Statistical Analysis of Road Traffic Accidents (RTAs) in Sri Lanka

D.S. Kodithuwakku[1], T.S.G. Peiris[2]

Abstract

Road Traffic Accidents (RTAs) are one of the most prominent public health problems as it is a leading cause of death by injury and all deaths globally. This study, therefore, intended to determine the significant factors associated with RTAs in Sri Lanka (2005 - 2019) and the impact of those factors using data obtained from the Department of Police, Sri Lanka. The leading causes for RTAs are overtaking, speed driving and diversion and about 80% RTAs are due to these factors. The percentage of RTAs due to alcohol consumption by the driver is around 9%. Both exploratory and confirmatory factors analysis found that the causes for RTAs can be classified into two independent factors namely, (i) negligence of pedestrians & drivers and (ii) lack of attention of the driver. These factors are invariant by the factor extraction method and the type of orthogonal rotation. The condition of road the surface, light condition of the road, the situation of weather, type of vehicle and age of the driver are significantly influential factors in fatal accidents. The highest percentage of fatal accidents have occurred when the road is wet and light condition is poor during night. The inferences derived from this study can be effectively used for policy decisions related to traffic in order to minimize RTAs in Sri Lanka. The study confirmed the benefits of data-driven decision-making for policy decision process.

Keywords: Data Driven Decision Making, Road Traffic Accidents, Severity of Accident

[1]Lecturer, Department of Social Statistics, University of Kelaniya, Sri Lanka.
[2]Professor in Applied Statistics, Department of Mathematics and Statistics, Faculty of Humanities and Sciences, SLIIT, Malabe Campus

Corresponding Author
D.S. Kodithuwakku, Lecturer, Department of Social Statistics, University of Kelaniya, Sri Lanka.
E-mail: dilushik@kln.ac.lk

1. Introduction

Road safety has been recognized by the 2030 Agenda for Sustainable Development as a precondition to ensure strengthened well-being and healthy lives and to create inclusive, sustainable and resilient cities (WHO, 2018). The WHO (2018) stated that the disabilities and injuries that occur due to Road Traffic Accidents (RTAs) and the safety of roads influence other issues related to community well-being as these factors contribute to the hindrance of human activity. Diseases like diabetes, chronic obstructive pulmonary diseases, strokes and heart diseases become widespread due to self-imposed limitations to walking, cycling and taking public transport by citizens and, increased motorization has also been linked to illnesses in the respiratory system. Ensuring road traffic safety and supporting active travel also contribute to reducing the number of preventable deaths (WHO, 2018). Thus, RTAs are a prominent health-care problem that has a major influence on the universal health situation. However, less attention has been given to analyzing past data related to RTAs. Despite traffic law implementations, technological advances, and improvements in the traffic education systems, the number of traffic fatalities has not decreased substantially over the year. This indicates that although efforts are being made in the right direction, there is a lack of overall understanding of all the contributing factors due to not investigating the data deeply.

Nevertheless, various authors have presented multiple explanations for the reduction in crashes seen in recent years (Nasiri et al., 2019; Demissie, 2017; Dhananjaya & Alibuhtto, 2016). Some factors that influence the reduction are the use of safety belts, safer vehicles, and better roads, strengthened funding for safety infrastructure improvements.

Sri Lankan Police have identified 25 causes for RTAs. However, the direct impact of each attribute has not been investigated via comprehensive statistical analysis. Thus, it is vital to identify the key factors out of the causes listed as recklessness of the driver, negligent of the driver, indiscipline driving, lack of knowledge (road rules and regulations, road conditions, weather conditions, conditions of the vehicle, knowledge regarding apparatus, controls, equipment), human error, fatigue or stress, road infrastructure defects, not planning the trip, duty poor health condition, lack of driving experience and skills, driving under the influence of alcohol, failure to check power, speeding, not wearing the safety belt, pedestrians not following road rules, not following the indications of traffic lights, not knowing the meanings of the different road signs markings signals, lack of skills driving during rains, winds, fog and mist, failure to obey road rules and regulations, driving after taking medication, trying to beat uptime, failure in respecting the rights of others, failure in recognizing civic responsibility, unsatisfactory enforcement by some police officer and not keeping the proper distance.

