

Analyse factorielle de variance: Principes d'expérimentation

Plans à groupes indépendants: organisation

† Organisation factorielle:
2 variables et plus
produit cartésien des
niveaux

Longueur de la liste	Type de mots	
	Abstraits	Concrets
10 mots	n_{11}	n_{12}
20 mots	n_{21}	n_{22}
30 mots	n_{31}	n_{32}

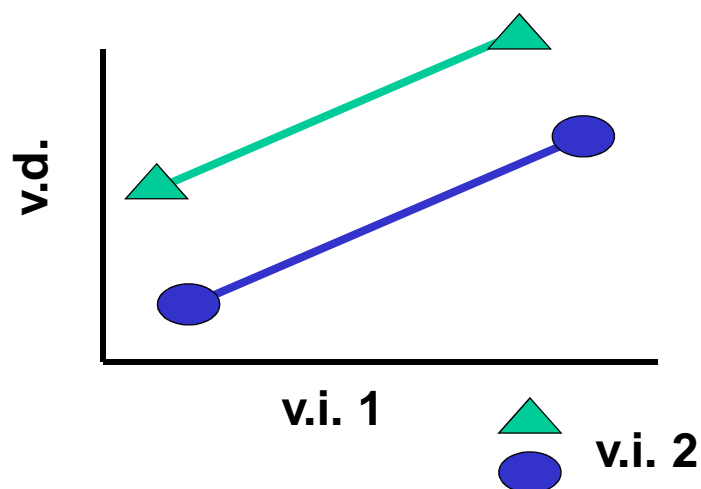
Plans factoriels à groupes indépendants: notion particulière

◆ Effet principal
effet spécifique de la variable indépendante

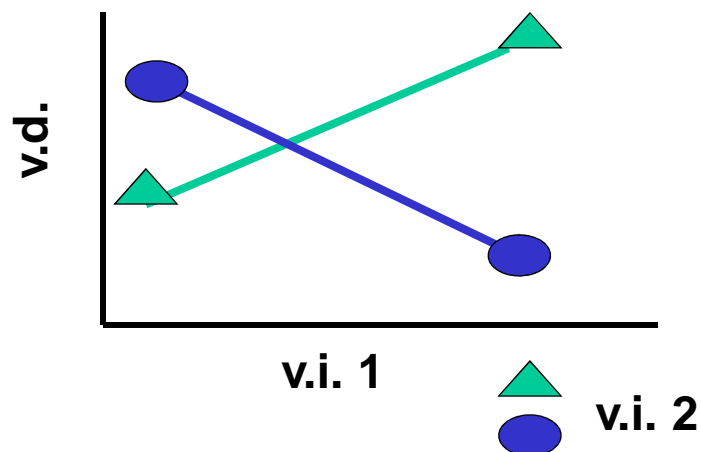
◆ Effet d'interaction

l'effet d'une variable indépendante n'est pas le même à tous les niveaux de l'autre variable indépendante

Graphique d'un effet principal



Graphique d'un effet d'interaction



Rappel (1)

- Laquelle des 2?
- Permutations:
assemblément d'objets quand la séquence d'assemblage fait une différence
- Combinaisons:
assemblément d'objets quand la séquence d'assemblage **ne** fait **pas** de différence

$$P_r^N = \frac{N!}{(N-r)!}$$

$$C_r^N = \frac{N!}{r!(N-r)!}$$

Limite des plans factoriels: multiplication des facteurs

- Si pour 2 facteurs, il y a 1 interaction;
- pour 3 facteurs, il y a 4 interactions;
- pour 4 facteurs, il y a 11 interactions;
- pour 5 facteurs, il y a 26 interactions.
- pour X ..., ...

Rappel (2)

Au début de tout se trouve la variance,

$$\frac{\sum (X - \bar{X})^2}{N-1}$$

**soit l'éloignement d'une mesure (Xi)
par rapport à sa tendance centrale (\bar{X})**

Rappel (3)

Cette variance sera décomposée en deux sources de variance, chacune décomposable.

