Université Mohamed Khider de Biskra. **Faculté des Sciences Exactes et de la Vie**
Département: Biologie **Module: Bio statistique**
Année Univ: 2021-2022. **Master 1 LMD(Bioch+Micro+BA+BV)**
Le 16/06/2022. **(60 Minutes)**

Examen de Rattrapage S2

Exercice 01:(10)

On a étudié les longueurs respectives des 2 (deux) paires d'ailes d'une espèce de guêpe (Vespa sp) sur un échantillon de 07 individus. Soit la variable indépendante : la longueur d'une aile de la première paire et la variable dépendante : celle de l'aile de la deuxième paire mesurée sur le même individu. On a obtenu les résultats suivants:

Individus	1	2	3	4	5	6	7
La longueur d'une aile de la première paire	294	271	314	356	383	369	402
La longueur d'une aile de la deuxième paire	624	661	728	782	819	869	1023

1) Déterminer l'équation de la régression linéaire simple.
2) Calculer le coefficient de corrélation entre la longueur d'une aile de la première paire et la longueur d'une aile de la deuxième paire (r), et interpréter ce coefficient.
3) Calculer le coefficient de détermination de la régression linéaire R, interpréter ce coefficient.
4) Tester s'il y a une liaison significative au seuil de confiance c'est 05%.

$t_5^{5\%} = 2.571$	$t_7^{5\%} = 2.365$	$t_6^{5\%} = 2.447$	$t_5^{5\%} = 2.571$

Exercice 01:(10 points)

Un producteur de sauce italienne souhaite comparer 4 conditionneuses : C1, C2, C3 et C4, quant à la teneur en viande des boites qu'elles remplissent. Il souhaite ainsi savoir si les teneurs en viande sont identiques d'une journée à l'autre. Il étudie 4 journées de fabrication : J_1, J_2, J_3 et J_4. Par jour et par conditionneuse, 2 boîtes sont analysées. On suppose ici que les conditions de réalisation d'ANOVA sont satisfaites. Les résultats sont

	C1	C2	C3	C4
J_1	13	15	13.5	13.2
	12.4	15.4	12.3	15.4
J_2	12	15.1	13.1	12.7
	11.8	15.4	12	13
J_3	15	15	15.1	13
	14.8	14.6	14.7	12.2
J_4	13.2	14.2	11.5	13.2
	12	14.6	12.2	12.5

1) Quelles sont les variables étudiés dans cette expérience, et déterminer ses natures, on déterminera le nombre total N, ainsi que nombre des répétitions.
2) Déterminer tous les hypothèses qui se trouvent dans cette expérience.
3) Tracer la table des moyennes cellulaires, et les moyennes marginales, et la moyenne totale.
4) Tracer la table d'analyse de la variance.
5) Que peut-on dire sur l'effet des facteurs étudiés ?

Correction

Exercice N°1 :

1) $\bar{x} = 341{,}285$. (0, 25 point)

$\bar{y} = 786{,}571$. (0, 25 point)

$\sum x^2 = 829663$. (0, 25 point)

$\sum y^2 = 4440256$. (0, 25 point)

$\sum xy = 1915155$. (0, 25 point)

$\sum x^2 - n(\bar{x})^2 (0,25 \text{ point}) = 14334{,}841$.(0, 25 point)

$\sum y^2 - n(\bar{y})^2 (0,25 \text{ point}) = 109398{,}433$.(0, 25 point)

Alors

$\hat{\alpha} = \frac{\sum xy - n(\bar{x})(\bar{y})}{\sum \quad x^2 - n(\bar{x})^2} (0{,}5 \text{ point}) = \frac{1915155 - 7(341{,}285)(786{,}571)}{14334{,}841}$=2,514. (0,5 point)

$\hat{\beta} = \bar{y} - \hat{\alpha}\bar{x} (0{,}5 \text{ point}) = 786{,}571 - 2{,}514(341{,}285) = -71{,}419$. (0,5 point)

2) Coefficient de corrélation

$$r = \hat{\alpha}\frac{S_X}{S_Y} = (0{,}5 \text{ point}) 2{,}514\sqrt{\frac{14334{,}841}{109398{,}433}} = 0{,}910 > 0 (0{,}5 \text{ point}).$$

Le coefficient r est positive ce qui signifier que la relation entre X et Y est croissante (0,25 point), mais très bien (0,25 point).

