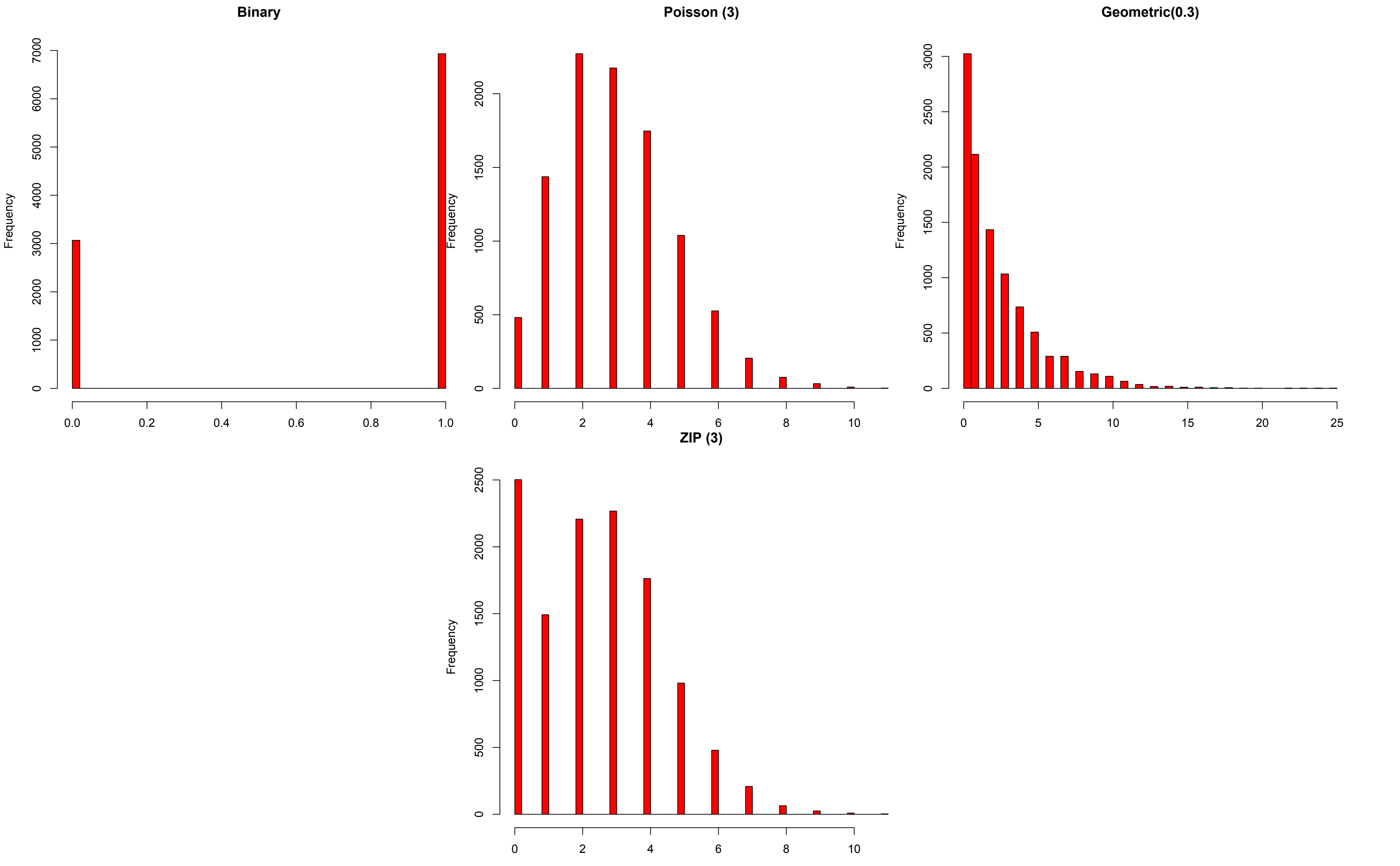
Notes about GLMs

Data are often not normal



We can model the response as function of predictors using link functions

- Defaults in GLMs

- Binomial data (logit link)

$$logit(p_i) = \log\left(p_i/(1-p_i\right) = BX_i$$

- Poisson, Negative Binomial, Gamma, Lognormal (log link)

$$log(u_i) = BX_i$$

- Note that these formulas don't include additional error (like regression)

