



Research Article



Investigating the Factors Affecting the Severity of Single-Vehicle Crashes on Urban Roads using Bayesian Binary Probit Regression

Mohammad Rahmaninezhad Asil*, Iraj Bargegol

Department of Civil Engineering, Faculty of Engineering, University of Guilan, Rasht, Iran

Keywords

Crash Severity,
Single-vehicle Crashes,
Urban Roads,
Bayesian Binary Probit
Regression.

Abstract

Traffic crashes are a global challenge, and many people lose their lives in crashes every year. Single-vehicle accidents have high injury and death rates and are therefore a major concern. In this study, the effect of various factors on the severity of this type of crash on urban roads of Rasht, in Iran, has been investigated. Rasht is the capital of Guilan province and is one of the cities that many tourists visit every year. Investigating Single-vehicle crashes and predicting the factors affecting them facilitates the possibility of prevention and provision of medical facilities. In this paper, a Bayesian Binary Probit Regression model was used for this purpose. The results of this study showed that Single-vehicle crashes during non-peak traffic hours, which are from 21 to 6 o'clock, increase the severity of the collision and also the results stated that spring and summer seasons, Iranian-made passenger cars and motorcycles, male drivers, and inability to control vehicles increase the severity of a single-vehicle crash.

1. Introduction

Annual traffic accidents cause a lot of human and financial losses to countries [1]. According to the World Health Organization, traffic accidents are the eighth leading cause of death globally, which is the seventh leading cause by 2030 if the current trend continues [2]. The organization states that about 1.25 million people worldwide die annually due to traffic crashes, and the injuries caused by these crashes are the leading cause of death for people aged 15 to 29 [3, 4]. According to World Health Organization estimates, 32.1 per 100,000 people die in road crashes in Iran, compared to 10.9 per 100,000 for the United States and 5.4 per 100,000 for Australia and Austria. Also, based on Iran Legal Medicine Organization reports [5], an average of 16,680 people have been killed and 323,000 injured in traffic crashes in Iran in the last four years. As the capital of Guilan province, the city of Rasht is one of the touristic provinces of northern Iran,

with over 5 million passengers entering this province annually [6, 7]. According to Iran Legal Medicine Organization reports, traffic crashes in Rasht have killed 122 people and injured 8,676 in the past four years. Single-vehicle crashes account for 4.5% of all crashes, 2% of which result in death and 46.2% of which result in injury, while other crashes, which are 95.5%, lead to 0.7% died and 59% were injured. According to the information, this type of accident is one of the most dangerous types of crashes due to its higher possibility of injury severity. In this research, the factors affecting the severity of single-vehicle crashes have been investigated [8, 9].

2. Literature Review

Many studies have been conducted on the severity of single-vehicle crashes over the years. In 2010, Soyoung Jung et al. assessed the effects of rainfall on the severity of Wisconsin interstate highways single-vehicle crashes. To

* Corresponding Author: Mohammad Rahmaninezhad Asil
E-mail address: mo_rahmaninejad@msc.guilan.ac.ir

Received: 16 November 2021; Revised: 15 December 2021; Accepted: 14 January 2022
<https://doi.org/10.52547/crpase.8.2224>

Academic Editor: **Mahdi Feizbahr**

Please cite this article as: M.Rahmaninezhad Asil, I.Bargegol, Investigating the Factors Affecting the Severity of Single-Vehicle Crashes on Urban Roads using Bayesian Binary Probit Regression, Computational Research Progress in Applied Science & Engineering, CRPASE: Transactions of Civil and Environmental Engineering 8 (2022) 1–6, Article ID: 2224.

evaluate the influential factors in this study, the logistic regression method was used, which showed that rainfall intensity, wind speed, road curvature, driver gender, and seat belt use have significant increasing effects on the severity of the crashes [10]. In 2019, Qinzhong Hou et al. examined the severity of drivers' injuries in single-vehicle crashes on freeways in Heilongjiang Province. To analyze the severity of crashes, he has used the Multinomial Logit Model. Results showed that female and elderly drivers, a night without lighting and fixed object collisions, increase the severity of crashes and also stated that increasing driving experience reduces the severity of crashes [11]. In 2019, Mo Zhou et al. investigated single-vehicle crashes due to the higher severity of these sorts of crashes compared to other types. To analyze the results, the Probit Model has been used. The results showed that old drivers (aged 65 years and older) using alcohol and improper maneuvers on horizontal curves increase the severity of crashes [12]. In 2019, Zhenning Li et al. examined the factors influencing the severity of single-vehicle crashes in New Mexico City. This research has been done using the mixed logit model, which showed that urban roads, wet pavement surface, overtaking, refusal of seat belts, and drivers over 65 years of age adversely affect the severity of crashes [13].

