Special Spatial Distributions

◆ Spatial correlation is not catered for in the Poisson-gamma model. That model only allows extra –variation in risk

◆ Alternative models are log linear in the risk

  – This allows the inclusion of other terms in the model (v and u)

◆ Example:

$$y_i \sim Poiss(e_i \theta_i)$$
$$\log(\theta_i) = \alpha_0 + v_i + u_i$$

Uncorrelated and Correlated heterogeneity (UH, CH)

- The two UH and CH components are

- These have normal distributions

- The second distribution (CAR) is conditional autoregressive and defines *spatial correlation* between regions

$$v_i + u_i$$

$$v_i \sim N(0, \tau_v)$$
$$u_i \sim N(\bar{u}_{\delta_i}, \tau_u / n_{\delta_i})$$

where δ_i is a neigborhood of the i th region and n_{δ_i} is the number of neighbors

Log linear models

◆ Assume a Poisson data likelihood and we want to model the log of the relative risk....in general we have GLMM:

$$\log(\theta_i) = x_i'\beta + z_i'\gamma$$

◆ Linear predictor with covariates $x_i'\beta$

◆ Random effect predictor $z_i'\gamma$

◆ Special (simple) case:

$$\log(\theta_i) = \alpha_0 + v_i + u_i$$

CAR models

◆ On WinBUGS the CAR model can be fitted using the **car.normal()** distribution

◆ This requires an adjacency vector and a list of the number of neighbors for each region

b[1:J]~car.normal(adj[],wei[],num[],tau.b)

New Example

◆ South Carolina congenital abnormality deaths 1998 by county

◆ With number of deaths and number of births and % poverty per county

◆ This is a binomial example:

$$y_i \sim bin(p_i, n_i)$$
$$\log it(p_i) = \alpha_0 + \beta x_i + v_i + u_i$$

SCbinomial

```
#In this case we examine the congenital data
 for 1998 with a % of pop classified by the US
 #census as in poverty
# Data file is SCcongen_bin_pov.txt
#spatial file:below

model
        {
         for( i in 1 : J) {
            y[i] ~ dbin(p[i], n[i])
            logit(p[i]) <- al0+al1* X[i]+b[i]+h[i]
        h[i]~dnorm(0.0,tau.h)
            }
        al0~dflat()
        al1~dnorm(0.0,tau.al1)
        b[1:J]~car.normal(adj[],wei[],num[],tau.b)

        for (i in 1:sumNumNeigh){
        wei[i]<-1.0
        }

        tau.h~dgamma(0.001,0.001)
        tau.al1~dgamma(0.001,0.001)
        tau.b~dgamma(0.001,0.001)

        }
# main data is in SCcongen98_data.txt
Main data ➜click on one of the arrows to open data⬅
```

SCcongen_bin_pov.txt

The following data repesents the number of cases of
congenital abnormality death for 1998 in the counties of South Carolina,
the number of births in South Carolina, and a census indicator
of percentage of the population below a poverty level
within each county for the same period.

cg98	birth98	%poverty
0.0	364.0	13.6
5.0	1797.0	13.8
0.0	177.0	32.3
16.0	2145.0	11.0
0.0	208.0	24.2
0.0	338.0	19.9
3.0	1668.0	12.3
5.0	2017.0	13.3
0.0	168.0	17.0
7.0	4403.0	15.4
1.0	686.0	14.1
5.0	476.0	15.7
1.0	550.0	18.0
2.0	409.0	24.3
1.0	519.0	21.8
3.0	907.0	20.2
2.0	489.0	24.9
5.0	1152.0	12.0
2.0	265.0	17.4
1.0	335.0	18.1
3.0	1867.0	18.7
2.0	736.0	17.5
14.0	5176.0	10.6
7.0	963.0	13.8
1.0	311.0	22.8
9.0	2324.0	13.7
1.0	280.0	21.0
3.0	665.0	12.6
1.0	809.0	14.0
3.0	864.0	14.0
1.0	258.0	26.9

