Safety Science 54 (2013) 38–42

Contents lists available at SciVerse ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

The Poisson–Weibull generalized linear model for analyzing motor vehicle crash data

Lingzi Cheng [a], Srinivas Reddy Geedipally [b,1], Dominique Lord [a,*]

[a] Zachry Department of Civil Engineering, Texas A&M University, 3136 TAMU College Station, TX 77843-3136, USA
[b] Texas Transportation Institute, 110N Davis Drive, Suite 101, Room 124, Arlington, TX 76013-1877, USA

ARTICLE INFO

Article history:
Received 25 July 2012
Received in revised form 31 August 2012
Accepted 4 November 2012
Available online 23 December 2012

Keywords:
Poisson–Weibull
Poisson–gamma
Generalized linear model
Traffic crashes

ABSTRACT

Over the last 20–30 years, there has been a significant amount of tools and statistical methods that have been proposed for analyzing crash data. Yet, the Poisson–gamma (PG) is still the most commonly used and widely acceptable model. This paper documents the application of the Poisson–Weibull (PW) generalized linear model (GLM) for modeling motor vehicle crashes. The objectives of this study were to evaluate the application of the PW GLM for analyzing this kind of dataset and compare the results with the traditional PG model. To accomplish the objectives of the study, the modeling performance of the PW model was first examined using a simulated dataset and then several PW and PG GLMs were developed and compared using two observed crash datasets. The results of this study show that the PW GLM performs as well as the PG GLM in terms of goodness-of-fit statistics.

1. Introduction

Over the last 20–30 years, there has been a significant amount of tools and statistical methods that have been proposed for analyzing crash data. These tools are needed because of the nature of the data, which are characterized by the random and (assumed) independent discrete non-negative events. Lord and Mannering (2010) have recently documented the latest regression models that have been developed or applied for modeling crash data. Statistical models can be used for various purposes, such as establishing relationships between motor vehicle crashes and different covariates (i.e., understanding the system), predicting values or screening variables.

As documented in the literature, crash data are characterized by unique attributes. They include datasets that contain a limited number of observations which can also be subjected to low sample mean values (see Lord and Mannering, 2010; Maher and Summersgill, 1996, and references herein). It has been noted that these unique attributes could negatively influence regression models by providing biased estimates of regression coefficients or predicted values (Lord, 2006). As a result, a few researchers have examined the application of alternative regression models that could be used for minimizing these problems. Some of the models include the Poisson–lognormal, the negative binomial (NB) bootstrap maximum likelihood estimation (MLE) method and the NB-Lindley model (Miaou et al., 2003; Zhang et al., 2007; Geedipally et al., 2012).

In the line of the discussion above, this paper documents the application of the Poisson–Weibull (PW) regression model for analyzing crash data. Before examining the stability of this model for datasets characterized by small sample sizes and low sample mean values, there is a need to determine whether or not the PW model performs as well as the traditional NB or Poisson–gamma (PG) model. If it does not, there is no need to evaluate it in the context of the issues described above. Although not very popular in traffic safety (due to the complexity of parameterization – described below), the PW model has been applied in the other areas (see, e.g., Wagner and Steenbakkers, 1989; Phoenix et al., 1997). In highway safety, the PW model has only been applied once (Maher and Mountain, 2009). Thus, there is a need to fully describe its modeling performance using simulated and observed datasets.

The study objective was accomplished using both simulated and observed datasets. First, the general performance of the PW model was shown using the simulated data. Then the modeling performance of the PW model was compared to the PG model using two observed datasets. Several goodness-of-fit (GOF) measures were used for evaluating the models.

The next section describes the characteristics of the PW distribution and its generalized linear model (GLM).

* Corresponding author. Tel.: +1 (979) 458 3949; fax: +1 (979) 845 6481.
E-mail addresses: lingzi_cheng@tamu.edu (L. Cheng), srinivas-g@ttimail.tamu.edu (S.R. Geedipally), d-lord@tamu.edu (D. Lord).
[1] Tel.: +1 (817) 462 0519x11519.

