U.D.L Sidi Bel Abbès
Faculté des Sciences Exactes
Département : Probabilités-Statistique
Master 2 : SA / PA

Module : ADD
Responsable : M. HAMMAD
Mercredi 11/01/2023
Durée :1h30

EXAMEN FINAL

Question de Cours (04 points).

Montrer sur le schéma de dualité et avec les applications nécessaires comment passer de **l'ACP** à **l'ACP Normée** si $M = D_{1|\sigma^2}$.

Exercice (16 points).

L'objectif de cet exercice est de réaliser une analyse en composantes principales normée sur un tableau de données comportant cinq individus et deux variables quantitatives, les individus sont affectées de masse $p_i = \frac{1}{5}$, $i = \overline{1,5}$.

x^1	x^2
0	1
1	2
2	2
3	3
4	2

01,50 – Calculer le centre de gravité du nuage des individus.

01 – Calculer les écartes-types des variables x^1 et x^2.

02,50 – Donner le tableau Z de données centrées réduites. Déduire la matrice de corrélations R.

03 – Calculer les éléments propres de R.

01 – Vérifier que l'inertie totale est égale à la somme des valeurs propres. Que-signifie celà ?.

01 – Calculer les pourcentages d'inertie pour chaque axe de représentation.

02,50 – Trouver les coordonnées de projection des individus sur les deux axes.

02 – Quelles sont les propriétés à constater sur les éléments propres obtenus lorsque une ACP normée est effectuée sur un tableau de données ne comportant que deux variables quantitatives ?.

01,50 – Représenter le nuage des points individus sur le plan factoriel.

U.D.L Sidi Bel Abbès
Faculté des Sciences Exactes
Département : Probabilités-Statistique
Master 2 : SA / PA

Module : ADD
Responsable : M. HAMMAD
Mercredi 11/01/2023
Durée :1h30

EXAMEN FINAL

Question de Cours (04 points).

Montrer sur le schéma de dualité et avec les applications nécessaires comment passer de **l'ACP** à **l'ACP Normée** si $M = D_{1|\sigma^2}$.

Exercice (16 points).

L'objectif de cet exercice est de réaliser une analyse en composantes principales normée sur un tableau de données comportant cinq individus et deux variables quantitatives, les individus sont affectées de masse $p_i = \frac{1}{5}$, $i = \overline{1,5}$.

x^1	x^2
0	1
1	2
2	2
3	3
4	2

01,50 – Calculer le centre de gravité du nuage des individus.
01 – Calculer les écartes-types des variables x^1 et x^2.
02,50 – Donner le tableau Z de données centrées réduites. Déduire la matrice de corrélations R.
03 – Calculer les éléments propres de R.
01 – Vérifier que l'inertie totale est égale à la somme des valeurs propres. Que-signifie celà ?.
01 – Calculer les pourcentages d'inertie pour chaque axe de représentation.
02,50 – Trouver les coordonnées de projection des individus sur les deux axes.
02 – Quelles sont les propriétés à constater sur les éléments propres obtenus lorsque une ACP normée est effectuée sur un tableau de données ne comportant que deux variables quantitatives ?.
01,50 – Représenter le nuage des points individus sur le plan factoriel.

(1)

Exercice 1

$$\bar{x}^1 = \frac{0+1+2+3+4}{5} = \frac{10}{5} = 2 \quad ; \quad \bar{x}^2 = \frac{1+2+2+3+2}{5} = \frac{1}{5} = 2$$

$$g = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\,\mathbb{1} \quad \text{avec} \quad \mathbb{1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$X_c = \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$Z = \cancel{\ } \cdot D_{1/\sigma} \cdot X_c \quad \text{avec} \quad D_{1/\sigma} = \begin{pmatrix} 1/\sigma_1 & 0 \\ 0 & 1/\sigma_2 \end{pmatrix}$$