Start   My Computer   Local Disk (C:)   8:13 P
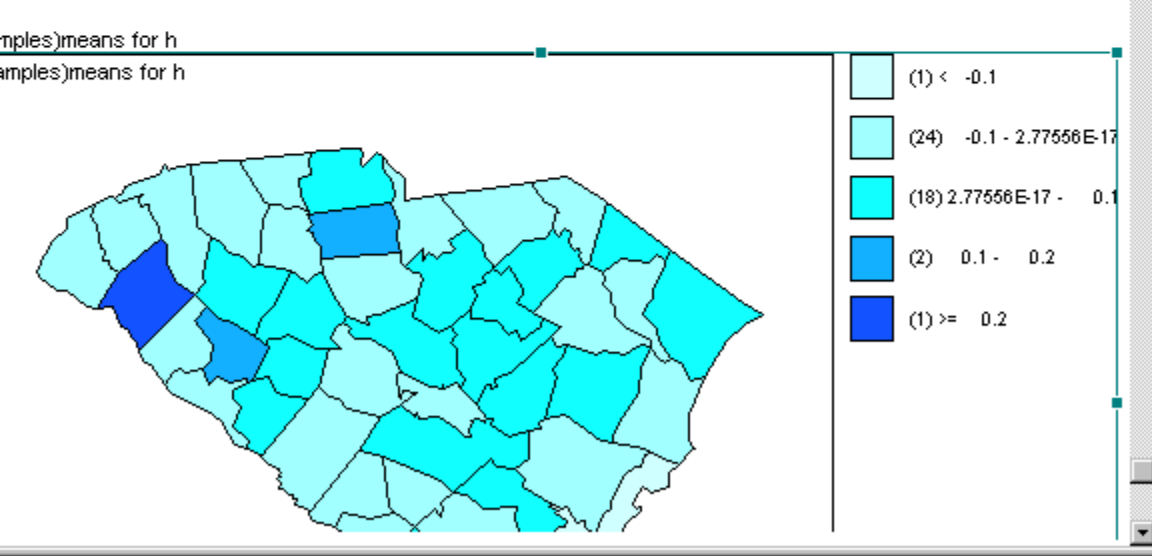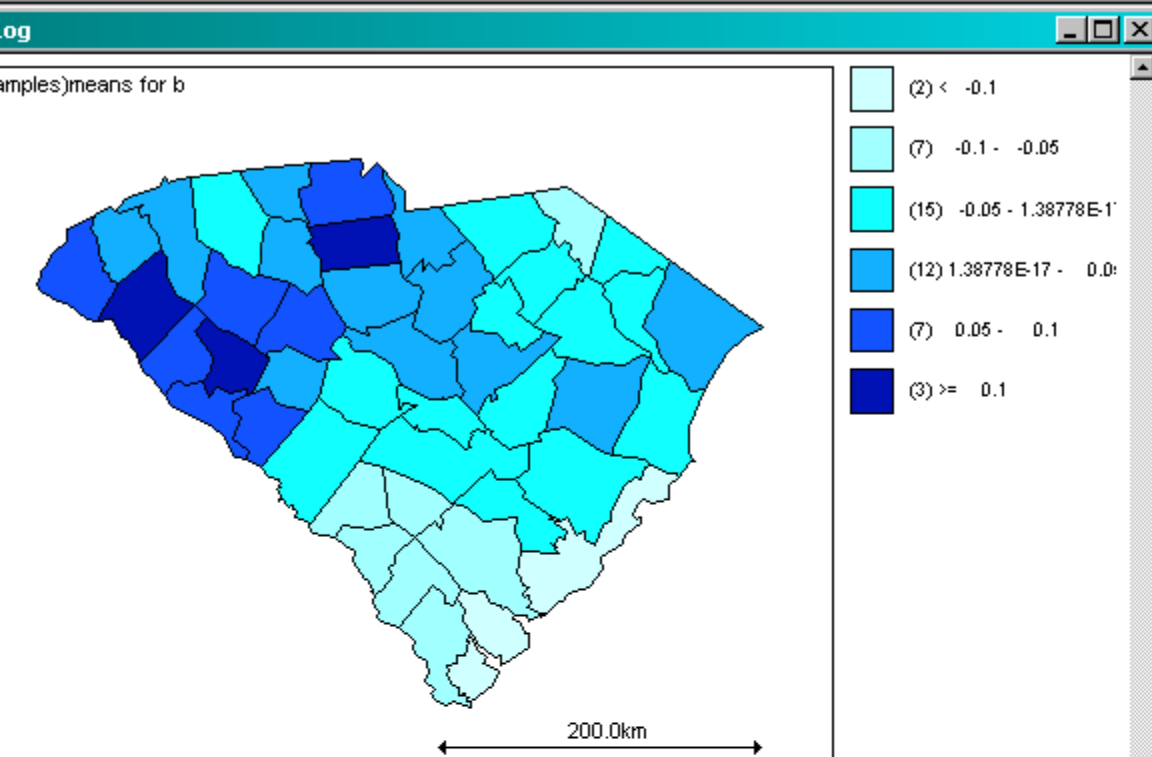
main data is in SCcongen98_data.txt

