

Examen de « Outils Élémentaires de statistiques appliquées » (VI22GUST)

- *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.*
- *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.*

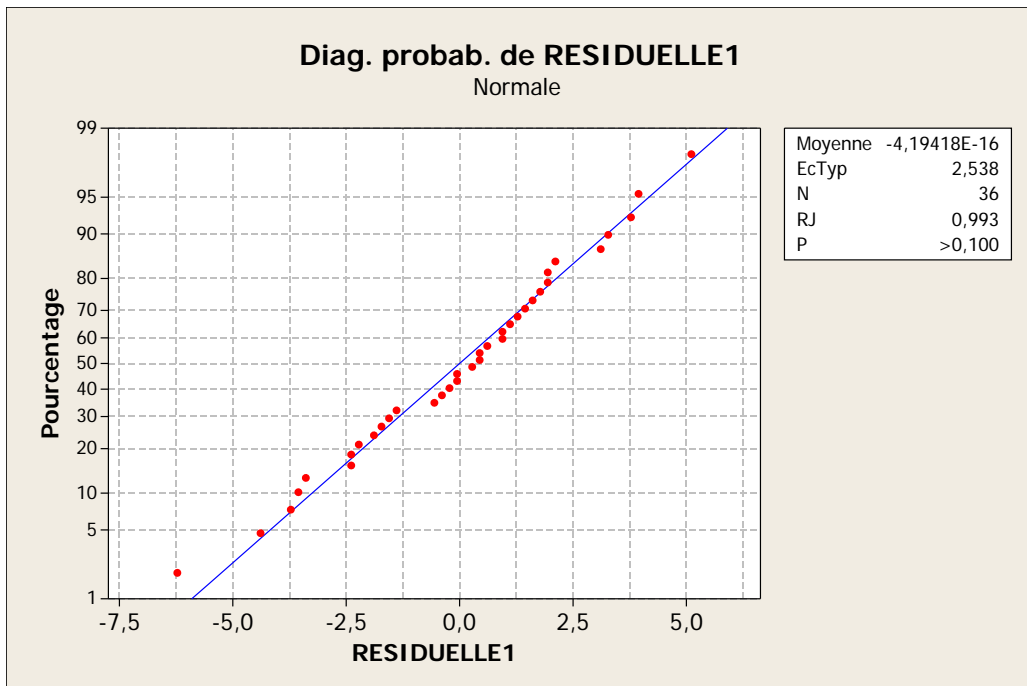
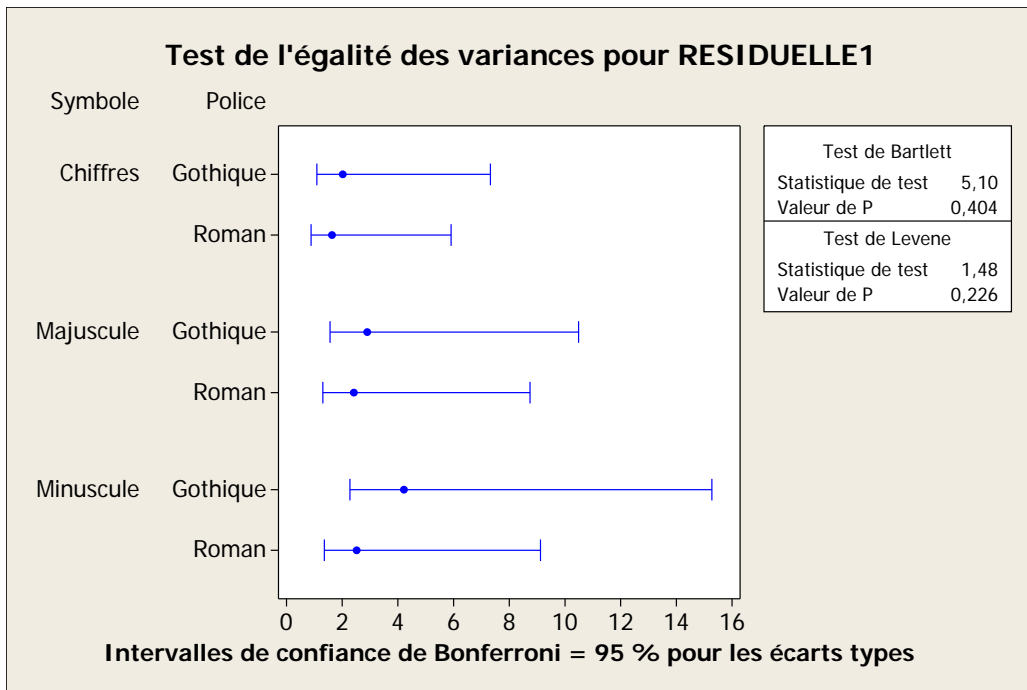
Durée de l'épreuve : 2 heures

