Models

Kasper Schou Telkamp

Section for Dynamical Systems

DTU Compute
Department of Applied Mathematics and Computer Science

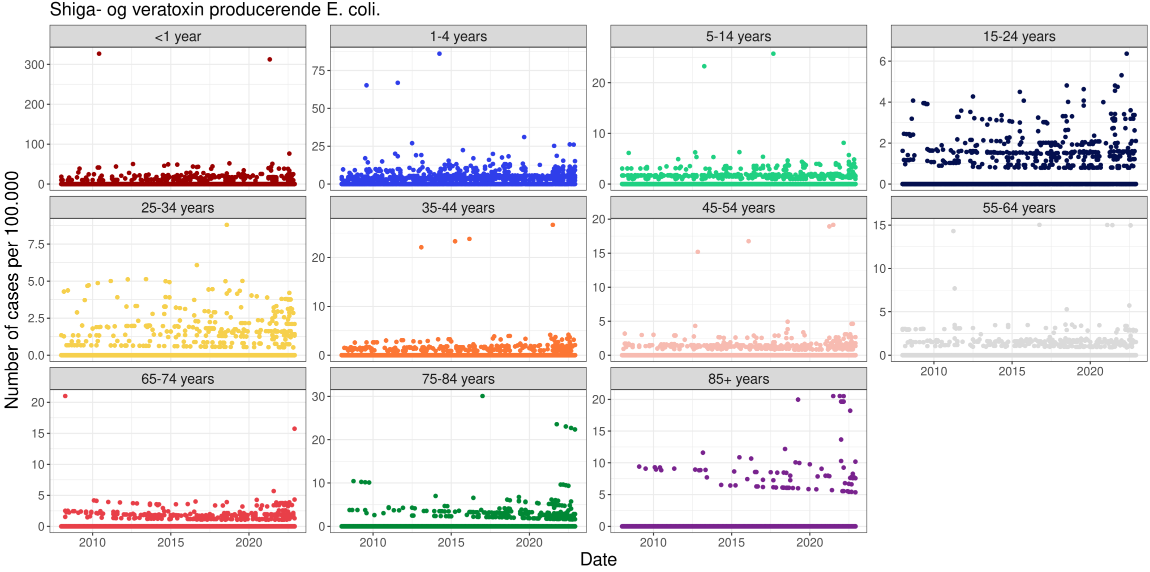
Outline

- Data exploration
  - VTEC / STEC
- Hierachical Poisson-Normal model
  - Fomulation
  - Implementation
- Hierachical Poisson-Gamma model
  - Formulation
  - Probability function for Y
  - 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

VTEC / STEC


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
Number of cases per 100.000
Date
2010
2015
2020

Fomulation

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

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

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;
}
```

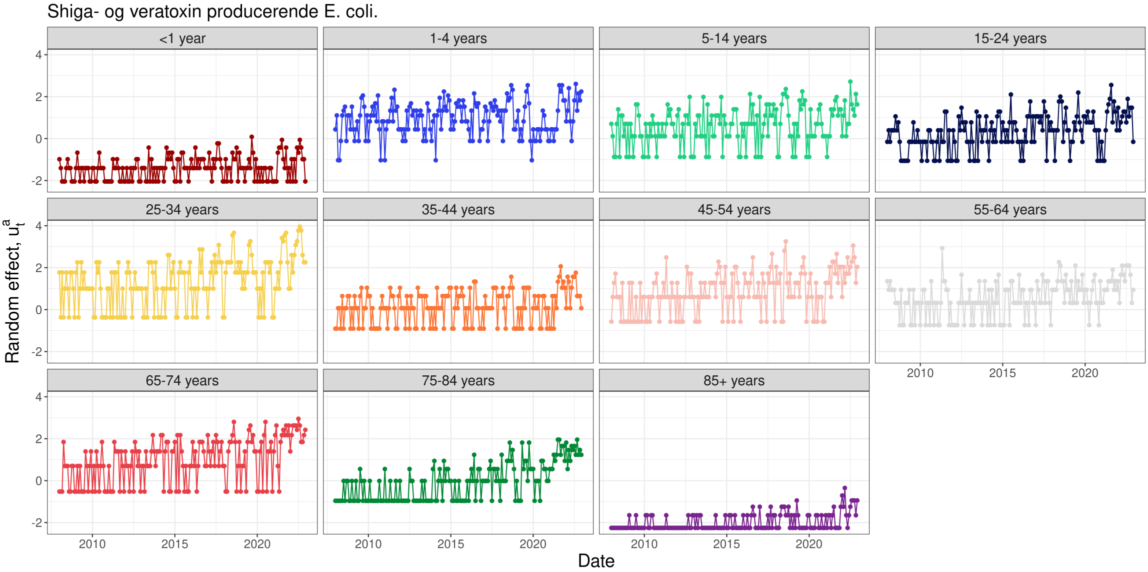
Implementation - Call from R

