territorium 29 (II), 2022, 83-91
journal homepage: https://territorium.riscos.pt/numeros-publicados/
DOI: https://doi.org/10.14195/1647-7723_29-2_6
Artigo científico / Scientific article

A NONLINEAR LOGIT REGRESSION MODEL FOR DETERMINING DRIVERS' ACCIDENT PROBABILITY IN ALGERIA*

UM MODELO DE REGRESSÃO DE LOGIT NÃO-LINEAR PARA DETERMINAR A PROBABILIDADE DE ACIDENTE DE MOTORISTAS NA ARGÉLIA

Bencherif Houria
University Batna 2, Laboratory of Mangement and Transport Logistics (Algeria)
Institute of Health and Safety
h.bencherif@univ-batna2.dz

Lounansa Ramdhane
University Batna 1 (Algeria)
Faculty of Economic and Business Sciences Management
r.lounansa@univ-batna.dz

ABSTRACT

The aim of the study is to estimate a nonlinear regression logit model to calculate the probability of a road accident occurring in Algeria. Referring to the literature, this probability depends on several variables related to the driver, the vehicle and the road environment. Thus, the objective is to show which explanatory variables favour the occurrence of a traffic accident and to what extent. In addition, we seek to estimate the probability that a driver with given characteristics, driving a vehicle with given characteristics, and having a driving licence for a given length of time may have an accident. The structure of the estimation is based on disaggregated data collected following the analysis of files proposed by an Algerian insurance company. The results obtained make it possible to determine the category of variables that has a significance and an important role in explaining the probability of accidents occurring.

Keywords: Road accident, driver, nonlinear regression model, logit, probability, explanatory variables.

RESUMO

O objetivo do estudo é estimar um modelo logit de regressão não linear para calcular a probabilidade de ocorrência de acidente rodoviário na Argélia. Referindo-se à literatura, essa probabilidade depende de diversas variáveis relacionadas ao motorista, ao veículo e ao ambiente viário. Assim, o objetivo é mostrar quais variáveis explicativas favorecem a ocorrência de um acidente de trânsito e em que medida. Além disso, buscamos estimar a probabilidade de acidente um motorista com determinadas características, dirigindo um veículo com determinadas características e possuindo carteira de habilitação com determinada antiguidade. A estrutura da estimativa é baseada em dados desagregados recolhidos na sequência da análise de arquivos proposta pela seguradora argelina. Os resultados obtidos permitem determinar qual a categoria de variáveis com significância e um papel importante na explicação da probabilidade de ocorrência dos acidentes.

Palavras-chave: Acidente de viação, motorista, modelo de regressão não linear, logit, probabilidade, variáveis explicativas.

* *O texto deste artigo foi submetido em 30-11-2021, sujeito a revisão por pares a 17-12-2021 e aceite para publicação em 12-03-2022.*
*Este artigo é parte integrante da Revista **Territorium**, n.º 29 (II), 2022,* © Riscos, ISSN: 0872-8941.

Introduction

Road accidents are a major economic, social and public health problem according to the World Health Organization (WHO). Every year the lives of approximately 1.3 million people are cut short as a result of a road traffic crash. Between 20 and 50 million more people suffer non-fatal injuries, with many incurring a disability as a result of their injury (WHO, 2018). Like many countries around the world, the extent of road insecurity in Algeria is considerable. Each year 4,000 people lose their lives, 60,000 are injured in more than 40,000 accidents, in addition to economic losses of around one hundred billion dinars (Boubakour, 2011; Madani *et al.*, 2020; Hebbar *et al.*, 2020).

Three essential factors contribute to the occurrence of a road accident: the user, the driver and the structure of the road and its environment. Much research has attempted to establish the respective influences of humans, vehicles and the environment on accidents and to specify the nature of their interactions. The search for such an influence requires the use of statistical tools that make it possible to model the occurrence or non-occurrence of an accident in relation to different variables of different kinds. Thus, the research developed was based in particular on dichotomous models consisting in explaining the occurrence of the event that is considered as a function of a certain number of characteristics observed for the individuals in the sample.

Among these dichotomous models, the nonlinear regression logit model was used in this study to calculate the probability of accidents for drivers with regard to some quantitative and qualitative variables such as the age and gender of the driver, the age of the vehicle and the seniority of the driving license. The structure of the estimate is based on disaggregated data collected following an analysis of files proposed by the Algerian insurance companies.

