T. D. n° 3
Analyse de variance à deux facteurs : modèles aléatoires et mixtes

Exercice 1. Prise en compte d'un effet bloc[1]

Parmi trois variétés de blé, quelle est celle qui donne le meilleur rendement ? Pour répondre à cette question, on met en culture ces trois variétés, sur trois parcelles distinctes. Chaque parcelle est divisée en trois sous-parcelles, et ainsi chaque variété est cultivée sur chaque parcelle. Les apports d'engrais azoté, de fongicide et d'insecticide sont les mêmes et sont considérés comme optimaux pour chaque variété. Les mesures ont été reportées dans le tableau ci-dessous.
Le sol de chaque parcelle est censé être de même qualité mais on préfère le vérifier en prenant en compte un effet parcelle dans le modèle d'analyse de la variance. Ce type d'effet est appelé « effet bloc ». On ne s'intéresse pas à cet effet, mais on le prend en compte dans l'analyse de la variance pour mieux mettre en évidence les éventuels effets des autres effets.

Parcelle (bloc)	Variété	Rendement
1	A	61,00
2	A	70,92
3	A	55,94
1	L	70,80
2	L	66,44
3	L	68,12
1	M	55,30
2	M	57,30
3	M	54,39

1. Pourquoi la prise en compte de l'effet bloc peut-elle être utile pour conclure à un eventuel effet variété ?
2. Y a-t-il un effet variété au niveau de confiance de 5% ? On supprime l'effet bloc de l'analyse.
3. Pour quelle raison, en pratique, pourrait-on effectuer cette suppression ? Construire le tableau d'analyse de la variance à un facteur (le facteur variété).
4. Construire le test global de significativité de l'effet variété. Comment expliquez-vous que le test devienne significatif ?

[1]Cet exercice est issu du livre d'exercices de François Husson et de Jérôme Pagès intitulé *Statistiques générales pour utilisateurs*, éditions PUR.

Exercice 2. Viennoiseries[2]

Pendant la cuisson, les croissants absorbent la graisse en quantité variable. Nous avons relevé la quantité de graisse absorbée lors de la cuisson de six fournées de croissantes pour quatre types de graisse. Les mesures sont présentées dans le tableau ci-dessous.

Fournée/Graisse	1	2	3	4
1	64	78	75	55
2	72	91	93	66
3	68	97	78	49
4	77	82	71	64
5	56	85	63	70
6	95	77	76	68

1. Quels sont les modèles que vous pouvez utiliser pour analyser ces données ? Pour chacun d'entre eux, détaillez les hypothèses testées.
2. Nous nous intéressons uniquement à ces quatre types de graisse. Par contre, nous cherchons à savoir si il y a un effet des fournées en général sur la quantité de graisse utilisée. Pour cela, quel est donc le modèle à choisir parmi les trois précédents ?
   Nous allons maintenant travailler avec le modèle que nous venons de choisir.
3. Vérifiez que les conditions d'application de l'ANOVA II sont vérifiées.
4. Construisez, avec l'aide de MINITAB, le tableau d'analyse de la variance pour ce modèle.
5. La fournée a-t-elle un effet significatif sur la quantité de graisse absorbée ? Écrivez explicitement les hypothèses nulle et alternative de votre test et justifiez votre réponse.
6. La graisse a-t-elle un effet significatif sur la quantité de graisse absorbée ? Écrivez explicitement les hypothèses nulle et alternative de votre test et justifiez votre réponse.
7. Calculez η^2. Que concluez-vous ?
8. Évaluez la puissance dans chacun des tests du tableau de l'ANOVA II.
9. Quelles sont les comparaisons multiples que vous pouvez réaliser ? Réalisez-les et analysez-les.

Exercice 3. Comparaison de l'évaluation de la résistance d'un même type de ciment

Davies et Goldsmith[3] ont récolté les données d'une expérience dont le but était d'étudier les différentes sources de variabilité possibles de la résistance d'un ciment

[2]D'après Y. Dodge. *Analyse de régression appliquée.* Éditions Dunod, 1999.