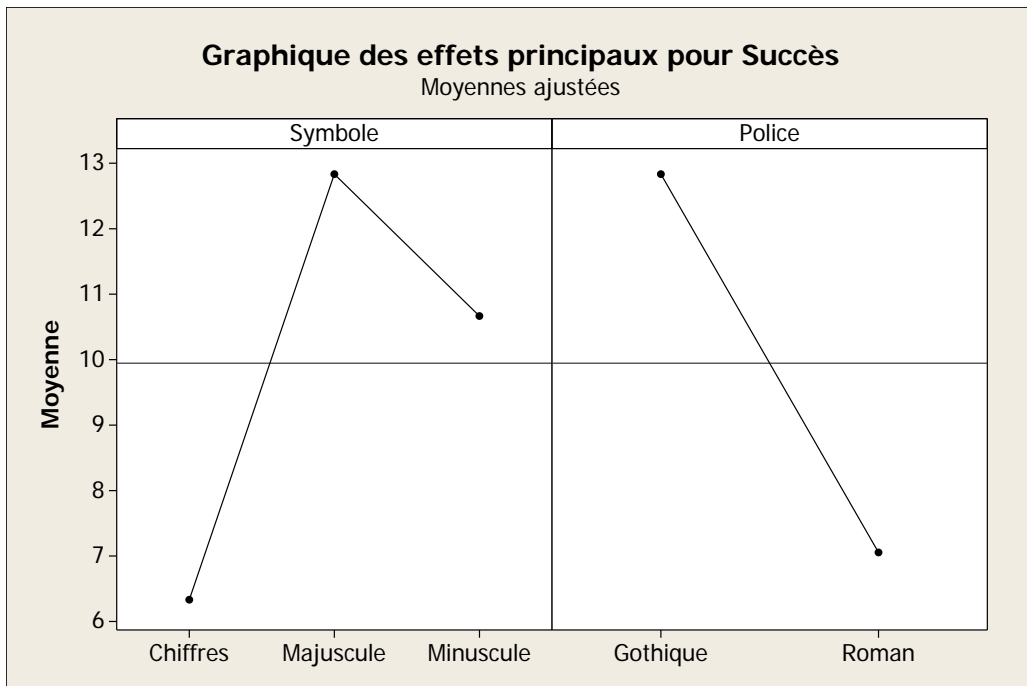
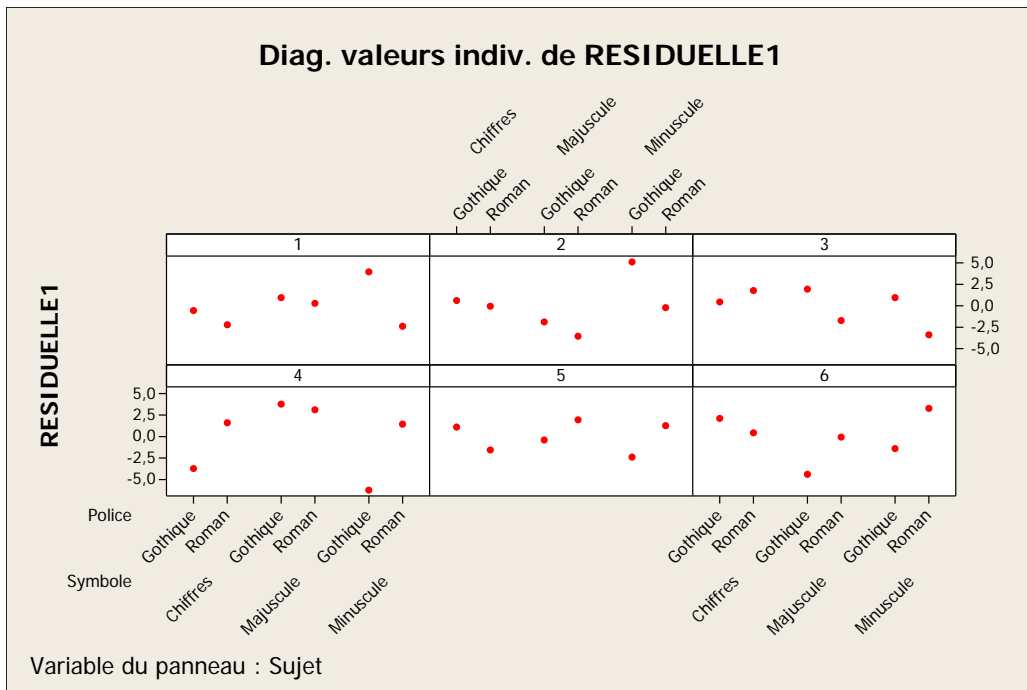
Exercice 1. Une expérience a été menée pour rechercher les effets de deux facteurs (variables indépendantes) sur la reconnaissance de symboles présentés brièvement sur un écran, mesurée par le nombre d'identifications correctes sur un nombre fixe d'essais. Les facteurs sont Symbole (avec les niveaux Chiffres, Minuscule, Majuscule) et la police de caractères (avec les niveaux Gothique et Roman). Chacun des participants de l'expérience est testé sous toutes les six combinaisons de deux facteurs traitements.

