İstatistikçiler Dergisi: İstatistik & Aktüerya

Journal of Statisticians: Statistics and Actuarial Sciences

IDIA 10, 2017, 1, 11-22

Geliş/Received:07.03.2017, Kabul/Accepted: 22.06.2017

www.istatistikciler.org

Araştırma Makalesi / Research Article

An analysis of road traffic accidents in Turkey using logit models

Ayfer Ezgi Yılmaz
Hacettepe University, Department of Statistics,
06800-Beytepe, Ankara, Turkey
ezgiyilmaz@hacettepe.edu.tr

Abstract

The aim of this paper is to evaluate how the variables, such as: age, gender, education of driver, weather condition, condition of day light, condition of settlement, number of vehicle involved, and formation of accident, affect the result of accidents (death or personal injury). Logit models are used to investigate the causes of traffic accidents for the years of 1998, 2004, 2010, and 2013 "Traffic Accident Statistics - *Road*" data of Turkey.

Keywords: contingency tables; logit models; odds ratio; traffic accident.

Öz

Bu makalede yaş, cinsiyet, sürücünün eğitim durumu, hava koşulları, gün durumu, yerleşim durumu, kazaya karışan araç sayısı ve kaza oluş biçimi gibi değişkenlerin kaza sonucu (ölümlü ya da yaralanmalı) üzerindeki etkisini araştırmak amaçlanmıştır. 1998, 2004, 2010 ve 2013 yıllarının "Trafik Kaza İstatistikleri – Yol" Türkiye verileri kullanılarak trafik kazalarının nedenleri lojit modeller kullanılarak incelenmiştir.

Anahtar Sözcükler: *olumsallık tabloları; lojit modeller; odds oranları; trafik kazaları.*

1. Introduction

Traffic and traffic accidents have become dramatically big problems all around the world. According to the World Health Organization report, road traffic injuries caused an estimated 1.25 million deaths each year worldwide in 2015, slightly down from 1.26 million in 2000. That means one person dies every 25 seconds [11].

As a consequence of increasing number of vehicles, traffic and traffic accidents have also become big problems in Turkey. According to the Traffic Accident Statistics-Road reports, 537 352 traffic accidents occurred in 2004. 4427 people died, and 136 437 people injured in the accidents. 1 207 354 traffic accidents occurred, 3 685 people died, and 274 829 people injured in 2013 [10]. The data on road traffic accidents are compiled by the Record of Traffic Accidents forms by filling the forms out. However, because there are also unreported accidents, unfortunately real fatality number is greater than the official number. Besides the deaths and injuries, traffic accidents also result in disabled persons and enormous economic losses.

The aim of this study is to evaluate how the variables, such as: age, gender, education of driver, weather condition, condition of day light, condition of settlement, number of vehicle involved and formation of accident, affect the result of accidents (death or personal injury) and also to investigate whether the precautions taken are useful or not.

Log-linear and logit models are reviewed in Section 2. The description of the "Traffic Accident Statistics-*Road* " data of Turkey and the results of analysis are given in Section 3, followed by discussion in Section 4.

2. Methods

Categorical variables have an importance in many fields often in the medical, social, and behavioral sciences. The tables that represent the categorical variables are called contingency tables. Analysis of categorical variables can be possible by means of log-linear models. Log-linear models that use the binary nature of response variables are referred as logit models. The main purpose of using these models is to investigate the effects of several independent or factor variables on a response.

2.1. Log-linear models for three-way tables

Let X, Y, and Z are the response variables of a three-way $R \times C \times K$ contingency table. The model consists of all interactions between variables is called saturated model. The saturated model for three-way tables is:

$$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}, \quad (1)$$

where m_{ijk} is the expected frequency corresponds to the i th layer ($i = 1, 2, ..., R$) and j th row ($j = 1, 2, ..., C$), and k th column cell ($k = 1, 2, ..., K$). λ is the overall effect parameter, λ_i^X is the effect of variable X at i, λ_j^Y is the effect of variable Y at j, and λ_k^Z is the effect of variable Z at k.

In Equation (1), each pair of variables is conditionally dependent. λ_{ij}^{XY}, λ_{ik}^{XZ}, and λ_{jk}^{YZ} in Equation (1) are the two-way interaction parameters. The model also contains three-way interaction parameter λ_{ijk}^{XYZ}. The saturated model has 0 degrees of freedom [3]. The expected frequencies of a saturated model are equal to the observed frequencies. The eight sub-models are shown in Table 1 [1, 3].

Table 1. Log-linear models for $R \times C \times K$ contingency tables

	Shorthand	Model	Degree of Freedom
0	[X][Y][Z]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z$	$RCK - R - C - K + 2$
1	[X][YZ]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{jk}^{YZ}$	$RCK - R - CK + 1$
2	[Y][XZ]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ}$	$RCK - C - RK + 1$
3	[Z][XY]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY}$	$RCK - K - RC + 1$
4	[XZ][YZ]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XZ} + \lambda_{jk}^{YZ}$	$K(R-1)(C-1)$
5	[XY][YZ]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{jk}^{YZ}$	$C(R-1)(K-1)$
6	[XY][XZ]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ}$	$R(C-1)(K-1)$

7	[XY][XZ][YZ]	$\log m_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$	$(R-1)(C-1)(K-1)$

The likelihood ratio test is a statistical test used to compare the fit of a null model and an alternative model [8]. The general definition of the log-likelihood ratio statistic (G^2) is:

$$G^2 = 2\sum_{i,j,k} n_{ijk} \log(n_{ijk} / m_{ijk}), \tag{2}$$

where n_{ijk} is the observed frequencies for i th layer, j th row, and k th column cell.

