

Week Nine

Last Week

- Exams

This Week: Generalized Linear Models for Count Data

Today:

- Exam Recap:
- Activity:
 - GLMs for count data
- Thursday: Lab
 - Separation

Next Week: Count Regression / Ordinal Regression

Probability Distributions for Count Data

Poisson Distribution:

$$Pr[X = k] = \frac{\lambda^k \exp(-\lambda)}{k!}$$

- Expectation
- Variance

Use `rpois()` to generate and visualize data with different λ parameters.

- $\lambda = [1, 5, 10, 20]$

Negative Binomial Distribution:

$$Pr[X = k] = \frac{\Gamma(k+n)}{\Gamma(n)k!} p^n (1-p)^k$$

- Expectation
- Variance

Alternatively, we can define

- the mean,
- the size (dispersion parameter), as

Use `rnbinom()` with `mu` and `size` to simulate data with some different combinations of the parameters.

- $\mu = [1, 5, 10, 20]$
- `size = [.75, 1, 10]`

Then plot figures and confirm that the mean and variance of the data match your expectations.

Count Regression

Recall that a GLM has three parts: random component, systematic component, and link function.

So with Poisson regression, it looks like this

Let's generate data with one continuous predictor

Fit a generalized linear model to your data.

Add the regression fit line to your figure