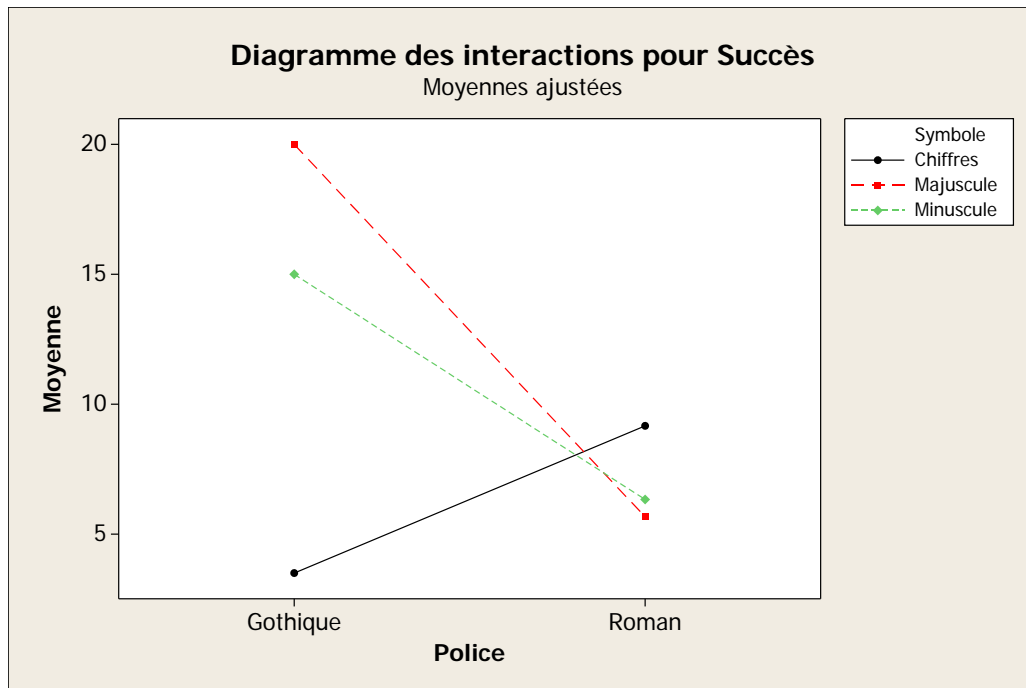
Les résultats sont présentés dans le tableau ci-dessous.

	Chiffres		Minuscule		Majuscule	
Sujet	Gothique	Roman	Gothique	Roman	Gothique	Roman
1	2	6	18	3	20	5
2	4	9	20	6	18	2
3	3	10	15	2	21	3
4	1	12	10	9	25	10
5	5	8	13	8	20	8
6	6	10	14	10	16	6

1. Écrire le modèle d'analyse de la variance relatif au modèle dont les résultats sont reproduits dans la suite. On précisera la nature des facteurs explicatifs ainsi que les hypothèses faites.
2. Indiquer pourquoi il faut considérer dans le modèle un troisième facteur « Sujets », de nature aléatoire.
3. Les hypothèses du modèle sont-elles vérifiées ? Calculer les estimations de tous les paramètres du modèle.
4. Procéder à l'étude du tableau de l'analyse de la variance.
5. Compte tenu de la nature des 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 et interpréter les résultats.







Modèle linéaire général : Succès en fonction de Symbole; Police; Sujet

Facteur	Type	Niveaux	Valeurs
Symbole	fixe	3	Chiffres; Majuscule; Minuscule
Police	fixe	2	Gothique; Roman
Sujet	aléatoire	6	1; 2; 3; 4; 5; 6

Analyse de la variance pour Succès, avec utilisation de la somme des carrés ajustée pour les tests

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Symbole	2	262,89	262,89	131,44	14,58	0,000
Police	1	300,44	300,44	300,44	33,32	0,000
Symbole*Police	2	637,56	637,56	318,78	35,35	0,000
Sujet	5	21,56	21,56	4,31	0,48	0,789
Erreur	25	225,44	225,44	9,02		
Total	35	1447,89				

S = 3,00296 R carré = 84,43 % R carré (ajust) = 78,20 %

Terme	Coeff	Coef ErT	T	P
Constante	9,9444	0,5005	19,87	0,000
Symbole				
Chiffres	-3,6111	0,7078	-5,10	0,000
Majuscule	2,8889	0,7078	4,08	0,000
Police				
Gothique	2,8889	0,5005	5,77	0,000
Symbole*Police				
Chiffres Gothique	-5,7222	0,7078	-8,08	0,000
Majuscule Gothique	4,2778	0,7078	6,04	0,000
Sujet				
1	-0,944	1,119	-0,84	0,407
2	-0,111	1,119	-0,10	0,922
3	-0,944	1,119	-0,84	0,407
4	1,222	1,119	1,09	0,285
5	0,389	1,119	0,35	0,731

Observations aberrantes pour Succès

Observation	Succès	Valeur ajustée	Ajust ErT	Valeur résiduelle	Valeur résiduelle normalisée
14	20,0000	14,8889	1,6599	5,1111	2,04 R
16	10,0000	16,2222	1,6599	-6,2222	-2,49 R

