ISARA Lyon Durée 1h

UP SFO- S4 – UE : Statistique

Epreuve N°2 - partie 6 (25%) – Mai 2014

Mme Bottollier Lemallaz CORRIGE EN BLEU

Conditions d’examen : **Documents**: Tables statistiques non annotées autorisées

**Calculatrice** : non programmable

Remarques importantes :

1. A la fin de l’épreuve vous devez rendre les 4 pages du sujet complétées de façon lisible et propre.
2. Sauf indication contraire, vous présenterez vos résultats avec 2 chiffres significatifs.

**Nom :……………………………………………………………….Prénom :……………………………………………….Note sur 20 :……….**

**Evaluation de connaissances de base (8 points) (15 min):**

Il y a 24 informations à donner sur la page 2, 1/3 de point par information

1°) Dans une étude à 3 facteurs avec répétitions,combien y aura-t-il de :

tests d’homoscédasticité ?…**… ….7**.

tests F dans le tableau de l’ANOVA ?….…**7**

tests de la ppds au minimum et au maximum ?..entre 0 et 3..

2°) Dans le modèle : xij = µ + i + j + ij

la formule du calcul du résiduij =  xij+ µ --………**……………………**

3°) Compléter le résultat donné par STATBOX :

|  |  |  |
| --- | --- | --- |
| Coefficient Beta 1 | PROB = 0,25 | H0 : **1 =0** H1 : ** 1**  **0**  On a **25%** de risque de se tromper en rejetant l’hypothèse nulle de **symétrie**. Ce risque est trop important (> 5%) On conclue que la distribution des **résidus** est **symétrique**. |

4°) On veut vérifier les hypothèses d’homoscédasticité des résidus selon les facteurs A, B et AB dans une étude à 2 facteurs étudiés A et B respectivement à 7 et 2 modalités avec 3 répétitions. Compléter le tableau :

|  |  |  |
| --- | --- | --- |
| Facteur A ou B ? |  | Valeur du critère théorique  pour  = 0.01 |
| Facteur  …**A**… | Pour le calcul du X² :  C (°)= …………**1.076**  **au 1/1000ème** | X² 0.99 (6) = 16.8  42 unités au total, 7 modalités donc nj= 6 par modalité. |
| Facteur  …**B**… |  | F 0.995 (20;20)= 3.32  42 unités au total, 2 modalités donc ni = 21 par modalité. |
| Pour l’interaction A\*B : critère calculé = 30  Hypothèses : 42 unités au total, 14 traitements (14 combinaisons) donc 14 variances résiduelles.  H0 : ²11. () =……………= ²72. ()  H1 : Au moins 1 des variances résiduelles est supérieure à une autre  X² 0.995 (13) < X² calculé < X² 0.999 (13)  Conclusion : On rejette H0, on a pu prouver que l’hypothèse d’homoscédasticité des variances résiduelles n’est pas respectée avec un risque d’erreur compris entre 0.1% et 0.5%.  OU critère théorique (pour  = 0.01) = 27.7  X² calculé > X² 0.99 (13) On rejette H0, on a pu prouver que l’hypothèse d’homoscédasticité des variances résiduelles n’est pas respectée avec un risque d’erreur < à 1% | | |

(°) = **1 + (1/ (3x6)) (7x1/5 – 1/35) = 1.076**

**Contexte pour les 2 exercices d’application suivants, ces exercices sont indépendants :** Afin de mesurer la performance de chevaux de sport afin de l’améliorer, une étude a été menée sur la mesure d'un constituant sanguin C (U ml-1) en fonction de la durée de l'entrainement (Ei) qu'ils ont eu (supérieure ou égale à 3 ans (i=1) ou inférieure à 3 ans (i=2)), et de 4 types d'alimentation (Aj) de l'animal (j = 1 à 4).

Dans cette étude il y aura **8**….traitements.

Si chaque traitement est répété n (r = 1 à n) fois, il y aura ……**8\*n**…..unités statistiques et le modèle s’écrira :

**…………………………… xijr = µ + i + j + i j + ijr ……………………………………………………………..**

**Exercice 1 (7points) (25 min) :**

Il y a 21 informations à donner sur la page 3, 1/3 de point par information

Dans cet exercice on suppose que pour chaque traitement on dispose de r répétitions (r = 1 à 3).