0925-7535/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ssci.2012.11.002

2. Background

This section describes the characteristics and parameterization of the PW GLM. First, the statistical background about the PW is provided. Then, the second part illustrates the parameterization of the PW model.

2.1. Characteristics of the PW GLM

The PW distribution is a mixture of Poisson and Weibull distributions, as the name implies. Similar to most Poisson-based distributions (e.g., Poisson–gamma, Poisson–lognormal, etc.), the PW model is also designed to accommodate the over-dispersion (McCullagh and Nelder, 1993; Raghavachari et al., 1997; Lord et al., 2005; Maher and Mountain, 2009).

First, the Poisson and Weibull distributions need to be defined, respectively. Crash data can be characterized as the product of Bernoulli trials with unequal probability of events (also known as Poisson trials). As the number of trials increases the distribution may approximately follow a Poisson process and the amount of dispersion is governed by the characteristics of this process (Lord et al., 2005). Thus, the number of crashes at ith roadway entity Y_i is assumed to be Poisson distributed with mean μ_i and independent over all entities:

$$Y_i|\mu_i \sim \text{Poisson}(\mu_i) \quad i = 1, 2, 3, \ldots, I \tag{1}$$

The Poisson mean μ_i is structured as:

$$\mu_i = \hat{\mu}_i \varepsilon_i = f(\mathbf{X}; \boldsymbol{\beta}) \cdot \varepsilon_i \tag{2}$$

and,

$$\hat{\mu}_i = f(\mathbf{X}; \boldsymbol{\beta}) = \exp(\beta_0 + \sum_{j=1}^{q} \beta_j X_j), j = 1, 2, 3, \ldots, q \tag{3}$$

where Xs are the independent variables; q represents the total number of independent variables; βs are the regression coefficients; and ε_i is the model error independent of all covariates (Cameron and Trivedi, 1998; Miaou and Lord, 2003; Lord and Park, 2010).

For the PW model, it is assumed that ε_i is independent of all covariates and Weibull distributed. This is used for capturing the extra variation that the traditional Poisson cannot fully handle. The Weibull probability density function (PDF) is given as follows:

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{x}{\lambda}\right)^{k}\right] \quad x > 0, k > 0, \lambda > 0 \tag{4}$$

where λ and k are scale and shape parameters, respectively. The PDF of the Weibull distribution has a wide variety of shapes depending on the k values and the shape can be similar to that of the gamma, gamma-like, exponential or approximate normal distributions. This characteristic gives the model a lot of flexibility to fit different kinds of data.

The mean and variance of the Weibull distribution are:

$$\begin{aligned} E(\varepsilon) &= \lambda\Gamma\left(1 + \tfrac{1}{k}\right) \\ Var(\varepsilon) &= \lambda^2\Gamma\left(1 + \tfrac{2}{k}\right) - \left[\lambda\Gamma\left(1 + \tfrac{1}{k}\right)\right]^2 \end{aligned} \tag{5}$$

Given the statistical characteristics of Poisson and Weibull distributions, the PW distribution is defined as the mixture of those two distributions such that

$$P(Y = y; \mu, \lambda, k) = \int \text{Poisson}(y; \hat{\mu}\varepsilon)\text{Weibull}(\varepsilon; \lambda, k)d\varepsilon \tag{6}$$

The mean or expected value of the PW distribution is given as:

$$E(Y) = \hat{\mu}E(\varepsilon) = \hat{\mu} \times \lambda\Gamma\left(1 + \frac{1}{k}\right) \tag{7}$$

and the variance is given by:

$$\begin{aligned} Var(Y) = \hat{\mu} \times \lambda\Gamma\left(1 + \frac{1}{k}\right) + \hat{\mu}^2 \times \lambda^2\Gamma\left(1 + \frac{2}{k}\right) - \hat{\mu}^2 \\ \times \left[\lambda\Gamma\left(1 + \frac{1}{k}\right)\right]^2 \end{aligned} \tag{8}$$

