

[illegible]

Outline

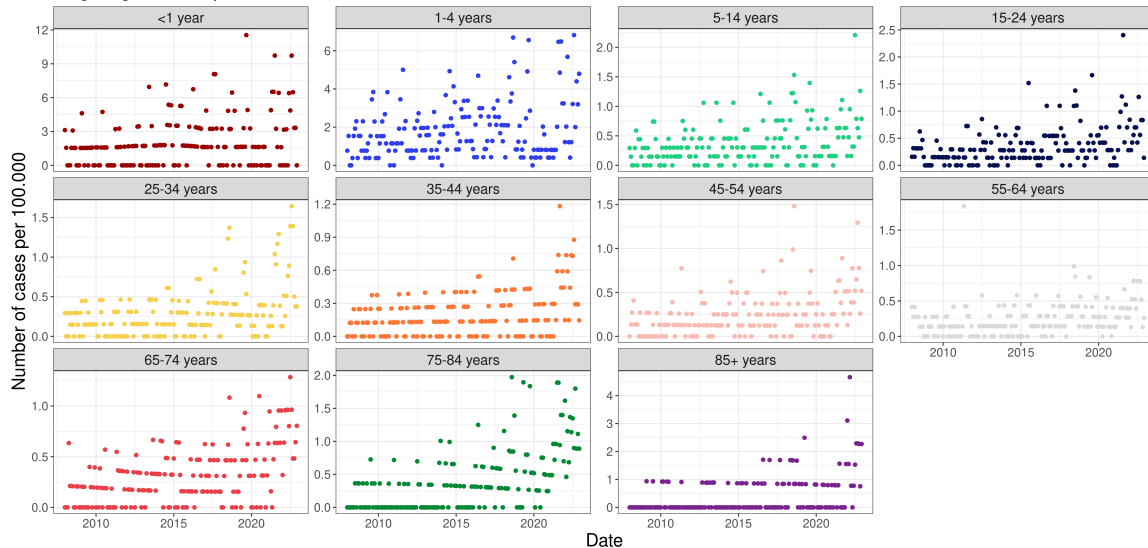
- Data exploration
 - VTEC / STEC
 - Formulation
 - Implementation
 - Results
- Hierarchical Poisson-Gamma model
 - Formulation
 - Probability function for Y
 - Inference on individual group means
 - Implementation
 - Results

Date	ageGroup	y_{it}	x_{it}
2008-01-01	<1 year	2	64137
2008-01-01	1-4 years	2	259910
2008-01-01	5-14 years	2	680529
2008-01-01	15-24 years	1	631724
...
2022-12-01	55-64 years	2	773073
2022-12-01	65-74 years	5	621965
2022-12-01	75-84 years	4	449423
2022-12-01	85+ years	3	132181

Data exploration

VTEC / STEC

Shiga- og veratoxin producerende E. coli.



The count observations are assumed to follow a Poisson distribution with intensities λ_{it} . Also, we shall assume that the count is proportional to the population size, x_{it} , within each age group, i , at a given time point, t . Hence, in terms of the canonical link for the Poisson distribution the model is

$$\log(\lambda_{it}) = \mathbf{X}_i^T \beta_{it} + \log(x_{it}) + u_{it} \quad (1)$$

Here \mathbf{X}_i is $T \times 11$ -dimensional, and β_{it} contains the corresponding fixed effect parameter. The random effects u_{it} are assumed to be Gaussian.

$$u_{it} = \epsilon_{it} \quad (2)$$

where $\epsilon_{it} \sim N(0, \sigma^2)$ is a white noise process, and σ is a model parameter.

Henceforth, the model can be formulated as a two-level hierarchical model

$$Y_{it}|u_{it} \sim \text{Pois}(\lambda_{it}) \quad (3a)$$

$$u_{it} \sim \text{N}(0, \sigma^2) \quad (3b)$$

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(x);          // Data vector transmitted from R
    DATA_MATRIX(X);          // Design matrix transmitted from R

    PARAMETER_VECTOR(u);      // Random effects

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

    vector<Type> lambda = exp(X*beta-log(x)+u);
    Type sigma_u = exp(log_sigma_u);

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

    int i;

