Risk Analysis, Vol. 30, No. 8, 2010 DOI: 10.1111/j.1539-6924.2010.01417.x

Extension of the Application of Conway-Maxwell-Poisson Models: Analyzing Traffic Crash Data Exhibiting Underdispersion

Dominique Lord,[1,*] Srinivas Reddy Geedipally,[1] and Seth D. Guikema[2]

The objective of this article is to evaluate the performance of the COM-Poisson GLM for analyzing crash data exhibiting underdispersion (when conditional on the mean). The COM-Poisson distribution, originally developed in 1962, has recently been reintroduced by statisticians for analyzing count data subjected to either over- or underdispersion. Over the last year, the COM-Poisson GLM has been evaluated in the context of crash data analysis and it has been shown that the model performs as well as the Poisson-gamma model for crash data exhibiting overdispersion. To accomplish the objective of this study, several COM-Poisson models were estimated using crash data collected at 162 railway-highway crossings in South Korea between 1998 and 2002. This data set has been shown to exhibit underdispersion when models linking crash data to various explanatory variables are estimated. The modeling results were compared to those produced from the Poisson and gamma probability models documented in a previous published study. The results of this research show that the COM-Poisson GLM can handle crash data when the modeling output shows signs of underdispersion. Finally, they also show that the model proposed in this study provides better statistical performance than the gamma probability and the traditional Poisson models, at least for this data set.

KEY WORDS: Conway-Maxwell-Poisson; gamma models; negative binomial models; regression models; underdispersion

1. INTRODUCTION

Over the last two decades, there has been a considerable amount of research performed on statistical methods for analyzing motor vehicle crashes.(1–10) Some of the most recent methods include the application of neural and Bayesian neural networks, latent class or mixture models, gamma probability, and support vector machine models, among others.(11–16) In highway safety, the traditional Poisson and mixed-Poisson models remain the most common probabilistic models utilized for analyzing crash data. This type of data has been found to often exhibit overdispersion (i.e., the variance is larger than the mean) and thus mixed-Poisson models (such as the Poisson-gamma or negative binomial) are generally preferred over the traditional Poisson model. Although very rare, crash data have sometimes shown characteristics of underdispersion (i.e., the variance is smaller than the mean value under an assumed probability model) especially in cases where the sample mean is very low.(15) It has been shown that Poisson and

[1]Zachry Department of Civil Engineering, Texas A&M University, 3136 TAMU, College Station, TX 77843-3136, USA.
[2]Department of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, MD 21218, USA.
*Address correspondence to Dominique Lord, Zachry Department of Civil Engineering, Texas A&M University, 3136 TAMU, College Station, TX 77843-3136, USA; tel: (979) 458-3949; fax: (979) 845-6481; d-lord@tamu.edu.

Poisson-gamma models experience important limitations when the sample mean value is low and the sample size is small.[17–20] In addition, other studies have demonstrated that both types of models have significant difficulties handling (or cannot handle) data characterized by underdispersion.[17,21] In the light of these limitations, several researchers have started examining the application of new and innovative methods for analyzing crash data.

As part of these new methods, the Conway-Maxwell-Poisson (COM-Poisson) distribution has very recently been reintroduced by statisticians for modeling count data that are characterized by either over- or underdispersion.[22–25] This distribution was first introduced in 1962, but has only been evaluated in the context of a generalized linear model (GLM) by Guikema and Coffelt, Lord *et al.*, and Sellers and Shmueli.[24,26–28] The COM-Poisson distribution has been used in many studies, such as analyzing word length, births, the prediction of purchase timing and quantity decisions, quarterly sales of clothing, Internet search engine visits, the timing of bid placement and extent of multiple bidding, modeling electric power system reliability, and motor vehicle crashes.[22,24,27,29–32] Only a handful of studies have applied the COM-Poisson distribution to observed or simulated data characterized by underdispersion.[23–25,28]

The objective of this article is to evaluate the performance of the COM-Poisson GLM for analyzing crash data exhibiting underdispersion, in cases where Poisson and Poisson-gamma models cannot be used. This article is continuation of the work done by Guikema and Coffelt and Geedipally *et al.* and Lord *et al.* on the COM-Poisson distribution and GLM.[24,25,27] To accomplish the objective of this study, several COM-Poisson models were estimated using crash data collected at 162 railway-highway crossings (RHX) in South Korea between 1998 and 2002. This data set has been identified as being characterized by underdispersion when the observations were modeled using regression methods.[15] To model such a data set, Oh *et al.* have proposed the gamma probability model.[15] The results of this study will show that the COM-Poisson GLM can handle crash data when the modeling output shows signs of underdispersion. The results also show that the model provides better statistical performance than the gamma probability and the traditional Poisson model.