[3]Davies, O.L. et Goldsmith, P.L. (Eds.), *Statistical Methods in Research and Production*, 4th edition, Oliver and Boyd, Edinburgh, 1972.

fabriqué à Portland. On note Y la variable associée à la résistance du ciment.

L'expérience s'est déroulée ainsi : plusieurs petits prélèvements d'un même type de ciment ont été mélangés à de l'eau et travaillés par trois personnes différentes, les « mélangeurs ». On a alors formé douze cubes à l'aide de chacune des préparations des « mélangeurs ». Puis on a donné ces 36 cubes à trois personnes chargées d'évaluer leur résistance, les « casseurs ». La répartition des 36 cubes entre ces « casseurs » a été faite de telle sorte que chaque « casseur » reçoive quatre cubes provenant de chacune des préparations des « mélangeurs » soit douze cubes au total.

Tous les tests de résistance ont été faits sur la même machine. L'objectif principal de cette expérience était d'étudier et de quantifier l'importance de la variabilité dans les tests de résistance qui pouvait provenir des différences individuelles entre les « mélangeurs » et les « casseurs ». Les données ci-dessous, exprimées dans les unités d'origine c'est-à-dire en livres par pouces carrés, ont été recopiées dans le tableau ci dessous.

	« Casseur » 1		« Casseur » 2		« Casseur » 3	
« Mélangeur » 1	5280	5520	4340	4400	4160	5180
	4760	5800	5020	6200	5320	4600
« Mélangeur » 2	4420	5280	5340	4880	4180	4800
	5580	4900	4960	6200	4600	4480
« Mélangeur » 3	5360	6160	5720	4760	4460	4930
	5680	5500	5620	5560	4680	5600

Partie I :

Dans l'expérience d'origine, on ne s'intéressait qu'aux différences ne pouvant être dues qu'à ces trois « mélangeurs » et à ces trois « casseurs ».

1. Écrire le modèle d'analyse de la variance relatif à cette étude. On précisera la nature des facteurs explicatifs ainsi que les hypothèses faites.
2. Les hypothèses du modèle sont-elles vérifiées ? Calculer les estimations de tous les paramètres du modèle.
3. Existe-t-il une interaction dans l'évaluation de la résistance du ciment entre les « mélangeurs » et les « casseurs » ?
4. Existe-t-il des différences dans l'évaluation de la résistance dues aux « mélangeurs » ?
5. Existe-t-il des différences dans l'évaluation de la résistance dues aux « casseurs » ?
6. Compte tenu de la nature des deux facteurs peut-on procéder à des comparaisons multiples ? Il y a-t-il un facteur pour lequel cette procédure se justifie, si oui procéder aux tests correspondants.

Partie II :

On souhaite désormais que les résultats de cette expérience ait une portée plus générale que celle de la première partie : on ne restreint plus l'étude à cette population de « casseurs » et de « mélangeurs ».

6. Écrire le modèle d'analyse de la variance relatif à cette étude. On précisera la nature des facteurs explicatifs ainsi que les hypothèses faites.
7. Les hypothèses du modèle sont-elles vérifiées ? Calculer les estimations de tous les paramètres du modèle.
8. Existe-t-il une interaction dans l'évaluation de la résistance du ciment entre les « mélangeurs » et les « casseurs » ?
9. Existe-t-il des différences dans l'évaluation de la résistance dues aux « mélangeurs » ?
10. Existe-t-il des différences dans l'évaluation de la résistance dues aux « casseurs » ?

Exercice 4. Quatre modèles de machines à écrire[4]

Une entreprise cherche à tester quatre modèles de machines à écrire. Pour faire ce test, elle demande à cinq secrétaires professionnelles de taper un texte pendant 5 minutes. À la fin du test, on compte le nombre moyen de mots tapés en une minute. On répète l'expérience le lendemain.
Les résultats (nombre moyen de mots par minute) sont présentés dans le tableau au verso.