2. Materials and Methods

2.1 Secondary Data

This study consists with panel data of RTAs in Sri Lanka which were gathered from the Sri Lanka Police. All the data are extent the period of 2005 - 2019. Sri Lanka Police have announced main seven causes of RTAs by combining the causes mentioned in table 1. They are speed driving, overtaking, diversion, mechanical faults of vehicle, alcohol consumption of driver, negligence of pedestrians and other reasons.

2.2 Statistical Analysis:

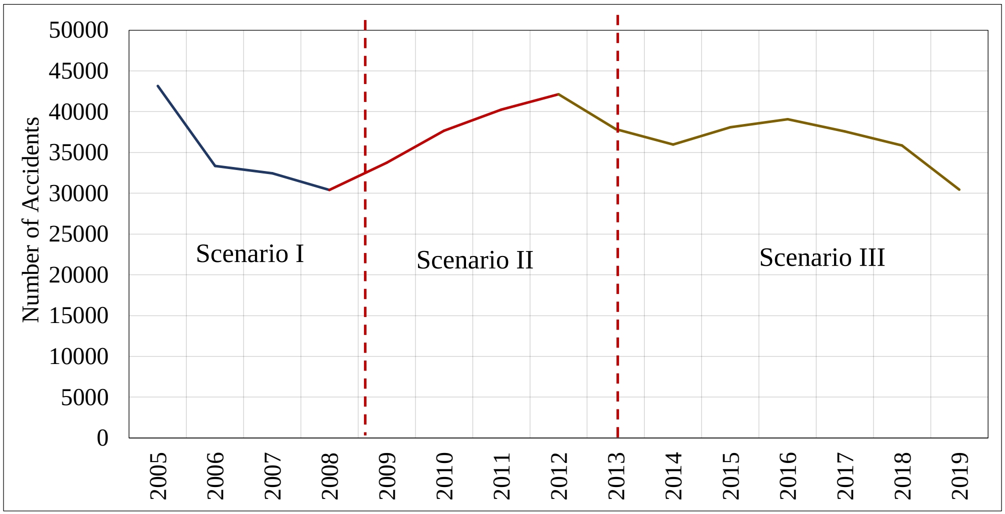
The basic exploratory data analyses and both Exploratory Factor Analysis (EFA) and the Confirmatory Factor Analysis (CFA) are used.

3. Results and Discussion

3.1 Identification of Three Scenarios

In terms of road accidents in Sri Lanka over the last 15 years, there can be seen considerable changes in different periods. Thus, the entire time period from 2005 to 2019 was considered as three main scenarios based on the trend in the annual number of accidents, as shown in figure 1. Scenario I, II and III can thus be considered as the periods: 2005 - 2008, 2009 - 2012 and 2013 – 2019, respectively.

Figure 1: RTAs in Sri Lanka (2005 - 2019)



Source: Survey Data, 2022

It was found that during the period from 2005 to 2019, the number of total accidents has varied from a minimum of 30,420 (2008) to a maximum of 43,171 (2005), with a mean of 36,539 and SE of mean of 1,006.

3.2 Comparison of RTAs in Three Scenarios

The useful descriptive statistics for the three scenarios are shown in table 2.