- $Var(x^1) = \|x^1\|^2_{D_p} = D_p(x^1, x^1) = \frac{1}{5}\sum_{i=1}^{5} x_{1i}^2 = \frac{1}{5}(4+1+1+4)$

$$= \frac{10}{5} = 2$$

$$\boxed{\sigma_1 = \sqrt{2}}$$

- $Var(x^2) = \|x^2\|^2_{D_p} = D_p(x^2, x^2) = \frac{1}{5}\sum_{i=1}^{5} x_{2i}^2 = \frac{1}{5}(1+1)$

$$= \frac{2}{5}$$

$$\boxed{\sigma_2 = \sqrt{2/5}}$$

Donc $D_{1/\sigma} = \begin{pmatrix} 1/\sqrt{2} & 0 \\ 0 & \sqrt{5}/\sqrt{2} \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & 0 \\ 0 & \sqrt{10}/2 \end{pmatrix}$

$$Z = D_{1/\sigma} \cdot X_c = \begin{pmatrix} -2/\sqrt{2} & -1/\sqrt{2} & 0 & 1/\sqrt{2} & 2/\sqrt{2} \\ -\sqrt{5}/\sqrt{2} & 0 & 0 & \sqrt{5}/\sqrt{2} & 0 \end{pmatrix}$$

$$Z = \frac{1}{2}\begin{pmatrix} -2\sqrt{2} & -\sqrt{2} & 0 & \sqrt{2} & 2\sqrt{2} \\ -\sqrt{10} & 0 & 0 & \sqrt{10} & 0 \end{pmatrix}$$

$$R = Z\, D_p\, Z' = \frac{1}{2}\cdot\frac{1}{5}\cdot\frac{1}{2}\begin{pmatrix} -2\sqrt{2} & -\sqrt{2} & 0 & \sqrt{2} & 2\sqrt{2} \\ -\sqrt{10} & 0 & 0 & \sqrt{10} & 0 \end{pmatrix} \cdot \begin{pmatrix} -2\sqrt{2} & -\sqrt{10} \\ -\sqrt{2} & 0 \\ 0 & 0 \\ \sqrt{2} & \sqrt{10} \\ 2\sqrt{2} & 0 \end{pmatrix}$$

2

$$= \frac{1}{20}\begin{pmatrix} 20 & 6\sqrt{5} \\ 6\sqrt{5} & 20 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \frac{3\sqrt{5}}{10} \\ \frac{3\sqrt{5}}{10} & 1 \end{pmatrix}$$

où

$$V = X_c \cdot D_p\, X_c' = \frac{1}{5}\begin{pmatrix} 10 & 3 \\ 3 & 2 \end{pmatrix}$$

$$R = D_{1/\sigma} \cdot V \cdot D_{1/\sigma} = \frac{1}{5}\begin{pmatrix} \sqrt{2}/2 & 0 \\ 0 & \sqrt{10}/2 \end{pmatrix}\begin{pmatrix} 10 & 3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} \sqrt{2}/2 & 0 \\ 0 & \sqrt{10}/2 \end{pmatrix}$$

$$= \frac{1}{5}\begin{pmatrix} \frac{10\cdot\sqrt{2}}{2} & 3\sqrt{2}/2 \\ \frac{3\sqrt{10}}{2} & \sqrt{10} \end{pmatrix}\begin{pmatrix} \sqrt{2}/2 & 0 \\ 0 & \sqrt{10}/2 \end{pmatrix}$$

$$= \frac{1}{5}\begin{pmatrix} 5 & \frac{3\sqrt{5}}{2} \\ \frac{3\sqrt{5}}{2} & 5 \end{pmatrix} = \begin{pmatrix} 1 & \frac{3\sqrt{5}}{10} \\ \frac{3\sqrt{5}}{10} & 1 \end{pmatrix}$$