n data ⇨

```
y=c(0,5,0,16,0,0,3,5,0,7,1,5,1,2,1,3,2,5,2,1,3,2,14,7,1,9,
1,3,1,7,0,1,0,4,2,7,3,14,1,7,7,0,4,10),
(364,1797,177,2145,208,338,1668,2017,168,4403,686,476,550,40
19,907,489,1152,265,335,1867,736,5176,963,311,2324,280,665,8
864,258,3028,83,478,433,523,796,1265,1379,4286,256,3288,1699,369,502,2122),
c(13.6,13.8,32.3,11,24.2,19.9,12.3,13.3,17,15.4,14.1,15.7,18,
3,21.8,20.2,24.9,12,17.4,18.1,18.7,17.5,10.6,13.8,22.8,13.7,
12.6,14,14,26.9,9.4,17.8,23.1,23.1,14.4,10.8,22.1,10.1,13.6,15.9,11.2,18.3,14,26.4,10.6))
```

atial Data ⇨

```
J=46, num = c(5, 5, 4, 5, 5, 4, 3, 6, 5, 4,
5, 6, 7, 5, 4, 4, 4, 6,
5, 6, 5, 3, 2, 7, 5, 7,
3, 5, 4, 7, 2, 9, 3, 6,
6, 7, 5, 4
```

```
= c(
30, 24, 23, 4,
38, 32, 19, 6,
15, 6, 5,
37, 30, 23, 1,
25, 15, 6, 3,
5, 3, 2,
25, 15,
38, 22, 18, 14, 10,
40, 38, 32, 14,
18, 15, 8,
44, 42,
44, 29, 20,
31, 29, 28, 16,
43, 38, 21, 9, 8,
25, 18, 10, 7, 5, 3,
31, 28, 21, 13,
34, 26, 21
```

Log _ □ ✕

amples)means for b

(2) < -0.1

(7) -0.1 - -0.05

(15) -0.05 - 1.38778E-1

(12) 1.38778E-17 - 0.0:

(7) 0.05 - 0.1

(3) >= 0.1

200.0km

amples)means for h
amples)means for h

(1) < -0.1

(24) -0.1 - 2.77556E-17

(18) 2.77556E-17 - 0.1

(2) 0.1 - 0.2

(1) >= 0.2

Sample Monitor Tool

node [▼]   chains [1] to [1]   percentile

beg [1]   end [1000000]   thin [1]

2.5
5
10
25
median
75
90
95
97.5

clear   set   trace   history   density

stats   coda   quantiles   bgr diag   auto cor

Map Tool

map   [SC   ▼]   cuts   pa

variable   [h   ]   num cuts   [default ▼]   bl

quantity   [mean(sample) ▼]   cut1   [-0.09999999]

cut points   ⦿ abs value   ○ percentile   cut2   [2.77555756]

threshold   [   ]   quantile   [   ]   cut3   [0.1]

cut4   [0.2]

beg [1]   end [100000]   cut5   [   ]

cut6   [   ]

