150 COMPARAISONS MULTIPLES (CONTRASTES) 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
150.6 ENSEMBLE DE COMPARAISONS QUELCONQUES 8
150.7 ENSEMBLE DE COMPARAISONS 2 A 2 8
150.8 EXEMPLE 8

150 Comparaisons multiples (contrastes)

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 un 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 quartiers d'une ville) par opposition à un critère ordonné (par exemple trois villes de latitude différentes, ordonnées du Nord vers le Sud).

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-α 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 Ho laisse ouvertes plusieurs alternatives entre lesquelles il faut pouvoir opérer un choix.

Un géologue compare le niveau de saturation de nappes phréatiques[1] en Condroz, en Hesbaye et en Ardennes.

Après avoir étudié 5 nappes dans chaque région, il rejette Ho : $\mu 1 = \mu 2 = \mu 3$ 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 Ho :

Ho : $\mu 1 = \mu 2$
Ho : $\mu 1 = \mu 3$
Ho : $\mu 2 = \mu 3$

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 Ho au seuil α, on choisira H_1 si Mx1 > Mx2 et réciproquement.

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

[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 rédhibitoires[2] : une forte augmentation du risque d'erreur α et une diminution, variable suivant les cas, de la puissance des tests 1 -β.

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

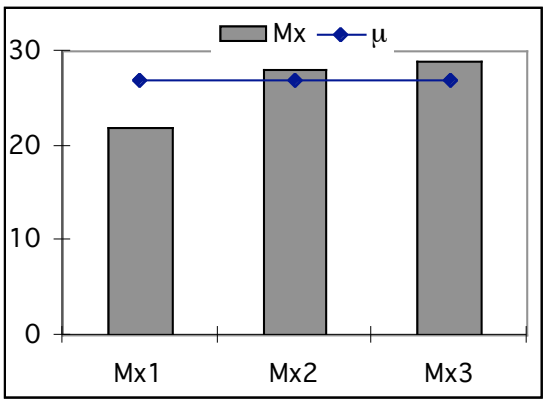

Figure 150-1 Représentation schématique d'une situation où une Ho vraie sera rejetée 2 fois par erreur.

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 régions avec 3 réplicats pour chacune. Dans une analyse globale, SCER a 15 -3 = 12 degrés de liberté. Dans une analyse ne comprenant que 2 régions, 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

Le temps de chômage des jeunes diplômés est estimé à Bruxelles (moyenne réelle $\mu 1$) en Flandre($\mu 2$), et en Wallonie($\mu 3$) pour 50 chômeurs pris au hasard pour représenter chaque région, un an après la mise en place du plan Rosetta

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

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

qui compare la métropole aux deux régions et ensuite

Ho μ2 -μ3 =0

Qui compare les deux régions entre elles.

Exemple 2

Imaginons un anthropologue qui mesure la longueur de la canine inférieure fossile pour 4 genres de singes : Semnopithèque[3] , Colobe[4], Gorille[5] et Chimpanzé[6].

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

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

dans un contraste qui peut aussi s'écrire

μ1 + μ2 - μ3 - μ4 = 0

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

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.

---

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

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

[5] 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.)

[6] 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.)

150.5 Tests

L'estimateur du contraste

λ: 2μ1 - μ2 - μ3 est L : 2Mx1 - Mx2 - Mx3

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 :

Ho : $\lambda=0$.

La variance du contraste est notée σ^2_L, elle est estimée par S^2_L :

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

$2\mu1 - \mu2 - \mu3$

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

150.6 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 -α d'être compris dans les limites de ± D.S :

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

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

150.7 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 absolue[7]) à la PPDS, il y a rejet de Ho : $\lambda = 0$

150.8 Exemple

La production de déchets ménagers (Kg/personne/an) est mesurée dans 20 familles de 4 personnes de revenus comparable. Cinq types d'habitat sont pris en considération : maison 4 façades avec (1) ou sans (2) potager ; immeuble à appartement au rez-de-chaussée (3) ou au 3° étage avec (4) ou sans (5) ascenseur. Quatre familles sont suivies dans chaque type d'habitat.

	1	2	3	4	
	241,3	287,8	271,4	277,1	23
	207,4	299,8	293,3	287,3	25
	243,2	296,5	275,9	267,3	25
	218,3	275,6	274,8	276,6	24
Moyennes	227,6	289,9	278,9	277,1	24
variances	308,5	116,8	96,5	66,8	9

[7] *Valeur absolue : représentation d'un nombre sans tenir compte de son éventuel signe négatif.*

Tableau 150-1 Production de déchets ménagers (Kg/personne/an) dans cinq types d'habitats. Données fictives, moyennes et variances.

