

TMA4300 INLA Practice

Author 1, Author 2, Author 3

The document is intended as a tool for students to practice using INLA with the R-INLA package. Students are welcome to work together in groups, and to ask questions as they progress in this worksheet. When coding, students should take turns:

1. *coding up various components of the samplers, and*
2. *looking over and double checking other group members' code as it is being written.*

Make sure your code is well commented and has understandable variable and function names.

Problem setup

As in the MCMC practice, assume we are interested in the response, relative to a baseline, of a patient after being assigned to one of two groups. We have n patients in a treatment group that receive medication for a condition, and n other patients in a control group that receive a placebo. Measurements are obtained for the n patients in each group, denoted by Y_{Ti} and Y_{Ci} for the treatment and control groups respectively, and for $i = 1, 2, \dots, n$.

The responses are modeled as, for $G \in \{T, C\}$ denoting the patient group,

$$Y_{Gi} = \mu_G + \epsilon_{Gi},$$

where we model μ_G as a Gaussian latent effect with prior $\mu_T, \mu_C \mid \nu^2 \stackrel{iid}{\sim} N(0, \nu^2)$, we assume Gaussian error, $\epsilon_{Ti}, \epsilon_{Ci} \mid \sigma^2 \stackrel{iid}{\sim} N(0, \sigma^2)$ for $i = 1, \dots, n$, and we place inverse gamma hyperpriors on both σ^2 and ν^2 so that $\sigma^2, \nu^2 \stackrel{iid}{\sim} \text{Inv-Gamma}(\alpha, \beta)$. Hence, our hyperpriors have density:

$$p(\sigma^2) = \frac{\beta^\alpha}{\Gamma(\alpha)} (1/\sigma^2)^{\alpha+1} \exp\{-\beta/\sigma^2\}$$
$$p(\nu^2) = \frac{\beta^\alpha}{\Gamma(\alpha)} (1/\nu^2)^{\alpha+1} \exp\{-\beta/\nu^2\}.$$

Let $\mathbf{Y}_T = Y_{T1}, \dots, Y_{Tn}$ and $\mathbf{Y}_C = Y_{C1}, \dots, Y_{Cn}$.

The following code simulates the data:

```
library(invgamma)

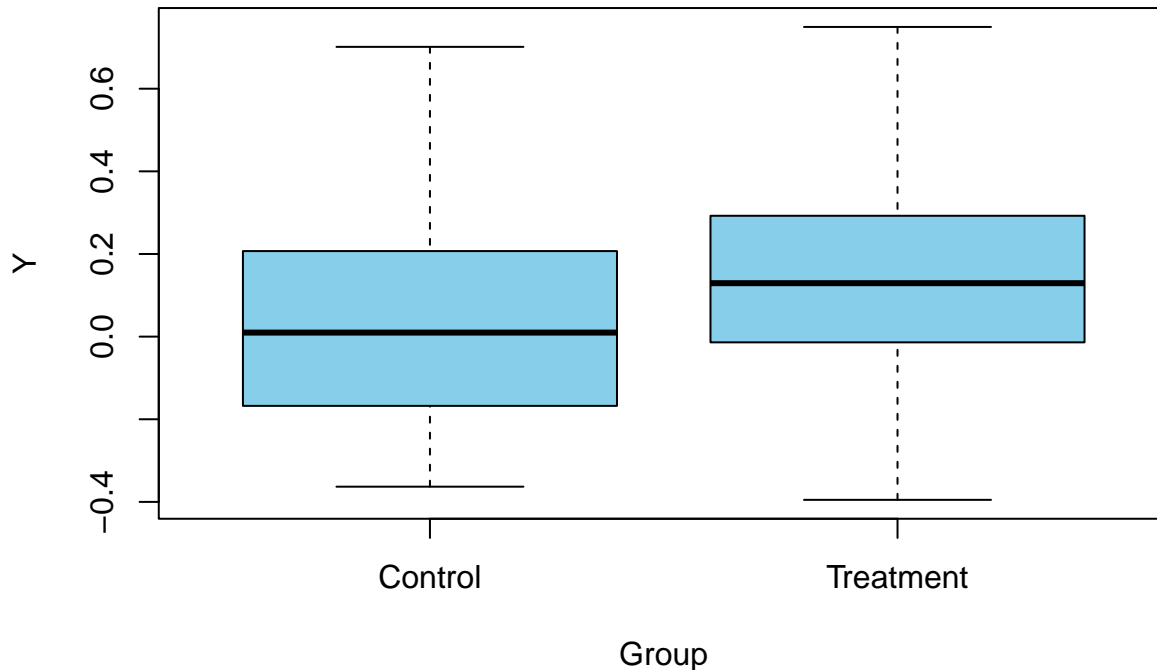
# simulate data based on true parameters
alpha = 2
beta = 0.05
set.seed(1)
sigma2 = rinvgamma(1, alpha, beta)
nu2 = rinvgamma(1, alpha, beta)
muT = rnorm(1, sd=sqrt(nu2))
muC = rnorm(1, sd=sqrt(nu2))
```

```

n=100
YT = rnorm(n, muT, sd=sqrt(sigma2))
YC = rnorm(n, muC, sd=sqrt(sigma2))

# make dataset and a boxplot of the responses in the 2 groups
dat = data.frame(Group=c(rep("Treatment", n), rep("Control", n)), Y=c(YT, YC))
boxplot(Y ~ Group, data=dat, col="skyblue")