2.2. Parameter estimation

As discussed above, the likelihood function for the PW GLM was given by Eq. (6). An important characteristic of the PW distribution is that the likelihood function does not have a closed form. However, this can be solved by using a hierarchical representation of this function. As discussed above, the PW model is conditional on the site-specific error term ε which explains the additional heterogeneity. Therefore, the PW model can be also written as:

$$\begin{aligned} &P(Y = y; \mu|\varepsilon) = \text{Poisson}(y; \hat{\mu}\varepsilon) \\ &\varepsilon \sim \text{Weibull}(\varepsilon; \lambda, k) \end{aligned} \tag{9}$$

The formulation above has a nice Bayesian interpretation and hierarchical structure. Therefore, the parameter estimation and inference can be obtained by using Markov chain Monte Carlo (MCMC) and software such as WinBUGS (Spiegelhalter et al., 2003). Since the Bayesian formulation requires priors for all unknown parameters, non-informative normal priors for βs and gamma priors for $\omega(= 1/\lambda^k)$ and k were used.

The total number of iterations was determined using the Gelman–Rubin (G–R) statistic (Gelman and Rubin, 1992; Brooks and Gelman, 1998). The G–R convergence statistic is generally used to verify that the simulation runs converged properly and is based on analyzing multiple simulated MCMC chains (usually 3–5 chains) by comparing the variances within each chain and the variance between chains. A total of three Markov chains were used in the coefficient estimation process so that the G–R diagnostic was performed accurately. For model comparison, it was suggested that convergence was achieved when the G–R statistic was less than 1.2 (Mitra and Washington, 2007). The first half of iterations was used as burn-in samples and was discarded. Thus, the remaining half of the iterations was used for estimating the coefficients.

3. Methodology

This section describes the functional form and the GOF statistics that were used for comparing the PW GLM with the PG GLM. The PG model is one of the most commonly used models in highway safety application (Hauer, 1997; Lord, 2006; Lord and Mannering, 2010). The characteristics of the PG GLM can be found in other documents (see, e.g., Hilbe, 2011).

The functional form used for model development is as follows:

$$\hat{\mu}_i = \beta_0 \times L_i \times F_i^{\beta_1} \times y \times e^{\sum_{j=2}^{n} X_{ij}\beta_j} \tag{10}$$

where μ_i is the estimated number of crashes per year for site i; F_i the vehicles per day (ADT) for segment i; L_i the length of segment i in miles; y the number of years of crash data; X_{ij} the a series of covariates (e.g., shoulder width, lane width, etc.) for site i; n the number of covariates; and β_s is the estimated coefficients.

The GOFs used in this study for comparing the PW and PG GLMs included the Deviance Information Criterion (DIC), the Mean Absolute Deviance (MAD), the Mean Squared Predictive Error (MSPE) and the Pearson Chi-square (χ^2). The DIC is often used as a GOF measurement when Bayesian estimation method is used (Mitra and Washington, 2007; Lord and Park, 2010; Geedipally et al., 2012). In general, a smaller DIC value refers to a better model fit. As a rule of thumb, differences in the DIC values more than 10 will rule out the model with a higher DIC and a difference between 5 and 10 can be considered substantial (Spiegelhalter et al., 2002). However, it is important to point out that the DIC is dependent on the structure of the model. Thus, as discussed by Mitra and Washington (2007), and Lord and Park (2008), it is recommended to use more than one GOF measure for evaluating regression models.

The MAD, MSPE and Pearson χ^2 are three common GOF measurements that are also utilized for assessing how well the model fit the data. Generally, a model with a smaller value shows a better statistical fit. Interested readers are referred to Oh et al. (2003) for additional information about these GOF measures.

