Tutorial 4: Log-linear models

1) Read in the data in the file falkxyzsmr(bios794).txt. This file contains data for the 26 census tracts of the Falkirk example and has the following columns:

grid reference, male expected, female expected,deprivation score, count, smr, x centroid, y centroid.

2) Attach the data.frame and assign sensible variable names to the columns (I suggest: gridref,mexp,fexp,dep,num,smr,x,y)

3) Exploratory analysis:

a) Plot the smoothed surface of the smr (use interp,contour and image)

b) Plot the deprivation surface

c) how does deprivation relate to the smr variation?

d) How does the expected surface relate to deprivation?

4) Log-linear modelling

a) Fit a log-linear model relating deprivation to the disease count . Assess its overall goodness of fit (*gof*) and examine residual diagnostics.

b) Spatial trend: instead of fitting deprivation add x and y as variables, refit the model, and check diagnostics. Is this a better fit?

c) Spatial trend +deprivation: refit the deprivation with x and y and check *gof* and diagnostics.

Which is the best model?

d) Are there other models which could be fitted? if so fit them and check for *gof*.