Exercise 1 – Estimate BYM model using different hyperpriors

In the years 1986-1990 all oral cavity cancer cases in Germany were recorded. Each case was assigned to one of the 544 German districts. For each district the number of expected oral cavity cancer cases is also available. The data should be analysed using a BYM model; additionally, the sensitivity of the model with regard to different hyperpriors should be assessed.

- Load the library `INLA`, source `map.germany.txt` and `germany.txt` and read in the dataset `oral.txt`. Additionally, download the file `germany.graph.txt` which contains the information about neighbouring districts and try to understand its structure.
- Create a `data.frame()` called `oraldata` including the variables `Y, E, region, region.struct`.
- The data should be analysed using hyperpriors (a) and (c) from the lecture notes. Define the model formula for,e.g., hyperprior (a) using
```
hypera<-Y~f(region.struct,model="besag",graph.file="germany.graph.txt",param=c(1,0.005))+
          f(region,model="iid",param=c(1,0.005)).
```
  Run `inla()` for both hyperpriors using
```
mhypera<-inla(hypera,family="poisson",data=oraldata,E=E,
         control.predictor=list(compute=TRUE)).
```
- Extract the posterior mean of the relative risk using
  `rrhypera<-exp(mhypera$summary.linear.predictor$mean)`.
  Compare the results for both hyperpriors using a `summary()` of the estimated relative risks and the corresponding maps. Is there a difference in the results when using different hyperpriors?
- Compare the BYM relative risk estimates with the SMRs from Exercise sheet 4, Exercise 1 by mapping both. What is the effect of applying the BYM model to the data?

Exercise 2 – Compute Bayesian p-values for BYM estimates of relative risk

The Bayesian p-value of the relative risk is the probability that the logarithmized relative risk is larger than 0 and the relative risk is larger than 1, respectively.

- The Bayesian p-values for the relative risk in each area can be computed by adding the option `cdf=c(0)` to `control.predictor` in the `inla()` call. Estimate the BYM model `mhypera` again using the modified function call.
- Calculate the Bayesian p-values from the obtained results and map them using appropriate cutpoints as input for the function `map.germany()`. Compare BYM estimates and Bayesian p-values by plotting them in one `R` window.

Exercise 3 – Ecological regression - Lip cancer in Scotland

The dataset `scotland.txt` contains the observed and expected number of lip cancer cases in 56 counties in Scotland. Since it is suspected that people working in agriculture, fishing or forestry have a higher risk for the disease, the percentage of the population employed in these sectors is given for each conty.

- Fit a BYM model for the Scottish lip cancer data using `R` code similar to Exercise 1. Add the percentage of the population engaged in agriculture, fishing, or forestry as a linear covariate.
- Is there an association between the percentage of people working in the agricultural sector and the risk for lip cancer?

Exercise 4 − (*) Analyse simulated data - Oral cavity cancer in Germany

To assess to ability of `INLA` to estimate a true relative risk surface in the right way, data can be simulated using the expected number of oral cavity cancer cases and a prespecified value of the relative risk for all/some districts.

- Assume the relative risk is equal to 1 over the whole study region. Sample a number of observed cases for each district using the expected number of cases as rate λ of a Poisson distribution (`rpois()`). Estimate a BYM model for the simulated data and look at the obtained relative risk estimates.
- Repeat the simulation procedure, but assume that the relative risk is 1 in regions 1 to 200, but 2 in the remaining regions. What do you see when plotting the relative risk estimates?