4. Data description

Two datasets are used in this study for comparing the modeling performance of the PW and PG GLMs. This section describes the characteristics of these two datasets.

4.1. Indiana data

Five years (1995–1999) of crash and traffic data collected at 338 rural interstate road segments in the state of Indiana were used for the model development. This dataset was utilized by Anastasopoulos et al. (2008), Washington et al. (2011) and Geedipally et al. (2012) for developing tobit and NB-Lindley models. Table 1 summarizes the key traffic and geometric variables. For a complete and detailed description of the variables, the interested reader is referred to Washington et al. (2011).

4.2. Texas data

Crash data from 1997 to 2001 collected at 4-lane rural undivided in Texas were provided by the Texas Department of Public Safety (DPS) and the Texas Department of Transportation (TxDOT). This dataset was used by Lord et al. (2008a,b) for the project titled "NCHRP 17-29: Methodology to Predict the Safety Performance of Rural Multilane Highways" and the evaluation of the Conway–Maxwell–Poisson GLM, respectively. The final database included 1499 segments ($\geqslant$0.1 mile). Table 2 presents the summary statistics of the data.

5. Modeling results

This section presents the modeling results for the PW and PG GLMs based on simulated and observed data.

5.1. Simulated data

This section presents the modeling results for the simulated data that are intended to illustrate the general performance of the PW model. The simulation design was carried out in several steps:

(1) Generate a true mean value (ρ_i) for roadway entity i from a fixed sample population mean (δ):

$$\rho_i = \delta \tag{11}$$

(2) Generate a value (ε_i) from a Weibull distribution with parameters k and λ. Here, the mean of the Weibull distribution is equal to 1. The values for k and λ are determined based on the dispersion parameter (DP) of the NB model ($Var(Y) = \mu + \alpha\mu^2$, where α is the DP) in order to better illustrate the level of dispersion based on a regression model commonly used in highway safety:

$$\varepsilon_i \sim \text{Weibull}(k, \lambda) \tag{12}$$

(3) Calculate the mean (μ_i) for roadway entity i:

$$\mu_i = \rho_i \times \varepsilon_i \tag{13}$$

(4) Generate a discrete value (Y_i) for entity i from a Poisson distribution with mean equal to μ_i:

$$Y_i \sim \text{Poisson}(\mu_i) \tag{14}$$

(5) Repeat steps 1–4 "N" times, where "N" is the required number of observations (which is also the sample size).

To avoid the potential bias associated with the small sample size and low sample mean (see Lord, 2006), the following values were considered:

(a) Sample size "N" = 300.
(b) Sample population mean "δ" = 10.
(c) DP "α" = 0.5, 1, 2, 3, 5.

The theoretical values for k and λ were calculated and labeled as "true" parameters for each dispersion level. By following the steps above, crashes or counts are then simulated using the PW distribution. Once the crash dataset was created, the parameters were re-estimated using the PW model and then compared with the "true" parameters.

Table 3 presents the modeling results for the simulated dataset. This table shows that PW model was able to reproduce the "true" parameter values. All parameter were statistically significant at the 5% level.

5.2. Indiana data

Table 4 summarizes the coefficient estimates and GOF statistics of the PG and PW models for the Indiana data. The segment length was treated as an offset and thus the crash frequency increases linearly with the increase in segment length. For both PG and PW models, the estimated coefficient for the traffic flow is less than 1. This indicates that with the increase in the traffic flow, the crash risk increases as the flow increases. It should be noted that the 95% marginal posterior credible intervals for each coefficient estimate did not include the origin. Table 4 shows that all the estimated coefficients between the two models have the same sign, and most of their coefficients are very close. The standard errors for both

Table 1
Indiana data summary.

