Original Research Paper

Road Safety When Drivers Use Alcohol and Marijuana: Confounded Poisson Distribution Helps to Understand

1Ramalingam Shanmugam, 2Ram Tripathi and 3Karan Singh

1Honorary Professor of International Studies, School of Health Administration, Texas State University, 601 University Drive, San Marcos, TX 78666, USA
2Department of Management Science and Statistics, University of Texas at San Antonio, San Antonio, TX 78249, USA
3Professor and Chair, Department of Epidemiology and Biostatistics, UT Health Northeast, 11937 U.S. Hwy 271, Tyler, Texas 75708-3154, USA

Article history
Received: 19-02-2017
Revised: 06-05-2017
Accepted: 25-05-2017

Corresponding Author:
Ramalingam Shanmugam
Honorary Professor of International Studies, School of Health Administration, Texas State University, 601 University Drive, San Marcos, TX 78666, USA
Email: rs25@txstate.edu

Abstract: Self-admitted users of alcohol and/or marijuana while driving haunt road safety regulating agencies in general and community groups such as Mothers Against Drunk Driving (MADD) in particular. Motivated by such data in a survey, by 6,612 drivers above 18 years old in USA during years 2013-2015, we thought of interpreting how the road safety is in the scenarios. Needless to emphasize the importance of any breach of road safety to nursing professionals working in emergency unit of a hospital as the victims of road accidents arrive there to get a preemptive service with urgency. To be educated on the issues, this article develops and uses an appropriate probability model, which is named Confounded Poisson Distribution (CPD). Statistical properties of CPD are derived and applied to analyze the data and to demystify the patterns in road safety among impaired drivers in the age brackets 18 through 75+ years.

Keywords: Confounded Poisson Distribution, Road Safety, Psychologic Temptation

Introduction

Mannering and Bhat (2014) narrated the evolution of methodological applications that are used in road accidents research. They alerted about issues of unobserved heterogeneity and endogeneity of drivers in accidents data analysis with an intention to improve road safety.

There has been a progress in developing suitable statistical methodologies to assess road safety. In spite of such progress over the years, new methodologies keep appearing in the literature to address complex issues in road accidents. Such efforts could lead to effective road safety policies. It is with this aim, we prepared this article with an innovative statistical modeling of perceptions on impaired drivers. Occurrences of highway accidents are too complex to configure without involving heterogeneity among drivers with external stimuli, which are more often than not alcohol and marijuana. This is an important current topic, given an increase in marijuana availability is occurring in the U.S. due to the enactment of medical and recreational marijuana laws.

This article offers an interesting statistical approach to analyze the perceptions on impaired driving. It builds on a relatively recent (Shanmugam, 2011) contribution and this article adds something new to the literature. Accidents analysts are not able to access to all micro information in data about stimuli. The absence of such information is impediment to perform traditional statistical analyses and it is referred to as unobserved heterogeneity. Such fundamental issues are well articulated in Mannering et al. (2016), who eloquently advocated the importance of addressing unobserved heterogeneity of impaired drivers in highway accident data analysis. A main research direction they offered is to build in the heterogeneity as correlated parameters. We have exactly constructed this concept as the foundation of this article.

Road safety is a top priority not only to the National Safety Board and State agencies, but also to others including Mothers Against Drunk Driving (MADD). However, the fourth major cause of deaths in the USA is traffic accidents (http://cdc.gov/nchs/fastats).
Mainly, the root cause of the fatalities is driving impairment. Blood Alcohol Concentration (BAC) is a serious cause of impairment. BAC is quantified by a breath test, oral fluid test, or blood test. After a lapse of a few hours, the BAC diminishes and the driver might recover full driving skill.

Another cause of impaired driving is the use of marijuana (https://drugabuse.gov).

Legal or illegal use of marijuana is increasing. Marijuana, also known in botanical literature as cannabis, is used to cure mental and physical defects. Some short term effects of using marijuana are memory loss, impaired motor skills and feelings of paranoia, whereas, long term effects of its use are addiction and behavioral problems (Ashton, 2001).

Since 1996, sixteen states and the District of Columbia in the USA have decriminalized the use of marijuana for medical reasons (Li et al., 2011). Marijuana alleviates symptoms of cancer, HIV/AIDS, multiple sclerosis, Alzheimer’s disease, post-traumatic stress disorder, epilepsy, Crohn’s disease and glaucoma. Cannabis is an effective alternative to narcotic painkillers. It is not clear how long after its use marijuana might impair driving, because its chemical remains are detectable in body fluids for weeks after intoxication (Li et al., 2011; Hartman and Huestis, 2013).

