

**NOM/MATRICULE :** \_\_\_\_\_

## **EXAMEN 2**

### **Consignes :**

L'examen 2 compte pour une valeur de 27.5% de la note finale. Vous avez de 8h30 à 10h20 pour compléter l'examen. Cet examen est divisé en trois parties : 1) partie "papier", à répondre directement sur l'examen, 2) partie SAS à réaliser sur ordinateur dans un programme SAS et 3) une partie R à réaliser dans sur ordinateur dans un programme R.

Si vous terminez l'examen avant la fin, vous pouvez m'en aviser en levant la main afin que vous puissiez copier vos programmes sur une clé USB. Sinon, vous aurez entre 10h20 et 10h30 pour me remettre vos programmes informatiques par courrier électronique à [denis.talbot@mat.ulaval.ca](mailto:denis.talbot@mat.ulaval.ca).

Conseils pour la partie informatique : Si vous bloquez à une étape d'un numéro à plusieurs étapes, continuez avec les autres étapes en faisant comme si tout avait fonctionné. Si vous ne parvenez pas à réaliser certaines étapes, ajoutez en commentaire que vous savez que ça ne fonctionne pas et mentionnez les avenues essayées.

### **PARTIE 1 : Partie papier      40 points**

#### **NUMÉRO 1 : 20 points**

Pour chacune des situations suivantes, identifiez un type de représentation graphique qui serait approprié pour répondre aux objectifs et justifiez brièvement votre choix.

a) Un chercheur a placé des cages pour petits animaux en forêt et en lisière de la forêt durant une nuit. Le lendemain, il compte la proportion des cages qui sont vides et les pourcentages des cages qui contiennent des petits animaux, par type d'animal (oiseaux, rongeurs, autre). Le chercheur s'intéresse à savoir si les taux et les types de captures sont similaires en forêt et en lisière de forêt. Ses données consistent donc à deux tableaux de fréquences relatives indiquant les pourcentages de cages vides, comportant des oiseaux, comportant des rongeurs ou comportant d'autres animaux, un pour la forêt et un pour la lisière de la forêt.

b) Une étudiante en psychologie s'intéresse au lien existant entre les théories de l'intelligence adoptées par les étudiants au secondaire et leurs intentions de décrocher. Les théories de l'intelligence sont mesurées selon un score continu opposant deux types de théories, i.e. l'intelligence est quelque chose de fixe versus l'intelligence peut évoluer dans le temps. Les intentions de décrocher sont mesurées en deux catégories : les élèves sont certains de poursuivre leurs études secondaires ou ont des doutes. Elle dispose des données individuelles pour ces deux variables pour plusieurs étudiants.

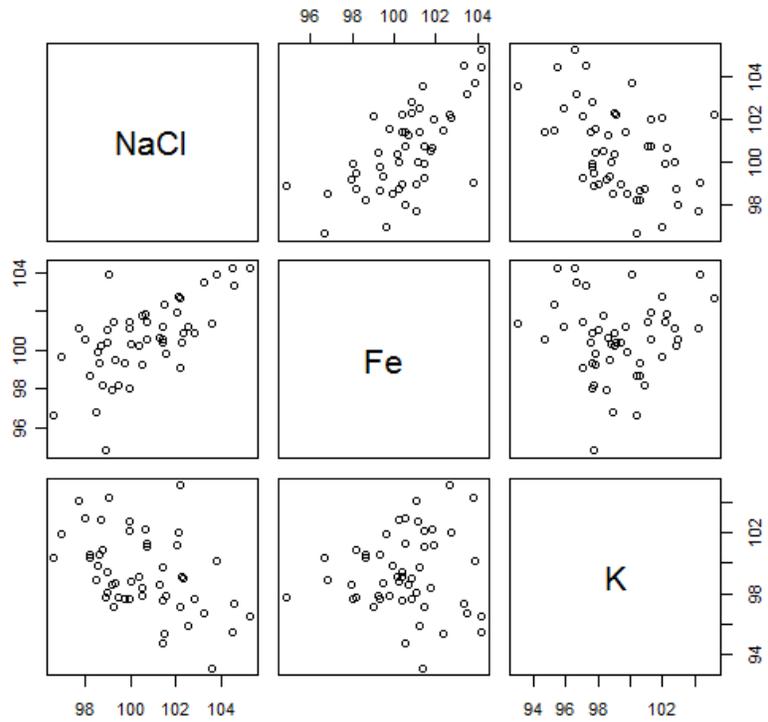
c) Une chercheuse en réadaptation veut savoir si le développement langagier des enfants ayant un retard de langage à l'âge de 3 ans est différent du développement langagier des enfants n'ayant pas de retard de langage à l'âge de 3 ans. Pour ce faire, elle a suivi le développement langagier de deux groupes d'enfants : un avec retard de langage et un sans retard de langage. Les enfants avaient tous initialement entre 3 et 4 ans et ont été suivis jusqu'à leur 10 ans. À chaque année, la chercheuse a évalué le développement de chaque enfant selon un score ayant une valeur numérique continue.

