A quick overview of GLMs

ISTA 410 / INFO 510: Bayesian Modeling and Inference

U. of Arizona School of Information

October 28, 2020

Outline

Last time:

- Covariance between parameters

Today (and forthcoming):

- More linear and generalized linear models
- Mixtures and nonlinear models

Generalized linear models

GLMs in a nutshell

Basic idea of a GLM:

- Want the mechanics of a linear regression, but outcomes aren't normally distributed
  - outcomes may be discrete/categorical
  - outcomes may have heavier tails than a normal distribution
- So, use an outcome distribution dependent on an expectation parameter $E[y]$ and model

$$g(E[y_i]) = \boldsymbol{\beta} \cdot (x_i)$$

- What's g? The link function

Link functions

Link functions:

- Transform the linear model so that it takes on sensible values
- e.g., probabilities lie in $[0, 1]$, rates lie in $[0, \infty)$
- Most common include:
    - logit (common for binomial outcomes)
    - log (common for Poisson outcomes)
    - probit (similar to logit, but different tails)

Logistic regression

Most familiar GLM: logistic regression

- Binomial outcome, logit link
- Underlying parameter

$$y_i \sim \mathrm{Binomial}(p, n_i)$$
$$\mathrm{logit}(p) = \alpha + \boldsymbol{\beta} \cdot \mathrm{x}$$

(We saw this last week with the UC Berkeley admissions data)

Example: varying-intercepts Poisson regression

Way back when:

- Kidney cancer data
- Strange pattern:
- Approach back then: fully separate model

$$
\begin{aligned}
y_i &\sim \mathrm{Poisson}(n_i \lambda_i) \\
\lambda_i &\sim \mathrm{Gamma}(\alpha_0, \beta_0)
\end{aligned}
$$

Example: varying-intercepts Poisson regression

Extension: add county-level predictors

- Download data set of census-derived demographic data
- Join on to kidney cancer death data frame
- Fit a Poisson GLM with varying intercepts

Why varying intercepts? Helps deal with high variability

The model

Here is an example of the model:

$$
\begin{aligned}
y_i &\sim \mathrm{Poisson}(n_i \lambda_i) \\
\log(\lambda_i) &= \alpha_i + \boldsymbol{\beta} \cdot x \\
\alpha_i &\sim \mathrm{Normal}(0, \sigma) \\
\sigma &\sim \mathrm{HalfCauchy}(5) \\
\beta_i &\sim \mathrm{Normal}(0, 0.3)
\end{aligned}
$$

The model

Here is an example of the model:

$$
\begin{aligned}
y_i &\sim \mathrm{Poisson}(n_i \lambda_i) \\
\log(\lambda_i) &= \alpha_i + \boldsymbol{\beta} \cdot x \\
\alpha_i &\sim \mathrm{Normal}(0, \sigma) \\
\sigma &\sim \mathrm{HalfCauchy}(5) \\
\beta_i &\sim \mathrm{Normal}(0, 0.3)
\end{aligned}
$$

Notes:

- Can work with a longer-tailed prior on σ
- Varying intercepts needed for handling of variability (see also section 16.4)
- What predictors go in x?

Offsets

- Useful in some cases to fix a known coefficient; the predictor is now known as an *offset*
- Particularly in Poisson models: used to model exposure
- Idea: what we are trying to estimate is the rate of kidney cancer deaths, but Poisson variables give a count
- Need to account for varying population

Also appears in section 16.4's Poisson model.

Robust regression

Another purpose for GLMs

- Most obvious application of GLMs: allow regression with different outcome types (binomial, multinomial, Poisson count)
- Another application: robust regression
  - robustness to outliers – can we accommodate some extreme examples or greater-than-expected variation?
  - sensitivity analysis – is the model sensitive to the normal assumption?

Overdispersion in GLMs

Overdispersion: data displays more variation than expected from the outcome distribution

We have seen some of this already:

- our Poisson model used varying intercepts as an "overdispersion" parameter
- See also: model for police stops (BDA3 sec. 16.4)

Why does this happen? Common outcome distributions for GLMs do not have independent mean & variance

- Poisson: variance equal to mean
- Binomial: variance equal to $np(1-p)$

Substituting an overdispersed distribution

In our previous Poisson model we insert overdispersion by including varying intercepts (thought of as an extra error term)

Another approach: substitute an outcome model with an additional dispersion parameter

- Normal $\rightarrow$ Student t
- Poisson $\rightarrow$ negative binomial
- Binomial $\rightarrow$ beta-binomial

All of the above can be thought of as mixture models, where the dispersion parameter is sampled first followed by the outcome variable.

Summary

GLMs allow more flexibility with respect to outcome types and dispersion than conventional normal models; can require some care.

Next week:

- Nonlinear models and mixtures
- More general graphical models (maybe)