IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 06, Issue 12 (December. 2016), ||V1|| PP 59-66

A model for the cost of accidents considering drivers' alertness criteria

Shahin Shabani[1], Azadeh Ghanbarpour[2], Pedram Esmaeeli[3]
[1] *(Department of Civil Engineering, Payame Noor University, PO Box 19395-3697 Tehran, IRAN*
[2] *(Department of Civil Engineering, Payame Noor University, PO Box 19395-3697 Tehran, IRAN*
[3] *(Department of Civil Engineering, Science and Research Branch of Islamic Azad University, Tehran, IRAN*

Abstract: - Human factors are the main reasons of the traffic accidents. Among the human characteristics, alertness, directly affects the attention and concentration of the drivers which have impact on the severity and consequently, imposed cost of the accidents. In this study, focusing on the human factors, a cost model is generated based on the drivers' alertness parameters. In determining the alertness factors, considered parameters are; driver life style, behavioral habits, social and personal issues. Obtaining a 90 percent reliability level, variables like, age, sex, daily average work hour, illness and sleepiness indexes are statistically meaningful. Mean square method is used to assess the costs in order to generate a linear regression model. The output of the model represents the coefficient of the sex and the daily average work hour, 10 and 7 respectively which shows the most influence on the accident costs. Other considered variables are: sleepiness index, drug usage index, age and illness index with coefficient less than the two main variables. Sensitivity analysis of the model towards the independent variables shows the changes in age variable as the most influential factor.

Keywords: - *regression model, alertness criteria, traffic accident, sensitivity analysis*

I. INTRODUCTION

Highway transportation, because of its dominant role in general transportation in a country has a significant effect on economic development. With the increased traffic in cities and related accidents, the different costs imposed to the transportation system and the society seems inevitable. In Iran, according to statistics derived from a study done by transportation research center, more than 6 percent of the gross national product goes to the irreparable loss due to accidents.One approach to improve and develop the transportation systems is to study the factors involved in accidents that leads to give solutions to decrease the number of accidents and thus the amount of loss. Most of the studies focuses on diagnosis the reasons and ameliorate the consequences of the accidents while in this article, the driver as the most important counterpart of the traffic system is studied and factors related to human alertness are discussed. The quality of driving is mainly depended on the driver's skills and his perception. The psychomotor skills are not included in this study and factors affecting the driver's alertness are: drivers' life styles, behavior habits, social and personal issues like job, age and sex and also soporific drug usage. As a sample, the Pittsburgh sleep quality index that is used in designing the questionnaire is attached. Another factor which enormously affects a driver performance is tiredness. All factors mentioned, have included in accidents severity and the respective costs. Focusing on the degree of effectiveness of the factors above, a relation between the dependent variable (accidents' cost according to accidents severity) and the independent variables like driver physical conditions, drug usage and chronic tiredness due to influence of some of the illnesses which also reduces the drivers' alertness is concluded and studied. Defined parameters are extracted from standard international indexes and used in designing the questionnaires. The concluded formula is practiced through some defined accidents and the degree of effectiveness for each factor in the resulted model is discussed. In short, the importance of this article is for decision makers to put emphasis on development the facilities to help drivers' alertness mechanisms.Map of this paper is as follows; after reviewing the previous related documents, the method description which is applied to this study is defined and the variables definitions are given. The next step is to collect the data of the survey and generate the summarized tables. Testing the data statistically is a core part of this study and meaningfulness of the final variables are granted. The model is generated based on the concluded data and finally a sensitivity analysis is done.

