Discussion on the Influence of Weather on Expressway Traffic Accidents in time based on Survival Analysis Model

Chen Chen[1, a,*], Yishuai Tian[2, b]

[1]Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

[2]College of Hydraulic & Environmental Engineering, China Three Gorges University, Yichang 443002, China

[a]qazplm123098chen@yeah.net, [b]mohanic0307@gmail.com

*Corresponding author

Keywords: Survival analysis, Traffic accident risk function, Weather impact.

Abstract: to explore the influence of weather on highway traffic accidents in time span, 2219 pieces of effective traffic accident data of Kunshi Expressway in 2009 and 2012 were extracted in this paper. A survival analysis model is established to analyze the traffic accident data under sunny, rainy and cloudy days, and its risk function is discussed. The results show that 17 hours is the high incidence period of accidents, and 23 hours is the period with the highest risk of traffic accidents. Rainy days have the greatest impact on traffic accidents, followed by cloudy and sunny days. Based on the survival analysis model, this paper puts forward relevant suggestions for traffic management.

1. Introduction

In recent decades, China has witnessed rapid economic and social development and continuous expansion of transportation infrastructure. From 2005 to 2014, the mileage of roads increased from 19.305 million km to 4.463 million km. However, adverse weather conditions such as rain, snow, fog, and ice will worsen road traffic, reduce traffic efficiency, and easily cause major accidents, resulting in serious casualties and a lot of traffic, especially for high-speed highways. Long-distance and other characteristics, so it is very important to study the influence of weather on highway traffic accidents in the time span.[1]

In the study of traffic safety, there have been many studies on the special meteorological environment of an expressway, such as rainy and cloudy days. For example, COOLS [2] and others have analyzed and studied the measured data and obtained the influence of meteorological environment such as rain and snow on traffic density. Abdel-Aty [3] based on the analysis of road traffic accident data in Florida from 2003 to 2007, we found the effect of foggy weather on traffic accidents and found that frontal collisions and rear-end collisions are the most common in foggy weather, and are more likely to occur without lighting at night. Although the content of the research is rich, the research on the time is less. Therefore, based on the survival analysis, this paper discusses the impact of expressway traffic accidents on the time span, establishes a survival analysis model, and analyzes the traffic accident data under sunny, rainy and cloudy days, so as to provide relevant suggestions for highway safety control. [4]

2. Based on the establishment of survival analysis of Kunming-Shijiazhuang Expressway

2.1 Project Overview

Kunshi Expressway in China starts from Shihuguan in the eastern suburb of Kunming and ends in Shilin Scenic spot in Shilin County. The total length is 78.07 km, with a geographical location of 24 °50km north latitude 25 °02' north longitude 102 °43km-103 °20' east longitude. Kunshi Expressway in China is a key highway construction project in the large-scale development of western China, which was completed and opened to traffic on November 16, 2003. Access to China's Yunnan

Province traffic accident data 2009 Murray 2012 analysis report, obtained 2129 valid traffic accident data. After sorting out, it is found that there are four kinds of weather conditions in traffic accidents: rainy, sunny, cloudy and foggy. As there are only two cases of foggy weather, it does not have the meaning of statistical analysis. Therefore, the influence of foggy weather is not considered this time.

2.2 Establishment of Survival Analysis

Survival analysis has two analysis modes, namely Kaplan-Meier analysis, usually referred to as KM analysis. The process of survival analysis uses the multiplicative active limit method to estimate the survival rate. At the same time, it can also test an influencing factor. It is the most basic survival analysis method, which usually describes the influence of a variable. And Cox regression analysis, also known as the proportional hazard model, is an important model in survival analysis, which can analyze the multi-quantitative relationship among many risk factors with irregular survival time and complete or truncated state. This time, Kaplan-Meier was used for survival analysis. [5]

1. The probability of traffic accident distribution is expressed by the distribution function according to the probability statistics and the probability distribution of the accident. The function is as follows.

$$F(t) = P(T \le t) = \int_0^t f(x)\,dx$$

$$f(x) = \frac{dF(t)}{dt} = \lim_{\Delta t \to 0} \frac{P(t,\ t+\Delta t)}{\Delta t}$$

Where $F(t)$ represents distribution function, P represents probability, T represents the number of accidents, $f(x)$ Indicates the probability density of the moment。

