U.D.L Sidi Bel Abbès **Mercredi 08/12/2021**
Faculté des Sciences Exactes
Département de Probabilité-Statistiques

TD-ACP

Module Analyse des données

Exercice 1. Les notes de trois matières Math, Physique et Chimie dans un teste de niveau chez quatre étudians Reda, Hanane, Aicha et Mourad ont été consignés dans le tableau de données $X = (x_i)$ suivant :

$$\begin{pmatrix} 10 & 6 & 8 \\ 6 & 11 & 10 \\ 12 & 13 & 8 \\ 12 & 10 & 14 \end{pmatrix}$$

$\lambda_1 = 30$, $\lambda_2 = 24$ sont les valeurs propres de la matrice $4V$.
On associe à chaque étudiant un poid $p_i = \frac{1}{4}$, $\forall i = 1, 4$.

1. Énoncer les démarches de calcul dans une ACP.
2. Appliquer à l'exemple si-dessus.
3. Donner une représentation conjointe des résultats dans le plan principal.
4. Représenter le nuage des variables dans le plan des facteurs principals.

Exercice 2. Soit $X = \sqrt{10}\begin{pmatrix} 2 & 2 & 3 \\ 3 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 1 & 4 \\ 2 & 1 & 3 \end{pmatrix}$

le tableau de données correspondant à des mesures effectuées sur 5 individus de poids statistiques égaux pour les trois variables X^1, X^2 et X^3.

1. Calculer le vecteur moyen des individus et déduire le tableau centré.
2. Calculer la matrice des variances-covariances $\mathcal{V}$ et déduire le vecteur $(\sigma_1, \sigma_2, \sigma_3)'$ des écarts types des variables.
3. Déduire le tableau Z de données centrées réduites et la matrice de corrélation $\mathcal{R}$.

   On veut faire une ACP centrée avec des poids uniformes.
4. Sur quelle matrice faut-il travailler ?.

5. Calculer les valeurs propres λ_j, $j=\overline{1,3}$ avec $\lambda_1 > \lambda_2 > \lambda_3$. Vérifier que $\lambda_1 + \lambda_3 = 2\lambda_2$.
6. Calculer les composantes principales C^1 et C^2 et les coordonnées de projection des variables sur le plan principal.
7. illustrer les résultats dans le plan principal.

<u>Exercice 3.</u> Soit le tableau des données de huit points de $\mathbf{R}^2$ défini par

$$X = \begin{pmatrix} 2 & 5 & 8 & 3 & -2 & -5 & -8 & -3 \\ -4 & 0 & 4 & 4 & 4 & 0 & -4 & -4 \end{pmatrix}.$$

1. Préciser le tableau des données centrées et déduire la matrice des variances-covariances V.
2. Vérifier que, pour une matrice quelconque $V = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, le vecteur $U_j = \begin{pmatrix} d-\lambda_j \\ -c \end{pmatrix}$, $\lambda_j \in \mathbb{R}$ vérifie l'égalité
$$(V-\lambda_j I_2)U_j = \begin{pmatrix} det(V-\lambda_j I_2) \\ 0 \end{pmatrix}.$$
3. Si λ_j est une valeur propre pour V, déduire que U_j est un vecteur propre de la matrice V.
4. Soient $\lambda_1 = 30$ et $\lambda_2 = 7.5$ les valeurs propres de la matrice V, on utilisant la question précédente, déduire les vecteurs propres associés U_1 et U_2
5. Calculer les deux composantes principales C^1 et C^2 et représenter graphiquement le nuage des huit points dans le plan principal.

   <u>Indication</u> : Pour simplifier les calcules, on peut extraire les deux vecteurs propres $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$.
6. Calculer les coordonnées de projection des variables sur le plan principal.