Models

Kasper Schou Telkamp

Section for Dynamical Systems

DTU Compute
Department of Applied Mathematics and Computer Science

Outline

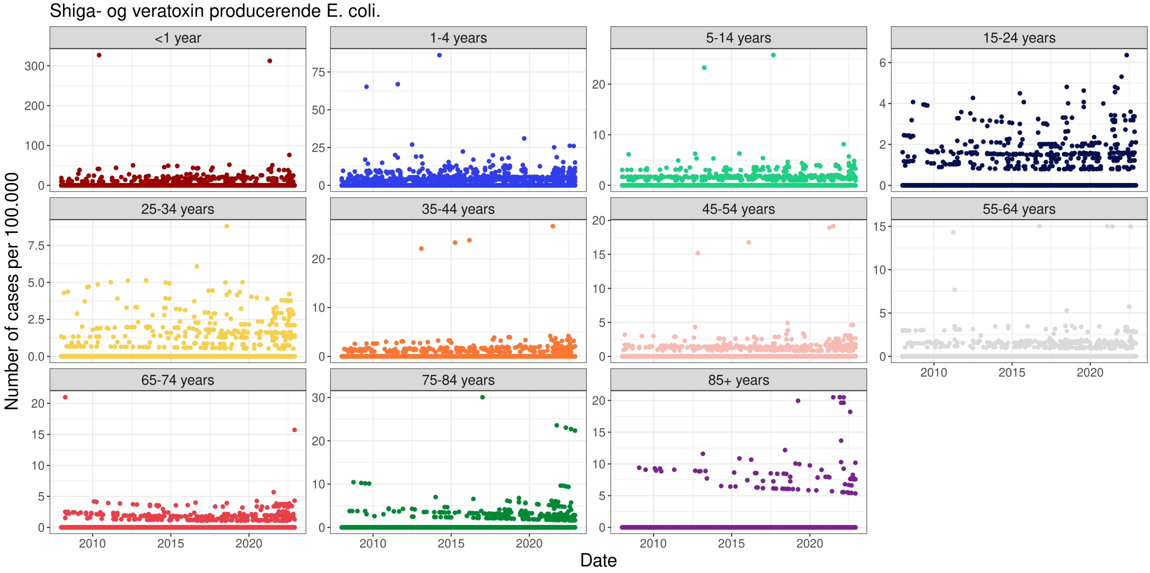
- Data exploration
  - VTEC / STEC
- Hierachical Poisson-Normal model
  - Fomulation
  - Implementation
- Hierachical Poisson-Gamma model
  - Formulation
  - Inference on individual group means

VTEC / STEC

Date	ageGroup	y	n
2008-01-01	<1 year	2	10120
2009-01-01	<1 year	1	10288
2010-01-01	<1 year	1	10654
2011-01-01	<1 year	0	11199
...	...	...	...
2019-12-01	85+ years	1	14153
2020-12-01	85+ years	0	14613
2021-12-01	85+ years	2	14976
2022-12-01	85+ years	3	15203

Data exploration

VTEC / STEC



Fomulation

$$Y_t^a | u_t^a \sim \text{Pois}\big(w_t^a \lambda_a \exp(u_t^a)\big) \tag{1a}$$

$$u_t^a \sim \text{N}(0, \sigma^2) \tag{1b}$$

Implementation - Objective function in C++

```
#include <TMB.hpp>                // Links in the TMB libraries
template<class Type>
Type objective_function<Type>::operator() ()
{
  DATA_VECTOR(y);                         // Data vector transmitted from R
  DATA_VECTOR(w)                  // Data vector transmitted from R
  DATA_FACTOR(ageGroup);          // Data factor transmitted from R

  PARAMETER_VECTOR(u);                // Random effects

  // Parameters
  PARAMETER_VECTOR(lambda);         // Parameter value transmitted from R
  PARAMETER(log_sigma_u);                 // Parameter value transmitted from R

  Type sigma_u = exp(log_sigma_u);

  int nobs = y.size();
  Type mean_ran = Type(0);

  int j;

  Type f = 0;                   // Declare the "objective function" (neg. log. likelihood)
  for(int i=0; i < nobs; i++){
    f -= dnorm(u[i],mean_ran,sigma_u,true);
    j = ageGroup[i];
    f -= dpois(y[i],exp(log(lambda[j])-log(w[j]))*exp(u[i]),true);
  }

  return f;
}
```

