

# Negative Binomial

## Parametrisation

The negative Binomial distribution is

$$\text{Prob}(y) = \frac{\Gamma(y+n)}{\Gamma(n)\Gamma(y+1)} p^n (1-p)^y$$

for responses  $y = 0, 1, 2, \dots$ , where

$n$ : number of successful trials (*size*), or dispersion parameter. Must be strictly positive, need not be integer.

$p$ : probability of success in each trial.

## Link-function

The mean and variance of  $y$  are given as

$$\mu = n \frac{1-p}{p} \quad \text{and} \quad \sigma^2 = \mu \left(1 + \frac{\mu}{n}\right)$$

and the mean is linked to the linear predictor by

$$\mu = E \exp(\eta)$$

where the hyperparameter  $n$  (*size*) plays the role of an overdispersion parameter.  $E$  represents known constant and  $\log(E)$  is the offset of  $\eta$ .

## Hyperparameters

The overdispersion parameter  $n$  (*size*) is represented as

$$\theta = \log(n)$$

and the prior is defined on  $\theta$ .

## Specification

- family = `nbinomial`
- Required arguments:  $y$  and  $E$  (default  $E = 1$ ).

## Hyperparameter specification and default values

hyper

theta

```
name    size
short.name  size
initial 2.30258509299405
fixed   FALSE
prior   loggamma
param   1 1
to.theta
```

```

      from.theta

survival FALSE

discrete TRUE

link default log

pdf nbinomial

```

## Example

In the following example we estimate the parameters in a simulated example with negative binomial responses and assign the hyperparameter  $\theta$  a Gaussian prior with mean 0 and precision 0.01

```

n=100
a = 1
b = 1
E = rep(1,n)
z = rnorm(n)
eta = a + b*z
mu = E*exp(eta)
size = 15
prob = size/(size + mu)
y = rnbinom(n, size=size, prob = prob)

data = list(y=y,z=z)
formula = y ~ 1+z
result = inla(formula, family = "nbinomial", data = data, E=E,
              control.data = list(hyper = list(
                                theta = list(
                                  prior="gaussian",
                                  param = c(0,0.01))))))

summary(result)

```

## Notes

As  $n \rightarrow \infty$ , the negative Binomial converges to the Poisson distribution. For numerical reasons, if  $n$  is too large:

$$\frac{\mu}{n} < 10^{-4},$$

then the Poisson limit is used.