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

Test de Hartley	
S2max/S2min	4,620
H 3;5	50,7
Non significatif	

	SCE	DL	CM	F	
totale	12689,61	19			
factorielle	10633,84	4	2658,46	19,40	0,
résiduelle	2055,76	15	137,05		

Tableau 150-2 Production de déchets ménagers (Kg/personne/an) dans cinq types d'habitats. Test de Hartley et ANOVA1.

Les résultats montrent une homogénéité de variance avec un risque d'erreur inconnu et qu'au moins une moyennes est différentes des autres avec un risque d'erreur quasi nul.

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

Des contrastes groupés peuvent être réalisées par des contrastes de Scheffé pour répondre à des questions logiques. Les coefficients et le résultat des tests sont présentés dans la table suivante :

1	POTAGER	~~POTAGER~~	~~POTAGER~~	~~POTAGER~~	~~POTAGER~~
2	MAISON	MAISON	APPART	APPART	APPART
3			REZ	3°	3°
4			~~Escalier~~	~~Escalier~~	escalier
5				Ascenseur	~~Ascenseur~~
6	POTAGER	~~POTAGER~~			
	227,6	289,9	278,9	277,1	247,7
1	4	-1	-1	-1	-1
2	3	3	-2	-2	-2
3			2	-1	-1
4			1	1	-2
5				1	-1
6	1	-1			

n	4
na-1	4
N-na	15
F4;15;0,95	3,06
Cme = CMR	137,05

	Σci2	D	DS	
1	20	183,3	91,5	S
2	30	54,8	112,1	
3	6	33,0	50,1	
4	6	60,6	50,1	S
5	2	29,4	28,9	S
6	2	62,4	28,9	S

Tableau 150-3 Comparaisons globales réalisées par des contrastes de Scheffé .

Le premier contraste (4, -1, -1, -1, -1) compare la présence de potager à tous les autres habitats. Il correspond à la différence de moyennes suivante :

$$\lambda_1 = \mu_1 - \frac{\mu_2 + \mu_3 + \mu_4 + \mu_5}{4}$$

Tous les coefficients sont ensuite multipliés par 4 pour simplifier le calcul.

$$\lambda_1 = 4\mu_1 - \mu_2 - \mu_3 - \mu_4 - \mu_5$$

la valeur du seuil 5% est obtenue en calculant Σci^2, S^2_L puis

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

soit pour le contraste λ1

$$DS_{\lambda 1} = \sqrt{(5-1) \times \frac{137{,}05}{4} \times 20 \times 3{,}06} = 91{,}5$$

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

L1 = ABS(4x227,6 - 289,9 - 278,9 - 277,1 - 247,7) = 183,3

La valeur observée 183,3 est comparée au seuil (DS) = 91,5

La valeur observée est supérieure et il y a RHo.avec une confiance de 95% (probabilité du F inclue dans le calcul de DS).

On peut donc conclure avec une confiance de 95% que la production de déchet de l'habitat avec potager est inférieure à la moyenne de tous les habitats sans potager.

De même, il y a rejet de Ho lorsqu'on compare la production de déchet des maisons 4 façades à celle des appartements :

$$\lambda_2 = 3\mu_1 + 3\mu_2 - 2\mu_3 - 2\mu_4 - 2\mu_5$$

Les coefficients sont équilibrés en fonction du nombre de moyennes intervenant avec un signe positif ou négatif : s'il n'y a pas de différence entre le groupe 1, représenté par 2 niveaux, et le groupe 2, représenté par 3 niveaux, la valeur attendue de $\lambda 2$ est 0.

La valeur observée du contraste est calculée à partir des moyennes observées

L2 = ABS(3 x 227,6 + 3 x 289,9 - 2 x 278,9 - 2 x 277,1 - 2 x 247,7) = 54,8

La valeur absolue 54,8 est comparée au seuil DS 112,1
La valeur observée est inférieure au seuil et il y a AHo.

On peut donc conclure avec une probabilité inconnue de se tromper que l'on a pas mis en évidence de différence globale de production de déchet entre les maisons 4 façades et les appartements.

La procédure se poursuit avec les autres contrastes, indiquant (L3) que l'on ne montre pas de différence globale entre les appartements rez-de-chaussée et ceux du 3° étage (risque β inconnu), mais qu'on en montre une (L4) entre les appartements situés au 3° étage sans ascenseur et les autres ainsi qu'entre les appartements situés au 3° étage avec et sans ascenseur (L5) (risque $\alpha < 5\%$). C'est donc le facteur « escalier » qui contribuerait à limiter la production de déchets. Notons encore la différence entre les maisons 4 façades avec et sans potager (L6, risque $\alpha < 5\%$).

En résumé l'étude montre essentiellement la diminution de production de déchets liée au potager et au fait de devoir monter les achats au 3° étage par l'escalier.

Ces données sont fictives.