2. Calculate the survival function: when the survival probability is low, the survival curve is steep. When the survival probability is high, the survival curve is flat.

$$s(t) = P(T > t) = \int_t^{\infty} f(x)\,dx = 1\text{-}F(t)$$

3. Traffic accident risk function: The risk function can also be expressed by the distribution function $F(t)$ and the probability density function $f(t)$. The formula is as follows:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t+\Delta t \mid T \ge t)}{\Delta t} = \frac{f(t)}{s(t)} = -\frac{d}{d}s(t) = \frac{f(t)}{1\text{-}F(t)}$$

The cumulative risk function curve is obtained by integrating the risk function. The higher the position, the higher the probability of ending the traffic accident event in Δt.

4. Nonparametric $Kaplan-Meier$ method. Different values and they are directly arranged from small to large $t_1 < t_2 < \cdots < t_k$. The survival function of the traffic accident duration is the estimation function of $s(t)$ as shown in the equation:

$$s(t) = \prod_{t_j \le t} \frac{n_j - d_j}{n_j}$$

Where n_j is the number of samples that existed before the moment t_j, that is, the sample size that traffic accident continues to be; d_j is the number of samples for which congestion has ended during this unit time period t_j; $\frac{n_j - d_j}{n_j}$ is the probability of survival for the moment t_j

$$\hat{S}(\mathrm{t}) = \frac{\Omega_t}{\Omega}.$$

Here $\hat{S}(\mathrm{t})$ is an estimate of $S(\mathrm{t})$, indicating the probability that the congestion duration is longer than t if a congestion event has occurred. The risk probability formula is expressed as follows:

$$\hat{h}(t)=\frac{A}{B\Delta t}=\frac{C}{B}$$

Then get the average risk rate: the number of samples ending in the unit time interval of a unit divided by the number of samples in the interval where the conceptual congestion does not end, the formula is as follows:

$$\hat{\mathrm{h}}(t)=\frac{A}{B\text{-}\frac{1}{2}D}$$

Where A is Number of samples ending in congestion at the beginning of time T;B is Number of samples with no congestion after time T;C is Number of samples congested per unit time after time T;D is Number of samples ending in the interval. Δt is Interval width.

3. Results and result analysis

3.1 Analysis based on Survival function

Based on the SPSS data analysis platform, this paper analyzes the survival analysis of the three kinds of weather for the accident in one day, based on the analysis report of 2009 Mel 2012 traffic accident data in Yunnan Province, China, and based on the SPSS data analysis platform. As shown in figure 1.

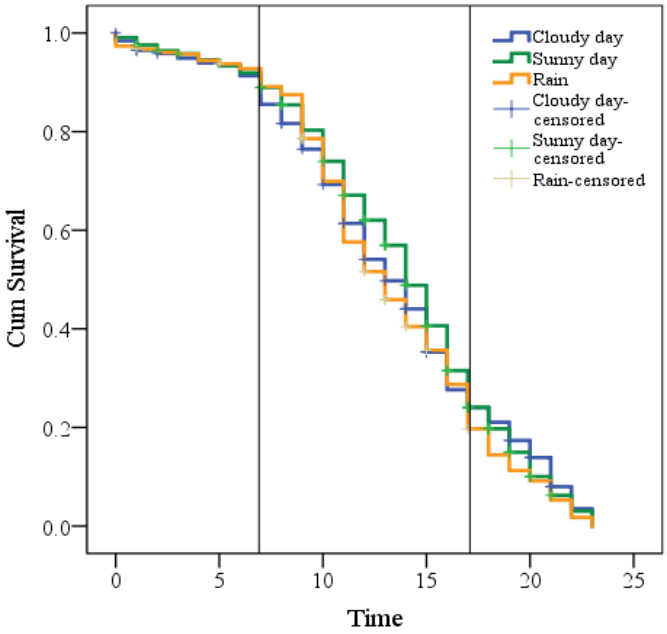

Fig. 1 Survival index analysis chart

As can be seen from figure 1, there are roughly three stages in a day, and the survival index of the three kinds of weather decreases slowly during the 7-hour period of 0muri. It shows that there is no obvious difference in the impact on the occurrence of the accident during the 7-hour period of 0muri. Mainly due to the late weather, the number of driving is less, at this time the probability of accidents is less. In 7Mel 17 hours, the decline rate of survival index of the three kinds of weather decreased rapidly, and in each hour, the survival index of sunny days was higher than that of cloudy days and greater than that of rainy days. It can be explained that the accident rate is high during the 17-hour period, with the greatest impact on rainy days and the least impact on sunny days. At 23 hours, the survival index decreased slowly, mainly because the time became late, the number of traffic decreased, and the frequency of accidents entered a stable period.

