The Influence of Speed and Age on Crash Rate

Bambang Hermanto¹, Leksmono Suryo Putranto¹, dan Dadang M. Ma'soem²

¹Program Studi Doktor Teknik Sipil, Universitas Tarumanagara, Jl. Let. Jend S. Parman No.1 Jakarta 11440
²Program Studi Doktor Teknik Sipil, Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi No.229, Bandung, 40154

Corresponding Author*: bambang.328202003@stu.untar.ac.id, leksmonop@ft.untar.ac.id

ABSTRACT

Purpose: Accidents are one of the parameters that show whether the road safety condition is good or not. In Indonesia, the number of road accidents is still relatively high, so a serious handling of this matter is needed. There are many factors that influence road accidents, but this study will only involve two factors, namely age and speed. The accidents in this study are divided into two, namely minor accidents and serious accidents.

Design/methodology/approach: The method used in this study is a multinomial logit model using a two scenario modelling approach.

Findings: The results showed that the age variable tends to increase the risk of minor accidents compared to severe accidents. This indicates that the older a person is, the higher the risk of a minor accident. However, the speed variable is the opposite, tending to increase the risk of severe accidents when compared to minor accidents. This shows that the more often a person drives at a high speed tends to increase the risk of serious accidents that cause death. Both variables together contribute significantly to the increased risk of accidents.

Paper type: This paper can be categorized as research paper.

Keywords: Age, Accident, Multinomial Logit, Slight Accident, Severe Accident

Received : November 4th
Revised : February 18th
Published : March 31st

I. INTRODUCTION

Road safety is not only a national issue, but also a global problem. More than 1.2 million people die each year on the world's roads, and between 20 and 50 million suffer non-fatal injuries. In most regions of the world, this epidemic of road traffic injuries is on the rise. Tens of millions more victims are seriously injured or disabled each year. Victims who die or are severely injured/disabled will impact their personal and family lives for a long time, even forever. More than 90% of the world's fatalities occur on the roads of low- and middle-income countries, which have only 48% of the world's vehicles. Based on data from the Indonesian National Police, in 2020 there were 100,028 incidents. This figure is lower than in 2019. However, over the last 6 (six) years, there has been an increase in the number of incidents by 3% per year (Polri, Korlantas, 2020).

Many studies have linked the relationship between speed and age to crashes. However, almost all of these studies only stop at the conclusion that accidents or age have an effect on accidents. However, in this study, we will categorise accidents into minor and major accidents. What is meant by a minor accident in this study is an accident that only causes material losses, without human casualties. While what is meant by serious accidents in this study is an accident that causes human casualties, ranging from minor injuries or serious injuries or death.

This study is expected to provide an overview of the relationship between speed and age on crash rates. The accident rate is divided into two categories, namely minor accidents and serious accidents. The object of this research is public transport drivers and the samples in this study are AKAP (Inter-Provincial City-to-Province) drivers with routes or majors in DKI Jakarta from/to provinces in Indonesia with proportions according to the number of vehicles registered at the Ministry of Transportation.
A. Factors Causing Accidents

Based on the NTSC report, driving habits are the most common cause of accidents. As many as 40% of accidents caused by drivers are due to driving habits, 27% are due to driver competence and the third causal factor is 14% due to fatigue (KNKT, 2022).

The factors that cause accidents are divided into three groups, namely vehicle factors, road and environmental factors and human factors. One of the sources of human factors that cause accidents is driving behaviour. The six biggest factors in road traffic accidents are caused by behaviour, namely not being aware of traffic ahead, failing to maintain a safe distance, being careless when turning, being careless when overtaking, exceeding the speed limit, and ignoring the right of way of pedestrians. One of the behavioural causes of accidents is driving speed. Based on data from the Indonesian police, speed is one of the causes of accidents that contributes 8.64% (Polri, Korlantas, 2020).