2.2. Logit models

Log-linear and logit models can be used to analyze the association of the variables of a contingency table. Log-linear models examine the association between the variables without any distinction of dependent or independent variables. However, logit models are suggested to examine the effects of independent variables on a categorical response variable. Logit models can be derived from the log-linear models [1].

Logit models pair each response with a baseline (reference) category. The baseline category is often the first or the last one. When the baseline category is J, the logits are defined as

$$\log \frac{\pi_j}{\pi_J} \qquad j = 1, 2, ..., J-1. \tag{3}$$

Let Z be a categorical response with K categories. For a three-way contingency table, the logits are defined as

$$l\,\text{og} \frac{m_{ijk}}{m_{ijK}} = \log m_{ijk} - \log m_{ijK}. \tag{4}$$

The logit models can be derived from the best fitting log-linear model. The logit form of the saturated model can be written as the following.

$$\log \frac{m_{ijk}}{m_{ijK}} = \alpha + \tau_i^X + \tau_j^Y + \tau_{ij}^{XY}. \tag{5}$$

The models in Table 1 can be rewritten in the logit form. The equivalent form of logit models to log-linear models for a three-way table with binary response variable Z is shown in Table 2.

Table 2. The equivalent log-linear and logit models for three-way tables

	Log-linear Model	Logit Model
0	[X][Y][Z]	α
1	[X][YZ]	$\alpha + \tau_j^Y$
2	[Y][XZ]	$\alpha + \tau_i^X$
3	[Z][XY]	α
4	[XZ][YZ]	$\alpha + \tau_i^X + \tau_j^Y$
5	[XY][YZ]	$\alpha + \tau_j^Y$
6	[XY][XZ]	$\alpha + \tau_i^X$
7	[XY][XZ][YZ]	$\alpha + \tau_i^X + \tau_j^Y$

2.3. Odds Ratio

Each log-linear and logit model can be used to interpret the association between the variables through odds ratios. Odds ratios of successfully fitting models can be used to interpret the contingency tables.

For a probability of success (π), the odds are:

$$\Omega = \pi/(1-\pi). \tag{6}$$

The odds are nonnegative values. When $\Omega > 1$, it means that a success is more likely than a failure. The logit formulation in Equation (4) is equal to the log-odds [1].

The ratio of two odds is called odds ratio [1]. The odds ratio for 2×2 tables is:

$$\theta = \frac{\Omega_1}{\Omega_2} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}. \tag{7}$$

$\theta = 1$ corresponds to independence of two variables. When $\theta > 1$, it means that the subjects in row 1 are more likely to have a success than the subjects in row 2 [1]. The subset of $(R-1)(C-1)$ local odds ratio for $R \times C$ tables is:

$$\theta_{ij} = \frac{\pi_{ij}\pi_{i+1,j+1}}{\pi_{i+1,j}\pi_{i,j+1}}, \tag{8}$$

3. Data and Results

"Traffic Accident Statistics -Road" yearbooks of Turkey is used to investigate how the variables shown in Table 3 affect the results of accidents (death or personal injury). Data are taken from the Bulletins of General Directorate of Public Security for the years of 1998 and 2004; and Turkish Statistical Institute for the years 2010 and 2013 [4-7, 9-10]. Logit models are used to investigate the causes of traffic accidents. The results are compared on yearly basis.

Table 3. The percentages of road traffic accident data on yearly basis

Variables	**Categories**	**1998** (%)	**2004** (%)	**2010** (%)	**2013** (%)
Age	0-9	3.6	2.2	0.3	6.9
	10-14	7.1	1.7	1.3	4.7
	15-17	10.7	3.4	3.7	5.4
	18-20	4.3	7.1	7.5	8.0
	21-24	17.9	11.1	11.8	10.2
	25-64	21.4	73.1	72.3	59.3
	65+	25.0	1.6	3.2	5.6
Gender	Male	96.7	97.4	98.4	94.2
	Female	3.3	2.6	1.6	5.8
Education of driver	Primary school	58.3	56.2	44.1	35.3
	Secondary school	13.1	12.8	12.3	11.9
	Primary education	0.0	2.1	6.6	10.9
	High school	19.6	19.4	25.2	29.0
	University	9.0	9.4	11.5	12.9

