

**Examen du 6 Novembre 2012**  
**Durée 3 heures: documents autorisés**

---

*NB: Chaque étudiant enregistre son programme dans un fichier sauvé sous le nom: "nom\_prénom.sas"  
Ne pas oublier d'écrire en tête de programme, en commentaire, le nom et le prénom. A la fin de  
l'examen le fichier sera enregistré sur la clé USB fournie  
Ecrivez, en commentaire, où commence chaque exercice et la réponse à chaque question.*

---

Télécharger le fichier *sleep\_animals.txt* qui se trouve sur ma page personnelle, à l'adresse:

<http://math.univ-lyon1.fr/~gciuperca/enseign.html>

rubrique "Fichiers de données".

**EXERCICE 1**

Le fichier *sleep\_animals.txt* contient des informations sur la durée du sommeil de certains animaux. Les variables dans le fichier sont, par ordre: *specie*, *body\_weight*, *brain\_weight*, *slow\_sleep*, *paradoxal\_sleep*, *total\_sleep*, *span*, *time*, *predation*, *exposure*, *danger*. Les variables *specie*, *predation*, *exposure*, *danger* sont de type caractère et les autres de type numérique. Les valeurs manquantes sont symbolisées par *-999.0*.

- 1) Créer un tableau SAS, appelé *exo1*, à partir du fichier *sleep\_animals.txt*, contenant toutes les observations.
- 2) Remplacer les données manquantes *-999.0* avec le symbole "." (point).
- 3) A partir du tableau créé dans 2), on demande de créer une nouvelle variable comme la différence entre *total\_sleep* et *slow\_sleep + paradoxal\_sleep*. Que constatez-vous?
- 4) A partir du tableau créé dans 3), créer une nouvelle variable qui a quatre modalités: A si la variable *body\_weight* prend des valeurs  $\in ]0, 100[$ , B si la variable *body\_weight* prend des valeurs  $\in [100, 500[$ , C si la variable *body\_weight* prend des valeurs  $\in [500, 1000[$ , D si la variable *body\_weight* prend des valeurs  $\geq 1000$ .
- 5) A partir du tableau créé dans 4), créer deux nouveaux tableaux SAS, appelés *numeriques* et *caracteres*, contenant, respectivement, les variables numériques et caractères. A partir de ces deux nouveaux tableaux créer des fichiers texte externes, avec le même nom.
- 6) A partir du tableau créé dans 4), réaliser l'histogramme et le box-plot pour la variable *brain\_weight*.
- 7) Calculer les corrélations entre les variables *body\_weight*, *brain\_weight*, *slow\_sleep*, *paradoxal\_sleep*.
- 8) Réaliser une analyse descriptive univariée (moyenne, variance, écart-type, min, max, ...) pour chacune des variables *body\_weight*, *brain\_weight*, *slow\_sleep*, *paradoxal\_sleep*.
- 9) Réaliser un test statistique pour tester si la variable *total\_sleep* suit une loi Normale.
- 10) Réaliser une analyse descriptive pour les variables *predation*, *exposure*, *danger*.
- 11) A partir du tableau créé dans 4), représenter en graphique la variable *paradoxal\_sleep* fonction de la variable *total\_sleep*, les points du graphique prenant les valeurs correspondantes de la nouvelle variable créée (à quatre modalités). Représenter le même graphique, sauf que maintenant les points prennent les valeurs de *specie*.

**EXERCICE 2** (à utiliser PROC GPLOT)

On utilise le tableau créé à l'Exercice 1, question 4).

- 1) Tracer sur le même graphique, en abscisse la variable *body\_weight* et en ordonnée les variables *slow\_sleep*, *paradoxal\_sleep*, en respectant les échelles de mesure. Les points du graphique seront joints par interpolation.
- 2) Tracer le graphique de la variable *body\_weight* en abscisse, fonction de la variable *total\_sleep* en ordonnée. Les points de représentation sont des étoiles et la courbe, obtenue par interpolation, sera de couleur rouge. Donner un titre à ce graphique et pour les axes spécifier les noms des variables (le nom du *body\_weight* sera en vert et le nom du *total\_sleep* sera en en bleu).

**EXERCICE 3** (à utiliser PROC IML)

On utilise le tableau *exo1* créé à l'Exercice 1, question 1).

- 1) Lire, dans une matrice, qu'on va noter X, toutes les observations et les variables numériques du tableau SAS *exo1*.
- 2) Afficher l'élément de la matrice X qui se trouve sur la ligne 15 et la colonne 3.
- 3) créer une nouvelle variable Z qui a quatre valeurs: 1 si la variable *body\_weight* prend des valeurs  $\in ]0, 100[$ , 2 si la variable *body\_weight* prend des valeurs  $\in [100, 500[$ , 3 si la variable *body\_weight* prend des valeurs  $\in [500, 1000[$ , 4 si la variable *body\_weight* prend des valeurs  $\geq 1000$ .
- 4) Créer une nouvelle matrice, notée Y, en concatnant la matrice X et le vecteur Z.
- 5) Créer un tableau SAS appelé *tableN* qui va contenir toutes les variables de la matrice X et le vecteur Z.

**EXERCICE 4** (à utiliser les macros SAS)

- 1) Créer une macro-variable *mex* associée au nom du tableau *exo1*.
- 2) Programmer une macro paramétrée qui permet de faire le graphe d'une variable en fonction d'une autre, en identifiant les points par un identificateur (une autre variable). Ces variables se trouvent dans le tableaux *Emex*.
- 3) En utilisant la macro créé précédemment, réaliser le graphique, pour les données de l'Exercice 1, de la variable *paradoxal\_sleep* fonction de la variable *total\_sleep*, les points du graphique prenant les valeurs correspondantes de la variable *specie*.