R indique une observation ayant une valeur résiduelle normalisée importante

Espérance mathématique des carrés moyens, en utilisant la somme des carrés ajustée

Source	Espérance mathématique des carrés moyens pour chaque terme
1 Symbole	(5) + Q[1; 3]
2 Police	(5) + Q[2; 3]
3 Symbole*Police	(5) + Q[3]
4 Sujet	(5) + 6,0000 (4)
5 Erreur	(5)

Termes d'erreur pour les tests, en utilisant la somme des carrés ajustée

Synthèse du
Erreur Erreur carré moyen

	Source	- DL	- CM	de l'erreur
1	Symbole	25,00	9,02	(5)
2	Police	25,00	9,02	(5)
3	Symbole*Police	25,00	9,02	(5)
4	Sujet	25,00	9,02	(5)

Composantes de la variance, avec la somme des carrés ajustée

	Valeur
Source estimée	
Sujet	-0,7844
Erreur	9,0178

Moyennes issues des moindres carrés pour Succès

	Moyenne
Symbole	
Chiffres	6,333
Majuscule	12,833
Minuscule	10,667
Police	
Gothique	12,833
Roman	7,056
Symbole*Police	
Chiffres Gothique	3,500
Chiffres Roman	9,167
Majuscule Gothique	20,000
Majuscule Roman	5,667
Minuscule Gothique	15,000
Minuscule Roman	6,333

Intervalles de confiance simultanés de Tukey = 95,0 %

Variable de réponse Succès

Toutes les comparaisons deux à deux sur les niveaux de Symbole

Symbole = Chiffres soustrait de :

Symbole	Inférieur	Centré	Supérieur	
Majuscule	3,449	6,500	9,551	(-----*-----)
Minuscule	1,282	4,333	7,385	(-----*-----)

-----+-----+-----+-----+-----
 -4,0 0,0 4,0 8,0

Symbole = Majuscule soustrait de :

Symbole	Inférieur	Centré	Supérieur	---+-----+-----+-----+---
Minuscule	-5,218	-2,167	0,8848	(-----*-----)
				---+-----+-----+-----+---
				-4,0 0,0 4,0 8,0

Tests de simultan  it   de Tukey

Variable de r  ponse Succ  s

Toutes les comparaisons deux    deux sur les niveaux de Symbole

Symbole = Chiffres soustrait de :

	Diff��rence	Erreur type	Valeur	Valeur de
Symbole	des moyennes	de la	de T	p ajust��e
		diff��rence		
Majuscule	6,500	1,226	5,302	0,0001
Minuscule	4,333	1,226	3,535	0,0045

Symbole = Majuscule soustrait de :

	Diff��rence	Erreur type	Valeur	Valeur de
Symbole	des moyennes	de la	de T	p ajust��e
		diff��rence		
Minuscule	-2,167	1,226	-1,767	0,2011

Test de l'  galit   des variances : RESIDUELLE1 en fonction de Symbole; Police

Intervalles de confiance de Bonferroni = 95 % pour les   carts types

Symbole	Police	N	Inf��rieur	EcTyp	Sup��rieur
Chiffres	Gothique	6	1,09031	2,02118	7,3207
Chiffres	Roman	6	0,88029	1,63186	5,9106
Majuscule	Gothique	6	1,56207	2,89572	10,4882
Majuscule	Roman	6	1,30246	2,41446	8,7451
Minuscule	Gothique	6	2,27496	4,21725	15,2748
Minuscule	Roman	6	1,35836	2,51808	9,1204

Test de Bartlett (loi normale)

R  sultat du test = 5,10 ; valeur p = 0,404

Test de Levene (toute loi de distribution continue)

R  sultat du test = 1,48 ; valeur p = 0,226

.....

Exercice 2. Box et Cox ont reproduit les résultats d'une expérience conçue pour évaluer les effets d'agents toxiques dans certaines conditions expérimentales. Des groupes de quatre animaux ont été aléatoirement attribués à ces trois poisons pour chacune des quatre situations, appelées traitements A, B, C et D, étudiées dans l'expérience.

Les temps de survie, observés sur 48 animaux différents et exprimés en dizaines d'heures, ont été reproduits dans le tableau ci-dessous.

	Traitement			
Poison	A	B	C	D
I	0,31	0,82	0,43	0,45
	0,45	1,10	0,45	0,71
	0,46	0,88	0,63	0,70
	0,43	0,72	0,76	0,62
II	0,36	0,92	0,44	0,56
	0,29	0,61	0,35	0,82
	0,40	0,49	0,31	0,71
	0,23	0,84	0,40	0,38
III	0,22	0,30	0,13	0,20
	0,21	0,37	0,25	0,36
	0,08	0,38	0,24	0,31
	0,33	0,29	0,32	0,33

1. Donner le modèle complet de l'analyse de la variance. Donner les effets principaux et interactions qu'il est possible de tester. Les conditions d'utilisation du modèle sont-elles remplies ?
2. Effectuer l'analyse de la variance en indiquant comment sont obtenus les degrés de liberté.
3. Donner, le cas échéant, les effets et les interactions qui ont un effet significatif avec le risque de mauvaise décision associé. Analyser ces effets à l'aide de graphiques et / ou, si cela est possible, des comparaisons multiples.
4. Évaluer, le cas échéant, le risque de mauvaise décision pour les effets principaux et les interactions qui ne sont pas significatifs.
5. Donner un bref commentaire de l'ensemble des résultats.