On donne : xijr² = 18362242 xijr = 20646

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tableau de moyennes par traitement EiAj | A1 | A2 | A3 | A4 | moyenne |
| E1 | 819 | 519 | 878 | 839 | 763,75 |
| E2 | 905 | 989 | 990 | 943 | 956,75 |
| moyenne | 862 | 754 | 934 | 891 |  |

Construisez le tableau de l’analyse de la variance, **résultat du critère calculé au 1/100ème obligatoirement**, hypothèses et conclusions pour un risque d’erreur égal à 5%. On supposera que les conditions portant sur les résidus sont remplies pour faire cette analyse.

Il est préférable de passer par les raccourcis de calculs pour ne pas perdre de temps

(somme \*moyenne = somme² /n ou = moyenne² \*n)



|  |  |
| --- | --- |
| N = 24 C = | 24 (20646/24)² = **17760721.50** |

H0: CME / CMe = 1 H1: CME / CMe > 1

H0: CMA / CMe = 1 H1: CMA/ CMe > 1

H0: CME\*A/ CMe = 1 H1: CME\*A/ CMe > 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *variations* | *SCE* | *ddl* | *CM* | *F calculé* | *F 0,95* |
| Facteur E | 223494,00 | 1 | 223494,00 | 30,31 | 4.49 |
| Facteur A | 106060,50 | 3 | 35353,50 | 4,79 | 3.24 |
| Interaction | 153990,00 | 3 | 51330,00 | 6,96 | 3.24 |
| résiduelle | 117976,00 | 16 | 7373,50 |  |  |
| Totale | 601520,50 | 23 |  |  |  |

Conclusions : Tous les F calculés > F théoriques, tous les effets de facteurs et d’interaction ont été mis en évidence.

Avec moins de 5% de risque d’erreur.

**Exercice 2 (5 points) (20 min) :**

Il y a 15 informations à donner sur la page 4, 1/3 de point par information

Dans cet exercice on suppose que pour chaque traitement on disposait de 2 répétitions. On sait que la performance de l'animal sera maximale quand le constituant sanguin sera à son taux minimum.

Le tableau de l’ANOVA a donné les résultats arrondis suivants :

Il y a 16 unités expérimentales

|  |  |  |  |
| --- | --- | --- | --- |
| *variations* | *ddl (à compléter)* | *CM* | *p-value* |
| E | 1 | 175142 | 0,001 |
| A | 3 | 17775 | 0,115 |
| E\*A | 3 | 31983 | 0,032 |
| résiduelle | 8 | 6536 |  |
| total | 15 |  |  |

On donne les moyennes des répétitions pour chaque combinaison:

|  |  |  |
| --- | --- | --- |
| moyenne | E1 | E2 |
| A1 | 853,0 | 909,5 |
| A2 | 553,0 | 1007,5 |
| A3 | 882,0 | 981,0 |
| A4 | 795,0 | 1022,0 |

1°) Sur l’extrait de résultat ci-dessous, calculer le résidu (1/10ème) correspondant :

Conformément à l’énoncé au bas de la page 2 il s’agit de la combinaison E2A1, 1ère répétition

|  |  |
| --- | --- |
| Constituant sanguin mesuré x211 | Résidu e211 |
| 840 | **840 – 909.5 = - 69.5** |

2°) Calculer la valeur de la ppds(°) au 1/100ème pour un risque égal à 0,05 et réaliser le groupement des moyennes sous forme de tableau(x) pour répondre concrètement à la problématique de cette étude.

(°)ppds = t 1-/2() \* d(m)

L’interaction étant significative on va faire uniquement le regroupement des 8 moyennes des 8 traitements (8 combinaisons).

|  |
| --- |
| L’ANOVA donne CMe = 6536 avec un ddl = 8 t0,975(8) = 2.306  Pour chaque combinaison E\*A on a nij. = 2 |
| d(m) = Racine (6536\*(1/2+ 1/2))= 80.846 |
| **ppds= 2.306 \* 80.846 = 186.43** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Traitements | moyenne | Groupement (3 groupes) | | | Conclusion de l’étude : la performance maximale des chevaux de course s’observe pour un taux minimum, sur des animaux dont la durée d’entrainement est > ou = à 3 ans et qui sont nourris avec l’alimentation de type 2. |
| E1A2 | 553 | A |  |  |
| E1A4 | 795 |  | B |  |
| E1A1 | 853 |  | B | C |
| E1A3 | 882 |  | B | C |
| E2A1 | 909,5 |  | B | C |
| E2A3 | 981 |  | B | C |
| E2A2 | 1007,5 |  |  | C |
| E2A4 | 1022 |  |  | C |