

EXAMEN : INTRODUCTION AUX LOGICIELS MATHÉMATIQUES

Seuls les notes prises en TD et les documents distribués en TD sont autorisés.

Pour commencer,

Copier-collez le dossier *TPnote* situé dans le répertoire *X:\GM 3A\IntroLogiciels* dans votre répertoire de travail. Le dossier *TPnote* doit contenir les dossiers *DonnéesSAS* et *AuSecours* ainsi que le fichier *sleep.txt*.

Exercice 1 MATLAB :

Soit f une fonction continue sur un intervalle $[a, b]$. Pour un entier n non nul donné, soit $h = \frac{b-a}{n}$. On construit la subdivision uniforme $x_i = a + (i-1)h$ pour $i = 1, \dots, n+1$ et on calcule les ordonnées $y_i = f(x_i)$ correspondantes. L'objectif est d'approcher $I = \int_a^b f(t)dt$ par la formule des trapèzes

$$J_1(n) = \frac{h}{2} \left[y_1 + y_{n+1} + 2 \sum_{k=2}^n y_k \right].$$

1. Ecrire un fichier `f.m` qui contiendra la fonction f . Exemple $f(x) = e^x$.

```
>> f([0.5 1])
ans =
    1.6487    2.7183
```

2. Ecrire une fonction `trap.m`, qui utilisera la fonction f et permettra le calcul approché de l'intégrale par `trap(namefun, a, b, n)`. Test avec $a = 0$, $b = 1$, $n = 10$. *boucle interdite !*

```
>> trap('f', 0, 1, 10)
ans =
    1.7197
```

3. En faisant varier n , étudier $\log |J_1(n) - I|$ en fonction de $\log(n)$. On tracera en particulier la courbe correspondante. $a = 0$, $b = 1$, $f(x) = e^x$.

Pour cette étude, on partira d'un tableau $tabn = [1, 2, 5, 7, 10, 15, 20, 35, 50, 75, 100]$ et on construira le tableau correspondant $taberr$ puis dessinera $\log(taberr)$ en fonction de $\log(tabN)$. Sauvegarde dans `VI1.m` où `VI` sont vos initiales, en indiquant en commentaire l'ordre de la méthode, c'est à dire le coefficient α tel que $erreur \simeq C/n^\alpha$.

4. On souhaite accélérer la convergence en définissant la nouvelle approximation

$$J_2(n) = (64J_1(4n) - 20J_1(2n) + J_1(n))/45.$$

Ecrire un programme `VI2.m` qui calcule ces nouvelles approximations pour différentes valeurs de n et étudier l'ordre de la méthode. A vous de choisir l'échantillonnage des valeurs de n . Autres données identiques à la question précédente.

Pour votre curiosité intellectuelle, cette méthode est inspirée de l'extrapolation de Romberg.

Exercice 2 Logiciel R

- Votre code R sera enregistré dans un seul script sous le nom *ExoR_Nom.R*
- On évitera dans la mesure du possible d'utiliser des boucles dans la programmation.

Le jeu de données sur lequel vous allez travailler porte sur la morphologie et la composition du sommeil de différentes espèces de mammifères. Le fichier *sleep.txt*, présent dans le dossier *TPnote*, compte 62 observations. Chaque observation est constitué des 6 variables suivantes :

- `id` : identifiant de l'espèce,
- `body_weight` : poids d'un individu en kg,
- `brain_weight` : poids du cerveau en grammes,
- `slow_sleep` : nombre d'heures de sommeil sans rêve par jour,
- `paradoxal_sleep` : nombre d'heures de sommeil paradoxal par jour,
- `total_sleep` : nombre total d'heures de sommeil par jour.

1. Importer les données du fichier `sleep.txt` contenu dans le dossier *TPnote* dans un tableau nommé `sommeil`.
2. Donner une instruction permettant de compter le nombre de mammifères pesant plus de 1 kg.
3. Représenter par un nuage de points la variable `paradoxal_sleep` en fonction de la variable `slow_sleep`. Vous tracerez sur le même graphique la droite d'équation $y = 0,2697 + 0,2026x$ (droite de régression). L'équation de la droite devra apparaître sur le graphique.
4. Créer un tableau `sommeilbis` en supprimant la première colonne du tableau `sommeil`.
5. Ecrire une fonction `suppress` qui supprime toutes les lignes d'un tableau `tab` qui comportent au moins une donnée manquante. Vous appliquerez cette fonction au tableau `sommeilbis` et le tableau obtenu sera nommé `sommeilter`.
6. Soit $\mathbf{X} = (x_{ij})$ un tableau de réels de taille $n \times p$. On appelle tableau centré-réduit associé à \mathbf{X} le tableau \mathbf{Z} de taille $n \times p$ défini par :

$$\mathbf{Z} = (z_{ij}) \text{ avec } z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

où \bar{x}_j et s_j désignent respectivement la moyenne et l'écart-type des composantes de la $j^{\text{ème}}$ colonne du tableau \mathbf{X} .

