Zero-inflated models: Poisson and Binomial

Parametrisation

There is support two types of zero-inflated models, which we name type 0 and type 1. These are defined for both the Binomial, the Poisson and the negative Binomial likelihood. For simplicity we will describe only the Poisson as the other two cases are similar.

Type 0

The (type 0) likelihood is defined as

$$\text{Prob}(y \mid \ldots) = p \times 1_{[y=0]} + (1-p) \times \text{Poisson}(y \mid y > 0)$$

where p is a hyperparameter where

$$p = \frac{\exp(\theta)}{1 + \exp(\theta)}$$

and θ is the internal representation of p; meaning that the initial value and prior is given for θ. This is model is called `zeroinflatedpoisson0` (and `zeroinflatedbinomial0`).

Type 1

The (type 1) likelihood is defined as

$$\text{Prob}(y \mid \ldots) = p \times 1_{[y=0]} + (1-p) \times \text{Poisson}(y)$$

where p is a hyperparameter where

$$p = \frac{\exp(\theta)}{1 + \exp(\theta)}$$

and θ is the internal representation of p; meaning that the initial value and prior is given for θ. This is model is called `zeroinflatedpoisson1` (and `zeroinflatedbinomial1`).

Link-function

As for the Poisson, the Binomial and the negative Binomial.

Hyperparameters

For Poisson and the Binomial, there is one hyperparameter; where

$$p = \frac{\exp(\theta)}{1 + \exp(\theta)}$$

and the prior and initial value is is given for θ.

For the negative Binomial, there are two hyperparameters. The overdispersion parameter n is represented as

$$\theta_1 = \log(n)$$

and the prior is defined on θ_1. The zero-inflation parameter p, is represented as

$$p = \frac{\exp(\theta_2)}{1 + \exp(\theta_2)}$$

and the prior and initial value is is given for θ_2.

Specification

- family = `zeroinflatedbinomial0`
- family = `zeroinflatedbinomial1`
- family = `zeroinflatednbinomial0`
- family = `zeroinflatednbinomial1`
- family = `zeroinflatedpoisson0`
- family = `zeroinflatedpoisson1`
- Required arguments: As for the Binomial, the negative Binomial and Poisson likelihood.

Hyperparameter spesification and default values

Zeroinflated Binomial Type 0

hyper

theta

name probability
short.name p
initial 0
fixed FALSE
prior gaussian
param c(0, 1)

survival FALSE

discrete FALSE

Zeroinflated Binomial Type 1

hyper

theta

name probability
short.name p
initial 0
fixed FALSE
prior gaussian
param c(0, 1)

survival FALSE

discrete FALSE

Zeroinflated NegBinomial Type 0

hyper

theta1

- **name** size
- **short.name** size
- **initial** 2.30258509299405
- **fixed** FALSE
- **prior** loggamma
- **param** c(1, 0.01)

theta2

- **name** probability
- **short.name** p
- **initial** -1
- **fixed** FALSE
- **prior** gaussian
- **param** c(0, 1)

survival FALSE

discrete FALSE

Zeroinflated NegBinomial Type 1

hyper

theta1

- **name** size
- **short.name** size
- **initial** 2.30258509299405
- **fixed** FALSE
- **prior** loggamma
- **param** c(1, 0.01)

theta2

- **name** probability
- **short.name** p
- **initial** -1
- **fixed** FALSE
- **prior** gaussian
- **param** c(0, 1)

survival FALSE

discrete FALSE

Zeroinflated Poisson Type 0

hyper

theta

name probability
short.name p
initial 0
fixed FALSE
prior gaussian
param c(0, 1)

survival FALSE

discrete FALSE

Zeroinflated Poisson Type 1

hyper

theta

name probability
short.name p
initial 0
fixed FALSE
prior gaussian
param c(0, 1)

survival FALSE

discrete FALSE

Example

In the following example we estimate the parameters in a simulated example for both type 0 and type 1.

Poisson

```
## type 0
n=100
a = 1
b = 1
z = rnorm(n)
eta = a + b*z
p = 0.2
E = sample(c(1,5,10,15), size=n, replace=TRUE)
lambda = E*exp(eta)

## first sample y|y>0
y = rpois(n, lambda = lambda)
is.zero = (y == 0)
```

```
while(sum(is.zero) > 0)
{
    y[is.zero] = rpois(sum(is.zero), lambda[is.zero])
    is.zero = (y == 0)
}
## then set some of these to zero
y[ rbinom(n, size=1, prob=p) == 1 ] = 0

data = list(y=y,z=z)
formula = y ~ 1+z
result0 = inla(formula, family = "zeroinflatedpoisson0", data = data, E=E)
summary(result0)

## type 1
y = rpois(n, lambda = lambda)
y[ rbinom(n, size=1, prob=p) == 1 ] = 0
data = list(y=y,z=z)
formula = y ~ 1+z
result1 = inla(formula, family = "zeroinflatedpoisson1", data = data, E=E)
summary(result1)
```

Binomial

```
## type 0
n=100
a = 1
b = 1
z = rnorm(n)
eta = a + b*z
p = 0.2
Ntrials = sample(c(1,5,10,15), size=n, replace=TRUE)
prob = exp(eta)/(1 + exp(eta))

y = rbinom(n, size = Ntrials, prob = prob)
is.zero = (y == 0)
while(sum(is.zero) > 0)
{
    y[is.zero] = rbinom(sum(is.zero), size = Ntrials[is.zero], prob = prob[is.zero])
    is.zero = (y == 0)
}
y[ rbinom(n, size=1, prob=p) == 1 ] = 0
data = list(y=y,z=z)
formula = y ~ 1+z
result0 = inla(formula, family = "zeroinflatedbinomial0", data = data, Ntrials = Ntrials)
summary(result0)

## type 1
y = rbinom(n, size = Ntrials, prob = prob)
y[ rbinom(n, size=1, prob=p) == 1 ] = 0
data = list(y=y,z=z)
formula = y ~ 1+z
```

```
result1 = inla(formula, family = "zeroinflatedbinomial1", data = data, Ntrials=Ntrials)
summary(result1)
```

Notes

None.

Extentions

There are some extentions available which currently is only implemented for the cases where its needed/requested.

Type 2 Is like Type 1 but where (for the Poisson)

$$p = 1 - \left(\frac{E \exp(x)}{1 + E \exp(x)} \right)^{\alpha}$$

where $\alpha > 0$ is the hyperparameter instead of p (and $E \exp(x)$ is the mean). Available as `zeroinflatedpoisson2` and for negative binomial as `zeroinflatednbinomial2`.

The internal representation is $\theta = \log(\alpha)$ and prior is defined on $\log(\alpha)$.

Zeroinflated Poisson Type 2

hyper

theta

name probability
short.name p
initial 0
fixed FALSE
prior gaussian
param c(0, 1)

survival FALSE

discrete FALSE

Zeroinflated Negative Binomial Type 2

hyper

theta1

name size
short.name size
initial 2.30258509299405
fixed FALSE
prior loggamma
param c(1, 0.01)

theta2

name alpha
short.name a
initial 0.693147180559945
fixed FALSE
prior gaussian
param c(0, 1)

survival FALSE

discrete FALSE