

Evaluation IFI3
Analyse de Données
Ecole des Mines d'Albi Carmaux

Vous êtes invités à porter vos réponses sur les feuilles d'énoncé, et à les rendre en fin d'examen en vous assurant que votre identité est déclinée sur chaque page.

Cocher lisiblement la (ou les) réponse(s) qui vous semble(nt) être juste(s).

Le barème est défini pour chaque question à la fin de l'énoncé.

I- Culture générale (3 pts)

1. Pourquoi l'analyse de données relève-t-elle du domaine des statistiques ?

Il s'agit de statistiques descriptives cherchant à fournir une image particulière d'une population. 0,5

2. Peut-on affirmer qu'un traitement de données élaboré participe de la production de connaissances ?

a) oui.....

b) non.....

0,5

3. Citer deux exemples concrets d'application de l'analyse de données dans des champs différents

Ex. Sciences sociales → caractéristiques d'une population 1,5

Sciences ingénieur → analyse de résultats expérimentaux

4. Une enquête se conçoit et s'administre en plusieurs étapes. Pouvez vous rappeler ce que sont ces étapes et leur fonction principale ?

- 0,5
- { Conception des variables et du dictionnaire
 - { du questionnaire.
 - { recensement
 - { Analyse

5. Le logiciel utilisé en TP s'appelle 'LeSphinx', cela a-t-il un sens particulier ?

0

Hi! Hi!

II- Connaissances en statistique exploratoire (6 pts)

6. Un tableau de contingence est un tableau à deux dimensions, les individus en ligne et les caractères en colonne :

- a) vrai..... b) faux.....



7. Le calcul d'un centre de gravité en ACP est basé sur une pondération des caractères :

- a) vrai..... b) faux.....



8. Une métrique euclidienne permet de calculer les corrélations entre caractères dans un espace à repère orthonormé :

- a) vrai..... b) faux.....



9. Une variable ordinale ne peut être qu'une variable qualitative :

- a) vrai..... b) faux.....



10. L'Analyse en Composantes Principales s'applique à des caractères :

- a) qualitatifs
b) quantitatifs
c) entiers disjonctifs

11. Les méthodes factorielles font un traitement univarié des données :

- a) vrai b) faux

12. Les composantes principales d'une ACP sont représentées graphiquement dans un cercle de corrélation :

- a) vrai b) faux

13. Les marges d'un tableau de contingence sont calculées pour chaque variable, en fonction des modalités :

- a) vrai b) faux

14. Les nouveaux caractères d'une ACP sont des combinaisons linéaires des anciens caractères :

- a) vrai b) faux

15. L'inertie d'un nuage de points en ACP est une mesure globale et agrégée des distances entre les individus :

- a) vrai b) faux

16. Expliquer la notion de dépendance linéaire entre deux variables qualitatives :

Il s'agit de mesurer un écart entre effectif réel et théorique, modalité par modalité. Les effectifs théoriques étant représentatifs du fait de hasard, c'est-à-dire, étant calculés sur une règle de proportionnalité / marges.

III- Mise en pratique (10 pts)

17. On soumet le tableau de données suivant à une ACP (poids des individus identiques, métrique identité), quel est le nombre de nouveaux caractères ? Pourquoi ?

$$X = \begin{bmatrix} -1 & 7 \\ 1 & 1 \\ 4 & -8 \\ 0 & 4 \end{bmatrix} \quad X_c = \begin{bmatrix} -2 & 6 \\ 0 & 0 \\ 3 & -9 \\ -1 & 3 \end{bmatrix}$$

$\frac{2}{4} \quad \frac{4}{4} \quad \frac{4}{4}$

$(x-3)$

3

1 nouveau caractère

18. Compléter le tableau de contingence ci dessous. Faire un test statistique pour déterminer avec un minimum de risques le degré de dépendance des deux variables.

2

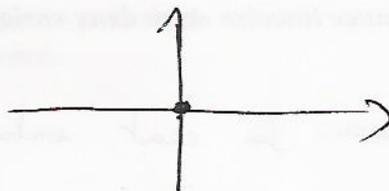
	Variable 2				marge
Variable 1	15	10	5	20	50
	45	30	15	60	150
	15	10	5	20	50
marge	75	50	25	100	250

Tableau croisé des effectifs observés

15	10	5	20
45	30	15	60
15	10	5	20

Les deux variables sont parfaitement dépendantes (usk 0)

19. Quelle serait la forme du plan principal de l'ACP du tableau de la question 18 ?



Tous les points confondus au centre.

1

20. Dans l'exercice d'ACP du TP, Les familles « manuelles » (ma2 à ma5) sont alignées sur le plan principal, cela a-t-il une signification ?

1

oui, il y a des proportions qui sont conservées sur les parts alternatives.