

Homework 9 - Stat 534
Due Friday April 7, 2017

1. I would like you to try and reproduce some of the results in Table 9.10 on page 395 in Waller and Gotway. They fit several models to the Scottish lipcancer data set. The ones I am interested in are models $S + OD$ and the *MGLM*. I want you to use `glmmPQL` to see if you can get close to their fitted results for $S + OD$. I provide you with some code for fitting the nonlinear least squares model to see what you can come up with as an approximation to their *MGLM*. We did something similar with the Virginia lead level data earlier (take a look back at page 50 or thereabouts). I have spent hours every time I teach this course trying to come close to their results. So I have two possible outcomes in mind here: (1) you (or some of you) are able to figure out where I am going wrong which I would love to see, and/or (2) you get a birds-eye view of the issues related to trying to incorporate spatial correlation structures into generalized linear models.

The data set they provided on their website is attached. The last 2 columns are the transformed spatial locations they discuss in the text. If you look at the SAS code they used they divided these by 1000. I encourage you to work together on this one. It will go much smoother. I would like for you all to be ready to discuss this on Friday. I have been a bit easy on the exact time homeworks are due, accepting several late over the course of the semester. But I want to be able to talk about this in class on Friday so you need to have them ready by class time.

```
# R code for MGLM.
# expct is what I called the expected counts.
# emp2 is the percentage of outside workers divided by 10 (aff/10)
log.model<-function(x1,expct,b0,b1){exp(b0 + log(expct)+ b1*x1)}
require(nlme)
lipcancer.poi<-gnls(observed~log.model(emp2,expct,b0,b1),
data=lipcancer,start=c(b0=-.63,b1=0.74),corr=YOURCHOICE,
weights=varPower(form=~fitted(.),fixed=.5))
summary(lipcancer.poi)
```