```



Before you begin, make sure to install R-INLA with the following command:

```

install.packages("INLA", repos=c(getOption("repos"),
                                INLA="https://inla.r-inla-download.org/R/stable"), dep=TRUE)

```

1

Fit the above model in INLA assuming $\alpha = 2$ and $\beta = 0.05$. You may use the following formula:

```
library(INLA)
```

```

## Loading required package: Matrix
## Loading required package: foreach
## Loading required package: parallel
## Loading required package: sp

## This is INLA_22.12.16 built 2022-12-23 13:36:23 UTC.
## - See www.r-inla.org/contact-us for how to get help.
## - To enable PARDISO sparse library; see inla.pardiso()

# reorganize data
dat$Group[dat$Group == "Treatment"] = 1
dat$Group[dat$Group == "Control"] = 0
dat = rbind(dat[dat$Group==0,],

```

```

      dat[dat$Group==1,])
form = Y ~ -1 + f(Group, model="iid", hyper=list(prec=list(prior="loggamma", param=c(2, .05))))

```

For more information on the log gamma hyperprior on the log precision (i.e.~an inverse gamma prior on the variance), see `{r, eval=FALSE} inla.doc("loggamma")`. In addition, since we are interested in the quantity $\mu_T - \mu_C$, we can define this linear combination of parameters as a quantity of interest for INLA to calculate the posterior marginal of using the following code:

```

lincomb = inla.make.lincomb(Group=c(-1, 1))

```

This can be used with the `lincomb` argument of the `inla()` function. Give a summary of the fit model. Plot all marginals, and compare estimates of the mean of each distribution to the true values generated above (see, for example, `?inla.qmarginal`, `?inla.mmarginal` and `?inla.tmarginal`). What do the results imply about the treatment? Does it significantly improve (increase) patient outcomes relative to the placebo? Make sure to set all priors correctly (see `?control.family` and the `control.family` argument of `?inla`).

2

How do your results compare to the MCMC results from the MCMC practice exercise? Do the marginals agree with the histograms generated from the MCMC sampler? Plot them together to compare them.

3

How do the computation times between INLA and MCMC compare? You may use `proc.time()[3]` to get the current time in seconds, taking differences to get total time taken. What if n were 1000 instead of 100? 10000? Make a plot of the computation time of INLA and the MCMC sampler versus n from $n = 100$ to $n = 10000$ on a log scale.