

Assignment 3 - Question 3

Question 3

Variance components model, taken from Box & Tiao (1973).

Read the complete WinBUGS example “Dyes: variance components model” (Examples Vol. 1). To show the advantages of the Bayesian approach, Box & Tiao also simulated data from a model with a known between batch variance of 4 and a within batch variance of 16, presented in *dyestuff.txt*.

- (a) Use WinBUGS to plot estimated posterior densities of the two variance components, including the joint posterior (by scatterplot), and give appropriate 95% intervals.

Here is the code, again using *R2OpenBUGS*.

```
#define the model
dyesmodel <- function(){
  for(i in 1 : batches) {
    m[i] ~ dnorm(theta, tau.btw)
    for(j in 1 : samples) {
      y[j , i] ~ dnorm(m[i], tau.with)
    }
  }
  sigma2.with <- 1 / tau.with
  sigma2.btw <- 1 / tau.btw
  tau.with ~ dgamma(0.001, 0.001)
  tau.btw ~ dgamma(0.001, 0.001)
  theta ~ dflat()
}
# write the model code out to a file
write.model(dyesmodel, "dyesmodel.txt")
model.file1 = paste(getwd(),"dyesmodel.txt", sep="/")
coda.file = paste(getwd(),"Q3Coda", sep="/")

#prepare the data for input into OpenBUGS
batches <- 6
samples <- 5
y <- as.matrix(dyedata)
data <- list ("batches","samples","y")

#initialization of variables
inits <- function(){
  list(theta=0, tau.with=1, tau.btw=1)
}

WINE="/opt/local/bin/wine"
WINEPATH="/opt/local/bin/winepath"
OpenBUGS.pgm=paste0("/Users/benjamin/Applications/wine/",
                     "drive_c/ProgramFiles/OpenBUGS/OpenBUGS323/OpenBUGS.exe")

#these are the parameters to save
parameters = c("sigma2.with","sigma2.btw","theta")
```

```

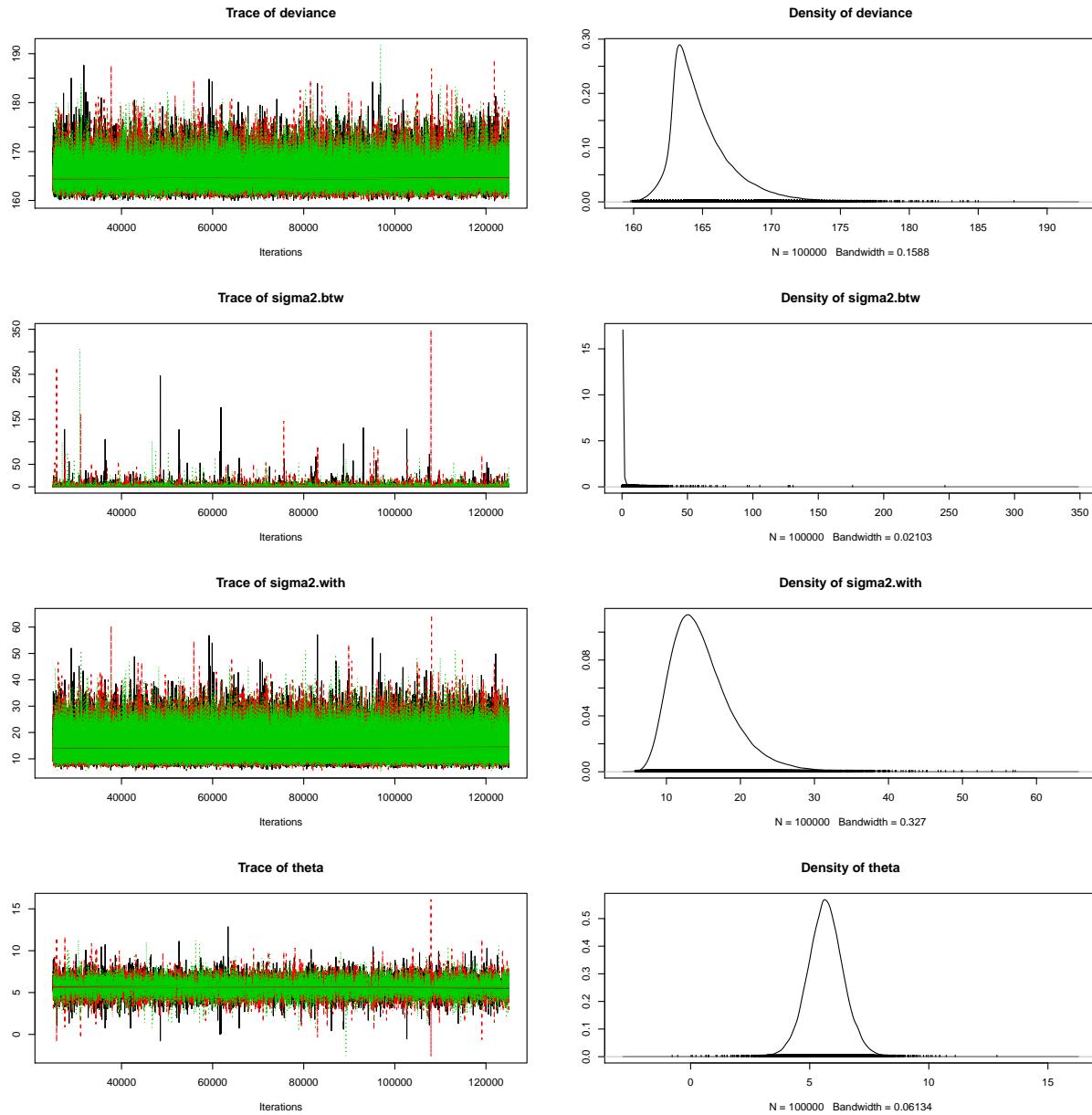
#run the model
dyes.sim <- bugs(data,
                  inits,
                  model.file = model.file1,
                  parameters=parameters,
                  n.burnin = 25000,
                  n.chains = 3,
                  n.iter = 125000,
                  OpenBUGS.pgm=OpenBUGS.pgm,
                  WINE=WINE,
                  WINEPATH=WINEPATH,
                  useWINE=T,
                  codaPkg = T,
                  working.directory = coda.file,
                  debug = F)

samples <- read.openbugs(paste0(coda.file, "/"))
samples2 <- as.matrix(samples[1])

