

L'ANALYSE DE LA VARIANCE

Ph. COURCOUX
ENITIAA
Domaine de la Géraudière
44072 NANTES

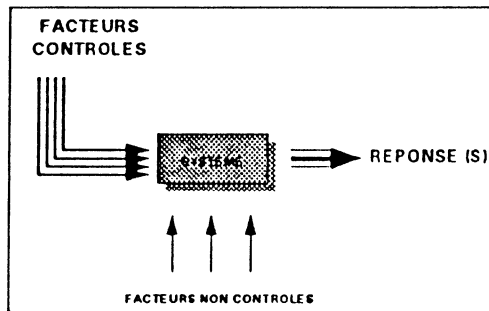
PLAN

Analyse de la variance à un facteur	p. 2
Analyse de la variance à deux facteurs	p. 9
Tests de comparaisons multiples	p. 19

"L'analyse de la variance consiste à analyser des données qui dépendent de plusieurs types d'effets opérant simultanément, afin de quantifier ces effets, et éventuellement, d'en évaluer l'importance "

M.SCHEFFE.

On utilise l'analyse de variance pour analyser un système expérimental sous la dépendance d'un ou plusieurs facteurs contrôlés. L'objet de l'analyse est de comparer la variabilité due aux facteurs à la variabilité résiduelle (ou bruit ou erreur expérimentale).



En amont du traitement, on a donc un plan expérimental (plan d'expériences) qui prévoit le nombre d'expériences à faire pour chaque niveau (ou modalité) du ou des facteurs dont on souhaite évaluer l'effet.

Le plan est dit équilibré si il y a même nombre d'observations par niveau de facteur.

I - ANALYSE DE LA VARIANCE A 1 FACTEUR.

I.1 - GENERALITES.

C'est le cas le plus simple rencontré lorsqu'il n'y a qu'un facteur agissant sur les résultats. Le schéma est alors le suivant :

	niveaux du facteur			
	G ₁	G ₂G _i	G _I
observations	x ₁₁	x ₂₁	x _{I1}
	x ₁₂	x ₂₂	x _{I2}
	x _{1n1}	x _{2n2}	x _{InI}
Nb d'observations	n ₁	n ₂	n _I
Moyennes	\bar{x}_1	\bar{x}_2	\bar{x}_I

La moyenne empirique générale peut être calculée,

soit par :
$$\bar{x} = \frac{1}{n} \sum_i \sum_j x_{ij}$$
 avec $n = \sum_i n_i$

ou par :
$$\bar{x} = \frac{1}{n} \sum_i n_i \bar{x}_i$$

Exemple

Considérons une étude portant sur l'appréciation sensorielle de la texture de trois viandes. Il a été prévu une séance de dégustation par viande et 5 dégustateurs différents ont été convoqués à chaque séance. On ne considérera que le caractère fibreux de la viande. Pour ce caractère, les échantillons ont été notés en utilisant une échelle en 15 points.

Comme il n'a pas été possible de réunir les mêmes juges à chacune des séances, l'influence du facteur "Juge" sur les valeurs données ne pourra être étudiée. Les 15 évaluations seront considérées comme provenant de 15 dégustateurs différents, l'attribution d'un juge à un échantillon se faisant aléatoirement. On parle d'un plan en randomisation totale.

Les valeurs collectées sont les suivantes :

notes pour le caractère fibreux

viande		
A	B	C
3	10	13
5	8	11
6	5	7
3	7	11
3	5	8
$\bar{x}_1 = 4$	$\bar{x}_2 = 7$	$\bar{x}_3 = 10$

$$\bar{x} = 7$$

I.2 - MODELE.

Le modèle de l'analyse de la variance à un facteur peut s'écrire :

$X_{ij} = \mu + \alpha_i + \epsilon_{ij}$

où dans l'exemple des viandes

X_{ij} est la variable à expliquer*Texture fibreuse d'une viande*
 α_i est l'effet du i^{ieme} niveau du facteur.....*caractère fibreux de la viande i*
 μ est l'effet moyen général.....*caractère fibreux potentiel*

et ε_{ij} est la variable aléatoire résiduelledue à l'ensemble des causes qui déterminent la note fibreuse d'une viande autre que la nature de la viande (âge de l'animal, condition de conservation, appréciation du juge.....)

Hypothèses . On suppose que, pour tout couple (i,j) :

$$\begin{aligned} &\text{Les } \varepsilon_{ij} \text{ sont indépendants} \\ &E(\varepsilon_{ij}) = 0 \\ &\text{Var}(\varepsilon_{ij}) = \sigma^2 \\ &\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

I.3 - ESTIMATION DES PARAMETRES DU MODELE (μ, α) .

