

Examen final de « Chimiométrie-Plans d'expériences »

- *Le cours, les exercices de travaux dirigés, leurs corrigés ainsi que les notes de cours sont autorisés. Tout autre document est interdit.*
- *Les téléphones portables sont formellement interdits.*
- *Les calculatrices sont autorisées.*
- *Tous les tests seront effectués au seuil de signification $\alpha = 5 \%$.*
- *Les deux exercices sont indépendants et sont à traiter.*
- *Afin de pouvoir traiter les questions, plusieurs résultats numériques et graphiques ont été intégrés au document.*
- *On prendra un soin particulier à préciser quelles sont les hypothèses testées.*

Durée de l'épreuve : 2 heures

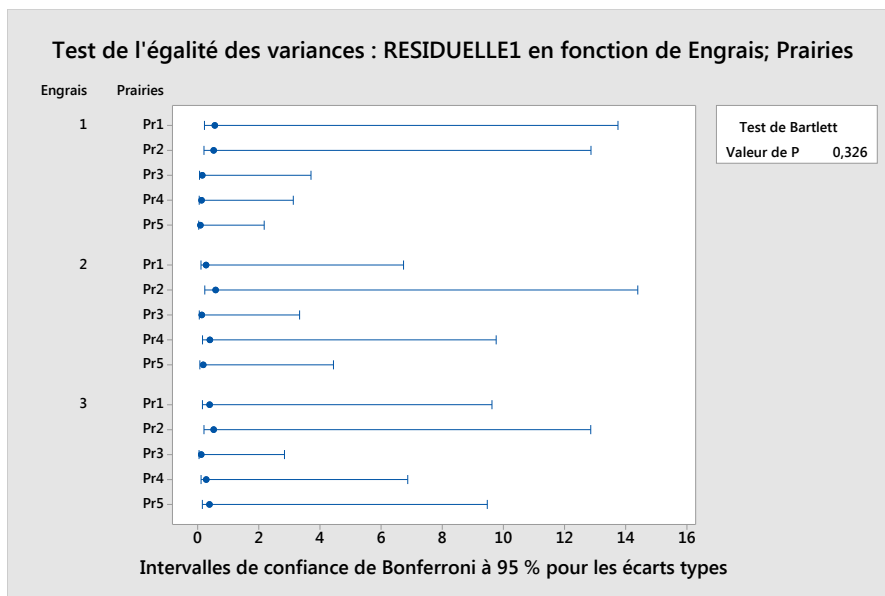
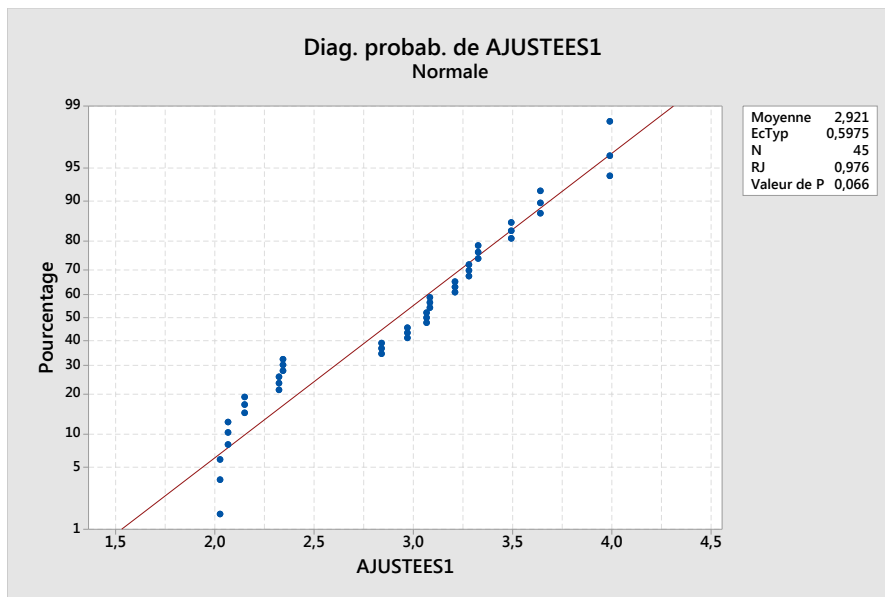
Exercice 1. Rendement fourrager.