Une nouvelle notation sera utilisée

$$SC_{\text{total}} = SC_{\text{traitement}} + SC_{\text{erreur}}$$

Analyse factorielle de variance: Principes statistiques (1)

$$SC_{\text{total}} = SC_{\text{traitement}} + SC_{\text{erreur}}$$

↓ ↓

$$SC_A + SC_B + SC_{A \times B}$$

$$SC_{\text{erreur}} = SC_{\text{total}} - SC_{\text{traitement}}$$

N.B.: traitement s'appelle maintenant
cellules dans Howell

**Analyse factorielle
de variance:
Principes statistiques (2)**

2 soustractions

$$\mathbf{SC}_{\text{erreur}} = \mathbf{SC}_{\text{total}} - \mathbf{SC}_{\text{traitement}}$$

$$\mathbf{SC}_{\text{AxB}} = \mathbf{SC}_{\text{traitement}} - \mathbf{SC}_{\text{A}} - \mathbf{SC}_{\text{B}}$$

**Calcul de l'analyse de variance:
les sommes de carré (1)**

$$\mathbf{SC}_{\text{Total}} = \sum \left(\mathbf{X} - \bar{\mathbf{X}} \right)^2 = \sum \mathbf{X}^2 - \frac{(\sum \mathbf{X})^2}{\mathbf{N}}$$

$$\mathbf{SC}_{\text{traitement}} = \sum \left(\bar{\mathbf{X}}_j - \bar{\mathbf{X}} \right)^2 = \frac{\sum \mathbf{T}_j^2}{\mathbf{n}} - \frac{(\sum \mathbf{X})^2}{\mathbf{N}}$$

$$\mathbf{SC}_{\text{erreur}} = \sum \left(\mathbf{X}_{ij} - \bar{\mathbf{X}}_j \right)^2 = \sum \mathbf{X}^2 - \frac{\sum \mathbf{T}_j^2}{\mathbf{n}}$$

Calcul de l'analyse de variance: les sommes de carré (2)

Longueur de la liste	Type de mots	
	Abstrait	Concret
10 mots	n_{11}	n_{12}
20 mots	n_{21}	n_{22}
30 mots	n_{31}	n_{32}

$$SC_{\text{traitement}} = \sum \left(\bar{X}_j - \bar{X} \right)^2 = \frac{\sum T_j^2}{n} - \frac{(\sum X)^2}{N}$$

$$SC_A = \sum \left(\bar{X}_a - \bar{X} \right)^2 = \frac{\sum T_a^2}{nb} - \frac{(\sum X)^2}{N}$$

$$SC_B = \sum \left(\bar{X}_b - \bar{X} \right)^2 = \frac{\sum T_b^2}{na} - \frac{(\sum X)^2}{N}$$

Calcul de l'analyse de variance: effet de l'interaction

◆ Rappel

$$SC_{A \times B} = SC_{\text{traitement}} - SC_A - SC_B$$

◆ Interprétation d'une interaction significative

- graphique
- comparaisons multiples
- calculs d'effet simple d'une variable indépendante aux niveaux de l'autre variable indépendante

Comparaison des variances: le rapport F

$$F_{(k-1, N-k)} = \frac{CM_{\text{traitement}}}{CM_{\text{erreur}}}$$

où k-1 et N-k sont les dl du rapport F

Calcul de l'analyse de variance: des sommes de carré aux carrés moyens

$$CM_{\text{traitement}} = \frac{SC_{\text{traitement}}}{k-1}$$

$$CM_{\text{erreur}} = \frac{SC_{\text{erreur}}}{N-k}$$

**Analyse factorielle
de variance:
Principes statistiques (3)
les degrés de liberté (dl)**

$$N_{\text{total}} = N - 1 \quad k = a \times b$$

$$dl_A = a - 1 \quad dl_B = b - 1$$

$dl_{\text{interaction}}$ = multiplication des dl des facteurs
composant l'interaction

$$dl_{\text{erreur}} = N - k$$

**Analyse factorielle
de variance:
Principes statistiques (4)
les carrés moyens (CM)**

chaque carré moyen s'obtient par la
division de sa SC par son dl

Le test F est un rapport de variance,
soit le carré moyen d'un traitement
divisé par le carré moyen d'erreur

Rappel (4): les étapes des tests statistiques

- 1) Formulation de l'hypothèse nulle et de l'hypothèse alternative
- 2) Sélection d'un test et du seuil alpha
- 3) Détermination des dl et de la valeur critère
- 4) Calcul
- 5) Décision eu égard à l'hypothèse nulle
- 6) Formulation d'une conclusion

