Generalised Linear Modelling 1 – some options

Starting point:

- Clarity about what is your response variable and what are explanatory variables
- Scientific questions about the relationship between response and explanatory variables
- Data in a long-form dataframe or tibble

Generalised Linear Modelling 1 – some options

Starting point:

- Clarity about what is your response variable and what are explanatory variables
- Clarity about how you expect your response variable to be distributed
- Scientific questions about the relationship between response and explanatory variables
- Data in a long-form dataframe or tibble

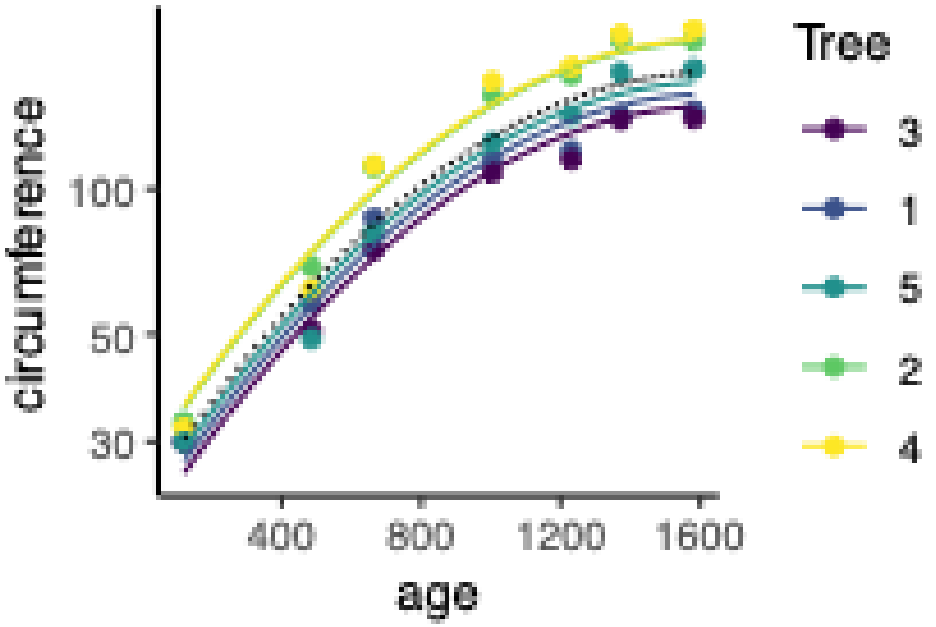
Generalised Linear Modelling 1 – some options

Starting point:

- Clarity about what is your response variable and what are explanatory variables
- Clarity about how you expect your response variable to be distributed
- Scientific questions about the relationship between response and explanatory variables
- Data in a long-form dataframe or tibble

Types of response variable

- Have considered different kinds of explanatory variable
  - Numeric, categorical, fixed, random
- Have silently assumed things about the response variable
  - Numeric, continuous
  - Normally distributed (have at least checked that one)
- What if I'm counting pigeons in my garden?
  - Counts: Numeric, categorical – integers only
  - Don't know how many pigeons there could be
  - Never negative
  - Is an average number (e.g. 2)