plot   get cuts   set cuts

Specification Tool   ✕
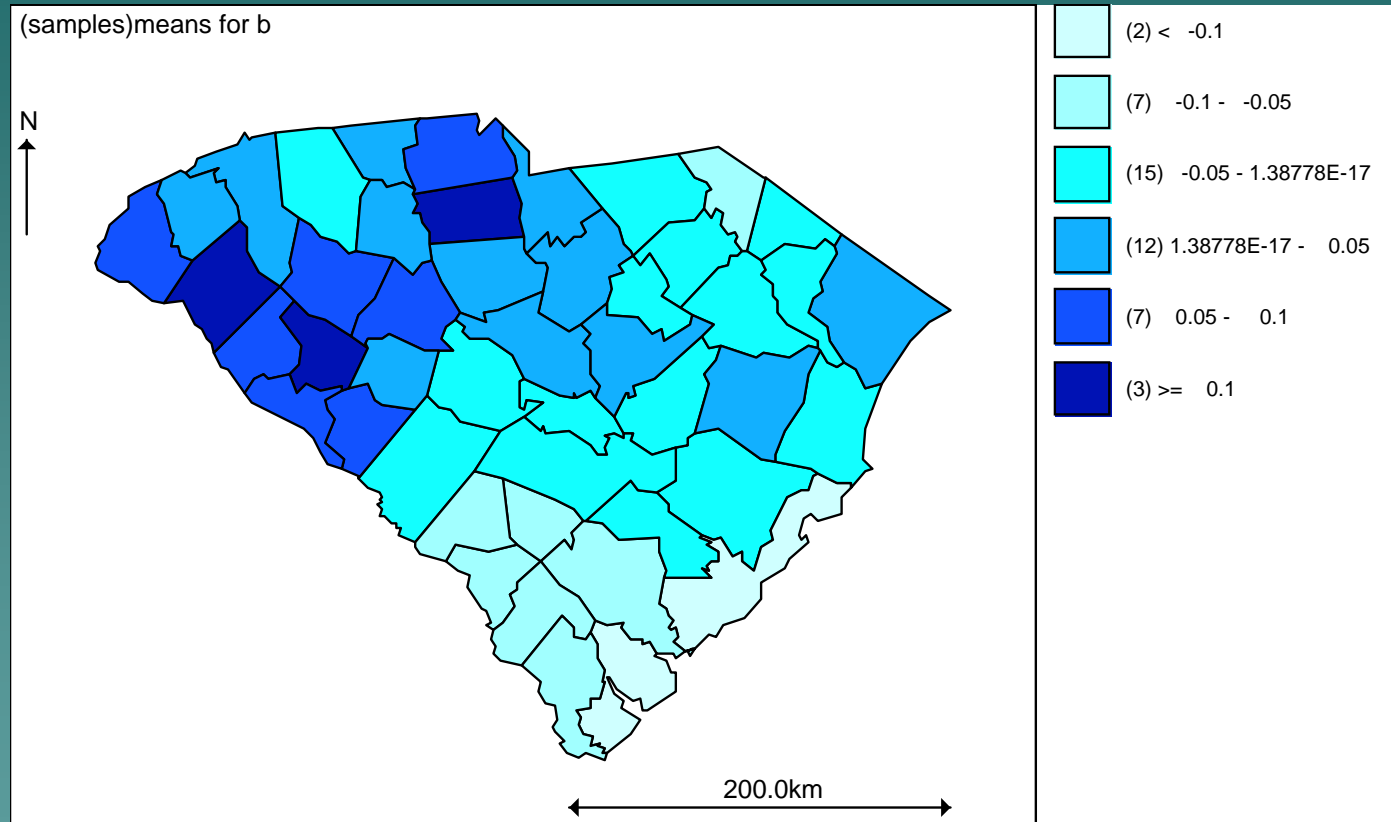
check model   load data