En choisissant de travailler par rapport à l'effet moyen général, c'est-à-dire en posant comme contrainte :

$$\sum_i \alpha_i = 0 \quad \text{dans le cas d'un plan équilibré}$$

ou
$$\sum_i (n_i \alpha_i) = 0 \quad \text{dans le cas d'un plan non équilibré,}$$

on estime les paramètres μ et α_i ($i=1, \dots, I$) du modèle par : $\mu = \bar{\bar{x}}$ et $\alpha_i = \bar{x}_i - \bar{\bar{x}}$ pour tout i

I.4 - DECOMPOSITION DE LA VARIABILITE.

Décomposons l'élément x_{ij} en deux termes :

$$x_{ij} = \bar{x}_i + (x_{ij} - \bar{x}_i)$$

c'est-à-dire en moyenne de la $i^{\text{ème}}$ colonne (parfois écrite \bar{x}_i) plus écart de la valeur individuelle à la moyenne de la colonne.

La moyenne de la colonne i peut de la même manière être décomposée en deux termes :

$$\bar{x}_i = \bar{\bar{x}} + (\bar{x}_i - \bar{\bar{x}})$$

soit en moyenne globale $\bar{\bar{x}}$ (parfois écrite $\bar{x}_{..}$) plus la différence de la moyenne de colonne à la moyenne globale

On obtient ainsi :

$$x_{ij} = \bar{\bar{x}} + (\bar{x}_i - \bar{\bar{x}}) + (x_{ij} - \bar{x}_i)$$

Ce développement effectué pour chaque case du tableau de données conduit à l'écriture matricielle suivante :

$$\begin{bmatrix} 3 & 10 & 13 \\ 5 & 8 & 11 \\ 6 & 5 & 7 \\ 3 & 7 & 11 \\ 3 & 5 & 8 \end{bmatrix} = \begin{bmatrix} 7 & 7 & 7 \\ 7 & 7 & 7 \\ 7 & 7 & 7 \\ 7 & 7 & 7 \\ 7 & 7 & 7 \end{bmatrix} + \begin{bmatrix} -3 & 0 & 3 \\ -3 & 0 & 3 \\ -3 & 0 & 3 \\ -3 & 0 & 3 \\ -3 & 0 & 3 \end{bmatrix} + \begin{bmatrix} -1 & 3 & 3 \\ 1 & 1 & 1 \\ 2 & -2 & -3 \\ -1 & 0 & 1 \\ -1 & -2 & -2 \end{bmatrix}$$

$$X = M + B + W$$

$$\text{Données} = \text{Moyennes globales} + \text{Ecart Inter-colonnes} + \text{Ecart Intra-colonne}$$

On a, matriciellement : $(X - M) = B + W$

ou algébriquement : $(x_{ij} - \bar{x}) = (\bar{x}_i - \bar{x}) + (x_{ij} - \bar{x}_i)$ (1)

les termes de cette équation étant les estimations de :

$$(X_{ij} - \mu) = \alpha_i + \epsilon_{ij}$$

En élevant au carré et en sommant, pour toutes les observations, les termes de l'équation (1) on a :

$$\sum_i \sum_j (x_{ij} - \bar{x})^2 = \sum_i n_i (\bar{x}_i - \bar{x})^2 + \sum_i \sum_j (x_{ij} - \bar{x}_i)^2$$

Explicitement cette égalité se traduit par :

Somme des carrés des écarts totaux	=	Somme des carrés des écarts inter-niveaux	+	Somme des carrés des écarts intra-niveau
------------------------------------	---	---	---	--

Dans la suite, les termes de cette égalité seront notés respectivement

SCE_{Totale}

SCE_{Inter}

et SCE_{Intra} .

Vérification sur l'exemple :

$$\begin{aligned} SCE_{\text{Totale}} &= 140 \\ SCE_{\text{Inter}} &= 90 \\ SCE_{\text{Intra}} &= 50 \end{aligned}$$

I.5 - TEST D'HYPOTHESE.

On souhaite tester les hypothèses :

H_0 : Il n'y a pas d'effet "Produit",
i.e.: les moyennes pour les différents produits (niveaux du facteur) sont égales.
 soit $\alpha_1 = \alpha_2 = \dots = \alpha_I$

contre H_1 : Il y a un effet dû au produit,
 ou deux moyennes au moins sont différentes.

Le principe du test est le suivant :

Si les différences entre produits (entre niveaux du facteur ou entre colonnes du tableau de données) sont grandes par rapport aux écarts intra-produit, alors on conclura qu'il y a un effet différentiel en fonction du produit.

Il s'agit donc de comparer la variabilité inter-niveaux à la variabilité intra-niveau du facteur, en faisant intervenir les quantités SCE_{Inter} et SCE_{Intra} .

