T. D. n° 1
Analyse de variance à un facteur

Exercice 1. D'après Prum, *Modèle linéaire. Comparaison de groupes et régression*, Les éditions INSERM, 1996.
On souhaite comparer trois traitements, notés A, B et C contre l'asthme : le traitement B est un nouveau traitement, que l'on souhaite mettre en compétition avec les traitements classiques A et C. On répartit par tirage au sort les patients venant consulter dans un centre de soin, et on leur affecte l'un des trois traitements. On mesure sur chaque patient la durée, en jours, séparant de la prochaine crise d'asthme. Voici les mesures :

Traitement A	Traitement B	Traitement C
26; 27; 35; 36 38; 38; 41; 42 45; 50; 65	29; 42; 44; 44 45; 48; 48; 52 56; 56; 58; 58 60; 61; 63; 63 69	26; 26; 30; 30 33; 36; 38; 38 39; 46; 47; 51 51; 56; 75

Peut-on conclure que les traitements ont une efficacité différente pour le critère « temps séparant la prochaine crise » ?
Pour répondre à cette question, vous allez procéder à une ANOVA à un facteur.

a) Pourquoi faut-il faire ici une ANOVA à un facteur ? Et non pas une analyse de régression ?
b) Écrire le modèle de l'ANOVA à un facteur.
c) Quelles sont les hypothèses que l'on fait ? Les vérifier par la suite.
d) Calculer à l'aide de Minitab le tableau de l'ANOVA à un facteur.
e) Réaliser le test de Fisher au seuil de 5% puis de 1%. Qu'en déduisez-vous ?
f) Donner une estimation de la variance σ^2.
g) Faire des comparaisons multiples. Que concluez-vous ?

Exercice 2. D'après Dagnelie, *Statistique théorique et appliquée*, De-Boeck Université, 1998.
Un même traitement a été appliqué à des graines de sapin de Douglas [1] au cours de périodes de durées différentes, allant de 1 à 12 jours, quatre groupes de 25 graines étant observés pour chaque durée d'application du traitement. En outre, quatre groupes de 25 graines non traités ont été observés. Le tableau suivant indique les

[1] Sapin de Douglas ou Faux Tsuga, Pseudotsuga menziesii. Contrairement à ce que laisse entendre son nom, il n'appartient pas au genre des Sapins (Abies), mais il se raccroche aux Pseudotsuga. Origine : Amérique du Nord (découvert par Menzies en 1795), introduit en France par Douglas, en 1827. Il s'est bien répandu, couvrant 2,4% de la forêt française. Habitat : sols acides, humides. Durée de vie : 500 ans. Taille maximale : 50 m. Port : conique, branches verticillées. Tronc droit.

nombres de germinations qui ont été obtenues pour chacun des 52 groupes de graines.

0	1	2	3	4	5	6	7	8	9	10	11	12
9	6	11	14	18	18	17	20	16	18	13	12	9
5	16	11	15	14	15	20	13	17	17	18	9	11
7	10	16	17	14	15	20	20	16	9	14	10	9
6	14	21	13	19	20	20	14	16	16	9	20	13

Peut-on conclure que les traitements ont une efficacité différente pour le critère « temps (en nombre de jours) du traitement » ?
Pour répondre à cette question, vous allez procéder à une ANOVA à un facteur.

a) Pourquoi faut-il faire ici une ANOVA à un facteur ? Et non pas une analyse de régression ?

b) Écrire le modèle de l'ANOVA à un facteur.

c) Quelles sont les hypothèses que l'on fait ? Les vérifier par la suite.

d) Calculer à l'aide de Minitab le tableau de l'ANOVA à un facteur.

e) Réaliser le test de Fisher au seuil de 5% puis de 1%. Qu'en déduisez-vous ?

f) Donner une estimation de la variance σ^2.

g) Faire des comparaisons multiples. Que concluez-vous ?

