

150	ANOVA A UN CRITERE DE CLASSIFICATION FIXE.....	2
150.1	UTILITE	2
150.2	COMPARAISONS 2 A 2	3
150.3	CONFIANCE ET PUISSANCE	4
150.4	CONTRASTES.....	5
150.5	TESTS	7
140.5.1	<i>Ensemble de comparaisons quelconques</i>	<i>8</i>
140.5.2	<i>Ensemble de comparaisons 2 à 2</i>	<i>8</i>
150.6	EXEMPLE.....	8

150 ANOVA à un critère de classification fixe

150.1 Utilité

Les contrastes sont des comparaisons de moyenne qui suivent le rejet d'hypothèse nulle concernant le critère fixe et non ordonné d'une ANOVA, permettant d'identifier précisément quels sont les niveaux du critère qui présentent une différence de moyenne significative.

En effet, le fait d'avoir conclu que toutes les moyennes ne sont pas égales ne permet pas d'établir qu'elles soient toutes différentes.

RHo permet plusieurs combinaisons d'inégalités :

$$H_1 : \mu_1 < \mu_2 = \mu_3$$

$$H_2 : \mu_1 = \mu_2 < \mu_3$$

$$H_3 : \mu_1 = \mu_3 > \mu_2$$

$$H_4 : \mu_1 < \mu_2 < \mu_3$$

etc, etc... d'autant plus nombreuses que le nombre de moyennes à comparer est grand.

Non ordonné signifie que les niveaux du facteur fixe peuvent être analysés sans ordre déterminé (par exemple : trois régimes alimentaires) par opposition à un critère ordonné (par exemple trois aquariums de température 15,20 et 25°C).

Nous le voyons ici dans le contexte de l'ANOVA I, mais les contrastes peuvent s'intégrer à des analyses à plusieurs critères, pour autant que ceux-ci soient fixes.

Les contrastes permettent de faire face à deux problèmes :

1. Maintenir la confiance $1-\alpha$ en prenant en compte le risque d'erreur « en cascade » dans une série de tests qui impliquent le même échantillon : 1 comparé à 2, 1 comparé à 3, 1 comparé à 4 etc..
2. Bénéficier de la puissance maximale de l'analyse en prenant en compte la variance globale (CMR) mesurée dans l'ensemble des échantillons . Autrement dit, les variances des échantillons 3 et 4, par exemple, contribuent à l'estimation de la variance nécessaire pour comparer les échantillons 1 et 2.

150.2 Comparaisons 2 à 2

Lorsque l'ANOVA compare plus de deux moyennes et que le critère de classification n'est pas ordonné, le rejet de H_0 laisse ouvertes plusieurs alternatives entre lesquelles il faut pouvoir opérer un choix.

Un géologue compare le niveau de saturation de nappes phréatiques¹ en Condroz, en Hesbaye et en Ardennes.

Après avoir étudié 5 nappes dans chaque région, il rejette $H_0 : \mu_1 = \mu_2 = \mu_3$ et conclut que le niveau est différent d'une région à l'autre. A ce stade, il n'a pas montré que les 3 régions étaient différentes : il souhaite comparer Condroz et Hesbaye, Condroz et Ardennes, Hesbaye et Ardennes.

Ceci revient à tester 3 H_0 :

$$H_0 : \mu_1 = \mu_2$$

$$H_0 : \mu_1 = \mu_3$$

$$H_0 : \mu_2 = \mu_3$$

A chacune de ces hypothèses nulles correspondent deux hypothèses alternatives :

$$H_1 : \mu_1 > \mu_2$$

$$H_2 : \mu_1 < \mu_2$$

Etc...

En cas de rejet de H_0 au seuil α , on choisira H_1 si $Mx_1 > Mx_2$ et réciproquement.

Si a priori on se pose une seule hypothèse alternative, on gagnera en puissance en faisant le test au seuil 2α tout en conservant une confiance $1 - \alpha$.

¹ Nappe phréatique, nappe d'eau souterraine, formée par l'infiltration des eaux de pluie et alimentant des puits et des sources.

150.3 Confiance et puissance

Multiplier des analyses sur des groupes pris deux à deux diminue gravement la confiance et la puissance des tests.

