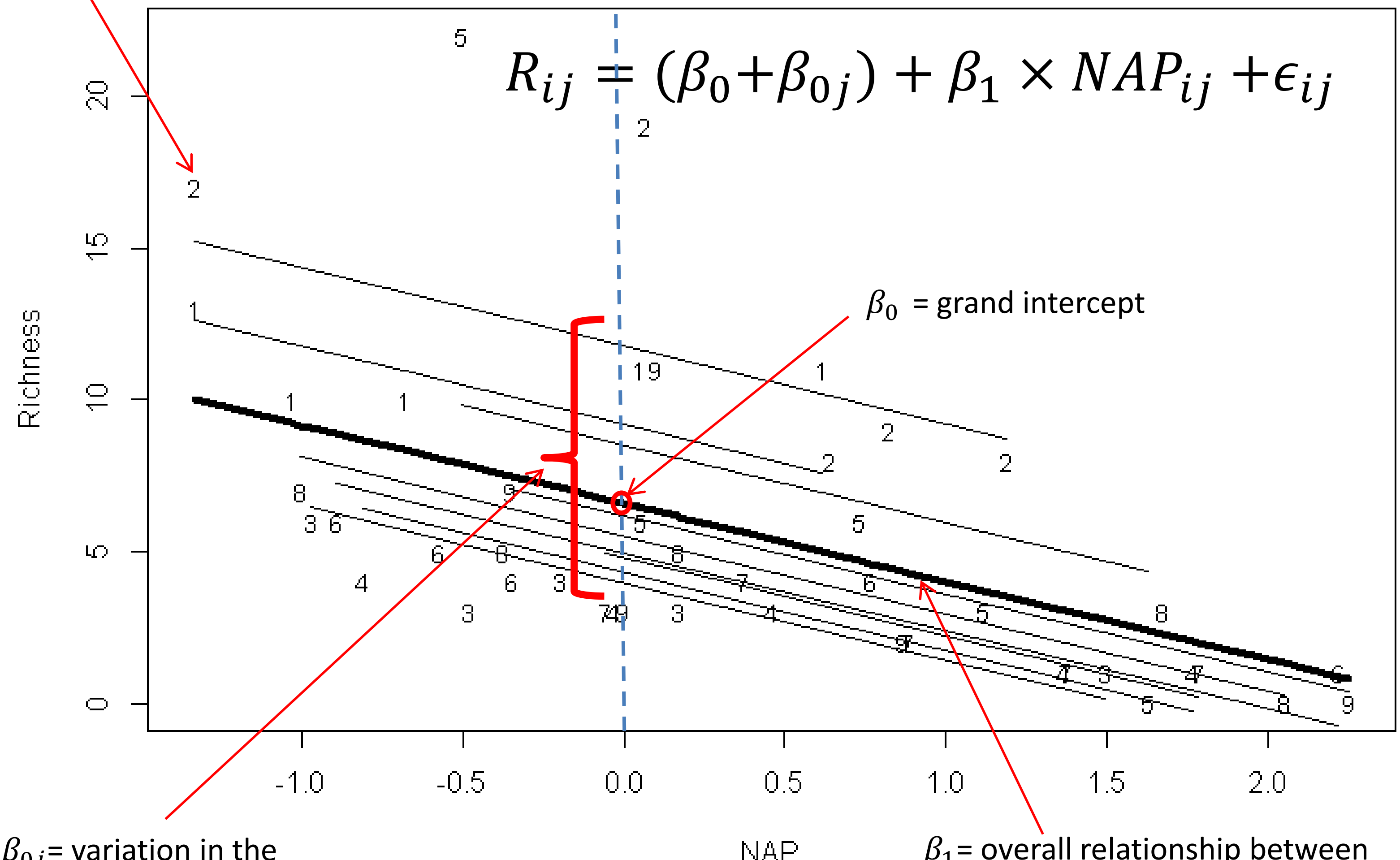
9. Random regression models (random slopes)

