Spatiotemporal models in environmental epidemiology

Lecture 2: Spatial Epidemiology

Reference: Chapter 2

Spatial Epidemiology

Epidemiology: The study of the distribution, causes and control of diseases in human populations.

Disease risk depends on the classic epidemiological triad of person (genetics/behavior), place and time – spatial epidemiology focuses on the second of these.

Place is a surrogate for exposures present at that location, e.g. environmental exposures in water/air/soil, or the lifestyle characteristics of those living in particular areas.

Overview of Epidemiological Framework

- The *incidence proportion* measures the proportion of people who develop the disease during a specified period.
- The *prevalence proportion* is the proportion of people with the disease at a certain time.
- The *risk* is the probability of developing the disease within a specified time interval – estimated by the incidence proportion (note: a probability so between 0 and 1).
- The *relative risk* is the ratio of risks under two exposure distributions.

Precise definitions of the outcomes and exposures under study are required.
The majority of epidemiological studies are observational in nature (interventions provide an example of an experimental study).

Observational Study Types

Cohort studies select a study population and obtain exposure information, and then the population is followed over time to determine incidence. Requires large numbers of individuals (since diseases are usually statistically rare), and long study duration (for most exposures).
Case-control studies begin by identifying "cases" of the disease and a set of "controls", exposure is then determined. Although subject to selection bias, can overcome the difficulties of cohort studies.
Cross-sectional studies determine the exposure and disease outcome on a sample of individuals at a particular point of time.
Ecological studies use data on groups, areas in a spatial setting. No direct linkage between individual disease and exposures/confounders.
Semi-ecological studies collect individual-level data on disease outcome and confounders, and supplement with ecological exposure information.

Types of Data

An important distinction is whether the data arise as:

- *Point data* in which "exact" residential locations exist for cases and non-cases, or
- *Count data* in which aggregation (typically over administrative units) has been carried out. These data are *ecological* in nature, in that they are collected across groups, in spatial studies the groups are geographical areas.

We will only consider non-infectious diseases, though many of the issues transfer to infectious diseases.

Spatial Examples

Example 1: Exposure to Sunlight

In the absence of a direct measurement we might use Northings as a surrogate for exposure to sunlight.

However, if we have an available measure (average hours of sunshine by location) then we would not want to include Northings in the model.

Example 2: Confounding by Location

Often, spatial analyses will not contain all of the information on confounders and spatial location will be included in the model to act as a surrogate for the unmeasured variables.

If we are interested in estimating the association between a health outcome and an environmental pollutant then great care must be taken in modeling space: including a complex term for the spatial model will dilute the estimate of the effect of pollution (since this has spatial structure), and including a very simple term may not be sufficiently subtle to control for the unmeasured confounders.

Socio-economic confounding

- In spatial epidemiological applications that use count data, population data are obtained from the census and so while one can control for known factors such as age and gender (and sometimes race), information is not available on other possible confounders.
- In such situations it has become common to control for a measure of *socio-economic status.*
- Across various scales of aggregation, measures of deprivation have been shown to be powerful predictors of a variety of health outcomes.
- Deprivation may be viewed as a surrogate for individual-level characteristics such as smoking, diet and alcohol consumption.
- Could be true area-level effects, however, for example, access to health care services.
- Relationship between health, socio-economic status and exposure to environmental pollution is complex since ill-health may cause deprivation (e.g. lose job) so that Y causes Z.

Carstairs Index

A number of area-level indices of deprivation have been created in the UK (e.g., Carstairs, Jarmen, Townsend). In the US income and education may be used as surrogates.

The *Carstairs index* has been extensively used by the Small Area Health Studies Unit (SAHSU), a number of whose studies we shall use as illustration in this course. This index measures (from the census) the proportion of individuals within each ED who: are unemployed; live in overcrowded accommodation; lack a car; have a head of the household who is in low social class.

These variables are standardized across the country and then added together to give a continuous area-based measure with high values indicating increased deprivation.

Important Point: since control is at the ecological, and not the individual, level, the control is not likely to be strong – casting doubt in situations in which *small* relative risks are observed.

- The basic measures for assessing the health of a community and its needs for health services are the size and composition of the human population and the counts of vital events occurring within them (births, deaths, morbid disorders, etc). Such measures are known as *vital statistics*.
- Indices, that is summary statistics, derived from such raw data in terms of rates, ratios and proportions are of use to many workers in the health field. Understanding these indices and their uses should help towards efficient provision and use of resources and to appropriate preventative measures targeted at susceptible sections of the community.
- By themselves, the raw data are of little use until they are transformed to standard indices for valid comparisons.

- There is only one accurate way to estimate the size of the relevant population, and describe its demographic characteristics, and that is to count it at a particular point in time. This is known as a *census* and is usually done once every ten years.
- The second major problem of measurement in public health is to record the demographic events that occur within the community: births, deaths, marriages and migration both in and out of the area. If these are measured accurately and continuously then the changes in population from year to year can be determined, and the frequent need for censuses eliminated.
- Population size estimates between census years are only an approximation. Some estimates, such as the numbers in different occupations, are so unreliable that mortality data for them is only tabulated in census years.