The article is organized as follows. The next section provides a brief background on the COM-Poisson and gamma probability models. The third section describes the methodology used for estimating and comparing the various models. The fourth section presents the characteristics of the data used in this study. The fifth section summarizes the results of the parameter estimates and comparison analysis. The last section offers important concluding remarks as well as ideas for further research.

2. BACKGROUND

This section provides a brief description of the characteristics of the COM-Poisson and the gamma probability models, respectively.

2.1. COM-Poisson Model

The COM-Poisson distribution is a generalization of the Poisson distribution and was first introduced by Conway and Maxwell for modeling queues and service rates.[26] Shmueli *et al.* further elucidated the statistical properties of the COM-Poisson distribution using the formulation given by Conway and Maxwell, and Kadane *et al.* developed the conjugate distributions for the parameters of the COM-Poisson distribution.[22,23,26] Its probability mass function (PMF) can be given by Equations (1) and (2):

$$P(Y = y) = \frac{1}{Z(\lambda, \nu)} \frac{\lambda^y}{(y!)^\nu} \tag{1}$$

$$Z(\lambda, \nu) = \sum_{n=0}^{\infty} \frac{\lambda^n}{(n!)^\nu}, \tag{2}$$

where Y is a discrete count; λ is a centering parameter that is approximately the mean of the observations in many cases; and ν is defined as the shape parameter of the COM-Poisson distribution. The centering parameter λ is approximately the mean when ν is close to one; it differs substantially from the mean for small ν. Given that ν would be expected to be small for overdispersed data, this would make a COM model based on the original COM formulation difficult to interpret and use for overdispersed data.

To circumvent this problem, Guikema and Coffelt proposed a reparameterization of the COM-Poisson distribution by substituting $\mu = \lambda^{1/\nu}$ to provide a clear centering parameter.[24] This new formulation of the COM-Poisson is summarized in Equations (3) and (4).

$$P(Y = y) = \frac{1}{S(\mu, \nu)} \left(\frac{\mu^y}{y!} \right)^\nu \tag{3}$$

$$S(\mu, \nu) = \sum_{n=0}^{\infty} \left(\frac{\mu^n}{n!} \right)^\nu \tag{4}$$

The mean and variance of Y are given in terms of the new formulation as $E[Y] = \frac{1}{\nu} \frac{\partial \log S}{\partial \log \mu}$ and $V[Y] = \frac{1}{\nu^2} \frac{\partial^2 \log S}{\partial \log^2 \mu}$, with asymptotic approximations $E[Y] \approx \mu + 1/2\nu - 1/2$ and $\mathrm{Var}[Y] \approx \mu/\nu$ especially accurate once $\mu > 10$. With this new parameterization, the integral part of μ is now the mode, leaving μ as a reasonable centering parameter. The substitution $\mu = \lambda^{1/\nu}$ also allows ν to keep its role as a shape parameter. That is, if $\nu < 1$, the variance is greater than the mean, while $\nu > 1$ leads to underdispersion.

Guikema and Coffelt developed a COM-Poisson GLM framework using Bayesian framework in WinBUGS for modeling discrete count data.[24,33] Equations (5) and (6) describe this modeling framework. The framework is a dual-link GLM in which both the mean and the variance depend on the covariates. In Equations (5) and (6), x_i and z_j are covariates, and there are assumed to be p covariates used in the centering link function and q covariates used in the shape link function (similar to the varying dispersion parameter of the Poisson-gamma model proposed by Miaou and Lord, Hauer, and Heydecker and Wu[6,34,35]). The sets of parameters used in the two link functions do not necessarily have to be identical.

$$\ln(\mu) = \beta_0 + \sum_{i=1}^{p} \beta_i x_i \tag{5}$$

$$\ln(\nu) = \gamma_0 + \sum_{j=1}^{q} \gamma_j z_j \tag{6}$$

The GLM framework can model underdispersed data sets, overdispersed data sets, and data sets that contain intermingled underdispersed and overdispersed counts (for dual-link models only, since the dispersion characteristic is captured using the covariate-dependent shape parameter). The variance is allowed to depend on the covariate values, which can be important if high (or low) values of some covariates tend to be variance-decreasing while high (or low) values of other covariates tend to be variance-increasing. The parameters have a direct link to either the mean or the variance, providing insight into the behavior and driving factors in the problem, and the mean and variance of the predicted counts are readily approximated based on the covariate values and regression parameter estimates.