Random intercepts model

R_{ij} = individual observation (coded for Beach ID)

$$R_{ij} = (\beta_0 + \beta_{0j}) + \beta_1 \times NAP_{ij} + \epsilon_{ij}$$

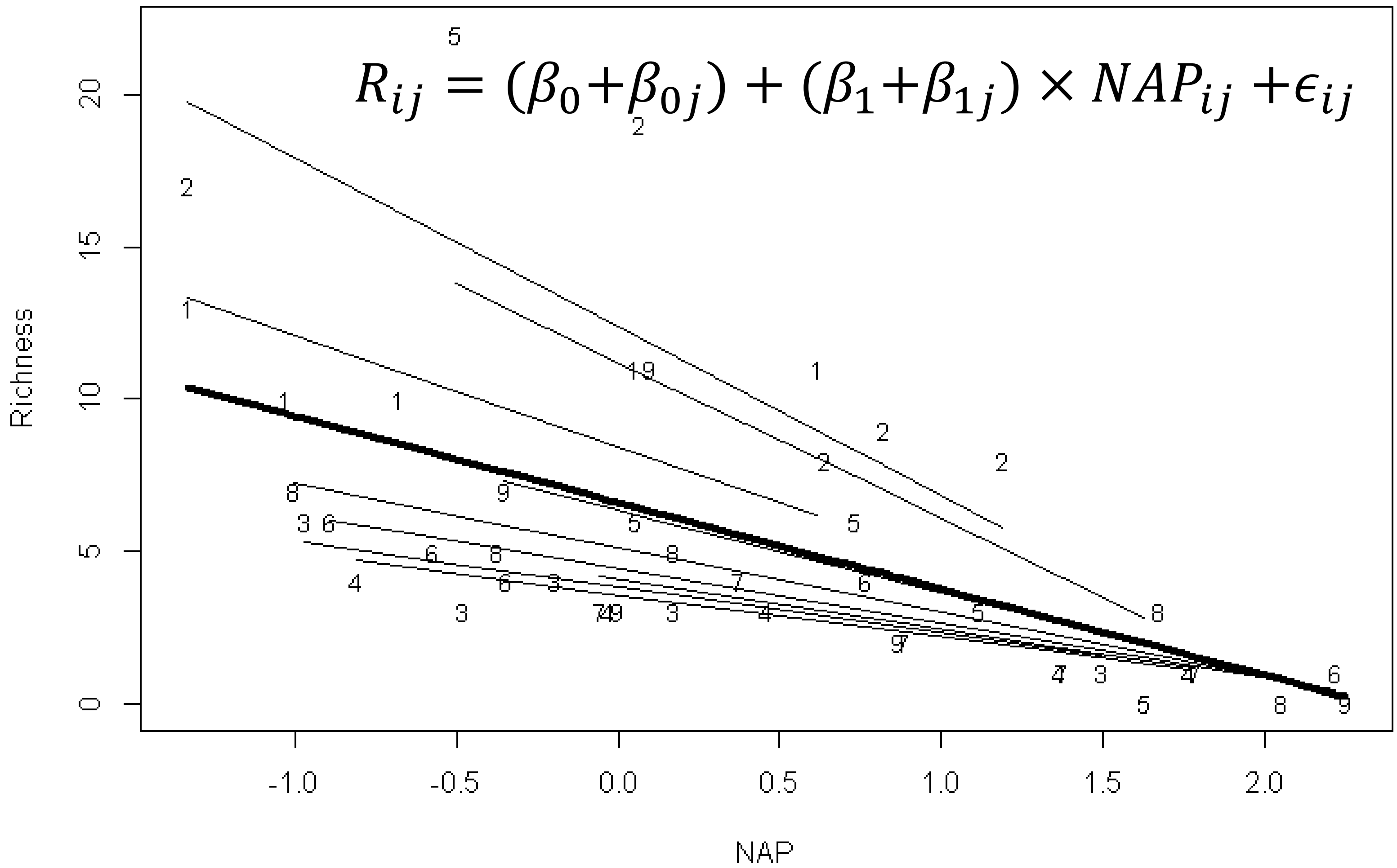
β_0 = grand intercept



β_{0j}= variation in the intercepts among the beaches

β_1= overall relationship between Richness & NAP