Reprenons la décomposition :

$$SCE_{Totale} = SCE_{Inter} + SCE_{Intra}$$

équation de l'analyse de la variance

A chacune de ces quantités SCE, on va associer le nombre de degrés de liberté (ddl) adéquat :

source de variation	SCE	ddl
Inter-niveaux	SCE Inter	I - 1
Intra-niveau	SCE Intra	n - I
Totale	SCE Totale	n - 1

On définit le carré moyen inter-groupes par :

$$CM_{Inter} = \frac{SCE_{Inter}}{I - 1}$$

De même, on définit le carré moyen intra-groupes par :

$$CM_{Intra} = \frac{SCE_{Intra}}{n - 1}$$

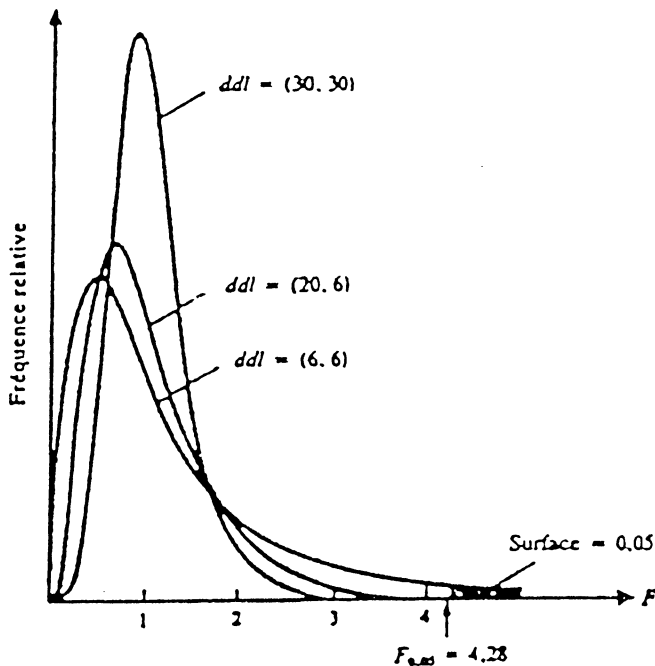
Pour tester H_0 contre H_1 , on évalue la quantité :

$$F = \frac{CM_{Inter}}{CM_{Intra}}$$

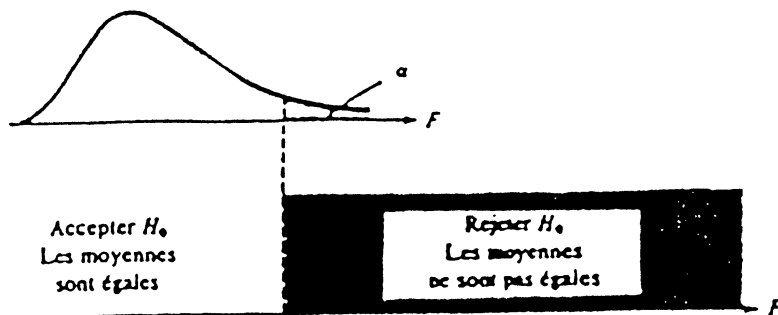
Si l'hypothèse H_0 est vraie, la valeur F est voisine de 1. Plus on s'éloigne de cette hypothèse, en faveur de l'hypothèse H_1 , plus le rapport F augmente. Il reste à déterminer à partir de quelle valeur observée de F on rejettera H_0 . Cette prise de décision est fondée sur la connaissance de la loi de F sous H_0 .

Dans la mesure où les résidus du modèle d'analyse de la variance suivent une loi normale et si H_0 est vraie, on sait que F est l'observation d'une variable qui suit une loi de Fisher ayant $(I-1)$ ddl au numérateur et $(n-I)$ ddl au dénominateur (notée $F_{(I-1, n-I)}$).

L'allure de la fonction de densité de la distribution de Fisher est la suivante :



Pour un niveau de signification donné α (risque de 1er espèce), on adopte la stratégie de décision suivante:



Exemple

TABLEAU D'ANALYSE DE LA VARIANCE

source de variation	SCE	ddl	CM	F
Type de viande	90	2	45	10,80
Residuelle (Intra-produit)	50	12	4,17	($p < 0,01$)
Totale	140	14		

EXEMPLE

Les teneurs en tannins de 4 barils de vin californien (Cabernet Sauvignon) provenant de producteurs différents (A, B, C, D) sont comparés. Ces teneurs sont évaluées sur une échelle de 0 à 30 (30 représentant la plus grande teneur).