Traffic variable	Minimum	Maximum	Average (std. dev.)	Total
Crashes (for 5 years)	0	329	16.97(36.30)	5737
Average daily traffic (ADT) (veh./day)	9442	143,422	30237.56(28776.43)	_
Minimum friction on road segment (FR) (0–100 scale)	15.90	48.20	30.51(6.67)	_
Pavement surface (PS) (asphalt = 1, concrete = 0)	0	1	0.77(0.42)	_
Median width (MW) (ft)	16	194.7	66.98(34.71)	_
Median barrier (MB) (present = 1, absent = 0)	0	1	0.16(0.37)	_
Interior rumble strips (IRS) (Present = 1, absent = 0)	0	1	0.72(0.45)	_
Segment length (L) (miles)	0.009	11.53	0.89(1.48)	300.09

Table 2
Texas data summary.

Traffic variable	Minimum	Maximum	Average (std. dev.)	Total
Crashes (for 5 years)	0	97	2.84(5.69)	4253
Average daily traffic (ADT) (veh./day)	42	24,800	6613.61(4010.01)	_
Lane width (LW) (ft)	9.75	16.5	12.57(1.59)	_
Total shoulder width (SW) (ft)	0	40	9.96(8.02)	_
Curve density (CD) (curves/mile)	0	18.07	1.43(2.35)	_
Segment length (L) (miles)	0.1	6.28	0.55(0.67)	830.49

Table 3
Modeling results for the simulated data.

Dispersion Level	α = 0.5		α = 1.0		α = 2.0		α = 3.0		α = 5.0	
Parameters	k	ω	k	ω	k	ω	k	ω	k	ω
Theoretical values	1.436	0.871	1.000	1.000	0.721	0.936	0.607	0.935	0.500	1.414
Estimated values (std. dev.)	1.446	0.873	1.000	0.993	0.731	0.931	0.607	0.936	0.506	1.414
	(0.080)	(0.061)	(0.053)	(0.065)	(0.038)	(0.060)	(0.032)	(0.061)	(0.029)	(0.084)

[1] $\omega = 1/\lambda^k$ where k, λ are the shape and scale parameters respectively in Eq. (4). The Weibull distribution was re-parameterized since WinBUGS uses a slightly different parameterization than the one discussed above. This did not influence the model performance.

Table 4
Modeling results for the Indiana data.

Variable	PG		PW	
	Value	S.E.	Value	S.E.
Intercept (ln β_0)	−4.627	1.354	−4.022	1.377
ADT (β_1)	0.7029	0.1254	0.6428	0.1243
FR (β_2)	−0.02589	0.01048	−0.02713	0.01128
PS (β_3)	0.4226	0.1874	0.4267	0.1992
MW (β_4)	−0.005169	0.001906	−0.005489	0.002007
MB (β_5)	−3.035	0.3047	−2.99	0.3093
IRS (β_6)	−0.3901	0.1866	−0.4113	0.1987
ϕ	1.089	0.1392		
ω[a]			0.9959	0.316
k			0.9805	0.07021
DIC	1486.47		1450.67	
Pearson χ^2	1009.406		1003.086	
MAD	6.919		7.014	
MSPE	209.909		231.476	

[a] $\omega = 1/\lambda^k$.

Table 5
Modeling results for the Texas data.

Variable	PG		PW	
	Value	S.E.	Value	S.E.
Intercept (ln β_0)	−6.3590	0.3907	−6.2280	0.4298
ADT (β_1)	0.9770	0.0424	0.9682	0.0430
LW (β_2)	−0.0532	0.0169	−0.0509	0.0163
SW (β_3)	−0.0100	0.0033	−0.0101	0.0033
CD (β_4)	0.0674	0.0121	0.0694	0.0122
ϕ	2.5510	0.2338		
ω[a]			1.0130	0.3068
k			1.6290	0.0817
DIC	4784.45		4771.61	
Pearson χ^2	1798.14		1869.32	
MAD	1.70		1.69	
MSPE	11.24		11.43	

[a] $\omega = 1/\lambda^k$.

models are also very similar. Though the PG and PW model have the same hierarchical structure, there is an extra parameter in the PW model. This is because the shape and scale parameters are assumed to be the same in the PG model, but not for the PW model. However, this extra parameter does not add any additional burden with regards to coding and running the PW model in WinBUGS (Spiegelhalter et al., 2003) compared to the PG model.