Recently, Sellers and Shmueli derived the likelihood function for the COM-Poisson GLM. This derivation greatly simplifies the estimation of the parameters of a COM GLM when full posteriors are not needed for the parameters, as opposed to the Bayesian estimating method.[24,28] However, by the time this article was prepared, the MLE formulation did not allow for a varying shape parameter, as described in Equation (6). The interested reader can find the MLE codes in R for the COM-Poisson GLM at: http://cran.r-project.org/web/packages/compoisson/index.html.[36]

2.2. Gamma Probability Model

The gamma probability distribution can be used for analyzing underdispersed and overdispersed data sets. Oh *et al.* used this distribution to analyze crashes collected at RHX in South Korea.[15] They found the gamma probability model to provide a good statistical fit for the railway-highway crossing crash data under study. The gamma probability model for the count data is given by:[15,37]

$$P(y_i = k) = Gamma(\alpha k, \lambda_i) - Gamma(\alpha k + \alpha, \lambda_i), \tag{7}$$

where λ_i is the mean of the crashes at ith site and is given by $\lambda_i = \exp(\mathbf{x}_i \beta)$:

$$Gamma(\alpha k, \lambda_i) = 1 \quad \text{if } k = 0; \tag{8a}$$

$$Gamma(\alpha k, \lambda_i) = \frac{1}{\Gamma(\alpha k)} \int_0^{\lambda_i} t^{\alpha k - 1} e^{-t} dt \quad \text{if } k > 0, \tag{8b}$$

where α is the dispersion parameter; for $\alpha > 1$, the model shows underdispersion; for $\alpha < 1$, the model exhibits overdispersion; for $\alpha = 1$, it is equidispersion, which means that the gamma model reduces to the Poisson model. It should be noted that the gamma probability model assumes a dual-state process, one for $y_i = 0$ and one for $y_i > 0$; this is the formulation of a hurdle function (note: $\lambda_i = 0$ since y_i is always equal to zero).[37] As discussed in previous studies, this kind of process may not be appropriate for analyzing crash data.[38,39]

3. METHODOLOGY

This section briefly describes the methodology used for comparing the different models. The same functional form used by Oh *et al.* was used for fitting all the models:[15]

$$\mu_i = \exp\left(\beta_0 + \beta_1 \ln(F_i) + \sum_{j=1}^{n} \beta_j x_j\right), \tag{9}$$

where μ_i is the mean number of crashes for site i; F_i is the average daily vehicle traffic on site i (vehicles/day); x_j is the estimated covariates, such as average daily railway traffic, detector distance, etc; and β_i is the estimated regression coefficients

Different methods were used for evaluating the goodness of fit (GOF) and predictive performance of the models. The methods used in this research include the following.

3.1. Akaike Information Criterion (AIC)

The AIC is a measure of the goodness of fit of an estimated statistical model and is defined as:[40]

$$\mathrm{AIC} = -2\log L + 2p, \tag{10}$$

where L is the maximized value of the likelihood function for the estimated model, and p is the number of parameters in the statistical model. The AIC methodology attempts to find the model that best explains the data with a minimum of free parameters and thus it penalizes models with a large number of parameters. The model with the lowest AIC is considered to be the best model among all available models.

3.2. Mean Prediction Bias (MPB)

MPB provides a measure of the magnitude and direction of the average model bias.[41] If the MPB is positive then the model overpredicts crashes and if the MPB is negative then the model underpredicts crashes. It is computed using the following equation:

$$\text{Mean Prediction Bias (MPB)} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i), \tag{11}$$

where n is the sample size, and $\hat{y}_i$ and y_i are the predicted and observed crashes at site i, respectively.

