

Examen

Nom et Prenom :

Année Universitaire : 2015 - 2016

Date : 12/01/2016

Filière : Ingénieur

Durée : 90 min

Semestre : S3

Période : P2

Module : M3.5 - Statistiques et Analyse de Données

Elément de Module : M3.5.2 - Analyse de Données

Professeur : H. Benbrahim

Consignes aux élèves ingénieurs:

- Toute tentative de fraude sera sanctionnée par la note **zéro**.
- **Tout document est non-autorisé.**
- La clarté et la simplicité des réponses est obligatoire.
- **IL FAUT TOUJOURS SPECIFIER LE NUMERO DE LA FIGURE QUE VOUS AVEZ UTILISE POUR DONNER OU JUSTIFIER LE RESULTAT.**
- **IL Y A DES FIGURES QUI MANQUENT (2 OU 3 OU 4 FIGURES). C'EST FAIT EXPRES. SI VOUS NE TROUVEZ PAS UNE FIGURE, MENTIONNEZ QUE VOUS AVEZ BESOIN DE TELLE OU TELLE FIGURE.**

Considérons le tableau des mesures de pH et de conductivité d'une rivière, en 4 points différents, et lors de mesures aux quatre saisons.

Station	pH	Cs	O2	DB	Ox	N4	N3	P4	TE	DE
1p	8,8	274,4	91,3	2,4	1,6	,1	3,7	,1	9,8	39,0
1e	8,5	302,3	78,9	2,4	2,1	,1	3,3	,2	12,7	58,9
1a	8,7	293,0	75,5	1,6	,6	,1	7,0	,0	1,0	23,8
1h	8,2	302,3	83,0	1,6	1,4	,2	2,0	,2	2,9	112,1
2p	8,5	293,0	10,8	7,8	3,8	3,1	3,0	1,8	10,8	150,1
2e	7,8	353,4	16,6	21,6	6,6	10,8	,9	4,4	12,7	76,0
2a	8,2	395,3	31,5	37,1	9,2	13,8	2,4	7,8	2,9	59,9
2h	8,5	334,8	83,0	9,8	3,3	2,8	5,1	1,9	2,9	239,4
3p	8,8	269,7	93,8	3,4	1,7	,4	4,4	,1	10,8	188,1
3e	8,0	358,1	38,2	15,5	2,9	8,7	8,5	5,4	14,7	95,0
3a	8,3	325,5	69,7	7,3	2,2	3,0	14,5	4,4	2,0	75,1
3h	8,5	344,1	83,0	9,0	3,2	3,1	5,3	3,4	2,9	299,3
4p	8,8	227,9	83,0	1,8	1,0	,1	3,0	,2	10,8	287,9
4e	8,4	265,1	68,1	8,8	1,8	,6	4,1	,7	12,7	294,5
4a	8,9	251,1	87,2	2,9	,6	,1	4,0	,5	3,9	172,0
4h	8,4	269,7	83,0	1,3	,9	,0	2,4	,2	2,9	456,0

Partie I: Statistique Descriptive:

I-1 Comment peut-on lire et surtout interpréter toutes les valeurs de la colonne **PH** de la figure 1 ?
Utiliser aussi la boîte à moustache de **PH** pour l'interprétation.

Partie II: Régression Linéaire Simple:

On aimerait construire une droite de régression pour expliquer **PH** en fonction de toutes les autres variables.

2.1 Y a-t-il une liaison linéaire entre PH et les autres variables? Justifiez :

La régression en entrée a convergé et a abouti au modèle donné en Annexe.

2.2 Quelle est la qualité du modèle ? et quelle est son interprétation ?

2.3 Expliquez à un enfant l'apport de ce résultat exprimé en terme de PH.

2.4 La statistique F (D dans la table d'ANOVA) pour la régression est égale à 1,948. Expliquez précisément quelle est l'hypothèse testée et quelle est la conclusion ?

2.5 Donnez l'équation de la droite avant justification de la significativité :

**2.6 En considérant un risque d'erreur de 10%, cette régression est-elle significative?
Donner la nouvelle équation de la régression, justifier.**

2.7 Analyser les résidus et vérifier si les hypothèses de validation du modèle de régression sont vérifiées en justifiant par les différentes figures données en annexe.

2.8 Quelle est la valeur de PH pour l'exemple suivant :

(Cs=8.8) (O2=274.4) (91.3) (DB=2.4) (Ox=1.6) (N4=0.1) (N3=3.7) (P4=0.1) (TE=9.8) (DE=39)

Partie III : Analyse en Composantes Principales

1. Les opérations de centrage et de réduction sont-elles nécessaire ? Justifier
2. L'analyse factorielle est-elle pertinente ? justifier votre réponse en utilisant tous les critères possibles.
3. L'extraction de facteurs a été faite à l'aide de l'Analyse en composantes principales
 - a) Quels sont les pourcentages d'inerties expliquées par l'ACP obtenue par l'extraction de :

- 3 facteurs- 4 facteurs- 5 facteurs
 - b) quels nombres de facteurs retenez vous? justifiez avec toutes les justifications possible.
4. Pour des raisons quelconques, 3 facteurs ont été retenus.
 - a) Est-il nécessaire de faire une rotation?Justifiez :
 - b) donnez une interprétation de ses facteurs (dépendamment de votre choix si rotation aura lieu ou non)
5. Quelles groupes de données peut on générer après cette étude ACP ?
6. Comment peut-on décrire les points suivants:

Cas 7 :

Cas 11 :

Cas 16 :