A *rate* is defined as the number of events, for example deaths or cases of disease, per unit of population, in a particular time span. To calculate a rate we require:

- a defined period of time
- a defined population, with an accurate estimate of the size of the population during the defined period
- the number of events occurring over the period.

$$\text{rate} = \frac{\text{No. events}}{\text{total person-time at risk}}$$

For a fixed time period Δt, an average size of the population at risk during that period $\bar{N}$ and the number of events A the rate is

$$\text{rate} = \frac{A}{\bar{N} \times \Delta t}$$

- The *incidence rate* refers to the number of new cases of a particular disease that develop during a specified time interval.
- For a fixed time period Δt, a population at risk of size N and the number of new cases of disease A the incidence rate is

$$\text{incidence} = \frac{A}{N \times \Delta t}$$

- An incidence rate lies between 0 and ∞, and a yearly incidence rate measures the number of cases per person-year.
- Measuring the population at risk may be difficult because of changes in the population over the time period.
- Since it may not be possible to measure disease-free periods precisely, we often calculate the incidence using the average size of the population. This is reasonably accurate if the population size is stable and the incidence rate is low.

- The *prevalence* refers to the number of cases of disease that exist at a specified point in time. It is the proportion of the population who have a disease at a given time.
- Prevalences may also be calculated over a time period; for example the number of events within a time period.

$$\text{prevalence} = \frac{\text{Number of cases}}{\text{Number at risk}}$$

- Diseases with high incidence rates may have low prevalences if they are rapidly fatal.

- The *crude mortality rate* is usually calculated as deaths per 1000 population per year.
- Let D be the number of deaths in a given time period of length Δt, and $\bar{N}$ be the average size of the population at risk during that period (often approximated by the number in the population at the mid-point of the time period).
- Then the crude mortality rate is given by

$$\mathrm{r} = \frac{d}{\bar{N} \times \Delta t} \times 1000$$

- Rates may be required for particular sections of a community and these are referred to as *specific rates*; that is, where the populations are *specified*; that is the denominators.
- For example, age-specific or age- and sex-specific rates may be used for comparison of different populations.
- Other common specific rates are area, occupation or social class specific (and combinations of these).

- infant mortality rate number of deaths under one year of age after live birth, divided by the number of live births
- neonatal mortality rate number of deaths at 0 to 27 days after live birth, divided by the number of live births
- stillbirth rate number of stillbirths, divided by the total number of births, live and still
- perinatal mortality rate number of stillbirths and deaths at days 0 to 6, divided by the total number of births
- birth rate number of live births per year, divided by total population
- fertility rate number of live births per year, divided by number of women 15-44 (ie of childbearing age)

Table: Mortality Data by Age Group for South West of England 2003. *Source: ONS Mortality Statistics Series DH1 no. 36*

Age Group	Population (1000s)	% in age group	Deaths	Age-specific mortality rate
0-4	259.7	5.2	244	
5-14	609.3	12.2	59	
15-24	591.0	11.8	251	
25-44	1319.3	26.4	1219	
45-64	1282.2	25.6	5944	
65-74	470.7	9.4	8668	
75+	467.1	9.3	39480	

Table: Mortality Data by Age Group for England 2003. *Source: ONS Mortality Statistics Series DH1 no. 36*

Age Group	Population (1000s)	% in age group	Deaths	Age-specific mortality rate
0-4	2848.2	5.7	3682	1.3
5-14	6299.8	12.6	725	0.1
15-24	6304.0	12.6	2634	0.4
25-44	14485.7	29.1	14001	1.0
45-64	11971.3	24.0	62755	5.2
65-74	4158.6	8.3	87714	21.1
75+	3788.3	7.6	331898	87.6
All	49855.9	100	503409	10.1

- When comparing populations, we can eliminate the effects of, for example, different age structures by looking at age-specific rates.
- However, this can be cumbersome, and it is often easier to compare a single summary figure.

- For direct standardisation, we use a standard population structure for reference.
- We then calculate the overall mortality rate that this reference population would have observed if it had the age-specific mortality rates of the population of interest.
- Suppose the reference population has population counts $N'_k; k = 1, \ldots, K$ in each age-group k.
- We calculate the age-specific mortality rates r_k for the population of interest.
- The directly standardised rate is given by

$$\text{directly standardised rate} = \frac{\sum_{k=1}^{K} N'_k r_k}{\sum_{k=1}^{K} N'_k}$$

- For indirect standardisation, we take the age-specific rates from the reference population and convert them into the mortality rate we would observe if those reference rates were true for the age-structure of the population of interest.
- This gives us the expected rate for the population of interest, if age-specific mortality rates were the same as for the reference population.
- We calculate the age-specific mortality rates r'_k for the reference population. Suppose the population of interest has population counts $N_k; k = 1, \ldots, K$ in each age-group k.
- The expected rate of deaths in the population of interest is

$$\text{expected rate} = \frac{\sum_{k=1}^{K} N_k r'_k}{\sum_{k=1}^{K} N_k}$$

and the expected number of deaths in the population of interest is

$$E = \sum_{k=1}^{K} N_k r'_k$$

- We can compare the expected number of deaths, using the indirect standardisation method, with the observed number using the *standardised mortality ratio* (SMR).
- Let O be the observed number of deaths in the population of interest, and E be the expected number of deaths when indirectly standardised with respect to some reference population.

$$\text{SMR} = \frac{O}{E} \times 100$$

- The SMR is a ratio, not a rate or a percentage. An SMR of 100 means that the population of interest has the same number of deaths as we would expect from the reference population.
- If it is greater than 100, we have more deaths than expected; if it is less than 100 we have less.