3) Coefficient de détermination de la régression linéaire

$$R = r^2 (0{,}5 \text{ point}) = (0{,}910)^2 = 0{,}828. (0{,}25 \text{ point})$$

Alors le taux de corrélation c'est 82,81%.(0,25 point)

4) Test de Student

On calcule tous d'abord la variance de l'estimateur de $\hat{\alpha}$

$$Var(\hat{\alpha}) = \frac{S^2}{\sum x^2 - n(\bar{x})^2} (0{,}5 \text{ point}) = \frac{\frac{\sum y^2 - n(\bar{y})^2 - (\hat{\alpha})^2\left(\sum x^2 - n(\bar{x})^2\right)}{n-2} (0{,}5 \text{ point})}{\sum x^2 - n(\bar{x})^2 \ (0{,}5 \text{ point})}$$
$$= \frac{\frac{109398{,}433 - (2{,}514)^2(14334{,}841)}{7-2}}{14334{,}841} = \frac{15265{,}934 (0{,}5 \text{ point})}{14334{,}841} = 0{,}262. (0{,}5 \text{ point})$$

On veut tester si X et Y sont significativement liés, pour cela on doit faire l'hypothèse suivante

$$\begin{cases} H0: \hat{\alpha} = 0, \\ H1: \hat{\alpha} \neq 0. \end{cases} (0{,}25 \text{ point})$$

Alors

$$T = \frac{\widehat{\alpha} - 0}{\sqrt{Var(\widehat{\alpha})}} (0{,}25 \text{ point}) = \frac{2{,}514}{\sqrt{0{,}262}} = 4{,}919 (0{,}25 \text{ point}).$$

On calcule maintenant le fractile t=2,571(0,25 point).

Alors $T > t$. Alors on accepte H1, donc il y a une liaison significative. (0,25 point)

Exercice N°2 :

1) Les variables sont : Qualitatives (0,25 point): Facteur A : Jours (p=4) (0,25 point), et Facteur B : Conditions (q=4) (0,25 point).
   Quantitative (0,25 point): Teneur de viande (0,25 point), avec n=2, alors N=npq=2(4)(4)=32 (0,25 point).

2) H_0 : (Il n' y a pas un effet de jours sur Teneur des viande) (0,25 point).
   H_1 : (Il y a un effet de jours sur Teneur des viande)

   Hypothèses pour facteur B= Conditions:

   H'_0 : (Il n' y a pas un effet de Conditions sur Teneur des viande) (0,25 point).
   H'_1 : (Il y a un effet de Condition sur Teneur des viande)

   Hypothèses pour interaction: Hypothèses pour facteur d'interaction AB

   H''_0 : (Il n' y a pas un effet d'interaction sur Teneur). (0,25 point)
   H''_1 : (Il y a un effet d'interaction sur Teneur).

3) Tableau des moyennes :

	C1 (0,25 point)	C2 (0,25 point)	C3 (0,25 point)	C4 (0,25 point)	Moyennes marginales(0,25 point)
J1	12.7	15.2	12.9	14.3	13.775
J2	11.9	15.25	12.55	12.85	13.137
J3	14.9	14.8	14.9	12.6	14.3
J4	12.6	14.4	11.85	12.85	12.925
Moyennes marginales	13.025	14.912	13.05	13.15	13.534

$\sum \bar{x}_i^2 = 733.877$ (0,25 point)

$\sum \bar{x}_j^2 = 735.243$ (0,25 point)

$\sum\sum \bar{x}_{ij}^2 = 2953.512$ (0,25 point)

$\sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{k=1}^{2} x_{ijk}^2 = 5912.930$ (0,25 point)

$$N(\bar{X})^2 = 5861.412 (0{,}25 \text{ point})$$

$$SCE_T = \sum\sum\sum x_{ijk}^2 - N(\bar{X})^2 (0{,}25 \text{ point}) = 5912.930 - 5861.412 = 51.518 (0{,}25 \text{ point})$$

$SCE_R = \sum \sum \sum x_{ijk}^2 - - n \sum \sum \bar{x}_{ij}^2$ (0,25 point) = 5.906 (0,25 point)

$SCE_A = qn \sum \overline{x_i^2} - N(\bar{X})^2$(0,25 point) = 9.604 (0,25 point)

$SCE_B = pn\sum \overline{x_J^2} - N(\bar{X})^2(0,25 \text{ point}) = 20.532\ (0,25 \text{ point})$

$SCE_{AB} = SCE_T - SCE_R - SCE_A - SCE_B(0,25 \text{ point}) = 15.476\ (0,25 \text{ point})$

4) Tableau analyse de variance

Source des variations	SCE	DDL	CM $(0,5 \text{ point})$	Test de Fisher $(0,5 \text{ point})$
Total	51.518	N-1=31	1.661	
Facteur A	9.604	p-1=3	3.201	$F_{obs}^{A} = 8.674$
Facteur B	20.532	q-1=3	6.844	$F_{obs}^{B} = 18.547$
Facteur AB	15.476	(p-1)(q-1)=9	1.719	$F_{obs}^{AB} = 4.658$
Résiduelle	5.906	pq(n-1)=16	0.369	

5) Conclusion :

$F_\alpha^A = F(3,16) = 3.24\ (0,25 \text{ point})$

$F_\alpha^B = F(3,16) = 3.24\ (0,25 \text{ point})$

$F_\alpha^{AB} = F(9,16) = 2.54\ (0,25 \text{ point})$

Alors on remarque ici que
$F_{obs}^{AB} = 4.658 > F_\alpha^{AB} = F(9,16) = 2.54\ (0,5 \text{ point})$

Alors on rejette H''_0, alors il y a un effet d'interaction, c'est-à-dire que tous les moyennes cellulaires sont significativement différentes. $(0,5 \text{ point})$