

Bayesian Model for Vehicle Crashes on Two-Lane Primary Roads in Iowa

Daniel Ries
Michael D. Pawlovich
Alicia Carriquiry
Zachary Hans

March 24, 2017

Abstract

Enter a short summary here. What topic do you want to investigate and why? What experiment did you perform? What were your main results and conclusion?

Contents

1	Introduction and Data Description	1
2	Model Specification	1
3	Variable Selection	2

1 Introduction and Data Description

In this report we present a model as well as results to produce Safety Performance Functions (SPF) for two lane primary roads in Iowa. The model we develop and estimate uses data from 2005-2014. We have information on over 70 road characteristics such as speed limit, median width, volume and access control (to name a few) for each road segment. We also have the total number of crashes, number of fatal crashes, number of crashes resulting in major injuries and minor injuries, and number of crashes resulting in only property damage. Because fatal crashes are relatively uncommon, we will combine fatal and major crashes into one category. There are 94,370 observations at 10,055 unique road segments.

2 Model Specification

To model the number of crashes as a function of observed covariates, we use a generalized linear mixed effects model with Poisson response and a log link. The random effect comes from the correlation induced between years at specific road segments. Because we have measurements at the same location over the

course of 10 years, we include random effects to allow dependence across years. We also include an offset term as road segments are not all equal length.

Denote:

- Y_{ij} : number of crashes in segment i during year j
- t_i : length of segment i
- x_{ij} : p -dimensional vector of covariates for segment i and year j
- β : p -dimensional parameter vector for covariate coefficients
- v_i : random intercept for segment i , induces correlation between years for segment i

The model is characterized by:

$$\begin{aligned} Y_{ij} | \lambda_{ij}, t_i &\overset{iid}{\sim} \text{Poisson}(\lambda_{ij} t_i), \\ \log(\lambda_{ij}) | \beta, v_i &= x_{ij}^T \beta + v_i, \\ v_i &\overset{iid}{\sim} N(0, \sigma_v^2). \end{aligned}$$

We specify the priors as:

$$\begin{aligned} \beta &\overset{iid}{\sim} N(0, 100), \\ \sigma_v^2 &\sim IG(0.1, 0.1). \end{aligned}$$

3 Variable Selection

There