Age, for example, will almost always need controlling for – different disease risks in different area may reflect differences in the age population. There are a number of ways to control for confounding, and one method is direct or indirect standardization.
Let Y_{ij} denote the number of cases, within some specified period (in years) of S (which is assumed the same for each individual in the study population) in area i and confounder stratum j, and N_{ij} be the corresponding population at risk, $i = 1, ..., m$, $j = 1, ..., J$.
Let Z_j and M_j denote the number of cases and population in stratum j in a "reference", or standard, population.

The risk of disease in confounder stratum j in area i, over the time period T, is $\widehat{p}_{ij} = Y_{ij}/N_{ij}$. The rate of disease is $r_{ij} = 1000 \times Y_{ij}/[N_{ij} \times T]$ per 1000 person years. Note that a rate does not need lie between 0 and 1. The crude rate in area i is given by $1000 \times Y_i / \left\{ \sum_{j=1}^{J} N_{ij} \times S \right\}$ per 1000 person years.

The directly standardized rate in area i is given by $\sum_{j=1}^{J} r_{ij} w_j$, per 1000 person years, where $w_j = M_j / \sum_j M_j$ is the proportion of the population in stratum j (these weights may be based on the world, or a uniform, population.

The directly standardized rate is a weighted average of the stratum-specific risks, and corresponds to a "counter-factual" argument in which the estimated rates within the study region are applied to the standard population.

If $q_j = 1000 \times Z_j / [M_j \times T]$ is a standard disease rate in stratum j then the comparative mortality/morbidity figure (CMF) for area i is given by:

$$\text{CMF}_i = \frac{\sum_{j=1}^{J} r_{ij} w_j}{\sum_{j=1}^{J} q_j w_j}$$

In small-area studies in particular the CMF is rarely used since it is very unstable, due to small counts by stratum in area i, Y_{ij}.

The method of indirect standardization produces the standardized mortality/morbidity ratio (SMR):

$$\frac{Y_i}{\sum_{j=1}^{J} N_{ij}\widehat{q}_j}$$

where $Y_i = \sum_j Y_{ij}$ is the total number of cases in area i, and $\widehat{q}_j = Z_j/M_j$ is a reference risk.
The indirectly standardized rate compares the total number of cases in an area to those that would result if the rates in the reference population were applied to the population of area i.
Which reference rates to use? In a regression setting dangerous to use internal standardization in which $\widehat{q}_j = \sum_i Y_{ij} / \sum_i N_{ij}$.
External standardization uses risks/rates from another area.

Expected Numbers

The expected numbers $E_i = \sum_{j=1}^{J} N_{ij} q_j$ follow from assuming the proportionality model

$$p_{ij} = \theta_i q_j$$

where θ_i is the relative risk associated with area i (this assumption removes the need to estimate J risks in each area). Since

$$\mathsf{E}[Y_{ij}] = N_{ij} \theta_i q_j$$

we obtain

$$\mathsf{E}[Y_i] = \sum_{j=1}^{J} N_{ij} \theta_i q_j = \theta_i E_i.$$

The SMR is therefore given by

$$\mathsf{SMR}_i = \frac{Y_i}{E_i}.$$

If incidence is measured then also known as the Standardized Incidence Ratio (SIR). Control for confounding may also be carried out using regression modeling.

In routinely carried out investigations the constituent data are often subject to errors.

Population data

- Population registers are the gold standard but counts from the census are those that are typically routinely-available.
- Census counts should be treated as estimates, however, since inaccuracies, in particular underenumeration, are common.
- For inter-censual years, as well as births and deaths, migration must also be considered.
- The *geography*, that is, the geographical areas of the study variables, may also change across censuses which causes complications.

Health data.

- For any health event there is always the possibility of diagnostic error or misclassification.
- For other events such as cancers, case registrations may be subject to double counting and under registration.

In both instances *local knowledge* is invaluable.

Exposure data

- Exposure misclassification is always a problem in epidemiological studies.
- Often the exposure variable is measured at distinct locations within the study region, and some value is imputed for all of the individuals/areas in the study.
- A measure of uncertainty in the exposure variable for each individual/area is invaluable as an aid to examine the sensitivity to observed relative risks.

Wakefield and Elliott (1999, Statistics in Medicine) contains more discussion of these aspects.
In terms of combining the population, health and exposure data, this is easiest if such data are *nested*, that is, the geographical units are non-overlapping.

- A GIS is a computer-based set of tools for collecting, editing, storing, integrating, displaying and analyzing spatially referenced data.
- A GIS allows linkage and querying of geographically indexed information. So for example, for a set of geographical residential locations a GIS can be used to retrieve characteristics of the neighborhood within the locations lies (e.g. census-based measures such as population characteristics and distributions of income/education), and the proximity to point (e.g. incinerator) and line (e.g. road) sources.
- Buffering – a specific type of spatial query in which an area is defined within a specific distance of a particular point, line or area.

- Time activity modeling of exposures – we may trace the pathway of an individual, or simulate the movements of a population group through a particular space-time concentration field, in order to obtain an integrated exposure.
- In this course, we will not use any GIS tools, but use capabilities within `R` and `WinBUGS/GeoBUGS` to display maps.

Examples

Figure 1 shows a map of Washington state with various features superimposed; this was created with the Maptitude GIS.
Figure 2 smoothed relative risk estimates for bladder cancer.
Figure 3 shows 16 monitor sites in London – a GIS was used to extract mortality and population data within 1km of the monitors, and the association with SO_2 was estimated.