Literature review

Road safety research took its real impetus during the 1970s in most developed countries, faced with increasingly catastrophic road reports. Much work has been done by different researchers who were interested in the problem of road accidents (Bouzigues, 1995; De Lapparent. 2008; GCESD, 2010; Carnis and Lassarre, 2019). Indeed, mathematicians, statisticians, economists such as Oppe, Carré, Lassarre, Gaudry and Elvik have contributed to improving investigations and analyses of road accidents. In addition, they contributed to raising awareness regarding road safety challenges. In this sense, the work carried out called for epidemiological, psychological or economic studies, field surveys or based on statistical analyzes and the application of models (Aloulou and Naouar, 2016; Youssouf Azmi, 2017; Chen and Wei, 2019).

One of the advantages of modeling traffic accidents is to go beyond the stage of simply describing the evolution of road risk indicators, such as the number of accidents and victims reported, to promote a synthetic road risk through research and the development of models that better integrate the complexity of this phenomenon.

We first cite the trivariate probit model that was used in France by Maatigto model risky driving behaviors for a sample of drivers in Paris region in 2010. This involves testing the existence of causal links between three drivers' risk behaviors: use of mobile phones while driving, driving under the influence of alcohol and the use of parking spaces reserved for the disabled (Maatig, 2010). The results show that cell phone use while driving increases with driving under the influence of alcohol and with the use of parking spaces reserved for the disabled. This study has shown that these risky behaviors require preventive measures and must be covered by insurers in order to estimate road risk.

A multinomial logit analysis was developed in Turkey by Çelika and Oktay in 2014 to determine risk factors affecting the severity of road accidents. The analysis covered 11,771 traffic accidents reported by police between January 2008 and December 2013. The estimation results reveal that the main factors that increase the likelihood of fatal injuries are: drivers over 65 years of age , accidents involving one vehicle, accidents occurring on national roads, highways or provincial roads, and the presence of pedestrian crossings. The results also indicate that accidents involving private vehicles or those occurring during peak evening hours in the presence of traffic lights reduce the likelihood of fatal injuries (Çelik and Oktay, 2014).

In the same year, an ordered generalized model was used by Abegaza *et al.*, to examine factors that may influence the severity of road accident injuries in Ethiopia. Data was collected from June 2012 to July 2013 on one of the main and busiest highways in the country. During the study period, a total of 819 road accidents were recorded and investigated. The result of the model estimation shows that alcohol consumption, drowsiness at the wheel, driving at night in the absence of light, weather conditions and being in minibuses or vans are aggravating factors. In this study, risky driving behaviors that may increase the severity of traffic accidents are speeding, alcohol consumption, sleep and fatigue (Abegaza *et al.*, 2014).

In 2015 in Canada, the study of Shamsunnahar looked at the evolution of the severity of road collisions using advanced econometric models. The aim is to examine the influence of the characteristics of drivers, vehicles, roadways, environmental factors as well as road collisions themselves, on the fatalities resulting from these collisions. This work pays particular attention to the severity analysis of drivers' injuries (Shamsunnahar, 2015).

The econometric models developed in this research are estimated using data contained in police databases at regional and national scales of various developed countries. The results of this work have been proposed to enlighten decision-makers on certain measures to be taken, particularly those geared towards drivers and passengers of vehicles in order to reduce fatalities linked to road collisions.

To estimate the severity levels of road accidents in Tunisia, Aloulou and Naouar carried out a work in 2016 using the multinomial Logit model. The severity estimate was linked to three components of the traffic system namely: the driver, the vehicle and its condition of use, and the infrastructure. To describe these components several quantitative and qualitative variables were identified and measured. This involves estimating the probability that a driver with given characteristics and driving in a given traffic environment will be the victim of an accident of a given level of severity. The data collected comes from the study of the survey sheets proposed by the national traffic observatory in Tunisia. The results show that all the variables used are statistically and theoretically significant and explain, to different degrees, the severity of an accident (Aloulou and Naouar, 2016).

In order to detect the characteristics of elderly drivers in physical accidents in comparison with drivers of other age groups in Quebec in 2017, work was carried out by Yousfi. To do this, the author used data from road accidents that occurred in Quebec between 2000 and 2011 (Azmi, 2017). The study consisted of implementing multinomial logistic regression on certain variables of interest such as the presence of road signs, accident severity and driver's condition. The results show a greater involvement of older drivers in crashes at crossings due to disregarding a stop sign or when making a left turn, compared to other drivers. Taking into account the physical vulnerability of the elderly, the risk of being the victim of a road accident with serious or fatal injuries increases significantly for this category of drivers.

