

STATS 253 SUMMER 2015
HOMEWORK 3: BAYESIAN MODELING
Due: FRIDAY, JULY 31

Policies

- Please submit all homework files through the online homework system on the course website. You can find more instructions there.
- You may work on this homework in pairs. If you choose to do so, one person should submit all of the files (writeup, code, predictions, etc.). However, both partners should submit a statement indicating their contributions.

Description

The goal of this assignment is to get you thinking about Bayesian modeling and give you practice using a modeling language like **BUGS**. The data set below is only a suggestion. If you have your own data set (such as the one you plan to analyze for your final project), please feel free to build a model for that data set instead of the one suggested below, as long as you complete all the tasks (building a hierarchical model, using convergence diagnostics, and interpreting the posterior).

The default data set for this project is the Scottish lip cancer data. This data set contains the number of lip cancer incidences in the 56 Scottish counties between 1975-1980. The goal is to study the presence or absence of spatial dependence in the lip cancer data by building a hierarchical Bayesian model.

The shapefile is contained in **scotlip.zip**.

Tasks

1. Load the **scotlip** shapefile into R. You can do this using **readShapeSpatial** in the **maptools** package. Make a plot of the cancer incidences by county, using the **spplot** function.
2. You can obtain a regular data frame by calling **as.data.frame** on the resulting **SpatialPolygonsDataFrame** object. You can also obtain a neighbor list object by calling **poly2nb** on the object. Please see this document <https://cran.r-project.org/web/packages/spdep/vignettes/nb.pdf> for more information on creating neighbor list objects.
3. Use **JAGS** (or **BUGS** or **STAN**) to build a Bayesian model for the data. If you are working with the lip cancer data, a Poisson model probably makes sense because the data are counts. Make sure you incorporate some element of spatial autocorrelation into your model, either by autoregression or covariance modeling. However, you should note that the usual CAR model is a cyclic graph and so cannot be fit by **JAGS**. Describe your final model in your report. (You may have to experiment with several models before finding one that you like.

4. Use **JAGS** to generate samples from the posterior. Use traceplots and the Gelman-Rubin diagnostic (`gelman.diag`) to determine how long to run the chain.
5. Make plots of the posterior distribution and interpret them. Be sure to comment on what they say about the presence (or absence) of spatial autocorrelation.