Table of Contents

EPITA ING2 S8 2018

PARTIEL ANALYSE DES DONNEES
Durée(1h30)

Notes de cours ne sont pas autorisées
Calculatrice non programmable autorisée

EXERCICE1:

Une urne contient 4 boules vertes et 12 boules blanches. On dispose par ailleurs d'une réserve de boules de chaque couleur. On considère des tirages successifs d'une boule de l'urne.

Après chaque tirage, on remet la boule dans l'urne avec elle 3 boules supplémentaires prélevées dans la réserve de la même couleur que la boule tirée.

On note X_n la variable aléatoire qui prends la valeur 1 si on obtient une boule blanche au n ième tirage et la valeur 0 si on obtient une boule verte

Soit $S_n = \sum_{i=1}^{n} X_i \quad \forall n \in N^*$

1) Déterminer la loi de X_1 et calculer $E(X_1)$
2) Calculer les probabilités conditionnelles : $P(X_2 = 0 / X_1 = 0)$
   Et $P(X_2 = 0 / X_1 = 1)$
3) Déterminer la loi de X_2 et $E(X_2)$
4) Calculer la probabilité conditionnelle :
   $P(X_{n+1} = 1 / S_n = k) \quad 0 \le k \le n$

EXERCICE 2

On considère la matrice des données définie par :

$$X = \begin{pmatrix} -6 & 2 & -5 \\ -1 & -2 & 0 \\ 1 & 0 & 2 \\ 1 & 4 & -2 \\ -1 & 2 & -4 \\ 0 & -6 & -3 \end{pmatrix}$$

Exo 2

On considère la matrice des données définie par :

$$X = \begin{pmatrix} -6 & 2 & -5 \\ -1 & -2 & 0 \\ 1 & 0 & 2 \\ 1 & 4 & -2 \\ -1 & 2 & -4 \\ 0 & -6 & -3 \end{pmatrix}$$

La métrique choisie dans l'espace des individus est la matrice identité.

Les six individus ont le même poids égal à $\frac{1}{6}$.

Calculer la moyenne de chaque variable et en déduire le centre de gravité du nuage

$$\bar{X^{(i)}} = \sum_{j=1}^{n} P_j * X_j^{(i)}$$

$$\Rightarrow \bar{X^{(1)}} = \frac{1}{6} * (-6 - 1 + 1 + 1 - 1 + 0) = -1$$

$$\Rightarrow \bar{X^{(1)}} = \frac{1}{6} * (2 - 2 + 0 + 4 + 2 - 6) = 0$$

$$\Rightarrow \bar{X^{(1)}} = \frac{1}{6} * (-5 + 0 + 2 - 2 - 4 - 3) = -2$$

Le centre de gravité est donc $(-1, 0, -2)$

Calculer la matrice des données centrées Y

$$Y = X^{(i)} - \bar{X^{(i)}}$$

$$Y = \begin{pmatrix} -5 & 2 & -3 \\ 0 & -2 & 2 \\ 2 & 0 & 4 \\ -2 & 4 & 0 \\ 0 & 2 & -2 \\ 1 & -6 & -1 \end{pmatrix}$$

Calculer la matrice de variance - covariance V

$$V = {}^tYDY = \frac{1}{6}\,{}^tYY$$

$$^tY = \begin{pmatrix} -5 & 0 & 2 & -2 & 0 & 1 \\ 2 & -2 & 0 & 4 & 2 & -6 \\ -3 & 2 & 4 & 0 & -2 & -1 \end{pmatrix}$$

$$\Rightarrow V = \begin{pmatrix} \frac{17}{3} & -\frac{4}{3} & \frac{11}{3} \\ -\frac{4}{3} & \frac{32}{3} & -\frac{4}{3} \\ \frac{11}{3} & -\frac{4}{3} & \frac{17}{3} \end{pmatrix}$$

Déterminer les valeurs propres de la matrice V et calculer le pourcentage d'inertie

$$\Rightarrow Pv(\lambda) = det(V - \lambda I)$$

$$\Rightarrow Pv(\lambda) = \begin{pmatrix} \frac{17}{3}-\lambda & -\frac{4}{3} & \frac{11}{3} \\ -\frac{4}{3} & \frac{32}{3}-\lambda & -\frac{4}{3} \\ \frac{11}{3} & -\frac{4}{3} & \frac{17}{3}-\lambda \end{pmatrix} = \begin{pmatrix} 8-\lambda & -\frac{4}{3} & \frac{11}{3} \\ 8-\lambda & \frac{32}{3}-\lambda & -\frac{4}{3} \\ 8-\lambda & -\frac{4}{3} & \frac{17}{3}-\lambda \end{pmatrix} (C_1 \leftarrow C_1 + C_2 + C_3)$$