Many researchers have used Bayesian Statistics to model crash severity in recent years. In 2006, Matthieu de Lapparent studied the severity of motorcycle crashes in urban areas. In this study, the Empirical Bayesian method was used for the purpose of analysis, and the results showed that women motorcyclists between 30 and 50 years old and use high-powered motorcycles are more susceptible to severe injuries [14]. In 2018, Mohamed M. Ahmed et al. investigated the impact of various factors on the severity of heavy vehicle crashes on rural roads. They used Bayesian Binary Logit Model for their analysis, and the results showed that in the presence of a heavy vehicle, the probability of severe collision on state and interstate highways increases by 2.3 and 4.5 times, respectively. Also, the crash severity increases significantly in icy and snowy surface conditions [15]. In 2020, Quan Yuan et al. used Bayesian Bivariate Probit Regression to investigate the factors influencing the severity of vehicle crashes at intersections. The results showed that the type of vehicle and daylight crashes increase the probability of injury, and also the number of lanes and the speed of the vehicle latently affect the severity of the accident [16]. In 2014, Rongjie Yu used the hierarchical Bayesian binary probit model to investigate the factors affecting the severity of crashes on freeways. The results of this study showed that high-speed changes in the time just before the accident cause more severe crashes. Also, this study showed that the Bayesian binary probit model performs better than the binary probit model [17]. In 2017, Xuecai Xu et al. examined the severity of pedestrian crashes at intersections using the Bayesian spatial logit model. Most injuries to pedestrians are pedestrians over 65 years of age, head injuries during an accident, sidewalk crashes, and crashes near a tram or light rail stations [18]. In 2019, Qiang Zeng et al. analyzed the severity of freeway crashes using the Bayesian spatial generalized ordered logit model, a study conducted on the Kaiyang Freeway in China. The results of this study showed that driver experience, season, traffic

volume, time to respond to emergency medical services, and the type of accident have significant effects on the severity of the accident. In addition, the type of vehicle, season, time of day, weather conditions, the vertical grade of the road, bridge, traffic volume, and type of accident have significant effects on moderate and severe crashes [19].

In 2019, Lian Zhu et al. analyzed the severity of curved road crashes using Bayesian Network Analysis. The results of this analysis showed that the crashes that occur on curved roads are more severe than the crashes on straight roads. In this study, the role of factors such as driver behavior, road characteristics, and factors related to vehicles and environmental characteristics have been studied. The results stated factors such as point of impact, site of location, accident side of road, and alcohol/drugs condition. They are very important in crashes in horizontal curves [20]. In 2018, Zhenning Li et al. examined the severity of driver injuries at intersections, using cluster analysis and hierarchical Bayesian models. This study showed that the time of the accident, weather, light condition, type of road, type of vehicle, driver age, driver use of drugs and alcohol, use of seat belts, and factors such as left-turn affect the severity of crashes [21]. In 2018, Qixuan Chen et al. investigated the role of using the drug in fatal crashes in the United States. They used the Bayesian multilevel model to analyze the results. The results of this study showed that older drivers, women, drivers who have used seat belts in crashes, have valid licenses, and drivers of heavy vehicles are less likely to have a severe accident due to drug use. The results of this study also showed that the relationship between alcohol and drug consumption is the opposite relationship [22]. In 2020, Qiang Zeng et al. examined the role of weather conditions on the severity of freeway crashes. They used the Bayesian spatial generalized ordered logit model for modeling. The results of this study stated that with increasing rainfall, the probability of light and severe crashes decreases, and the probability of medium crashes increases. This study also shows that factors such as vehicle type, driving skills, accident time, type of accident, medical service reaction time, and the presence of vertical and horizontal curves affect the severity of the accident [23].