Formation of accident	Crashed from reciprocal, behind and side (CRBS)	57.6	26.0	19.2	44.6
	Collision with standing vehicle (CSV)	3.8	6.3	5.5	2.2
	Collision with stationary object (CSO)	8.2	18.4	22.5	6.7
	Hitting pedestrian (HP)	3.2	3.9	3.3	18.8
	Hitting animal (HA)	0.2	0.7	0.7	0.5
	Overturn (OT)	10.3	16.3	18.8	11.6
	Running off road (RR)	16.6	28.3	29.7	17.9
	Dropped from the vehicle (DV)	0.2	0.1	0.3	0.7
Condition of daylight	Daytime	62.8	60.9	61.8	66.6
	Night	34.0	35.5	34.7	30.3
	Twilight	3.1	3.6	3.6	3.1
Condition settlement	Inhabited area	55.2	57.3	61.2	74.5
	Uninhabited area	44.8	42.7	38.8	25.5
Number of vehicle involved	Single	38.5	36.2	35.9	50.5
	Two	55.2	57.3	57.1	44.4
	Multi	6.3	6.4	7.0	5.1
Weather condition	Sunny	74.0	759	75.3	89.7
	Cloudy or windy	11.4	12.6	12.5	0.2
	Rainy	12.6	8.9	10.9	8.1
	Snowy	1.4	1.8	0.9	1.0
	Foggy	0.7	0.7	0.4	1.0
Result of accident	Death	3.3	3.3	2.1	1.9
	Personal injury	96.7	96.7	97.9	98.1

The percentages of road traffic accident data on yearly basis is given in Table 3. Most of the people (drivers, passengers, and pedestrians) involved in the accidents are 25 to 65 years old. Males are more likely to involve in road traffic accidents than females. The percentage of women involved in the accidents increases in 2013.

When compare with the other education levels, the drivers graduated from primary school involve in the accidents more. However, this percentage decreases over the years. In spite of that, the percentage of the drivers graduated from high school or university involved in the accident increases.

When most common format of accident is crashed from reciprocal, behind, and side in 1998 and 2013; running off road is the most observing format in 2004 and 2010. The percentage of hitting pedestrian increases over the years. Most of the road accidents happen during the daytime, in inhabited area and at sunny days. Around 2% of the traffic accidents result in death in 2013. When the percentage of the accidents happens in inhabited area decreases form 1998 to 2013, the percentage of the accidents happens in uninhabited area increases.

These percentages give us some ideas about the reasons of the accidents but we cannot draw any conclusions about how the other variables affect the result of accidents. Logit models can be used for more detailed inferences. In this study, seven different three-way contingency tables are constituted. Result of accident is accepted as dependent variable and condition of settlement as layer.

The examined model terms are given as follows:

C1: Settlement (S) x Daylight (D) x Result of accident (R)
C2: Settlement (S) x Formation of Accident (F) x Result of accident (R)
C3: Settlement (S) x Number of Vehicle Involved (V) x Result of accident (R)
C4: Settlement (S) x Weather (W) x Result of accident (R)
C5: Settlement (S) x Education (E) x Result of accident (R)
C6: Settlement (S) x Age (A) x Result of accident (R)
C7: Settlement (S) x Gender (G) x Result of accident (R)

The log-linear models in Table 1 are applied to the seven contingency tables (conditions) and the results are summarized in the Appendix. The best fitting models are determined by using the Akaike Information Criteria [2]. Then, the logit models based on the best fitting log-linear model are applied to data and the results are summarized in Table 4. Regarding to the presented results, all models are fit the data sufficiently well (P-*value*> 0.01). The logits are calculated for death/personal injury.

The odds estimated from the parameter estimates of the best fitting models are summarized in Figure 1-7. Figure 1 shows the summarized estimated odds of daylight on yearly basis. The odds of death in inhabited area during the daytime instead of personal injury in 2013 can be calculated from Equation (9). The odds of death instead of personal injury in uninhabited area during the daytime in 2013 can be calculated from Equation (10).

$$\log\left(\frac{InhabitedArea, Daytime, Death}{InhabitedArea, Daytime, PersonalInjury}\right) = \alpha + \tau_1^S + \tau_1^D = 0.009 \quad (9)$$

$$\log\left(\frac{UninhabitedArea, Daytime, Death}{UninhabitedArea, Daytime, PersonalInjury}\right) = \alpha + \tau_2^S + \tau_1^D = 0.041 \quad (10)$$

From Figure 1, the odds of death instead of personal injury decreases till 2010, but increases again in 2013. Consequently, the precautions taken till 2010 are useful. The odds of death at night is higher than the other daylight categories.

Table 4. The results of the logit models for each condition on yearly basis

Condition	Year	Model	G^2	*d.f.*	P-*value*
C1	1998	$\log(\theta) = \alpha + \tau_i^S + \tau_j^D$	5.057	4	0.281
	2004	$\log(\theta)=\alpha+\tau_i^S+\tau_j^D$	2.211	2	0.331
	2010	$\log(\theta)=\alpha+\tau_i^S+\tau_j^D$	1.623	2	0.444
	2013	$\log(\theta)=\alpha+\tau_i^S+\tau_j^D$	3.914	2	0.141
C2	1998	$\log(\theta)=\alpha+\tau_i^S+\tau_j^F+\tau_{ij}^{SF}$	0*	0	-
	2004	$\log(\theta)=\alpha+\tau_i^S+\tau_j^F+\tau_{ij}^{SF}$	0*	0	-
	2010	$\log(\theta) = \alpha + \tau_i^S + \tau_j^F + \tau_{ij}^{SF}$	0*	0	-
	2013	$\log(\theta)=\alpha+\tau_i^S+\tau_j^F+\tau_{ij}^{SF}$	0*	0	-
C3	1998	$\log(\theta)=\alpha+\tau_i^S$	0.000	4	1.000
	2004	$\log(\theta)=\alpha+\tau_i^S+\tau_j^V$	0*	0	-
	2010	$\log(\theta)=\alpha+\tau_i^S+\tau_j^V$	0*	0	-
	2013	$\log(\theta) = \alpha$	0.000	5	1.000
C4	1998	$\log(\theta)=\alpha+\tau_j^W$	5.314	8	0.724
	2004	$\log(\theta)=\alpha+\tau_j^W$	3.739	8	0.880
	2010	$\log(\theta)=\alpha+\tau_j^W$	13.497	8	0.096
	2013	$\log(\theta)=\alpha+\tau_i^S+\tau_j^W$	3.2291	4	0.510
C5	1998	$\log(\theta)=\alpha+\tau_i^S+\tau_j^E$	3.128	4	0.534
	2004	$\log(\theta)=\alpha+\tau_i^S+\tau_j^E$	8.584	4	0.072
	2010	$\log(\theta) = \alpha + \tau_i^S + \tau_j^E$	10.650	4	0.030
	2013	$\log(\theta)=\alpha+\tau_i^S+\tau_j^E$	11.671	4	0.020