$\lambda_1 \neq \lambda_2$

$(R - \lambda_1 \cdot I_2)\, U_1 = 0_{\mathbb{R}^2}$ avec

$$\det(R - \lambda I_2) = \begin{vmatrix} 1-\lambda & 3\sqrt{5}/10 \\ 3\sqrt{5}/10 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - \left(\frac{3\sqrt{5}}{10}\right)^2$$

$$= \left(1-\lambda-\frac{3\sqrt{5}}{10}\right)\left(1-\lambda+\frac{3\sqrt{5}}{10}\right)$$

$$\begin{cases} \lambda_1 = 1 + \dfrac{3\sqrt{5}}{10} = \dfrac{10+3\sqrt{5}}{10} \simeq 1{,}6708 \\ \lambda_2 = 1 - \dfrac{3\sqrt{5}}{10} = \dfrac{10-3\sqrt{5}}{10} \simeq 0{,}3291 \end{cases}$$

3

$\lambda_1 + \lambda_2 = Tr(R) = 2.$

Donc Pour $\lambda_1 = \dfrac{10+3\sqrt{5}}{10}$

$$(R - \lambda_1 I_2)\, U_1 = 0_{R^2} \iff \begin{pmatrix} -\frac{3\sqrt{5}}{10} & \frac{3\sqrt{5}}{10} \\ \frac{3\sqrt{5}}{10} & -\frac{3\sqrt{5}}{10} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff x = y \iff U_1 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\frac{\sqrt{2}}{2} \simeq 0{,}70$

Pour $\lambda_2 = \dfrac{10-3\sqrt{5}}{10}$

$$(R - \lambda_2 I_2)\, U_2 = 0_{R^2} \iff \begin{pmatrix} \frac{3\sqrt{5}}{10} & 3\sqrt{5}/10 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff y = -x \iff U_2 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

L'inertie pour l'axe 1 et 2 :

$$\frac{\lambda_1}{Tr(R)} = \frac{1{,}6708}{2} \simeq 0{,}8354 \text{ Soit } (84\,\%)$$

$$\frac{\lambda_2}{Tr(R)} = \frac{0{,}3291}{2} \simeq 0{,}16455 \text{ Soit } 16\,\%$$

<u>Remarque</u> : $Tr(R) = \lambda_1 + \lambda_2 = 2$: l'inertie totale (Somme des v.p)
Cela veut dire que l'inertie totale égale au nombre de variables actives.

4

$$C^1 = Z' U_1 = \frac{1}{2}\begin{pmatrix} -2\sqrt{2} & -\sqrt{10} \\ -\sqrt{2} & 0 \\ 0 & 0 \\ \sqrt{2} & \sqrt{10} \\ 2\sqrt{2} & 0 \end{pmatrix} \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1-\frac{\sqrt{5}}{2} \\ -1/2 \\ 0 \\ \frac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix}$$

$$C^2 = Z' U_2 = \frac{1}{2}\begin{pmatrix} -2\sqrt{2} & -\sqrt{10} \\ -\sqrt{2} & 0 \\ 0 & 0 \\ \sqrt{2} & \sqrt{10} \\ 2\sqrt{2} & 0 \end{pmatrix} \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1+\frac{\sqrt{5}}{2} \\ -1/2 \\ 0 \\ \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}$$

1 $\left(-1-\frac{\sqrt{5}}{2} ; -1+\frac{\sqrt{5}}{2}\right)$; 2 $\left(-\frac{1}{2}, -\frac{1}{2}\right)$

3 $(0, 0)$; 4 $\left(\frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right)$; 5 $(1, 1)$

- Lorsque on réalise une ACP normée sur le tableau de donnée, ne comportant que deux variables quantitatives x^1 et x^2, on observe les propriétés remarquables suivantes pour les valeurs propres et vecteurs propres 1/ les vecteurs propres sont de composante égales $\pm \sqrt{2}/2$.