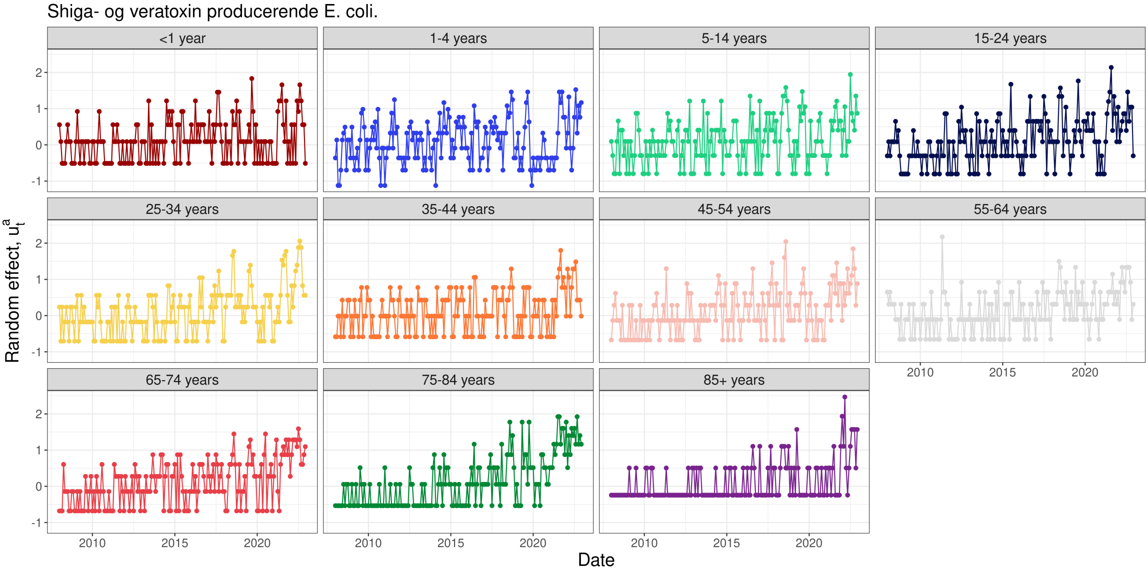
Implementation - Call from R

```r
# Import libraries
library(readr)
library(dplyr)
library(TMB)

# Import the data
dat <- read_rds(file = "../../data/processed/dat.rds")

# Only consider some of the data
y <- dat %>%
  filter(caseDef == "Shiga- og veratoxin producerende E. coli.") %>%
  group_by(Date, ageGroup) %>%
  mutate(y = sum(cases)) %>%
  select(Date, ageGroup, y, n)

compile(file = "PoissonLognormal.cpp")  # Compile the C++ file
dyn.load(dynlib("PoissonLognormal"))    # Dynamically link the C++ code

# Function and derivative
PoisLN <- MakeADFun(
  data = list(y = y$y, ageGroup = y$ageGroup, w = y$n),
  parameters = list(u = rep(1, length(y$y)),
                    lambda = rep(1, nlevels(y$ageGroup)),
                    log_sigma_u = log(1)),
  random = "u",
  DLL = "PoissonLognormal"
)

opt <- nlminb(start = PoisLN$par, PoisLN$fn, PoisLN$gr, lower = c(0.01, 0.01))
```