Table 1: Basic statistics of each scenarios by types of accidents

Type of accident	Scenario I		Scenario II		Scenario III	
	Mean	**SE of Mean**	**Mean**	**SE of Mean**	**Mean**	**SE of Mean**
Fatal	2,135 (6.1%)	24.3	2,405 (6.3%)	81.2	2,640 (7.2%)	121
Grievous	4,919 (14.1%)	71.7	6,424 (16.7%)	413	7,922 (21.8%)	281
Minor	12,430 (35.7%)	661	12,908 (33.6%)	724	12,918 (35.5%)	472
Damage	15,363 (44.1%)	2,186	16,708 (43.5%)	687	12,937 (35.5%)	706
Total	34,847		38,445		36,417	

Note: () represents the percentage contribution with respect to the total

According to table 1, of the three scenarios, the highest number of fatal, grievous and minor accidents have occurred during scenario III. Moreover, the SE of the mean of both grievous and minor accidents are higher in scenario II, and SE of mean of fatal accidents are higher in scenario III. Furthermore, the highest number of damage accidents are in scenario II, but the SE of the mean is lowest in scenario II. The percentage increases in fatal and grievous accidents in scenario III compared to scenario II are 10% and 23%, respectively. In contrast, damage types of accidents have reduced by 23% during scenario III compared to scenario II. Minor types of accidents are almost the same in both scenarios.

3.3 Main Reasons for RTAs

Though Sri Lanka Police has idenitified 25 causes for RTAs, the reasons for RTAs have been classified into seven, as shown in table 2. Furthermore, the category of 'others' has not been well defined. Table 2 describes the mean values of each observed variable and the SE of the mean to recognize the scattering of RTAs data over 15 years (2005 - 2019).

Table 2: Basic statistics of RTAs (per year) among the seven categories during 2005 – 2019

Reasons	Mean	SE of Mean	% of contribution with respect to the total of first six (excluding others)
Overtaking	5,861	16.5	31.6
Speed driving	4,606	13.0	24.8
Diversion	5,024	14.1	27.1
Alcohol consumption of driver	1,610	4.5	8.7
Mechanical faults of vehicle	453	1.3	2.4
Negligence of pedestrians	1,006	2.8	5.4
Others	16,971	47.8	-
Total	35,531		

Source: Survey Data, 2022

Of the 25 causes identified by Sri Lanka Police, the other category in table 2 represents the various minor causes of RTAs which do not belong to the six major reasons described in the first column in table 2. Furthermore, though the number in the other category is very high (47.8% of the total), it is considered as least impact on the society compared to the impact of the other six causes.

Column 4 in Table 2 represents the percentage contribution concerning the total of major causes for RTAs. It can be seen that, nearly more than 80% RTAs were due to overtaking, diversion, and speed driving. Furthermore, it should be noted that RTAs due to alcohol consumption of the driver is significantly less (8.7%).

3.4 Use of Exploratory Factor Analysis (EFA)

As it was found that there is a significant association among six variables shown in table 2, it is necessary to find any common factors linked to those six variables and consequently, EFA was carried out. The significance of Bartlett's Test of Sphericity (Table 3) ratified that the observed correlation matrix is significantly different from the identity matrix, and thus, common factors due to inter-correlation can be investigated. Results in table 3 also indicated that the KMO statistic (0.780) is greater than 0.6, confirming that data satisfying sample adequacy for EFA.

Table 3: Results of KMO and Bartlett's test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.780
Bartlett's Test of Sphericity	Approx. Chi-Square	92.496
	Df	15
	Sig.	0.00

Source: Survey Data, 2022

The results of the JB normality test for each variable concluded that all the variables are not significantly deviate from normal distribution at the 5% level. Thus, the common factors can be extracted using Maximum Likelihood Factoring (MLF) in addition to Principal Axis Factoring (PAF) and Principle Component Analysis (PCF). Furthermore, as the variances of the six variables are unevenly distributed, EFA was carried out for the standardized data, and thus, eigenvalue analysis was carried out for the correlation matrix.

3.5 Eigen Analysis for Correlation Matrix Analysis

Eigen analysis for the correlation matrix in table 4 indicated that only one eigenvalue is greater than one. However, the second eigenvalue is also close to 1. Thus, according to Kaiser's rule (Kaiser, 1960), two factors can be used to explain the observed correlation matrix of six variables. These two factors accounted for 91% of the total observed variance of the six variables confirming that the initial 6 dimensions system can be reduced to a 2 dimensions system.