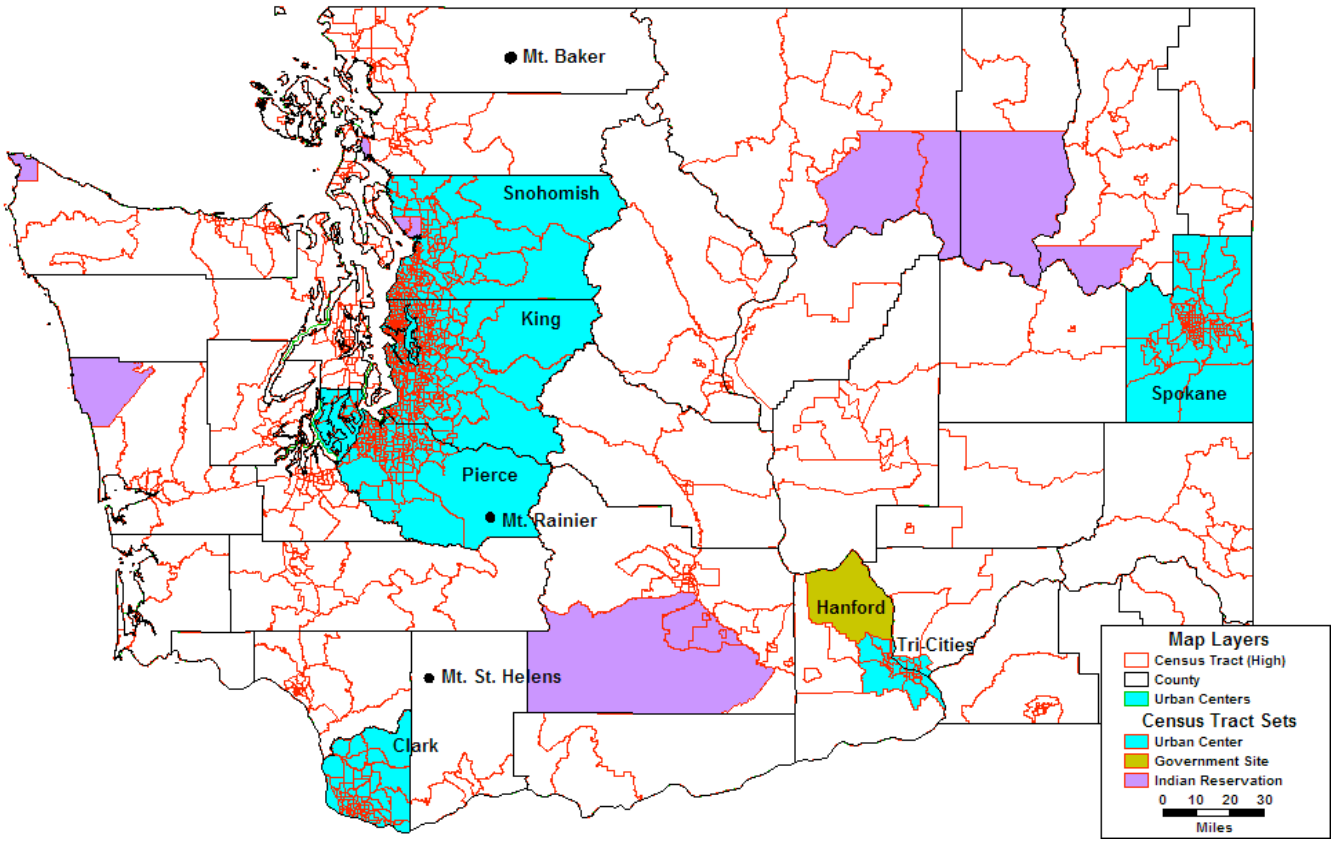

Figure: Features of Washington state, created using a GIS.