Marijuana users frequently combine its use with drinking alcohol. The effect of marijuana does not dissipate away but rather stays longer in organs. Consequently, marijuana’s effect could worsen the impairment when the driver is drunk (i.e., has a significant BAC). In this article, we explore the data (in Arnold and Tefft, 2016) on perceptions of road safety in the USA and look at whether marijuana plus alcohol has a more negative effect on road safety than either alcohol alone or marijuana alone. We develop and use a probability model, which is named Confounded Poisson Distribution (CPD). We also investigate the statistical properties of CPD in detail.

Data Clues and Literature Review about Road Safety

First, let us look at data about self-admitted users of alcohol and/or marijuana by the drivers in the age brackets: 18-24, 25-39, 40-59, 60-74 and 75+ years in USA during 2013-2015 in Table 1 for data clues. The self-admitted users (per 100,000) increases until age 59 years and then decreases; the overall average is 1.322. Interestingly, the average number of users of alcohol alone, marijuana alone and both alcohol and marijuana are 208, 89.5 and 44.6, respectively. Marijuana is less used compared to use of alcohol. However, contrary to any intuitive common sense, a simultaneous consumption of alcohol and marijuana is less frequent. The aim of this article is to explain this apparent non-triviality by a statistical methodology. In other words, does the consumption of alcohol or marijuana simply kicks in an extra psychologic temptation to go for both? If so, how do we capture it and explain this phenomenon? To get further insight and confirmation of this non-triviality, let us examine whether it repeats if we make group categories in terms of seniors (that is, age is above 60 years) versus non-seniors (i.e., age falls within 18-59 years) as done in Table 2. Again, contrary to an intuitive common sense, a simultaneous consumption of alcohol and marijuana is less frequent among drivers on the road. We set out to explain this non-triviality by a statistical methodology. Now, let us probe how the relations among just alcohol, just marijuana and alcohol-marijuana combinations are occurring among impaired drivers. Drinking alcohol alone (at a correlation of 0.86) is the most highly correlated and the next is using marijuana alone (at a correlation level of 0.46) in the perception. The least correlated (at a correlation of 0.21) is the combination of alcohol and marijuana in the perception about road safety. These data clues suggest that the psychologic temptation to use both is too important to miss in a scientific inquiry about the road safety and hence, a new probability model for the data is sought and obtained in the next section.

Confounded Poisson Distribution with its Statistical Properties and Methodology

There is a need to capture the confounded users of marijuana and alcohol with respect to road safety. It has been investigated to some extent (not probabilistically) by Ramaekers et al. (2000) and Sivakumar and Krishnaraj (2012). It is common to apply probability ideas to demystify the pattern of accidents due to alcohol and/or marijuana (Davis, 2003; Fuchs et al., 2012). Let \( Y = y \) be a random number of users who might be impaired drivers due to alcohol and/or marijuana use. Suppose that the parameters \( \theta > 0 \) portrays the expected number of users of either alcohol or marijuana but not both. Then, the probability pattern of road safety is explainable using the usual Poisson distribution, \( p(y|\theta) = e^{-\theta}\theta^y/y! \), \( y = 0, 1, 2, \ldots, \infty; 0 < \theta < \infty \) with the mean \( \mu_y = E(y|\theta) = \theta \) and variance, \( \sigma_y = \mu_y \).

A challenge is in developing an appropriate distribution to model the number of users of both alcohol and marijuana. Impaired drivers due to alcohol or marijuana are tempted to use both with a temptation level. Because it is not measured in the data, it ought to be a parameter, \( \delta \geq 0 \) in our discussions. This psychologic temptation could kick in and it is the basis to infuse it as a weight factor, \( w(y, \delta) = \delta[y-1] \) in the probability modeling. Such modeling is recognized as size biased sampling in the statistics literature (see Patil and Rao, 1978 for details). Hence, we suggest using:

\[
\begin{align*}
\delta \geq 0 & \\
\theta > 0 & \\
\mu_y & = E(y|\theta) = \theta \\
\sigma_y & = \mu_y.
\end{align*}
\]
To capture and explain the psychologic temptation, δ ≥ 0 to use both marijuana and alcohol, as an addition to using alcohol or marijuana by impaired drivers. We name the probability distribution in (1) a Confounded Poisson Distribution (CPD) in (1).

Shanmugam (2011) introduced the CPD in another context to generalize the Modified Power Series Distribution (MPSD). When only alcohol or marijuana but not both is the cause of breaching road safety, the psychologic temptation to use both is non-existent by default (that is, δ = 0). Consequently, the CPD (1) reduces to the usual Poisson Distribution (PD) which portrays the pattern of traffic accidents due to alcohol or marijuana, where the parameter θ portrays the expected number of users of either alcohol or marijuana by impaired drivers.

