facteur B $\Leftrightarrow (\forall j)\ \beta_j = 0$ contre
l'hypothèse alternative : $\mathbf{H}_{\mathbf{1}\mathrm{B}}$: il existe un effet principal du facteur B $\Leftrightarrow (\exists j)\ \beta_j \neq 0$

Théorème : sous l'hypothèse $\mathbf{H}_{\mathbf{0}\mathrm{B}}$, la statistique $\mathbf{F_B} = \dfrac{\mathbf{MCB}}{\mathbf{MC(BS)}}$ suit une loi de Fischer à c – 1=2 et $(n$ - 1$)(c$ - 1$)$=6 degrés de liberté.

Le test2 est alors défini au seuil de signification α par la règle de décision suivante :
si $\mathbf{F}_{B\ obs} \geq \lambda_\alpha$ =5,14 alors on rejette l'hypothèse nulle
où λ_α est donné par l'équation : $\alpha = P(\mathbf{F}_B \geq \lambda_\alpha)$.
Or $\mathbf{F}_{B\ obs}$ =16,827 > λ_α=5,14 on en conclue qu'il y a un effet principal du facteur *B*

Test 3 : Test de l'effet d'interaction des facteurs *A* et *B*
hypothèse nulle $\mathbf{H}_{\mathbf{0}\mathrm{AB}}$: pas d'effet d'interaction $\Leftrightarrow (\forall i,j)\ (\alpha\beta)_{ij} = 0$ contre
l'hypothèse alternative : $\mathbf{H}_{\mathbf{1}\mathrm{AB}}$: il existe une interaction $\Leftrightarrow (\exists i,j)\ (\alpha\beta)_{ij} \neq 0$

Théorème : sous l'hypothèse $\mathbf{H_0}$AB, la statistique $\mathbf{F_{AB}} = \dfrac{\mathbf{MC(AB)}}{\mathbf{MCR}}$ suit une loi de Fischer à $(r - 1)(c - 1)=2$ et $(n - 1)(c - 1)(r - 1)=6$ degrés de liberté.

Le test3 est alors défini au seuil de signification α par la règle de décision suivante :

si $\mathbf{F}_{AB\ obs} \geq \lambda_\alpha = 5{,}14$ alors on rejette l'hypothèse nulle

où λ_α est donné par l'équation : $\alpha = P(\mathbf{F}_{AB} \geq \lambda_\alpha)$.

Or $\mathbf{F}_{AB\ obs} = 2{,}334 < \lambda_\alpha = 5{,}14$ on en conclue qu'il n'y a pas d'effet d'interaction des facteurs *A* et *B*

Condition de validation du modèle :

Calculons la matrice de variances covariances empirique (et aussi la matrice des corrélations) pour examiner la condition de validité du modèle mixte univarié. Les résultats sont présentés sous forme de blocs homogènes (qui devraient l'être).

Matrice de Variances Covariances

	A	*Test1*			*Test2*		
A	*B*	M_1	M_2	M_6	M_1	M_2	M_6
Test1	M_1	28,750					
	M_2	24,500	22,500				
	M_6	17,875	17,000	13,188			
Test2	M_1	16,250	18,750	15,875	25,250		
	M_2	24,250	18,750	20,125	27,750	32,750	
	M_6	19,250	19,750	15,875	21,750	25,750	20,250

Matrice des corrélations.

	A	*Test 1*			*Test 2*		
A	***B***	M_1	M_2	M_6	M_1	M_2	M_6
Test 1	M_1	1					
	M_2	0,9633	1				
	M_6	0,9180	0,9869	1			
Test 2	M_1	0,6031	0,7866	0,8700	1		
	M_2	0,7903	0,7276	0,9684	0,9650	1	
	M_6	0,7978	0,9253	0,9714	0,9619	0,9999	1

Le premier bloc (en haut et à gauche) est le bloc des variables issues des modalités de *B* alors que *A* prend la modalité « test1 ». Le deuxième bloc (en bas et à droite) est le bloc des variables issues des modalités de *B* alors que *A* prend la modalité « test2 ». Et le troisième bloc (en bas et à gauche) est le bloc des variables issues des modalités de *B* et aussi des modalités de *A* (test1 et test2) à la fois. Dans ce dernier bloc, on distingue les éléments de la diagonale, dont les modalités de *B* sont les mêmes des éléments hors diagonale.

Nous constatons qu'à l'intérieur des blocs nous avons, (plus ou moins) une homogénéité des covariances (et aussi des corrélations). En ce qui concerne la condition d'homoscédasticité (homogénéité des variances), nous ne constatons pas une différence importante sur la diagonale principale de la matrice, donc nous pouvons penser que cette condition est également acquise.