A supposer que l'on répète le calcul de l'ANOVA 1 pour plusieurs tableaux partiels ne comprenant chacun que deux régions, nous obtiendrons le résultat souhaité, mais entaché de deux défauts rédhibitoires² : une forte augmentation du risque d'erreur α et une diminution, variable suivant les cas, de la puissance des tests $1 - \beta$.

Premier défaut : la confiance $1 - \alpha$ diminue parce que les analyses sont réalisées indépendamment les unes des autres alors que les données ne le sont pas.

Par exemple, supposons que dans la réalité, $\mu_1 = \mu_2 = \mu_3$, mais que dans l'expérience, Mx_1 soit particulièrement basse (ce qui a un risque α de se produire).

Dans le test $H_0 : \mu_1 = \mu_2$ nous rejeterons H_0 à tort, situation prévue par le risque α . Mais sachant cela, la probabilité de rejeter H_0 le test $H_0 : \mu_1 = \mu_3$ est presque 1, et sur 3 tests, on se trompera 2 fois, soit 66% d'erreur, ce qui est largement supérieur à α .

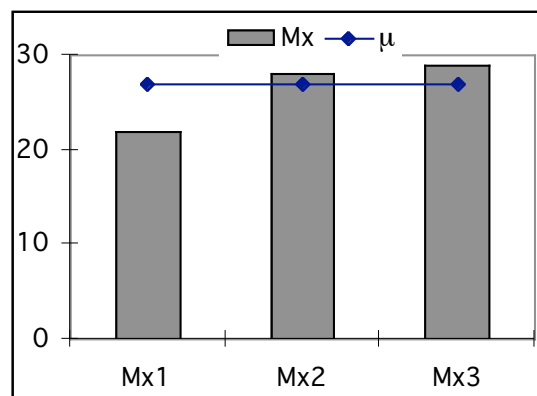


Figure 150-1 Représentation schématique d'une situation où une H_0 vraie sera rejetée 2 fois par erreur.

Le risque d'erreur α global d'un ensemble de tests non indépendants peut être très grand.

Deuxième défaut : la puissance diminue parce qu'elle est liée au nombre de degrés de liberté de SCER, qui est moindre dans les comparaisons 2 à 2 que dans les comparaisons globales.

² Rédhibitoire : qui constitue un obstacle radical

Par exemple, supposons que l'on compare 5 régimes avec 3 réplicats pour chacun. Dans une analyse globale, SCER a $15 - 3 = 12$ degrés de liberté. Dans une analyse ne comprenant que 2 régimes, SCER a $6 - 2 = 4$ degrés de liberté.

Autrement dit, dans une analyse sur 2 groupes on se prive de l'estimation de la variance réalisée dans les trois autres groupes : CMR est donc moins bien estimé et le test est moins puissant.

La puissance du test est liée à la précision de l'estimation de la variance et donc au nombre total d'observations qui interviennent dans son estimation

150.4 Contrastes

Un contraste est une combinaison linéaire de moyennes permettant la comparaison de 2 moyennes à partir d'un ensemble de plus de 2 moyennes.

La comparaison $\mu_1 - \mu_2$ est un cas particulier de contraste.

Un contraste peut être plus complexe :

$$\mu_1 - \frac{\mu_2 + \mu_3}{2}$$

est une combinaison de 2 moyennes impliquant 3 moyennes : il compare la moyenne de la population 1 avec la moyenne des populations 2 et 3, prises ensemble.

On peut l'écrire :

$$\mu_1 - 0.5.\mu_2 - 0.5.\mu_3$$

ou plus simplement :

$$2\mu_1 - \mu_2 - \mu_3$$

Il y a souvent des raisons logiques d'effectuer ces regroupements.

Exemple 1

Imaginons qu'un chimiste mesure la solubilité d'acides aminés dans un solvant organique. Il obtient des mesures pour la glycine³ (solubilité moyenne μ_1), la phénylalanine⁴ (μ_2) et le tryptophane⁵ (μ_3).

³ Glycine : le plus petit des acides aminés (a.a.) constitutif des protéines. Sa chaîne latérale est limitée à un atome d'hydrogène.