Nous mesurons la rendement fourrager, dans cinq prairies. Chaque prairie a été divisée elle-même en neuf parcelles. À l'intérieur d'une prairie, trois parcelles ont été traitées avec l'engrais 1, puis les trois suivantes avec l'engrais 2 et les trois dernières avec l'engrais 3. Les résultats, portant sur 45 parcelles différentes, ont été reportés dans le tableau suivant.

1. Écrire le modèle d'analyse de la variance relatif au modèle. Préciser la nature des facteurs explicatifs ainsi que les conditions d'utilisation à vérifier.
2. Les conditions d'utilisation du modèle sont-elles remplies ?
3. Procéder à l'étude du tableau de l'analyse de la variance.
4. Quelles sont les estimations des coefficients du modèle ?
5. Compte tenu de la nature des facteurs pouvez-vous procéder à des comparaisons multiples ? Il y a-t-il un (ou des) facteur(s) pour le(s)quel(s) cette procédure se justifie ? Si oui procéder aux tests correspondants et interpréter les résultats.
6. Donner et calculer dans le cas échéant les risques d'erreur associés à chacune des décisions prises à la lecture du tableau de l'analyse de la variance obtenu à la question 3..

Engrais	Prairie	Rendement	Engrais	Prairie	Rendement
1	1	2,06	2	3	2,37
1	1	2,99	2	4	3,92
1	1	1,98	2	4	3,43
1	2	1,59	2	4	3,13
1	2	2,63	2	5	3,25
1	2	1,98	2	5	3,11
1	3	2,92	2	5	2,89
1	3	2,85	3	1	3,78
1	3	3,14	3	1	3,08
1	4	2,95	3	1	3,12
1	4	2,70	3	2	4,52
1	4	2,87	3	2	3,98
1	5	2,25	3	2	3,47
1	5	2,08	3	3	3,01
1	5	2,12	3	3	2,99
2	1	2,91	3	3	3,20
2	1	3,27	3	4	3,57
2	1	3,45	3	4	3,01
2	2	1,57	3	4	3,26
2	2	1,82	3	5	3,75
2	2	2,69	3	5	3,21
2	3	2,43	3	5	3,96
2	3	2,17			

Voici les sorties de Minitab



Test de l'égalité des variances : RESIDUELLE1 en fonction de Engrais; Prairies

Méthode

Hypothèse nulle Toutes les variances sont égales
 Hypothèse alternative Au moins une variance est différente
 Seuil de signification $\alpha = 0,05$

La méthode de Bartlett est utilisée.
 Elle est exacte uniquement pour les données normales.

Tests

Méthode	Statistique du test	Valeur de P
Bartlett	15,80	0,326

Modèle linéaire général : Rendement en fonction de Engrais; Prairies

Méthode

Codage de facteur (-1; 0; +1)

Informations sur les facteurs

Facteur	Type	Niveaux	Valeurs
Engrais	Fixe	3	1; 2; 3
Prairies	Fixe	5	Pr1; Pr2; Pr3; Pr4; Pr5

Analyse de variance

Source	DL	SomCar		Valeur F	Valeur de p
		ajust	CM ajust		
Engrais	2	7,497	3,7487	28,95	0,000
Prairies	4	1,373	0,3433	2,65	0,052
Engrais*Prairies	8	6,837	0,8546	6,60	0,000
Erreur	30	3,884	0,1295		
Total	44	19,592			

Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
0,359830	80,17%	70,92%	55,39%

Coefficients

Terme	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	2,9207	0,0536	54,45	0,000	
Engrais					
1	-0,4467	0,0759	-5,89	0,000	1,33
2	-0,0933	0,0759	-1,23	0,228	1,33
Prairies					
Pr1	0,039	0,107	0,37	0,716	1,60
Pr2	-0,226	0,107	-2,11	0,043	1,60
Pr3	-0,134	0,107	-1,25	0,221	1,60
Pr4	0,284	0,107	2,65	0,013	1,60
Engrais*Prairies					
1 Pr1	-0,170	0,152	-1,12	0,271	2,13
1 Pr2	-0,181	0,152	-1,19	0,242	2,13
1 Pr3	0,630	0,152	4,15	0,000	2,13
1 Pr4	0,082	0,152	0,54	0,592	2,13
2 Pr1	0,343	0,152	2,26	0,031	2,13
2 Pr2	-0,574	0,152	-3,79	0,001	2,13
2 Pr3	-0,370	0,152	-2,44	0,021	2,13
2 Pr4	0,382	0,152	2,52	0,017	2,13