Sans utiliser la fonction `scale`, écrire une fonction `CentrageReduction` qui permet de contruire le tableau centré-réduit associé à un tableau \mathbf{X} (Votre fonction peut faire appel à d'autres fonctions si nécessaire). Vous appliquerez la fonction `CentrageReduction` au tableau `sommeilter`.

Rappel

Soit y_1, \dots, y_n une série statistique et \bar{y} sa moyenne.

$$\text{L'écart-type de la série est défini par } s = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2 - (\bar{y})^2}.$$

Exercice 3 SAS

- Votre programme SAS sera enregistré sous le nom *ExoSAS_nom.sas*
- Créer une bibliothèque permanente *sos* pour avoir accès aux tables contenues dans le dossier *AuSecours*. Cette bibliothèque vous permettra d'avoir de l'aide en cas d'échec à certaines questions (☞ **HELP**).

Dans le cadre d'une surveillance de la population angevine, le CHU d'Angers a effectué, entre 2000 et 2002, une enquête sur le ronflement auprès de 100 patients. Les résultats de cette enquête sont reportés dans les tables SAS *enquete1* et *enquete2* contenues dans le dossier *DonnéesSAS*.

La table SAS *enquete1* comporte 5 variables :

- **ID** : Identifiant de l'individu
- **AGE** : Age de l'individu en années
- **POIDS** : Poids de l'individu en kg
- **TAILLE** : Taille de l'individu en cm
- **SEXE** : Sexe de la personne (homme/femme)

La table SAS *enquete2* comporte 4 variables :

- **ID** : Identifiant de l'individu
- **ALCOOL** : Nombre de verres bus par jour (en équivalent verre de vin rouge)
- **RONFLEMENT** : Diagnostic de ronflement (oui/non)
- **FUMEUR** : Indique si le patient est fumeur (oui/non)

1. Créer une bibliothèque permanente *donnees* pour avoir accès aux tables contenues dans le dossier *DonnéesSAS*.
2. Créer une table SAS *enquete* dans la bibliothèque *work* en fusionnant les tables *enquete1* et *enquete2* de la bibliothèque *donnees*.

☞ **HELP** : Si vous n'avez pas réussi à faire la question 2, soumettre le code ci-dessous et passer à la question 3.

```
DATA work.enquete;  
    SET sos.enquete;  
RUN;
```

On travaille désormais sur la table SAS *enquete* de la bibliothèque *work*.

3. A l'aide d'une procédure, créer un format pour afficher la variable `alcool` sous forme de classes définies de la façon suivante :
 - **nulle** : si la consommation d'alcool est de 0 verre
 - **faible** : si la consommation d'alcool est de 1 ou 2 verres
 - **modérée** : si la consommation d'alcool est comprise entre 3 et 5 verres

- élevée : si la consommation est supérieure ou égale à 6 verres

Vous appliquerez ce format à la variable `alcool` de la table `enquete` de la bibliothèque `work`.

- ☞ **HELP** : Si vous n'avez pas réussi à faire la question 3, soumettre le code ci-dessous et passer à la question 4.

```
DATA work.enquete;  
    SET sos.enquetebis;  
RUN;
```

4. Représenter graphiquement le poids des patients en fonction de leur taille. On fera apparaître un titre au-dessus du graphique.
5. Tracer le diagramme en bâtons de la répartition des patients en fonction de leur consommation d'alcool. Chaque bâton sera subdivisé en deux secteurs selon que les patients ronflent ou non.
6. Créer une table SAS `hommes_buveurs` contenant les données concernant les hommes dont la consommation journalière d'alcool est élevée. Vous supprimerez les variables `sexe` et `alcool` de cette table et vous exporterez son contenu dans un fichier `buveurs.txt` que vous placerez dans le dossier `TPnote`.