In 2019, Chen and Fan has used a multinomial logit model to study and identify significant contributing factors that determine the severity of pedestrian-vehicle crashes in North Carolina, United States. Accident data were taken from the Road Safety Information System database from 2005 to 2012. The results show that factors that significantly increase the likelihood of death and disabling injuries include driver's physical condition, type of vehicle (motorcycle and heavy truck), age of pedestrians (26 to 65 and over 65), weekends, characteristics of the roadway; road class and speed limits. The developed model and the analysis' results provide information on the development of effective countermeasures to reduce the severity of vehicle-pedestrian collisions and improve safety performance of traffic system.

Finally, an ordered logit regression model is proposed in order to understand the influencing factors and the causes of accidents related to the transport of hazardous materials in China by Ma, Zhou and Yang in 2020. The aim of the work is to improve driver safety awareness and help traffic professionals develop effective countermeasures. In total, 343 data concerning accidents involving hazardous materials were collected from the chemical accident information network between April 2018 and May 2019.Four independent variables were retained according to the characteristics of the drivers, hazardous materials, vehicles and road environment. The results of the estimate show that the factors linked to non-compliance with regulations, dangerous driving, and mechanical faults considerably increase the severity of accidents linked to transporting dangerous goods.

It is clear that the models developed in these studies made use of large amounts of data related to the three components of the road traffic system, namely the road user, the vehicle and the road and its environment. Estimating the severity of road accidents is a major research avenue in this field. In addition, the availability and quality of these data constitute an important stage in the design of models and largely determine the quality of the obtained results and their uses.

Method

The study presented in this article is part of the continuation of research carried out in this area in which we opted for the use of the nonlinear regression Logit model to evaluate the probability of accident occurrence from a set of variables. Indeed, among all the models found in the literature we opted for the Logit model that can be applied to the available Algerian data. The use of other models requires specific and detailed information that are not currently available. It is important to note that the availability of data in Algeria is a considered as a serious problem for public authorities and especially for researchers (Himouri, 2005; Bencherif, 2017; Madani, 2020).

The Logit model is a dichotomous nonlinear regression model whose application requires the existence of qualitative and dichotomous variables explained by a set of qualitative and quantitative variables. By dichotomous model, we mean a statistical model which is part of nonlinear regression models in which the explained variable can only take two modalities (dichotomous variable) (Gourieroux, 1984; Thomas, 2000; Hurlin, 2003). Then, it is generally a question of explaining the occurrence or not of an event or a choice.

We observe whether a certain event has occurred and we ask:

$$Yi = \begin{cases} 1 \text{ if the event occurs} \\ 0 \text{ if the event did not occur} \end{cases}$$

Note that the choice of coding (0, 1) is traditionally retained for dichotomous models. Indeed, this makes it possible to define the occurrence probability of the event as the expectation of the coded variable Yi, since:

$$E[Yi] = Pr(Yi = 1) \times 1 + Pr(Yi = 0) \times 0 = Pr(Yi = 1)$$

- E [Yi] = Expectation of the coded variable Yi
- Pr (Yi) = Probability of occurrence of the coded variable Yi

The objective of this work is to use the nonlinear regression logit model to calculate the probability of the occurrence of a road accident for drivers with regard to some quantitative and qualitative variables. The study sample is made up of 1,200 policyholders drawn from the files of the Algerian insurance company. The information in each file was processed to determine whether the insured had an accident. The study period is from July 7 to August 7, 2019. During this period we analyzed the files of the year 2018. It presents some of the files recorded during this year (127,652 policyholders, 22,887 accidents declared including 1045 physical accidents and 21,842 material accidents After data collection, we proceeded to the classification and coding of variables to facilitate analysis by the Eviews software.

Thus, the first step has allowed determining the independent variable of our model that can take two values:

- 1 if the insured had one accident or more;
- 0 if the insured did not have an accident.

The second step determines the quantitative and qualitative independent variables defined as follows:

- The quantitative explanatory variables concern : age of the driver, seniority of the driving license and age of the vehicle;
- Qualitative explanatory variables relate to driver gender.

Each variable contains several classes and can take the value 1 if it belongs to a class and the value 0 if not.

Results and discussion

In this section we will present descriptive analysis of the sample and estimation of the model and discussion

Descriptive analysis of the sample

We proceeded to the classification and coding of variables in order to facilitate analysis by the Eviews 4.0 software.