compile   num of chains   [1]
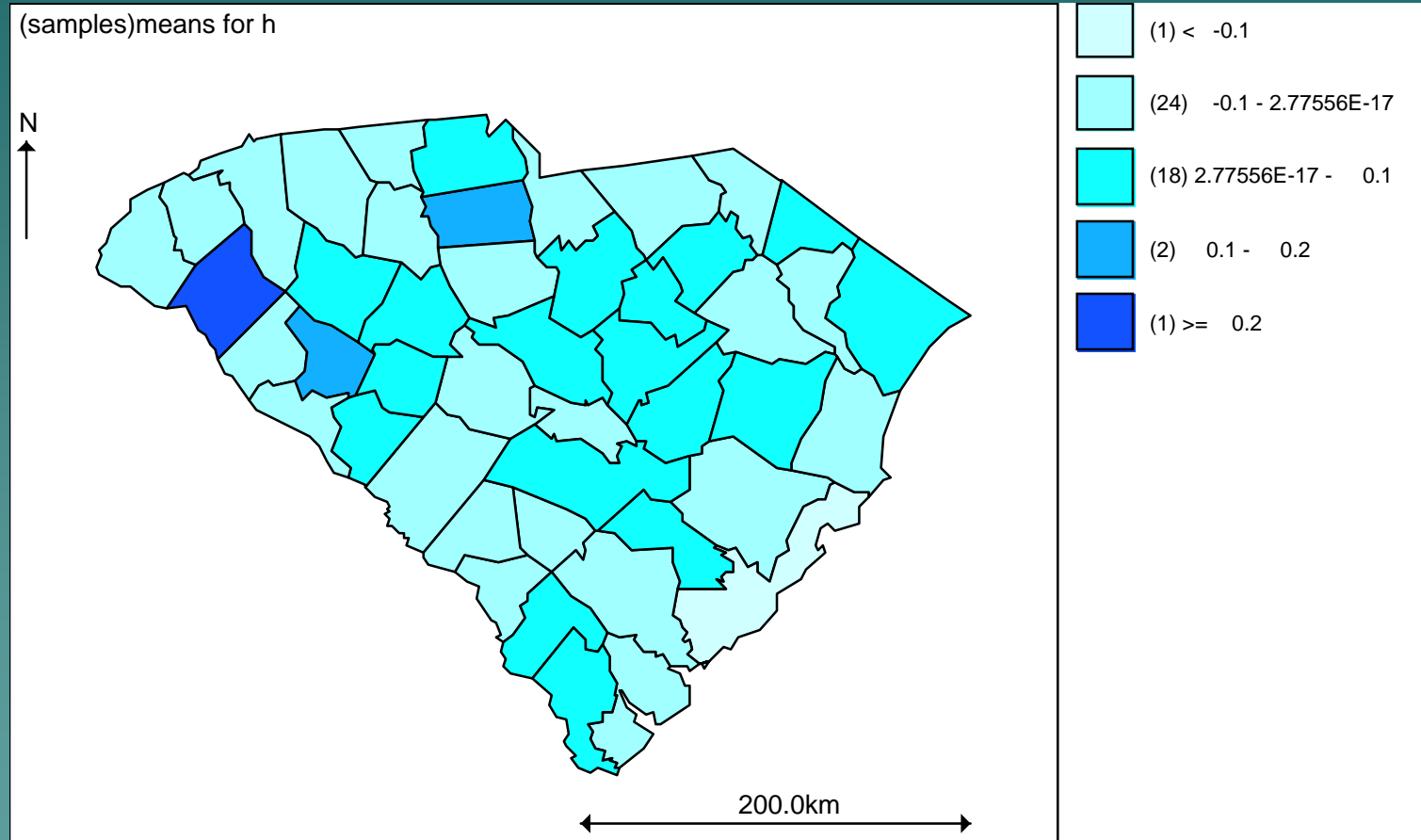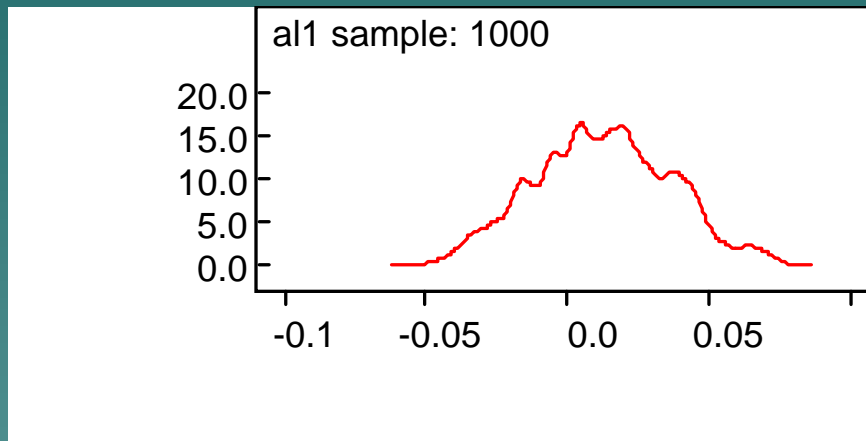
Update Tool   ✕

updates [10000]   refresh [100]   hain   [1] ▲▼

update   thin [1]   iteration [11000]

☐ over relax   ☐ adapting

Start   My Computer   Local Disk (C   79°   8:26 P

CAR (CH) component

(samples)means for b

(2) < -0.1
(7) -0.1 - -0.05
(15) -0.05 - 1.38778E-17
(12) 1.38778E-17 - 0.05
(7) 0.05 - 0.1
(3) >= 0.1

200.0km

UH component

(samples)means for h

N

	(1) < -0.1
	(24) -0.1 - 2.77556E-17
	(18) 2.77556E-17 - 0.1
	(2) 0.1 - 0.2
	(1) >= 0.2

200.0km

Is %poverty significant?