C6	1998	$\log(\theta)=\alpha+\tau_i^S+\tau_j^A$	4.682	6	0.583
	2004	$\log(\theta)=\alpha+\tau_i^S+\tau_j^A$	3.678	6	0.720
	2010	$\log(\theta)=\alpha+\tau_i^S+\tau_j^A$	4.050	6	0.670
	2013	$\log(\theta)=\alpha+\tau_i^S+\tau_j^A$	6.909	6	0.329
C7	1998	$\log(\theta)=\alpha+\tau_i^S+\tau_j^G$	1.409	1	0.235
	2004	$\log(\theta)=\alpha+\tau_i^S+\tau_j^G$	0.156	1	0.693
	2010	$\log(\theta)=\alpha+\tau_i^S+\tau_j^G$	0.018	1	0.894
	2013	$\log(\theta)=\alpha+\tau_i^S+\tau_j^G$	0.100	1	0.752

* Calculated under the saturated model

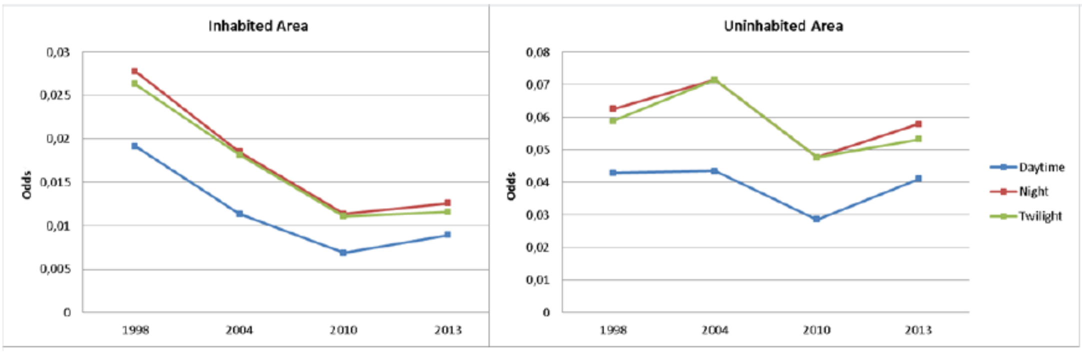

Figure 1. The estimated odds of daylight on yearly basis

Figure 2 shows the estimated odds of formation of accident on yearly basis. Figure 2 shows that, the odds of death instead of personal injury is lower for hitting the pedestrian and higher for the dropped from the vehicle. The odds of death increase in the hitting pedestrian accidents in 2013. Hitting the pedestrian and dropped from the vehicle accidents are the highest risky accidents in uninhabited area. The odds of death is higher in uninhabited area than inhabited area.

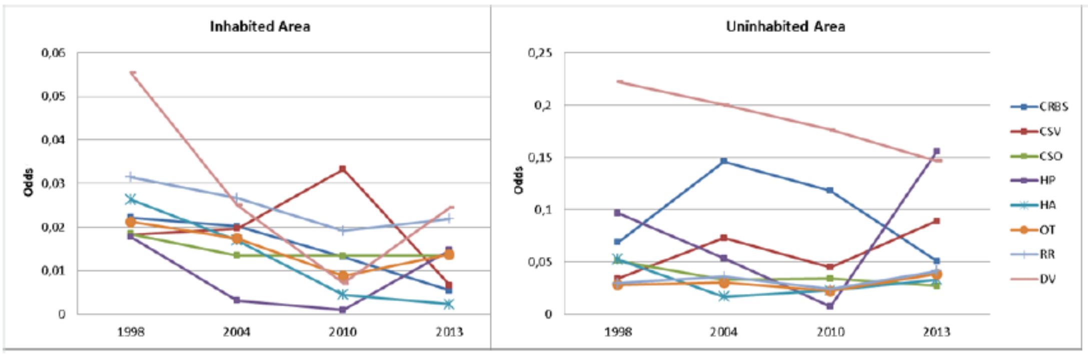

Figure 2. The estimated odds of formation of accident on yearly basis

Figure 3 shows the estimated odds of number of vehicles involved on yearly basis. From Figure 3, the odds are similar for all kind of accidents in both inhabited area and uninhabited area. Although the odds of death in inhabited area decreases in 2010, it has the highest value in 2013.

