

Statistique bayésienne avec R

Exercice chirurgie

Julien JACQUES

On commence par définir les données

```
data <- list(n = operation, y=deces, N = length(deces))
```

puis on définit les initialisations et les modèles

```
inits1 <- list(list(theta = runif(1)),list(theta = runif(1)),list(theta = runif(1)))
library(rjags)
m1 <- jags.model('Rcode/modelechirurgie.txt', data = data, inits = inits1, n.chains = 3, quiet=TRUE)

inits2 <- list(list(theta = runif(12)),list(theta = runif(12)),list(theta = runif(12)))
library(rjags)
m2 <- jags.model('Rcode/modelechirurgie2.txt', data = data, inits = inits2, n.chains = 3, quiet=TRUE)

inits3 <- list( list(mu=-5,sd=1),list(mu=0,sd=2),list(mu=5,sd=3))
library(rjags)
m3 <- jags.model('Rcode/modelechirurgie3.txt', data = data, inits = inits3, n.chains = 3, quiet=TRUE)
```

Puis on lance les itérations MCMC

```
update(m1, 3000,progress.bar="none")
mcmc1 <- coda.samples(m1, variable.names = c("theta"), n.iter = 2000,progress.bar="none")
update(m2, 3000,progress.bar="none")
mcmc2 <- coda.samples(m2, variable.names = c("theta"), n.iter = 2000,progress.bar="none")
update(m3, 3000,progress.bar="none")
mcmc3 <- coda.samples(m3, variable.names = c("theta"), n.iter = 2000,progress.bar="none")
```

On peut alors comparer les critères DIC (il faudrait au préalable vérifier la convergence des chaînes)

```
dic.samples(m1,n.iter=1000)
```

```
## Mean deviance: 111.8
## penalty 1.007
## Penalized deviance: 112.8
```

```
dic.samples(m2,n.iter=1000)
```

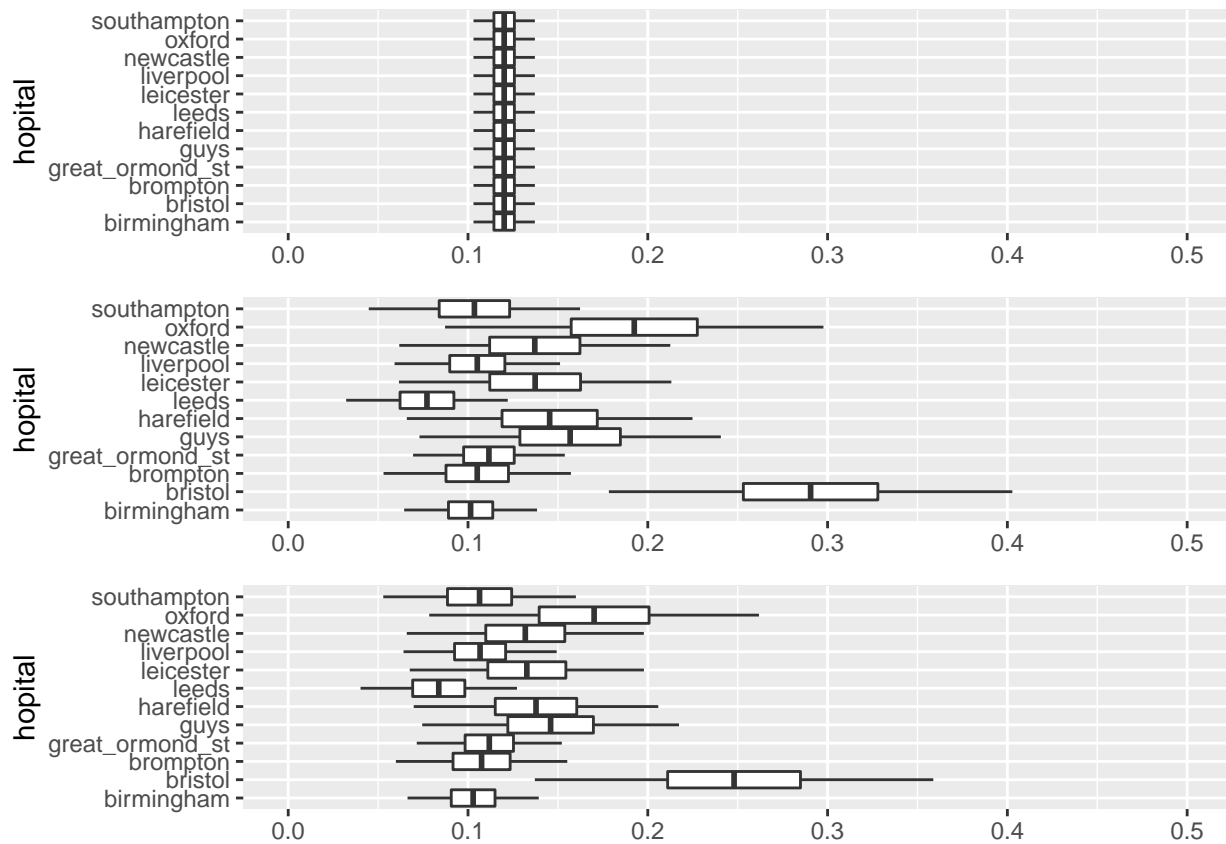
```
## Mean deviance: 73.62
## penalty 11.89
## Penalized deviance: 85.51
```

```
dic.samples(m3,n.iter=1000)
```

```
## Mean deviance: 73.98
## penalty 10.35
## Penalized deviance: 84.33
```

Et afficher les estimations des θ_n par les trois modèles

```
library(ggplot2)
df1=data.frame(mean=summary(mcmc1)$statistics[,1],sd=summary(mcmc1)$statistics[,2],hopital=hopital)
bp1=ggplot(df1,aes(x=hopital),ylim=c(0,0.5))+geom_boxplot(aes(lower=mean-sd,upper=mean+sd,middle=mean,ymin=mean-3*sd,ymax=mean+3*sd))
df2=data.frame(mean=summary(mcmc2)$statistics[,1],sd=summary(mcmc2)$statistics[,2],hopital=hopital)
bp2=ggplot(df2,aes(x=hopital))+geom_boxplot(aes(lower=mean-sd,upper=mean+sd,middle=mean,ymin=mean-3*sd,ymax=mean+3*sd))
df3=data.frame(mean=summary(mcmc3)$statistics[,1],sd=summary(mcmc3)$statistics[,2],hopital=hopital)
bp3=ggplot(df3,aes(x=hopital))+geom_boxplot(aes(lower=mean-sd,upper=mean+sd,middle=mean,ymin=mean-3*sd,ymax=mean+3*sd))
library(gridExtra)
grid.arrange(bp1+ coord_flip(), bp2+ coord_flip(), bp3+ coord_flip(), ncol=1, nrow = 3)
```



Remarques :

- le modèle avec effets aléatoires (paramètres échangeables) conduit à des intervalles de crédibilité légèrement plus étroits que le modèles avec effets fixes (paramètres indépendants), du fait que tous les paramètres ont la même posterior
- les estimations du modèle à effet aléatoire ont tendance à se rapprocher plus de l'effet moyen