Les scores sont les suivants :

vin A	vin B	vin C	vin D
8	9	8	1
6	7	5	2
5	6	6	1
7	8	6	0
6	8	7	0
7	7	7	2
7	8	7	0
8	8	5	6
5	7	8	
6	9	59	
7	77		
7			
79			
$n = 12$	$n = 10$	$n = 9$	$n = 7$
$X = 6.58$	$X = 7.7$	$X = 6.56$	$X = 0.86$

$$T = 221$$

$$N = 38$$

D'après M. O'MAHONY - *Sensory Evaluation of Food*

Ces observations permettent de construire le tableau d'analyse de la variance suivant ;

Source	SCE	ddl	CM	F
Totale	253.71	37		
Entre les vins	219.61	3	73.2	73.2***
Erreur	34.10	34	1.0	

Pour 3 et 34 degrés de liberté, les tables de Fisher donnent :

$$\text{Prob}(F > 2.88) = 0.05$$

$$\text{Prob}(F > 4.42) = 0.01$$

$$\text{Prob}(F > 7.05) = 0.001$$

La valeur F calculée est de 73.2. On en déduit qu'il existe une différence de teneur moyenne en tannin entre barils très significative.

Un test LSD, de comparaisons multiples, au niveau de signification de 0.001, est utilisé afin de répondre à la question : Lequel (ou lesquels) de ces vins diffère(nt) des autres ?

$$\text{LSD} = t \sqrt{MSE \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

	Vin			
	D	C	A	B
Moyenne	0.86	6.56	6.58	7.7
n	7	9	12	10

Les vins A, B et C ne sont pas significativement différents les uns des autres quant à leur teneur en tannins, alors que le vin D obtient des scores significativement plus petits que les autres ($\alpha = 0.001$).

II - ANALYSE DE LA VARIANCE A 2 FACTEURS CROISES.

Dans l'exemple traité en ANOVA à un facteur, portant sur la texture de plusieurs viandes, les séries de mesures en colonne étaient considérées comme étant indépendantes. Ce dispositif complètement randomisé présente un inconvénient majeur: les conditions expérimentales d'évaluation sont différentes d'un type de viande à un autre (juges différents, séances différentes).

Il aurait été préférable de faire intervenir cinq juges, au cours d'une même séance, en leur présentant dans un ordre aléatoire un échantillon de chaque viande. Les séries de mesures sont dans ce cas liées et l'effet "Juge" peut être évalué.

Le plan d'expérience considéré est un dispositif à deux facteurs croisés, complet et sans répétition. Nous considérerons dans un deuxième temps le cas de dispositifs avec répétitions.

D'une manière générale, l'analyse à deux facteurs croisés est mise en oeuvre dans les situations se présentant comme suit :

	1 i I
1	
.	
.	
j	
.	
J	n_{ij}
	$n_{i.}$

A et B étant deux facteurs respectivement à I et J niveaux.
 n_{ij} étant le nombre d'observations pour le niveau i de A et le niveau j de B

Vocabulaire

- * On dit que le dispositif factoriel est croisé lorsque les 2 facteurs jouent le même rôle. Si l'un est subordonné à l'autre, le modèle est dit hiérarchisé.
- * Un modèle est dit complet quand il y a au moins une observation par couple (i,j) ($n_{ij} \geq 1$).
- * Si le nombre d'observations $n_{ij} = 1$, pour tout couple (i,j), le dispositif est dit sans répétition. si $n_{ij} > 1$, il s'agit d'un dispositif avec répétitions. Seul un plan d'expérience avec répétitions permet de tester l'interaction entre les facteurs.
- * Lorsque, pour tout couple (i,j), $n_{ij} = r$, constant, le dispositif est dit équilibré.

II.1 - DISPOSITIF COMPLET SANS REPETITION.

Reprenons la comparaison du caractère fibreux de trois viandes A, B et C et supposons que les mêmes juges aient évalué une série de 3 échantillons . Les valeurs collectées se présenteraient ainsi :

notes pour le caractère fibreux

		viande			
		A	B	C	
juges	1	3	10	13	$\bar{x}_{.1} = 8.7$
	2	5	8	11	$\bar{x}_{.2} = 8$
	3	6	5	7	$\bar{x}_{.3} = 6$
	4	3	7	11	$\bar{x}_{.4} = 7$
	5	3	5	8	$\bar{x}_{.5} = 5.3$
		$\bar{x}_1 = 4$	$\bar{x}_2 = 7$	$\bar{x}_3 = 10$	$\bar{\bar{x}} = 7$

Si l'effet des facteurs est additif, le modèle est le suivant :

$$X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

où *dans l'exemple de la viande*

X est la variable à expliquer*Texture fibreuse d'une viande*
 α_i est l'effet du $i^{i\text{ème}}$ niveau du facteur A...*caractère fibreux de la viande i*
 β_j est l'effet du $j^{j\text{ème}}$ niveau du facteur B...*effet lié au juge j*
 μ est l'effet moyen général.....*caractère fibreux potentiel*

et ϵ_{ij} est la variable aléatoire résiduelle*dûe à l'ensemble des causes qui déterminent la note fibreuse d'une viande autre que la nature de la viande et l'appréciation du juge.*

Avec les hypothèses :

Pour tout couple (i,j), ϵ_{ij} est distribuée selon une loi normale de moyenne nulle et de variance σ^2

Les variables aléatoires ϵ_{ij} sont indépendantes 2 à 2.