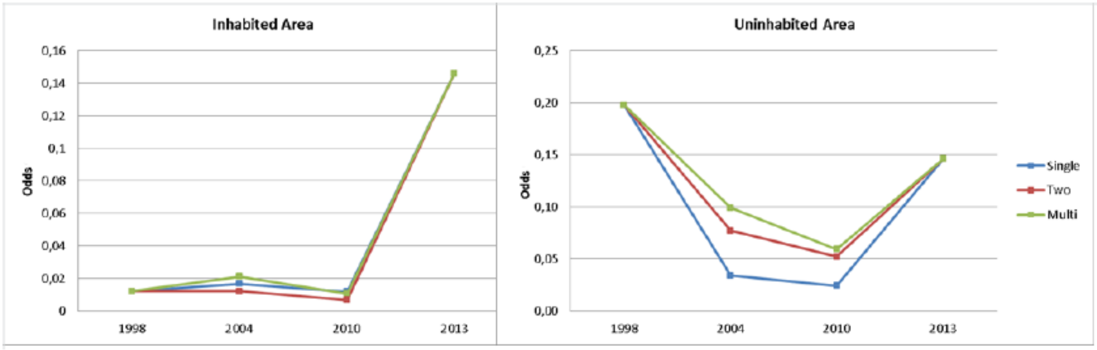

Figure 3. The estimated odds of number of vehicle involved on yearly basis

Figure 4 shows the estimated odds of weather condition involved on yearly basis. From Figure 4, it can be said that the odds of death decreases from 1998 to 2010. In 2013, cloudy or windy weathers have the highest, snowy and rainy weathers have the lowest risks. The reason is the drivers drive the vehicle slowly and more carefully when the weather is snowy or rainy.

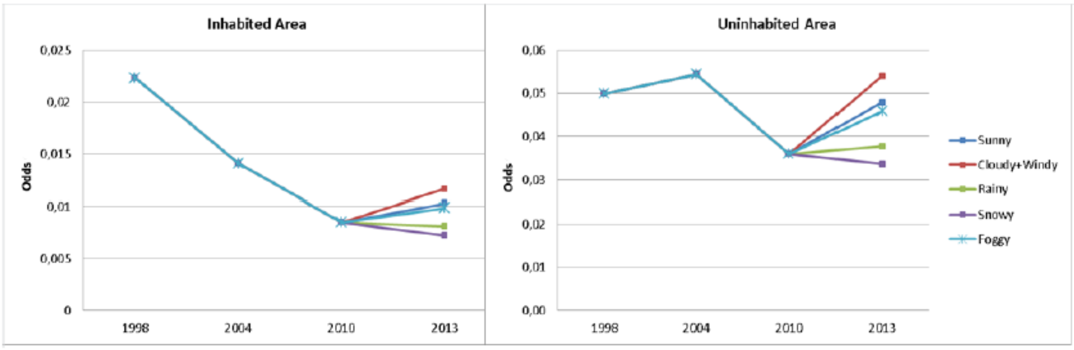

Figure 4. The estimated odds of weather conditions on yearly basis

Figure 5 shows the estimated odds of education on yearly basis. Figure 5 shows that, the drivers have primary education have the highest risk of death. For all the education categories, the odds of death in 2013 is higher than the odds in 2010. Besides, the odds in uninhabited area are higher than inhabited area.

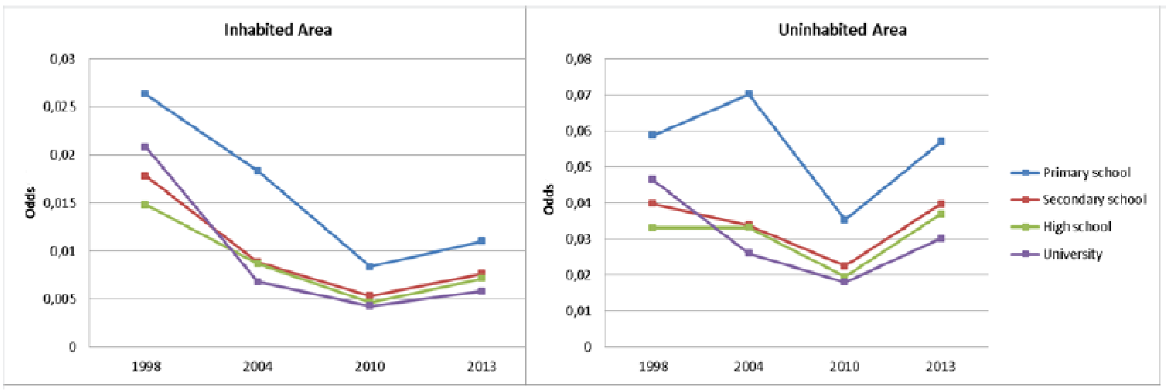

Figure 5. The estimated odds of education on yearly basis

Figure 6 shows the estimated odds of age on yearly basis. Figure 6 shows that, the odds of death for the people older than 21 increases. It has the highest value for the people over 65 years old. The odds of dead of a person younger than 15 decreases in 2013.