Equation de régression

$$\begin{aligned}
 \text{Rendement} = & 2,9207 \\
 & - 0,4467 \text{ Engrais}_1 - 0,0933 \text{ Engrais}_2 + 0,5400 \text{ Engrais}_3 \\
 & + 0,039 \text{ Prairies}_{\text{Pr1}} - 0,226 \text{ Prairies}_{\text{Pr2}} - 0,134 \text{ Prairies}_{\text{Pr3}} \\
 & + 0,284 \text{ Prairies}_{\text{Pr4}} + 0,037 \text{ Prairies}_{\text{Pr5}} \\
 & - 0,170 \text{ Engrais*Prairies}_1 \text{ Pr1} \\
 & - 0,181 \text{ Engrais*Prairies}_1 \text{ Pr2} + 0,630 \text{ Engrais*Prairies}_1 \text{ Pr3} \\
 & + 0,082 \text{ Engrais*Prairies}_1 \text{ Pr4} - 0,361 \text{ Engrais*Prairies}_1 \text{ Pr5} \\
 & + 0,343 \text{ Engrais*Prairies}_2 \text{ Pr1} - 0,574 \text{ Engrais*Prairies}_2 \text{ Pr2} \\
 & - 0,370 \text{ Engrais*Prairies}_2 \text{ Pr3} + 0,382 \text{ Engrais*Prairies}_2 \text{ Pr4} \\
 & + 0,219 \text{ Engrais*Prairies}_2 \text{ Pr5} - 0,173 \text{ Engrais*Prairies}_3 \text{ Pr1} \\
 & + 0,756 \text{ Engrais*Prairies}_3 \text{ Pr2} - 0,260 \text{ Engrais*Prairies}_3 \text{ Pr3} \\
 & - 0,464 \text{ Engrais*Prairies}_3 \text{ Pr4} + 0,142 \text{ Engrais*Prairies}_3 \text{ Pr5}
 \end{aligned}$$

Ajustements et diagnostics pour les observations aberrantes

Observation	Rendement	Valeur ajustée	Résiduelle	Val. résid. norm.	
2	2,990	2,343	0,647	2,20	R
21	2,690	2,027	0,663	2,26	R

R : Valeur résiduelle élevée

Comparaisons pour Rendement

Comparaisons deux à deux de Tukey : réponse = Rendement, terme = Engrais

Informations de groupement avec la méthode de Tukey
et un niveau de confiance de 95 %

Engrais	N	Moyenne	Groupement
3	15	3,46067	A
2	15	2,82733	B
1	15	2,47400	C

Les moyennes ne partageant aucune lettre sont significativement différentes.

Tests de simultanéité de Tukey pour les différences des moyennes

Différence des niveaux Engrais	Différence des moyennes	Erreur type de la différence	IC simultané à 95 %	Valeur de T	Valeur de p ajustée
2 - 1	0,353	0,131	(0,029; 0,678)	2,69	0,030
3 - 1	0,987	0,131	(0,662; 1,311)	7,51	0,000
3 - 2	0,633	0,131	(0,309; 0,958)	4,82	0,000

Exercice 2. Consommation d'électricité (KWh par habitant).

La consommation d'énergie électrique mesure la production des centrales électriques et des centrales électrocalogènes moins la transmission et la distribution, les pertes de transmission et l'utilisation énergétique des centrales électriques et électrocalogènes. La banque mondiale donne la consommation en France sur les 30 dernières années. Les données sont reproduites dans le tableau suivant :

Année	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
Consommation	4700	4931	5244	5493	5634	5721	5843	5951	6360	6476
Année	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Consommation	6453	6531	6619	6905	6833	7016	7149	7238	7347	7302
Année	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Consommation	7531	7665	7655	7536	7520	7650	7340	7736	7292	7342