Exercice 3. D'après Frontier, Davoult, Gentilhomme, Lagadeuc, *Statistiques pour les sciences de la vie et de l'environnement*, Dunod, 2001.
Pour étudier la maturation gonadique d'une espèce d'échinoderme, des données de poids sec de gonades (en mg) sont récoltées en trois sites différents, pour les mâles et pour les femelles.

• Mâles

Site 1	21,4	20,7	18,3	35,5	23,5	38,7	41,7	35,7	36,2	20,0		
Site 2	75,1	58,9	61,5	58,7	38,4	26,4	38,0	42,5	24,5	32,6	35,5	41,9
Site 3	9,0	18,6	21,7	20,5	22,6	30,4	14,0	17,0	22,1	12,7	27,0	13,9

• Femelles

Site 1	36,7	46,2	45,7	41,2	38,9	44,8	39,7	45,4	29,9	30,7		
Site 2	61,5	53,1	46,3	32,4	23,9	38,0	25,3	51,2	33,7	56,7	48,5	51,4
	23,7	46,0	28,3	32,7								
Site 3	17,5	12,7	20,0	19,1	12,6	16,6	21,8	22,6	22,9	21,0	13,4	14,4

a) Pour chaque sexe, le poids gonadique varie-t-il significativement selon le site ? Pour répondre à cette question, vous serez amené à faire une ANOVA à un facteur pour chaque sexe. Pourquoi procéder à une ANOVA à un facteur ? Écrire le modèle statistique associé. Est-ce que les hypothèses de l'ANOVA à un facteur sont vérifiées ? Dans le cas où votre dernière réponse est négative, que pouvez-vous envisager comme alternative ?

b) Entre quels sites existe-t-il une différence significative ? Pour répondre à cette question, quelle est la procédure sous Minitab que vous devez utiliser ?

c) Tester simultanément les significativités des effets du site, du sexe et de leur interaction. Avec le cours que vous avez à disposition, pouvez-vous répondre à cette question ? Si oui, pourquoi ? Si non, pourquoi ?

d) Interpréter les résultats obtenus à la question précédente.

Exercice 4. D'après Parreins, Techniques Statistiques, moyens rationnels de choix et de décision, Dunod Technique, 1974.
Le contrôle laitier de Haute-Savoie a étudié en 1961 la production de lait, fourni par une même race de vaches, dans plusieurs exploitations : $A, B, \cdots, F, G$.
La quantité de lait produite, indiquée dans le tableau ci-dessous, est-elle liée à l'exploitation ? Eventuellement comment peut-on classer les exploitations ?

A	B	C	D	E	F	G
3 380	2 670	2 400	2 700	3, 380	3, 020	2 780
3 190	2 490	2 330	2 200	3 200	2 970	2 710
2 940	2 460	2 170	2 090	2 810	2 570	2 600
2 600	2 340	2 100	1 940	2 580	2 540	2 560
2 490	2 300	1 990	1 300	2 520	2 290	2 510
2 490	2 130	1 670	1 140	2 430	2 130	2 270
2 450	1 250	1 020		2 100		
2 420						
2 230						

Exercice 5. D'après Parreins, Techniques Statistiques, moyens rationnels de choix et de décision, Dunod Technique, 1974.
On irradie avec des doses croissantes de rayons X (exprimées en roentgens) des espèces vivantes et on calcule le pourcentage de décès :

	Doses en roentgens				
	0	500	1 000	2 000	3 000
1	2, 5	5	10	30	44
2	2, 7	4, 5	6, 5	27	41
3	3, 5	3, 5	8	26	43

a) Les variances sont-elles égales ? Pour répondre à cette question, utilisez une procédure que vous commencez à bien connaître sous Minitab. Dans le cas où elles ne sont pas égales, que pouvez-vous faire ? En quoi cela est-il limitant pour vous ?

b) On transformera les données pour les stabiliser en utilisant la transformation :

$$\phi = \arcsin\left(\sqrt{x}\right).$$

Quel est l'avantage de cette transformation ?

c) Après analyse,on arrive à la conclusion que la dose 500 roentgens n'est pas significativement plus mortelle que la dose zéro. Que doit-on conclure pour les autres doses ?