GLMMs

- Including additional variation turns GLMs -> GLMMs

$$logit(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = BX_i + e_i$$

$$log(u_i) = BX_i + e_i$$

$$e_i \sim Normal(0, \sigma)$$

- More data hungry, but flexible
  - Random effects allow us to turn ordinary GLMMs into time series models or models with spatial effects

Where have we seen this before?

$$X_{t+1} = BX_t + e_t$$

$$e_t \sim Normal(0, q)$$

$$logit(p_t) = X_t$$

$$Y_t \sim Bernoulli(p_t)$$

- We could construct a DLM with binomial response (or any other distribution)

Univariate -> multivariate

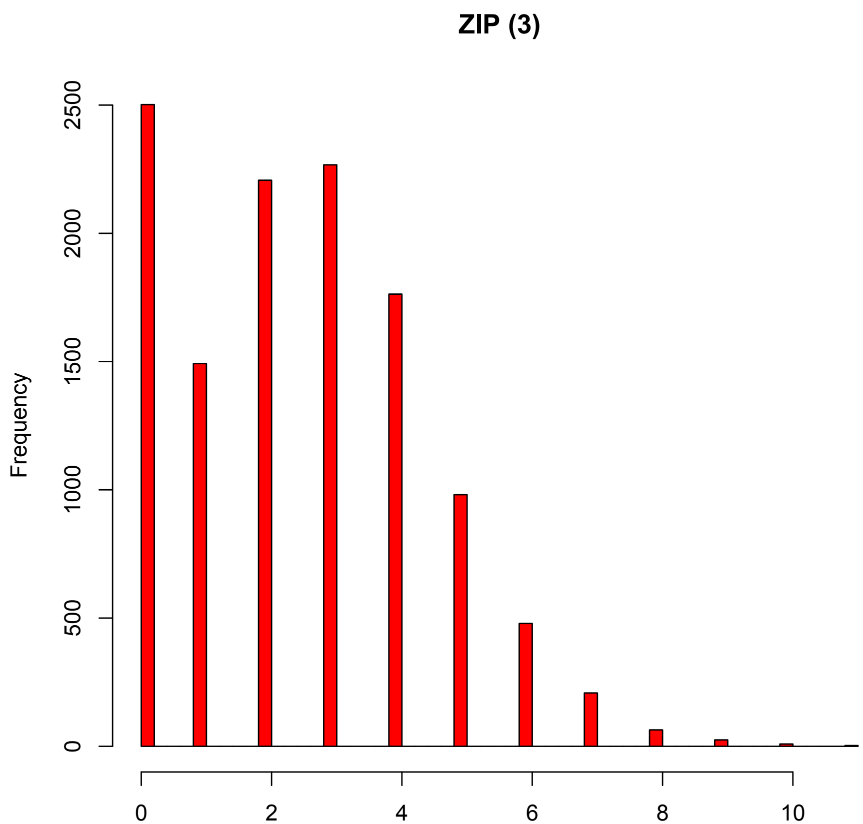
- For population *i*

$$X_{i,t+1} = BX_{i,t} + e_{i,t}$$

- As in MARSS models, we need to think about how to model the deviations
  - Independent and shared variance across pops?
  - Independent and unique variance across pops?
  - "equalvarcov"
  - Unconstrained
  - Model covariance as spatially correlated

Delta-GLMs

- Density of marine fishes almost always fits this pattern (zero inflated)



Delta-GLM or 'hurdle models'

- Breaks the response into 2 parts
  - Presence / absence
  - Positive density
- 2 separate GLMs
  - May include different covariates
- If we include random effects / shared terms, they usually aren't correlated across models
  - Different data + different link functions = weird interpretation
  - Results from both models combined for estimates of total density (nwfscDeltaGLM)