II. REVIEW OF THE LITERATURE

Studies on highway accidents shows that road crashes, kill 1.3 million and injure some 50 million people every year; urban congestion reduces global GDP by more than 8 percent and more than 50 million dollars are imposed to the region due to highway accidents in 2014 which will affect the economy of the region undoubtedly and transport lending accounts for 21 percent of the Bank's total active portfolio (more detail on this issue can be found at world bank reports 2014). Researches on the detriment due to accidents shows about one percent of GNP goes to the direct loss which is considerably important economically [1]. This direct loss due to traffic accidents in Spain is between 1 to 2 percent of GNP [2]. The most relevant factor in human behavior which leads to the most severe accidents is the loss of the alertness and the sleep disorders are among the most effective reasons to blemish this item whether it is due to tiredness or lack of sleep [3]. Dawson in his studies shows that amongst all traffic accidents, 20 percent is directly because of drivers' sleep disorders. In a research done by Stanford University with the information prepared by National Research Commission, the direct cost of traffic accidents is more than 15.9 billion dollars a year and between 50 to 100 billion dollars indirectly which included destruction of the public assets, hospitalization and even death [4]. The most importanti reason on losses due to accidents is because of the lack of experience in driving which has showed that would be improved through training and educating. These solutions even reduce the cost of losses related to accidents [5]. Recent studies on finding solutions to reduce the correlated costs of the accidents that are because of some cognitive or perceptive disorders are done and concluded that reaction time in drivers with perceptive disorders are relatively more than normal drivers [6]. The type of the accident and the amount of loss it caused are depended on the severity of the accident which leads to death in accidents with drivers who used drugs and had some mental disorders [7]. Other criteria affecting the parameters included in accident cost model are the speed and acceleration patterns used by drivers [8] in which it represents a model for both the severity of the accident and the risky behavior of the drivers. Lots of other studies have done in finding the relation between the sex and the occurrence of the accident or its severity but the one done by Peden [9], [10] proves that fatality in men drivers are twice the women and therefore the cost and consequences of accidents get worse.

III. METHODOLOGY OF THE STUDY

The main methods used in this study are theoretical and practical. Gathering the information is done through interviewing and filling out questionnaires which were designed especially for this purpose. For the data related to accident types from the property or life loss point of view, four levels assumed; eventuate to death, severe property loss, minor property loss and only property loss. For the factors related to drivers' alertness, a logit model developed and related variables are presented. Having introduced an abbreviation for each variable, a regression model is developed for dependent variable. After generating the model, statistical test is done to validate the most suitable model and by using try and error method, final model is concluded. In generating the regression model, minimum square method is applied. Sensitivity analysis is done by changing each independent variable about 10 percent and monitoring the changes in accident cost as the dependent variable.

For this study, from 180 questionnaire prepared (the reason of this number is because of estimated relation proposed in the book of "Statistical and Econometric Methods for Transportation Data Analysis" written by Simon P. Washington Matthew G. Karlaftis Fred L. Mannering that the number of samples must be around 15 times of the number of independent variables. Since in this study, number of independent variables are 10, thus about 150 samples are needed. Considering 20% of samples not qualified, valid numbers of samples are 180). 147 subjects participated in the survey and entered the analyze process and after controlling the independency of the data, generation and evaluation of the model carried out.

IV. INTRODUCTION OF THE VARIABLES AND GENERATING THE MODEL

In general, the indexes relating to driver's alertness are analyzed from three points of views. The first view, relates the sleepiness and tiredness of driver to the parameters like, job, work hours, age and the time of driving. The second view includes accuracy and attentiveness which depends on distractions, reaction time and genetic qualities of people. In third view, underlying disease and usage of drugs with side effects are considered. In evaluating the third view, firstly, the diseases which lead to mental disorders are studied and secondly the diseases don't lead to sleep disorders but the drugs in prescription leads to tiredness and sleepiness are considered.In this article, the cost imposed to the traffic system is dependent variable and age, sex and driving experience (in year), average driving time in day (in hours), disease index (Im), drug usage index (Id) and the sleepiness index (Is) are the independent variable of the model.Gathering the data in this study was done through following steps;

4.1. test of the variable independency

As is evident from table (1), the regression of variables is done and the small quantity of the correlation proves that the variables don't have linear relationship and they don't describe each other and in fact they are independent.