Interestingly, when the number of alcohol and/or marijuana users in road is already at a specified number, y, for an additional one more user in road, the probability is dynamically changing by an amount:

\[
p(y+1|θ, δ) = \frac{θe^{−θ}(1+δ)(y+1)}{(y+1)(1+δ)(y-1)} p(y|θ, δ)
\]

The amplifier, \( \frac{θe^{−θ}(1+δ)(y+1)}{(y+1)(1+δ)(y-1)} \) varies not only due to the expected number of users of alcohol or marijuana but also in what neighborhood of the number, y of users. Realize that \( p(y=0|θ, δ) \) portrays the probability of full road safety. The jump-start deterioration probability from the full road safety to a non-safety is then portrayed by the probability inflator, \( θe^{−θ} \) in Fig. 1 (see Foster and George, 1994 for definition and details about risk). Realize that under this concept, the probability of breaching the full safety is:

\[
p(y=1|θ, δ) = θe^{−θ} p(y=0|θ, δ)
\]

The risk inflation level is \( θe^{−θ} \), which is controlled by expected number of users of alcohol, \( θ > 0 \) and the psychologic temptation, \( δ ≥ 0 \). The Fig. 1 depicts their nonlinear behavior. The mean of the CPD (1) is:

\[
μ = E(y|θ, δ) = θe^{−θ} \left( 1 + \frac{2θe^{−θ}}{1+θe^{−θ}} \right)
\]

after algebraic simplifications. Without using marijuana (that is, \( δ = 0 \)), the mean (2) reduces to \( E(y|θ, δ = 0) = θ \). A graph (Fig. 2) illustrates users of marijuana and alcohol in road. Without the marijuana users (that is \( δ = 0 \)), the expected number \( z = E(y|θ, δ = 0) = 0 \) of users of alcohol alone is simply an Euclidean plate (that is, \( z = 0 \)). Otherwise (that is with the users of marijuana), the expected users with driving impairment due to alcohol is twisted nonlinearly as in Fig. 2.

Furthermore, to visualize the scenarios of with and without marijuana users when the driver is already intoxicated by alcohol, why not reparametrize the mean (2) and express it as:

\[
σ^2 = μ^2 + μ^2 δ \left[ \frac{4 + θe^{−θ}(θe^{−θ} - 2)}{1 + θe^{−θ}(θe^{−θ} + 2)} \right]
\]

where, \( σ^2 = var(y|θ, δ) \) is the variance. In the regular Poisson chance mechanism, the variance, \( σ^2 \) is identical to the mean \( μ \) and hence, their relationship is an Euclidean plate (Fig. 2).
Fig. 1. Risk inflation level, $\theta e^{-\delta \theta}$ for road safety

However, in the case of Confounded Poisson Distribution (CPD) in (1), the variance, $\sigma^2$ is bent quadratic ally (see the bent Poisson plate in Fig. 2) by a nonlinear force $\mu_y^2 \delta \left(\frac{4 + \theta e^{-\delta \theta} (\theta e^{-\delta \theta} - 2)}{1 + \theta e^{-\delta \theta} (\theta e^{-\delta \theta} + 2)}\right)$ involving the mean $\mu_y$. When the temptation parameter, $\delta$ is absent (that is, $\delta = 0$), the nonlinear force is nullified and consequently, no bending of Poisson plate occurs. Its implication is then that there is more heterogeneity in the relationship between the variance and the mean in the case of Confounded Poisson Distribution (CPD) in (1). Otherwise, there is a constant Poisson heterogeneity. In other words, the temptation parameter, $\delta$ quantifies the force causing to drift from the constant Poisson heterogeneity.

Next, let us examine the road safety level with zero tolerance (that is, no fatal accident due to any cause) of driving impairment is most desired by the public and traffic regulating agencies. There have been movements
by the mothers against drunk driving in USA and elsewhere in other nations. To configure the chance for such zero tolerance of impaired drivers due to users of alcohol or marijuana, we need to obtain the Survival Function (SF) of the Confounded Poisson Distribution (CPD) in (1).