	Secrétaires				
Machines à écrire	1	2	3	4	5
1	33	31	34	34	31
	36	31	36	33	31
2	32	37	40	33	35
	35	35	36	36	36
3	37	35	34	31	37
	39	35	37	35	40
4	29	31	33	31	33
	31	33	34	27	33

Soit le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + \gamma_{i,j} + \epsilon_{i,j,k}$$

avec $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4, 5$ et $k = 1, 2$ et où α correspond à l'influence de la machine à écrire, β l'influence de la secrétaire et γ l'interaction entre la machine i et la secrétaire j.

[4]Cet exercice est issu du livre d'exercices de François Husson et de Jérôme Pagès intitulé *Statistiques générales pour utilisateurs*, éditions PUR.

1. Décrire le modèle, et ses conditions d'utilisation, qui permet de tester les trois hypothèses nulles suivantes :
   (i) Que cherche-t-on à tester avec les deux hypothèses suivantes ?
   $$\mathcal{H}_0 \ : \quad \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$$
   contre
   $$\mathcal{H}_1 \ : \quad \text{les valeurs des } \alpha_i \text{ ne sont pas toutes égales, } i = 1, 2, 3, 4.$$
   (ii) Que cherche-t-on à tester avec les deux hypothèses suivantes ?
   $$\mathcal{H}_0 \ : \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$$
   contre
   $$\mathcal{H}_1 \ : \quad \text{les valeurs des } \beta_j \text{ ne sont pas toutes égales, } j = 1, 2, 3, 4, 5.$$
   (iii) Que cherche-t-on à tester avec les deux hypothèses suivantes ?
   $$\mathcal{H}_0 \ : \quad \gamma_{1,1} = \gamma_{1,2} = \cdots = \gamma_{1,5} = \gamma_{2,1} = \cdots = \gamma_{4,5}$$
   contre
   $$\mathcal{H}_1 \ : \quad \text{les valeurs des } \gamma_{i,j} \text{ ne sont pas toutes égales, } i = 1, 2, 3, 4 \text{ et } j = 1, 2, 3, 4, 5.$$
2. Décrire le modèle, et ses conditions d'utilisation, qui permet de tester les trois hypothèses nulles suivantes :
   (i) Que cherche-t-on à tester avec les deux hypothèses suivantes ?
   $$\mathcal{H}_0 \ : \quad \sigma_A^2 = 0$$
   contre
   $$\mathcal{H}_1 \ : \quad \sigma_A^2 \neq 0.$$
   (ii) Que cherche-t-on à tester avec les deux hypothèses suivantes ?
   $$\mathcal{H}_0 \ : \quad \sigma_B^2 = 0$$
   contre
   $$\mathcal{H}_1 \ : \quad \sigma_B^2 \neq 0.$$
   (iii) Que cherche-t-on à tester avec les deux hypothèses suivantes ?
   $$\mathcal{H}_0 \ : \quad \sigma_{AB}^2 = 0$$
   contre
   $$\mathcal{H}_1 \ : \quad \sigma_{AB}^2 \neq 0.$$
3. Décrire le modèle, et ses conditions d'utilisation, qui permet de tester les trois hypothèses nulles suivantes :
   (i) Que cherche-t-on à tester avec les deux hypothèses suivantes ?
   $$\mathcal{H}_0 \ : \quad \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$$

contre

$$\mathcal{H}_1 \ : \quad \text{les valeurs des } \alpha_i \text{ ne sont pas toutes égales, } i = 1, 2, 3, 4.$$

(ii) Que cherche-t-on à tester avec les deux hypothèses suivantes ?

$$\mathcal{H}_0 \ : \quad \sigma_B^2 = 0$$

contre

$$\mathcal{H}_1 \ : \quad \sigma_B^2 \neq 0.$$

(iii) Que cherche-t-on à tester avec les deux hypothèses suivantes ?

$$\mathcal{H}_0 \ : \quad \sigma_{AB}^2 = 0$$

contre

$$\mathcal{H}_1 \ : \quad \sigma_{AB}^2 \neq 0.$$

4. Quelles différences doit-on faire dans l'interprétation des résultats des obtenus aux questions 1., 2. et 3.