

154 et 012^{ème} promotions –3^{ème} année

2012 - 2013

RMO : Nathalie Robert.....Intervenant : N. Robert.....

Intitulé du Module : Modèle linéaire (apprentissage) Page 1 sur 11

Date :18/04/13 Durée : 2 h Sans document Avec calculatrice

Nom de l'étudiant : Prénom : N° liste promo : Place :

Lisez attentivement et tranquillement l'énoncé avant de répondre. Quand nécessaire, utilisez les tables jointes. La qualité de présentation des formules de calcul intervient dans la notation ainsi que l'argumentation des réponses. La plage pour répondre peut se situer sur deux pages.

Exercice 1

Une conserverie s'intéresse à l'influence de la variété de haricot vert sur le diamètre moyen du haricot dans une zone de production dans laquelle elle envisage de s'approvisionner. Des échantillons aléatoires ont été prélevés sur des parcelles similaires du point de vue des conditions de culture (sol et pratique culturale). L'effectif prélevé dépend de l'importance de chaque variété dans la zone de production.

Les données sont synthétisées dans le tableau suivant :

| variété | moyenne (mm) | effectif |
|-----------|--------------|----------|
| variété A | 6,37 | 35 |
| variété B | 5,41 | 42 |
| variété C | 6,95 | 33 |

L'ensemble des échantillons est caractérisé par les valeurs suivantes : moyenne 6,18 mm, écart-type 1.57 mm.

1) Quel type de modèle allez-vous choisir et pourquoi ? Ecrivez ce modèle.

2) Vous allez construire le tableau d'analyse de la variance puis l'interpréter. Le risque de première espèce choisi pour cette étude vaut 5%.

2.1) Le calcul de la somme de carrés d'écarts résiduelle n'a pu être terminé. En effet la dernière valeur pour chaque variété n'a pu être introduite dans le calcul. Le résultat intermédiaire vaut 208,92.

Nom de l'étudiant : **Prénom :** **Place :**

- a) Dans quelle unité est exprimée cette somme de carrés d'écarts ?

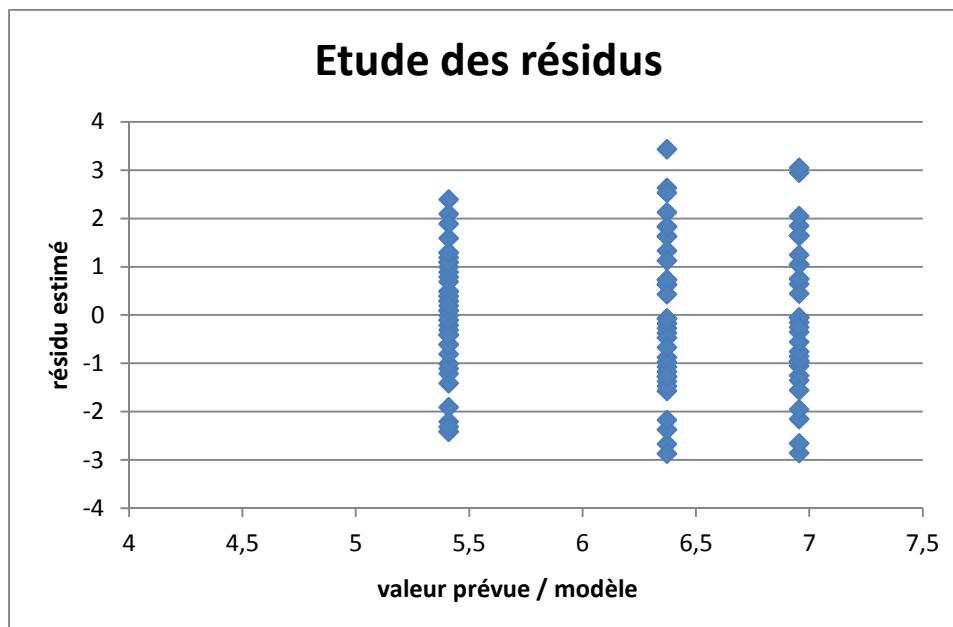
- b) La dernière valeur pour le diamètre vaut 5,7 pour la variété A, 3,1 pour la variété B et 8,2 pour la variété C.

Calculez les résidus correspondants en les présentant sous forme symbolique. Donnez ensuite la valeur de la somme de carrés d'écarts résiduelle après avoir présentée la formule symbolique du calcul.

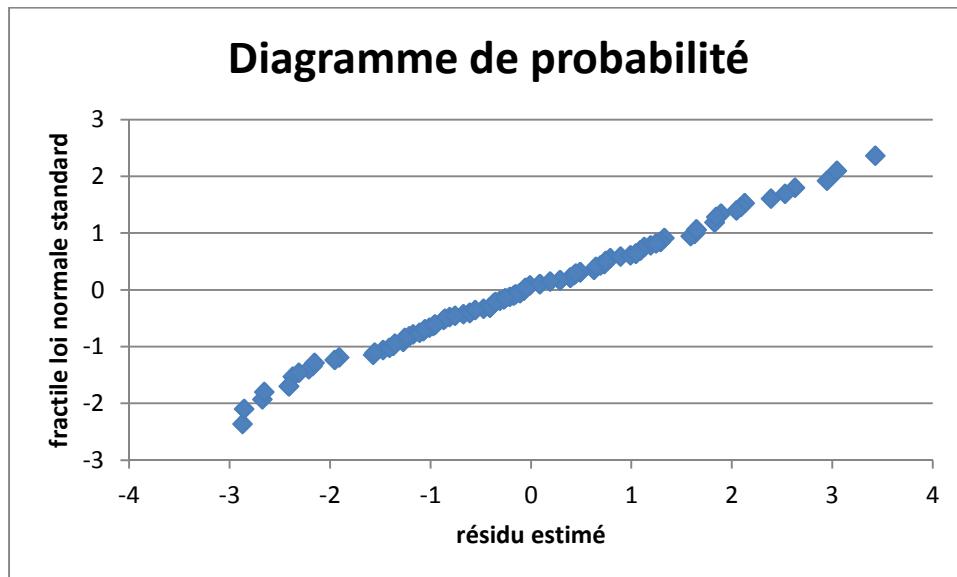
2.2) Fournissez le tableau d'analyse de la variance. Détaillez les calculs réalisés.