Table 1.independency test of the independent variables

		I_m	I_d	I_s	Job time
I_m	Pearson Correlation	1	0.339	-0.389	-0.656
	Sig (2tailed)	-	0.015	0.005	0.00
	N	147	147	147	147
I_d	Pearson Correlation	0.339	1	-0.744	-0.22
	Sig (2tailed)	0.015	-	0.00	0.121
	N	147	147	147	147
I_s	Pearson Correlation	-0.389	-0.744	1	0.549
	Sig (2tailed)	0.005	0.00	-	0.00
	N	147	147	147	147
Job time	Pearson Correlation	-0.656	-0.22	0.549	1
	Sig (2tailed)	0.00	0.121	0.00	-
	N	147	147	147	147

4.2. Group frequency analysis and the accidents cost

In order to group the data, the information of 301 accidents have been registered which gives an average of 2.04 accidents per capita. In table (2), the accidents were grouped and the abundance of each group is showed. The frequency of the imposed cost of accidents is showed in table (3) with assumption that each dollar equals 10,000 rials.

Table 2.The frequency of the severity of the accident groups

Accident severity	With death	Severe injury	Minor injury	Only property loss	Total
frequency	0	19	61	221	301

Table 3.The imposed cost due to accidents

Cost ($)	<100	100<-<300	300<-<500	>500	Total
frequency	53	131	107	10	301

4.3. Allocation of the data to the model parameters

To generate a regression model for the imposed costs, the first step is to encode the variables. It should be mentioned that the dependent variable in this model has a numerical type and shows the cost imposed to system due to accidents. All the coded variables defined in table (4) with their types.

Table 4.The code and the type of the independent variables

variable	Variable code	Type of the variable
age	Age	numerical
sex	Sex	virtual
Education level	G	ordinal
Having night shifts	N	Virtual
Average daily work hour	WH	Numerical
Experienced drivers	B	Numerical
Average time driving in day	D	Numerical
Disease index	I_m	Numerical
Drug usage index	I_d	Numerical
Sleepiness index	I_s	Numerical
Accident time	T	virtual

4.4. Generating the model and discussions

In a regression equation with minimum square method, the disorder sentence is estimated through the method mentioned which is the difference between mathematical model and statistical one. The assumptions are as follows: the disorder sentence is random which shows the model has correctly fitted and the disorder sentence has a normal distribution which eases the analysis of the model, the mathematical expectation of the disorder sentences are zero that means the total deviation of the regression line is zero and has a fixed variance. Having the mentioned assumptions, the model is generated after fitting the data and several try and error process that is showed in table (5) below.

Table 5.Estimated model for the accidents imposed cost

Independent variable	state	Variable code	Estimated value	p-value
Constant number	-	-	80.18	0.005
Age	Continuous	-	2.31	0.05
sex	Male/female	Male:1/female:0	10.45	0.08
Education level	Under diploma/diploma/BSc/MSc and more	0,1,2,3	Statically meaningless	0.74
Having night shifts	-	No:0/ yes:1	Statically meaningless	0.32
Average daily work hour	Continuous	-	7.21	0.07
Driving experience	Continuous	-	Statically meaningless	0.27
Daily average driving time	Continuous	-	Statically meaningless	0.24
Disease index	Continuous	-	2.18	0.04
Drug usage index	Continuous	-	3.54	0.01
Sleepiness index	Continuous	-	4.43	0.06
Time of the accident	-	0 to 6/ 6 to 12/ 12 to 18/ 18to 24	Statically meaningless	0.45

Considering the parameters in the table above and the estimated model, the concluded model based on the driver alertness parameters explained is as equation (1):

$$C=80.18+2.31Age+10.45Sex+7.21WH+2.18I_m+3.54I_d+4.43I_s \quad (1)$$

All the parameters in the model had been described in table (4). To have a 90 percent confidence level, the variance analysis is done for the model and the result is as table (6).

Table 6.The variance analysis of the accident cost model

Source	DF	SS	MS	F	P
Regression	5	25461	28756	7.78	0.019
Residual Error	6	37145	12548	89.88%	
Total	11	28456	R-square		

As can be seen in the tables above, the variables like: age, sex, average daily work hour, disease index, drug usage and sleepiness indexes are statistically meaningful in the model and have an enormous impact on an accident occurrence and its severity which leads to the subsequent costs.The data related to the variance analysis also shows that the model can powerfully predict the accident costs. The amount of the r-square (R2) equals 89.88 percent means that the dependent variable (accident cost) is acceptably explicable by the model presented.

4.5. Validity of the model

To validate the generated model, 20 questionnaires distributed to the new interviewees and the data collected were compared to the main model. The compared data are as shown in diagram (1) below.