    Type f = 0;               // Declare the "objective function"
    for(int t=0; t < nobs; t++){
        f -= dnorm(u[t],mean_ran,sigma_u,true);
        f -= dpois(y[t],lambda[t],true);
    }

    return f;
}
7

```

Implementation - Call from 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) %>% reframe(y = sum(cases), n = sum(n))

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

# Make months into integers
y.design <- y %>% mutate(Month = as.integer(format(Date, "%m")))

# Construct the design matrix
designMatrix <- model.matrix(y ~ -1 + ageGroup, data = y.design)

# Construct initial parameters
beta <- rep(1, nlevels(y$ageGroup))

# Function and derivative
PoisLN <- MakeADFun(
  data = list(y = y$y, x = y$n, X = designMatrix),
  parameters = list(u = rep(1, length(y$y)), beta = beta, log_sigma_u = log(1)),
  random = "u", DLL = "PoissonNormal")

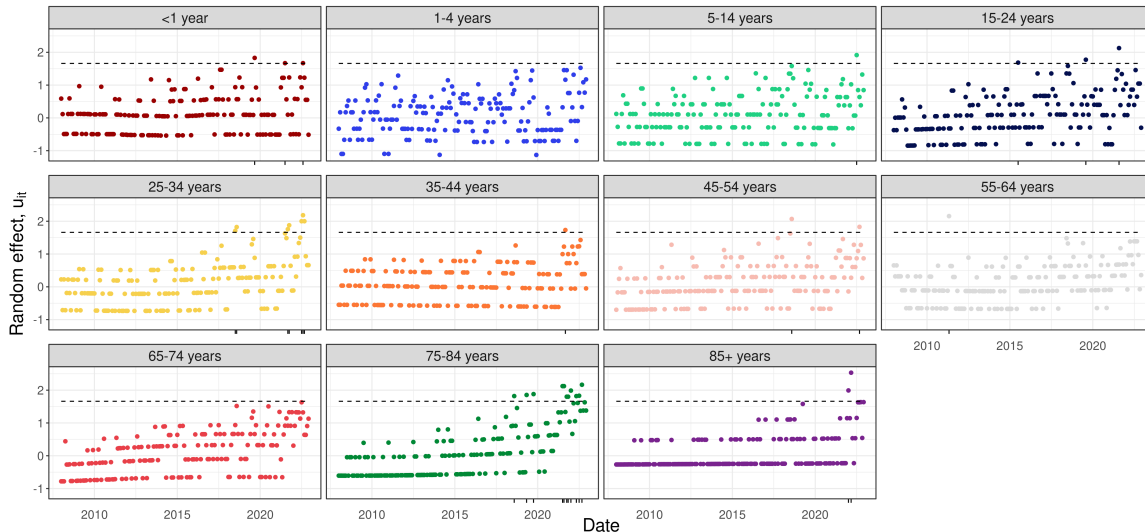
opt <- nlminb(start = PoisLN$par, PoisLN$fn, PoisLN$gr, lower = c(0.01, 0.01))
```


Parameter	Estimate	Std. Error
$\beta_{<1year}$	10.83	0.11
$\beta_{1-4years}$	13.64	0.09
$\beta_{5-14years}$	13.94	0.09
$\beta_{15-24years}$	14.04	0.09
$\beta_{25-34years}$	13.78	0.10
$\beta_{35-44years}$	13.53	0.10
$\beta_{45-54years}$	13.84	0.10
$\beta_{55-64years}$	13.70	0.10
$\beta_{65-74years}$	13.57	0.10
$\beta_{75-84years}$	12.61	0.11
$\beta_{85+years}$	10.49	0.14
σ	1.01	0.03

Results

Shiga- og veratoxin producerende E. coli.

Hierarchical Poisson Normal model



Likewise, in the compound Poisson-Gamma model the count observations are assumed to follow a Poisson distribution, but this time the intensities, λ_{it} , the model are defined as

$$\log(\lambda_{it}) = \mathbf{X}_i^T \beta_{it} + \log(x_{it}) + \log(u_{it}) \quad (4)$$

Here \mathbf{X}_i is $T \times 11$ -dimensional, and β_{it} contains the corresponding fixed effect parameter. Additionally, the random effects u_{it} are assumed to be Gamma distributed.

Subsequently, the model can be formulated as a two-level hierarchical model

$$Y_{it}|u_{it} \sim \text{Pois}(\lambda_{it}) \quad (5a)$$

$$u_{it} \sim G(1/\phi_i, \phi_i) \quad (5b)$$

$$\begin{aligned} P[Y = y_i] &= g_Y(y; \lambda, \phi) \\ &= \frac{\lambda^y}{y! \Gamma(1/\phi) \phi^{1/\phi}} \frac{\phi^{y+1/\phi} \Gamma(y + 1/\phi)}{(\lambda\phi + 1)^{y+1/\phi}} \\ &= \frac{\Gamma(y + 1/\phi)}{\Gamma(1/\phi) y!} \frac{1}{(\lambda\phi + 1)^{1/\phi}} \left(\frac{\lambda\phi}{\lambda\phi + 1} \right)^y \\ &= \binom{y + 1/\phi - 1}{y} \frac{1}{(\lambda\phi + 1)^{1/\phi}} \left(\frac{\lambda\phi}{\lambda\phi + 1} \right)^y, \text{ for } y = 0, 1, 2, \dots \end{aligned} \tag{6}$$

where we have used the convention

$$\binom{z}{y} = \frac{\Gamma(z + 1)}{\Gamma(z + 1 - y) y!} \tag{7}$$

The marginal distribution of Y is a negative binomial distribution, $Y \sim \text{NB}(1/\phi, 1/(\lambda\phi + 1))$

