Université Mohamed Khider, Biskra 2020/2021
Faculté des Sc. Exactes et Sc. de la Nature et la Vie
Master 1: **Modèle linéaire**

Série N°2 : Analye Factorielles des Correspondances (AFC)

Je vous conseil d'effectuer vos calculs à l'aide du logiciel Workplace!.

Exercice 1 *Reprenons l'exercice 1 du TD 1.*

On interroge 1587 *étudiants de M2 sur la catégorie socioprofessionnelle de leurs parents. Les étudiants suivent différents cursus: écoles d'ingénieurs, écoles de commerce, universités scientifiques. Les résultats sont les suivants :*

	Ouvriers	*Employés*	*Cadres*	*Professions libérales*
Écoles d'ingénieurs	*50*	*280*	*120*	*20*
Écoles de commerce	*8*	*29*	*210*	*350*
Universités Scientifiques	*150*	*230*	*100*	*40*

On veut étudier l'influence du milieu socioprofessionnel des parents sur le type d'étude des enfants. On rappel que, dans l'exercice 1 du TD 1, nous avons déjà vérifié, par le test du Khi-deux qu'il existe une dépendance entre le milieu socioprofessionnel des parents et le type d'étude des enfants. On peut alors effectuer une analyse factorielle des correspondances sur les données.

1-Donner les centre de gravités, g_r et g_c, associes aux profils-linges et profils-colonnes.

2-Donner les matrices diagonales des profils-linges D_r et des profils-colonnes D_c.

3-Donner les matrices des profils-linges X_r et profils-colonnes X_c.

4-Calculer les deux matrices $A_r := X_r^t X_c^t$ et $A_c := X_c^t X_r^t$.

5-Calculer les valeurs propres et les vecteurs propres de A_r et de A_c. Que peut-on déduire?

6-Que représente g_r (resp. g_c) pour la matrice A_r (resp. A_c)?

* **

Exercice 2 *Suite de l'exercice 1.*

Soit $Y_r := X_r - \mathbf{1}_3 g_r^t$ et $Y_c := X_c - \mathbf{1}_4 g_c^t$ les matrices centrées des profils-linges les profils-colonnes, respectivement. Les matrices de variance-covariances (pondérées) des profils-lignes et des profils-colonnes sont définies, respectivement, par

$$V_r := Y_r^t D_r Y_r \text{ et } V_c := Y_c^t D_c Y_c.$$

L'analyse factorielle des correspondances est basée essentiellement sur les deux matrices $V_r D_c^{-1}$ et $V_c D_r^{-1}$.

1-Déduire de la question 2, de l'exercice 1, les valeurs propres et les vecteurs propres de $V_r D_c^{-1}$ et $V_c D_r^{-1}$.

2-Que représente g_r (resp. g_c) pour la matrice $V_r D_c^{-1}$ (resp. $V_c D_r^{-1}$)?.

3-Donner une base $D_c^{-1}-$orthonormée (resp. $D_c^{-1}-$orthonormée) de $\mathbb{R}^4$ (resp. de $\mathbb{R}^3$) basée sur les vecteurs propres de la matrice $V_r D_c^{-1}$ (resp. $V_c D_r^{-1}$).

4-Donner les axes principaux des profils-linges X_r et des profils-colonnes X_c.

5-Calculer l'inertie totale du profils-lignes X_r par rapport à son centre de gravité g_r, et déduire celle de profils-colonnes par rapport à son centre de gravité g_r.

6-Quelles sont les pourcentages d'inerties par rapport aux axes principaux pour les deux profils?

7-Calculer les inerties du profils-lignes X_r et des profils-colonnes par rapport à leurs axes principaux.

8-Quelles sont les pourcentages d'inerties par rapport aux axes principaux pour les deux profils?

* *

Exercice 3 *Suite de l'exercice 1, et 2.*

1-Déterminer les composantes principales des profils-lignes et déduire celles des profils-colonnes.

2-Déduire les coordonnées des lignes de X_r et X_c.

3-Quelle sont les valeurs de l'espérance et la variance de chaque composante principale pour les deux profils? Quelles sont les valeurs les covariances entre les composantes principales? Que peut-on déduire?

4-Déterminer les contributions absolues et relatives, de chaque ligne, aux inerties des axes principaux.

5. Déduire, de la question 16, les coordonnées des lignes des deux nuages de points X_r et X_c des les nouvelles bases associées.

6.Dans le même plan définit pas les deux premiers axes principaux, représenter les deux nuages de point associes à X_r et X_c.
7-Analyser et discuter les résultats obtenus.