L'équation de décomposition de la variance s'écrit dans ce cas :

$$\sum_i \sum_j (x_{ij} - \bar{x}_{..})^2 = \sum_i \sum_j (\bar{x}_{i.} - \bar{x}_{..})^2 + \sum_i \sum_j (\bar{x}_{.j} - \bar{x}_{..})^2 + \sum_i \sum_j (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2$$

ou encore

$$\sum_i \sum_j (x_{ij} - \bar{x}_{..})^2 = J \sum_i (\bar{x}_{i.} - \bar{x}_{..})^2 + I \sum_j (\bar{x}_{.j} - \bar{x}_{..})^2 + \sum_i \sum_j (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2$$

soit

$$SCE_T = SCE_A + SCE_B + SCE_R$$

T comme Totale

A comme liée au facteur A

R comme Résiduelle

B comme liée au facteur B

D'où le tableau d'analyse de la variance à 2 facteurs croisés sans répétition :

source de variation	SCE	ddl
Effet principal de A	$SCE_A = J \sum_i (\bar{x}_i - \bar{x}_.)^2$	I - 1
Effet principal de B	$SCE_B = I \sum_j (\bar{x}_j - \bar{x}_.)^2$	J - 1
Résiduelle	$SCE_R = SCE_T - (SCE_A + SCE_B)$	(I-1)(J-1)
Totale	$SCE_T = \sum_i \sum_j (\bar{x}_{ij} - \bar{x}_.)^2$	n - 1

Exemple :

source de variation	SCE	ddl	CM	F
Type de viande	90.00	2	45.00	13.21
Effet Juge	22.73	4	5.68	1.66
Residuelle	27.27	8	3.41	
Totale	140	14		

1er Test

contre H_0 : Absence d'effet du facteur A, c'est-à-dire $\alpha_1 = \alpha_2 = \dots = \alpha_I = 0$
 H_1 : Au moins 2 termes α_i différents

La statistique du test $F = \frac{SCE_A / I - 1}{SCE_R / (I - 1)(J - 1)}$ est distribuée, sous H_0 , comme une loi de

Fisher à (I-1) et (I-1)(J-1) degrés de liberté.

Exemple : $F_{obs} = 13.21 >> F_{\alpha, 2, 8} = 4.46$ avec $\alpha = 5\%$.

On rejette donc H_0 avec un risque de 5% , c'est-à-dire on conclue à l'existence d'un effet "Produit"

2ième Test

contre H_0 : Absence d'effet du facteur B, c'est-à-dire $\beta_1 = \beta_2 = \dots = \beta_J = 0$
 H_1 : Au moins 2 termes β_j différents

La statistique du test $F = \frac{SCE_B / J - 1}{SCE_R / (I - 1)(J - 1)}$ est distribuée, sous H_0 , comme une loi de Fisher à $(J-1)$ et $(I-1)(J-1)$ degrés de liberté.

Exemple : $F_{\text{obs}} = 1.66 < F_{\alpha, 4, 8} = 3.84$ avec $\alpha = 5\%$.

On ne rejette donc pas H_0 , c'est-à-dire qu'aucun effet "Juge" n'a pu être décelé.

EXEMPLE

On considère les données collectées par I. LEGUERINEL pour la caractérisation sensorielle de différents cidres.

La question ici est de savoir, d'une part, si les 10 cidres concernés par l'étude sont significativement différents en ce qui concerne leur caractéristiques sensorielles ou au contraire comparables, et d'autre part si les juges sont homogènes lorsqu'ils notent un même produit. Il s'agit donc de tester l'existence de deux effets : l'effet *produit* et l'effet *juge*.

L'analyse de la variance étant une méthode unidimensionnelle, ce test ne pourra être conduit que pour chacune des variables, prises séparément.

Considérons comme variable à expliquer la :

SAVEUR SUCREE

Tableau d'analyse de la variance. Modèle additif à deux facteurs (juge, produit)

Source de variation	SCE	ddl	CM	Fobs	Niv. signif.
Liée aux cidres	88.52	9	9.83	26.70	0.00
Liée aux juges	1.85	6	0.31	0.90	0.50
Résiduelle	18.51	54	0.34		
Totale	108.88	69			

