



Examen CCF2 de Statistiques

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

IMPORTANT : Attention, il faut traiter trois exercices que vous choisirez comme vous voulez sur les six donnés. Ces six exercices sont indépendants. Des niveaux sont indiqués pour chacun d'eux à titre indicatif.

AVERTISSEMENT

L'ensemble des fichiers de données nécessaires pour cet examen ('M1IGAPASA11data.txt', 'CCF1L2biomecaA11.txt', 'doncompar.txt' et 'catcheur.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 (débutant).

On étudie le fichier de données 'M1IGAPASA11data.txt'.

- (1) Analyser la variable 'age'.
- (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 ?

Pour toute la suite de l'énoncé, on n'indiquera pas sur la copie la façon dont ont été obtenus les calculs et les figures ; seuls les résultats des calculs et des figures doivent être présentés.

Exercice 2 (débutant).

On s'intéresse à des notes de L2, contenu dans le fichier 'CCF1L2biomecaA11.txt'.

- (1) Quelles sont les différentes variables contenus dans ce fichier ?
- (2) (a) Analyser la variable 'prenom'. *Ceux-ci ont été mis dans le désordre pour respecter la confidentialité des notes !*
 - (b) Cette analyse est-elle pertinente ? Pourquoi ?
 - (c) Calculer les effectifs des prénoms en ne conservant que ceux qui apparaissent plus de 3 fois ?
- (3) (a) La variable 'groupe' contient les groupes de TD. Analysez-la.
 - (b) Commentez succinctement cette analyse.
- (4) Analyser la variable 'note'.
- (5) (a) Étudier le croisement de la variable 'groupe' et de la variable 'note'.
 - (b) Commentez cette analyse.
 - (c)
 - (i) L'enseignant responsable des TD 1 et 5 constate que ses deux groupes sont différents des trois autres groupes de TD l'un des deux travaillerait beaucoup, l'autre très peu. Expliquez-cela grâce à un croisement adapté !
 - (ii) Que font les commandes suivantes :

```
CCF1 <- read.table("CCF1L2biomecaA11.txt", h = T)
CCF1 <- CCF1[CCF1[, 2] != "TD1" & CCF1[, 2] != "TD5", ]
```

- (iii) En tapant

```
CCF1 <- CCF1[CCF1[, 2] != "TD1" & CCF1[, 2] != "TD5", ]
CCF1$groupe <- factor(CCF1$groupe)
```

refaites l'analyse de la question 5a en enlevant les deux groupes TD1 et DT5 et commentez-la !

Exercice 3 (débutant).

On étudie le fichier 'doncompar.txt' qui contient les masses, les tailles et les types de 69 individus. Ces individus sont soit des catcheurs¹ soit des étudiants de M1 APA. Le type de l'individu est donné dans la variable 'type' qui prend

- soit la modalité : 'catcheur' pour les catcheurs,
- soit la modalité : 'etudiant' pour les étudiants.

- (1) (a) Pour la variable 'taille', donner les statistiques par groupes des deux populations étudiées (étudiante et catcheuse) et tracer les boîtes à moustache par groupe. On pourra utiliser la fonction 'determin.qualiquanti.R'
 - (b) Commentez.
 - (c) Étudier le croisement de la variable 'taille' et de la variable 'type'.
- (2) (a) Croiser les variables 'masse' et 'taille'. Commentez et expliquer pourquoi cette analyse n'est pas pertinente.

1. Il provient du site web : <http://chezalv.chez.com/catch/catcheur.htm>. Attention, les masses et tailles ne sont que les masses et tailles "officielles" qu'imposent les fédérations de catch. En réalité, ces masses et tailles plus faibles.

- (b) Réaliser un graphique représentant, *pour chaque population (catcheuse ou étudiante)*, la relation entre la masse et la taille (on choisira la masse en variable explicative, c'est-à-dire en abscisse).

Attention

- *Avec Rcmdr* :

Il faut utiliser le menu déroulant "Graphes" du Rcmdr, puis l'option "Nuage de points". Choisir dans le bouton "graphe par groupe" la variable "type" et décocher l'option "courbe de lissage".

- *Sans Rcmdr* :

Il faut taper dans "Rgui" une suite de commandes du type

```
library(car)
```

```
scatterplot(taille ~ masse | type, smooth = FALSE, data = doncompar)
```

Représenter grossièrement ce graphique sur votre copie. Qu'observe-t-on ? Commentez et comparez avec la méthode de la question 2a.

Exercice 4 (confirmé).

- (1) On étudie une loi binomiale de paramètres $n = 15$ et $p = 0.8$.

Calculer les probabilités :

- (a) $P(X \leq 4)$
- (b) $P(X > 1)$
- (c) $P(2 \leq X \leq 10)$
- (d) $P(X \leq -2)$

- (2) On étudie une loi normale centrée réduite (c'est-à-dire, de moyenne nulle et d'écart-type égal à 1).

Calculer les probabilités :

- (a) $P(X \leq -2.5)$
- (b) $P(X > 0)$
- (c) $P(0.2 \leq X \leq 3.2)$

Exercice 5 (confirmé).

On étudie le fichier 'catcheur.txt' qui contient, en autres, les masses et les tailles de 27 catcheurs, déjà étudiés en partie dans l'exercice 3 (qu'il n'est pas nécessaire d'avoir traité).

On s'intéresse dans cet exercice à la variable 'taille', exprimée en cm.

- (1) Donner la moyenne et l'écart-type de la population à partir de l'échantillon.
- (2) Donner les intervalles de confiance de la moyenne de la population aux niveaux de confiance 0.95 et 0.99.
- (3) On suppose que la taille suit une loi normale dont les paramètres ont été estimés dans question 1.
 - (a) Tracer cette loi normale.
 - (b) Quelle est la probabilité pour un catcheur d'avoir une taille supérieure à 198.

Exercice 6 (confirmé).

On reprend les études des exercices 3 et 5 (qu'il n'est pas nécessaire d'avoir traités). On étudie le fichier 'doncompar.txt' qui contient les masses, les tailles et les types de 69 individus. Ces individus sont soit des catcheurs soit des étudiants de M1 APA. Le type de l'individu est donné dans la variable 'type' qui prend

- soit la modalité : 'catcheur' pour les catcheurs,
- soit la modalité : 'etudiant' pour les étudiants.

On s'intéresse dans cet exercice à la variable 'taille', exprimé en cm.

- (1) Donner la moyenne et l'écart-type de la population catcheuse et de la population étudiante.

On pourra utiliser des commandes du type :

```
mean(doncompar$taille[doncompar$type=="catcheur"])  
mean(doncompar$taille[doncompar$type=="etudiant"])
```

- (2) Donner les intervalles de confiance de la moyenne de la population aux niveaux de confiance 0.95 et 0.99, pour chacune des deux populations.
- (3) On suppose que la taille suit une loi normale dont les paramètres ont été estimés dans question 1, pour chacune des populations. Tracez ces deux lois normales et commentez
 - (a) Quelle est la probabilité pour un catcheur d'avoir une taille supérieure à 198.
 - (b) Quelle est la probabilité pour un étudiant d'avoir une taille supérieure à 198.
 - (c) Commentez !

Corrigé

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