Random slopes (& intercepts) model

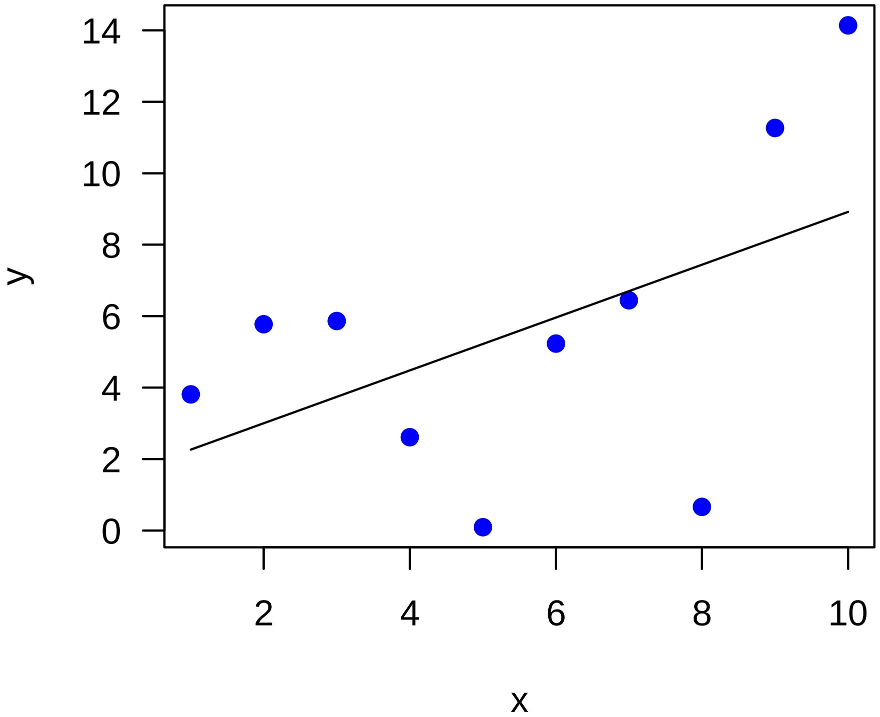


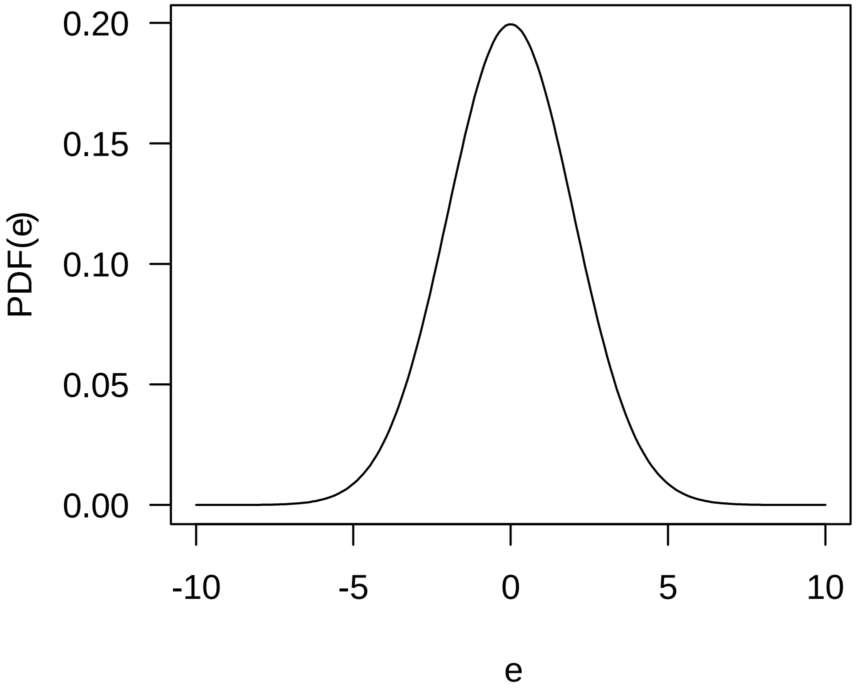
$$R_{ij} = (\beta_0 + \beta_{0j}) + (\beta_1 + \beta_{1j}) \times NAP_{ij} + \epsilon_{ij}$$

Linear model

$$y_{ij} = \beta_0 + \beta_1 \times x_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} \sim \mathrm{N}(0, \sigma_e^2)$$



