

# Interpreting GLMs

October 26, 2018

## Abstract

This document will briefly discuss some generalized linear models, and give examples of how to interpret them.

## 1 Poisson regression models

Poisson regression is sometimes used to model count data. The canonical form of a Poisson (log-linear) regression model is

$$\begin{aligned}\log(E[Y|X]) &= \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \\ Y &\sim \text{Poisson}(E[Y|X])\end{aligned}$$

The Poisson distribution has only one parameter, the mean, which is also the variance of the distribution. So in estimating  $E[Y|X]$ , we are also estimating  $\text{Var}(Y|X)$ . This is a potential drawback to the Poisson model, because there is no variance parameter to estimate, and so incorrect models can give wildly inaccurate standard errors (frequently unrealistically small). A better model is a quasi-Poisson model, for which the variance is proportional to the mean, but not necessarily equal to it. The negative binomial regression model is also commonly used to address over-dispersed count data where the variance exceeds the mean.

The canonical link function for Poisson outcomes is the natural logarithm. When we use a log-link, we can write

$$E[Y|X] = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}.$$

We can interpret  $\beta_0$  as follows: for observations which are 0 on all of the predictors, we estimate that the mean (expected) value of the outcome will be  $e^{\beta_0}$ .

We can interpret  $\beta_1$  as follows: adjusting for the other predictors, a one-unit difference in  $X_1$  predicts a  $(e^{\beta_1} - 1) \times 100\%$  difference in the outcome.

Generally, when using a log-link, we assume that differences in the predictors are associated with multiplicative differences in the outcome.

Some advantages to using an exponential link are

1. the model is mathematically more tractable and simpler to fit
2. the model parameters are easy to interpret
3. the mean of  $Y$  is guaranteed to be positive for all values of  $X$ , which is required by the Poisson distribution

We can fit a Poisson log-linear regression by writing  
`glm(Y ~ X, family = poisson(link = 'log'))`

To fit a quasi-Poisson model, write

```
glm(Y ~ X, family = quasipoisson(link = 'log'))
```

To fit a negative binomial regression model, write (after loading the MASS library)

```
glm.nb(Y ~ X, link='log')
```

To fit a Poisson regression with an identity link (where coefficients are interpreted as expected differences in the outcome associated with unit differences in the predictor), write

```
glm(Y ~ X, family = poisson(link = 'identity'))
```

To fit a Poisson regression with a square root link, which is vaguely like a compromise between an identity link and a log link (and is harder to interpret than either), write

```
glm(Y ~ X, family = poisson(link = 'sqrt'))
```

To fit a Poisson log-linear model with a random intercept and slope, write

```
glmer(Y ~ X + (X|grp), family = poisson(link = 'log'))
```

## 2 Dichotomous regression models

When predicting either successes and failures, or proportions, we can use a model with a binomial outcome. Here we'll focus on models where the data is represented as individual successes and failures. The canonical model for these data is logistic regression, where

$$\text{logit}(E[Y|X]) \equiv \log \left( \frac{P(Y = 1|X)}{1 - P(Y = 1|X)} \right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$$
$$Y \sim \text{Binomial}(1, E[Y|X])$$

We can rewrite this model as

$$\text{odds}(Y) = \frac{P(Y = 1|X)}{1 - P(Y = 1|X)} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}$$

or

$$P(Y = 1|X) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}$$

We can interpret  $\beta_0$  as follows: for observations which are 0 on all of the predictors, we estimate that the mean value of the outcome will be  $\frac{e^{\beta_0}}{1 + e^{\beta_0}}$ . That is, we estimate that the probability of the outcome being a 'success' (assuming 'success' is coded as a 1) will be  $\frac{e^{\beta_0}}{1 + e^{\beta_0}}$ .

We can interpret  $\beta_1$  as follows: adjusting for the other predictors, a one-unit difference in  $X_1$  predicts a  $\beta_1$  difference in the log-odds of the outcome being one, or a  $(e^{\beta_1} - 1) \times 100\%$  difference in the odds of the outcome. Unfortunately, the change in probability of a unit change depends on where the starting point is, so there is no easy way to interpret these coefficients in terms of direct probability. One can calculate the estimated change for specific units, however, and look at the distribution of those changes.

Other possible link functions include the probit (which uses a Normal CDF to link  $\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$  to  $P(Y = 1|X)$ ), or the complementary log-log (which allows  $P(Y = 1|X)$  to be asymmetric in the predictors), among others.

We can fit a logistic regression model by writing

```
glm(Y ~ X, family = binomial(link = 'logit'))
```

We can fit a probit regression model by writing

```
glm(Y ~ X, family = binomial(link = 'probit'))
```

We can fit a complementary log-log model by writing

```
glm(Y ~ X, family = binomial(link = 'cloglog'))
```

We can allow a random slope and intercept by writing

```
glmer(Y ~ X + (X|grp), family = binomial(link = 'logit'))
```