Intervalles de confiance (LSD) à 95% autour des moyennes par niveau du facteur

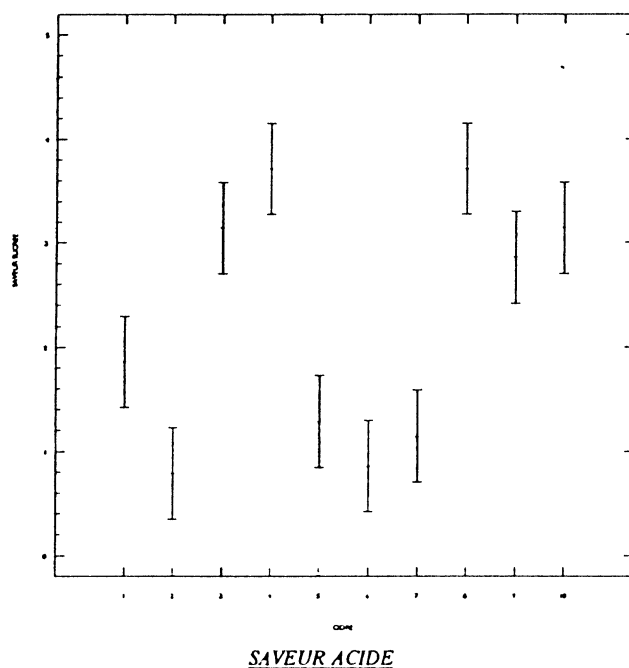
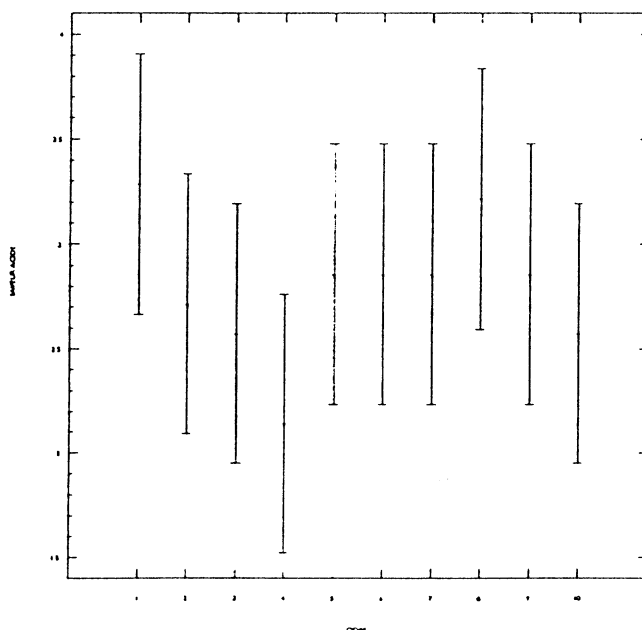


Tableau d'analyse de la variance. Modèle additif à deux facteurs (juge, produit)

Source de variation	SCE	ddl	CM	Fobs	Niv. signif.
Liée aux cidres	6.75	9	0.75	1.77	0.10
Liée aux juges	17.62	6	2.94	6.93	0.00
Résiduelle	22.88	54	0.42		
Totale	47.25	69			

Intervalles de confiance (LSD) à 95% autour des moyennes par niveau du facteur.



Remarque :

Envisageons le cas où, pour savoir si les cidres ont des saveurs acide différentes, on ait omis le facteur "juge".

Le modèle est alors un modèle d'analyse de la variance à un facteur (les cidres). Le tableau de cette analyse est le suivant :

Source de variation	SCE	ddl	CM	Fobs	Niv. signif.
"Cidre" (inter)	6.75	9	0.75	1.11	0.37
Résiduelle (intra)	40.50	60	0.68		
Totale	47.25	69			

II.2 - DISPOSITIF COMPLET AVEC REPETITIONS.

Considérons un plan d'expérience équilibré, avec plusieurs répétitions par croisement de deux facteurs A et B.

Ces répétitions rendent possible une évaluation de l'interaction des deux facteurs.

Plan à 2 facteurs croisés SANS répétition

	A1	A2	A3
B1	x	x	x
B2	x	x	x
B3	x	x	x
B4	x	x	x
B5	x	x	x

Plan à 2 facteurs croisés équilibré AVEC répétitions

	A1	A2	A3
B1	xxx	xxx	xxx
B2	xxx	xxx	xxx
B3	xxx	xxx	xxx
B4	xxx	xxx	xxx
B5	xxx	xxx	xxx

Le tableau de données se présente alors sous la forme suivante :

	1		i		I	moy.
1						$\bar{x}_{.1}$
	x_{1j1} x_{1j2} x_{1jk} $\bar{x}_{1j.}$		x_{ij1} x_{ij2} x_{ijk} $\bar{x}_{ij.}$		x_{Ij1} x_{Ij2} x_{Ijk} $\bar{x}_{Ij.}$	$\bar{x}_{.j}$
J						$\bar{x}_{.J}$
moy.	$\bar{x}_{.1}$		$\bar{x}_{.i}$		$\bar{x}_{.I}$	$\bar{x}_{..}$

Le modèle d'analyse de la variance associé est :

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

- où
- μ est l'effet moyen général
 - α_i ($i=1, \dots, I$) est l'effet principal du facteur A,
 - β_j ($j=1, \dots, J$) est l'effet principal du facteur B,
 - $(\alpha\beta)_{ij}$ est l'interaction des niveaux i et j des 2 facteurs,
 - ϵ_{ijk} est la variable aléatoire résiduelle attachée à la $k^{i\text{eme}}$ ($k=1, \dots, K$) observation pour la combinaison i,j des facteurs.