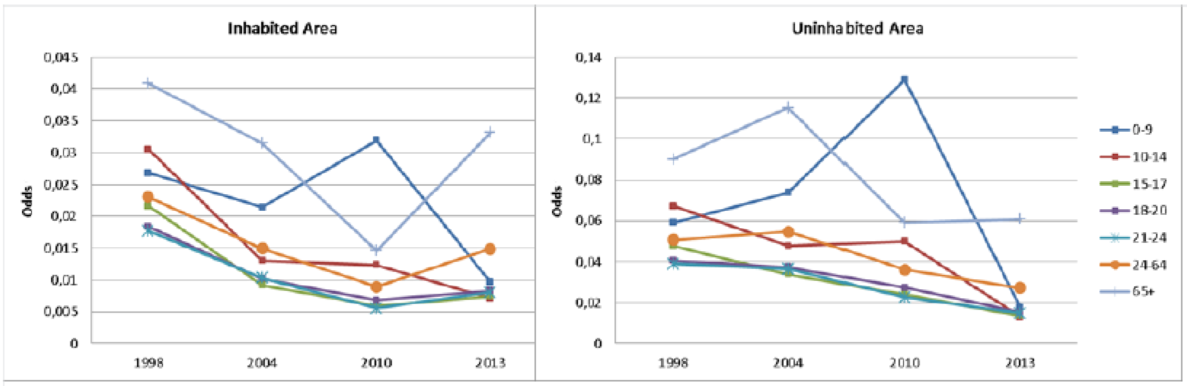

Figure 6. The estimated odds of age on yearly basis

Figure 7 shows the estimated odds of gender on yearly basis. Figure 7 shows that, males are at higher risk of death than females. In 2013, the odds of a female drivers' death in inhabited area decrease.

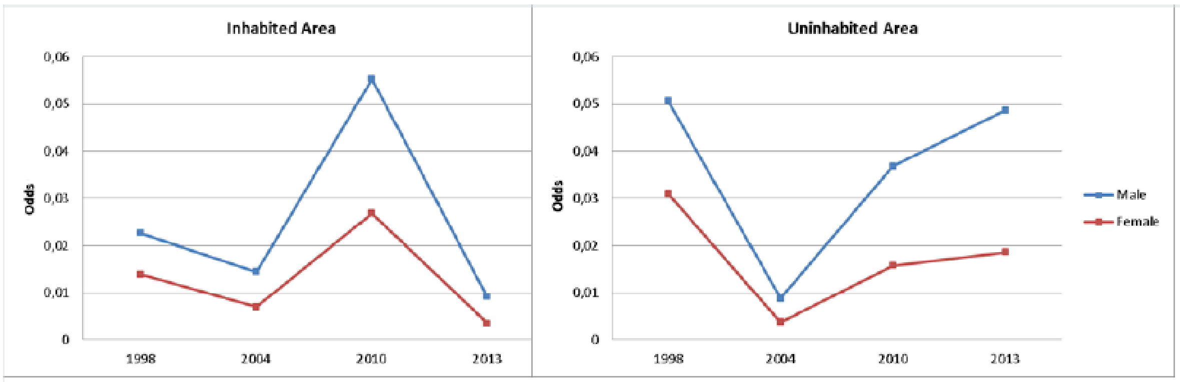

Figure 7. The estimated odds of gender on yearly basis

From the odds in Figure 1-7, it is possible to estimate the odds ratios. The odds ratios of death in uninhabited area instead of inhabited area for each variable are summarized in Table 5. For all the categories of condition of daylight, the odds of the accident result in death in uninhabited area is 2.24 times the odds for those in inhabited area in 1998. This odds ratio increases to 4.56 in 2013.

$$\theta = \log\left(\frac{UninhabitedArea, Daytime, Death}{UninhabitedArea, Daytime, PersonalInjury}\right) \Big/ \log\left(\frac{InhabitedArea, Daytime, Death}{InhabitedArea, Daytime, PersonalInjury}\right) = 0.041/0.009 = 4.56$$

Table 5. The estimated odds ratios on yearly basis

	1998	2004	2010	2013
Formation of Accident				
Crashed from reciprocal. behind and side	3.11	7.24	8.92	9.24
Collision with standing vehicle	1.88	3.71	1.35	13.32
Collision with stationary object	2.78	2.45	2.54	2.03
Hitting pedestrian	5.44	17.01	7.91	10.59
Hitting animal	2.00	1.00	5.05	14.13
Overturn	1.32	1.74	2.53	2.81
Running off road	0.94	1.35	1.26	1.89
Dropped from the vehicle	4.00	8.00	24.69	6.01
Number of vehicle involved				
Single	16.25	2.03	2.03	1.00
Two	16.25	6.29	7.68	1.00
Multi	16.25	4.68	5.44	1.00
*Weather Condition**	2.24	3.86	4.23	4.65
*Education of driver**	2.23	3.83	4.23	5.18
*Age**	2.20	3.44	4.04	1.84
Gender				
Male	2.24	0.61	0.67	5.31
Female	2.24	0.54	0.59	5.23

* The calculated odds ratios for all the categories of the variable are equal.

For the running of the road, the odds of the accident result in death in uninhabited area is 0.94, 1.35, 1.26, and 1.89 times the odds for those in inhabited area in the years of 1998, 2004, 2010, and 2013, respectively. For the hitting pedestrian, the odds ratio increases from 5.44 to 17.01 from 1998 to 2004. Although it decreases to 7.91 in 2010, increases to 10.59 in 2013. For the dropped from the vehicle, the odds ratio of accident result in death increases from 4.00 to 24.69 in from 1998 to 2010. However, it decreases to 6.01 in 2013.