3.3. Mean Absolute Deviance (MAD)

MAD provides a measure of the average misprediction of the model.[41] The model closer to zero is considered to be the best among all the available models. It is computed using the following equation:

$$\text{Mean Absolute Deviance (MAD)} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|. \tag{12}$$

3.4. Mean Squared Predictive Error (MSPE)

MSPE is typically used to assess the error associated with a validation or external data set.[41] The model closer to zero is considered to be the best among all the available models. It can be computed using Equation (12):

$$\text{Mean Squared Predictive Error(MSPE)} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2. \tag{13}$$

Since the models estimated by Oh *et al.* were done using the likelihood method, the coefficients of the COM-Poisson GLMs were also estimated using the MLE code developed by Sellers and Shmueli.[15,28] This way, all the model comparisons are performed using the same estimation method.

4. DATA DESCRIPTION

This section provides an overview of the characteristics of the data set used in this study. This data set was previously used to develop Poisson and gamma probability models by Oh *et al.*[15] The characteristics of the data set used in this study are described in Table I. Only the variables that are found to be significant in this study are presented in Table I. For the characteristics of all other variables, the reader is referred to Table 2 of Oh *et al.*[15] It should be noted that looking at the raw observations, the crash data exhibit overdispersion (mean = 0.33, variance = 0.36). The underdispersion is in fact noticed when the observed values are modeled conditional on the mean, as described in the next section.

Fig. 1 shows the comparison between the actual crash data distribution with the values predicted by the Poisson distribution. This figure illustrates that the Poisson distribution predicts almost the same number of crashes as the actual data.

Table I. Summary Statistics of the Data Set

Variables		Min.	Max.	Average (*SD*)	Frequency
Crashes		0	3	0.33 (0.60)	162
AADT		10	61199	4617 (10391.57)	162
Average daily railway traffic		32	203	70.29 (37.34)	162
Presence of a commercial area	1 (yes)	–	–	–	149 (91.98%)
	0 (no)	–	–	–	13 (8.02%)
Train detector distance		0	1329	824.5 (328.38)	162
Time duration between the activation of warning signals and gates		0	232	25.46 (25.71)	162
Presence of a speed hump	1 (yes)	–	–	–	134 (82.72%)
	0 (no)	–	–	–	28 (17.28%)
Presence of a track circuit controller	1 (yes)	–	–	–	113 (69.75%)
	0 (no)	–	–	–	49 (30.25%)
Presence of a guide	1 (yes)	–	–	–	126 (77.78%)
	0 (no)	–	–	–	36 (22.22%)

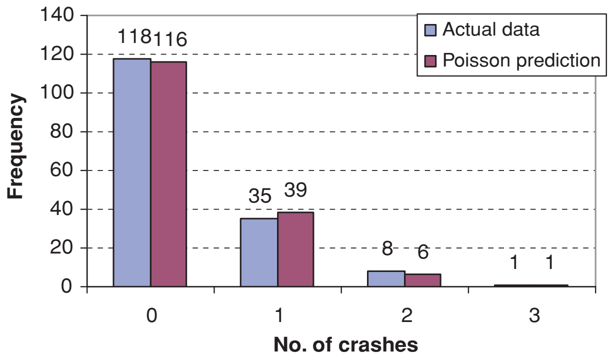


Fig. 1. Observed crash data versus values estimated using the Poisson distribution.

5. RESULTS

This section describes the results of the analysis. Several models were estimated using the variables documented in Oh *et al.*[15] To evaluate the characteristics of the variance function, Poisson-gamma models were first estimated using the six variables that were reported to be significant by the gamma probability model in the original study.[15] Figs. 2 and 3 show the output of the Poisson-gamma models for the MLE and Bayesian estimating methods, respectively. The Bayesian estimating method was used to confirm the results of the MLE method for this particular data set. For the MLE, Fig. 2 illustrates that the Poisson-gamma model cannot handle the data very well, as determined by the negative value of the dispersion parameter and its confidence interval (e.g., $\mathrm{Var}\,(Y) = \mu + \alpha\mu^2$, where $\alpha =$ the dispersion parameter of the Poisson-gamma model). In addition, the model provides unreliable parameter estimates, since most of the variables are not significant at the 10% level. For the Bayesian model, Fig. 3 shows that the inverse dispersion parameter of the Poisson-gamma model becomes unstable and tends toward infinity (i.e., reverts back to a Poisson model), both when vague ($\phi \sim \mathit{gamma}(0.01, 0.01)$) and less-vague hyperpriors ($\phi \sim \mathit{gamma}(0.2, 0.1)$) are used. Similar to the MLE, most of the parameter estimates were not significant at the 10% level. In sum, these plots confirm that the modeling results are characterized by underdispersion when the model is estimated using the six original explanatory variables or covariates.