Deux indices de la taille de l'effet

$$\eta^2 = \frac{\sigma_{\text{traitement}}^2}{\sigma_{\text{total}}^2} = \frac{SC_{\text{total}} - SC_{\text{erreur}}}{SC_{\text{total}}} = \frac{SC_{\text{traitement}}}{SC_{\text{total}}}$$

$$\omega^2 = \frac{\sigma_{\text{traitement}}^2}{\sigma_{\text{total}}^2} = \frac{SC_{\text{traitement}} - (k-1) CM_{\text{erreur}}}{SC_{\text{total}} + CM_{\text{erreur}}}$$

Présentation des résultats de l'analyse de variance en style APA

- Tableau, voir manuel p. 129
- Texte, voir p. 117
- exemple:

6^e éd.

*Les participants font plus d'erreurs
(M=10,0) quand ils voient les lettres pour
un court laps de temps que pour une
durée de présentation plus longue
(M=7,0), $F(1, 16) = 18.00, p < .05, \eta^2 = 0.21$.*

Un exemple de calcul

Du schème à 1 facteur
au schème à 2 facteurs

Problème 11.3

Jeunes/ - élevé	Jeunes/ + élevé	Vieux / - élevé	Vieux / + élevé
8	21	9	10
6	19	8	19
4	17	6	14
6	15	8	5
7	22	10	10
6	16	4	11
5	22	6	14
7	22	5	15
9	18	7	11
7	21	7	11
65	193	70	120

Problème 11.3: les carrés

Jeunes/ - élevé	Jeunes/ + élevé	Vieux / - élevé	Vieux / + élevé
64	441	81	100
36	361	64	361
16	289	36	196
36	225	64	25
49	484	100	100
36	256	16	121
25	484	36	196
49	484	25	225
81	324	49	121
49	441	49	121
6316			

$$SC_{\text{total}} = \sum \left(X - \bar{X} \right)^2 = \sum X^2 - \frac{(\sum X)^2}{N}$$

$$SC_{\text{traitement}} = \sum \left(\bar{X}_j - \bar{X} \right)^2 = \frac{\sum T_j^2}{n} - \frac{(\sum X)^2}{N}$$

=

=

	A1 Jeunes	A2 Vieux	Total
B1 - élevé			
B2 + élevé			
Total			

**Calcul de l'analyse de
variance:
les sommes de carré (2)**

$$SC_A = \sum \left(\bar{X}_a - \bar{X} \right)^2 = \frac{\sum T_a^2}{nb} - \frac{(\sum X)^2}{N}$$

$$SC_B = \sum \left(\bar{X}_b - \bar{X} \right)^2 = \frac{\sum T_b^2}{na} - \frac{(\sum X)^2}{N}$$

$$SC_{AxB} = SC_{\text{traitement}} - SC_A - SC_B$$

$$SC_A = \sum \left(\bar{X}_a - \bar{X} \right)^2 = \frac{\sum T_a^2}{nb} - \frac{(\sum X)^2}{N}$$

$$SC_B = \sum \left(\bar{X}_b - \bar{X} \right)^2 = \frac{\sum T_b^2}{na} - \frac{(\sum X)^2}{N}$$

$$SC_{AxB} =$$

Le tableau d'analyse de variance

Source de variation	SC	dl	CM	F	p
Âge					
Niveau de traitement					
Âge x Niveau de traitement					
Erreur					
Total					

Au-delà de l'analyse factorielle de variance

Comparaisons multiples: du test t aux autres

$$t_{(dl)} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} \longrightarrow q_{(dl,dl)} = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{\frac{CM_{\text{Erreur}}}{n}}}$$

Schémas à mesures répétées

- considérer l'erreur comme juchée dans les facteurs intersujets
- dl = (n-1)x(ab...)

Un autre exemple de calcul

Un schème à groupes indépendants
(intersujets) $2 \times 2 \times 3$.