For the all categories of number of vehicle involved, the odds of the accident result in death in uninhabited area is 16.23 times the odds for those in inhabited area in 1998. For the multi vehicle accidents, the odds of the accident result in death in uninhabited area is 7.68 times the odds for those in inhabited area in 2010.

For all the categories of weather condition, the odds of accident result in death in uninhabited area is 2.24 times the odds for those in inhabited area in 1998. This odds ratio increases to 4.65 in 2013. For all the categories of education, the odds of accident result in death in uninhabited area is 2.23 times the odds for those in inhabited area in 1998. This odds ratio increases 5.18 in 2013.

For all the categories of age, from 1998 to 2010, the odds ratio of accident result in death in uninhabited area than inhabited area increases from 2.20 to 4.04. This odds ratio decreases to 1.84 in 2013. The males' and females' odds of accident result in death in inhabited area are 2.24 times the odds for those in uninhabited area in 1998. Although, these odds ratios decrease to 0.67 for males and 0.59 for females in 2010, it increases to 5.31 for males and to 5.23 for females.

4. Discussions

Road traffic accidents increase dramatically worldwide. Many countries have redoubled their efforts to prevent traffic injuries but the results have not been as far as many have hoped. In that case, it is essential to investigate the effects of the traffic accidents. In this study, we analyzed the road traffic accidents data of Turkey on yearly basis and investigated whether the precautions which have been taken are sufficient.

From the results of the analyses, the accidents happen in the twilight or at night and dropped from the vehicle accidents are the most risky ones. When we look at the drivers' profile, most risky ones are primary educated males. The people aged over 65 years old have more risk of dying as a result of a road traffic accident. The odds of an accident results in death in inhabited area is lower than uninhabited area for all the categories of variables.

The lowest risky accidents are the accidents happen in daylight and hitting pedestrian or animal. When we look at the lowest risky drivers' profile, they are high school or university educated females. The people aged between 18 and 24 have less risk of dying as a result of a road traffic accident.

The results show that, the odds of the traffic accidents results in death decreases from 1998 to 2010, but increases again in 2013. The studies in the field of traffic have been helpful but still the risk of death is higher than the other European countries. Because the risk of death is higher, General Directorate of Public Security can take some extra precautions in uninhabited area.

There can be some extra precautions in order to prevent the accidents that happen in the twilight or at night. In order to get driving license, the driver applicants at least graduated from primary school. However, when we compare to the other education categories, these drivers are the most risky ones. Thus, this level can be changed from primary school to a higher level.

References

[1] A. Agresti, 2002, *Categorical Data Analysis*, 2nd ed., New York: John Wiley & Sons Inc.

[2] H. Akaike, 1974, A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19 (6), 716-723.

[3] R. Christensen, 1990, Log-linear Models, New York: Springer-Verlag.

[4] Road Safety Annual Report, 1998, *Republic of Turkey Ministry of Interior, General Directorate of Public Security*, Ankara, Turkey.

[5] Road Safety Annual Report, 2004, *Republic of Turkey Ministry of Interior, General Directorate of Public Security*, Ankara, Turkey.

[6] T. Saracbasi, S. Aktas, 2003, *Investigation of the traffic accident reasons by statistical models*. 2nd National Congress/Exhibition about Traffic and Road Safety, 7-9 May 2003, Ankara.

[7] T. Saracbasi, 2011, *An analytical view of traffic accident statistics*, Traffic and Road Safety Symposium, 23 November 2011, Ankara.

[8] R.R. Sokal, F.J. Rohlf, 1981, *Biometry: The Principles and Practice of Statistics in Biological Research*, New York: Freeman.

[9] Traffic Accident Statistics – Road, 2010, *General Directorate of Public Security and Turkish Statistical Institute*, TUIK, Ankara, Turkey.

[10] Traffic Accident Statistics – Road, 2013, *General Directorate of Public Security and Turkish Statistical Institute*, TUIK, Ankara, Turkey.

[11] WHO, 2013, "Global Status Report on Road Safety 2013: supporting a decade of action" (PDF) (official report). Geneva, Switzerland: World Health Organization (WHO).

Appendix The results of the seven log-linear models for each condition on yearly basis