In the initial step, a COM-Poisson model was estimated with all 31 explanatory variables documented in Table 2 of Oh *et al.*[15] Only six variables were found to be significant at the 10% confidence level. In the subsequent step, a Poisson model was estimated, also in R, by considering all 31 variables to compare this model with the COM-Poisson model in identifying significant variables and assess its statistical fit. Then, the final Poisson and COM-Poisson models were compared with the gamma probability model in Oh *et al.*[15]

The comparison analysis related to the significant variables is summarized in Table II. As seen in this table, each model gives different significant variables. The variables annual average daily traffic ($AADT$), *presence of a commercial area*, *train detector distance*, and *presence of a speed hump* are significant in all the three models. The *presence of a guide* is significant only in the COM-Poisson model, while the *presence of a track circuit controller* is significant in both the Poisson and COM-Poisson models. The variables *average daily railway traffic* and *warning time duration* are only significant in the gamma probability model.

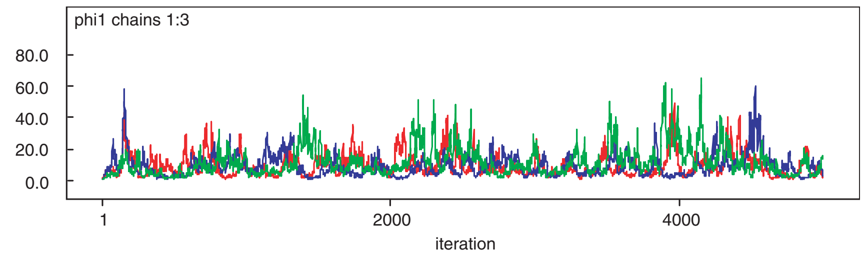
WARNING: Negative of Hessian not positive definite.

Analysis Of Parameter Estimates

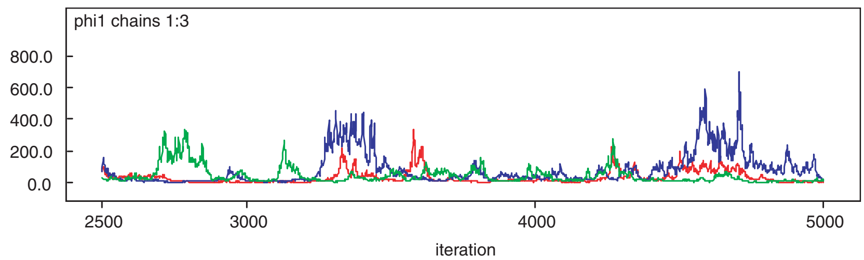
Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-2.7048	0.6447	-3.9683	-1.4413	17.60	<.0001
log_AADT_	1	0.1518	0.0636	0.0271	0.2765	5.69	0.0171
Railway_traffic	1	0.0027	0.0027	-0.0027	0.0080	0.95	0.3292
Presence_of_comm_are	1	0.7252	0.3188	0.1004	1.3501	5.17	0.0229
Distance_of_train_de	1	0.0005	0.0004	-0.0002	0.0012	2.00	0.1568
Warning_time_differe	1	0.0052	0.0021	0.0012	0.0092	6.36	0.0117
Presence_of_speed_hu	1	-0.4006	0.3190	-1.0258	0.2245	1.58	0.2091
Dispersion	0	-0.3333	0.0000	-0.3333	-0.3333		

NOTE: The negative binomial dispersion parameter was estimated by maximum likelihood.

Fig. 2. SAS output of the Poisson-gamma model.



(a) Nonvague hyperprior $\phi \sim gamma(0.2,0.1)$; $\phi = 10.26\,(8.10)$



(b) Vague hyperprior $\phi \sim gamma(0.01,0.01)$; $\phi = 50.78\,(70.78)$

Fig. 3. WinBUGS output (history for inverse dispersion parameter) of the Poisson-gamma model.

Table III compares the parameter estimates and the associated standard errors for the three models. Different measures of GOF are also presented in Table III to compare the models' performance. As a rule of thumb, if the difference in the AIC value is less than 10, then the models' fit is assumed to be not significantly different from one another. Thus, using the AIC values, the COM-Poisson is not significantly different from Poisson and gamma probability model. The MPB values, on the other hand, show that the COM-Poisson slightly underpredicts crashes whereas the Poisson model slightly overpredicts crashes. The value of MPB for the gamma probability model shows that it highly overpredicts crashes.