**NUMÉRO 3 : 20 points**

Interprétez brièvement les graphiques suivants :

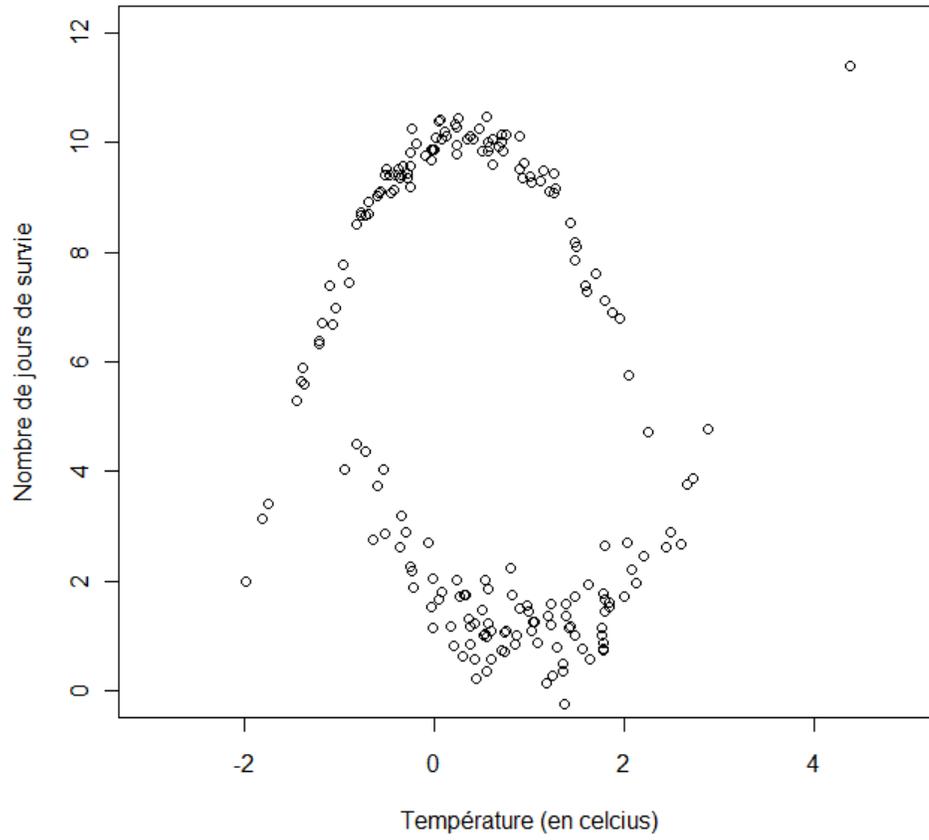
a)