Les données

A1						A2					
B1			B2			B1			B2		
C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
4	23	16	21	25	32	6	2	20	11	23	17
18	15	27	14	33	42	4	6	15	7	14	16
8	21	23	19	30	46	14	8	8	6	13	25
10	13	14	26	20	40	7	12	17	16	12	12

Les carrés des valeurs

A1						A2					
B 1			B2			B1			B 2		
C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
16	529	256	441	625	1024	36	4	400	121	529	289
324	225	729	196	1089	1764	16	36	225	49	196	256
64	441	529	361	900	2116	169	64	64	36	169	625
100	169	196	676	400	1600	49	144	289	256	144	144

Grand total
19080

SC_{total}

$SC_{\text{traitement}} =$

$SC_{\text{erreur}} =$

Calculs des totaux des effets principaux de A, B, et C

A1						A2						Grand total 830
B1			B2			B1			B 2			
C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	
4	23	16	21	25	32	6	2	20	11	23	17	
18	15	27	14	33	42	4	6	15	7	14	16	
8	21	23	19	30	46	13	8	8	6	13	25	
10	13	14	26	20	40	7	12	17	16	12	12	
40	72	80	80	108	160	30	28	60	40	62	70	
A1 = 540						A2 = 290						
B1 = 310						B2 = 520						
C1 = 40 + 80 + 30 + 40 = 190						C2 = 72 + 108 + 28 + 62 = 270						
C3 = 80 + 160 + 60 + 70 = 370												

$SC_A =$

$SC_B =$

$SC_C =$

A1						A2					
B1			B2			B1			B 2		
C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
4	23	16	21	25	32	6	2	20	11	23	17
18	15	27	14	33	42	4	6	15	7	14	16
8	21	23	19	30	46	13	8	8	6	13	25
10	13	14	26	20	40	7	12	17	16	12	12
40	72	80	80	108	160	30	28	60	40	62	70

$SC_{intercondAB} =$

$SC_{AxB} =$

A1						A2					
B1			B2			B1			B 2		
C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
4	23	16	21	25	32	6	2	20	11	23	17
18	15	27	14	33	42	4	6	15	7	14	16
8	21	23	19	30	46	13	8	8	6	13	25
10	13	14	26	20	40	7	12	17	16	12	12

40 72 80 80 108 160 30 28 60 40 62 70

Interaction AC

	C1	C2	C3	Total
A1	120	180	240	540
A2	70	90	130	290
Total C	190	270	370	830

$SC_{intercondAC} =$

$SC_{AxC} =$

A1						A2					
B1			B2			B1			B 2		
C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
4	23	16	21	25	32	6	2	20	11	23	17
18	15	27	14	33	42	4	6	15	7	14	16
8	21	23	19	30	46	13	8	8	6	13	25
10	13	14	26	20	40	7	12	17	16	12	12
40 72 80			80 108 160			30 28 60			40 62 70		

Trouvez comment le
tableau a été construit

Interaction BC

	C1	C2	C3	Total
B1	70	100	140	310
B2	120	170	230	520
Total C	190	270	370	830

$SC_{intercondBC} =$

$SC_{AxC} =$

Interaction AB

	B1	B2	Total A
A1	192	348	540
A2	118	172	290
Total B	310	520	830

Cette interaction est significative parce que la différence entre les niveaux de A (540 à 290) n'est pas la même à chaque niveau de B (192 à 118 pour B1 versus 348 à 172 pour B2) ou la différence entre les niveaux de B (310 à 520) n'est pas la même à chaque niveau de A (192 à 348 pour A1 versus 118 à 172 pour A2)

Interaction BC

	C1	C2	C3	Total
B1	70	100	140	310
B2	120	170	230	520
Total C	190	270	370	830

Cette interaction n'est pas significative parce que la différence entre les niveaux de B (310 à 520) est semblable à chaque niveau de C (70 à 120 pour C1 vs 100 à 170 pour C2 vs 140 à 230 pour C3) ou la différence entre les niveaux de C (190 à 270 à 370) est semblable à chaque niveau de B (70 à 100 à 140 pour B1 versus 120 à 170 à 230 pour B2)

Exemple d'effet simple

A_1B_1 9	A_1B_2 7	A_2B_1 5	A_2B_2 13	Σa^2	$\Sigma a\bar{X}$	$(\Sigma a\bar{X})^2$	$n(\Sigma a\bar{X})^2$

Le tableau d'analyse de variance

Source de variation	SC	dl	CM	F	p
Type de lettrage	5	1	5.0	2.00	>.05
Effet simple du T_exp. au niveau du lettrage normal	10	1	10.0	4.00	>.05
Effet simple du T_exp. au niveau du lettrage en ital.	160	1	160.0	64.00	<.05
Erreur	40	16	2.5		
Total	215	19			

Un autre exemple de calcul

Un schème à groupes indépendants
(intersujets) 3 x 2.