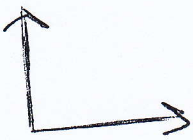
2/ les valeurs propres sont égales à $\lambda_1 = 1 + r(x^1, x^2)$
$\lambda_2 = 1 - r(x^1, x^2)$
où $r(x^1, x^2)$ est le coefficient de corrélation linéaire entre x^1 et x^2

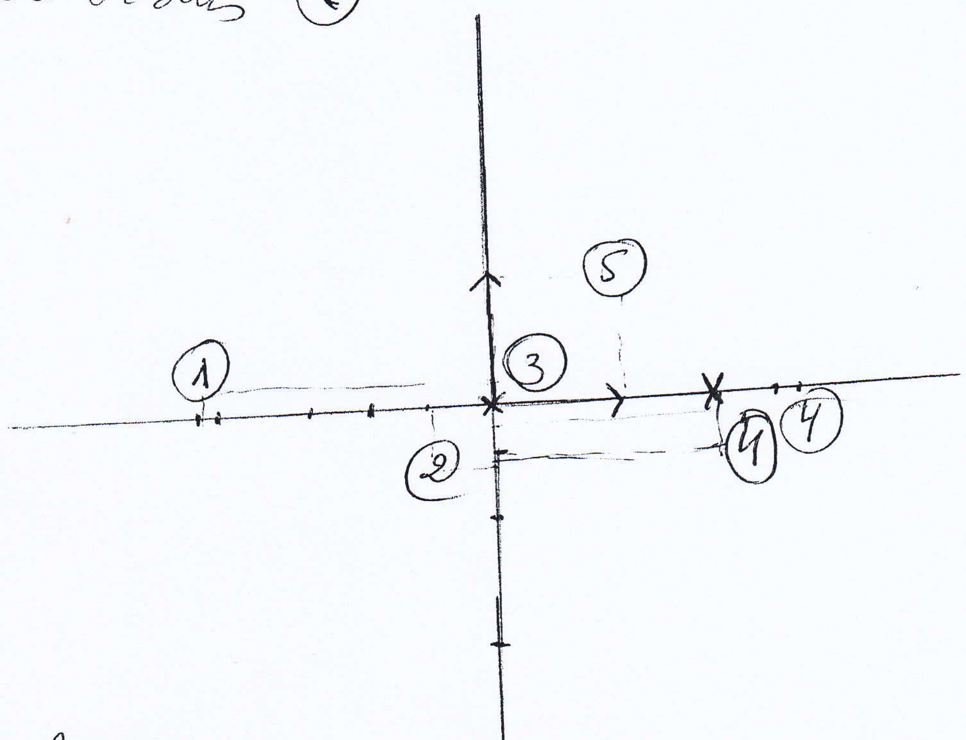
(5)

* Lorsque deux variables étudiées sont de corrélation nulle (non corrélés) $r(x^1, x^2) = 0$, l'analyse prédit deux axes de valeurs égales 1 de même inertie. Si au contraire les variables sont totalement corrélés i.e. $r(x^1, x^2) = 1$ alors un seul axe d'inertie égale à 2 suffit pour résumer parfaitement l'échantillon (les données).

Représentation graphique

des individus (i)





$\sqrt{5} \simeq 2,24$

Q. R :

$$\begin{tikzcd}
E & E \arrow[l, "D_{1/\sigma}"'] \arrow[d, shift right, "M = D_{1/\sigma^2}"'] & F^* \arrow[l, "X_c"'] \arrow[ll, bend right, "Z"'] \arrow[d, shift right, "W"'] \\
E^* \arrow[u, shift left, "R"] \arrow[r, "D_{1/\sigma}"] \arrow[rr, bend right, "X'_c \cdot D_{1/\sigma}"'] & E^* \arrow[r, "X'_c"] \arrow[u, shift right, "V"'] & F \arrow[u, shift right, "D_p"']
\end{tikzcd}$$

I_p

Si : $M = D_{1/\sigma^2}$

$= D_{1/\sigma} \cdot D_{1/\sigma}$

on a : $V = X_c \, D_p \, X'_c$

donc : $R = Z \, D_p \, Z'$

$= D_{1/\sigma} \cdot X \cdot D_p \cdot X' \cdot D_{1/\sigma}$