On suppose que

- les variables ϵ_{ijk} sont indépendantes,
- et pour tout (i,j,k) ϵ_{ijk} suit une loi normale de moyenne nulle et de variance σ^2 .

En procédant comme pour l'analyse de la variance à un critère de classification, l'équation de l'analyse de la variance peut être obtenue en élevant au carré les deux membres du modèle, et en sommant pour toutes les valeurs observées :

$$\sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{..})^2 = JK \sum_i (\bar{x}_{i.} - \bar{x}_{..})^2 + IK \sum_j (\bar{x}_{.j} - \bar{x}_{..})^2 + K \sum_i \sum_j (x_{ij.} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2 + \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{ij.})^2$$

La somme totale des carrés des écarts (SCE_T) se divise donc ici en quatre composantes que l'on peut désigner respectivement par :

SCE_A liée au premier facteur (par exemple le produit),
 SCE_B liée au second facteur (par exemple le juge),
 SCE_{AB} liée à l'interaction des deux facteurs et
 SCE_R qui est une somme résiduelle.

$$SCE_T = SCE_A + SCE_B + SCE_{AB} + SCE_R$$

A ces différentes sommes des carrés des écarts sont attachés des nombres de degrés de liberté, qui vérifient la relation :

$$IJK-1 = (I-1) + (J-1) + (I-1)(J-1) + (IJK-IJ)$$

où $IJK = n$ le nombre total d'observations.

Le tableau de l'analyse de la variance peut enfin être dressé :

source de variation	SCE	ddl
Effet principal de A	$SCE_A = JK \sum_i (\bar{x}_{i.} - \bar{x}_{..})^2$	I - 1
Effet principal de B	$SCE_B = IK \sum_j (\bar{x}_{.j} - \bar{x}_{..})^2$	J - 1
Interaction A X B	SCE_{AB} (par différence)	(I-1)(J-1)
Résiduelle	$SCE_R = \sum_i \sum_j \sum_k (\bar{x}_{ijk} - \bar{x}_{ij.})^2$	IJK - IJ
Totale	$SCE_T = \sum_i \sum_j \sum_k (\bar{x}_{ijk} - \bar{x}_{..})^2$	n - 1

La stratégie des tests d'hypothèses adoptée sera la suivante :

a- Test de l'absence d'interaction $H_0 : (\alpha\beta)_{ij} = 0$ pour tout i et j

La statistique du test $F = \frac{SCE_{AB} / (I-1)(J-1)}{SCE_R / (IJK - IJ)}$ est distribuée, sous H_0 , comme une loi de Fisher à $(I-1)(J-1)$ et $(IJK - IJ)$ degrés de liberté.

b- Test d'absence d'effet du facteur A $H_0 : \alpha_i = 0$ pour tout i

La statistique du test $F = \frac{SCE_A / (I-1)}{SCE_R / (IJK - IJ)}$ est distribuée, sous H_0 , comme une loi de Fisher à $(I-1)$ et $(IJK - IJ)$ degrés de liberté.

c- Test d'absence d'effet du facteur B $H_0 : \beta_j = 0$ pour tout j

La statistique du test $F = \frac{SCE_B / (J-1)}{SCE_R / (IJK - IJ)}$ est distribuée, sous H_0 , comme une loi de Fisher à $(J-1)$ et $(IJK - IJ)$ degrés de liberté.

EXEMPLE

La tenue à la fonte de 4 marques de savon est étudiée à l'aide d'un testeur mécanique. Deux techniciens ont réalisé 3 mesures de suite pour chaque marque.

Les résultats, exprimés en pourcentage, et les totaux marginaux sont fournis dans le tableau suivant :

Techniciens	Marques de savon				Totaux
	A	B	C	D	
1	60.6	73.9	66.1	76.7	822.5
	59.0	74.2	64.4	75.6	
	59.2	73.8	65.0	74.0	
2	59.9	73.9	63.5	73.8	817.9
	61.6	74.0	64.9	73.0	
	62.0	72.3	63.0	76.0	
Totaux	362.3	442.1	386.9	449.1	1640.4
Moyennes	60.4	73.7	64.5	74.9	
Sous-totaux pour les marques suivant le technicien					
1	178.8	221.9	195.5	226.3	
2	183.5	220.2	191.4	222.8	

D'après M. GACULA, J. C. JAGBIR
 Statistical Methods in Food and Consumer Research

Les résultats de l'analyse de la variance sont les suivants :

Source	DF	SS	MS	F ratio
Total	23	921.54		
Savon	3	894.68	298.23	160.33
Technicien	1	0.88	0.88	0.47
Savon x technicien	3	8.12	2.71	1.46
Erreur	16	17.86	1.86	

Pour un modèle à effets fixes, l'hypothèse nulle selon laquelle les différentes marques sont équivalentes est testée en formant le rapport : "MS Savons / MS erreur". La valeur F calculée conduit au rejet de cette hypothèse nulle.