3.2 Analysis based on the risk function

In order to analyze the impact of weather changes on driving risks in different time periods, and provide relevant suggestions. Based on the SPSS platform, the risk function of the influence of Kunshi highway on weather in the time span is calculated. The result is shown in figure 2

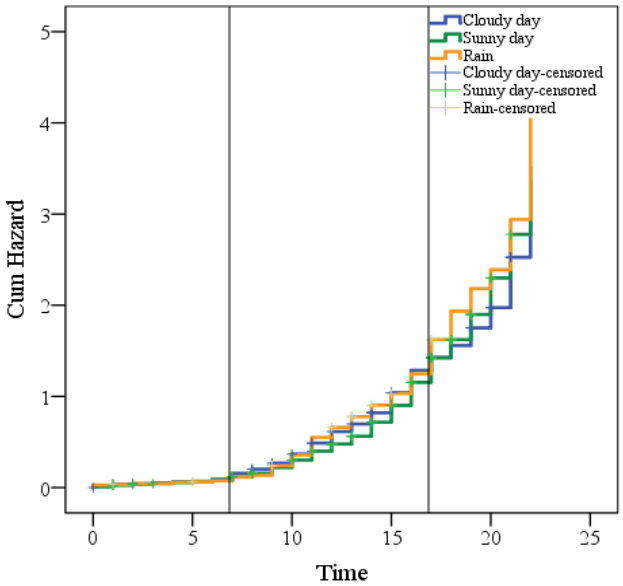

Fig.2 risk function diagram

As can be seen from figure 2. The function as a whole can be divided into three stages. When the time is 7 hours, the traffic flow is less and the accident rate is small, so the increment of the risk function is less, and the three lines coincide as a whole. During the 17-hour period, due to the increase of traffic flow during the day, the risk function will gradually increase, and the value of risk in rainy days is the highest among the three, followed by rainy and sunny days, which may be caused by blurred vision and skidding road caused by rain. During the 23-hour period, the risk function increased sharply, probably due to the darkening of the sky, affecting the driver's line of sight, and driving for a long time will make the driver feel tired, so the risk of driving will increase sharply.

4. Conclusion

By fetching the traffic accident data of Yunnan Province of China in the 2009 Mel 2012 analysis report, we have obtained 2129 valid traffic accident data. A survival analysis model is established to analyze the traffic accident data under sunny, rainy and cloudy days, and its risk function is discussed. In the survival analysis: 7 Mel 17 hours is the high incidence period of accidents, and is greatly affected by rainy days. In the risk analysis, 17Mel 23 hours is the period with the highest risk of traffic accidents, rainy days have the greatest impact on traffic accidents, followed by cloudy and sunny days. Based on the survival analysis model, it provides relevant suggestions for the safety management of expressways.

References

[1] Hamed, M.M. and B.M. Al-Eideh, An exploratory analysis of traffic accidents and vehicle ownership decisions using a random parameters logit model with heterogeneity in means. Analytic Methods in Accident Research, 2020. 25.

[2] Cools, M., E. Moons and G. Wets, Assessing the Impact of Weather on Traffic Intensity. WEATHER CLIMATE AND SOCIETY, 2010. 2(1): p. 60-68.

[3] Abdel-Aty, M., et al., A study on crashes related to visibility obstruction due to fog and smoke. 43(5): p. 1730-1737.

[4] Rok, K., N.P. Matjaž and T. Darja, Using the scanners and drone for comparison of point cloud accuracy at traffic accident analysis. Accident; analysis and prevention, 2020. 135.

[5] Guosen, Z., et al., OSpaad: An online tool to perform survival analysis by integrating gene expression profiling and long-term follow-up data of 1319 pancreatic carcinoma patients. Molecular carcinogenesis, 2020. 59(3).

[6] V, P.R., et al., Clinical utility of 177 Lu-DOTATATE PRRT in somatostatin receptor-positive metastatic medullary carcinoma of thyroid patients with assessment of efficacy, survival analysis, prognostic variables, and toxicity. Head & neck, 2020. 42(3).