Table 4: Eigenvalue analysis based on the correlation matrix of the 7 variables

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	4.480	74.669	74.669	4.480	74.669	74.669
2	.995	16.585	91.254	.995	16.585	91.254
3	.350	5.837	97.090			
4	.084	1.398	98.488			
5	.057	.943	99.432			
6	.034	.568	100.000			

Source: Survey Data, 2022

3.6 Extraction and Rotation of Factors

Factors were initially extracted using the PCF method followed by PAF and MLF. It was found that the results are invariant on the type of extraction method. The identified factors need to be rotated to get influential factors (Peiris, 2018). It was found that factors were invariant of the type of orthogonal rotation (Table 5).

Table 5: Summary of the 2 - factor model

Extraction Methods	Type of rotation	Factor 1	Factor 2
PCF	Varimax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver
	Quartimax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver
	Equamax	Speed driving	Overtaking Diversion

		Mechanical faults of vehicle Negligence of pedestrians	Alcohol consumption of driver
PAF	Varimax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver
	Quartimax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver
	Equamax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver
MLF	Varimax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver
	Quartimax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver
	Equamax	Speed driving Mechanical faults of vehicle Negligence of pedestrians	Overtaking Diversion Alcohol consumption of driver

Source: Survey Data, 2022

3.7 Factor Score Coefficient

Of the three types of extraction methods and three types of rotation methods, PCF with the Varimax rotation method is more popular and efficient (Peiris, 2018). Therefore, the factor score coefficient was obtained for the combination of PCF and the Varimax rotation method (Table 6).

Table 6: Factor score coefficient of selected factors from varimax (PCF)

Reasons	Component	
	1	2
Overtaking	.065	.232
Speed driving	.187	.113
Diversion	-.041	.363
Alcohol consumption of driver	-.329	.602
Mechanical faults of vehicle	.448	-.232
Negligence of pedestrians	.457	-.233

Source: Survey Data, 2022

As the EFA was carried out for the correlation matrix, initial variables were standardized prior EFA. Thus, the two factors (F1 and F2) can be defined as,

F1 = 0.187 * [(Speed driving - 4606.47) / 1875.40] +0.448* [(Mechanical fault of vehicle - 452.67) / 254.07] + 0.457* [(Negligence of pedestrians - 1006.00) / 438.17]

F2 = 0.232* [(Overtaking - 5860.73) / 1859.86] + 0.363* [(Diversion – 5024.27) / 790.29] + 0.602* [(Alcohol consumption of driver – 1609.53) / 326.84]

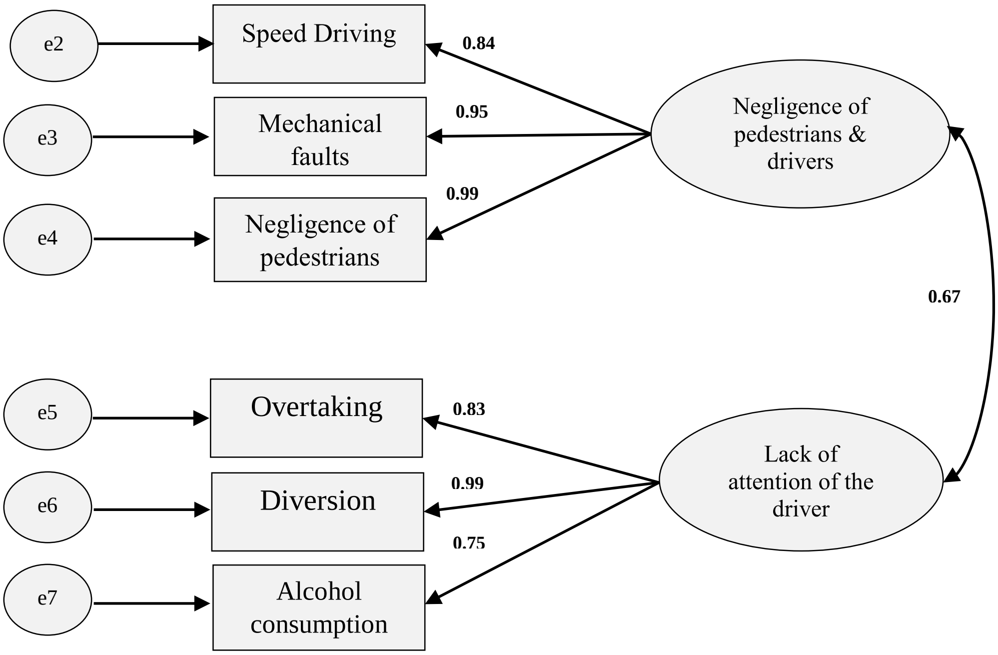
The two factors were named as, "negligence of pedestrians & drivers" and "lack of attention of the driver".