1. Un économiste souhaite utiliser le modèle linéaire simple pour modéliser le lien entre la consommation et l'année. Spécifier ce modèle (donner les coefficients) et bien identifier chacune des composantes du modèle (la variable explicative et la variable à expliquer) dans le contexte de ce problème.
2. Rappeler les conditions d'utilisation du modèle. Les conditions d'utilisation du modèle sont-elles remplies ?
3. Tester l'hypothèse nulle \mathcal{H}_0 suivante contre l'hypothèse alternative \mathcal{H}_1 avec un test approprié (vous donnerez le nom de ce test) :

$$\mathcal{H}_0 : \beta_1 = 0$$

contre

$$\mathcal{H}_1 : \beta_1 \neq 0.$$

4. Tester l'hypothèse nulle \mathcal{H}_0 suivante contre l'hypothèse alternative \mathcal{H}_1 avec un test approprié (vous donnerez le nom de ce test) :

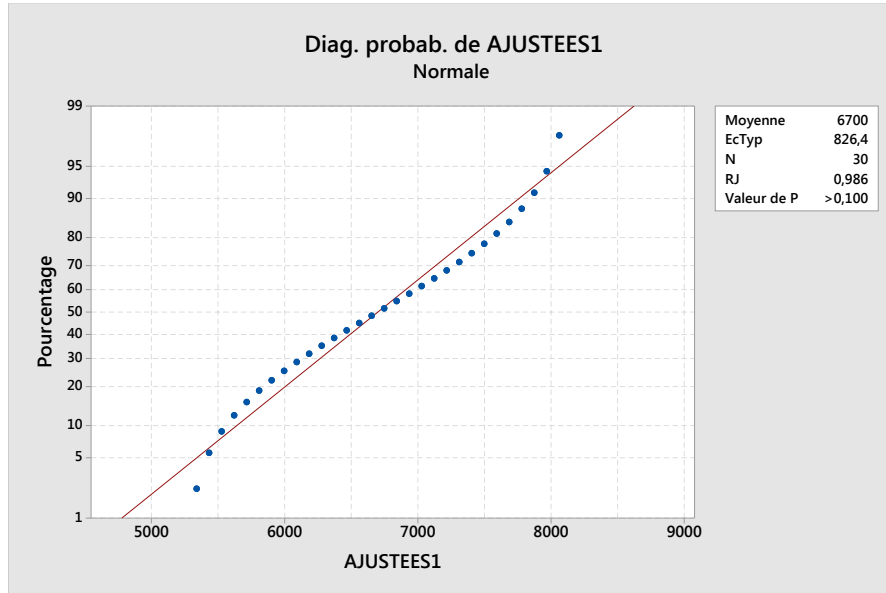
$$\mathcal{H}_0 : \beta_0 = 0$$

contre

$$\mathcal{H}_1 : \beta_0 \neq 0.$$

5. Donner la valeur de la variation qui est expliquée par la droite des moindres carrés ordinaires et la variation qui est inexpliquée par la même droite. En déduire le pourcentage de variation qui est expliqué par la droite des moindres carrés ordinaires.
6. Donner une estimation de la consommation en KWh pour les deux années suivantes : 2013 et 2014 ainsi qu'un intervalle de confiance pour ces deux années. (Servez-vous des sorties de Minitab)

Voici les sorties de Minitab



Analyse de régression : Consommation en fonction de Année

Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Régression	1	19803207	19803207	182,41	0,000
Année	1	19803207	19803207	182,41	0,000
Erreur	28	3039846	108566		
Total	29	22843053			

Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
329,493	86,69%	86,22%	83,83%

Coefficients

Terme	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	-180801	13883	-13,02	0,000	

Année 93,87 6,95 13,51 0,000 1,00

Equation de régression

Consommation = -180801 + 93,87 Année

Ajustements et diagnostics pour les observations aberrantes

Observation	Consommation	Valeur ajustée	Résiduelle	Val. résid. norm.	R
1	4700,0	5339,3	-639,3	-2,08	R
29	7292,0	7967,7	-675,7	-2,18	R
30	7342,0	8061,5	-719,5	-2,34	R

R : Valeur résiduelle élevée

Prévision pour Consommation

Equation de régression

Consommation = -180801 + 93,87 Année

Variable Configuration
Année 2013

Valeur ajustée	ErT ajust	IC à 95 %	IP à 95 %
8155,39	123,386	(7902,64; 8408,13)	(7434,68; 8876,10)

Variable Configuration
Année 2014

Valeur ajustée	ErT ajust	IC à 95 %	IP à 95 %
8249,26	129,499	(7983,99; 8514,52)	(7524,06; 8974,45)

.....