Node statistics

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
al1	0.01258	0.02437	0.003607	-0.03415	0.01262	0.06251	11001	1000

CAR and other models

◆ The CAR model is available on WinBUGS

◆ Other correlation models can also be used and are available: for example a full Multivariate Normal model can be used (similar to Bayesian Kriging). A proper CAR can also be used

Dependencies

◆ adjacency: 2,3 ,4 ,6, 8 neighbors of 1:

◆ hence
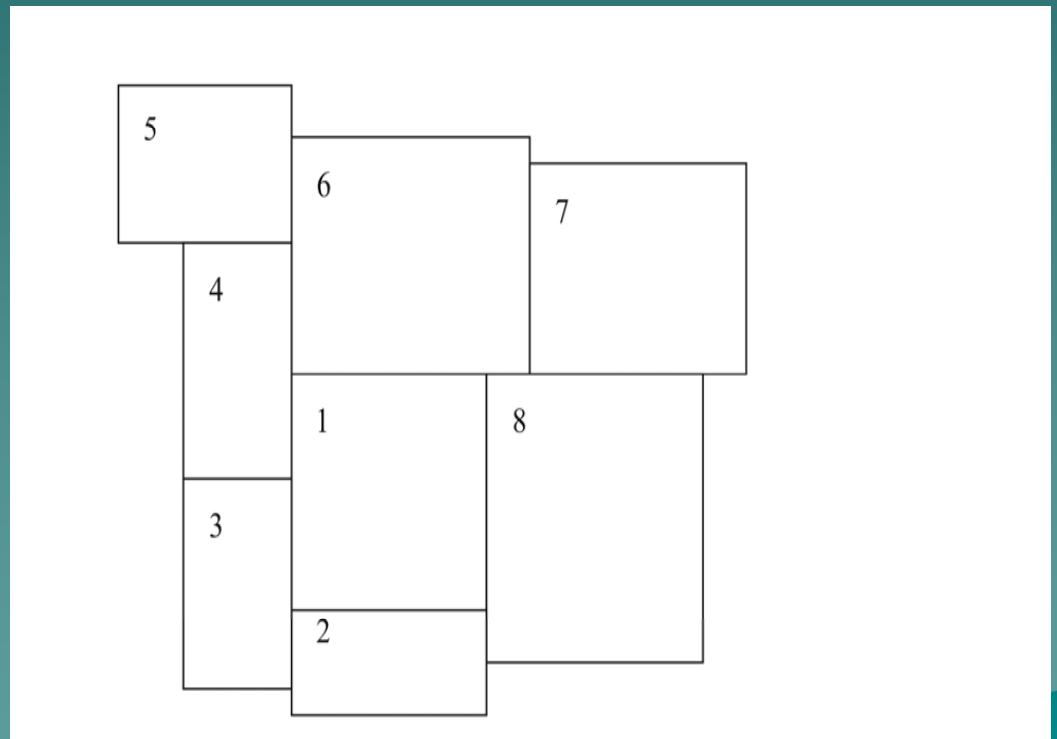
$$\overline{\mu}_{\delta_i} = \sum_{j \in \delta_i} \mu_j / n_{\delta_i}$$

$$\overline{\mu}_{\delta_1} = (\mu_2 + \mu_3 + \mu_4 + \mu_6 + \mu_8)/5$$

Other spatial depedencies

◆ Fully specified covariance models:

$$u \sim N_m(\mathbf{0}, \tau\Sigma)$$
$$\Sigma_{ij} = \text{cov}(u_i, u_j) = \exp(-\alpha d_{ij})$$
$$d_{ij} = \|s_i, s_j\|$$
where s_* is the location of u_*

Comments

◆ Advantage: more flexible spatial form (two parameters instead of one)

◆ Advantage: stationary

◆ Disadvantage: requires inversion of large m x m covariance matrix

MCAR models

◆ The CAR model is now available on WinBUGS for multiple diseases (MCAR models)

◆ Multiple disease analysis can be examine in a variety of ways

MCAR Scottish example
a) cryptorchidism, b) Hypospadia,
c) testicular cancer

Infectious Diseases

- Usually must consider temporal dimension (as well as space)
- Interest usually in spread and rate of transmission
- Simple example: description of spatial cross-section of disease

- Foot and Mouth disease (FMD): NW England 2001

FMD

◆ Foot and mouth disease (FMD) is a viral disease of cloven-footed animals and being extremely contagious it spreads rapidly.

◆ It reduces animal production, and can sometimes be fatal in young stock.

◆ Counts of new cases of FMD were available for parishes within the county of Cumbria, Northern England for the period February 2001-August 2001. The data were available as half monthly counts and so there were 13 time periods available for analysis.