ANALYSE DES DONNEES (Examen)

Master M1 MMD - MA, 23 mai 2017

Calculatrice autorisée, documents autorisés : 2 feuilles recto-verso.

Barême approximatif : 10 points pour chacun des deux exercices.

Exercice 1

On considère le tableau K suivant où a est un entier non nul :

I/J	j_1	j_2	j_3	j_4	j_5
i_1	a	a	a	0	0
i_2	a	a	0	a	0
i_3	0	a	0	a	a

On pose

$$I = \{i_1, i_2, i_3\} \text{ et } J = \{j_1, j_2, j_3, j_4, j_5\}.$$

On effectue l'analyse factorielle des correspondances (AFC) de K.

1. Déterminer les centres de gravité des nuages $\mathcal{N}(I)$ et $\mathcal{N}(J)$.
2. Déterminer la matrice des profils colonnes F_1 ainsi que la matrice des profils lignes F_2 de K.
3. Vérifier que le produit
$$F_1F_2 = \frac{1}{18}\begin{pmatrix} 11 & 5 & 2 \\ 5 & 8 & 5 \\ 2 & 5 & 11 \end{pmatrix}.$$
4. Quel est l'influence du réel a sur l'AFC de ce tableau ?
5. Quel est l'axe factoriel trivial, à quelle valeur propre est-il associé ?
6. Quelle est l'inertie du nuage $\mathcal{N}(J)$?
7. On pose
$$w_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ et } w_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$
Montrer que w_1 et w_2 sont des vecteurs propres de F_1F_2, en déduire les axes factoriels non triviaux u_1 et u_2 ainsi que les valeurs propres associées. On choisira u_1 de manière que la première coordonnée soit positive, de même pour u_2.
8. On note $\varphi_\alpha(i)$ l'abscisse de la projection du profil de la ligne i sur le α ème axe factoriel. Remplir le tableau suivant avec la contrainte $\varphi_\alpha(i_1) \geq 0$

I/J	φ_1	φ_2	φ_3
i_1			
i_2			
i_3			

9. On note ψ_α^j l'abscisse de la projection du profil de la colonne j sur le α ème axe factoriel. En utilisant les formules de transition, compléter le tableau suivant

I/J	j_1	j_2	j_3	j_4	j_5
ψ_1	$\frac{\sqrt{6}}{4}$				
ψ_2	$\frac{-\sqrt{2}}{4}$				

10. Représenter les deux nuages $\mathcal{N}(I)$ et $\mathcal{N}(J)$ simultanément dans le plan factoriel 1-2.
11. Calculer la contribution de i_1 à chacun des axes factoriels non triviaux ainsi que la qualité de représentation de i_1 dans le plan factoriel 1-2 c'est-à-dire $COR_1(i_1) + COR_2(i_1)$.

Exercice 2

On considère un tableau de contingence, noté $K = (k_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}$, qui croise deux variables qualitatives X et Y dont le nombre de modalités est p pour X et q pour Y. On pose

$$I = \{1, \cdots, p\} \text{ et } J = \{1, \cdots, q\}.$$

On effectue une AFC sur ce tableau. Soit α le rang d'un axe factoriel et λ_α la valeur propre associée à cet axe factoriel, on note F_α la composante principale associée à l'axe factoriel α pour le nuage $\mathcal{N}(J)$ et G_α la composante principale associée à l'axe factoriel α pour le nuage $\mathcal{N}(I)$.

1. Rappeler les formules de transitions entre F_α et G_α.
2. On suppose qu'il existe une valeur propre non triviale égale à 1 : $\lambda_\alpha = 1$.
   (a) On suppose qu'il existe des indices i_0, j_0 tels que
   $$\forall i \in I \setminus \{i_0\}, \quad F_\alpha(i) < F_\alpha(i_0) \text{ et } \forall j \in J \setminus \{j_0\}, \quad G_\alpha(j) < G_\alpha(j_0).$$
      i. Montrer que
      $$F_\alpha(i_0) = G_\alpha(j_0).$$
      ii. En déduire que
      $$f_j^{i_0} = \begin{cases} 1 & \text{si } j = j_0, \\ 0 & \text{sinon} \end{cases}$$
      iii. Que cela signifie t-il pour les modalités i_0 et j_0 ?

   (b) On note I_0 l'ensemble des indices pour lesquels $F_\alpha(i)$ atteint son maximum et J_0 l'ensemble des indices pour lesquels $G_\alpha(j)$ atteint son maximum. Pour simplifier on note $I_0 = \{1, \cdots, p_0\}$ et $J_0 = \{1, \cdots, q_0\}$, montrer que le tableau d'effectif K s'écrit

I/J	J_0	$J_1 \setminus J_0$
I_0	K_0	0
$I \setminus I_0$	0	K_1

   où K_0 et K_1 sont des tableaux d'effectifs.
3. On suppose que $p = q$ et que 1 est la seule valeur propre de F_1F_2.
   (a) Montrer que $F_1F_2 = I_p$ où I_p est la matrice identité.
   (b) Que peut-on en déduire pour les matrices F_1 et F_2 ?
   (c) Que conclure sur les variables X et Y ?