$$\Rightarrow Pv(\lambda) = (8-\lambda)\begin{pmatrix} 1 & -\frac{4}{3} & \frac{11}{3} \\ 1 & \frac{32}{3}-\lambda & -\frac{4}{3} \\ 1 & -\frac{4}{3} & \frac{17}{3}-\lambda \end{pmatrix} = \begin{pmatrix} 1 & -\frac{4}{3} & \frac{11}{3} \\ 0 & \frac{28}{3}-\lambda & -\frac{15}{3} \\ 1 & -\frac{4}{3} & \frac{17}{3}-\lambda \end{pmatrix} (L_2 \leftarrow L_2 - L_1)$$

$$\Rightarrow Pv(\lambda) = (8-\lambda)\begin{pmatrix} 1 & -\frac{4}{3} & \frac{11}{3} \\ 0 & \frac{28}{3}-\lambda & -\frac{15}{3} \\ 1 & -\frac{4}{3} & \frac{17}{3}-\lambda \end{pmatrix} = \begin{pmatrix} 1 & -\frac{4}{3} & \frac{11}{3} \\ 0 & \frac{28}{3}-\lambda & -\frac{15}{3} \\ 0 & -\frac{8}{3} & \frac{6}{3}-\lambda \end{pmatrix} (L_3 \leftarrow L_3 - L_1)$$

$$\Rightarrow Pv(\lambda) = (8-\lambda)\left[(8-\lambda) * \left(\frac{6}{3}-\lambda\right) - \left(-\frac{15}{3}\right) * \left(-\frac{8}{3}\right)\right]$$

$$\ldots$$

$$= -(\lambda-8)(\lambda-12)(\lambda-2)$$

Calcul du pourcentage d'inertie

$$= \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3} = \frac{12+8}{12+8+2} = \frac{10}{11}$$

$$\approx 0.91$$

Déterminer les deux facteurs principaux associés aux deux plus grandes valeurs propres

- $E_1 = Ker(v - 12I)$

$$\forall u \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_1 \Rightarrow \begin{cases} \left(\frac{17}{3} - 12\right)x - \frac{4}{3}y + \frac{11}{3}z = 0 \\ -\frac{4}{3}x + \left(\frac{32}{3} - 12\right)y - \frac{4}{3}z = 0 \\ \frac{11}{3}x - \frac{4}{3}y + \left(\frac{17}{3} - 12\right)z = 0 \end{cases} \Rightarrow \begin{cases} -\frac{19}{3}x - \frac{4}{3}y + \frac{11}{3}z = 0 & (1) \\ -\frac{4}{3}x - \frac{4}{3}y - \frac{4}{3}z = 0 & (2) \\ \frac{11}{3}x - \frac{4}{3}y - \frac{19}{3}z = 0 & (3) \end{cases} \Rightarrow \begin{cases} x = z \\ y = -2x \end{cases}$$

On a donc que : $E_1 = Vect \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

- $E_2 = Ker(v - 8I)$

$$\forall u \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_1 \Rightarrow \begin{cases} \left(\frac{17}{3} - 8\right)x - \frac{4}{3}y + \frac{11}{3}z = 0 \\ -\frac{4}{3}x + \left(\frac{32}{3} - 8\right)y - \frac{4}{3}z = 0 \\ \frac{11}{3}x - \frac{4}{3}y + \left(\frac{17}{3} - 8\right)z = 0 \end{cases} \Rightarrow \begin{cases} -\frac{7}{3}x - \frac{4}{3}y + \frac{11}{3}z = 0 & (1) \\ -\frac{4}{3}x + \frac{8}{3}y - \frac{4}{3}z = 0 & (2) \\ \frac{11}{3}x - \frac{4}{3}y - \frac{7}{3}z = 0 & (3) \end{cases} \Rightarrow \begin{cases} x = z \\ y = x \end{cases}$$

On a donc que : $E_2 = Vect \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

Déterminer les composantes principales et calculer les coefficients de corrélation linéaire

$$C^{(i)} = D * E_i * Y$$

- $C^{(1)}$

$$C^{(i)} = \left(\frac{1}{\sqrt{6}}\right) \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \begin{pmatrix} -5 & 2 & -3 \\ 0 & -2 & 2 \\ 2 & 0 & 4 \\ -2 & 4 & 0 \\ 0 & 2 & -2 \\ 1 & -6 & -1 \end{pmatrix} = \left(\frac{1}{\sqrt{6}}\right) \begin{pmatrix} -12 \\ 6 \\ 6 \\ -10 \\ -6 \\ 6 \end{pmatrix}$$

- $C^{(1)}$

$$C^{(2)} = \left(\frac{1}{\sqrt{6}}\right) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} -5 & 2 & -3 \\ 0 & -2 & 2 \\ 2 & 0 & 4 \\ -2 & 4 & 0 \\ 0 & 2 & -2 \\ 1 & -6 & -1 \end{pmatrix} = \left(\frac{1}{\sqrt{6}}\right) \begin{pmatrix} -6 \\ 0 \\ 6 \\ 2 \\ 0 \\ -6 \end{pmatrix}$$