



Examen CCF2 de Statistiques

Document autorisés : tout type (papier ou numérique), voire ordinateur personnel

AVERTISSEMENT

L'ensemble des fichiers de données nécessaires pour cet examen ('M1IGAPASA10data.txt' et 'notesCCF2A09.txt') est normalement disponible à la fois

- en ligne sur <http://utbmjb.chez-alice.fr/UFRSTAPS/index.html> à la rubrique habituelle (voir 'examen', en bas de la page) ;
- en cas de problème internet, sur le réseau de l'université Lyon I : il faut aller sur :
 - 'Poste de travail',
 - puis sur le répertoire 'P:' (appelé aussi '\\\teraetu\Enseignants'),
 - puis 'jerome.bastien',
 - enfin sur 'M1IGAPAS\examen\CCF2'.

Exercice 1.

On étudie le fichier de données 'M1IGAPASA10data.txt'.

- (1) Analyser la variable 'masse'.
- (2) Dans ce cas, quel est (sont) le(s) graphe(s) le mieux adapté(s) à la variable étudiée ?
- (3) Quelle commande permet-elle de le(s) tracer ?
- (4) Quelle commande vous permet-elle d'avoir accès à la moyenne, à l'écart-type et aux quartiles de cette variable ?

Exercice 2.

On étudie de nouveau le fichier de données 'M1IGAPASA10data.txt'. On s'intéresse à la latéralité de la main qui écrit, celle qui tient la fourchette et la latéralité de l'œil de visée, qui correspondent respectivement aux trois variables 'main.ecriture', 'main.fourchette' et 'oeil'.

- (1) (a) Étudier la variable 'main.ecriture'.
(b) Étudier la variable 'main.fourchette'.
(c) Étudier la variable 'oeil'.

- (2) (a) Étudier le croisement de la variable 'main.ecriture' et de la variable 'main.fourchette'.
 (b) Étudier le croisement de la variable 'main.ecriture' et de la variable 'oeil'.
 (c) Conclure

Exercice 3.

On étudie le fichier de données 'notesCCF2A09', qui rassemble divers types de données relatives à un ensemble d'étudiants à l'automne 2009 en STAPS. *Attention*, des données manquantes apparaissent sous la forme de NA ; elles correspondent à l'absence totale de l'étudiant en TD ou en CCF2.

- (1) (a) (i) On étudie le nombre d'absences relevées en TD, qui correspond à la variable 'nombre.absence'.

Attention, pour calculer l'écart-type, la moyenne et les quartiles on utilisera, à cause de la présence des NA :

```
sd(notesCCF2A09$nombre.absence,na.rm=T)
summary(notesCCF2A09$nombre.absence)
```

- (ii) *Question facultative*

On transforme la variable numérique 'nombre.absence' en variable catégorielle en tapant

```
nombre.absence.cat<-as.factor(notesCCF2A09$nombre.absence)
```

Étudiez cette nouvelle variable ! Quel est l'avantage de cette étude sur la précédente ?

- (b) Le nombre d'appels de présence réellement effectués est donné dans la variable 'nombre.appel' ? Est-ce que la question 1a est pertinente ? Pourquoi est-il préférable d'étudier la variable 'pourcentage.absence' ? Étudiez-la !
- (2) (a) Étudiez la note obtenue au CCF2 (variable 'note')
 (b) Commentez !
- (3) On s'intéresse à une éventuelle relation entre le pourcentage d'absence (variable 'pourcentage.absence') et la note obtenue au CCF2 (variable 'note'). Faites l'étude du croisement de ces deux variables et conclure sur l'absentéisme !
- (4) (a) Étudier le croisement de la variable 'groupe' et de la variable 'note'.
 (b) Conclure
- (5) *Question facultative*

On crée une variable 'cat.pourcentage.absence', qualitative, égale à 'TRUE' si l'étudiant a été présent au moins une fois en TD (pourcentage d'absence strictement inférieur à 100 et différent de NA) et égale à 'FALSE' sinon en tapant

```
cat.pourcentage.absence<-as.factor(notesCCF2A09$pourcentage.absence<100)
```

On crée de même une variable 'cat.note', qualitative, égale à 'TRUE' si l'étudiant a été présent au CC2 quelque soit sa note et égale à 'FALSE' sinon en tapant

```
cat.note<-as.factor(!is.na(notesCCF2A09$note))
```

Étudiez le croisement de ces deux nouvelles variables et concluez.

(6) *Question facultative*

Les étudiants, répartis en plusieurs groupes, ont subi des absences inégales d'enseignants, selon leur groupe. Pour remédier à cela, l'enseignant responsable propose de multiplier la note par un coefficient légèrement supérieur à 1 et différent selon les groupes selon la formule suivante :

$$\text{note.rectifiee} = \text{note} \times \text{coefficient},$$

où le coefficient dépend du pourcentage d'absence, noté a de la façon suivante :

$$\text{coefficient} = \begin{cases} K, & \text{si } a \in [0, 40], \\ -\frac{K-1}{60}a + K + \frac{K-1}{60} \times 40, & \text{si } a \in]40, 100], \end{cases}$$

avec $K = 1.1$ pour la groupe 1, $K = 1$ pour la groupe 2, $K = 1.05$ pour la groupe 3.

- Est-ce nécessaire de croiser la variable égale à K et la variable 'groupe' ?
- Tracer sommairement les valeurs de 'coefficient' en fonction de a , pour le groupe 2, par exemple.
- Que pensez-vous de cette façon de procéder ?
- Étudier le croisement de la variable 'groupe' et de la variable 'note.rectifiee'.
- Comparez cette analyse avec celle de la question 4.

Corrigé

Un corrigé sera disponible sur <http://utbmjb.chez-alice.fr/UFRSTAPS/index.html>