After algebraic simplifications (see Shanmugam, 2011 for details), the SF is obtained and it is:

\[
SF(m|\theta, \delta) = \frac{e^{-\theta} \theta^m}{m!} \cdot \frac{1}{1 + \delta(\theta e^{-\theta})^2}
\]

where, \(X_{2n-\text{df}}^2\) is the tail area of chi-squared distribution with \(2n\) Degrees of Freedom (DF) up to the percentile \(2\theta e^{-\theta}\). The cumulative chi-squared distribution is easily accessible in textbooks and software. The road safety level with zero tolerance as discussed above is then configured as:

\[
Pr(0|\theta, \delta) = 1 - SF(m = 1|\theta, \delta)
\]

(5.1)

Which is with the users of both marijuana and alcohol. Among single (either alcohol or marijuana) users (that is, \(\delta = 0\)), the road safety level with zero tolerance is:

\[
Pr(Y = 0|\theta, \delta = 0) = 1 - SF(m = 1|\theta, \delta = 0) = X_{2,0}^2
\]

(5.2)

Clearly, the road safety level \(X_{2,0}^2\) in (5.2) with just alcohol or marijuana users is higher than the road safety level \(X_{2,0}^2\) in (5.1) with users of both marijuana and alcohol. Hence, the Confounded Poisson Distribution (CPD) in (1) helps to capture and explain the worsening road safety due to users of alcohol and marijuana. We proved also that the heterogeneity increases along with the mean number of users.

Next, let us examine the odds ratio, \(\Omega(\theta, \delta)\) of road safety with the users of both marijuana and alcohol compared to either user. First, note that the odds of road safety is

\[
\text{Odds}_{\text{marijuana & alcohol}} = \frac{X_{2n-\text{df}}^2}{\delta(\theta e^{-\theta})^2 + 1 - X_{2n-\text{df}}^2}
\]

with the users of both marijuana and alcohol.

Secondly, the odds of road safety is

\[
\text{Odds}_{\text{alcohol}} = \frac{X_{2,0}^2}{1 - X_{2,0}^2}
\]

with the users of alcohol alone.

Hence, the odds ratio is:

\[
\Omega(\theta, \delta) = \frac{\text{Odds}_{\text{marijuana & alcohol}}}{\text{Odds}_{\text{alcohol}}}
\]

(6)

Which approaches one when the \textit{psychologic temptation} vanishes (that is, \(\delta \to 0\)). Otherwise (when, \(\delta \to 0\)), the odds ratio (6) shrinks.

Finally, the asymptotic maximum likelihood estimators of the parameters in the Confounded Poisson Distribution (CPD) in (1) (see Shanmugam, 2011 for details) are:

\[
\hat{\delta} = \frac{s_i^2 - \bar{Y}}{\bar{Y}^2(2\bar{Y} - s_i^2)} \quad \text{if} \ 2\bar{Y} \neq s_i^2
\]

(7)

In addition:

\[
\hat{\theta} = \frac{\bar{Y}}{(1 - \bar{Y})} \quad \text{if} \ 1 \neq \delta \bar{Y}
\]

(8)

Illustration of Road Safety based on US Data During 2013-2015

In this section, we demonstrate the expressions of section 2 using the self-admitted users of alcohol, marijuana or both while driving in USA during the years 2013-2015 as displayed in Table 1 (Source: Arnold and Tefft in http://AAA.foundation.org). There were 6,612 self-admitted users of alcohol, marijuana, or both in the age brackets: 18-24, 25-39, 40-59, 60-74 and 75+ years while driving. Among them, a sample average of 223 consumed both marijuana and alcohol with a sample variance of \(s^2 = 1,369.3\), while an average of 1,040 consumed only alcohol with a variance \(s^2 = 16,596\) and an average of 448 consumed only marijuana with a variance \(s^2 = 4,316\). Clearly, the alcohol users have more heterogeneity. The consumption of both alcohol and marijuana has lesser heterogeneity to the road safety than either alcohol or marijuana separately. This is counter-intuitive and it is captured with the help of CPD (1). Why is there such a counter-intuitiveness? We think that there must have been a \textit{psychologic} temptation to consume both and it is not measured in the data collection. It is therefore a parameter, \(\delta \geq 0\) in the data analysis. Hence, the Confounded Poisson Distribution (CPD) in (1) is appropriate to analyze and interpret the data evidences about the road safety.

By default, note that the parameter, \(\delta\) must be zero for the data in Table 1 on users of alcohol or marijuana but not both. Using the maximum likelihood estimator (8), the data based estimate of the expected number of users is \(\hat{\delta} \approx 208\) and \(\hat{\theta} \approx 89.52\) respectively in the case of using
alcohol or marijuana alone. Using the maximum likelihood estimator (7), the data based estimate of the psychologic temptation parameter and the expected number of users are $\hat{\theta} \approx 0.00052024$ and $\hat{\theta} \approx 45.65$ in the case of using both alcohol and marijuana.