The results of the classification obtained are:

1. Explained variable (Dependent) defined as follows
   - Y = 1 if the insured has had one or more accidents;
   - Y = 0 if the insured has not had an accident.
2. Quantitative and qualitative explanatory (independent) variables defined as follows:

   The quantitative explanatory variables concern three (3) variables:
   - The driver's age (DA)
   - The seniority of the driving license (SDL)
   - The vehicle age (VA)

Each variable contains several classes and can take the value 1, if it belongs to a class and the value 0, if not.

The overall driver's age includes five dichotomous variables:

$$DA1 \begin{cases} 1 \text{ if } 18 \leq DA \leq 29 \\ 0 \text{ otherwise} \end{cases}$$

DA2: $30 \leq AGC \leq 39$
DA3: $40 \leq DA \leq 49$
DA4: $50 \leq DA \leq 59$
DA5: ≥ 59 years old

According to the obtained results, drivers aged 18 to 29 are the category most exposed to road accidents with 44.71%. These results coincide with the results obtained at a national level that show that young drivers are the category at risk (fig.1).

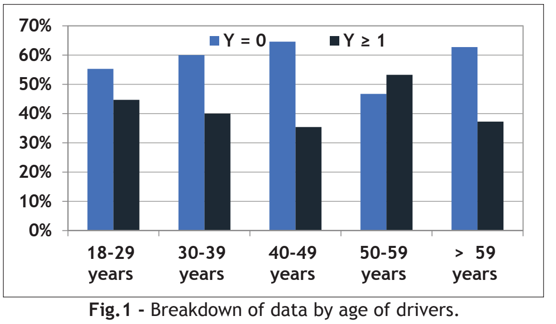

Fig.1 - Breakdown of data by age of drivers.

Fig.1 *- Repartição de dados por idade dos motoristas.*

- All data on the seniority of the driving licens:

$$SDL1 \begin{cases} 1 \text{ if } SDL < 2 \text{ years} \\ 0 \text{ otherwise} \end{cases}$$

SDL2: $3 \leq SDL \leq 5$
SDL3: $6 \leq SDL \leq 9$
SDL4: $10 \leq SDL \leq 14$
SDL5: ≥ 14 years old

The results show that accidents involve much more drivers with driving licenses of less than 2 years with 53.23% and those with licenses of 3-5 years with 50.99% (fig. 2). It is obvious that experience plays a predominant role in the risk of an accident.

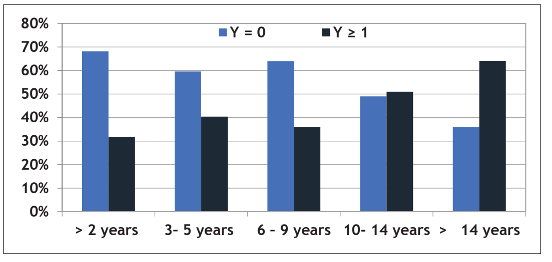

Fig. 2 - Breakdown of data according to length of time driving licence held.

***Fig. 2** - Repartição dos dados de acordo com a antiguidade da carta de condução.*

- All data on the age of the vehicle:

$$VA1 \begin{cases} 1 \text{ if VA Less than 5 years} \\ 0 \text{ otherwise} \end{cases}$$

VA2: 5 ≤ AV ≤ 9

VA3: 10 ≤ AV ≤ 14

VA4: 15 ≤ AV ≤ 19

VA5: ≥ 20 years

The results show that new vehicles are involved in the majority of accidents with 60% of cases, which confirms the results obtained at national level for the same year. The results also show that vehicles over 20 years old register a significant rate with 41.46% (fig. 3).

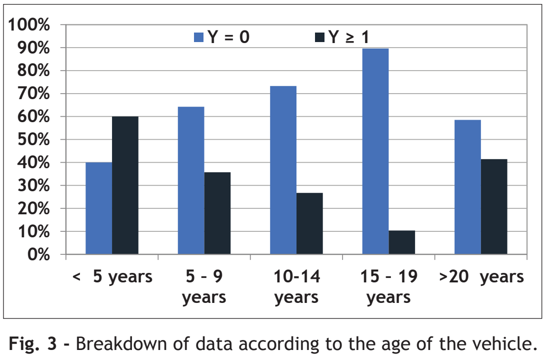

Fig. 3 - Breakdown of data according to the age of the vehicle.