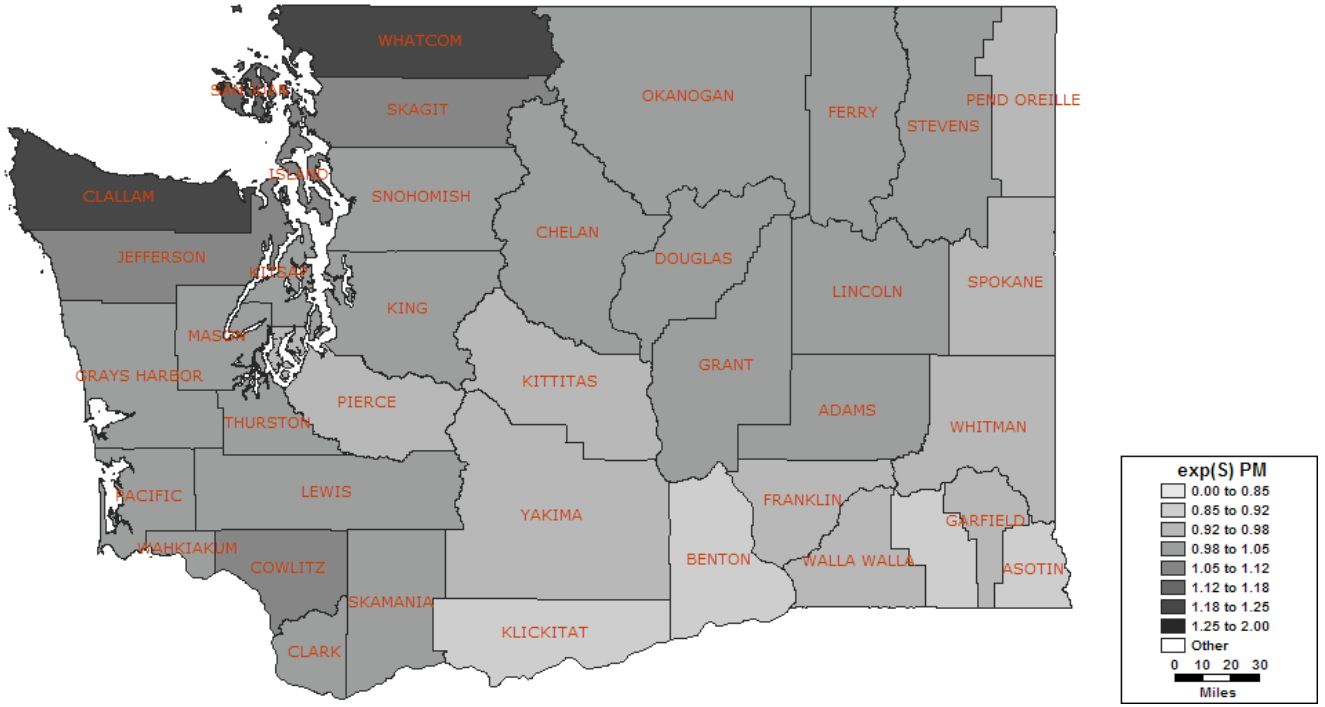

Figure: Smoothed relative risk estimates for bladder cancer in 1990–2000 for counties of Washington state.

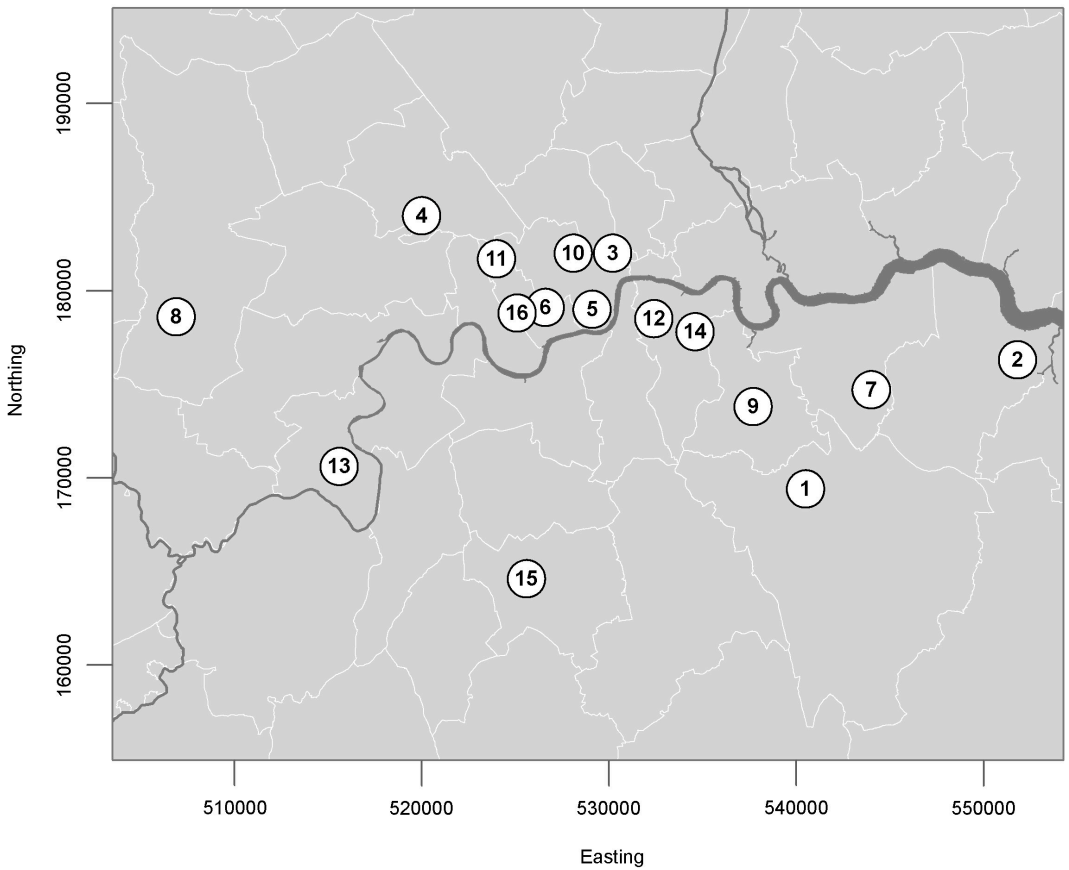

Figure: Air pollution monitor sites in London.

GLMs are a convenient family for fitting a range of data types – we will use the `glm` function in `R`. A GLM is defined by:

- The data arise as an independent sample from an exponential family probability distribution; this family includes the normal, binomial and Poisson distributions.
- A link function linking the mean function, $\mu = \mathrm{E}[Y]$ to a linear predictor $g(\mu) = \mathbf{x}\beta$; logistic regression and log-linear models form two common examples.

By assuming a linear predictor certain aspects of inference are simplified, both in terms of computation and properties of the resultant estimates.