		Year		**1998**			**2004**			**2010**			**2013**	
Condition	Model	*d.f.*	G^2	*P-value*	AIC	G^2	*P-value*	AIC	G^2	*P-value*	AIC	G^2	*P-value*	AIC
C1	0	7	119.38	0.00		461.84	0.00		628.05	0.00		1793.22	0.00	
	1	5	165.39	0.00		404.52	0.00		559.62	0.00		1708.27	0.00	
	2	6	39.05	0.00		88.73	0.00		100.20	0.00		124.58	0.00	
	3	5	194.75	0.00		425.41	0.00		594.00	0.00		1752.22	0.00	
	4	4	5.06	0.28	-**2.94**	31.41	0.00		31.76	0.00		39.63	0.00	
	5	4	34.42	0.00		52.30	0.00		66.14	0.00		83.58	0.00	
	6	3	160.75	0.00		368.09	0.00		525.56	0.00		1667.27	0.00	
	7	2	1.94	0.38	-2.06	2.21	0.33	-**1.79**	1.62	0.44	-**2.38**	3.91	0.14	**-0.09**
C2	0	22	4019.38	0.00		3145.70	0.00		5999.70	0.00		35716.75	0.00	
	1	21	3853.34	0.00		2973.52	0.00		5819.43	0.00		34070.81	0.00	
	2	15	276.91	0.00		377.62	0.00		416.88	0.00		141325.42	0.00	
	3	15	3988.04	0.00		3022.45	0.00		5893.89	0.00		35264.43	0.00	
	4	14	3822.00	0.00		2850.27	0.00		5713.62	0.00		33618.48	0.00	
	5	14	110.87	0.00		205.44	0.00		236.60	0.00		571.89	0.00	
	6	8	245.57	0.00		254.36	0.00		311.07	0.00		1765.50	0.00	
	7	7	40.70	0.00		63.66	0.00		95.78	0.00		316.56	0.00	
C3	0	7	1451.65	0.00		2682.42	0.00		14057.15	0.00		9081.88	0.00	
	1	5	1421.78	0.00		266.41	0.00		14046.82	0.00		8807.11	0.00	
	2	6	1291.33	0.00		2309.31	0.00		13774.07	0.00		7413.24	0.00	
	3	5	245.02	0.00		474.18	0.00		603.49	0.00		0.00	1.00	-**10.00**
	4	4	1261.45	0.00		2292.30	0.00		13733.74	0.00		7138.47	0.00	
	5	3	215.14	0.00		457.18	0.00		593.16	0.00		1635.12	0.00	
	6	4	0.00	1.00	-**8.00**	101.07	0.00		290.41	0.00		241.33	0.00	
	7	2	27.54	0.00		49.99	0.00		179.94	0.00		179.73	0.00	
C4	0	13	896.73	0.00		974.76	0.00		1092.34	0.00		2883.20	0.00	
	1	12	737.77	0.00		600.43	0.00		561.95	0.00		1225.83	0.00	
	2	9	164.28	0.00		0.00	0.00		0.00	0.00		0.00	0.00	
	3	9	891.60	0.00		960.86	0.00		1090.80	0.00		2880.41	0.00	
	4	8	732.64	0.00		586.53	0.00		560.41	0.00		123.04	0.00	
	5	8	5.31	0.72	-**2.63**	3.74	0.88	-**6.89**	13.50	0.10	-**2.50**	19.78	0.01	
	6	5	159.15	0.00		364.18	0.00		542.34	0.00		1674.36	0.00	
	7	4	1.62	0.81	0.76	1.74	0.78	-1.50	8.57	0.07	0.57	3.29	0.51	**-4.71**
C5	0	13	414.23	0.00		577.80	0.00		940.76	0.00		3735.93	0.00	
	1	12	253.90	0.00		204.69	0.00		634.09	0.00		1298.18	0.00	
	2	9	212.76	0.00		494.68	0.00		370.84	0.00		2643.52	0.00	
	3	9	362.28	0.00		459.21	0.00		886.12	0.00		3462.53	0.00	
	4	8	201.96	0.00		86.10	0.00		579.45	0.00		1024.76	0.00	
	5	8	52.43	0.00		121.58	0.00		64.47	0.00		205.79	0.00	
	6	5	160.82	0.00		376.09	0.00		316.20	0.00		2370.12	0.00	
	7	4	3.13	0.53	-**2.26**	8.58	0.07	**0.58**	10.65	0.03	**2.65**	11.67	0.02	**3.67**
C6	0	19	2016.63	0.00		2296.33	0.00		3753.78	0.00		2820.92	0.00	
	1	18	1856.30	0.00		1923.22	0.00		3233.30	0.00		2293.52	0.00	
	2	13	183.49	0.00		411.51	0.00		568.52	0.00		1412.46	0.00	
	3	13	1983.11	0.00		2218.28	0.00		3661.93	0.00		1897.34	0.00	
	4	12	1822.78	0.00		1845.17	0.00		3141.45	0.00		1369.95	0.00	
	5	7	149.97	0.00		333.46	0.00		476.67	0.00		488.89	0.00	
	6	12	23.16	0.03		38.40	0.00		48.05	0.00		885.06	0.00	
	7	6	4.68	0.59	-**3.36**	3.68	0.72	-**5.69**	4.05	0.67	-**5.31**	6.91	0.33	**-5.09**
C7	0	4	171.05	0.00		171.05	0.00		601.13	0.00		3360.66	0.00	
	1	3	166.90	0.00		166.90	0.00		550.83	0.00		3110.04	0.00	
	2	3	10.72	0.01		10.72	0.00		73.26	0.00		621.30	0.00	
	3	3	165.28	0.00		165.28	0.00		572.30	0.00		3213.40	0.00	
	4	2	14.95	0.00		4.95	0.00		44.44	0.00		472.99	0.00	
	5	2	16.57	0.00		6.57	0.00		22.97	0.00		99.44	0.00	
	6	2	161.13	0.00		161.13	0.00		522.01	0.00		2688.24	0.00	
	7	1	1.41	0.24	-**1.82**	1.41	0.69	-**1.41**	0.18	0.89	-**2.64**	0.10	0.75	**-1.90**