In addition to the speed factor, the age factor also influences the accident process, although the form of the relationship between age and road accidents varies from one researcher to another (Ryan, Legge, & Rosman, 1998). found the results that age has an influence on the occurrence of accidents. In 2019, it was found that the number of accidents was 85,057 accidents, the three largest percentages of the age of the accident perpetrators were the 30-39 age group, the 40-49 age group, and the 17-21 age group. These three age groups are productive ages. The first rank is the 30-39 age group with a percentage involvement of 17.47% or 14,858. The second place is the perpetrators in the 40-49 age group, which is 15.87% or 13,500. And, the third place is the 17-21 age group with a percentage of involvement of 15.48% or 13,170 (Polri, Korlantas, 2020).

B. Literature Review

There are several previous studies that conducted research on the relationship between speed and age on accident rates. Most of these studies only relate the effect of age or speed on accidents, but in this study, accidents will be divided into two groups, namely minor accidents and severe accidents. (Hermanto, Putranto, & Ma'soem, 2022) found that the driver has a very important role in the process of accidents. According to him, accidents occur because the condition of the driver's ability is below the condition of the task demands faced by the driver.

One of the contributing factors to the driver's ability condition is age. The relationship between age, speed and accidents is shown in Figure 1 below.

![Figure 1. Relationship between Ability and Task Demands to accidents.](source: Hermanto, Putranto, & Ma'soem, 2022)

(Wang, Zhou, Quddus, Fan, & Shou’en, 2018) found that a 1% increase in average speed on urban arterial roads was associated with a 0.7% increase in total crashes, and greater variation in speed was also associated with an increase in crash frequency. (Mohanty & Gupta, 2015) conducted research into the causes of urban and rural crashes. In this study, the variables used consisted of age, gender, driving habits. For driving habits, the variable used was speed while driving. The results showed that for urban areas, the main causes of accidents were caused...
The Influence of Speed and Age on Crash Rate  
Bambang Hermanto¹, Leksmono Suryo Putranto¹, dan Dadang M. Ma'soem²

by age, gender, speed, road conditions and lighting conditions. As for rural areas, the main cause of accidents is still dominated by speed.  
(Danciu, Popa, Micle, & Preda, 2012) He said that older drivers (over 65 years old) are more likely to be dangerous, because at this age, misperceptions of traffic conditions are greater. In addition, aggressive drivers will increase the potential for accidents. (Parker, Reason, Manstead, & Stradling, 1995) concluded that driving behaviour has a positive effect on road traffic accidents. Making mistakes in driving tends to cause active accidents (respondents crashing into other vehicles). Driving experience and age will reduce the risk of accidents.

II. METHODS

The method used in this research is multinominal logit analysis, which is a logistic regression used when the response variable is polychotomous or multinomial, nominal and ordinal scales with more than two categories. Regression models for response variables that are more than two categories must pay attention to the measurement scale. In this study using a logistic regression model with nominal scale response variables. The multinomial logit model is technically almost the same as the binary logit model, which compares the probability of one response with another response. Because there are more than two response variables, there is one response value that is used as a reference and usually uses the smallest response value (Y = 0). By comparing two probability values, the odds ratio value will be obtained with the following equation:

\[
\text{Odd Ratio} = \frac{P(y=1)}{1-P(y=1)} = e^{\theta(x)}
\]

The probability function of multinomial logistic regression with Y = 0 as the reference category for each category can be seen in the equation below.

\[
\pi_0 = \frac{1}{1+\exp g_1(x)+\exp g_2(x)}
\]

\[
\pi_1 = \frac{\exp g_1(x)}{1+\exp g_1(x)+\exp g_2(x)}
\]

\[
\pi_2 = \frac{\exp g_2(x)}{1+\exp g_1(x)+\exp g_2(x)}
\]

A. Sample

The number of samples in this study was 200 respondents. The data was collected by direct interview with the respondents and where the respondents were AKAP public transport drivers with a route to Jakarta. The time required for interviewing each respondent was around 10 to 15 minutes. The data obtained were recapitulated, then validity and reliability tests were carried out for the speed variable. All samples were used for the modelling process.