3. Research Method

In statistical inference, there are two methods for data analysis: the frequentist and Bayesian approaches. The use of Bayesian statistics has been increased during the past decades due to the significant progress in the theory and application of Bayesian methods.

3.1. Bayesian Perception

Bayesian inference is a probabilistic inferential method and dates back to Thomas Bayes (1764). In the last two decades, Bayesian inference has become more common due to computational power and significant advances in Monte Carlo Markov Chain (MCMC) methods for approximating high-dimensional integrals. In a Bayesian view, no fundamental distinction between observations and statistical model parameters is considered. To determine a Bayesian Model, suppose X and θ represent the observed data and model parameters, respectively. In Bayesian inference, a

possible probabilistic distribution of $P(X, \theta)$ must be determined on all random quantities. This combined distribution consists of two parts: the anterior distribution $P(\theta)$ and the likelihood function $P(X|\theta)$. Specifying these two parts leads to the following possible model:

$$P(X,\theta)=P(X|\theta)P(\theta) \tag{1}$$

Once the observed data X is known, and the Bayes theorem is applied, the distribution of θ under the condition X is determined as follows:

$$P(\theta|X) = \frac{P(X, \theta)}{P(\theta)} = \frac{P(X|\theta)P(\theta)}{\int P(X|\theta)P(\theta) d\theta} \tag{2}$$

$P(X|\theta)$ is called the posterior distribution θ , and the goal of all Bayesian perceptions is this posterior distribution. Any posterior distribution property such as torques, multiples, etc., is acceptable for Bayesian perception. As mentioned earlier, Bayesian theory is different from classical statistical theory in that it considers each unknown parameter as a random variable with a previous distribution. Prior distributions specify initial information about the θ parameter [24].

3.2. Probit Model

The Probit model, or unit probability, has a long history. With this method proposed in the 1930s, the probability of a unit value of a variable can be estimated. These probabilities are estimates of coefficients. According to a Probit Model, converting a linear sequence of independent variables to a cumulative probability value is a normal distribution. The ranked Probit Model is expressed as follows:

$$y_i^x = x_i'\beta + u_i \quad U_i \sim NID(0,1) \tag{3}$$

$$y = \begin{cases} 0 & \text{if } y_i^* < y_1 \\ 1 & \text{if } y_1 \leq y_i^* < y_2 \\ 2 & \text{if } y_2 \leq y_i^* \end{cases} \tag{4}$$

4. Results

4.1. Descriptive Analysis

In this research, single-vehicle crashes from 2015 to 2020 have been studied. The total number of crashes is 620 records, the information of which has been obtained from KAM-114 from the traffic department of Urban police in Rasht. Crashes are divided into two categories of no injury and injury/fatal according to the most severe injuries suffered by the driver or passengers. In this study, the impact

of 11 factors, the frequency of which is shown in Table 1, has been investigated. The frequency of crashes indicates that typically 100 single-vehicle crashes occur each year, but in 2020 the number of single-vehicle crashes decreased due to the COVID-19 epidemic. The distribution of accidents during hours indicates that the crashes were more severe between 21 and 6 in the morning. Due to the monthly distribution of crashes, collisions in the first six months of the year have higher injury severities despite the smaller number of crashes [25]. The frequency of crashes in terms of days of the week shows that crashes are more severe on weekends. This can be due to a decrease in traffic volume on weekends. Also, the frequency of crashes shows that most crashes occur in dry pavements and clear weather, and crashes in night conditions with light have the most injuries. One of the influential factors in crashes is the driver's personal characteristics. The frequency of crashes shows that the native driver, men, and drivers under 30 years old experience more injuries [26].

4.2. Results of Bayesian Binary Probit Model

In this research, 11 independent variables are introduced into the model as 20 sub-binary variables. In Bayesian models, determining the appropriate prior distribution can improve the model's performance. Table 2 shows the different previous distributions. Different models have been constructed using these distributions, and Deviance Information Criteria (DIC), Maximum Likelihood (ML), Bayesian Factor (BF) criteria have been calculated for each of them. The BF criterion is suitable for comparing models with the same variable combination, so this criterion has been used to find the best prior distribution. According to Table 2, the BF for the Flat distribution is 66, which is the highest number compared to the other distributions. In this study, the flat distribution, a non-informative distribution, was considered the prior distribution [27, 28].