Clearly, there is a finite non-zero but mild level of psychologic temptation to use both. Even with such a mild temptation level, the expected number of users decreases from 208 to 45.65 per 100,000. Using these estimates, we note that, for the road safety, the $Odds_{marijuana\&alcohol}$ with the users of both marijuana and alcohol is $2.12538 \times 10^{-20}$, while the $Odds_{alcohol}$ with the users of alcohol alone is $1.435 \times 10^{-20}$. Hence, their odds ratio is $1.435$, suggesting that the road safety is quite balanced with the users of both marijuana and alcohol and with the users of alcohol alone.

Now, let us examine whether the theme repeats when the data in Table 2 are looked upon in terms of seniors (age is 60 years or above) versus non-seniors (below 60 years of age). By default, note that the temptation parameter, $\delta$, must be zero for the data on users of alcohol or marijuana alone but not both. Using the maximum likelihood estimator (8), the data based estimate of the expected number of users is $\hat{\theta} = 243.5$ and $\hat{\theta} = 144.09$ respectively with the users of alcohol or marijuana alone. Using the maximum likelihood estimator (7), the data based estimate of the psychologic temptation parameter and the expected number of users are $\hat{\theta} \approx 0.0000057$ and $\hat{\theta} \approx 81.5$ with the users of both alcohol and marijuana. Clearly, there is a finite non-zero but mild level of psychologic temptation. Even with a mild level of psychologic temptation, the expected number of users decreases from 243.5 to 81.5.

Using the estimates of the expected number of users based on categorized data in terms of seniors versus non-seniors, we again note that, for the road safety, the $Odds_{marijuana\&alcohol}$ with the users of both marijuana and alcohol is $3.87896 \times 10^{-36}$, while the $Odds_{alcohol}$ with the users of alcohol alone is $3.8761 \times 10^{-36}$. Hence, their odds ratio is $0.998$, suggesting that the road safety is quite balanced with the users of both marijuana and alcohol and with the users of alcohol alone.

Contrary to a deductive logic, we learned surprisingly, from the data analytics, that the road safety is better with the users of both marijuana and alcohol than the road safety with the users of alcohol alone. These non-intuitive insights might not have been realized without the involvement of the Confounded Poisson Distribution (CPD) in (1). There may be much more still undiscovered mysterious findings with respect to the phenomenon called driving impairment which is a great concern to the public in general and mothers against drunk driving in particular.

Acknowledgement

The authors appreciate and thank their university for offering a research time to complete this work.

Funding

No special funding was involved for writing this article. It is not related to any funded research by any of the authors.

Author’s Contributions

All authors equally contributed in this work.

Conflict of Interest Statement

The authors have no conflicts of interest to disclose.

References

Arnold, L.S. and B.C. Tefft, 2016. Driving under the users of alcohol and marijuana: Beliefs and behaviors United States, 2013-2015.

Ashton, C.H., 2001. Pharmacology and effects of cannabis: A brief review. British J. Psychol., 178: 101-106.

Davis, G.A., 2003. Bayesian reconstruction of traffic accidents. Law, Probability Risk, 2: 69-89.

Foster, D.P. and E.I. George, 1994. The risk inflation criterion for multiple regression. Annals Stat., 22: 1947-1975.

Fuchs, P., T. Saska, R. Sousek and D. Valis, 2012. How to calculate the accident probability of dangerous substance transport? Archives Transport, 24: 273-284.
Hartman, R.L. and M.A. Huestis, 2013. Cannabis effects on driving skills. Clin. Chem., 15: 478-492. DOI: 10.1373/clinchem.2012.194381.

Li, M.C., J.E. Brady, C.J. DiMaggio, A.R. Lusardi and K.Y. Tzong et al., 2011. Marijuana use and motor vehicle crashes. Epidemiol. Rev., 34: 6572-6572. DOI: 10.1093/epirev/mxr017

Mannering, F.L. and C.R. Bhat, 2014. Analytic methods in accident research: Methodological frontier and future directions. Analytical Meth. Accident Res., 1: 1-22.

Mannering, F.L., V. Shankar and C.R. Bhat, 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytical Meth. Accident Res., 11: 1-16.

Patil, G.P. and C.R. Rao, 1978. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. Biometrics, 34: 179-189.

Ramaekers, J.G., H.W. Robbe and J.F. O’Hanlon, 2000. Marijuana, alcohol and actual driving performance. Human Psychopharmacology., 15: 551-535.

Shanmugam, R., 2011. Taylorized modified power series distributions with epileptic seizure incidence applications. Int. J. Stat. Econom, 7: 27-41.

Sivakumar, T. and R. Krishnaraj, 2012. Road traffic accidents due to drunken driving in India-challenges in prevention. Int. J. Res. Manage. Technol., 4: 401-406.