

Homework 4

Peter Tran

```
set.seed(04232002)

library(INLA)

## Loading required package: Matrix
## Loading required package: foreach
## Loading required package: parallel
## Loading required package: sp
## This is INLA_22.04.16 built 2022-04-16 20:01:07 UTC.
## - See www.r-inla.org/contact-us for how to get help.
## - To enable PARDISO sparse library; see inla.pardiso()

library(spdep)

## Loading required package: spData
## To access larger datasets in this package, install the spDataLarge
## package with: `install.packages('spDataLarge',
## repos='https://nowosad.github.io/drat/', type='source')`

## Loading required package: sf
## Linking to GEOS 3.9.1, GDAL 3.4.0, PROJ 8.2.0; sf_use_s2() is TRUE

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr   0.3.4
## v tibble  3.1.6      v dplyr  1.0.8
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x purrr::accumulate() masks foreach::accumulate()
## x tidyr::expand()      masks Matrix::expand()
## x dplyr::filter()      masks stats::filter()
## x dplyr::lag()          masks stats::lag()
## x tidyr::pack()         masks Matrix::pack()
## x tidyr::unpack()      masks Matrix::unpack()
## x purrr::when()         masks foreach::when()

library(tmap)
```

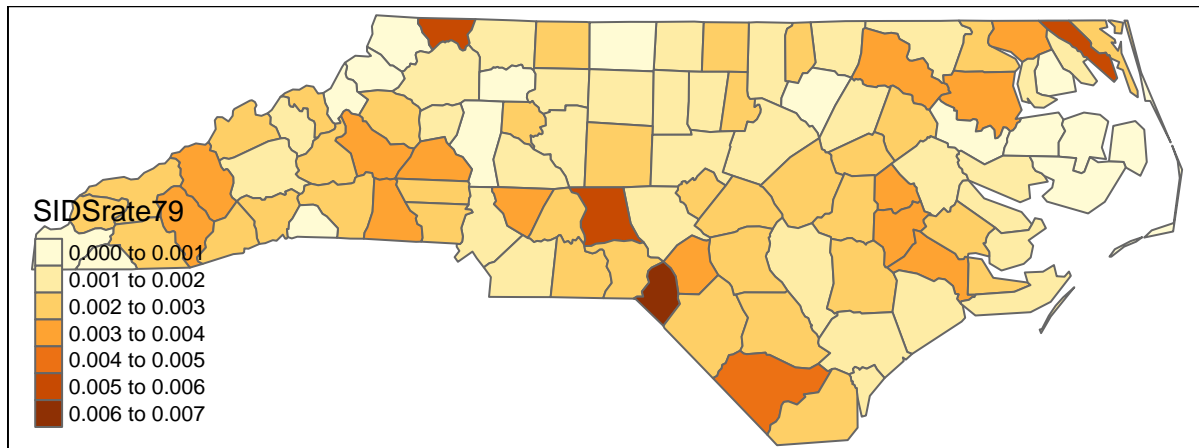
We will be examining the relationship between the number of SIDS cases in 1979 and the nonwhite birth rate during that year. Since this is a county level analysis, we may need to take into account correlation between counties, which we would use the BYM model to do so.

To begin, let's read in the data.

```
nc <- st_read(system.file("shapes/sids.shp", package="spData" )[1], quiet = T)
st_crs(nc) <- "+proj=longlat +datum=NAD27"
row.names(nc) <- as.character(nc$FIPSNO)
nc$SIDSrate79 <- nc$SID79 / nc$BIR79
nc$NWBIRrate79 <- nc$NWBIR79 / nc$BIR79
```

We can plot a choropleth map of the SIDS rate in that year to visually examine it for spatial autocorrelation.

```
tm_shape(nc) +
  tm_borders() +
  tm_fill("SIDSrate79")
```



It appears that may be the case. We can calculate Moran's I to tell us more, but for this assignment we will just assume there is spatial autocorrelation and use the BYM model with a relatively uninformative prior.

```
nb <- poly2nb(nc)
nb2INLA("nc.adj", nb)
g <- inla.read.graph(filename = "nc.adj")
file.remove("nc.adj")

## [1] TRUE

hyperPriors <- list(prec.unstruct = list(prior = "loggamma", param = c(1, 0.01)),
  prec.spatial = list(prior = "loggamma", param = c(1, 0.01)))

formula <- SID79 ~ NWBIRrate79 + f(seq(1, nrow(nc)), model = "bym", graph = g,
  hyper = hyperPriors)
model <- inla(formula, family = "poisson", data = nc, offset = log(nc$BIR79))
summary(model)

##
## Call:
## inla.core(formula = formula, family = family, contrasts = contrasts,
## data = data, quantiles = quantiles, E = E, offset = offset, " , "
## scale = scale, weights = weights, Ntrials = Ntrials, strata = strata,
## " , " lp.scale = lp.scale, link.covariates = link.covariates, verbose =
## verbose, " , " lincomb = lincomb, selection = selection, control.compute
## = control.compute, " , " control.predictor = control.predictor,
## control.family = control.family, " , " control.inla = control.inla,
## control.fixed = control.fixed, " , " control.mode = control.mode,
```

```
## control.expert = control.expert, ", " control.hazard = control.hazard,
## control.lincomb = control.lincomb, ", " control.update =
## control.update, control.lp.scale = control.lp.scale, ", "
## control.pardiso = control.pardiso, only.hyperparam = only.hyperparam,
## ", " inla.call = inla.call, inla.arg = inla.arg, num.threads =
## num.threads, ", " blas.num.threads = blas.num.threads, keep = keep,
## working.directory = working.directory, ", " silent = silent, inla.mode
## = inla.mode, safe = FALSE, debug = debug, ", " .parent.frame =
## .parent.frame)")
## Time used:
## Pre = 0.723, Running = 0.257, Post = 0.0489, Total = 1.03
## Fixed effects:
##          mean      sd 0.025quant 0.5quant 0.975quant mode kld
## (Intercept) -6.428 0.108      -6.644   -6.427      -6.221  NA   0
## NWBIRate79  0.651 0.300       0.069    0.648       1.249  NA   0
##
## Random effects:
## Name      Model
## seq(1, nrow(nc)) BYM model
##
## Model hyperparameters:
##                                     mean      sd 0.025quant
## Precision for seq(1, nrow(nc)) (iid component) 23.10 12.15      8.82
## Precision for seq(1, nrow(nc)) (spatial component) 78.36 81.12      6.32
##                                     0.5quant 0.975quant mode
## Precision for seq(1, nrow(nc)) (iid component) 20.13      54.82  NA
## Precision for seq(1, nrow(nc)) (spatial component) 53.85      296.02  NA
##
## Marginal log-Likelihood: -219.32
## is computed
## Posterior summaries for the linear predictor and the fitted values are computed
## (Posterior marginals needs also 'control.compute=list(return.marginals.predictor=TRUE)')
```