3.8 Confirmatory Factor Analysis (CFA)

In order to further justify the structure of the two factors identified using EFA, CFA was also carried out and results are shown in figure 3.

According to figure 3, all the variables of the two factors indicated high loading values (> 0.8). This justified the validity of the variables selected for the two factors using EFA. The statistics of the model justification is shown in table 8.

Figure 3: Standardized parameter estimates for 2 - factor CFA



Source: Survey Data, 2022

3.8.1 Evaluating Model Fit

Table 7: Model Chi-Square statistics (finalized model)

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	13	23.369	8	.003	2.921
Saturated model	21	.000	0		
Independence model	6	115.965	15	.000	7.731

Source: Survey Data, 2022

The results in table 7 indicate that the chi square statistic (χ_8^2 = 253.369, p = 0.003) is significant, confirming that the hypothesized model significantly fits the data. As this test is sensitive to sample size, various goodness-of -fit indices, have been suggested for the evolution the model fit. Some of the statistical indices generated from SPSS are shown in table 8.

Table 8: Model fit summary of finalized model

RMSEA	.070
NFI	.798
NNFI (TLI)	.815
CFI	.848
GFI	.742
AGFI	.822
PGFI	.283
PNFI	.426

Source: Survey Data, 2022

RMSEA is close to 0. All other indices in table 8 are higher than the critical value of the corresponding indices confirming that the hypothesized model fits the data adequately.

3.9 Association between Severity of Accident and Other Common Factors

In the database, the severity of the accident has been classified as fatal and non-as fatal, and therefore, the analysis of a 2-way frequency table was carried out to find the association between the severity of the accident and condition of the road, light condition of the road, status of weather, type of vehicle and age of the driver. The results of the Chi-square analysis and the percentage of fatal accidents for each level with the main variable are shown in table 9.

As all the Chi-square test statistics are significant in table 9 and as the corresponding p values are less than 5%, it can be concluded with 95% confidence that the condition of the road surface, light condition of the road, the situation of the weather, the type of the vehicle and the age of the driver are significantly influential factors on fatal accidents.

Table 9: Association between main variables and severity of accident

Main Variable	Category	% Fatal accident	Test statistic and p value
Road Characteristics			
Road surface	Dry	5.6	χ^2_1 = 250.507 p = .000
	Wet	7.7	
Light condition	Day light	4.7	χ^2_2 = 2582.049 p = .000
	Night, improper street lighting	8.4	
	Night, good street lighting	5.8	
Time & Environmental Characteristics			
Weather	Clear	5.6	χ^2_1 = 322.418 p = .000
	Humid	7.7	
Vehicle Characteristics			
Vehicle type	Two wheels	6.1	χ^2_2 = 62.542 p = .000
	Three wheels	5.6	
	More than three wheels	5.6	
Human & Accident Characteristics			
Age of driver	Less than 18 years	6.5	χ^2_3 = 52.202 p = .000
	18 - 40 years	5.7	
	40 - 60 years	5.6	
	More than 60 years	5.8	