5.3. Texas data

Table 5 summarizes the modeling results for the Texas data. The functional form used is similar to the one used for Indiana dataset. Like the previous dataset, the segment length was considered as an offset in the model. All the coefficients are significant at 5% level. In general, all the estimated coefficients are logical and consistent with the existing literature. The increase in the traffic flow and the horizontal curve density increases the crash risk, whereas the increase in lane width and shoulder width decreases the risk.

6. Discussion

There are two relevant points that need to be discussed. First, the results using the simulated data have shown that the PW model was able to reproduce the "true" parameters. Second, it was shown that the PW GLM offers potential for modeling traffic crash data and performs as well as the PG model. As detailed in the modeling results for two observed datasets, both models provided similar GOF statistics. In addition, the computational time for both the models in WinBUGS is very similar.

Given the fact that the PW GLM has so far not been applied often, there are many lines of research activities that could be investigated. First, as discussed above, crash data can sometimes be subjected to very low sample mean values or very sample sizes. Consequently, the PG model do not perform well with such datasets since the inverse dispersion parameter can be significantly biased (or mis-estimated). This can negatively influence the standard errors associated with the model's coefficient. Therefore, the further work is needed on examining the stability of the PW model when the data are characterized by low sample mean and small sample size.

Second, since the empirical Bayes (EB) method is now frequently used in highway safety, an EB modeling framework should be development for the PW model. The approach used by Zou et al. (2013) to develop the EB framework for the Sichel model could be utilized. Third, the use of the PW GLM to identify hazardous sites should also be investigated. Fourth, even though the analysis was executed by assuming a fixed shape parameter (independent of covariates) in the Weibull distribution, further work should be accomplished to examine the effects of a variable-dependent shape parameter. Finally, a well-defined likelihood function and the related moments for the PW model should be built. This way, the MLE method could be used for estimating PW GLMs.

7. Summary and conclusions

This paper has described the application of the PW GLM for modeling traffic crashes. Although the PW GLM has been used in other fields, it has only been applied once in traffic safety (Maher and Mountain, 2009). The PW model was developed using a Bayesian framework and its general modeling performance was examined using the simulated dataset. The comparison of the PW and PG models was performed using two observed crash datasets. The comparison analysis was carried out using the most common functional form used by transportation safety analysts, which link crashes to the traffic flows and geometric variables. The results of this study show that the PW GLMs perform as well as the PG model in terms of GOF statistics. If the PW model performs well and offers stability when the data are characterized by low sample mean and small sample size values, then it could potentially offer a better alternative over the PG model for modeling motor vehicle crashes.

References

Anastasopoulos, P., Tarko, A., Mannering, F.L., 2008. Tobit analysis of vehicle accident rates on interstate highways. Accident Analysis and Prevention 40 (2), 768–775.

Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7 (4), 434–455.

Cameron, A.C., Trivedi, P.K., 1998. Regression Analysis of Count Data. Cambridge University Press, Cambridge, UK.

Geedipally, S.R., Lord, D., Dhavala, S.S., 2012. The negative binomial-Lindley generalized linear model: characteristics and application using crash data. Accident Analysis and Prevention 45 (2), 258–265.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequence. Statistical Science 7 (4), 457–511.

Hauer, E., 1997. Observational before-after studies in road safety: estimating the effect of highway and traffic engineering measures on road safety. Elsevier Science Ltd., Oxford.

Hilbe, J.M., 2011. Negative Binomial Regression, second ed. Cambridge University Press, Cambridge, UK.

Lord, D., 2006. Modeling motor vehicle crashes using Poisson–gamma models: examining the effects of low sample mean values and small sample size on estimation of the fixed dispersion parameter. Accident Analysis and Prevention 38 (4), 751–766.