III - TESTS DE COMPARAISONS MULTIPLES

Utilisation :

Tests paramétriques qui interviennent après un test de comparaison globale dans le cas où l'hypothèse nulle a été rejetée; c'est-à-dire lorsqu'on a conclu à une différence significative entre modalités d'un facteur, au risque de première espèce α

Différents tests :

Il existe de nombreux tests de comparaisons multiples. Sans être exhaustif on peut citer :

- * le test de SCHEFFE
- * le test HSD⁽¹⁾ de TUKEY ("Studentized Range Test")
- * le test de NEWMAN-KEULS
- * le test de DUNCAN
- * le test LSD⁽²⁾ de FISHER
- * le test de DUNNETT

Ces tests sont ordonnés ici par ordre de puissance croissante, ou selon un autre critère, du plus conservatif⁽³⁾ au moins conservatif.

On sait que le problème de la comparaison de deux moyennes est résolu par le test "t" de Student: une méthode peut donc consister à comparer les moyennes deux à deux à l'aide de ce test. Mais cela nécessite, pour p moyennes, C_p^2 comparaisons.

Un certain nombre de méthodes permettent d'effectuer ces C_p^2 comparaisons comme par exemple la méthode de la plus petite différence significative (LSD de FISHER), peu correcte mais largement utilisée. Les tests de DUNCAN et de NEWMAN-KEULS sont "similaires au test LSD" mais n'en présentent pas les inconvénients" (GOUET J.P., 1974). Ces inconvénients sont de deux ordres. D'une part, il n'y a pas effectivement indépendance des échantillons comparés, puisque le résultat d'une comparaison est lié aux comparaisons entre moyennes intermédiaires. D'autre part, si on compare p moyennes 2 à 2 au niveau nominal α de 5%, le seuil au niveau global n'est plus nécessairement de 5%, mais augmente avec le nombre de moyennes à comparer.

Dans certains cas, il se peut que quelques comparaisons de moyennes soient intéressantes et d'autres inutiles. Ces situations justifient l'utilisation de tests plus spécifiques comme le test de DUNNETT.

Principe général :

On calcule pour chaque comparaison une valeur seuil r (ou range), fonction, entre autres, du carré moyen des erreurs de l'analyse de la variance. Si la différence entre les moyennes comparées est inférieure à ce seuil r , la différence est déclarée non significative. Traditionnellement, on prend le même niveau de signification α pour les tests de comparaisons multiples que celui pris pour l'ANOVA.

Les tests les plus puissants auront un "r" petit, les plus conservatifs un "r" élevé.

Certains tests de comparaisons multiples font intervenir des seuils r différents suivant le nombre de moyennes situées entre les deux moyennes à comparer. Parmi ces tests dits à plages multiples on trouve les tests de DUNCAN et de NEWMAN-KEULS.

(1) HSD : Honesty Significant Difference.

(2) LSD : Least Significant Difference.

(3) "conservative" en anglais : propriété à déceler une différence au risque que celle-ci n'existe pas.

Les différentes expressions de calcul des seuils r sont les suivantes :

LSD	$t \sqrt{2 CM_E / n}$	t -> table de Student n : nb d'obs.par échantillon CM_E : carré moyen résiduel
SCHEFFE	$S \sqrt{2 CM_E / n}$	S -> $\sqrt{(p-1)F}$ et F -> table de Fisher
DUNNETT (*)	$D \sqrt{2 CM_E / n}$	D -> table de Dunnett
NEWMAN-KEULS (*)	$Q \sqrt{2 CM_E / n}$	Q -> table du "Range Studentized"
DUNCAN (*)	$Q_D \sqrt{2 CM_E / n}$	Q_D -> table du "Range Studentized" adaptée
HSD de TUKEY (*)	$Q_{\max} \sqrt{2 CM_E / n}$	Q_{\max} -> table du "Range Studentized"

(*)test ne s'appliquant que dans le cas de dispositifs équilibrés.

Choix d'un test de comparaisons multiples :

Le choix d'un test n'est pas forcément aisé. Le tableau ci-dessous fournit quelques éléments pour faire ce choix

DUNNETT	pour comparer toutes les moyennes à l'une d'elles. comparaison à un contrôle
SHEFFE	comparaisons simples ou complexes.
HSD de TUKEY	risque α final maintenu au niveau du risque fixé a priori
NEWMAN-KEULS	comparaisons simples. "range" variables. ajustement des niveaux
DUNCAN	nominaux pour un seuil global α choisi.
LSD	comparaisons simples, dans le cas où il y a très peu de comparaisons à faire