Nom de l'étudiant : Prénom : Place :

2.3) Au vu des résultats suivants, pouvez-vous interpréter le test statistique présenté dans ce tableau ? Justifiez votre réponse.



Nom de l'étudiant : Prénom : Place :



Nom de l'étudiant : **Prénom :** **Place :**

2.4) Concluez sur l'étude.

3) Comparez les trois variétés pour leur diamètre moyen.

a) Quel test de comparaisons de moyennes devez-vous choisir et pourquoi ?

b) La valeur critique du test vaut 2,43 . A quoi correspond cette valeur ?

c) Faites la comparaison demandée et concluez.

Rappel : l'écart type de la différence de deux moyennes se calcule selon l'estimation de la variance des populations dont sont issus les deux échantillons. $s_d = \sqrt{\frac{s^2}{n_i} + \frac{s^2}{n_j}}$, s^2 étant

Nom de l'étudiant : Prénom : Place :

Exercice 2

En vue de pouvoir alerter la population en cas d'excès d'ozone dans l'air pouvant être potentiellement dangereux pour certaines personnes, un modèle de la quantité d'ozone en fonction de la température relevée à midi a été élaboré.

Les données utilisées pour estimer le modèle sont les suivantes :

| X température (°C) | Y ozone max (ug / ml) |
|--------------------------|-----------------------------|
| 23,8 | 115,4 |
| 16,3 | 76,8 |
| 27,2 | 113,8 |
| 25,1 | 115,4 |
| 27,5 | 125 |
| 19,4 | 83,6 |
| 19,8 | 75,2 |
| 32,2 | 136,8 |
| 20,7 | 102,8 |
| 12,8 | 57,9 |
| 28,4 | 119,1 |

Nom de l'étudiant : Prénom : Place :

Les résultats obtenus ($\alpha = 5\%$) sont :

| | Degré de liberté | Somme des carrés | Carré moyen | F | Valeur critique de F |
|------------|------------------|------------------|-------------|------------|----------------------|
| Régression | 1 | 5643,46004 | 5643,46004 | 96,8120923 | 4,0941E-06 |
| Résidus | 9 | 524,636325 | 58,292925 | | |
| Total | 10 | 6168,09636 | | | |

Le modèle estimé a pour équation $Y = 7,3867 + 4,1096X$

- 1) Il fait 36,4°C à midi à Bordeaux le 14 juillet. Faites la prévision par intervalle de confiance de la quantité d'ozone dans l'air pour cette journée.

Rappels

$$\sigma^2(\hat{\mu}_p) = \sigma^2 \left(\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)$$
$$\sigma^2(\hat{y}_p) = \sigma^2 \left(1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)$$

Exercice 3

Pour 34 ménages, on veut décrire la facture énergétique mensuelle en fonction du revenu mensuel du ménage, du nombre d'occupants dans le logement et de la superficie de celui-ci.

- 1) Quel type de modèle linéaire allez-vous utiliser et pourquoi ? Ecrivez ce modèle.

Nom de l'étudiant : Prénom : Place :

- 2) Quel est le nombre de paramètres de ce modèle ?

- 3) Quel est le nombre de degrés de liberté résiduels de ce modèle ?

- 4) Pour cette étude, le risque de première espèce (α) est choisi égal à **10%**.

Vous allez travailler à partir du tableau suivant :

| | Statistique | | | |
|-----------|--------------|-------------|-------|-------------|
| | Coefficients | Erreur-type | t | Probabilité |
| Constante | -358,44 | 198,74 | -1,80 | 0,081 |
| X1 | 0,08 | 0,14 | 0,55 | 0,585 |
| X2 | 55,09 | 29,05 | 1,90 | |
| X3 | 0,28 | 0,23 | 1,24 | 0,223 |

- 4.1) Quelle l'équation du modèle estimé ?

Nom de l'étudiant : **Prénom :** **Place :**

4.2) Exprimez littéralement l'hypothèse nulle du test portant sur la variable Nombre d'occupants (X2)

4.3) Réalisez ce test. Fournissez la formule du calcul.

4.4) Interprétez le tableau complété.

4.5) Fournissez l'intervalle de confiance pour le paramètre traduisant l'influence de la variable Nombre d'occupants sur le montant de la facture énergétique.

Nom de l'étudiant : **Prénom :** **Place :**

Exercice 4

Un fabricant d'alimentation pour bétail étudie l'intérêt de quatre aliments pour porcs. Il fournit ces aliments à des porcs de trois races différentes, les races les plus couramment élevées dans sa région d'activité. Par combinaison Aliment x Race, 30 porcs ont été nourris.

Les résultats de l'étude sont les suivants: effet Race non significatif, effet Aliment non significatif, effet Interaction Race x Aliment : significatif, au seuil α choisi.

- 1) Interprétez les résultats de l'étude du point de vue de l'effet des facteurs étudiés. A l'issue de l'étude, quelle recommandation faites-vous pour la vente de ces aliments ?

- 2) Ecrivez le modèle correspondant aux résultats présentés. Puis indiquez le nombre de lignes et le nombre de colonnes pour la matrice X correspondant à ce modèle.

Nom de l'étudiant : **Prénom :** **Place :**