***Fig. 3** - Repartição de dados de acordo com a idade do veículo.*

Qualitative explanatory variables concern one variable: sex of the driver

- All data on the gender of the insured that includes two dichotomous variables:
  - Female (0.1)
  - Male (0.1)

The results show that male drivers are often involved in crashes than women (46.11% vs. 22.32%) (fig. 4). This can be explained by the fact that men drive more than women and tend to adopt a risky driving style, which increases their exposure and further worsens their accident rate.

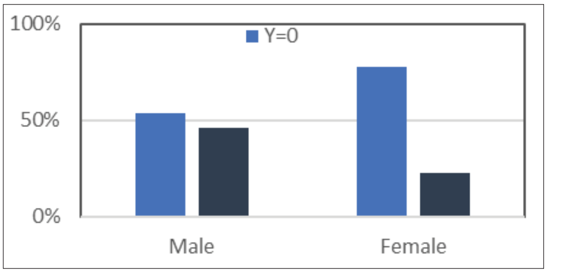

Fig. 4 - Breakdown of data by gender of the driver.

***Fig. 4** - Repartição de dados por gênero do motorista.*

The results of the descriptive analysis made it possible to know the reference variables (the classes most exposed to the risk of accident) which are:

- Driver's age: DA1;
- Seniority of the driving license: SDL1;
- Driver's gender: M:
- Vehicle age: VA1.

From these results, we proposed a model that allows us to calculate the probability of occurrence of an accident.

Estimation of the model and discussion

We will begin the last step devoted to estimating the proposed model based on the use of the Eviews 4.0 software. The proposed model is as follows:

Y* = a*DA + b*SDL + c*VA+ d*gender

Where Y *: Represents the number of unobserved accidents (to be estimated);

In addition, Yi represents the number of accidents observed with:

$$Yi = \begin{cases} 0 \text{ if the insured has not had an accident} \\ 1 \text{ if the insured has had at least one accident} \end{cases}$$

Which means:

$$Yi = \begin{cases} 0 & Yi < 1 \\ 1 & Yi \geq 1 \end{cases}$$

The results of the estimates obtained concerning the probability of having an accident or not were analyzed (Table I).

Table I - Estimated results of the preliminary model.

Tabela I - Resultados da estimativa do modelo preliminar.

Dependent Variable: Y
Method: ML - Binary Logit (Newton-Raphson)
Date: 05/11/2020 Time: 10:51
Sample: 1 1200
Included observations: 1200
Excluded observations: 0
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
DA2	-0.207040	0.175263	-1.181309	0.2375
DA3	-0.360702	0.183576	-1.964867	0.0494
DA4	0.134205	0.216819	0.618972	0.5359
DA5	-0.227787	0.258622	-0.880775	0.3784
VA2	-0.910411	0.175303	-5.193342	0.0000
VA3	-1.280151	0.196615	-6.510959	0.0000
VA5	-0.701968	0.212290	-3.306642	0.0009
VA4	-2.461856	0.277979	-8.856253	0.0000
F	-0.931610	0.186473	-4.995954	0.0000
SDL3	0.175855	0.181610	0.968310	0.3329
SDL2	0.313181	0.199070	1.573223	0.1157
SDL4	0.823460	0.197418	4.171147	0.0000
SDL5	1.263724	0.220968	5.719022	0.0000
C	0.239504	0.175001	1.368587	0.1711

Mean dependent var	0.416040	**S.D. dependent var**	0.493106
S.E. of regression	0.445709	**Akaike info criterion**	1.173222
Sum squared resid	235.0110	**Schwarz criterion**	1.232726
Log likelihood	-688.1736	**Hannan-Quinn criter.**	1.195640
Restr. log likelihood	-812.7410	**Avg. log likelihood**	-0.574915
LR statistic (13 df)	249.1348	**McFadden R-squared**	0.153268
Probability (LR stat)	0.000000		
Obs with Dep=0 Obs with Dep=1	700 500	Total obs	1200

The above table shows the presence of all variables with different categories that explain the dependent variable (probability of occurrence of the accident). We note that there are significant variables, meaning that the probability is less than 0.05 (the level of significance) and others are not significant, Prob more than 0.05, meaning that the significance variables have an impact on the dependent variable. It should be noted that the variables affect the probability of a traffic accident negatively and positively, according to the variation of parameters.