```

Now, plotting the traces, densities and scatter-plot.

```
plot(samples)
```



```

summary(samples)
##
## Iterations = 25001:125000
## Thinning interval = 1
## Number of chains = 3
## Sample size per chain = 1e+05
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean      SD Naive SE Time-series SE
## deviance   164.9326 2.2897 0.004180      0.017834
## sigma2.btw  0.5398 2.2694 0.004143      0.009786
## sigma2.with 14.8296 4.2291 0.007721      0.011738

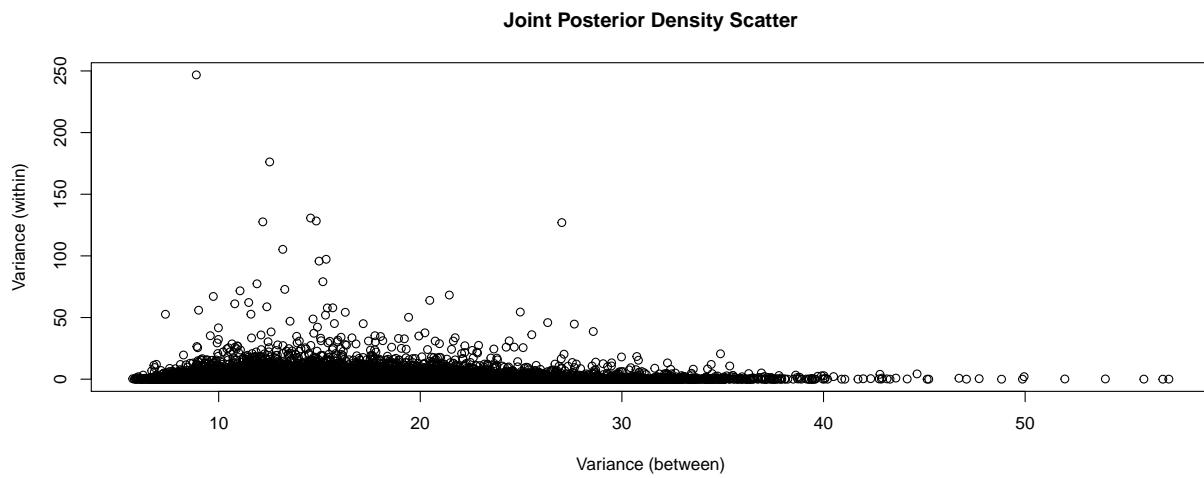
```

```

## theta      5.6640 0.7609 0.001389      0.009316
##
## 2. Quantiles for each variable:
##
##          2.5%     25%     50%     75%   97.5%
## deviance 1.619e+02 1.634e+02 164.40000 165.900 170.900
## sigma2.btw 8.594e-04 8.788e-03 0.05492 0.340 4.143
## sigma2.with 8.714e+00 1.186e+01 14.12000 17.010 25.070
## theta      4.167e+00 5.178e+00 5.66200 6.144 7.173

plot(samples2[, "sigma2.with"], samples2[, "sigma2.btw"], ylab = "Variance (within)", xlab = "Variance (be

```



We can see that the expected values for both variance parameters are close to those we expected. However, the between batch variance is most probably showing a difference; the assumed variance (4) is within the 95% CI, but only just, so the result is problematic if we try to reconcile it with our prior knowledge. The within sample variance estimate is much closer to what we expected though, so it presents less of an issue. By analysing the scatter-plot, we can see that both variance components can take on a range of possible values at a large remove from our expectation, but the quantile estimates confirm that the majority of samples are clustered nearer to where our prior knowledge thought they would be.

(b) Actual implementation of unconstrained priors for the variances is not straightforward but, without attempting this, you are asked to speculate on what would change in your results. Relate your discussion to anticipated results from classical methods. [E.g., SAS® software gives point estimates only, varying between values of 13.3 and 14.9 for the within batch component, and 0 or -1.3 for the between batch component.]

Not Attempted