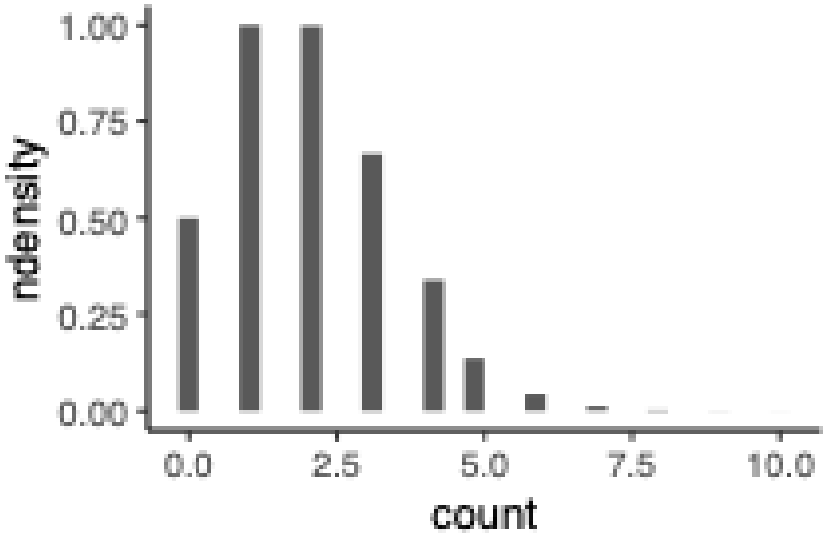


Types of response variable

- Have considered different kinds of explanatory variable
  - Numeric, categorical, fixed, random
- Have silently assumed things about the response variable
  - Numeric, continuous
  - Normally distributed (have at least checked that one)
- What if I'm counting pigeons in my garden?
  - Counts: Numeric, categorical – integers only
  - Don't know how many pigeons there could be
  - Never negative
  - Is an average number (e.g. 2)

Poisson models

`lambda`



Poisson distribution

```
rpois(1e5, lambda = 2)
```

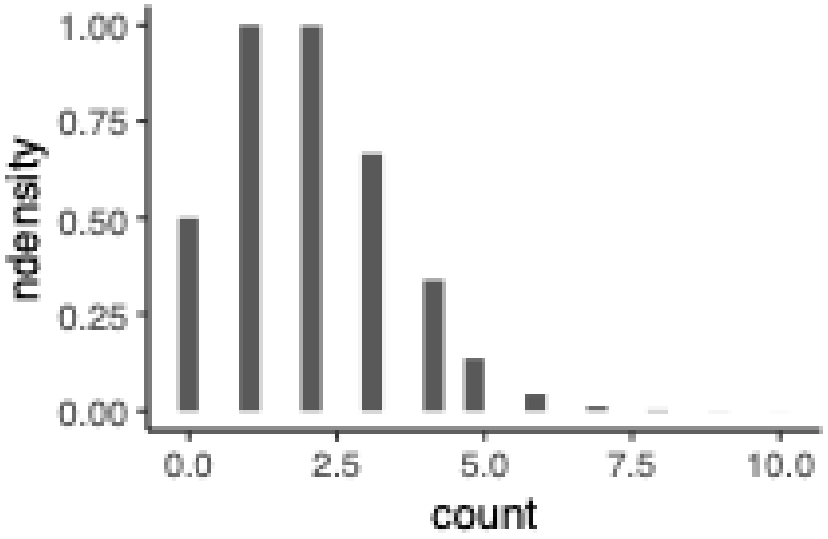
Types of response variable

- Have considered different kinds of explanatory variable
  - Numeric, categorical, fixed, random
- Have silently assumed things about the response variable
  - Numeric, continuous
  - Normally distributed (have at least checked that one)
- What if I'm counting pigeons in my garden?
  - Counts: Numeric, categorical – integers only
  - Don't know how many pigeons there could be
  - Never negative
  - Is an average number (e.g. 2)

Poisson models

`lambda`

- What if you you're classifying 'growth' vs. 'no growth'
  - Counts: Numeric, categorical – integers only
  - Do know how many plates I've checked (e.g. 4)
  - Never negative
  - Is a probability of growth/no growth (e.g. 0.5)



Poisson distribution

```
rpois(1e5, lambda = 2)
```

Types of response variable

- Have considered different kinds of explanatory variable
  - Numeric, categorical, fixed, random
- Have silently assumed things about the response variable
  - Numeric, continuous
  - Normally distributed (have at least checked that one)
- What if I'm counting pigeons in my garden?
  - Counts: Numeric, categorical – integers only
  - Don't know how many pigeons there could be
  - Never negative
  - Is an average number (e.g. 2)
- What if you you're classifying 'growth' vs. 'no growth'
  - Counts: Numeric, categorical – integers only
  - Do know how many plates I've checked (e.g. 4)
  - Never negative
  - Is a probability of growth/no growth (e.g. 0.5)
- Transform the variable?
  - Maybe
- Make different assumptions in the regression
  - Better
- Many other sorts of response variables
  - Time to event (survival analysis); multinomial models; negative binomial models…

Binomial distribution

```
rbinom(1e5, size = 4, prob = 0.5 )
```

Binomial models

Number of trials, `size`

Probability of 'success', `prob`

Fit with `glm(family = )`
`glmer(family = )`

`family = "poisson"`
`family = "binomial"`
…

Generalized linear models