Lord, D., Mannering, F.L., 2010. The statistical analysis of crash frequency data: a review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice 44 (5), 291–305.

Lord, D., Park, B.-J., 2010. Appendix D-negative binomial regression models and estimation methods. In: CrimeStat Version 3.3 Update Notes: Part 2: Regression Modeling, The National Institute of Justice, Washington, DC.

Lord, D., Geedipally, S.R., Persaud, B.N., Washington, S.P., Schalkwyk, I., Ivan, J.N., Lyon, C., Jonsson, T., 2008a. NCHRP Web-only document 126: methodology to predict the safety performance of rural multilane highways. <http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w126.pdf>.

Lord, D., Guikema, S.D., Geedipally, S.R., 2008b. Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes. Accident Analysis and Prevention 40 (3), 1123–1134.

Lord, D., Park, P.Y.-J., 2008. Investigating the effects of the fixed and varying dispersion parameters of Poisson–gamma models on empirical Bayes estimates. Accident Analysis and Prevention 40 (4), 1441–1457.

Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, Poisson–gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accident Analysis and Prevention 37 (1), 35–46.

Maher, M., Mountain, L., 2009. The sensitivity of estimates of regression to the mean. Accident Analysis and Prevention 41 (4), 861–868.

Maher, M.J., Summersgill, I., 1996. A comprehensive methodology for the fitting predictive accident models. Accident Analysis and Prevention 28 (3), 281–296.

Miaou, S.-P., Song, J.J., Mallick, B.K., 2003. Roadway traffic-crash mapping: a space-time modeling approach. Journal of Transportation and Statistics 6 (1), 33–57.

McCullagh, P., Nelder, J.A., 1983. Generalized Liner Models. Chapman and Hall, London, UK.

Miaou, S.-P., Lord, D., 2003. Modeling traffic-flow relationships at signalized intersection: dispersion parameter, functional form, and Bayes versus empirical Bayes. Transportation Research Record No.1840, 31–40.

Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. Accident Analysis and Prevention 39 (3), 459–468.

Oh, J., Lyon, C., Washington, S., Persaud, B., Bared, J., 2003. Validation of FHWA crash models for rural intersections: lessons learned. Transportation Research Record 1840, 41–49.

Phoenix, S.L., Ibnabdeljalil, M., Hui, C.-Y., 1997. Size effects in the distribution for strength of brittle matrix fibrous composites. International Journal of Solids Structures 34 (5), 545–568.

Raghavachari, M., Srinivasan, A., Sullo, P., 1997. Poisson mixture yield models for integrated circuits: a critical review. Microelectronics Reliability 37 (4), 565–580.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). Journal of the Statistical Society: Series B 64 (4), 583–639.

Spiegelhalter, D.J., Thomas, A., Best, N.G., Lun, D., 2003. WinBUGS Version 1.4.1 User Manual. MRC Biostatistics Unit, Cambridge. Available from: <http://www.mrcbsu. cam.ac.uk/bugs/welcome.shtml>.

Wagner, H.D., Steenbakkers, L.W., 1989. Stochastic strength and size effect in ultra-high strength polyethylene fibres. Philosophical Magazine Letters 59 (2), 77–85.

Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2011. Statistical and Econometric Methods for Transportation Data Analysis, second ed. Chapman and Hall/CRC, Boca Raton, FL.

Zhang, Y., Ye, Z., Lord, D., 2007. Estimating the dispersion parameter of the negative binomial distribution for analyzing crash data using a bootstrapped maximum likelihood method. Transportation Research Record 2019, 15–21.

Zou, Y., Lord, D., Zhang, Y., Peng, Y., 2013. Comparison of sichel and negative binomial models in estimating empirical Bayes estimates. Paper 13-2938. Presented at the 92nd Annual Meeting of the Transportation Research Board, 2013, Washington, D.C.