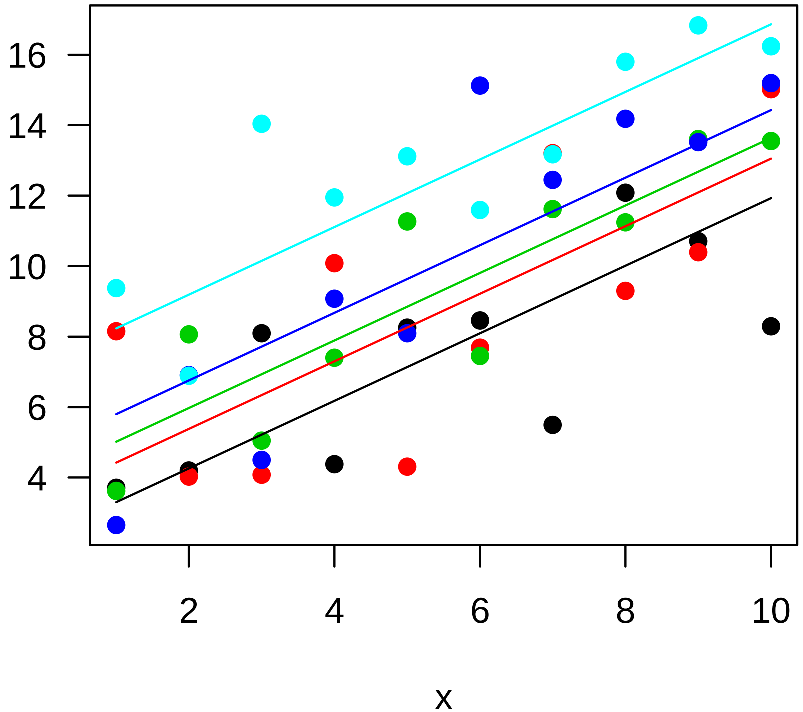


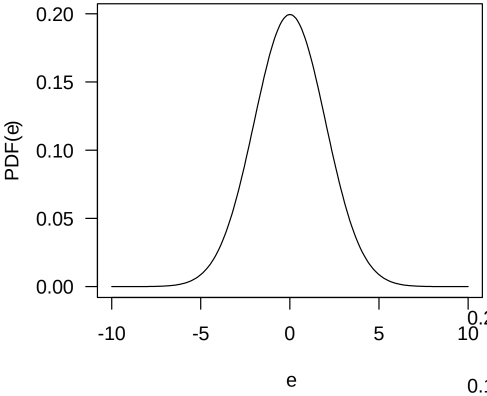
Linear mixed-effect model (random intercepts)

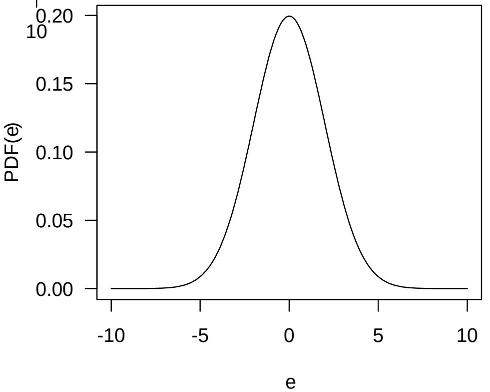
$$y_{ij} = (\boldsymbol{\beta_0} + \boldsymbol{\beta_{0j}}) + \beta_1 \times x_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} \sim \mathrm{N}(0, \sigma_e^2)$$