**Matrice de diagrammes de dispersions des concentrations de trois composés chimiques dans des aliments**



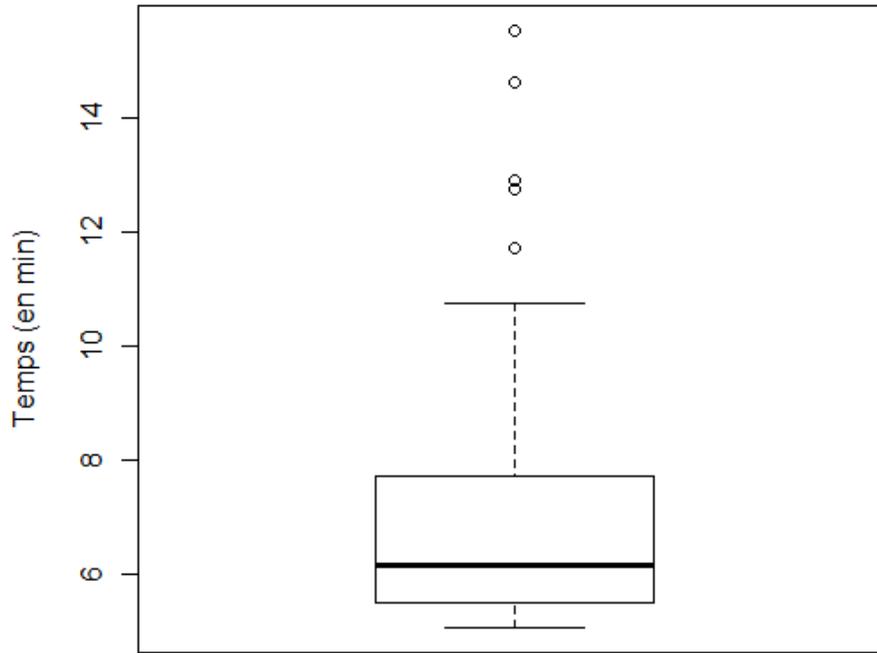
b)

**Diagramme de dispersion du temps de survie (en jours)  
de punaises selon la température**



c)

**Diagramme en boîte du temps effectué au test  
d'endurance de course à pied de Léger  
pour 100 étudiants du secondaire**



## **PARTIE 2 : Partie R      30 points**

### **NUMÉRO 1 (10 points) :**

Simulez 100 observations selon une loi du Khi-Deux à 5 degrés de libertés tronquée de sorte que toutes les observations aient des valeurs plus grandes que 3. Le résultat obtenu doit contenir exactement 100 observations et toutes ces observations doivent respecter le critère d'avoir des valeurs plus grandes que 3.

### **NUMÉRO 2 (20 points) :**

a) Simulez 1000 échantillons de 30 observations de la façon suivantes :

- $X_1, X_2, \dots, X_{15} \sim \text{Cauchy}$
- $X_{16}, X_{17}, \dots, X_{30} \sim \text{Student à 1 degré de liberté}$

b) Calculez la moyenne et la médiane de chacun des échantillons.

c) Vérifiez si la moyenne et la médiane simulées suivent approximativement une loi normale.

d) Commentez brièvement les résultats obtenus.

## **PARTIE 3 : Partie SAS      30 points**

Pour ce numéro, utilisez le jeu de données gnp disponible dans la librairie permanente sashelp de SAS. Vous pouvez directement utiliser ce jeu de données en SAS sous le nom sashelp.gnp (DATA = sashelp.gnp). Ce jeu de données rapporte, entre autre, le produit intérieur brut (*GNP*), la date (*date*) et les investissements effectués par les individus (*invest*). Utilisez la variable *date* comme numéro d'identification des observations.

a) Ajustez un modèle de régression linéaire simple du produit intérieur brut selon les investissements.

b) Vérifiez graphiquement les hypothèses du modèle de régression. (Si les hypothèses ne sont pas respectées, ne modifiez pas le modèle pour obtenir leur respect).

c) Interprétez brièvement l'ensemble des résultats, i.e. respect des hypothèses, corrections à effectuer et interprétation du modèle.

**SAUVEGARDEZ RÉGULIÈREMENT  
VOS PROGRAMMES**