Table III also shows that the shape parameter of COM-Poisson and gamma probability model clearly indicates the modeling output exhibits underdispersion. Thus, the Poisson model cannot be used for this data set, even if it fits the data relatively well. Although fitting a Poisson model to such data will not significantly influence the mean of the regression coefficients, it will have a significant effect on the standard errors. This can be seen in the comparison of the standard errors for each model. The Poisson model underestimates the standard errors and in turn produces inflated t-values and confidence intervals of the coefficients. It is important to note that the coefficients estimated for the COM-Poisson model in Table III are for centering parameter "λ" and not for

Table II. Significant Variables with Three Different Distributions

Variables	COM-Poisson	Poisson	Gamma
AADT	√	√	√
Average daily railway traffic	–	–	√
Presence of a commercial area	√	√	√
Train detector distance	√	√	√
Time duration between the activation of warning signals and gates	–	–	√
Presence of a track circuit controller	√	√	–
Presence of a guide	√	–	–
Presence of a speed hump	√	√	√

√ – Significant at 10% confidence level.

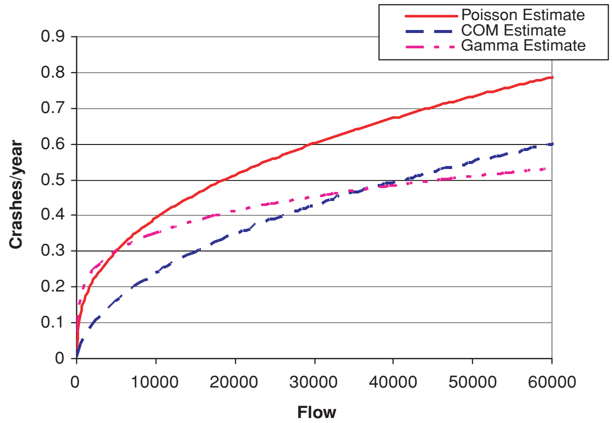


Fig. 4. Estimated values for gamma, Poisson, and COM-Poisson GLMs.

mean "$E[Y]$," as in the case of Poisson and gamma model. Finally, the MAD and MSPE values in Table III show that the COM-Poisson model fits the data much better than the gamma model.

Fig. 4 shows the estimated number of crashes for the Poisson, COM-Poisson, and gamma models as a function vehicular traffic flow. These values were estimated at the average value of all other continuous variables that are significant in the three models. Thus, a direct comparison between the curves should not be conducted, since each predicted value is estimated using different input variables. Fig. 4 illustrates that the trends shown by the Poisson and COM-Poisson models (i.e., the shape of the curve) are very similar, but are much different than the trend shown by the gamma model. The number of crashes predicted by the gamma model sharply increases with the increase in traffic flow until 2,000 and the rate per unit of exposure decreases significantly for larger flows.

6. SUMMARY AND CONCLUSIONS

The objective of this article was to evaluate the performance of the COM-Poisson GLM for analyzing crash data exhibiting underdispersion (when conditional on the mean). The COM-Poisson distribution, originally developed in 1962, has recently been reintroduced by statisticians for analyzing count data subjected to either over- or underdispersion. Over the last year, the COM-Poisson GLM has been evaluated in the context of crash data analysis and it has been shown that the model performs as well as

Table III. Parameter Estimates with Three Different Distributions (MLE)

Variables	COM-Poisson	Poisson	Gamma
Constant	−6.657 (1.206)[a]	−5.326 (0.906)[a]	−3.438 (1.008)[a]
Ln(ADT)	0.648 (0.139)	0.388 (0.076)	0.230 (0.076)
Average daily railway traffic	–	–	0.004 (0.0024)
Presence of commercial area	1.474 (0.513)	1.109 (0.367)	0.651 (0.287)
Train detector distance	0.0021 (0.0007)	0.0019 (0.0006)	0.001 (0.0004)
Time duration between the activation of warning signals and gates	–	–	0.004 (0.002)
Presence of track circuit controller	−1.305 (0.431)	−0.826 (0.335)	–
Presence of guide	−0.998 (0.512)	–	–
Presence of speed hump	−1.495 (0.531)	−1.033 (0.421)	−1.58 (0.859)
Shape parameter (ν_0)	2.349 (0.634)	–	2.062 (0.758)
AIC	210.70	196.55	211.38
MPB	−0.007	0.004	0.179
MAD	0.348	0.359	0.459
MSPE	0.236	0.252	0.308