Table 3 shows the model results. The final set of variables is chosen using backward elimination which is used in [29, 30]. The significance criterion for each variable is the lower and upper confidence interval limit. So that any variable whose lower and the upper confidence interval is also a sign of significance. Accordingly, non-peak hours, spring and summer, Iranian passenger vehicles and motorcycles, men, and the inability to control vehicles variables significantly impact the severity of the crash.

Table 1. Summary of Variables.

Variables	Category	Accident severity		Total accidents
		No injury accidents	Injury/Fatal accidents	
Day of week	Saturday	38	52	98
	Sunday	31	36	95
	Monday	61	34	90
	Tuesday	53	36	67
	Wednesday	46	46	92
	Thursday	47	42	89
	Friday	44	54	92
Season	Spring	70	78	148
	Summer	58	80	138
	Fall	105	67	172
	Winter	87	75	162
Year	2016	58	51	109
	2017	30	47	77
	2018	42	43	85
	2019	68	60	128
	2020	34	29	63
	2021	88	70	158
Surface Condition	Dry	258	238	496
	Wet	62	62	124
Weather Condition	Clear	243	223	466
	Cloudy	26	23	49
	Rainy	47	54	101
	Snowy	4	0	4
Light condition	daylight	182	125	307
	dark- lighted	127	149	276
	dark-not-lighted	11	26	37
Driver familiar	Familiar	263	251	514
	Unfamiliar	57	49	106
Vehicle type	Passenger Car–made in Iran (PCIR)	149	205	354
	Passenger Car–made in other countries (PCOC)	136	29	165
	Motorcycle	1	56	57
	Pickup	11	6	17
	Heavy Vehicle	23	4	27
Driver Age	< 30	100	115	215
	31-55	197	166	363
	>56	23	19	42
Driver Gender	Male	260	274	534
	Female	57	49	106

Table 2. Comparison of different prior distributions for model

Prior Distribution	DIC	Log (ML)	Log (BF)
Normal (0,10000)	733.3379	-432.9827	.
Normal (0,1000)	733.6469	-419.1988	13.78394
Normal (0,100)	733.993	-405.4671	27.51558
Normal (0,10)	733.679	-392.0635	40.91925
Normal (0,5)	733.092	-389.0813	43.90143
Normal (0,2)	733.0816	-385.6283	47.35439
Normal (0,1)	734.3161	-384.356	48.62672
Normal (0,0.5)	736.6612	-384.4305	48.55225
Uniform (-100,100)	733.5958	-430.2519	2.73079
Uniform (-10,10)	733.6672	-402.3161	30.66665
Flat	734.0873	-366.4287	66.55402

Table 3. Parameters Estimates.

Variables	Mean	Std. dev.	MCSE	Median	Equal-tailed	
					(95% cred. interval)	
Off-peak hours (21-6)	0.2676483	0.1150401	0.002483	0.2687438	0.039633	0.4901357
Spring	0.2754078	0.1365798	0.003071	0.2742052	0.0127428	0.5470013
Summer	0.435046	0.1417883	0.002987	0.4359112	0.154654	0.7089009
Dry	-0.1034311	0.1390183	0.006135	-0.1009687	-0.3775307	0.1653047
Dark-not-lighted	0.3929348	0.2391784	0.003285	0.3913925	-0.0667556	0.8626146
Cloudy	-0.0194693	0.2057808	0.005183	-0.0187105	-0.4202817	0.3806639
Motorcycle	3.225784	0.5497333	0.034969	3.186038	2.222183	4.40923
PCIR	0.9168676	0.3106742	0.029194	0.9143613	0.3091293	1.524991
Young driver (<30)	0.0490023	0.1176927	0.002002	0.048949	-0.1816533	0.2802509
Male	0.4090919	0.1629534	0.006946	0.4094307	0.0855972	0.7295428
Loosing vehicle's control	0.5206289	0.1208408	0.003107	0.5193659	0.2884966	0.7603037
Constant	-2.004139	0.3663015	0.035413	-2.012961	-2.69391	-1.28431