III. RESULTS AND DISCUSSION

Data was obtained through interviews between surveyors and respondents. The interviewed drivers were randomly selected for each route that had been determined. From the sample data results, for the age variable, the average age of the respondents was 44.17 years with a standard deviation of 8.64.

The speed variable was measured from the results of respondents’ answers related to speed behaviour when driving. The tool they used to measure this variable was the DSQ (Driving Style Questionaire). This questionnaire consisted of 15 questions of which 3 questions related to speed, 3 questions related to calmness, 2 questions related to social resistance, 3 questions related to focus, 2 questions related to trip planning and 2 trips related to deviant behaviour. The scale used is a Likert scale using 6 levels, namely 1 = very rarely/never; 2 = rarely; 3 = quite rarely; 4 = quite often; 5 = often; 6 = almost all the time/always. (French, West, Elander, & Wilding, 1993). A high crash variable score indicates that drivers tend to drive at high speeds and tend to break the speed limit. Low crash variable scores indicate the opposite. The statistical values for each variable are as shown in table 1 below.
The dependent variable in this study was respondents’ involvement in accidents over the past three years. Accident involvement is categorised into three groups. First, respondents who were not involved in any accidents during the last three years (Safety Driver). Second, respondents who during the last three years experienced minor accidents that only caused material losses (Almost Dangerous Driver), such as damage to the vehicle and did not result in human casualties. While the third group is respondents who have experienced serious accidents in the last three years that caused human casualties, either minor injuries, serious injuries or death (Dangerous Driver).

From the survey results, data was obtained for each group. The safety driver group is 46.5%, the almost dangerous driver group is 33% and the dangerous driver group is 27.5%. The percentage for each group is shown in table 2 below.

<table>
<thead>
<tr>
<th>Group</th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety Driver</td>
<td>93</td>
<td>46.50</td>
</tr>
<tr>
<td>Almost Dangerous Driver</td>
<td>66</td>
<td>33.00</td>
</tr>
<tr>
<td>Dangerous Driver</td>
<td>41</td>
<td>20.50</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100.00</td>
</tr>
</tbody>
</table>

A. Logit Model Formation

The modelling process is carried out with two scenarios, the first scenario, the model is only formed from one independent variable. This step is to determine how much influence a single independent variable has on the tendency of potential accidents. The second scenario, the model is formed jointly by two independent variables, namely age and speed. With this scenario, the amount of influence of the independent variables on the potential for accidents will be obtained after being combined by two independent variables. In the process of model building, the reference group is the first group, namely safety drivers.

B. First Scenario

The first stage in the modelling process is the Goodness of Fit test. This stage is carried out to determine whether the multinominal logit model is appropriate or not. The null hypothesis used is that the model is fit. Based on table 3 below, it can be seen that both models for age and speed have a significance value of more than 5%, so it can be concluded that the model is suitable for use.
The model fit test can also be done by comparing the chi-Square between the model that only involves constants and the model that includes independent variables. The calculation results are shown in table 4 below. From the calculation results obtained for the age model is significant at the 10% error level while for the speed model is significant at the 5% error level. This indicates that the logit model can proceed to the next stage.

To determine how much influence the variability of the independent variables had on the model, the Pseudo R_Square values were used as shown in Table 5 below. Based on this table, it can be seen that in both the age model and the speed model, the independent variables contribute less than 10%. This suggests that there are other independent variables that influence the model. Although both the age and speed variables contributed relatively little individually, this does not detract from the quality of the model.

To find out whether the independent variables should be included in the model, the chi-square test is used, namely by comparing the likelihood between the final model (model by including independent variables) and the reduced model (model without involving independent variables). The null hypothesis used is that there is no effect of independent variables on the model. Based on Table 6 below, it can be seen that for both the age variable and
speed variable models, both have significance values below 5%. It can be concluded that the independent variables individually influence the model.