⁴ Phénylalanine : a.a. constitutif des protéines, précurseur de l'adrénaline

Il y a une raison logique à tester particulièrement l'hypothèse :

$$H_0 : \mu_1 - \frac{\mu_2 + \mu_3}{2} = 0$$

qui compare la solubilité d'un a.a. à petite chaîne latérale à celle de la solubilité moyenne de 2 a.a. dont la grande chaîne latérale contient un cycle aromatique.

Exemple 2

Imaginons un anthropologue qui mesure la longueur de la canine inférieure pour 4 genres de singes : Semnopithèque⁶, Colobe⁷, Gorille⁸ et Chimpanzé⁹.

Il y a une raison logique à tester l'hypothèse :

$$H_0 : \frac{\mu_1 + \mu_2}{2} - \frac{\mu_3 + \mu_4}{2} = 0$$

dans un contraste qui peut aussi s'écrire

$$\mu_1 + \mu_2 - \mu_3 - \mu_4 = 0$$

et qui regroupe des petits singes d'une part et des grands singes d'autre part.

et de la thyroxine. Sa chaîne latérale contient un cycle aromatique hydrophobe.

⁵ *Tryptophane : a.a. constitutif des protéines. Sa chaîne latérale contient un cycle aromatique hydrophobe.*

⁶ *Semnopithèques : singe des forêts d'Asie, vivant en bandes. (Long. 75 cm sans la queue.)*

⁷ *Colobe : singe d'Afrique, au pouce réduit, au pelage long et soyeux (Long. 70 cm sans la queue.)*

⁸ *Gorille : singe anthropoïde de l'Afrique équatoriale. (Frugivore, il est le plus grand et le plus fort de tous les singes, sa taille atteignant 2 m et son poids pouvant dépasser 200 kg.)*

⁹ *Chimpanzé : singe anthropoïde de l'Afrique équatoriale, arboricole, sociable et s'apprivoisant facilement. (Haut. 1,40 m ; poids 75 kg ; longévité jusqu'à 50 ans, en liberté.)*

Dans un contraste, la moyenne de toutes les populations affectées d'un signe positif est comparée à la moyenne de toutes les populations affectées d'un signe négatif et la façon de regrouper les populations répond à un critère logique.

150.5 Tests

L'estimateur du contraste

$$\lambda: 2\mu_1 - \mu_2 - \mu_3 \text{ est } L: 2M_{x1} - M_{x2} - M_{x3}$$

étant entendu que nous ne considérons que les échantillons de taille égale.

Cet estimateur L est une variable aléatoire caractérisée par une moyenne et une variance.

Le plus fréquemment, par hypothèse nulle, la moyenne du contraste est nulle :

$$H_0 : \lambda = 0.$$

La variance du contraste est notée σ_L^2 , elle est estimée par S_L^2 :

$$S_L^2 = \frac{1}{n} CMe \sum_{i=1}^{n_a} c_i^2$$

Équation 150-1

expression dans laquelle c_i représente les coefficients de chaque moyenne, n_a le nombre de niveaux du critère de classification, n la taille de chaque échantillon. CMe représente le carré moyen utilisé pour tester globalement les moyennes.

Dans l'ANOVA1, et les anova plus complexes dont les critères sont tous fixes CMe est toujours le carré moyen résiduel CMR. Dans certains modèles d'ANOVA impliquant au moins un critère aléatoire, CMe est un carré moyen qui sera défini plus loin.

$$\sum_{i=1}^k c_i^2 = 2$$

Dans les comparaisons de moyennes 2 à 2, les coefficients sont toujours ± 1 et

Dans le contraste

$$\mu_1 - \frac{\mu_2 + \mu_3}{2}$$

qui s'écrit par facilité de calcul :

$$2\mu_1 - \mu_2 - \mu_3$$

$$\sum_{i=1}^k c_i^2 = 4 + 1 + 1 = 6$$

140.5.1 Ensemble de comparaisons quelconques

Selon Scheffé, l'ensemble des contrastes (logiques ou non) que l'on peut imaginer construire à partir de n_a moyennes a une probabilité $1 - \alpha$ d'être compris dans les limites de $\pm D.S$:

$$DS = \pm \sqrt{(n_a - 1) S_L^2 F_{n_a - 1; N - n_a; 1 - \alpha}} \quad \text{Équation 150-2}$$