5. Conclusion

This study shows that crashes are more severe during off-peak hours due to reducing traffic volume (it is no secret that land uses generate and attract less trips during these hours) [29, 30] that allows drivers to drive at free-flow speeds. Another reason can be the 'drivers' fatigue and drowsiness during these hours. Crashes are more severe in spring and summer. One of the reasons for this is the increase in the number of motorcyclists and tourists in these seasons in the first six months of the year, which can also be influential. Crashes with Iranian-made passenger cars and motorcycles can significantly increase the crash severity. One of the reasons for this is the low safety level of cars made in Iran, as they are not equipped with advanced brakes that can control the stability of the car in horizontal angles and also, not having airbags in the event of a collision that can endanger the lives of drivers and passengers. Motorcycle riders are also considered vulnerable road users with very severe crashes. Another factor influencing the severity of a single-vehicle accident is the driver's gender so that men increase the severity of the accident. The reasons for this are more recklessness of male drivers and more compliance of women with the rules. Also, due to the limitation of women using motorcycles and heavy vehicles, male users are considerably more than women [31].

References

[1] M. Vahedi Saheli, M. Effati, Examining the impact of land-use related factors on rural traffic collisions. *Journal of Injury and Violence Research* 11 (2019)

[2] World Health Organization, Road traffic injuries. (2021)

[3] I. Bargegol, M. Keyoumars, An investigation on pedestrian accident severity in urban streets using path analysis. *Computational Research Progress in Applied Science & Engineering (CRPASE)* 2 (2016).

[4] I. Bargegol, V. Najafi Moghaddam Gilani, F. Jamshidpour, Relationship between pedestrians' speed, density and flow rate of crossings through urban intersections (case study: rasht metropolis)(RESEARCH NOTE). *International Journal of Engineering* 30 (2017) 1814–1821.

[5] Iran Legal Medicine Organization reports. (2021)

[6] M. Effati, H. Behbahani, S. Mortezaei, M. Vahedi Saheli, Modelling and Analyzing the Severity of Two-Lane Highway Crashes Using the Spatial Data mining, Case Study: Old

Corridor of Qazvin-Loshan. *Journal of Civil and Environmental Engineering* (2020).

[7] M. Vahedi Saheli, M. Effati, Investigation of factors contributing to pedestrian crash severity in rural roads. *Journal of Injury and Violence Research* 11 (2019)

[8] M. Ghasedi, M. Sarfjoo, I. Bargegol, Prediction and analysis of the severity and number of suburban accidents using logit model, factor analysis and machine learning: a case study in a developing country. *SN Applied Sciences* 3 (2021) 1–16.

[9] B. Dimitrijevic, S. D. Khaled, R. Asadi, J. Lee, Short-Term Segment-Level Crash Risk Prediction Using Advanced Data Modeling with Proactive and Reactive Crash Data. *Applied Sciences* 12 (2022) 856.

[10] S. Jung, X. Qin, D. A. Noyce, Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident Analysis & Prevention* 42 (2010) 213–224.

[11] Q. Hou, X. Huo, J. Leng, Y. Cheng, Examination of driver injury severity in freeway single-vehicle crashes using a mixed logit model with heterogeneity-in-means. *Physica A: Statistical Mechanics and its Applications* 531 (2019) 121760.

[12] M. Zhou, H. C. Chin, Factors affecting the injury severity of out-of-control single-vehicle crashes in Singapore. *Accident Analysis & Prevention* 124 (2019) 104–112.

[13] Z. Li, Q. Wu, Y. Ci, C. Chen, X. Chen, G. Zhang, Using latent class analysis and mixed logit model to explore risk factors on driver injury severity in single-vehicle crashes. *Accident Analysis & Prevention* 129 (2019) 230–240.

[14] M. De Lapparent, Empirical Bayesian analysis of accident severity for motorcyclists in large French urban areas. *Accident Analysis & Prevention* 38 (2006) 260–268.

[15] M. M. Ahmed, R. Franke, K. Ksaibati, D. S. Shinstine, Effects of truck traffic on crash injury severity on rural highways in Wyoming using Bayesian binary logit models. *Accident Analysis & Prevention* 117 (2018) 106–113.

