

Corrigé Examen de Statistique Session 1

Durée 2 h – Documents et calculatrices autorisés

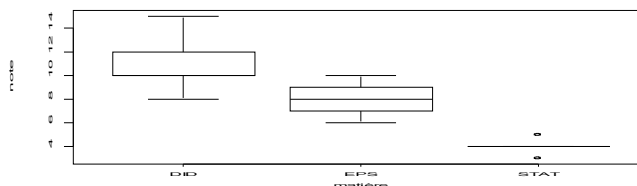
Exercice 1

Une enquête a porté sur le niveau de satisfaction des étudiants pour trois matières enseignées à l'IFEPSA.

Les indices de satisfaction, notés de 0 à 20, obtenus sont :

EPS : 6 10 8 9 7 Didactique : 8 12 10 10 15
Statistique : 3 4 4 5 4

1. Boxplot



2. Les étudiants préfèrent-ils certaines matières ?

3.

a) ANOVA à un facteur, H_0 « Les moyennes des notes par matière sont égales »

b)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
matière	2	123.33	61.667	18.5	0.0002157
Residuals	12	40.00	3.333		

c) Si H_0 est vraie, F suit la loi de Fisher à 2 et 12 ddl.

On a $F=18,5 > 3,89$, donc on rejette H_0 avec un risque de première espèce de 5%. Les moyennes sont significativement différentes.

d) Ce test n'est possible que si les hypothèses suivantes sont vérifiées :

1. Indépendance des observations.
2. Normalité des résidus.
3. Egalité des variances.

Exercice 2

Une enquête sur l'obtention des diplômes en fonction du sexe a donné les résultats suivant :

	Sans diplôme	Licence	Master
Fille	20	52	35
garçon	30	42	21

1. On construit les profils en lignes (distribution des diplômes par sexe).
2. A-t-on les mêmes diplômes que l'on soit une fille ou un garçon ?
3.
 - a) Test du χ^2 d'indépendance, H_0 « Sexe et diplôme sont indépendants »

b)

Tableau sous H_0

	sans licence	licence	master
F	26.75	50.29	29.96
G	23.25	43.71	26.04

$$\chi^2 = 5.6113$$

- c) Si H_0 est vraie, χ^2 suit la loi du χ^2 à $(3-1)(2-1)=2$ ddl. On a $\chi^2 = 5.6113 < 5,99$, on accepte H_0 . Les diplômes obtenus par les filles et les garçons ne sont pas significativement différents.
- d) Les effectifs théoriques doivent être supérieurs à 5 (ici le minimum est 23,25).

Exercice 3

On a procédé à des comparaisons multiples pour 5 traitements ordonnés de A à E.

On obtient les résultats suivants :

1. anova
2. TABLEAU

A	*		
B	*	*	
C		*	*
D		*	*
E			*

3. significativement différents de B : E
significativement différents de D : A
4. non significativement différents de C : B, D, E
non significativement différents de F : C, D

Exercice 4

Expliquer les valeurs caractéristiques (6 moustache du bas, 9 barre en bas de la boîte, 13 trait en gras, 17 barre en haut de la boîte, 18 moustache du haut, 30 point)

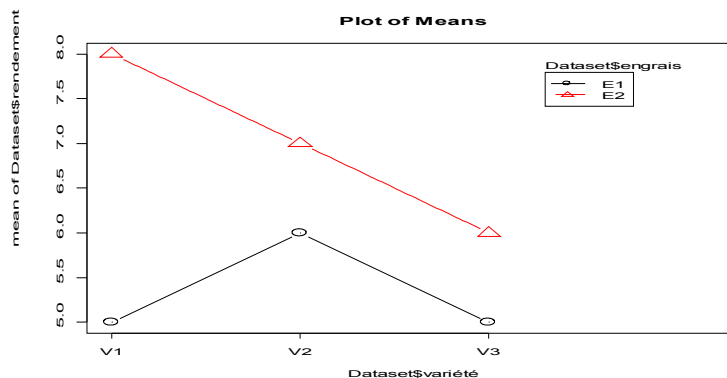
Pour les barres de la boîte, on retrouve les 3 quartiles dont la médiane :

Série représentée : 6 8 10 12 14 16 18 30
Q1=9 Médiane=13 Q3=17

Les moustaches ne dépassent pas $1,5 \text{ Inter-quartile} = 1,5 \times 8 = 12$ en s'arrêtant sur une valeur observée, soit en bas sur 6 et en haut sur 18 car $30 > 17 + 12 = 29$.

Le point représente un point très éloigné, 30.

Exercice 5



On observe un effet moyen faible de la variété (V1 et V2 6.5, V3 5,5), un effet moyen plus grand de l'engrais (E1 5,33 et E2 7), et une interaction.