$$\beta_{0j} \sim \mathrm{N}(0, \sigma_{int}^2)$$

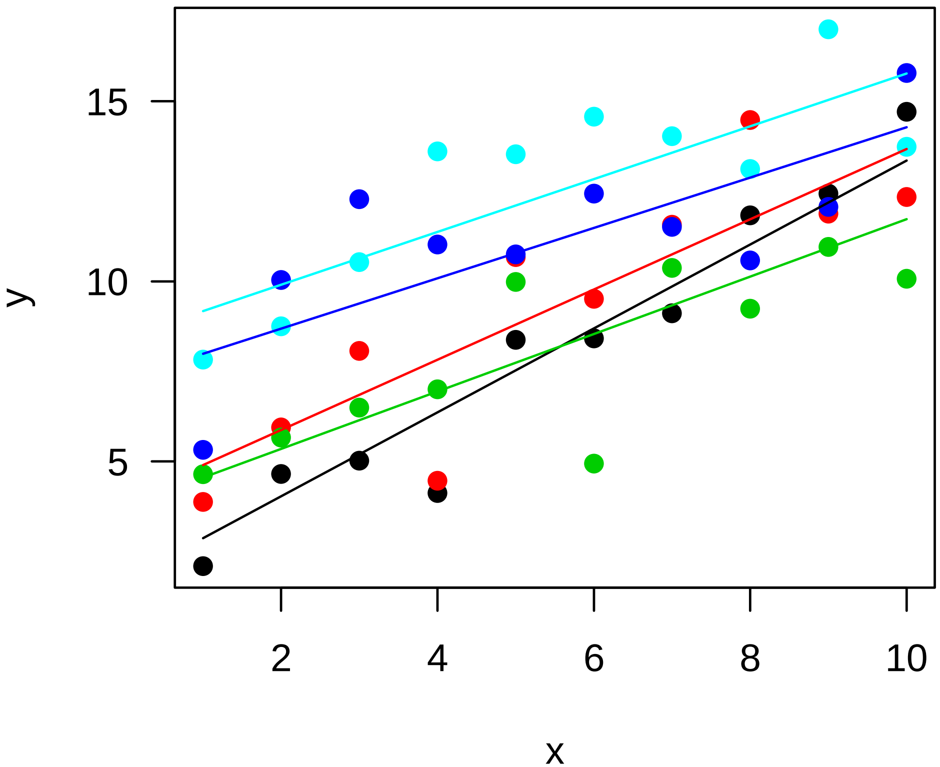






Linear ***mixed-effects*** model
(random intercept and slope model, aka. random regression)

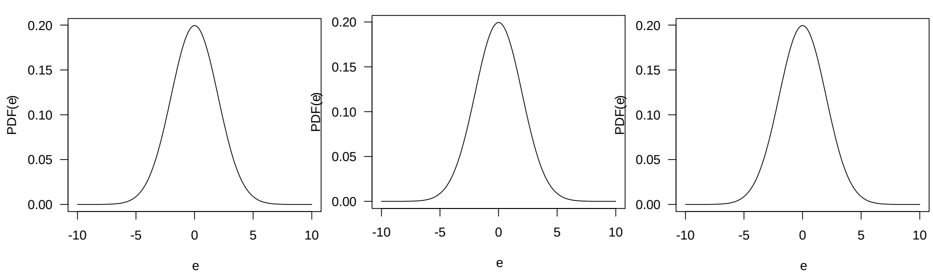
$$y_{ij} = \underbrace{(\beta_0 + \beta_{0j})}_{\text{intercept}} + \underbrace{(\beta_1 + \beta_{1j})}_{\text{slope of } x_1} x_{1ij} + \underbrace{\varepsilon_{ij}}_{\text{error}}$$



$$\epsilon_{ij} \sim \mathrm{N}(0, \sigma_e^2)$$

$$\beta_{0j} \sim \mathrm{N}(0, \sigma_{int}^2)$$

$$\beta_{1j} \sim N\left(0, \sigma_{\mathrm{slope[1]}}^2\right)$$



When do we use random regression?

- When we are interested in a general trend of the relationship (intercept and slope).
- But when observations are not independent, and slopes are not equal among groups.
- When we are interested in whether among-group slopes are equal.

RIKZ data set

- Include random slopes in the RIKZ data set

Do it on your own!

- Use the Cetaceans dataset and work through the model selection
- Start with this model:
  - Age ~ fStain*fSex + fStain*fLocation + + (1|fSpecies/fDolphinID)
- Try to validate this model and see what needs to be changed about the random structure
- Then finish the model selection (Stain is our fixed effect of interest, i.e. ‘treatment’)

Another variance het data set

- The ‘Biodiversity’ data set has similar variance heterogeneity problems
- You can work through this dataset too if you want
  - Concentration is the response variable
  - Biomass, Treatment, Nutrient are all predictors
  - Just make sure to make Treatment and Nutrient levels *factors* first

BONUS – generalized linear mixed models

- Data (specifically, the residuals from your data) are often not normally distributed
  - Poisson (counts)
  - Binomial (presence/absence)
  - Negative binomial
  - Gamma
- Mixed models can handle these too
- But this is an area of active research and there is lots of disagreement about the best way to run these models
- Read Chapter 13 in Zuur!

