




<b>Examen CFF2 de Statistiques</b>
------------------------------------

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

### AVERTISSEMENT

L'ensemble des fichiers de données nécessaires pour cet examen ('M1IGAPASA09data.txt', 'L3APA06.txt' et 'absenteisme.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 : enseignants sur '\Univ-lyon1\enseignement\homes'),
  - puis 'jerome.bastien',
  - enfin sur 'M2IGAPAS\examen\CFF2'.

- Pour les utilisateurs de Rcmdr, on prendra bien garde à utiliser la version 2.7 de  et non la version 2.9; on trouvera cette version, comme d'habitude, en faisant "démarrer", puis "programmes", puis "R" puis "R 2.7". Si cette version n'est pas installée sur votre ordinateur, il faut le redémarrer !

- Les fichiers de l'examen sont au format txt et non xls.

– Pour les utilisateurs de Rcmdr, on rappelle qu'il faut suivre les étapes suivantes :

(1) Dans le menu déroulant "Données" de Rcmdr, choisir l'option "Importer des données" puis "Depuis un fichier texte ou le presse-papier...". Dans la fenêtre de dialogue qui s'ouvre, donner un nom au jeu de données (à la place de Dataset, choisi par défaut), le nom du fichier texte sans extension. Laisser les autres champs avec les valeurs choisies par défaut.

(2) Employer la fenêtre qui s'ouvre alors pour retrouver le fichier à importer.

– Pour les utilisateurs de "Rgui", on exportera les données au format txt grâce aux instructions du type

```
dudu <- read.table("M1IGAPASA09data.txt", h = T)
```

**Exercice 1.**

On étudie le fichier de données 'M1IGAPASA09data.txt'.

- (1) Analyser la variable 'piece'. On rappelle que cette donnée correspond au résultat du jeu de pile ou face pour chaque étudiant du groupe étudié.
- (2) Commentez!

**Exercice 2.**

On étudie le fichier 'L3APA06.txt'.

- (1) Étudier le croisement de la variable 'sexe' et de la variable 'bacculaureat'.
- (2) Commentez!

**Exercice 3.**

Le jeu de données 'absenteisme.txt' contient 7 informations concernant 100 entreprises canadiennes (sélectionnées au hasard) :

- 'Salaire' : le salaire moyen des employés dans cette entreprise ;
- 'PctTempsPart' : le pourcentage de personnes employées à temps partiel ;
- 'PctSynd' : le pourcentage d'employés syndiqués ;
- 'TravailEquipe' : l'existence d'un travail par équipes se relayant (type 3\*8) ;
- 'SyndManRelation' : existence de bonnes relations entre le management et les syndicats ;
- 'Absent' : nombre moyen de jours d'absentéisme par employé dans l'année ;
- 'SqrtAbsent' : racine carrée du nombre précédent.

L'objectif de cette étude est de déterminer l'importance de l'absentéisme dans ces entreprises ainsi que ses déterminants.

- (1) (a) Analysez le nombre de jours d'absentéisme pour ces entreprises (variable 'Absent'). Représentez graphiquement la situation et indiquez quel graphique vous choisissez et pourquoi.
- (b) Analysez également la racine carrée du nombre de jours d'absentéisme (variable 'SqrtAbsent'). Quelle est la différence? Quels sont les avantages et désavantages d'une telle transformation?
- (2) Observe-t-on une différence d'absentéisme (variable 'Absent') selon la qualité des relations entre les syndicats et le management (variable 'SyndManRelation') (On étudiera le croisement entre ces deux variables)?
- (3) Existe-t-il une relation entre le salaire moyen (variable 'Salaire') des employés et l'absentéisme (variable 'Absent') (On étudiera le croisement entre ces deux variables)?
- (4) *Question facultative*
  - (a) Réalisez un graphique où l'on peut étudier la relation entre l'absentéisme et le salaire mais de façon différenciée dans les entreprises où les relations syndicats-management sont bonnes et dans celles où elles sont mauvaises. Recopiez grossièrement ce graphique sur votre copie et interprétez-le.

Indications :

- Avec *Rcmdr* : On peut obtenir ce dessin en allant dans le menu déroulant graphe et en prenant l'option nuage de points. On utilise comme axe X la variable 'Salaire' et Y la

variable 'Absent'. On précise que l'on souhaite un graphe par groupe et on différencie par rapport à la variable 'SyndManRelation'.

- *Sans Rcmdr* : Il faudra taper dans "Rgui"  
`scatterplot(Absent ~ Salaire | SyndManRelation, reg.line = lm,  
 by.groups = TRUE, data = X)`  
 où 'X' est le nom de la variable dans laquelle vous aurez stockez les données.

(b) On donne les résultats suivants :

- relations entre l'absentéisme et le salaire dans les entreprises où les relations syndicats-management sont bonnes :

Noms des indicateurs	Valeurs
pende $a$	-0.000167
ordonnée à l'origine $b$	9.15791
corrélation linéaire $r$	-0.404617
probabilité critique $p_c$	0.000913

- relations entre l'absentéisme et le salaire dans les entreprises où les relations syndicats-management sont mauvaises :

Noms des indicateurs	Valeurs
pende $a$	-0.00026
ordonnée à l'origine $b$	14.193252
corrélation linéaire $r$	-0.499093
probabilité critique $p_c$	0.001944

Commentez-les !

## Corrigé

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