In their original form, GLMs assume independent data, GLMMs extend this to allow dependence induced by random effects.
The link function now has

$$g(\mu_i) = \mathbf{x_i}\beta + \mathbf{b_i},$$

where b_i represents a *random effect*.
The random effects are then assigned a distribution, and in a spatial setting it is natural to assume

$$\mathbf{b} = (\mathbf{b_1}, ..., \mathbf{b_m})^{\mathbf{T}} \sim_{\mathbf{iid}} \mathbf{N_m}(\mathbf{0}, \mathbf{\Sigma}),$$

where $\mathbf{\Sigma}$ is an $m \times m$ covariance matrix with (i, j)-th element defining the covariance between random effects at locations i and j.
A simple choice model is $\mathbf{\Sigma_{ij}} = \sigma^{\mathbf{2}}\rho^{\mathbf{d_{ij}}}$, for $i, j = 1, ..., m$, with $\sigma^2 > 0, 0 < \rho < 1$ and d_{ij} the distance between the centroids of areas i and j. This model is *isotropic* since the covariance only depends on the distance between points.

Estimation of parameters may be based on specifying mean and variance of the data only, as in quasi-likelihood, or on specifying the complete probability distribution of the data, as in likelihood and Bayesian approaches.
The likelihood function is the probability distribution viewed as a function of the unknown parameter, and maximum likelihood estimation (MLE) the estimation criteria that chooses that value of the parameter that gives the highest probability to the observed data.
For most models the MLE is asymptotically normal which allows confidence intervals/tests to be constructed.

Example:

Poisson likelihood.
Suppose we have a count Y in an area with expected number E.
Assumed probability model for data, for fixed λ:

$$\Pr(Y = y|\lambda) = \frac{\mathrm{e}^{-E\lambda}(E\lambda)^y}{y!}$$

for $y = 0, 1, ...$ Here λ is the relative risk.
For fixed y we have the likelihood function:

$$l(\lambda) = \frac{\mathrm{e}^{-E\lambda}(E\lambda)^y}{y!} \propto \mathrm{e}^{-E\lambda}\lambda^y$$

for $\lambda > 0$.

Example: Seascale excess

Figure 4 gives an example for $y = 4$, $E = 0.25$ for which the MLE is $\widehat{\lambda} = 16 = \mathrm{e}^{\widehat{\alpha}} = \mathrm{e}^{2.773}$ with 95% asymptotic confidence interval

$$(\mathrm{e}^{2.773-1.96\times0.5}, \mathrm{e}^{2.773+1.96\times0.5}) = (6.0, 42.6).$$

Here is the `R` code for finding the MLE and the standard error:

```
> y <- 4; E <- 0.25
> summary(glm(y~1+offset(log(E)),family=poisson))
Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)    2.773      0.500   5.545 2.94e-08 ***
```

The "offset" is the known multiplier in the log-linear mean function:

$$\log \mu = \log E + \alpha$$

and $\sim$`1` denotes the intercept.
Notice that the parameter is on the linear predictor scale.

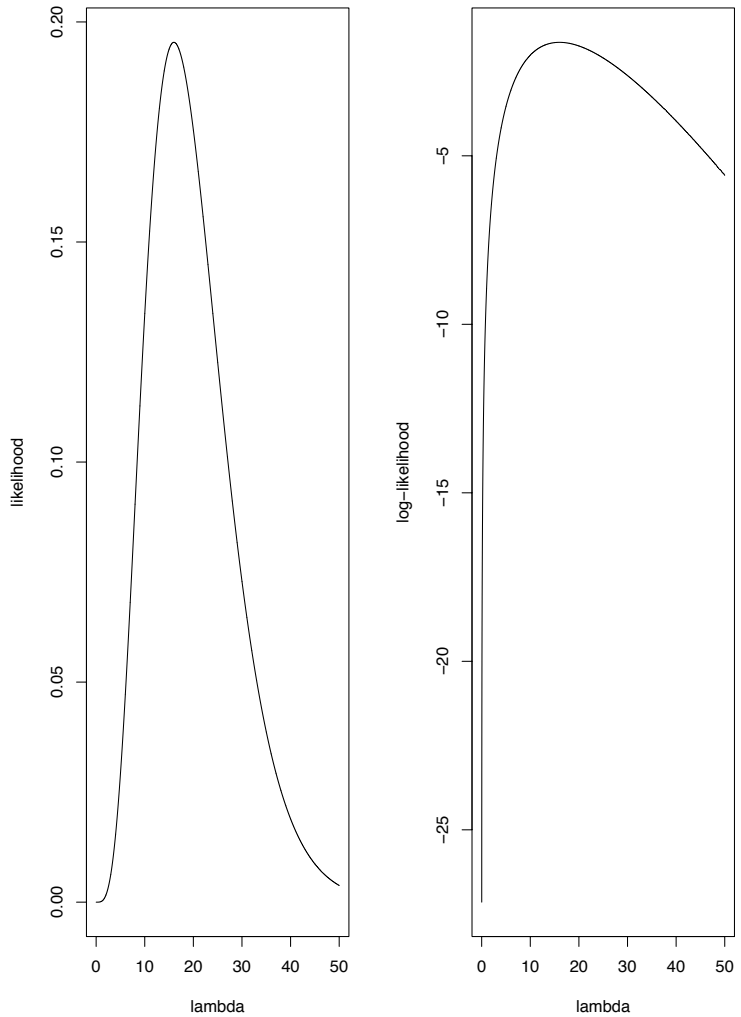

Figure: Likelihood (left) and log-likelihood (right) for Poisson data with $y = 4$, $E = 0.25$.

For the log-linear model

$$Y_i \sim_{ind} \text{Poisson}(\mu_i),$$

with

$$\log \mu_i = \log E_i + \alpha + \beta X_i$$

$i = 1, ..., n$ the MLEs for α and β are not available in closed form but reliable maximization routines are available in all statistical packages.

