# Introduction to Poisson Regression with Robust Standard Errors - Part 2

Christopher Liu, Siyu Zhao, Zheming Yu, Runnan Liu, Shiming Zhang

11/29/2018

San Diego State University, Stats 610

# Objectives of the Video

- Develop a theoretical basis for the Poisson Regression Model
- Understand the interpretation of the model's parameters
- Understand the uses and limitations of this specific regression model

# The Poisson Distribution

A random variable Y is said to have a Poisson distribution with parameter $\lambda$ if its probability is given by the probability mass function

$$Pr(Y = y) = \frac{e^{-\lambda}\lambda^{y}}{y!}$$

for $\lambda > 0$ and y = 0, 1, 2, ...

The mean and variance of this distribution can be shown to be

$$E(Y) = Var(Y) = \lambda$$

# Introduction to Poisson Regression

In Poisson Regression:

- Model used when the desired response variable, $Y_i$, is a count (eg. Number of vehicle accidents per year, number of visits to a website over a certain time span, etc)
- We can also have the response variable be $Y_i/\mathrm{t}$, the rate at which the event happens with t being an interval representing time, space, or some other grouping of interest

# Introduction to Poisson Regression

- The regression model with the log link function:

$$log(\lambda_i|X_i) = \beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip} = X_i\beta$$

  where $E(Y_i|X_i) = \lambda_i = e^{X_i\beta}$

- Predictor variables are estimated by maximizing the likelihood function:

$$L(\beta) = \prod_{i=1}^{n} f(Y_i) = \prod_{i=1}^{n} \frac{e^{-\lambda_i}\lambda_i^{Y_i}}{Y_i!}$$

# Introduction to Poisson Regression

Take the simple case:

$$log(\lambda_i|x) = \beta_0 + \beta_1 x$$

Consider the difference between the mean response given $(x + 1)$ and the mean response given $x$:

$$log(\lambda_i|x + 1) - log(\lambda_i|x)$$

$$= \beta_0 + \beta_1(x + 1) - (\beta_0 + \beta_1 x) = \beta_1$$

$$\implies \frac{(\lambda_i|x + 1)}{(\lambda_i|x)} = e^{\beta_1}$$

# Objectives of this Video

Objectives of this video:

- Demonstrate Poisson Regression with Robust Standard Errors on real data
- Discuss analysis/results
- Discuss basic diagnostic methods we can use