Effet de l'intensité de l'exercice et de
son moment sur la durée du sommeil.

	Exercice d'intensité					
	légère		modérée		élevée	
moment	matin	soir	matin	soir	matin	soir
	6,5	7,1	7,4	7,4	8,0	8,2
	7,3	7,9	6,8	8,0	7,6	8,5
	6,6	8,2	6,7	8,1	7,7	8,7
	7,4	7,7	7,3	8,2	6,6	9,5
	7,2	7,5	7,6	7,6	7,1	9,6
	6,8	7,6	7,4	8,0	7,2	9,4
Somme						
Moyenne						
Écart-type						

Somme de tous les X =

Somme de tous les X au carré =

Étape 1 de calcul des sommes de carrés

$$SC_{\text{Total}} = \sum (X - \bar{X})^2 = \sum X^2 - \frac{(\sum X)^2}{N} = 2143,18 - \frac{(276,4)^2}{36} = 21,04$$

$$SC_{\text{traitement}} = \sum (\bar{X}_j - \bar{X})^2 = \frac{\sum T_j^2}{n} - \frac{(\sum X)^2}{N} = \frac{41,8^2 + 46^2 + 43,2^2 + 47,3^2 + 44,2^2 + 53,9^2}{6} - 2122,14 = 15,47$$

$$SC_{\text{erreur}} = 21,04 - 15,47 = 5,57$$

	A1 exercices légers	A2 exercices modérés	A3 exercices élevés	Total
B1 matin				
B2 soir				
Total				

Étape 2

	A1 exercices légers	A2 exercices modérés	A3 exercices élevés	Total
B1 matin	41,8	43,2	44,2	129,2
B2 soir	46,0	47,3	53,9	147,2
Total	87,8	90,5	98,1	276,4

$$SC_A = \sum \left(\bar{X}_a - \bar{X} \right)^2 = \frac{\sum T_a^2}{nb} - \frac{(\sum X)^2}{N} = \frac{87,8^2 + 90,5^2 + 98,1^2}{6 \times 2} - \frac{2122,14}{6} =$$

$$2126,89 - 2122,14 = 4,75$$

Étape 2

	A1 exercices légers	A2 exercices modérés	A3 exercices élevés	Total
B1 matin	41,8	43,2	44,2	129,2
B2 soir	46,0	47,3	53,9	147,2
Total	87,8	90,5	98,1	276,4

$$SC_B = \sum \left(\bar{X}_b - \bar{X} \right)^2 = \frac{\sum T_b^2}{na} - \frac{(\sum X)^2}{N} = \frac{129,2^2 + 147,2^2}{6 \times 3} - \frac{2122,14}{6} =$$

$$2131,14 - 2122,14 = 9$$

Étape 2 des calculs de somme de carrés: effets indépendants et interaction

$$SC_A = \sum \left(\bar{X}_a - \bar{X} \right)^2 = \frac{\sum T_a^2}{nb} - \frac{(\sum X)^2}{N} = \frac{87,8^2 + 90,5^2 + 98,1^2}{6 \times 2} - 2122,14 =$$

$$2126,89 - 2122,14 = 4,75$$

$$SC_B = \sum \left(\bar{X}_b - \bar{X} \right)^2 = \frac{\sum T_b^2}{na} - \frac{(\sum X)^2}{N} = \frac{129,2^2 + 147,2^2}{6 \times 3} - 2122,14 =$$

$$2131,14 - 2122,14 = 9$$

$$SC_{AxB} = SC_{\text{traitement}} - SC_A - SC_B =$$

$$15,47 - 4,75 - 9 = 1,72$$

Le tableau d'analyse de variance

Source de variation	SC	dl	CM	F	p
Intensité de l'exercice	4,75	2	2,375	13,06	<.05
Moment	9,00	1	9,000	48,39	<.05
Intensité x moment	1,71	2	0,855	4,70	<.05
Erreur	5,58	30	0,186		
Total	21,04	35			

Examen de l'interaction significative

	A1 exercices légers	A2 exercices modérés	A3 exercices élevés	Total
B1 matin	41,8	43,2	44,2	129,2
B2 soir	46,0	47,3	53,9	147,2
Total	87,8	90,5	98,1	276,4