Modèle linéaire général : Durée en fonction de Poison; Traitement

Facteur	Type	Niveaux	Valeurs
Poison	fixe	3	I; II; III
Traitement	fixe	4	A; B; C; D

Analyse de la variance pour Durée, avec utilisation de la somme des carrés ajustée pour les tests

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Poison	2	1,02084	1,02084	0,51042	34,32	0,000
Traitement	3	0,74477	0,74477	0,24826	16,69	0,000
Poison*Traitement	6	0,18015	0,18015	0,03002	2,02	0,088
Erreur	36	0,53542	0,53542	0,01487		
Total	47	2,48118				

S = 0,121955 R carré = 78,42 % R carré (ajust) = 71,83 %

Terme	Coeff	Coef ErT	T	P
Constante	0,46562	0,01760	26,45	0,000
Poison				
I	0,15438	0,02489	6,20	0,000
II	0,04125	0,02489	1,66	0,106
Traitement				
A	-0,15146	0,03049	-4,97	0,000
B	0,17771	0,03049	5,83	0,000
C	-0,07313	0,03049	-2,40	0,022
Poison*Traitement				
I A	-0,05604	0,04312	-1,30	0,202
I B	0,08229	0,04312	1,91	0,064
I C	0,02062	0,04312	0,48	0,635
II A	-0,03542	0,04312	-0,82	0,417
II B	0,03042	0,04312	0,71	0,485
II C	-0,05875	0,04312	-1,36	0,181

Observations aberrantes pour Durée

Observation	Durée	Valeur ajustée	Ajust ErT	Valeur résiduelle	Valeur résiduelle normalisée
14	1,10000	0,88000	0,06098	0,22000	2,08 R
19	0,49000	0,71500	0,06098	-0,22500	-2,13 R
44	0,38000	0,61750	0,06098	-0,23750	-2,25 R

R indique une observation ayant une valeur résiduelle normalisée importante

Espérance mathématique des carrés moyens, en utilisant la somme des carrés ajustée

Espérance

Source	mathématique des carrés moyens pour chaque terme
1 Poison	(4) + Q[1; 3]
2 Traitement	(4) + Q[2; 3]
3 Poison*Traitement	(4) + Q[3]
4 Erreur	(4)

Termes d'erreur pour les tests, en utilisant la somme des carrés ajustée

Source	Erreur - DL	Erreur - CM	Synthèse du carré moyen de l'erreur
1 Poison	36,00	0,01487	(4)
2 Traitement	36,00	0,01487	(4)
3 Poison*Traitement	36,00	0,01487	(4)

Composantes de la variance, avec la somme des carrés ajustée

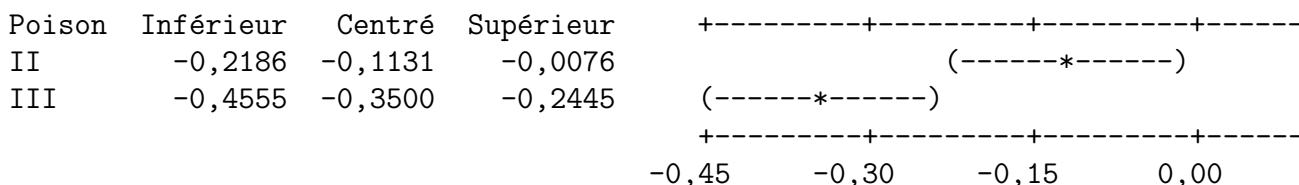
Source	Valeur estimée
Erreur	0,01487

Moyennes issues des moindres carrés pour Durée

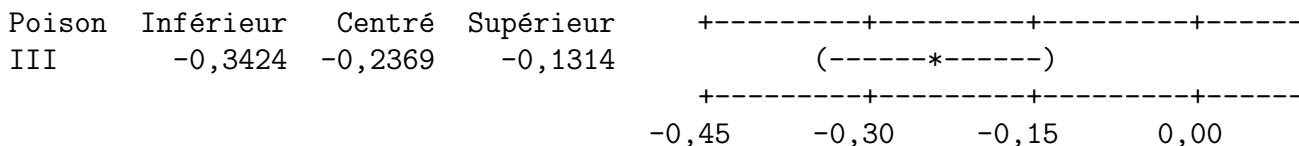
Poison	Moyenne	ErT moyenne
I	0,6200	0,03049
II	0,5069	0,03049
III	0,2700	0,03049
Traitement		
A	0,3142	0,03521
B	0,6433	0,03521
C	0,3925	0,03521
D	0,5125	0,03521
Poison*Traitement		
I A	0,4125	0,06098
I B	0,8800	0,06098
I C	0,5675	0,06098
I D	0,6200	0,06098
II A	0,3200	0,06098
II B	0,7150	0,06098
II C	0,3750	0,06098
II D	0,6175	0,06098
III A	0,2100	0,06098

