

Examen (1h)

Exercice 1 (ACP, 10 pts)

Considérons le tableau de données suivant :

	<i>Poids</i>	<i>Taille</i>	<i>Âge</i>
Farid	45	150	13
Adel	50	165	15
Tarek	60	160	14
Saïd	52	152	14

Tab.1

où les lignes représentent les individus (noms des collégiens) et les colonnes les variables (poids, taille, âge). On s'intéresse à une Analyse en Composantes Principales (ACP) basée sur la matrice de corrélation associée au tableau **Tab.1**, donnée par:

$$R := \begin{pmatrix} 1 & 0.433 & 0.327 \\ 0.433 & 1 & 0.875 \\ 0.327 & 0.875 & 1 \end{pmatrix}.$$

Les composantes principales associées sont:

	C_1	C_2	C_3
	2.147	-0.330	-0.283
	-1.632	-1.137	-0.059
	-1.005	1.247	-0.230
	0.489	0.219	0.573

- 1) Que vaut la moyenne de chaque composante principale C_k , $k = 1, 2, 3$? (0.5pt)
- 2) Calculer les variances de C_1 et C_2 puis "sans calcul" déduire celle de C_3 . (1.5pt)
- 3) Déterminer les valeurs propres, $\lambda_1 \geq \lambda_2 \geq \lambda_3$ associées à la matrice R ? (0.5pt)
- 4) En utilisant la définition de la corrélation, calculer la corrélation de C_1 avec C_2 ; puis "sans calcul" donner la valeur de la corrélation de C_1 avec C_3 . (1pt)

Soient

$$u_1 := (-0.430, 0.896, 0.103)^t; \quad u_2 := (-0.649, -0.227, -0.725)^t; \quad u_3 := (-0.626, -0.379, 0.680)^t,$$

les vecteurs propres normés, de la matrice R , associés aux valeurs propres $\lambda_1, \lambda_2, \lambda_3$.

- 5) Déterminer les axes principaux associés à cette ACP? (1.5pt)
- 6) Sans centrer-réduire la matrice des données, déduire les corrélations entre la variable-Poids (centré-réduit) et les trois composantes principales. (1.5pt)
- 7) Quels sont les pourcentages d'inerties expliquées par les deux premiers axes principaux? (1pt)
- 8) Quel est le pourcentage d'inertie expliquée par le premier plan principal? (0.5). Que conclut-on? (0.5pt)
- 9) Représenter "soigneusement" sur le premier plan principal les quatre collégiens? (1.5pt)

Exercice 2 (AFC, 10 pts)

- 1) Soient $u_k \neq g_r$ les vecteurs directeurs des axes principaux. Calculer le produit matriciel $g_r^t M_r u_k$ et déduire que

$$\sum_{j=1}^q u_k(j) = 0? \quad (2pts)$$

- 2) Donner la formule de l'inertie d'une modalité des profils-lignes X_r par rapport à g_r ? (2pts)
- 3) Soient E_1, \dots, E_p les axes principaux associés aux profils-colonnes X_c :

$$E_1^\perp = ?, \quad (E_2 \oplus \dots \oplus E_p)^\perp = ? \quad (2pts)$$

- 4) Quelle est la relation entre la statistique χ^2 et les valeurs propres associées à la matrice $A_r := X_r^t X_c^t$? (2pts)
- 5) Justifier: $0 \leq \chi^2 \leq n(\text{rg}(A_r) - 1)$, où n est l'effectif total. (2pts)