As for the significant variables, they are:

- The variable of Age of the driver:
    - with a significance of DA3 that relates to the category:
        - 40 ≤ DA ≤ 49.
- Vehicle age: all category are significant;
- Seniority of the driving license:
    - With signification of SDL4, that concerns the category:
        - 10 ≤ SDL ≤ 14 and SDL5 which concerns the category over 14 years old.
- In addition, the female gender is significant.

The tests used on the model parameters are the Wald test and the Likelihood Ratio Test (LRT):

Wald test

This involves testing the significance of the parameters of the Bi model in connection with the variables that can influence the phenomenon studied under the following hypothesis:

$$\begin{cases} Ho: Bj = 0 \\ H1: Bj \neq 0 \end{cases}$$

We reject Ho if the significance level is less than 0.05, i.e.Bj has statistical significance. We retain Ho if the significance level is greater than 0.05 that is to say Bj has no statistical significance. Thus, it is clear that all the variables presented in the preliminary model are not necessarily involved in the occurrence of the accident. At first glance, the estimates support that only the variables DA3, VA2, VA3, VA4, VA5, F, SDL4 and SDL5 have statistical significance.

LRT test

Likelihood Ratio Test tests the effect -or not- of the explanatory variables on the dependent variable (probability of occurrence of the accident). The null hypothesis shows that there is no effect; it means that the parameters (coefficients) of variables are equal to zero, while the alternative hypothesis says that the explanatory variables have an effect on the dependent variable meaning that the coefficients of the explanatory variables is different from zero, as shown by:

- Null hypothesis Ho:
    - B1= B2=…………………….…….Bk = 0
- Alternative hypothesis:
    - H1: Bi ≠ 0: i = 1, 2, …………………., k

$$LRT = -2\,[logL(\hat{B}_{ML} - logL(\hat{B}_{CML})] \rightarrow x_q^2$$

x_q^2: Chi distribution of with a degree of freedom q, q = k- & (k number of links);

$logL((\hat{B}_{M})$: Logarithm of likelihood under the Ho hypothesis;

$logL(\hat{B}_{CML})$: Logarithm under hypothesis H1

Bj : parameters of the variables presented in the previous table. We retain Ho if the level of significance α is:

$$LRT = -2\,[logL(\hat{B}_{ML} - logL(\hat{B}_{CML})] < x_q^2$$

We retain H1 if the significance level α:

$$LRT = -2\,[logL(\hat{B}_{ML} - logL(\hat{B}_{CML})] > x_q^2$$

That is, the LRT value is compared with the value in the chi-square table. If the calculated value is greater than the value in the table, we reject the zero hypothesis and retain it in the opposite case.

Through the results of the LRT test obtained, we noticed that the variables of the driver's age, the age of the vehicle and the length of the driving license have a statistical significance at 5%. That is to say, that it has an influence on the variable Y in the calculation of the probability of occurrence of the accident according to the category. After several iterations, the final model was created (Table II).

Table II - Estimated results of the final model.

Tabela II - Resultados da estimativa do modelo final.

Method: ML - Binary Logit (Newton-Raphson)
Date: 11/16/2020 Time: 18:57
Sample: 1 1200
Included observations: 1200
Excluded observations: 0
Convergence achieved after 3 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
DA3	-0.153494	0.127793	-1.201115	0.0050
VA3	0.986100	0.163844	6.018530	0.0000
SDL4	-0.876023	0.158443	-5.528938	0.0000
SDL5	-0.995563	0.182075	-5.467871	0.0000
Mean dependent var	0.417362	**S.D. dependent var**		0.493330
S.E. of regression	0.481356	**Akaike info criterion**		1.314106
Sum squared resid	276.6536	**Schwarz criterion**		1.331096
Log likelihood	-783.1495	**Hannan-Quinn criter.**		1.320506
Avg. log likelihood	-0.653714			
Obs with Dep=0	700	Total obs		1200
Obs with Dep=1	500			

After the simulation process, we can determine the categories of variables having a significance and an important role in the explanation of the accident's probability of occurrence. The variables are:

- Age of driver:
  - AD3 that relates to the category 40 ≤AD≤49;
- Vehicle age:
  - VA3 that concerns the category 10≤VA≤14;
- Seniority of the driving license:
  - Concerns the category 10 ≤ SDL ≤ 14;
  - SDL5: concerns the category over 14 years old.

