

# glm

November 17, 2020

Generalized Linear Models in R

## 1 1. Generalized Linear Models

- Family of models *including* linear models
- Generalization of linear models to included categorical and/or non-normally distributed dependent/outcome variable
- Examples:
  - Binomial logistic regression
  - Multinomial regression
  - Poisson regression

## 1.1 Note on Terminology - General Linear Model  $\neq$  Generalized Linear Model

- General linear model refers to models with a continuous outcome variable, and assumption of normality
  - ANOVA (and friends)
  - Linear regression
- Term **Generalized** Linear Model used to refer to a family of models for categorical and/or non-normally distributed outcome variables

## 2 2 Binomial Logistic Regression

- Linear regression assumes a continuous outcome variable
- If the outcome variable is *not* continuous, we need a different approach.
- In the case of a binary outcome variable, we model  $\Pr(y_i = 1)$

### 2.1 2.1 Logistic Regression vs. Linear Regression

Differences from linear regression: - Assumes outcome is bounded by 0 and 1 + that is  $0 \leq E(y_i) = \pi_i \leq 1$  - Variance of  $y$  is *not* constant (i.e., not the same for all  $y_i$ ) - Similarly, variance of  $\varepsilon$  is not constant - Computational differences (i.e., closed-form vs numerical methods)