

Statistiques
Master Statistique et économétrie
Exercices Mise à Niveau - Feuille n° 3

V. Monbet

Master 1 - 2012

Exercice 1

An outbreak of Salmonella-related illness was attributed to ice cream produced at a certain factory. Scientists measured the level of Salmonella in 9 randomly sampled batches of ice cream. The levels (in MPN/g) were:

0.593 0.142 0.329 0.691 0.231 0.793 0.519 0.392 0.418

Is there evidence that the mean level of Salmonella in the ice cream is greater than 0.3 MPN/g?

Exercice 2

La table `vaccsp.dat` montre le choix d'hébergement de vacances en fonction de la catégorie socio-professionnelle (CSP) du chef de famille. Elle est disponible sur

http://perso.univ-rennes1.fr/valerie.monbet/Cours_AD/vaccsp.dat

Le choix d'hébergement dépend-il de la CSP?

```
vaccsp = read.table(file="vaccsp.dat",header=FALSE,col.names=c('effectif','csp','heberg'))
```

Exercice 3

1. Charger les données sur la qualité de l'air. Ces données sont disponibles dans le logiciel R:

```
data(airquality)
summary(airquality)
```

On observe que les jours et les mois sont considérées comme des données quantitatives. On les redéfinit afin qu'elles soient qualitatives.

```
airquality$Month = as.factor(airquality$Month)
airquality$Day = as.factor(airquality$Day)
```

2. Visualiser les données (cf commandes suivantes) et commenter les graphiques.

```
pairs(airquality[,1:4])
dev.new()
par(mfrow=c(2,2))
boxplot(Ozone~Month,data=airquality) ; title("Ozone")
boxplot(Solar.R~Month,data=airquality); title("Rayonnement")
boxplot(Wind~Month,data=airquality);title("Vent")
boxplot(Temp~Month,data=airquality);title("Température")
```

3. Comparer la teneur de l'air en Ozone au mois de juin à celle du mois de juillet.
4. Comparer la différence de variabilité de la teneur de l'air en Ozone du mois de juin du mois de juillet est elle significative?
5. Quelles sont les variables qui selon vous pourrait permettre de prédire le taux d'ozone?
6. Ajuster un modèle linéaire pour expliquer le taux d'ozone en fonction des autres variables numériques. Interpréter les résultats

```
oz.lm1 = lm(Ozone~Solar.R+Wind+Temp,data=airquality)
summary(oz.lm1)
```

7. Si on se rappelle des nuages de points visualisés au début, on peut proposer un modèle non linéaire. Quel est le modèle ajusté ci-dessous? Ce modèle permet-il d'améliorer les résultats?

```
airquality$Wind2 = airquality$Wind^2
airquality$Temp2 = airquality$Temp^2
oz.lm2 = lm(Ozone~Solar.R+Wind+Temp+Wind2+Temp2+Wind:Temp,data=airquality)
summary(oz.lm2)
```

8. Est-il utile de considérer un effet mois?