EPITA_ ING2

PARTIEL ANALYSE DES DONNEES
Durée(1h30)

EXERCICE1 (Version1)

On considère un nombre entier n supérieur ou égal à 2 et une urne contenant n jetons numérotés de 1 à n . On extrait de cette urne successivement et sans remise deux jetons et on désigne alors par X_1 la variable aléatoire indiquant le numéro du premier jeton tiré et X_2 la variable aléatoire indiquant le numéro du deuxième jeton tiré.

1) Déterminer la probabilité $P(X_1 = i)$ pout tout i de $[[1, n]]$
2) Déterminer la probabilité conditionnelle $P_{(X_1=i)}(X_2 = j)$ pour tout couple $(i, j) \in [[1, n]]^2$
3) En déduire la probabilité $P(X_2 = j) \ \ \forall j \in [[1, n]]$ puis comparer Les lois de X_1 et X_2
4) Déterminer la loi du couple (X_1, X_2)
5) Calculer la covariance $Cov(X_1, X_2)$

EXERCICE1 (Version2)

On dispose de $n+1$ urnes numérotées de 0 à n $(n \geq 2)$.

Pour tout entier j compris entre 0 et n , l'urne numérotée j contient $(j+1)$ boules numérotées de 0 à j . On effectue dans ces urnes une suite de tirages d'une boule avec remise selon le protocole suivant :

Pour le premier tirage, on tire une boule au hasard dans l'urne numéro n

Pour les tirages suivants, si l'on a obtenu la boule numérotée j à un tirage, le tirage suivant s'effectue dans l'urne numérotée j .

On note X_k la variable aléatoire égale au numéro de la boule tirée lors du k ème tirage pour tout $k \in IN^*$

1) Déterminer la loi de X_1 , ainsi que son espérance.
2) Déterminer la probabilité conditionnelle $P_{(X_1=j)}(X_2=i)$ pour tout couple (i, j) d'entiers naturels.
3) Déterminer la loi du couple (X_1, X_2)
4) En déduire la loi marginale de X_2 et $E(X_2)$
5) Calculer la covariance $Cov(X_1, X_2)$

EXERCICE1 (Version3)

Un auto-stoppeur attend au péage d'une autoroute pendant une certaine période.
On admet que le nombre de véhicules franchissant le péage pendant cette période est une variable aléatoire qui suit une loi de poisson de paramètre λ .
A chaque fois qu'un véhicule franchit le péage, il lance une pièce truquée qui donne pile avec la probabilité p avec $0 < p < 1$ on pose $q = 1 - p$
On note N, X et Y respectivement le nombre de lancers, le nombre de fois où il a obtenu pile et le nombre de fois où il a obtenu face.

1) Déterminer la probabilité conditionnelle $P_{(N=j)}(X=i)$ pour tout couple (i, j) d'entiers naturels
2) En déduire la loi du couple (X, N)
3) Montrer que la loi marginale de X suit une loi de Poisson de paramètre λp
4) Expliquer sans calcul pourquoi Y suit une loi de Poisson de paramètre λq
5) Calculer la covariance $Cov(N, X)$

EXERCICE2 (Version1)

On considère la matrice des données X :

$$X = \begin{pmatrix} 5 & -2 & -3 \\ 1 & 3 & 2 \\ 3 & 5 & 4 \\ 7 & 1 & 4 \\ 5 & -1 & 2 \\ -3 & 0 & 3 \end{pmatrix}$$

Les variables sont les colonnes de cette matrice, notées $X^{(1)}, X^{(2)}, X^{(3)}$

les six individus sont les lignes de cette matrice.

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

Les six individus ont le même poids $p_i = \dfrac{1}{6} \quad \forall i = 1,\dots,6$

La métrique choisie dans l'espace des variables est la matrice des poids

.

1) Calculer la moyenne de chaque variable et en déduire le centre de gravité du nuage formé par les six individus.
2) Calculer la matrice des données centrées Y
3) Calculer la matrice de variance-covariance V
4) Déterminer les valeurs propres de la matrice V et calculer le pourcentage d'inertie
5) Déterminer les deux facteurs principaux associés aux deux plus grandes valeurs propres
6) Déterminer les deux premières composantes principales.
7) Calculer les coefficients de corrélation linéaire des variables initiales avec les deux Premières composantes principales.

<u>EXERCICE2(Version2)</u>

On considère la matrice des données X :

$$X = \begin{pmatrix} 7 & 0 & -1 \\ 3 & 5 & 4 \\ 5 & 7 & 6 \\ 9 & 3 & 6 \\ 7 & 1 & 4 \\ -1 & 2 & 5 \end{pmatrix}$$

Les variables sont les colonnes de cette matrice, notées $X^{(1)}, X^{(2)}, X^{(3)}$
Les six individus sont les lignes de cette matrice.
La métrique choisie dans l'espace des individus est la matrice identité.

Les six individus ont le même poids $p_i = \dfrac{1}{6} \quad \forall i = 1,\dots,6$

La métrique choisie dans l'espace des variables est la matrice des poids.

1) Calculer la moyenne de chaque variable et en déduire le centre de gravité du nuage formé par les six individus

2) Calculer la matrice des données centrée Y
3) Calculer la matrice de variance-covariance V
4) Déterminer les valeurs propres de la matrice V et calculer le pourcentage d'inertie
5) Déterminer les deux facteurs principaux associés aux deux plus grandes valeurs propres
6) Déterminer les deux premières composantes principales
7) Calculer les coefficients de corrélation linéaire des variables initiales avec les deux premières composantes principales.

EXERCICE2(Version3)

On considère la matrice des données X :

$$X = \begin{pmatrix} 9 & 2 & 1 \\ 5 & 7 & 6 \\ 7 & 9 & 8 \\ 11 & 5 & 8 \\ 9 & 3 & 6 \\ 1 & 4 & 7 \end{pmatrix}$$

Les variables sont les colonnes de cette matrice, notées $X^{(1)}, X^{(2)}, X^{(3)}$
Les six individus sont les lignes de cette matrice.
La métrique choisie dans l'espace des individus est la matrice identité.

Les six individus ont le même poids $p_i = \dfrac{1}{6} \quad \forall i = 1,\ldots,6$

La métrique choisie dans l'espace des variables est la matrice des poids.

1) Calculer la moyenne de chaque variable et en déduire le centre de gravité du nuage formé par les six individus
2) Calculer la matrice des données centrée Y
3) Calculer la matrice de variance-covariance V
4) Déterminer les valeurs propres de la matrice V et calculer le pourcentage d'inertie
5) Déterminer les deux facteurs principaux associés aux deux plus grandes valeurs propres
6) Déterminer les deux premières composantes principales
7) Calculer les coefficients de corrélation linéaire des variables initiales avec les deux premières composantes principales.