

Fiche TD N =° 05 (AFC)

Exercice 01:

Un chef de produit désire cibler la clientèle d'une nouvelle lessive écologique. Il voudrait notamment savoir quelle est la tranche d'âge la plus réceptive à ce produit. un échantillon de 391 personnes.

Tri croisé entre les différentes classes d'âge (six tranches codées 1 pour les 15-19 ans, 2 pour les 20-24 ans, 3 pour les 25-34 ans, 4 pour les 35-44 ans, 5 pour les 45-59 ans et 6 pour les 60 ans et plus) des répondants et une variable "Achat de produits écologiques" comportant 4 modalités (systématiquement, la plupart du temps, occasionnellement, jamais). Si on met ces données un tableau quel son nature ? (nbre ligne et colonne)

Exercice 01:

1. Définir L'AFC. et Expliquer comment on réalise une AFC .
2. C'est quoi la difference entre ACP et AFC.
3. Quel est le nombre maximum de valeurs propres calculées en AFC ?
4. On rappelle que l'analyse factorielle des correspondances (A.F.C.). revient à pratiquer une A.C.P. sur une matrice particulière munie d'une métrique particulière. Quelle sont ces matrice et métrique et qu'elle est la matrice de poids associée ?

Reponse: L'analyse factorielle des correspondances ou A.F.C. revient à pratiquer une A.C.P. sur une le tableau des profils lignes ou des profils colonnes muni de la métrique du Chideux et de la matrice de poids associée.

Formellement, il y a deux possibilités qui sont en dualité exacte, à savoir :

l'analyse des profils lignes : tableau de données : $L = D_r F$; métrique : $M = D_c^{-1}$; poids : D_r

l'analyse des profils colonnes : tableau de données : $C = D_c^{-1} F$; métrique : $M = D_r^{-1}$; poids : D_c

Exercice 02: Soit la matrice P des probabilités suivante, associée a une certaine matrice des données

de type (3,2) :
$$\begin{pmatrix} 10 & 10 \\ 5 & 15 \\ 15 & 5 \end{pmatrix}$$

- 1- Compléter la matrice P en calculant ses probabilités conjointes et marginales.
- 2- Diterminer les matrices des profils lignes et colonnes?
- 3- Calculer la distance χ^2 entre les modalités?

Exercice 03:

Le tableau suivant représente le type d'études poursuivies (université, classes préparatoires, autres) en fonction du parcours suivi au lycée (Lettres, Économie, Maths-Sciences, Technique).

	Univ	Prepa	Autres
Lettres	13.00	2.00	5.00
Economie	20.00	2.00	8.00
Math-Sciences	10.00	5.00	5.00
Tech	7.00	1.00	22.00

1. Quelle est la méthode d'analyse factorielle la plus adaptée à ces données ? justifier votre réponse.
2. Calculer la matrice X des profils-ligne et la matrice Y des profils-colonne. Discuter ces matrices.
3. Calculer la distance χ^2 entre les modalités?

- Rappeler la formule matricielle permettant de calculer la matrice à diagonaliser, notée S, pour réaliser l'ajustement du nuage des profils-lignes.
- L'opération de diagonalisation de la matrice S par un logiciel de mathématiques a permis d'obtenir $\lambda_0 = 1$ $\lambda_1 = 0.1998$ $\lambda_2 = 0.0493$ Que pouvez-vous dire sur la valeur propre λ_0 ?
- Déterminer l'inertie associée à chacun des axes factoriels non-triviaux. En déduire le pourcentage d'inertie associé à chacun des axes, puis l'inertie cumulée.
- Les deux premiers vecteurs propres issus de la diagonalisation de la matrice S sont respectivement $u_1 = (-0.6063 \ -0.1704 \ 0.7767)$ $u_2 = (0.7523 \ -0.6509 \ -0.1014)$ obtenir les coordonnées des profils-lignes ψ_1 , ψ_2

Exercice 04: On donne la répartition de 73128 algériens en fonction de deux variables : La catégorie socioprofessionnelle X et la représentation sociale Y.

	Serviable (y_1)	Généreux (y_2)	Honnête(y_3)	Discret (y_4)	
Paysan (x_1)	956	620	516	0	2092
Employé (x_2)	4012	6224	7284	6084	23604
Profession libérale (x_3)	2728	7776	3868	5332	19704
Commerçant et autres (x_4)	10376	4496	8704	4152	27728
	18072	19116	20372	15568	73128

On veut effectuer une analyse des données du tableau.

- Quelle est la méthode d'analyse factorielle la plus adaptée à ces données ? justifier votre réponse.
- Trouver les tableaux des profils lignes et des profils colonnes ?
- Donner le nombre maximum de valeurs propres ?
- Supposons que la somme des valeurs propres est égale à 0,113. Trouver la statistique de χ^2 .
- Supposons que nous avons trouvé les valeurs propres suivantes : $\lambda_1 = 0,098$; $\lambda_2 = 0,014$ et $\lambda_3 = 0,001$. Combien d'axes doit-on choisir ? Justifier votre réponse.

Exercice 05:

Soit le tableau k_{IJ} suivant, où $I = \{i_1, i_2, i_3\}$ et $J = \{A, B, C, D, E\}$.

I \ J	A	B	C	D	E
i_1	2	2	0	0	0
i_2	0	2	2	0	0
i_3	0	0	0	2	2

Dans ce qui suit, on effectue l'Analyse Factorielle des Correspondances (AFC) de k_{IJ} . On notera ψ_α^i (resp. ψ_α^j) l'abscisse de la projection du profil de la ligne i (resp. colonne j) sur le α ème axe factoriel issu de l'AFC de k_{IJ} et associé à la valeur propre λ_α .

- Calculer les lois marginales $f_{i.}$ et $f_{.j}$.
- Déterminer la matrice F1 (resp. F2) des profils-colonnes (resp. profils-lignes) de k_{IJ} .
- Donner le centre de gravité, noté g, du nuage $\mathcal{N}(I)$ des profils-lignes de k_{IJ} .
- On note $\chi_g(i)$ la distance du khi-deux entre g et le profil de la ligne i . Calculer $\chi^2(i_3)$. Par la suite on admettra que $\chi^2(i_1) = \chi^2(i_2) = \frac{5}{4}$
- Calculer l'inertie totale du nuage $\mathcal{N}(I)$
- Calculer les vecteurs propres de la matrice $F_2'F_1' = (F_1F_2)'$. En déduire le nombre r d'axes factoriels non triviaux
- On a $\psi_1 = 0$ et $var(\psi_1) = 1$, et en supposant le facteur ψ_1^I de la forme $(a, a, b)'$ avec $a > 0$, calculer a et b.
- Déduire du résultat précédent les valeurs de ϕ_1^j pour $j \in J$.

Exercice 06: