Problem Set 7

Due Dec 12, before class

1. Moments of zero-inflated count distributions

Calculate the expectation and variance of the zero-inflated Poisson and Negative-Binomial models.

2. Poisson latent variable and the cloglog model

Assume that $y_i^* \mid x_i \sim \text{Poisson}(e^{x_i^T \beta})$, and define $y_i = \mathbf{1}(y_i^* > 0)$ as the indicator that y_i^* is not zero. Show that $y_i \mid x_i$ follows a cloglog model,

$$\text{pr}(y_i = 1 \mid x_i) = g(x_i^T \beta),$$

where $g(z) = 1 - \exp(-e^z)$.

3. Likelihood for the zero-inflated Poisson regression

Write down the likelihood function for the Zero-inflated Poisson model, and derive the steps for Newton's method.

4. Data analysis

Zeileis et al. (2008) give a tutorial on count outcome regressions. Replicate their analysis and compare the results.

Reference:

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of*

statistical software, 27(8), 1-25.

Note: Read the tutorial to see how you can obtain the data. The tutorial introduces the hurdle model. Implement the hurdle model in addition to the Poisson, Negative-binomial, and zero-inflated models.