We can see that for each unit increase in the nonwhite birth rate, the number of SIDS cases increases by 0.651 (using our mean point estimate of the posterior distribution of that variable). 0 is not contained in the credible interval (though it is close), meaning that we can be reasonably confident that there really is an effect. I don't know if bayesians think of their point estimates that way.

Let's print out just those mean point estimates for $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\sigma}_v^2$, and $\hat{\sigma}_u^2$.

```
print(c("beta0hat", model$summary.fixed$mean[1]))

## [1] "beta0hat"          "-6.42835145167066"

print(c("beta1hat", model$summary.fixed$mean[2]))

## [1] "beta1hat"          "0.650557832302723"

print(c("sigma2vhat", 1 / model$summary.hyperpar$mean[1]))

## [1] "sigma2vhat"        "0.0432951822203119"

print(c("sigma2uhat", 1 / model$summary.hyperpar$mean[2]))

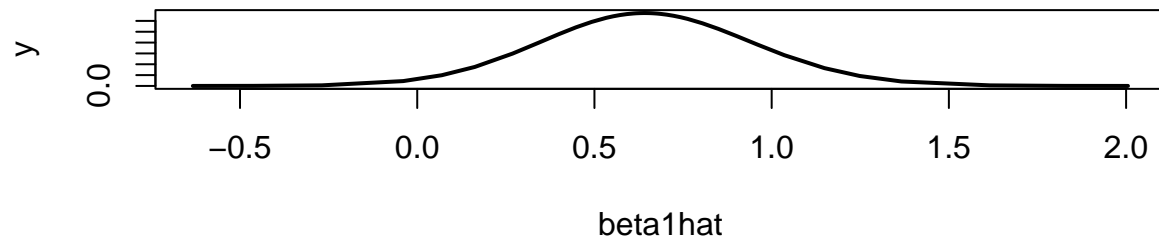
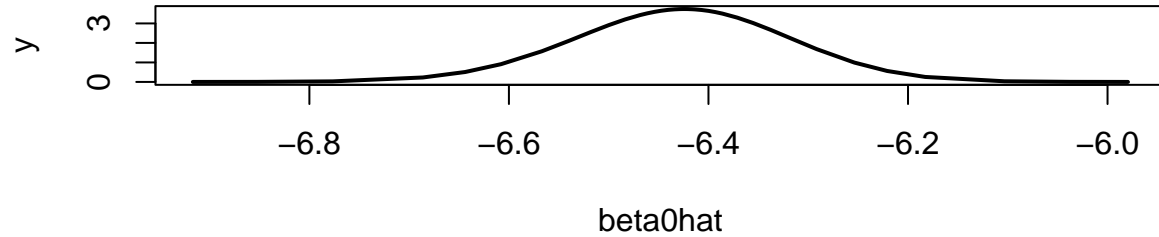
## [1] "sigma2uhat"        "0.0127622181742416"
```

Let's also plot the posterior distributions for both $\hat{\beta}_0$ and $\hat{\beta}_1$.

```

par(mfrow = c(2, 1))
plot(model$marginals.fixed$`(Intercept)`, type = "l", lwd = 2,
      xlab = "beta0hat")
plot(model$marginals.fixed$NWBIRate79, type = "l", lwd = 2,
      xlab = "beta1hat")

```



And then plot the posterior distributions for $\hat{\sigma}_v^2$ (iid component) and $\hat{\sigma}_u^2$ (spatial component).

```

par(mfrow = c(2, 1))
plot(model$marginals.hyperpar$`Precision for seq(1, nrow(nc)) (iid component)`,
      xlim = c(0, 1000), xlab = "Precision for iid component", type = "l")
plot(model$marginals.hyperpar$`Precision for seq(1, nrow(nc)) (spatial component)`,
      xlim = c(0, 1000), xlab = "Precision for spatial component", type = "l")

```