Example: Scottish Lip Cancer Data

Incidence rates of lip cancer in males in 56 counties of Scotland, registered in 1975–1980. These data were originally reported in the mapping atlas of Kemp, Boyle, Smans and Muir (1985).
The form of the data is:

- Observed and "expected" number of cases (based on the county age populations, details shortly) – allows the calculation of the standardized morbidity ratio, the ratio of the observed to the expected cases.
- A covariate measuring the proportion of the population engaged in agriculture, fishing, or forestry (AFF).
- The projections of the longitude and latitude of the area centroid, and the "position" of each county expressed as a list of adjacent counties.

County No. i	Obs Cases Y_i	Exp Cases E_i	Prop AFF	SMR	Project N (km)	Project E (km)	Adjacent Counties
1	9	1.4	0.16	6.43	834.7	162.2	5,9,19
2	39	8.7	0.16	4.48	852.4	385.8	7,10
3	11	3.0	0.10	3.67	946.1	294.0	12
4	9	2.5	0.24	3.60	650.5	377.9	18,20,28
5	15	4.3	0.10	3.49	870.9	220.7	1,12,19
6	8	2.4	0.24	3.33	1015.2	340.2	Island
7	26	8.1	0.10	3.21	842.0	325.0	2,10,13,16,17
8	7	2.3	0.07	3.04	1168.9	442.2	Island
9	6	2.0	0.07	3.00	781.4	194.5	1,17,19,23,29
...							
47	2	5.6	0.01	0.36	640.8	277.0	24,31,46,48,49,53
48	3	9.3	0.01	0.32	654.7	282.0	24,44,47,49
49	28	88.7	0.00	0.32	666.7	267.8	38,41,44,47,48,52,53,54
50	6	19.6	0.01	0.31	736.5	342.2	21,29
51	1	3.4	0.01	0.29	678.9	274.9	34,38,42,54
52	1	3.6	0.00	0.28	683.7	257.8	34,40,49,54
53	1	5.7	0.01	0.18	646.6	265.6	41,46,47,49
54	1	7.0	0.01	0.14	682.3	267.9	34,38,49,51,52
55	0	4.2	0.16	0.00	640.1	321.5	18,24,30,33,45,56
56	0	1.8	0.10	0.00	589.9	322.2	18,20,24,27,55

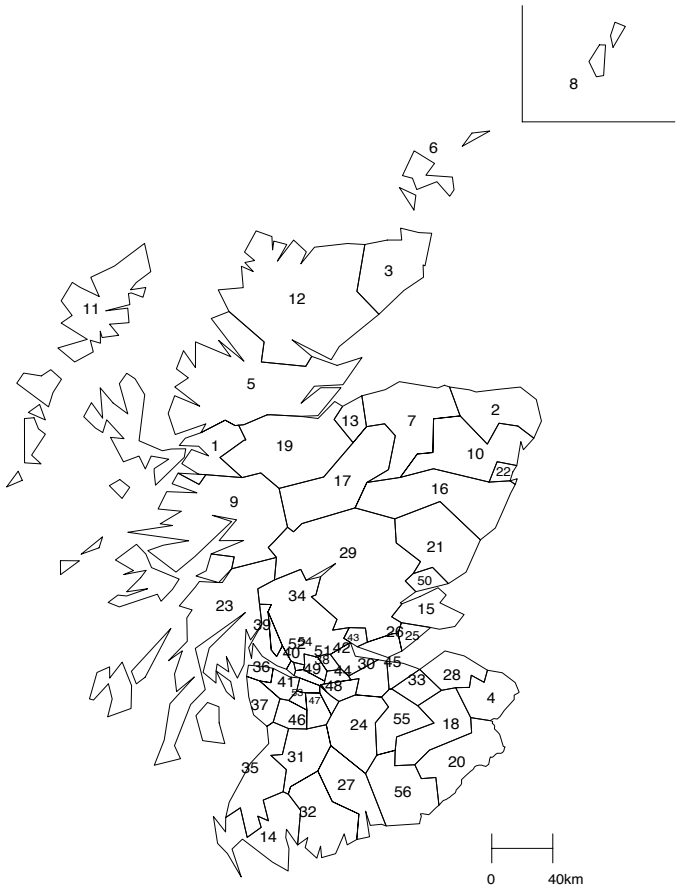

Figure: Labels for 56 counties of Scotland.

Figure: SMRs for male lip cancer in 56 counties of Scotland.

Example: Simple regression in the Scottish Lip Cancer Data

The file `scotdat.txt` contains the Scottish data as a `list`:

```
z <- list(N = 56, Y = c( 9, 39, 11 ... 1, 0, 0),
          E = c( 1.4, 8.7, 3.0... 7.0, 4.2, 1.8),
          X = c( 0.16, 0.16, 0.10 ... 0.01, 0.16, 0.10))
> source(``scotdat.txt'')
> SMR <- z$Y/z$E
> postscript("scot_smr.ps",horiz=F)
> par(mfrow=c(1,2)) # creates a 1 x 2 plot
> hist(SMR,xlab="SMR")
> plot(z$X,SMR,type="n")
> text(z$X,SMR)
> lines(lowess(z$X,SMR))
> dev.off()
```

This code creates a postscript file for Figure 5.

