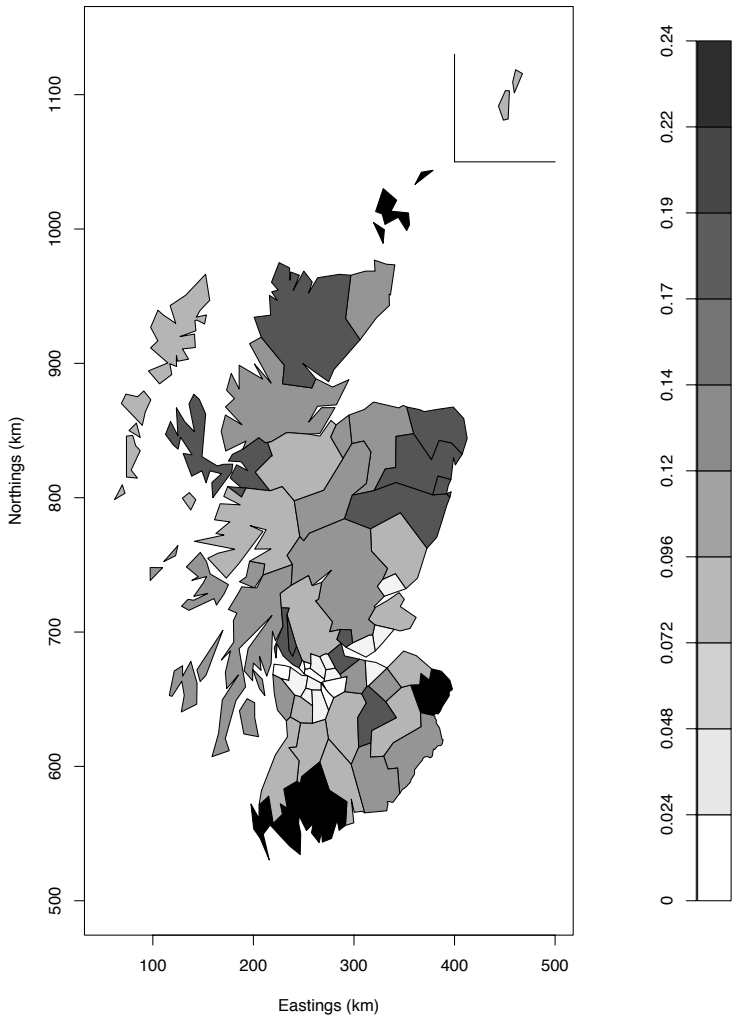
Spatiotemporal models in environmental epidemiology

Lecture 5 and 6

- In the usual implementation of regression models the standard errors are calculated under the assumption that the response data are independent.
- In a spatial context, and particularly when the areas are small, one would expect “residual” dependency between counts in areas that are geographically close, due to unmeasured risk factors that have spatial structure.
- The use of “residual” here acknowledges that known confounders and exposures are assumed to have been included in the regression model.
- Note that is doesn’t matter whether the exposure of interest has spatial structure, it’s the dependence in the residuals that is of concern.

The Scottish Lip Cancer Data

We return to this example, but now suppose that we are interested in the effect of AFF on risk. Figure **??** gives a map of the exposure.



Non-Spatial Poisson Model

The simple Poisson regression model

$$Y_i \sim \text{Poisson}(E_i \exp\{\beta_0 + \beta_1 x_i\}), \tag{1}$$

for $i = 1, ..., n$.

The ecological interpretation of $\exp(\beta_1)$ is the multiplicative change in risk between two areas whose area-level covariate x differ by one unit (ie all exposed versus all unexposed).

This model is naive in that it does not allow for overdispersion, and in particular residual spatial dependence.

The estimate (se) produced is $\widehat{\beta}_1 = 7.37$ (1.32).

The multiplicative difference in risk between areas with proportion in AFF 0.24 and 0 (the range of the observed data) is $\exp(7.37 \times 0.24) = 5.86$.

Non-Spatial Quasi-Likelihood Model

An easy way of extending the naive model is to assume the first two moments:

$$\begin{aligned} \mathsf{E}[Y_i|\beta] &= E_i \exp(\beta_0 + \beta_1 x_i) && (2) \\ \mathsf{V}(Y_i|\beta) &= \kappa \times \mathsf{E}[Y_i|\beta] && (3) \\ \mathsf{V}(Y_i, Y_j) &= 0. && (4) \end{aligned}$$

This model has the great advantage of being computationally easy to fit, and a simply interpretable parameter κ that represents the level of non-Poisson variability. Does not allow for residual spatial variability, however.

The estimate (se) produced is $\widehat{\beta}_1 = 7.37$ (1.32).

Lot of overdispersion here, $\widehat{\kappa} = 4.92$.

Non-Spatial Negative Binomial Model

An alternative parametric method of extending the Poisson model is to assume a negative binomial distribution for the counts.

As with disease mapping may assume gamma distributed random effects to give first two moments:

$$\begin{aligned} \mathbf{E}[Y_i|\beta] &= E_i \exp(\beta_0 + \beta_1 x_i) && (5) \\ \mathbf{V}(Y_i|\beta) &= \mathbf{E}[Y_i|\beta](1 + \mathbf{E}[Y_i|\beta]/\alpha) && (6) \\ \mathbf{COV}(Y_i, Y_j) &= 0. && (7) \end{aligned}$$

so we have a quadratic variance function. Also does not allow for residual spatial variability

The estimate (se) produced is $\widehat{\beta}_1 = 7.15$ (1.32). Reassuring that little change in the coefficient/se with a different variance model.