<table>
<thead>
<tr>
<th>Effect</th>
<th>-2 Log Likelihood of Reduced Model</th>
<th>Chi-Square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Models For Ages</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Konstanta</td>
<td>183,73</td>
<td>9,12</td>
<td>2</td>
<td>0,01</td>
</tr>
<tr>
<td>Usia</td>
<td>183,73</td>
<td>9,12</td>
<td>2</td>
<td>0,01</td>
</tr>
<tr>
<td><strong>Models for speed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Konstanta</td>
<td>104,60</td>
<td>21,34</td>
<td>2</td>
<td>0,00</td>
</tr>
<tr>
<td>Kecepatan</td>
<td>94,22</td>
<td>10,96</td>
<td>2</td>
<td>0,00</td>
</tr>
</tbody>
</table>

The last stage in the process of forming a logit model is to test the feasibility of the coefficients of the model. By using Wald test with the results as shown in table 7 below. For the age variable model, the results show that the constant (intercept) and the coefficient of the age variable are significant only in the Almost Dangerous Driver group equation or the group that tends to have the potential for minor accidents. As for the Dangerous Driver group equation, neither the constant (intercept) nor the coefficient of the age variable is significant.

For the speed variable model, the constant (intercept) is significant in both almost dangerous driver and dangerous driver equations. While the coefficient of the speed variable is only significant in the dangerous driver group.

<table>
<thead>
<tr>
<th>Group</th>
<th>B</th>
<th>Wald</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Models for Ages</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Almost Dangerous Driver</td>
<td>Constant</td>
<td>-2,56</td>
<td>7,79</td>
<td>0,01</td>
</tr>
<tr>
<td>Age</td>
<td>0,05</td>
<td>6,13</td>
<td>0,01</td>
<td>1,050</td>
</tr>
<tr>
<td>Dangerous Driver</td>
<td>Constant</td>
<td>-0,19</td>
<td>0,041</td>
<td>0,84</td>
</tr>
<tr>
<td>Age</td>
<td>-0,06</td>
<td>0,45</td>
<td>0,50</td>
<td>0,94</td>
</tr>
<tr>
<td><strong>Model for Speed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Almost Dangerous Driver</td>
<td>Constant</td>
<td>-0,85</td>
<td>2,96</td>
<td>0,08</td>
</tr>
<tr>
<td>Speed</td>
<td>0,08</td>
<td>0,07</td>
<td>0,27</td>
<td>1,08</td>
</tr>
</tbody>
</table>
Based on table 7 above, the logit equation for the model on the age variable there is only one equation formed, namely in the dangerous driver group as follows:

\[ \pi_3 = -2.56 + 0.05(\text{usia}) \]  \hspace{1cm} (5)

While in the model for the independent variable speed there are two equations formed, which are as follows:

\[ \pi_2 = -0.85 \]  
\[ \pi_3 = -2.59 + 0.26(\text{kecepatan}) \] \hspace{1cm} (6)  
\[ \pi_3 = -2.59 + 0.26(\text{kecepatan}) \] \hspace{1cm} (7)

C. Second Scenario

In the second scenario, the variables age and speed are included in the logit model. By entering them together, there will be a change in the interaction between the independent variables and the dependent variable as well as between the independent variables. This interaction will of course affect the logit model. The stages in the model building process are the same as in the previous scenario. The first is the model fit test (Goodness of fit). The null hypothesis is that the model is fit. Based on the chi-squared values, both Pearson and Deviance, both have a significance value greater than 5%, so it can be concluded that the model is suitable for use. The chi-squared values for both calculations are shown in Table 8 below.

**Table 8. Model Conformity Test**

<table>
<thead>
<tr>
<th></th>
<th>Chi-Square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson</td>
<td>259.35</td>
<td>256</td>
<td>0.43</td>
</tr>
<tr>
<td>Deviance</td>
<td>268.50</td>
<td>256</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Apart from using table 8 above, the model test can also be done by looking at table 9 below. By using the null hypothesis that all independent variables have no effect. Based on the calculation results seen in table 9 below, the significance value is less than 5%. So it can be concluded that all independent variables have an influence on the logit model.

