Exercices avec SAS/IML

Exercice 1

1. Créer un vecteur suivant.

$$X = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

2. Afficher le type, la transposition, le nombre de lignes et le nombre de colonnes du vecteur X.

3. Créer les matrices suivantes en utilisant X.

$$A = \begin{bmatrix} 2^1 \\ 2^2 \\ 2^3 \end{bmatrix}, \quad B = \begin{bmatrix} 2^2 \\ 2^4 \\ 2^6 \end{bmatrix}, \quad C = \begin{bmatrix} 2^1 & 2^2 & 2^3 \\ 2^2 & 2^4 & 2^6 \\ 2^3 & 2^6 & 2^9 \end{bmatrix}.$$

4. Afficher les matrices suivantes.
$C, \quad AB', \quad AB' - C.$

5. Notons $D = AB' - C$. Afficher :
a) l'élément maximum de D
b) l'élément minimum de la seconde ligne de D
c) les éléments maximum de la colonne 1 et de la colonne 3 de D
d) les indices des maximum de la colonne 1 et de la colonne 3 de D

6. Comparer C et D terme à terme et afficher :
a) le nombre d'éléments de D strictement supérieurs à C et le nombre d'éléments égaux
b) le nombre d'éléments de D qui sont strictement supérieurs à C sur la deuxième ligne
c) le nombre d'éléments de D qui sont strictement supérieurs à C pour les colonnes 1 et 3 confondues

Exercice 2

1. Créer sous forme d'un vecteur V, un 100-échantillon i.i.d. $(X_1, X_2, \ldots, X_{100})$ de loi uniforme sur $[-1, 1]$.
2. Créer un vecteur « Stat » de dimension 2 contenant la moyenne et la variance empirique de l'échantillon.
3. Créer à partir de V un vecteur U contenant les $Y_i = \frac{X_i - 3}{5}, \quad i = 1, \ldots, 100$.

Exercice 3

1. Créer sous forme d'un vecteur X, un 20-échantillon i.i.d. $(X_1, \ldots, X_{20})$ de loi uniforme sur $[-5, 5]$.
2. Créer un vecteur $Y = (Y_1, \ldots, Y_{20})$ défini par le modèle linéaire suivant
$Y_i = 2X_i + 1 + \varepsilon_i, \quad i = 1, \ldots, 20,$
où $\varepsilon_1, \ldots, \varepsilon_{20}$ sont des variables aléatoires i.i.d. de loi gaussienne $\mathcal{N}(0, 1/4)$.

3. En utilisant la méthode des moindres carrés ordinaire, estimer à partir des vecteurs simulés X et Y les coefficients du modèle linéaire.

4. Écrire avec SAS/IML une procédure `estmco` qui prend les arguments en entrée, `y_dep` la variable dépendante et `x_exp` la (ou les) variable(s) explicative(s) et qui génère trois vecteurs en sortie.

1) `est` contenant les coefficients estimés en utilisant la méthode des moindres carrés ordinaire ;

2) `y_pred` pour le vecteur des valeurs prédites de Y ;

3) `res` pour le vecteur des résidus.

Rappelle de la méthode des moindres carrés ordinaire :

Soit $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ le vecteur de variable dépendante. Soient $X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$, le vecteur de variable explicative. Le modèle linéaire

$$Y_i = \mu + \beta X_i + \varepsilon_i, \quad i = 1, \ldots, n$$

peut s'écrire en forme matricielle $Y = Z\theta + \varepsilon$ avec

$$\theta = \begin{bmatrix} \mu \\ \beta \end{bmatrix} \quad \text{et} \quad Z = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}.$$

Pour estimer les coefficients, il suffit d'écrire $\hat{\theta} = (Z'Z)^{-1}Z'Y$.

Exercice 4

Sous IML, écrire des fonctions pour évaluer :

1. Si un nombre donné est un entier naturel ou non. Cette fonction renvoie 1 si c'est le cas et 0 dans le cas contraire. Appliquer pour $m = 7.2$.

2. La somme des n premiers entiers élevés à la puissance k. Appliquer pour $n = 10$ et $k = 3$.

3. Factorielle n : $n!$ Appliquer pour $n = 10$.

4. C_n^k. Appliquer pour $n = 20$ et $k = 15$.

5. A^n, A étant une matrice carrée. Appliquer pour $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ et $n = 10$.

Exercice 5

Sous IML, écrire des fonctions qui créent :

1. Une réalisation de la variable X de loi $P(X = -1) = P(X = 1) = 0.5$.

Rappel : On sait que si U est une variable aléatoire uniforme sur $[0, 1]$, alors $X = (-1)I_{U<0.5} + (1)I_{U\geq 0.5}$ a pour loi celle décrite ci-dessus.

2. Une réalisation de la variable Y de loi exponentielle de paramètre $a > 0$.

Rappel : On sait que si U est une variable aléatoire uniforme sur $[0, 1]$, alors $X = -\frac{1}{a}\ln(1-U)$ suit la loi exponentielle de paramètre a.

3. Puis, écrire un module IML pour simuler n variables $(Z_1, \ldots, Z_n)$ de loi de Laplace de paramètre a. Appliquer pour $n = 100$ et $a = 2$.

Rappel : On sait que si X et Y sont deux variables aléatoires de loi respectivement celle décrite en 1 et exponentielle de paramètre a, alors $Z = X \times Y$ suit une loi de Laplace de paramètre a.

Exercice 6

Écrire avec SAS/IML des procédures ou fonctions qui réalisent les fonctionnements suivants.

1. (divlimo.sas) Diviser chaque ligne d'une matrice par sa moyenne.
2. (profils.sas) Calculer les profils lignes et les profils colonnes d'une matrice.
3. (dnombr.sas) Compter les éléments différents d'une liste (un vecteur de type quelconque).
4. (cleardat.sas) Supprimer les lignes d'une matrice qui sont pleines de zéros et/ou de données manquantes.
5. (taille.sas) Indiquer le nombre de symboles (chiffres et signe) pour tout entier < 100000.
6. (manquant.sas) Donner le nombre total de données manquantes et les numéros des lignes où elles sont localisées. Ce traitement doit être applicable à une matrice de type quelconque.
7. (complet.sas) Compléter les données manquantes numériques (.) des lignes d'une matrice E à l'aide des lignes d'une matrice complète W selon une règle d'affectation fournie par un vecteur CODE.
8. (disjbur.sas) A partir d'une matrice de réponses numériques, évaluer le tableau disjonctif complet et la table de Burt à l'aide de la commande « design ».
9. (tirpour.sas) Tirer au hasard un pourcentage donné de lignes d'une matrice de type quelconque.
10. (tir.sas) Tirer au hasard p entiers distincts compris entre 1 et n
11. (ar1.sas) Générer un processus AR(1) et en faire une représentation graphique à l'aide de la commande « pgraf ».