We carry out likelihood analyses using the `glm` function and the log-linear mean function

$$\log \mathsf{E}[Y_i] = \log E_i + \alpha + \beta X_i, \quad i = 1, ..., 56.$$

$\exp(\beta)$ represents the difference in area-level relative risk between areas with all the population in AFF and zero of the population in AFF.

```
> summary(glm(Y~offset(log(E))+X,data=z,family=poisson(link
Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.54227    0.06952   -7.80 6.21e-15 ***
X            7.37322    0.59557   12.38  < 2e-16 ***

(Dispersion parameter for poisson family taken to be 1)
```

So $\widehat{\alpha} = -0.542$ (0.070) and $\widehat{\beta} = 7.37$ (0.60) – the relative risk describing the area-based association between incidence and AFF is exp(7.37) = 1588!!!

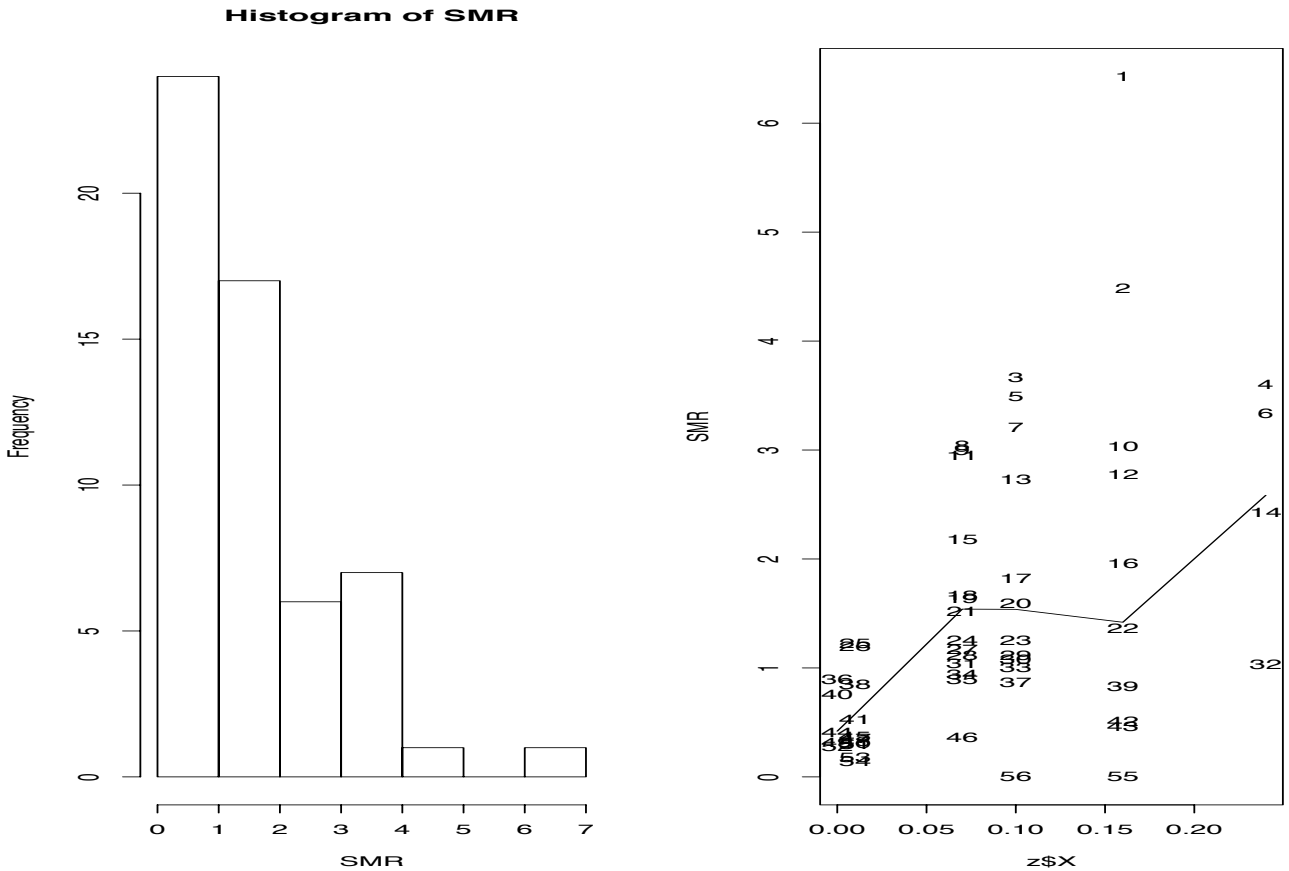

Figure: Histogram of SMRs for the Scottish data, and SMR versus proportion in AFF X.

The Poisson model is restrictive in the sense that the variance is constrained to equal the mean.
In a quasi-likelihood approach we assume

$$\text{var}(Y_i) = \kappa \text{E}[Y_i]$$

where κ allows *overdispersion* and is estimated as

$$\widehat{\kappa} = \frac{1}{n-p} \sum_{i=1} \frac{(Y_i - \widehat{\mu}_i)^2}{\widehat{\mu}_i^2}$$

where n is the number of counts, and p the number of estimated parameters.
Point estimates are identical to those from likelihood, but standard errors are multiplied by $\widehat{\kappa}^{1/2}$.

To fit a quasi-likelihood model:

```
> summary(glm(Y~offset(log(E))+X,data=z,family=quasipoisson
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.5423     0.1542  -3.517 0.000893 ***
X             7.3732     1.3208   5.583 7.89e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
(Dispersion parameter for quasipoisson family taken to be 4
```

So $\widehat{\alpha} = -0.542$ (0.15) and $\widehat{\beta} = 7.37$ (1.32) – identical point estimates and standard errors multiplied by $\sqrt{4.92} = 2.22$; large overdispersion here, and the Poisson model is clearly inadequate.