Dès que la valeur observée du contraste est supérieure (en valeur absolue) à D.S., il y a rejet de $H_0 : \lambda = 0$

140.5.2 Ensemble de comparaisons 2 à 2

Dans le cas de comparaisons de moyennes 2 à 2, la formule se simplifie et ne doit être calculée qu'une seule fois pour toutes les comparaisons. La valeur obtenue s'appelle la PPDS : plus petite différence significative :

$$PPDS = \pm \sqrt{(n_a - 1) \frac{2CMe}{n} F_{n_a - 1; N - n_a; 1 - \alpha}} \quad \text{Équation 150-3}$$

Dès que la valeur observée du contraste est supérieure (en valeur absolue¹⁰) à la PPDS, il y a rejet de $H_0 : \lambda = 0$

150.6 Exemple

Lors d'une étude sur l'effet de la teneur en calcium dans la nourriture des brebis, le poids frais (en g) du muscle semi -tendineux est mesuré sur les animaux des six groupes : un témoin qui reçoit l'alimentation normale, un additif alimentaire à faible dose et à forte dose, du calcium en comprimé, du calcium en solution liquide à faible dose et à forte dose.

Les moyennes et les variances obtenues sur 7 brebis par régime sont présentées dans la table suivante :

¹⁰ Valeur absolue : représentation d'un nombre sans tenir compte de son éventuel signe négatif.

	moyenne	variance
Témoin	140.0	3.0
Additif dose 1	143.4	7.0
Additif dose 2	145.1	3.2
Comprimé	139.3	5.7
liquide dose1	147.6	5.6
liquide dose2	147.9	3.8
Général	143.9	4.7
H_{6;6 obs}		2.3
H_{6;6 ;0,95}		13,7
Hartley :		AHo

Tableau 150-1 Effet du calcium sur le poids du muscle semi -tendineux : moyennes et variances. Test de Hartley.

Les résultats sont analysés par ANOVA1 et présentés dans la table suivante :

	sce	dl	cm	f	p
tot	644.8	41			
fact	475.4	5	95.1	20.2	1E -09
res	169.4	36	4.7		

Tableau 150-2 Effet du calcium sur le poids du muscle semi -tendineux : ANOVA1.

Le résultat montre avec un risque d'erreur quasi nul que certaines moyennes sont différentes des autres.

Il n'y a pas de variable quantitative permettant de traiter le problème par la régression.

Les comparaisons multiples de Scheffé peuvent être réalisées 2 à 2, en référence à une PPDS calculée au seuil $\alpha = 5\%$. Le CMe est ici le CMR.

$$PPDS = \pm \sqrt{(n_a - 1) \frac{2CMR}{n} F_{n_a-1; N-n_a; 1-\alpha}}$$

$$PPDS = \pm \sqrt{5 \times \frac{2 \times 4.7}{7} \times 2.48} = 4.1$$

Dans certaines tables, il faudra approximer $F_{5;36} = 2.48$ par la moyenne de $F_{5;30} = 2.53$ et $F_{5;40} = 2.45$ soit 2.49.

ce qui donne les résultats présentés dans la table suivante :

	140.0	143.4	145.1	139.3	147.6	147.9
140.0		3.4	5.1	0.7	7.6	7.9
143.4	3.4		1.7	4.1	4.3	4.5
145.1	5.1	1.7		5.8	2.6	2.8
139.3	0.7	4.1	5.8		8.4	8.6
147.6	7.6	4.3	2.6	8.4		0.2
147.9	7.9	4.5	2.8	8.6	0.2	

Tableau 150-3 Effet du calcium sur le poids du muscle semi -tendineux : comparaisons pairées. Les résultats supérieurs (en valeur absolue) à la PPDS sont indiqués en gras.