[16] Q. Yuan, X. Xu, M. Xu, J. Zhao, Y. Li, The role of striking and struck vehicles in side crashes between vehicles: Bayesian bivariate probit analysis in China. *Accident Analysis & Prevention* 134 (2020) p.105324.

[17] R. Yu, M. Abdel-Aty, Using hierarchical Bayesian binary probit models to analyze crash injury severity on high speed facilities with real-time traffic data. *Accident Analysis & Prevention* 62 (2014) 161–167.

[18] W. Cheng, G. S. Gill, T. Sakrani, M. Dasu, J. Zhou, Predicting motorcycle crash injury severity using weather data and alternative Bayesian multivariate crash frequency models. *Accident Analysis & Prevention* 108 (2017) 172–180.

[19] Q. Zeng, W. Gu, X. Zhang, H. Wen, J. Lee, W. Hao, Analyzing freeway crash severity using a Bayesian spatial generalized

- ordered logit model with conditional autoregressive priors. *Accident Analysis & Prevention* 127 (2019) 87–95.
- [20] L. Zhu, L. Lu, W. Zhang, Y. Zhao, M. Song, Analysis of accident severity for curved roadways based on Bayesian networks. *Sustainability* 11 (2019) 2223.
- [21] Z. Li, C. Chen, Y. Ci, G. Zhang, Q. Wu, C. Liu, Z. S. Qian, Examining driver injury severity in intersection-related crashes using cluster analysis and hierarchical Bayesian models. *Accident Analysis & Prevention* 120 (2018) 139–151.
- [22] Q. Chen, S. Z. Williams, Y. Liu, S. T. Chihuri, G. Li, Multiple imputation of missing marijuana data in the Fatality Analysis Reporting System using a Bayesian multilevel model. *Accident Analysis & Prevention* 120 (2018) 262–269.
- [23] Q. Zeng, W. Hao, J. Lee, F. Chen, Investigating the impacts of real-time weather conditions on freeway crash severity: a Bayesian spatial analysis. *International journal of environmental research and public health* 17 (2020) 2768.
- [24] S. M. Hosseini, V. Najafi Moghaddam Gilani, H. Tahmasbi Amoli, M. Nikookar, A. Orouei, Presentation of analytical methods for better decision making about the most important factor influencing rural accidents. *Mathematical Problems in Engineering* (2021).
- [25] I. Bargegol, V. N. M. Gilani, M. Ghasedi, M. Ghorbanzadeh, Delay modeling of un-signalized roundabouts using neural network and regression. *Computational Research Progress in Applied Science & Engineering* 2 (2016) 28–34.
- [26] S. M. Hosseini, V. N. M. Gilani, Analysis of factors affecting urban road accidents in rasht metropolis. *Eng Transactions* 1 (2020) 1–4.
- [27] V. N. M. Gilani, M. Ghasedi, M. Ghorbanzadeh, M. J. Samet, Estimation delay variation and probability of occurrence of different level of services based on random variations of vehicles entering signalized intersections. In: *IOP Conference Series: Materials Science and Engineering* 245, IOP Publishing (2017) 042023
- [28] S. Pourfalatoun, E. E. Miller, User perceptions of automated Truck-Mounted attenuators: Implications on work zone safety. *Traffic injury prevention* (2021) 1–8.
- [29] M. Effati, M. Vahedi Saheli, Examining the influence of rural land uses and accessibility-related factors to estimate pedestrian safety: The use of GIS and machine learning techniques. *International Journal of Transportation Science and Technology* (2021)
- [30] M. Vahedi Saheli, M. Effati, Segment-Based Count Regression Geospatial Modeling of the Effect of Roadside Land Uses on Pedestrian Crash Frequency in Rural Roads. *International Journal of Intelligent Transportation Systems Research* 19 (2021) 347–365.
- [31] I. Bargegol, V. N. M. Gilani, Estimating delay of vehicles in nearside legs of the signalized intersections under expectation method in under-saturation conditions for isolated intersection. *Trends Journal of Sciences Research* 2 (2015) 121–125.