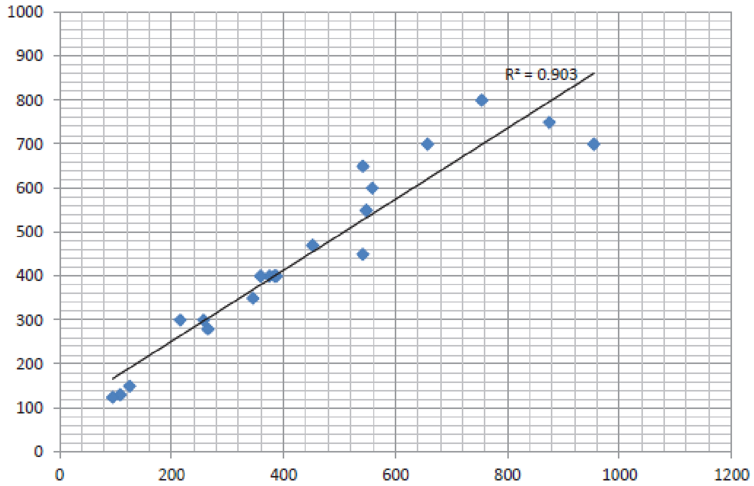

Diagram 1: Real data as opposed to the predicted data of the model

As can be seen in the diagram, the real data concluded from the validated data are in vertical column while the predicted data concluded from the model are in horizontal axle. As is driven from the diagram, the r-square value is equal 0.903 which shows the high fitness of the predicted values to the real ones. Actually, this result proves the validity of the model and the fact that it doesn't fully depend on the applied data.

4.6. Sensitivity analysis of the model

To see the importance of the independent variables on the result of the dependent variable, the sensitivity analysis is done. It means that the greater the change in dependent variable due to changes in independent variable is, the more sensitive the dependent variable to the independent variable would be. In table (7), sensitivity analysis of the model based on the variables is shown.

Table 7: Sensitivity analysis of the model

Independent variable	The amount of increase in dependent variable through 10 percent increase in independent variable	The amount of decrease in dependent variable through 10 percent decrease in independent variable
Age	4.7	4.4
Average daily work hour	2.3	2.8
Disease index	1.5	1.8
Drug usage index	2.2	1.9
Sleepiness index	1.3	1.1

As can be seen in table (7), the generated model is more sensible to the age variable than other variables.

V. CONCLUSIONS

Using regression model for predicting the cost imposed to the traffic system due to accidents and analyzing the model and validating it using statistical method, these results are concluded:

a. Paying attention to the generated model, the variable of average daily work hour is meaning less and the variables related to driver's alertness are meaningful.
b. Paying attention to the analysis of variance, the R^2 is equal 89.88 and p-value of the regression is 0.019 which prove that the dependent variable is suitably explained by the independent variables.
c. Final model shows that the variables of; age, sex, average daily work hour, disease index, drug usage index and sleepiness are statistically meaningful and the other variables are meaningless. This means the direct impact of drivers' alertness criteria on the reduction of the accident costs.
d. Studying the meaningful variables p-value shows that the drug usage and disease indexes are the two most influential variables to the final cost of accidents.
e. Considering the r-square of the model equal 0.903 proves the compatibility of the real and predicted values.

Other valuable studies to continue this research, would be the consideration of the other human factors in assessing the severity of the traffic accidents.

1. REFERENCES

[1] Kaplan, S. and C. G. Prato. (2012) "Risk factors associated with bus accident severity at United States: A generalized ordered logit model", Journal of Safety Research, 43 (3), pp. 171-180.

[2] Stradling JR., Barbour C., (2000) "Prevalence of Sleepiness and its relation to autonomic evidence of arousals and increased inspiratory effort in community based population of men and women", Journal of Sleep 9, pp.381-388.

[3] Dawson, D., Noy, Y.I., Härmä, M., Akerstedt, T., Belenky, G. (2011) "Fatigue modeling: Practices and principles in real world settings", Accidents, Analysis and Prevention, 43, pp. 549-564.

[4] Roth, T., Roehrs TA., (2012) "Etiologies and squeal of excessive daytime sleepiness" Clinther18, pp. 562-576.