The final mathematical structure of the model is as follows:

$$Y = -0.1534937363*DA3 + 0.9861000808*VA3 - 0.8760226909*SDL4 - 0.9955634705*SDL5$$

Therefore, the proposed model makes it possible to calculate the probability of having an accident depending on the age of the vehicle, the length of the driving license and the age of the vehicle:

- The more the driver's age decreases, the more the probability of having an accident increases;
- The more seniority of the driving license increases, the more the accident probability decreases;
- The older the vehicle, the more likely the driver is to have an accident.

Conclusion

Road insecurity is a major public health problem around the world, hence the need to strengthen research, particularly in low and middle-income countries in order to better understand the situation and target priority actions to be undertaken (WHO, 2018). Moreover, the development of research in this field has been influenced by the complexity of the phenomenon and by the diversity of researchers, hence the birth of different analytical approaches. In this regard, the use of statistical tools and models helps in understanding the situation and in guiding decision-makers.

In this work, we used the logit nonlinear regression model for the study of road safety based on insurance data. To our knowledge, nonlinear regression logit models have not been used in Algeria to calculate the probability of the occurrence of a road accident. The results obtained from this study have enabled us to identify risk factors among drivers that increase the probability of having an accident, which will allow us to better target road safety information campaigns. In this sense, we have proposed the following to the various decision-makers and road safety actors at different levels:

- Pay particular attention to young drivers by stepping up training, awareness and control measures;
- Give more importance to driving experience as an essential element for driver safety. This can be achieved by revising the current system of obtaining a driving license by opting for a gradual system;
- Strengthen the continuous monitoring systems for older vehicles in order to periodically guarantee their safety.

The results obtained show that the current data available on road accidents can be used to conduct road safety studies in Algeria. However, it would be important to know how to extract useful information from these data, particularly based on the statistical tool. Indeed, the results show that if the latter is well mastered, it can play a fundamental role in identifying the problem and proposing appropriate actions. We have seen at this level that there is still significant room for improvement in the area.

These results may be of great practical interest for public authorities concerned by this phenomenon. They constitute an important source of information that can guide authorities on the most adequate preventive actions to be taken and the most suitable orientations to the road safety policy. These authorities can increase hazard perception through awareness campaigns for young drivers that have new driving licenses and old vehicles.

Several perspectives could be the subject of different studies. The first is to integrate new variables in this model such as the characteristics of the pedestrians, the powers of vehicles and their categories and the characteristics of the road and its environment to have more information on the operation of the traffic system and the different interactions that exist between its components. A second perspective is to use the same model for law enforcement data to be able to make comparisons.

Bibliographic references

Abegaza, T., Berhaneb, Y., Workuc,W., Assratd., A., & Assefae, A. (2014). Effects of excessive speeding and falling asleep while driving on crash injury severity in Ethiopia: A generalized ordered logit model analysis. *Journal of Accident Analysis and Prevention*, vol. 71, 15-21.

Aloulou, F., and Naouar, S. (2016). Micro econometric analysis of road accidents in Tunisia. *Economic review,* Vol. 67, p. 1211-1230. Available at: https://www.cairn.info/revue-economique-2016-6-page-1211.htm

Bencherif, H. (2017). The production of accident data – the role of law enforcement in Algeria. *6th IRTAD Conference, Better data for a improvement of road safety*. 10-11-12 October, Marrakech, Morocco. Available at: https://irtad2017.ma/storage/app/uploads/public/59e/08d/45c/59e08d45c321f766952985.pdf

Boubakour, F. (2011). *Research project on the economic analysis and evaluation of road accidents in Algeria (Anevara - 2007-2010)*. Management-Transport-Logistics Laboratory, University, Batna1. Available at: http://siara.dz.online.fr/anevara/articles.php?lng=fretpg=12

Bouzigues, J-B. (1995). *Road violence*. The European Union and road safety. European Continent Edition. ISBN 2-84176-005-7, 319 p.

Carnis, L., Lassarre, S. (2019). Road safety policy and management. *Road safety in France, When research takes stock and outlines prospects*. Harmattan, 75-93. Available at: https://hal.archives-ouvertes.fr/hal-02531074

Carré, J-R., Lassarre, S., & Ramos, M. (1993). Modeling of road insecurity. Volume 1. INRETS-DERA. Paradigme. France.Çelik, A-K. and Oktay, E. (2014). A multinomial logit analysis of risk factors influencing road traffic injury severities in the Erzurum and Kars Provinces of Turkey. *Accident Analysis & Prevention*, vol. 72, 66-77. DOI: https://doi.org/10.1016/j.aap.2014.06.010.