III	B	0,3350	0,06098
III	C	0,2350	0,06098
III	D	0,3000	0,06098

Intervalles de confiance simultanés de Tukey = 95,0 %
 Variable de réponse Durée
 Toutes les comparaisons deux à deux sur les niveaux de Poison
 Poison = I soustrait de :



Poison = II soustrait de :



Tests de simultanéité de Tukey
 Variable de réponse Durée
 Toutes les comparaisons deux à deux sur les niveaux de Poison
 Poison = I soustrait de :

Poison	Différence des moyennes	Erreur type de la différence	Valeur de T	Valeur de p ajustée
II	-0,1131	0,04312	-2,624	0,0332
III	-0,3500	0,04312	-8,117	0,0000

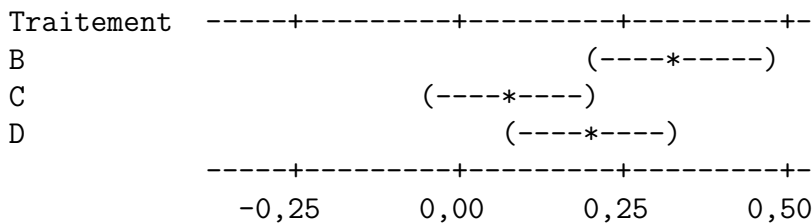
Poison = II soustrait de :

Poison	Différence des moyennes	Erreur type de la différence	Valeur de T	Valeur de p ajustée
III	-0,2369	0,04312	-5,494	0,0000

Intervalles de confiance simultanés de Tukey = 95,0 %
 Variable de réponse Durée

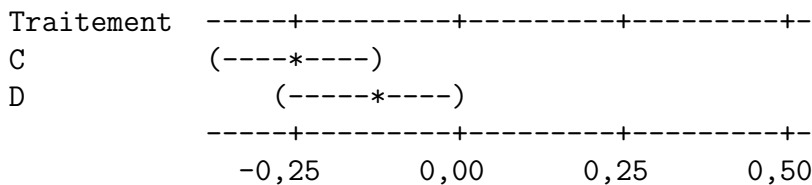
Toutes les comparaisons deux à deux sur les niveaux de Traitement
 Traitement = A soustrait de :

Traitement	Inférieur	Centré	Supérieur
B	0,19503	0,32917	0,4633
C	-0,05580	0,07833	0,2125
D	0,06420	0,19833	0,3325

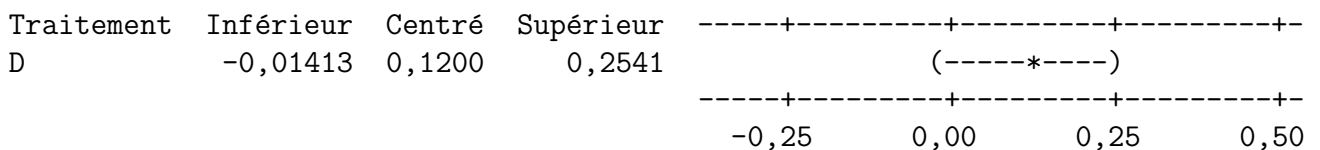


Traitement = B soustrait de :

Traitement	Inférieur	Centré	Supérieur
C	-0,3850	-0,2508	-0,1167
D	-0,2650	-0,1308	0,0033



Traitement = C soustrait de :



Tests de simultanéité de Tukey

Variable de réponse Durée

Toutes les comparaisons deux à deux sur les niveaux de Traitement
 Traitement = A soustrait de :

Traitement	Différence des moyennes	Erreur type de la différence	Valeur de T	Valeur de p ajustée
B	0,32917	0,04979	6,611	0,0000
C	0,07833	0,04979	1,573	0,4062

D 0,19833 0,04979 3,984 0,0017

Traitement = B soustrait de :

	Différence	Erreur type	Valeur	Valeur de
Traitement	des moyennes	de la	de T	p ajustée
		différence		
C	-0,2508	0,04979	-5,038	0,0001
D	-0,1308	0,04979	-2,628	0,0580

Traitement = C soustrait de :

	Différence	Erreur type	Valeur	Valeur de
Traitement	des moyennes	de la	de T	p ajustée
		différence		
D	0,1200	0,04979	2,410	0,0932