Source: Survey Data, 2022

The percentage of fatal accidents when the road surface is wet (7.7%) is significantly higher than that when the road is dry (5.6%). The percentage of fatal accidents during the night with improper street lighting (8.4%) is significantly higher than that of during the night with good street lighting (5.8%) as well as that of during the daylight (4.7%). These results were obtained by comparing two binomial distributions separately. The rate of fatal accidents when humid weather (7.7%) is significantly higher than that in clear weather (5.6%). The percentage of fatal accidents by two wheels vehicles (6.1%) is significantly higher than three wheels (5.6%) and four wheels vehicles (5.6%). When comparing three binomial distributions separately, the percentage of fatal accidents when the driver's age is between 18 - 40 years (5.7%) is significantly higher than when the driver's age is between 40 - 60 years. Moreover, the percentage of fatal accidents when the driver's age is above 60 years (5.8%) is significantly higher than when driver's age falls between 40 - 60 years. Similarly, the percentage of fatal accidents when the driver's age is less than 18 years (6.5%) is significantly higher than when the driver's age is more than 60 years.

4. Conclusions and Recommendations

4.1 Conclusions

The percentage of fatal accidents has increased from 6.1% (2005 - 2008) to 7.2% (2013 - 2019). In contrast, damages have dropped from 44.1% to 35.5% during the same time period. The percentage of grievous accidents has increased from 14.1% (2005 - 2008) to 21.8% (2013 - 2019). However, minor accidents have dropped from 35.7% to 35.5% during the same period.

The two common factors were identified irrespective of factor extraction and rotation methods. (i) factor 1: negligence of pedestrians & drivers (ii) factor 2: lack driver's attention. Moreover, the other critical factors associated with fatal accidents are wet road surfaces, nights with improper street lighting, rural areas, humid weather, two wheels vehicles and the age of drivers who are less than 18 years.

4.2 Recommendations

Pedestrians should be educated on the using sidewalks, safe road crossing procedures, and watchfulness while crossing roads.

Traffic signalization, pedestrian bridges and pavement tunnels are good alternatives to prevent pedestrian accidents as they allow pedestrians to cross the road without coming into contact with vehicles.

Lane driving should be strictly followed, and drivers should be encouraged to use signals.

Implementation of proper road maintenance, properly designed road labels, erect road signs and warning signals that suit all climate changes, maintenance of proper road conditions, traffic conditions, and other special road condition requirements and the construction of bumping in appropriate locations along with the removal of all unauthorized bumping on the roads are other alternate steps which could be taken to prevent pedestrian accidents.

References

Demissie, M. (2017). Risk Factors Associated with Serious and Fatal Road Traffic Accidents in Manzini City - Swaziland (Master Dissertation, School of Public Health, University of the Western Cape). Retrieved March 20, 2021 from https://pdfs.semanticscholar.org/bbf2/496ae05630297426426e9cb9c113e6717381.pdf_ga=2.175560935.497661902.1591724915-1509627908.1590092377,

Dhananjaya, S.A.T., & Alibuhtto, M.C. (2016). Factors Influencing Road Accidents in Sri Lanka: A Logistic Regression Approach. 5th Annual Science Research Sessions-2016, 157 - 173.

Kaiser, H. E. (1960). The application of electronic computers to factor analysis. Education & Psychological Measurement, 20, 141-151

Nasiri, N., Nazari, P., Kamali, A., Sharifi, A., & Sharifi, H. (2019). Factors Contributing to Fatal Road Traffic Accidents in the South of Kerman during the period from 2013 to 2017. JOHE 2019, 8(1), 6 - 11. https://doi:10.29252/johe.8.1.6

Peiris, T.S.G. (2018). Handbook on Analysis of Multivariate Data using SPSS. Central Print Shop - Piliyandala. ISBN 978-955-90-2763-8.

Sri Lanka Police. (n.d.). Retrieved April 22, 2020, from https://www.police.lk/index.php/traffic-police/56-traffic-police-road-traffic-accidents

WHO - World Health Organization. (2018). Global Status Report on Road Safety: Summary. Geneva, Switzerland. Retrieved March 20, 2021, from https://www.who.int/violence_injury_prevention/road_safety_status/ 2018/English-Summary-GSRRS2018.pdf