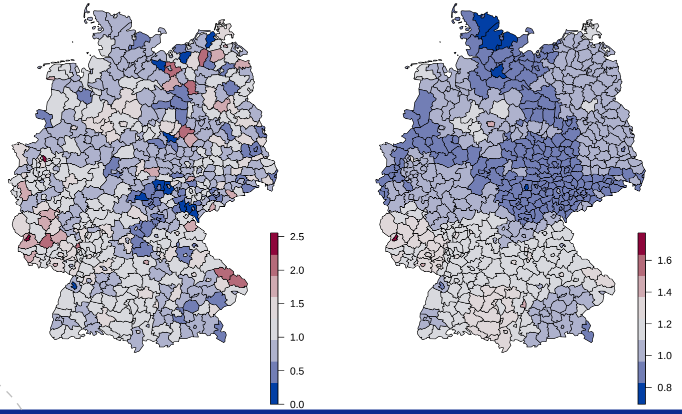
Disease mapping in Germany

We observed larynx cancer mortality counts for males in 544 district of Germany from 1986 to 1990 as well as level of smoking consumption (100 possible values).



The data

y_i: The cancer mortality count at location i.
E_i: An offset; expected number of cases in district i.
c_i: A covariate (level of smoking consumption) at location i
$\boldsymbol{s}_i$: spatial location i (here, district).

Here $i = 1, \ldots, 544$.

The model

$$y_i \mid \eta_i \sim \text{Poisson}(E_i \exp(\eta_i))$$
$$\eta_i = \mu + f(c_i) + f_s(s_i) + f_u(s_i)$$

— $f(c_i)$ is as smooth effect of the covariate $\boldsymbol{c}$

$$\boldsymbol{f} = \{f_1, \ldots, f_{100}\} \sim \text{RW2}(\tau_c)$$

The model

$$y_i \mid \eta_i \sim \text{Poisson}(E_i \exp(\eta_i))$$
$$\eta_i = \mu + f(c_i) + f_s(s_i) + f_u(s_i)$$

— $f(c_i)$ is as smooth effect of the covariate $\boldsymbol{c}$

$$\boldsymbol{f} = \{f_1, \ldots, f_{100}\} \sim \text{RW2}(\tau_c)$$

— $f_s(s_i)$ is a spatial effect modelled as an intrinsic GMRF

$$f_s(s) \mid f_s(s'), s \neq s', \tau_s \sim \mathcal{N}\left(\frac{1}{n_s} \sum_{s \sim s'} f_s(s'), \frac{\tau_s}{n_s}\right)$$

The model

$$y_i \mid \eta_i \sim \text{Poisson}(E_i \exp(\eta_i))$$
$$\eta_i = \mu + f(c_i) + f_s(s_i) + f_u(s_i)$$

— $f(c_i)$ is as smooth effect of the covariate $\boldsymbol{c}$

$$\boldsymbol{f} = \{f_1, \ldots, f_{100}\} \sim \text{RW2}(\tau_c)$$

— $f_s(s_i)$ is a spatial effect modelled as an intrinsic GMRF

$$f_s(s) \mid f_s(s'), s \neq s', \tau_s \sim \mathcal{N}\left(\frac{1}{n_s}\sum_{s\sim s'} f_s(s'), \frac{\tau_s}{n_s}\right)$$

— $f_u(s_i)$ is an unstructured random effect

$$f_s(s) \mid \boldsymbol{f}_u = \{f_u(s_1), \ldots, f_u(s_{544})\} \sim \mathcal{N}(0, \tau_u \mathbf{I})$$

Constraints

For identifiability of the intercept, a sum-to-zero constraint is default for all intrinsic models, so

$$\sum_j f_s(s_j) = 0$$
$$\sum_i f_i = 0.$$

Exercise: Model fitting and comparison

Fit the following models

1. The standard BYM model, i.e. only including μ, f_s, f_u.
2. BYM + linear covariate
3. BYM + non-linear covariate (random walk 2)
4. BYM + space-varying regression (i.e. the effect of the covariate is still linear but depends on the region).

Exercise: Model fitting and comparison

Fit the following models

1. The standard BYM model, i.e. only including μ, f_s, f_u.
2. BYM + linear covariate
3. BYM + non-linear covariate (random walk 2)
4. BYM + space-varying regression (i.e. the effect of the covariate is still linear but depends on the region).

Compare the models using

— The deviance information criterion

— The log-score: $\text{LS} = -\sum_i \log(\text{CPO}_i)$