Un exemple de calcul

I. Lettré, une chercheuse de réputation internationale, présente des mots simples de type CVC (ex.: bol) en lettrage conventionnel ou en italique en des temps d'exposition courts et très courts. Elle enregistre le nombre d'erreurs d'identification de mots faites par 20 participants et participantes

Le Tableau d'analyse de variance

Source de variation	SC	dl	CM	F	p
Type de lettrage					
Temps d'exposition					
Type de lettrage x Temps d'exposition					
Erreur					
Total					

Trois hypothèses (1)

Hypothèse nulle # 1: Il n'y aura pas plus d'erreurs dans un type de lettrage que dans l'autre

$$F_{\text{critère}} (1, 16) =$$

Hypothèse alternative # 1: Il y aura plus d'erreurs avec le lettrage en italique qu'avec le lettrage normal

Hypothèse nulle # 2: Il y aura autant d'erreurs au temps d'exposition court qu'au temps d'exposition très court $F_{\text{critère}} (1, 16) =$

Hypothèse alternative # 2: Il y aura plus d'erreurs au temps d'exposition très court qu'au temps d'exposition court

Trois hypothèses (2)

Hypothèse nulle # 3: Il y aura autant d'erreurs dans chacune des quatre conditions du plan factoriel composé par le type de lettrage et la durée du temps d'exposition

$$F_{\text{critère}}(1, 16) =$$

Hypothèse alternative # 3: Il y aura plus d'erreurs dans une des quatre conditions du plan factoriel composé par le type de lettrage et la durée du temps d'exposition

Les sommes de carré dans le calcul de l'analyse de variance: rappel

$$SC_{\text{Total}} = \sum \left(X - \bar{X} \right)^2 = \sum X^2 - \frac{(\sum X)^2}{N}$$

$$SC_{\text{traitement}} = \sum \left(\bar{X}_j - \bar{X} \right)^2 = \frac{\sum T_j^2}{n} - \frac{(\sum X)^2}{N}$$

$$SC_{\text{erreur}} = \sum \left(X_{ij} - \bar{X}_j \right)^2 = \sum X^2 - \frac{\sum T_j^2}{n}$$

Lettrage normal (A1)		Lettrage Italique (A2)					
90 mec (B1)	45 msec (B2)	90 msec (B1)	45 msec (B2)	A1B1 ²	A1B2 ²	A2B1 ²	A2B2 ²
7	5	3	11				
8	6	4	12				
9	7	5	13				
10	8	6	14				
11	9	7	15				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Calcul de l'analyse de variance:
sommes de carré (1)**

$$\sum \left(\bar{X}_j - \bar{X} \right)^2 = \frac{\sum T_j^2}{n} - \frac{(\sum X)^2}{N}$$

	A1 Lettrage normal	A2 Lettrage Italique	Total
B1 90 ms			
B2 45 ms			
Total			

**Calcul de l'analyse de variance:
les sommes de carré (2)**

$$SC_A = \sum \left(\bar{X}_a - \bar{X} \right)^2 = \frac{\sum T_a^2}{nb} - \frac{(\sum X)^2}{N}$$

$$SC_B = \sum \left(\bar{X}_b - \bar{X} \right)^2 = \frac{\sum T_b^2}{na} - \frac{(\sum X)^2}{N}$$

$$SC_{AxB} = SC_{\text{traitement}} - SC_A - SC_B$$

$$SC_A = \sum \left(\bar{X}_a - \bar{X} \right)^2 = \frac{\sum T_a^2}{nb} - \frac{(\sum X)^2}{N}$$

$$SC_B = \sum \left(\bar{X}_b - \bar{X} \right)^2 = \frac{\sum T_b^2}{na} - \frac{(\sum X)^2}{N}$$

$$SC_{A \times B} = SC_{\text{traitement}} - SC_A - SC_B =$$

Le tableau d'analyse de variance

Source de variation	Type de lettrage	Temps d'exposition	Type de Temps	Erreur	Total
	12	40		215	19

Les dl aussi

La colonne

Les SC

Les CM s'obtiennent par division des SC et de

Plus petit que .05 signifie rejet de l'hypothèse nulle, plus grand son acceptance

Le total représente la variance totale de la variable dépendant

Le CM de l'erreur

η^2 ou ω^2