**Table 9. Test Model (Simultaneous Test)**

<table>
<thead>
<tr>
<th>Model</th>
<th>Model Fitting Criteria</th>
<th>Likelihood Ratio Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2 Log Likelihood</td>
<td>Chi-Square</td>
</tr>
<tr>
<td>Intercept Only</td>
<td>346.52</td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>326.26</td>
<td>20.26</td>
</tr>
</tbody>
</table>

To determine the amount of influence of the variability of the independent variables on the model, it can be seen in the Pseudo R-Square value shown in table 10 below. There are 3 (three) Pseudo R-Square values, the Nagelkerke value reads 0.11 which means that the independent variable contributes 11% variability to the model, as well as for the other two values.
After testing the model, we then tested the independent variables that should be included in the model. Firstly, a partial test was conducted to determine whether each parameter of the model has a significant influence on crash involvement. The null hypothesis is that each parameter has no influence on crash involvement. The hypothesis will be rejected if the significance value is greater than 5%. Based on table 11 below, it can be seen that all parameters have significance values below 5%, which means the null hypothesis is rejected. Therefore, it can be concluded that all parameters in the model have a significant influence on crash involvement.

Tabel 11. Partial Test for Logit Parameters

<table>
<thead>
<tr>
<th>Effect</th>
<th>-2 Log Likelihood of Reduced Model</th>
<th>Chi-Square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>338,99</td>
<td>12,73</td>
<td>2</td>
<td>0,00</td>
</tr>
<tr>
<td>Age</td>
<td>337,34</td>
<td>11,08</td>
<td>2</td>
<td>0,00</td>
</tr>
<tr>
<td>Speed</td>
<td>335,56</td>
<td>9,30</td>
<td>2</td>
<td>0,01</td>
</tr>
</tbody>
</table>

After going through the stages above, the logit model equation will be obtained. In this study, there are three categories, so there is one category that will be used as a reference category, namely the first category (safety driver). Based on the calculation results, the results are as shown in table 12 below.

Tabel 12. Parameter Estimation

<table>
<thead>
<tr>
<th>Group</th>
<th>B</th>
<th>Wald</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Almost Dangerous Driver</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0,06</td>
<td>8,04</td>
<td>0,00</td>
<td>1,02</td>
</tr>
<tr>
<td>Speed</td>
<td>0,14</td>
<td>3,39</td>
<td>0,06</td>
<td>0,99</td>
</tr>
<tr>
<td>Constant</td>
<td>-2,89</td>
<td>4,79</td>
<td>0,03</td>
<td></td>
</tr>
<tr>
<td>Dangerous Driver</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0,01</td>
<td>0,06</td>
<td>0,81</td>
<td>0,96</td>
</tr>
<tr>
<td>Speed</td>
<td>0,26</td>
<td>9,87</td>
<td>0,00</td>
<td>1,10</td>
</tr>
</tbody>
</table>
From table 12 above, two logit equations are obtained, where the reference is the first group (safety drivers), with the following equation:

a. Kelompok kedua (Almost dangerous driver)
$$\pi_2 = -3.95 + 0.06 \text{(usia)} + 0.14 \text{(kecepatan)}$$

b. Kelompok ketiga (Dangerous driver)
$$\pi_3 = -2.89 + 0.26 \text{(kecepatan)}$$

D. Discussion

Observations regarding the effect of age and speed on crash rates were made in two scenarios. In the first scenario, the model included only one independent variable. This scenario was conducted to determine how much influence the independent variables individually have on the increased risk of crashes for each crash level. The second scenario, the model is composed by two independent variables together. Including the two independent variables together will certainly give an effect caused by the interaction of the two independent variables. There are two types of effects in the model analysis, the first is a positive effect, which means that if the independent variable increases, the risk of accidents will decrease and vice versa. While the positive effect means that if the independent variable increases, the risk of accidents also increases. The analysis is more focused on the odds ratio value which in the parameter estimation table is shown in the exp (B) column.