```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
$\log(\lambda_{<1year})$	410814.56	...
$\log(\lambda_{1-4years})$	305337.95	28427.34
$\log(\lambda_{5-14years})$	578647.21	64256.49
$\log(\lambda_{15-24years})$	782223.95	96931.77
$\log(\lambda_{25-34years})$	148651.67	13670.29
$\log(\lambda_{35-44years})$	738191.24	113778.60
$\log(\lambda_{45-54years})$	305727.09	31939.10
$\log(\lambda_{55-64years})$	457519.57	53292.63
$\log(\lambda_{65-74years})$	166840.39	16968.32
$\log(\lambda_{75-84years})$	280202.19	54641.59
$\log(\lambda_{85+years})$	917515.87	...
$\log(\sigma_u)$	0.45	0.02

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
4
2
0
-2
2010
2015
2020

Formulation

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

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

Probability function for Y

$$
\begin{aligned}
P[Y = y] &= g_Y(y; \lambda, \beta) \\
&= \frac{\lambda^y}{y!\Gamma(1/\beta)\beta^{1/\beta}} \frac{\beta^{y+1/\beta}\Gamma(y+1/\beta)}{(\lambda\beta+1)^{y+1/\beta}} \\
&= \frac{\Gamma(y+1/\beta)}{\Gamma(1/\beta)y!} \frac{1}{(\lambda\beta+1)^{1/\beta}} \left(\frac{\lambda\beta}{\lambda\beta+1}\right)^y \\
&= \binom{y+1/\beta-1}{y} \frac{1}{(\lambda\beta+1)^{1/\beta}} \left(\frac{\lambda\beta}{\lambda\beta+1}\right)^y \; , \; for\ y = 0, 1, 2, \ldots
\end{aligned}
\tag{3}
$$

where we have used the convention

$$
\binom{z}{y} = \frac{\Gamma(z+1)}{\Gamma(z+1-y)y!} \tag{4}
$$

The marginal distribution of Y is a negative binomial distribution, $Y \sim \mathrm{NB}\left(1/\beta, 1/(\lambda\beta+1)\right)$

Proof

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

$$f_{Y|u}(y; \lambda, u) = \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \tag{5}$$

and the probability density function for the distribution of u is

$$f_u(u; \beta) = \frac{1}{\beta \Gamma(1/\beta)} \left(\frac{u}{\beta}\right)^{1/\beta - 1} \exp(-u/\beta) \tag{6}$$

Proof

Given (5) and (6), the probability function for the marginal distribution of Y is determined from

$$\begin{aligned} g_Y(y;\lambda,\beta) &= \int_{u=0}^{\infty} f_{Y|u}(y;\lambda,u) f_u(u;\beta)\, du \\ &= \int_{u=0}^{\infty} \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \frac{1}{\beta\Gamma(1/\beta)} \left(\frac{u}{\beta}\right)^{1/\beta-1} \exp(-u/\beta)\, du \\ &= \frac{\lambda^y}{y!\Gamma(1/\beta)\beta^{1/\beta}} \int_{u=0}^{\infty} u^{y+1/\beta} \exp\big(-u(\lambda\beta+1)/\beta\big)\, du \end{aligned} \tag{7}$$

Proof

In (7) it is noted that the integrand is the *kernel* in the probability density function for a Gamma distribution, $\mathrm{G}\left(y + 1/\beta, \beta/(\lambda\beta + 1)\right)$. As the integral of the density shall equal one, we find by adjusting the norming constant that

$$\int_{u=0}^{\infty} u^{y+1/\beta} \exp\left(- u/\Big(\beta/(\lambda\beta + 1)\Big)\right) du = \frac{\beta^{y+1/\beta}\Gamma(y + 1/\beta)}{(\lambda\beta + 1)^{y+1/\beta}} \tag{8}$$

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,

$$u|Y = y \sim \mathrm{G}\left(y + 1/\beta, \beta/(\lambda\beta + 1)\right) \tag{9}$$

with mean

$$\mathrm{E}[u|Y = y] = \frac{y\beta + 1}{\lambda\beta + 1} \tag{10}$$

Proof

The conditional distribution is found using Bayes Theorem

$$
\begin{aligned}
g_u(u|Y=y) &= \frac{f_{y,u}(y,u)}{g_Y(y;\lambda,\phi)} \\
&= \frac{f_{y|u}(y;u)g_u(u)}{g_Y(y;\lambda,\beta)} \\
&= \frac{1}{g_Y(y;\lambda,\beta)}\left(\frac{(\lambda u)^y}{y!}\exp(-\lambda u)\frac{1}{\beta\Gamma(1/\beta)}\left(\frac{u}{\beta}\right)^{1/\beta-1}\exp(-u/\beta)\right) \\
&\propto u^{y+1/\beta-1}\exp\big(-u(\lambda\beta+1)/\beta\big)
\end{aligned} \tag{11}
$$

We identify the *kernel* of the probability density function

$$
u^y\exp\big(-u(\lambda+1)/\phi\big) \tag{12}
$$

as the kernel of a Gamma distribution, $\mathrm{G}(y+1/\beta,\beta/(\lambda\beta+1))$