Les tests indiquent qu'au seuil $\alpha = 5\%$, le témoin apparaît différent de l'additif à la dose 2, et de la forme liquide à dose 1 et 2 ; l'additif dose 1 est différent du groupe comprimé et de la forme liquide à dose 1 et 2, l'additif dose 2 est différent du comprimé ; le comprimé est différent de tous les autres, excepté le témoin, les formes liquides sont différentes du témoin, du comprimé et de l'additif dose 1.

Pour faire court et par abus de langage, on préférera dire : « le témoin apparaît différent de l'additif à la dose 2 » plutôt que « le poids du muscle semi -tendineux est plus élevé dans le groupe traité par l'additif alimentaire à forte dose que dans le groupe témoin »

Au risque d'erreur β (inconnu), l'expérience ne met pas en évidence de différence pour l'additif dose 1, ni pour le comprimé, par rapport au groupe témoin. Elle ne met pas non plus en évidence de différence d'effet entre les deux doses de la forme liquide.

Des comparaisons plus globales peuvent être réalisées par des contrastes de Scheffé. Les coefficients et le résultat des tests sont présentés dans la table suivante :

		témoin traités	Solide liquide	additif compr.
	Mx_i	c_i	c_i	c_i
Témoin	140.0	5	0	0
Add. 1	143.4	-1	2	-1
Add. 2	145.1	-1	2	-1
Compr.	139.3	-1	2	2
Liqu. 1	147.6	-1	-3	0
Liqu.2	147.9	-1	-3	0
$\sum c_i^2$		30	30	6
S^2L		20.2	20.2	4.0
Seuil α 5%		15.7	15.7	7.0
L		-23.3	-30.9	-9.9
		Rho	Rho	Rho

Tableau 150-4 Comparaisons globales réalisées par des contrastes de Scheffé .

Le premier contraste (5, -1, -1, -1, -1, -1) compare le témoin à tous les traités. Il correspond à la différence de moyennes suivante :

$$\lambda_1 = \mu_1 - \frac{\mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6}{5}$$

Tous les coefficients sont multipliés par 5 pour simplifier le calcul : étant donné qu'ils sont pris en compte dans le calcul de L et de S^2L , le facteur d'échelle ne change rien au test, puisqu'il est pris en considération dans le terme $\sum ci^2$

$$\lambda_1 = 5\mu_1 - \mu_2 - \mu_3 - \mu_4 - \mu_5 - \mu_6$$

la valeur du seuil 5% est obtenue en calculant $\sum ci^2$, s^2L puis

$$DS = \sqrt{(n_a - 1)S_L^2 F_{n_a - 1; N - n_a; 1 - \alpha}}$$

$$DS = \sqrt{5 \times 20.2 \times 2.48} = 15.8$$

la valeur observée du contraste est obtenue en effectuant le calcul :

$$L_1 = 5 \times 140 - 143.4 - 145.1 - 139.3 - 147.6 - 147.9 = -23.3$$

La valeur absolue 23.3 est comparée au seuil 15.8.

Elle est supérieure et il y a RHo.

On peut donc conclure avec une confiance de 95% que le témoin est différent de la moyenne de tous les traitements.

De même, il y a rejet de H_0 lorsqu'on compare le mode d'administration solide au mode d'administration liquide :

$$\lambda_2 = 2\mu_2 + 2\mu_3 + 2\mu_4 - 3\mu_5 - 3\mu_6$$

$$L_2 = 2 \times 143.4 + 2 \times 145.1 + 2 \times 139.3 - 3 \times 147.6 - 3 \times 147.9 = -30.9$$

La valeur absolue 30.9 est comparée au seuil 15.7.

La valeur observée est supérieure au seuil et il y a Rho.

On peut donc conclure avec une confiance de 95% que la moyenne des «mode d'administration solide» est différente de la moyenne des «mode d'administration liquide» .

L'additif est comparé au comprimé dans le troisième contraste.

$$\lambda_3 = -\mu_2 - \mu_3 + 2\mu_4$$

$$L_3 = -143.4 - 145.1 + 2 \times 139.3 = -9.9$$

La valeur absolue 9.9 est comparée au seuil 7.0.

Elle est supérieure et il y a Rho.

On peut donc conclure avec une confiance de 95% que la moyenne des «additifs» est différente de la moyenne des « comprimés».