Based on the model formed, the age variable individually only influences minor accidents. The older the age of a driver, the higher the risk of minor accidents. Indeed, the increase in the risk of accidents is only 0.05, but if a driver already has a high risk of accidents, then an older age will increase the risk of minor accidents by 1.05 times every additional 1 year of a driver’s age. Some studies that include the age variable in their research give varying results. (Chu, Wu, Zhang, & Ozkan, 2019) found that as people get older, they are more likely to commit driving offences. This behaviour will certainly increase the risk of accidents. However, there are some researchers who have come to different conclusions, for example (Busuhazy, et al., 2020), concluded that the older a person is, the lower the risk of accidents. Findings with similar results were also obtained by (Lady, Rizqandini, & Trenggonowati, 2020) which concluded that with increasing age the risk of accidents decreases.

Previous studies only discussed the effect of age on accidents, but in this study, accidents are divided into two categories, namely light and heavy accidents. From the previous discussion, age has a positive effect on minor accidents, but in the third group (dangerous drivers), the effect of age is negative with logit values of 0.94 (first scheme) and 0.99 (second scheme). This indicates that the older a person is, the lower the risk of serious accidents. However, in terms of significance, this conclusion is not significant. It is possible that if the sample is increased, it will give a different result.

The results above show that if in previous studies the effect of age on accidents varies, some are positive and some are negative, but the results of this study show that age has a positive effect on minor accidents and age has a negative effect on serious accidents.

Regarding the effect of speed on crash rates, based on the modelling results above, it can be seen that speed has a negative effect on the risk of minor crashes but a positive effect on serious crashes. The negative effect on minor crashes does not mean more safety, but more likely to increase severe crashes. This can be seen when the risk of severe accidents is compared to the risk of minor accidents with the following calculation:

$$\frac{P.\text{Kecepaan Ringan}}{P.\text{Tidak Celaka}} = e^B = 0.99$$ (10)

$$\frac{P.\text{Kecepaan Berat}}{P.\text{Tidak Celaka}} = e^B = 1.10$$ (11)

If equations (10) and (11) are substituted, we get:

$$\frac{P.\text{Kecepaan Berat}}{P.\text{Kecepaan Ringan}} = \frac{1.10}{0.99} = 1.11$$ (12)

Equation 12 shows that the potential for serious crashes to occur is greater than that for minor crashes due to a driver’s speed behaviour.

There have been several studies that have included the variable speed in relation to crashes and all results show that speed is positively correlated with crash occurrence. (Mohanty & Gupta, 2015) concluded that one of the causes of accidents outside the city is speed, even when combined with other factors. (Aarts & Van Schagen, 2006) stated that vehicles travelling faster than the surrounding speed have a high potential for accidents, the higher the difference, the higher the risk of accidents. (Elvik, 2005) found that changes in average speed can
increase crash risk and casualty severity. A 10 per cent change in average speed will increase the risk of fatality even more.

IV. CONCLUSION

A. Conclusion

Age and speed are two variables that strongly influence crashes. Their influence on crashes is not only linked individually but when combined they have a greater influence. Age has a different impact on minor and major crashes. The higher the age of a driver, the higher the risk of a minor crash. Whereas driving at a higher speed increases the risk of a serious crash. Speed will increase the potential for severe crashes more than minor crashes. However, the combination of both variables tends to increase crash risk. The difference in the effect of age and speed on the two types of accidents certainly gives different meanings in handling them. This suggests that speed in driving will increase the risk of death.

B. Suggestion

This research is a development of previous research related to the effect of age or speed on accidents. However, in this study, accidents are divided into two categories, namely minor accidents and severe accidents. This separation of accidents will make it easier to distinguish the tendency of a strong influence on the risk of death. However, due to the limited data available, this research is still lacking and in the future it is necessary to develop or expand both sample size and variable coverage, so that it will increase the strength of the model formed.

REFERENCES