BONUS – generalized linear mixed models

- Packages that can run GLMMs (there are others):
  - lme4 (using the glmer() command)
  - glmmPQL
  - glmmML
- Unfortunately, because behind-the-scenes math is not agreed upon, they can sometimes give different results (& so be **very** cautious about strongly believing results close to p =0.05)

Review

Cell PRESS

Generalized linear mixed models: a practical guide for ecology and evolution

Benjamin M. Bolker[1], Mollie E. Brooks[1], Connie J. Clark[1], Shane W. Geange[2], John R. Poulsen[1], M. Henry H. Stevens[3] and Jada-Simone S. White[1]

[1] Department of Botany and Zoology, University of Florida, PO Box 118525, Gainesville, FL 32611-8525, USA
[2] School of Biological Sciences, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand
[3] Department of Botany, Miami University, Oxford, OH 45056, USA

Really impressive supplementary too, with worked examples!

PeerJ

Using observation-level random effects to model overdispersion in count data in ecology and evolution

Xavier A. Harrison

Institute of Zoology, Zoological Society of London, London, UK

ABSTRACT

Overdispersion is common in models of count data in ecology and evolutionary

PeerJ

A comparison of observation-level random effect and Beta-Binomial models for modelling overdispersion in Binomial data in ecology & evolution

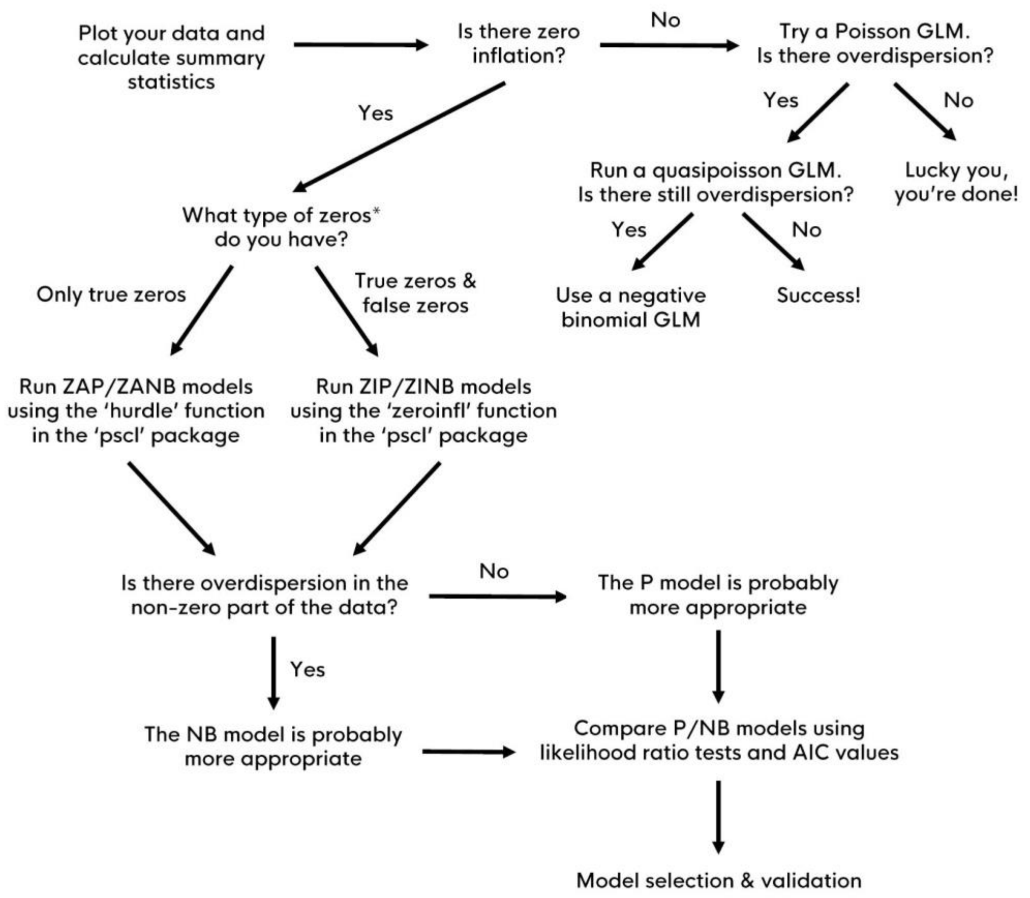
Xavier A. Harrison

Institute of Zoology, Zoological Society of London, UK

ABSTRACT

Overdispersion is a common feature of models of biological data, but researchers

Analysing count data in R



*True (structural) zeros arise when there is nothing to count. False (sampling) zeros arise when there could be things to be counted but they were not detected.

From: Natalie Pilakouta