Results

Parameter	Estimate	Std. Error
$\log(\lambda_{<1year})$	8438.14	276.06
$\log(\lambda_{1-4years})$	34772.54	905.76
$\log(\lambda_{5-14years})$	18464.41	528.24
$\log(\lambda_{15-24years})$	19674.21	584.60
$\log(\lambda_{25-34years})$	15130.53	473.09
$\log(\lambda_{35-44years})$	11256.22	353.22
$\log(\lambda_{45-54years})$	13601.67	408.42
$\log(\lambda_{55-64years})$	13508.84	405.98
$\log(\lambda_{65-74years})$	14753.44	442.70
$\log(\lambda_{75-84years})$	10622.91	345.14
$\log(\lambda_{85+years})$	3586.71	158.77
$\log(\sigma_u)$	0.01	0.01

Results


Shiga- og veratoxin producerende E. coli.
<1 year
1-4 years
5-14 years
15-24 years
25-34 years
35-44 years
45-54 years
55-64 years
65-74 years
75-84 years
85+ years
Random effect, u_t^a
Date
2010
2015
2020
2
1
0
-1

Formulation

$$Y_i | u_i \sim \mathrm{Pois}(\lambda_i u_i) \tag{2a}$$

$$u_i \sim \mathrm{G}(1, \phi) \tag{2b}$$

Probability function for Y

$$
\begin{aligned}
P[Y = y] &= g_Y(y; \lambda, \phi) \\
&= \frac{\phi^y \lambda^y}{(\lambda\phi + 1)^{y+1}}
\end{aligned}
\tag{3}
$$

Proof

The probability function for the conditional distribution of Y for given u

$$f_{Y|u}(y; \lambda, u) = \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \tag{4}$$

and the probability density function for the distribution of u is

$$f_u(u; \phi) = \frac{1}{\phi} \exp(-u/\phi) \tag{5}$$

Proof

Given (4) and (5), the probability function for the marginal distribution of Y is determined from

$$\begin{aligned} g_Y(y;\lambda,\phi) &= \int_{u=0}^{\infty} f_{Y|u}(y;\lambda,u) f_u(u;\phi)\,du \\ &= \int_{u=0}^{\infty} \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \frac{1}{\phi} \exp(-u/\phi)\,du \\ &= \frac{\lambda^y}{y!\phi} \int_{u=0}^{\infty} u^y \exp(-u(\lambda\phi+1)/\phi)\,du \end{aligned} \tag{6}$$

Proof

In (6) it is noted that the integrand is the *kernel* in the probability density function for a Gamma distribution, $G(y+1, \phi/(\lambda\phi+1))$. As the integral of the density shall equal one, we find by adjusting the norming constant that

$$\int_{u=0}^{\infty} u^y \exp\left(-u/\Big(\phi/(\lambda\phi+1)\Big)\right) du = \frac{\phi^{y+1}\Gamma(y+1)}{(\lambda\phi+1)^{y+1}} \tag{7}$$

and then (3) follows

Inference on individual group means

Consider the hierarchical Poisson-Gamma model in (2), and assume that a value $Y = y$ has been observed. Then the conditional distribution of u for given $Y = y$ is a Gamma distribution,

$$u|Y = y \sim G(y + 1, \phi/(\lambda\phi + 1)) \tag{8}$$

with mean

$$\mathrm{E}[u|Y = y] = \frac{y + 1}{} \tag{9}$$

Proof

The conditional distribution is found using Bayes Theorem

$$
\begin{aligned}
g_u(u|Y=y) &= \frac{f_{y,u}(y,u)}{g_Y(y;\lambda,\phi)} \\
&= \frac{f_{y|u}(y;u)g_u(u)}{g_Y(y;\lambda,\phi)} \\
&= \frac{1}{g_Y(y;\lambda,\phi)}\left(\frac{(\lambda u)^y}{y!}\exp(-\lambda u)\frac{1}{\phi}\exp(-u/\phi)\right) \\
&\propto u^y \exp\big(-u(\lambda+1)/\phi\big)
\end{aligned} \tag{10}
$$

Proof

We identify the *kernel* of the probability density function

$$u^y \exp\big(-u(\lambda+1)/\phi\big) \tag{11}$$

as the kernel of a Gamma distribution, $G(y+1, \phi/(\lambda\phi+1))$