[5] Rosenberg ML., Martinez R., (2006) "Graduated licensure: a win-win proposition for teen drivers and parents" Journal of Pediatrics 98(5), pp. 959-60.

[6] Gregersen NP. (2011) " Systematic cooperation between the driving schools and parents in the drivers' education system, an experiment", Accidents, Analysis and Prevention, 26(4): 453-61.

[7] Kennedy, BP., Isaac, NE., Graham, JD. (2008) "The role of the heavy drinking in the risk of traffic fatalities" Journal of Risk Analysis 16(4), pp. 565- 575.

[8] Nicholas, H. (2009) "Review of software to fit generalized estimating equation regression models" The American statistician.

[9] Peden, S., (2014) "world report on road traffic injury prevention" Road safety report.

[10] Ohayon, M.M., Carskadon, M.A., Guilleminault, C., Vitiello, M.V. (2004) "Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan" Journal of Sleep 27, pp. 1255–1273.

2. ATTACHMENT

A sample of questionnaires used for the sleepiness index is as follows. This form adopted from Pittsburgh Sleep Quality Index (PSQI) (Ohayon et al, 2004). Other indexes are extracted from related international sources and translated into Persian.

Name_____________________________________ Date_________

Sleep Quality Assessment (PSQI)

What is PSQI, and what is it measuring?

The Pittsburgh Sleep Quality Index (PSQI) is an effective instrument used to measure the quality and patterns of sleep in adults. It differentiates "poor" from "good" sleep quality by measuring seven areas (components): subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medications, and daytime dysfunction over the last month.

INSTRUCTIONS:

The following questions relate to your usual sleep habits during the past month only. Your answers should indicate the most accurate reply for the majority of days and nights in the past month. Please answer all questions.

During the past month,

1.	When have you usually gone to bed?		_____________ _____________ _____________						
2.	How long (in minutes) has it taken you to fall asleep each night?		_____________ _____________ _____________						
3.	What time have you usually gotten up in the morning?		_____________ _____________ _____________						
4.	A. How many hours of actual sleep did you get at night?		_____________ _____________ _____________						
	B. How many hours were you in bed?		_____________ _____________ _____________						
5	During the past month, how				No	Le	O	Thre	

.	often have you had trouble sleeping because you				t during	ss than	nceor	e or more	
					the past	once a week	twicea	times a week	
					month (0)	(1)	week(2)	(3)	
	A. Cannot get to sleep within 30 minutes								
	B. Wake up in the middle of the night or early morning								
	C. Have to get up to use the bathroom								
	D. Cannot breathe comfortably								
	E. Cough or snore loudly								
	F. Feel too cold								
	G. Feel too hot								
	H. Have bad dreams								
	I. Have pain								
	J. Other reason (s), please describe, including how often you have had trouble sleeping because of this reason (s):								
6.	During the past month, how often have you taken medicine (prescribed or "over the	counter") to help you sleep?							

7.	During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?					
8.	During the past month, how much of a problem has it been for you to keep up enthusiasm to get things done?					
9.	During the past month, how would you rate your sleep quality overall?	Very good	Fairly good	Fairly bad	Very bad (3)	
		(0)	(1)	(2)		
	Scoring					
Component 1	#9 Score		C1 ___________			
Component 2	#2 Score (<15min (0), 16-30min (1), 31-60 min (2), >60min (3))					
Component 3	+ #5a Score (if sum is equal 0=0; 1-2=1; 3-4=2; 5-6=3)		C2 ___________			
	#4 Score (>7(0), 6-7 (1), 5-6 (2), <5 (3)		C3 ___________			
Component 4	(total # of hours asleep) / (total # of hours in bed) x 100					
Component 5	>85%=0, 75%-84%=!, 65%-74%=2, <65%=3		C4 ___________			
	# sum of scores 5b to 5j (0=0; 1-9=1; 10-18=2; 19-27=3)		C5 ___________			
Component 6	#6 Score		C6 ___________			
Component 7	#7 Score + #8 score (0=0; 1-2=1; 3-4=2; 5-6=3)		C7 ___________			
	Add the seven component scores together ___________________		***Global PSQI*** ________________			