Chen, Z. and Wei, F. (2019). A multinomial logit model of pedestrian-vehicle crash severity in North Carolina, *International Journal of Transportation Science and Technology*, vol.8 (1), 43-52.

De Lapparent, M. (2006). Empirical Bayesian Analysis of Accident Severity for Motor-cyclists in Large French Urban Areas. *Journal of Accident Analysis and Prevention*, vol. 38(2), 260-268.

De Lapparent M. (2008). Willingness to Use Safety Belt and Levels of Injury in Car Accidents. *Journal of Accident Analysis and Prevention*, vol.40(3), 1023-1032.

Elvik, R., Høye, A. (2018). The potential for reducing the number of killed or seriously injured road users in Norway in the period 2018-2030. 90p. Available at: https://www.toi.no/getfile.php?mmfileid=48658.

Elvik, R., Høye, A. (2021). *Explaining the decline in traffic fatalities and serious injuries in Norway after 2000*. 94 p. Available at: https://www.toi.no/getfile.php/1355883-8906626/Publikasjoner/T%C3%98I%20rapporter/2021/1816-2021/1816-2021-elektronisk.pdf

Gaudry, M. (1993). An overview of DRAG, a comprehensive road safety model. *Economic News*, vol. 69(3), 203-211.

GCESD - GENERAL COUNCIL FOR THE ENVIRONMENT AND SUSTAINABLE DEVELOPMENT (2010). *Study on road safety measures in Europe*. Report N.° 007055-01, 50 p.

Gourieroux, C. (1984). *Econometrics of qualitative variables*. Economica. 2ed.. 300 p.

Hebbar, C., Tires, H., &Hassani, M-M. (2020). Road risk and road safety in Algeria. *Territorium - International Journal of Risks*, issue 27 (I),133-144. DOI: https://doi.org/10.14195/1647-7723_27-1_10

Himouri, S. (2005). *Theoretical model for monitoring road insecurity in Algeria 1970-2002 (Doctoral thesis).* University of Science and Technology, Oran. 114 p.

Hurlin, C. (2003). *Econometrics of qualitative variables. Course handouts*. University of Orleans, 59 p. Available at: http://frederik.ducrozet.free.fr/econometrics/qualitatives/Qualitatif_Chapitre1

Lassarre, S. (1994). *Methodological framing of a modeling for monitoring road insecurity*. Synthèse N°26, INRETS, 52 p.

Ma, C., Zhou, J., & Yang, D. (2020). Causation Analysis of Hazardous Material Road Transportation Accidents Based on the Ordered Logit Regression Model. *International journal of environmental research and public health*, *17*(4), 1259. DOI: https://doi.org/10.3390/ijerph17041259

Maatig, M. (2010). Multivariate modeling of the risk behaviors of car drivers. *42nd Statistics Days*. Available at: https://hal.inria.fr/inria-00494711/document

Madani, A., Chella, T., & Bouder, A. (2020). Road accidents in Algeria, the need for a diagnosis. *Review of cognitive pathways in the social and human* *sciences*, vol. 03(9), 105-122. Available at: https://www.researchgate.net/publication/338528364_Les_accidents_de_la_route_en_Algerie_necessite_d'un_diagnostic/citation/download

Oppe, S. (1991). Macroscopic models for traffic and traffic safety. *Journal of Accident Analysis and Prevention*, vol. 21, 225-232.

Shamsunnahar, Y. (2015). *Advanced Econometric Approaches to Modeling Driver Injury Severity* (Doctorat thesis). Civil Engineering and Applied Mechanics, McGill University, Montreal, Canada. 210 p. Available at: http://www.people.cecs.ucf.edu/neluru/StudentThesis/Shamsunnahar_ethesis.pdf

Thomas, A. (2000). *Econometrics of qualitative variables*. Manual and corrected exercises. Dunod, 179 p. Available at: https://hal.inrae.fr/hal-02836310

WHO – WORLD HEALTH ORGANIZATION (2018). *Global Status Report on RoadSafety*, 424 p. Available at:file:///C:/Users/TEMPDE~1.002/AppData/Local/Temp/9789241565684-eng-1.pdf

Youssouf Azmi, Y. (2017). *Characteristics of road accidents, with at least one injured, in elderly drivers that occurred in Quebec from 2000 to 2011. (Management Sciences Thesis)*. School of Higher Commercial Studies of Montreal. 114p. Available at: http://biblios.hec.ca/biblio/memoires/m62462.pdf