[a]Standard error.

the Poisson-gamma model for crash data exhibiting overdispersion. To accomplish the objective of this study, several COM-Poisson models were estimated using crash data collected at 162 railway-highway crossings in South Korea between 1998 and 2002. This data set has been shown to exhibit underdispersion when models linking crash data to various explanatory variables are estimated. The modeling results were compared to those produced from the Poisson and gamma probability models documented in Oh *et al.*[15]

The results of this study show that the COM-Poisson GLM can handle crash data when the modeling output shows signs of underdispersion. They also show that the model analyzed in this study provides better statistical performance than the gamma probability and the traditional Poisson models, at least for this data set. Similarly, the COM-Poisson GLM offers a more defensible approach than the gamma probability model, since the former does not assume that the observed data follow a dual-state generating process, one of which has a long-term equal to zero. Although interesting results were found in this study, further research should be conducted using different data sets to confirm those found in this research. The proposed analysis should also include looking at the robustness of the COM-Poisson GLM for data characterized by very low sample values and small sample size, since underdispersion is usually observed for data characterized by these two conditions (see some interesting findings in Geedipally *et al.*[25]). Finally, COM-Poisson models with a dual-link should be estimated for data sets exhibiting similar characteristics as the one used in this research. Given the changes in the nature of the variance function when variables are included or excluded, it is possible that such data sets could contain intermingled over- and underdispersed counts.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Simon Washington from Arizona State University and Dr. Jutaek Oh from the Korean Transport Institute for providing the data used in their study. Their answers to our numerous questions are also greatly appreciated.

REFERENCES

1. Abdel-Aty M, Addella MF. Linking roadway geometrics and real-time traffic characteristics to model daytime freeway crashes: Generalized estimating equations for correlated data. Transportation Research Record 1897, 2004; 106–115.
2. Maycock G, Hall RD. Accidents at 4-arm roundabouts. Crowthorne, Berks, UK: Transport Research Laboratory, Laboratory Report LR 1120, 1984.
3. Miaou S-P. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis & Prevention, 1994; 26(4):471–482.
4. Maher MJ, Summersgill IA. Comprehensive methodology for the fitting of predictive accident models. Accident Analysis & Prevention, 1996; 28(3):281–296.
5. Poch M, Mannering FL. Negative binomial analysis of intersection accident frequency. Journal of Transportation Engineering, 1996; 122(2):105–113.
6. Miaou S-P, Lord D. Modeling traffic crash-flow relationships for intersections: Dispersion parameter, functional form, and full Bayes versus empirical Bayes. Transportation Research Record 1840, 2003; 31–40.
7. Lord D, Manar A, Vizioli A. Modeling crash-flow-density and crash-flow-V/C ratio for rural and urban freeway segments. Accident Analysis & Prevention, 2005; 37(1):185–199.
8. Lord D, Bonneson JA. Development of accident modification factors for rural frontage road segments in Texas. Transportation Research Record 2023, 2007; 20–27.
9. Malyshkina NV, Mannering FL, Tarko AP. Markov switching negative binomial models: An application to vehicle accident frequencies. Accident Analysis & Prevention, 2009; 41(2):217–226.
10. Lord D, Mannering F. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transport Res Part A, 2010, doi:10.1016/j.tra.2010.02.001.
11. Abdelwahab HT, Abdel-Aty M. Artificial neural networks and logit models for traffic safety analysis of toll plazas. Transportation Research Record 1784, 2002; 115–125.
12. Xie Y, Lord D, Zhang Y. Predicting motor vehicle collisions using Bayesian neural networks: An empirical analysis. Accident Analysis & Prevention, 2008; 39(5):922–933.
13. Depaire B, Wets G, Vanhoof K. Traffic accident segmentation by means of latent class clustering. Accident Analysis & Prevention, 2008; 40(4):1257–1266.
14. Park BJ, Lord D. Application of finite mixture models for vehicle crash data analysis. Accident Analysis & Prevention, 2009; 41(4):683–691.
15. Oh J, Washington SP, Nam D. Accident prediction model for railway-highway interfaces. Accident Analysis and Prevention, 2006; 38(2):346–356.
16. Li X, Lord D, Zhang Y, Xie Y. Predicting motor vehicle crashes using support vector machine models. Accident Analysis & Prevention, 2008; 40(4):1611–1618.
17. Clark SJ, Perry JN. Estimation of the negative binomial parameter κ by maximum quasi-likelihood. Biometrics, 1989; 45:309–316.
18. Piegorsch WW. Maximum likelihood estimation for the negative binomial dispersion parameter. Biometrics, 1990; 46:863–867.
19. Dean CB. Modified pseudo-likelihood estimator of the overdispersion parameter in poisson mixture models. Journal of Applied Statistics, 1994; 21(6):523–532.
20. Lord D. Modeling motor vehicle crashes using Poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. Accident Analysis & Prevention, 2006; 38(4):751–766.
21. Saha K, Paul S. Bias-corrected maximum likelihood estimator of the negative binomial dispersion parameter. Biometrics, 2005; 61(1):179–185.
22. Shmueli G, Minka TP, Kadane JB, Borle S, Boatwright P. A useful distribution for fitting discrete data: Revival of the