Test de l'égalité des variances : RESIDUELLE1 en fonction de Poison; Traitement

Intervalles de confiance de Bonferroni = 95 % pour les écarts types

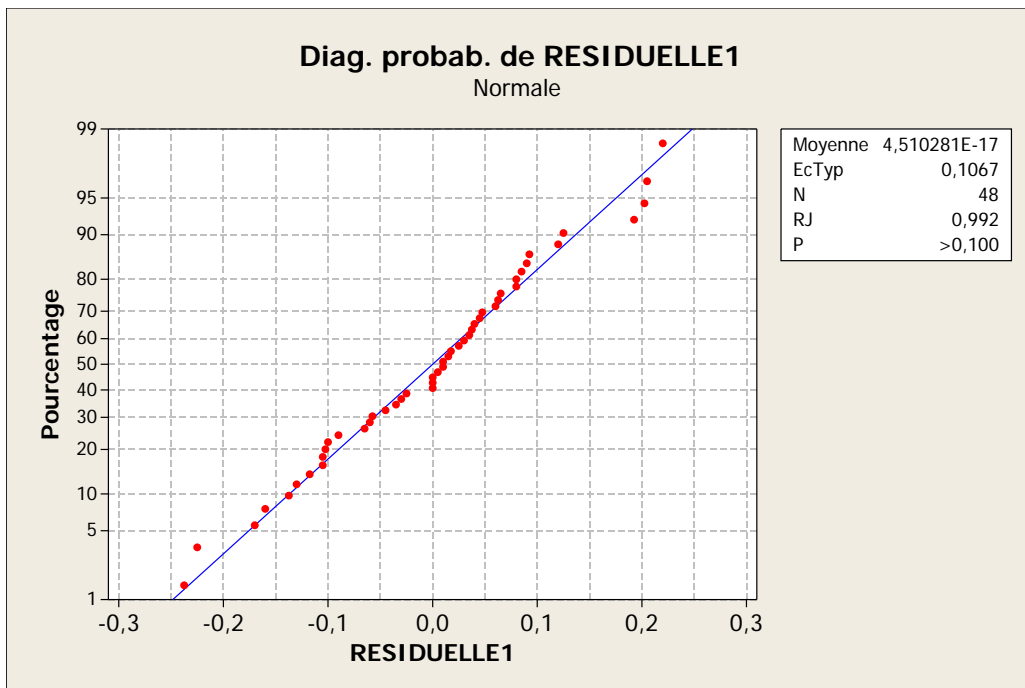
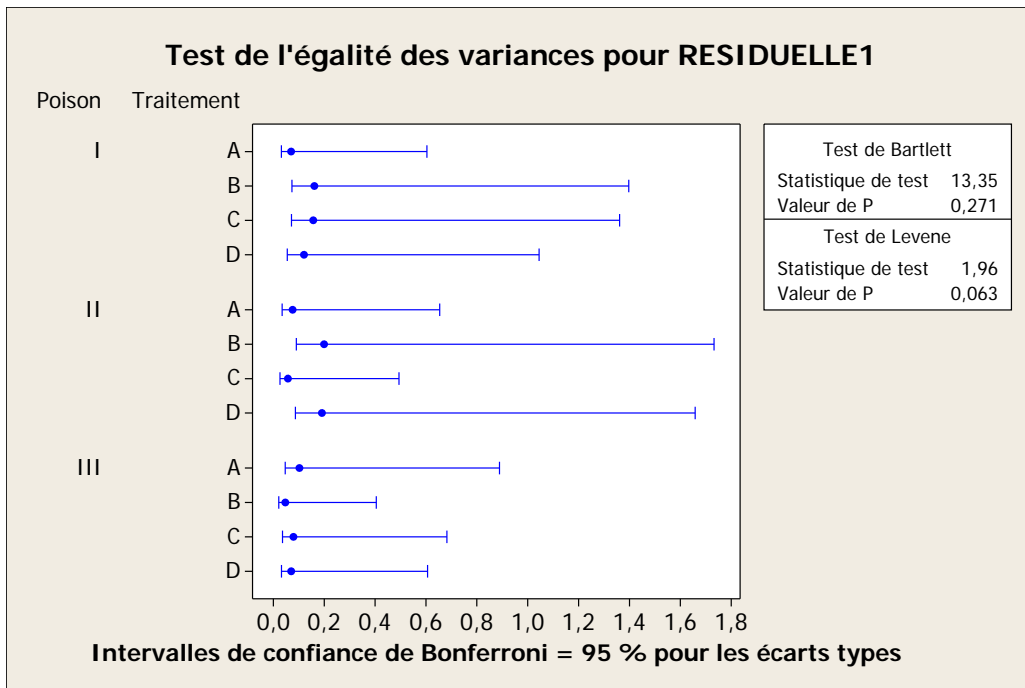
Poison	Traitement	N	Inférieur	EcTyp	Supérieur
I	A	4	0,0313706	0,069462	0,60340
I	B	4	0,0726347	0,160831	1,39710
I	C	4	0,0707740	0,156711	1,36131
I	D	4	0,0543198	0,120277	1,04482
II	A	4	0,0339968	0,075277	0,65391
II	B	4	0,0900604	0,199416	1,73227
II	C	4	0,0256803	0,056862	0,49395
II	D	4	0,0861934	0,190853	1,65789
III	A	4	0,0462039	0,102307	0,88871
III	B	4	0,0210218	0,046547	0,40434
III	C	4	0,0354650	0,078528	0,68215
III	D	4	0,0315058	0,069761	0,60600