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

$$f_{Y|u}(y; \lambda, u) = \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \quad (8)$$

and the probability density function for the distribution of u is

$$f_u(u; \phi) = \frac{1}{\phi \Gamma(1/\phi)} \left(\frac{u}{\phi} \right)^{1/\phi-1} \exp(-u/\phi) \quad (9)$$

Given (8) and (9), the probability function for the marginal distribution of Y is determined from

$$\begin{aligned} g_Y(y; \lambda, \phi) &= \int_{u=0}^{\infty} f_{Y|u}(y; \lambda, u) f_u(u; \phi) du \\ &= \int_{u=0}^{\infty} \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \frac{1}{\phi \Gamma(1/\phi)} \left(\frac{u}{\phi}\right)^{1/\phi-1} \exp(-u/\phi) du \\ &= \frac{\lambda^y}{y! \Gamma(1/\phi) \phi^{1/\phi}} \int_{u=0}^{\infty} u^{y+1/\phi} \exp(-u(\lambda\phi + 1)/\phi) du \end{aligned} \quad (10)$$

In (10) it is noted that the integrand is the *kernel* in the probability density function for a Gamma distribution, $G(y + 1/\phi, \phi/(\lambda\phi + 1))$. As the integral of the density shall equal one, we find by adjusting the norming constant that

$$\int_{u=0}^{\infty} u^{y+1/\phi} \exp\left(-u/\left(\phi/(\lambda\phi + 1)\right)\right) du = \frac{\phi^{y+1/\phi} \Gamma(y + 1/\phi)}{(\lambda\phi + 1)^{y+1/\phi}} \quad (11)$$

and then (6) follows

Inference on individual group means

Consider the hierarchical Poisson-Gamma model in (5), and assume that a value $Y = y$ has been observed. Then the conditional distribution of u for given $Y = y$ is a Gamma distribution,

$$u|Y = y \sim G(y + 1/\phi, \phi/(\lambda\phi + 1)) \quad (12)$$

with mean

$$E[u|Y = y] = \frac{y\phi + 1}{\lambda\phi + 1} \quad (13)$$

and variance

$$V[u|Y = y] = \frac{(y\phi^2 + \phi)}{(\lambda\phi + 1)^2} \quad (14)$$

The conditional distribution is found using Bayes Theorem

$$\begin{aligned}
 g_u(u|Y=y) &= \frac{f_{y,u}(y,u)}{g_Y(y;\lambda,\phi)} \\
 &= \frac{f_{y|u}(y;u)g_u(u)}{g_Y(y;\lambda,\phi)} \\
 &= \frac{1}{g_Y(y;\lambda,\phi)} \left(\frac{(\lambda u)^y}{y!} \exp(-\lambda u) \frac{1}{\phi \Gamma(1/\phi)} \left(\frac{u}{\phi} \right)^{1/\phi-1} \exp(-u/\phi) \right) \\
 &\propto u^{y+1/\phi-1} \exp(-u(\lambda\phi+1)/\phi)
 \end{aligned} \tag{15}$$

We identify the *kernel* of the probability density function

$$u^{y+1/\phi-1} \exp(-u(\lambda\phi+1)/\phi) \tag{16}$$

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

Implementation - Define negative likelihood function in R

```
# Define the negative likelihood function for the marginal distribution of Y
nll.age <- function(theta, data){
  # Extract counts
  y <- data$y
  # Extract agegroups
  ageGroup <- data$ageGroup
  # Extract number of agegroups
  n.ageGroup <- n_distinct(data$ageGroup)

  # Define parameters
  lambda <- theta[1:n.ageGroup]
  beta <- theta[(n.ageGroup+1):(n.ageGroup*2)]

  # Construct the size and probability for the negative binomial distribution
  r <- 1/beta
  p <- 1/(lambda*beta+1)

  # Initialize the log-likelihood
  ll <- 0
  for(i in 1:nrow(data)){
    ll = ll + dnbinom(x = y[i],
                      size = r[ageGroup[i]],
                      prob = p[ageGroup[i]],
                      log = TRUE)
  }

  # Return the negative log-likelihood
  -ll
}
```

Hierarchical Poisson-Gamma model

Results

ageGroup	λ	β
<1 year	1.24	0.33
1-4 years	4.57	0.36
5-14 years	2.47	0.42
15-24 years	2.57	0.48
25-34 years	2.08	0.57
35-44 years	1.47	0.20
45-54 years	1.88	0.39
55-64 years	1.77	0.25
65-74 years	1.89	0.40
75-84 years	1.45	1.08
85+ years	0.50	0.97

Results

Shiga- og veratoxin producerende E. coli.

Compound Poisson Gamma model

