

# CBinomial

## Parametrisation

The clustered/clumped-Binomial distribution arrives from a transformation of Binomial observations. Let  $z$  be Binomial distributed

$$\text{Prob}(z) = \binom{n}{z} p^z (1-p)^{n-z}$$

for  $z = 0, 1, 2, \dots, n$ , where

$n$ : number of trials.

$p$ : probability of success in each trial.

Then binary CBinomial distribution is the distribution for  $y$ , where

$$y = \begin{cases} 0 & z = 0 \\ 1 & z > 0 \end{cases}$$

It then follows that  $\text{Prob}(y = 0) = (1-p)^n$  and  $\text{Prob}(y = 1) = 1 - (1-p)^n$ , i.e.  $y$  is Binomial distributed with size 1 and probability for success  $1 - (1-p)^n$ . In general we have  $k$  independent experiments,  $y_1, \dots, y_k$ , and let  $w = y_1 + \dots + y_k$ . Then  $w$  is CBinomial( $k, n, p$ ) distributed, i.e.  $w$  is Binomial( $k, 1 - (1-p)^n$ ).

## Link-function

The probability  $p$  is by default linked to the linear predictor by

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

but other choices are also available.

## Hyperparameters

None.

## Hyperparameter specification and default values

**hyper**

**survival** FALSE

**discrete** TRUE

**link** default logit cauchit probit cloglog loglog

**status** experimental

**pdf** cbinomial

## Specification

- family = **cbinomial**
- Required arguments: the response  $w$  and the parameters  $k$  and  $n$  (keyword **Ntrials**, where the argument is a **two-column matrix**: **Ntrials** = **cbind(k,n)**)

## Example

In the following example we estimate the parameters in a simulated example with CBinomial responses.

```
N=1000
a = -1
b = 1
z = rnorm(N)
eta = a + b*z
n = sample(c(1,5,10,15), size=N, replace=TRUE)
p = exp(eta)/(1 + exp(eta))
prob = 1.0 - (1-p)^n
k = sample(c(1,5,10,15), size=N, replace=TRUE)
y = rbinom(N, size=k, prob = prob)

data = list(y=y,z=z)
formula = y ~ 1+z
result = inla(formula, family = "cbinomial", data = data,
              Ntrials=cbind(k, n), verbose=TRUE)
summary(result)
```

## Notes

None.