Conway-Maxwell-Poisson distribution. Journal of the Royal Statistical Society: Part C, 2005; 54:127–142.
23. Kadane JB, Shmueli G, Minka TP, Borle S, Boatwright P. Conjugate analysis of the Conway-Maxwell-Poisson distribution. Bayesian Analysis, 2006; 1:363–374.
24. Guikema SD, Coffelt J. A flexible count data regression model for risk analysis. Risk Analysis, 2008; 28(1):213–223.
25. Geedipally S, Guikema SD, Dhavala S, Lord D. Characterizing the performance of a Bayesian Conway-Maxwell Poisson GLM. Presented at the 2008 Joint Statistical Meetings, Denver, CO, August 3–7, 2008.
26. Conway RW, Maxwell WL. A queuing model with state dependent service rates. Journal of Industrial Engineering, 1962; 12:132–136.
27. Lord D, Guikema SD, Geedipally SR. Application of the Conway-Maxwell-Poisson generalized linear model for analyzing motor vehicle crashes. Accident Analysis & Prevention, 2008; 40(3):1123–1134.
28. Sellers KF, Shmueli G. A flexible regression model for count data. Annals of Applied Statistics, 2010, in press. Available at: http://imstat.org/aoas/next_issue.html.
29. Ridout MS, Besbeas P. An empirical model for underdispersed count data. Statistical Modelling, 2004; 4:77–89.
30. Boatwright P, Borle S, Kadane JB. A model of the joint distribution of purchase quantity and timing. Journal of the American Statistical Association, 2003; 98:564–572.
31. Telang R, Boatwright P, Mukhopadhyay TA. Mixture model for Internet search-engine visits. Journal of Marketing, 2004; 41(5):206–214.
32. Borle S, Boatwright P, Kadane JB. The timing of bid placement and extent of multiple bidding: An empirical investigation using eBay online auctions. Statistical Science, 2006; 21(2):194–205.
33. Spiegelhalter DJ, Thomas A, Best NG, Lun D. WinBUGS Version 1.4.1 User Manual. Cambridge: MRC Biostatistics Unit, 2003. Available from: <http://www.mrcbsu.cam.ac.uk/bugs/welcome.shtml>.
34. Hauer E. Overdispersion in modelling accidents on road sections and in empirical Bayes estimation. Accident Analysis & Prevention, 2001; 33(6):799–808.
35. Heydecker BG, Wu J. Identification of sites for road accident remedial work by Bayesien statistical methods: An example of uncertain inference. Advances in Engineering Software, 2001; 32:859–869.
36. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2006. Available at: http://www.Rproject.org. Accessed July 2009.
37. Cameron AC, Trivedi PK. Regression Analysis of Count Data. Econometric. Cambridge University Press, Boston Society Monograph No. 30, 1998.
38. Lord D, Washington SP, Ivan JN. Poisson, Poisson-gamma and zero inflated regression models of motor vehicle crashes: Balancing statistical fit and theory. Accident Analysis & Prevention, 2005; 37(1):35–46.
39. Lord D, Washington SP, Ivan JN. Further notes on the application of zero inflated models in highway safety. Accident Analysis & Prevention, 2007; 39(1):53–57.
40. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 1974; 19(6):716–772.
41. Oh J, Lyon C, Washington SP, Persaud BN, Bared J. Validation of the FHWA crash models for rural intersections: Lessons learned. Transportation Research Record 1840, 2003; 41–49.