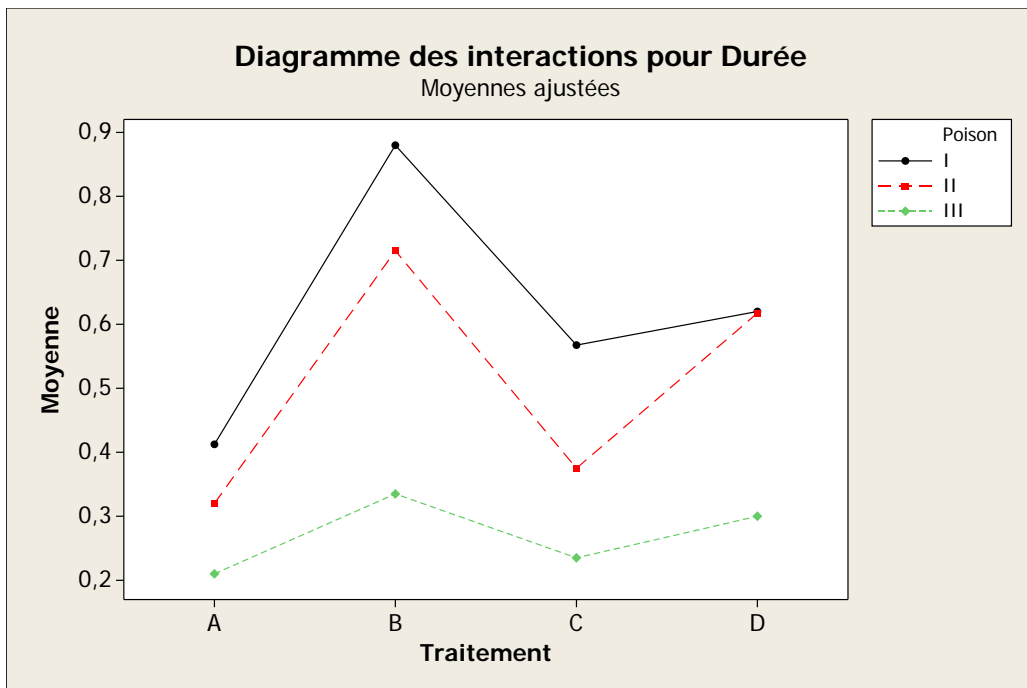
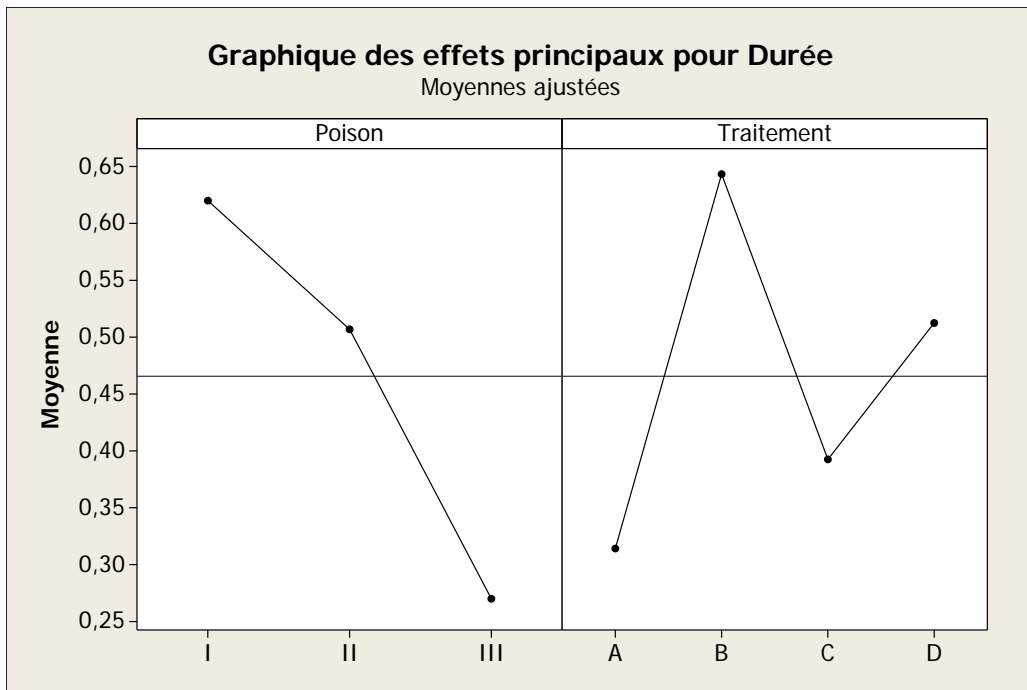
Test de Bartlett (loi normale)

Résultat du test = 13,35 ; valeur p = 0,271

Test de Levene (toute loi de distribution continue)

Résultat du test = 1,96 ; valeur p = 0,063





.....